

Long-term stability of TES satellite radiance measurements

T. C. Connor¹, M. W. Shephard², V. H. Payne¹, K. E. Cady-Pereira¹, S. S. Kulawik³, M. Luo³, G. Osterman³, and M. Lampel⁴

¹Atmospheric and Environmental Research, Inc., Lexington, Massachusetts, USA

²Environment Canada, Toronto, Ontario, Canada

³Jet Propulsion Laboratory, Pasadena, California, USA

⁴Raytheon Company, Pasadena, California, USA

Received: 4 January 2011 – Published in Atmos. Meas. Tech. Discuss.: 14 March 2011

Revised: 12 July 2011 – Accepted: 13 July 2011 – Published: 25 July 2011

Abstract. The utilization of Tropospheric Emission Spectrometer (TES) Level 2 (L2) retrieval products for the purpose of assessing long term changes in atmospheric trace gas composition requires knowledge of the overall radiometric stability of the Level 1B (L1B) radiances. The purpose of this study is to evaluate the stability of the radiometric calibration of the TES instrument by analyzing the difference between measured and calculated brightness temperatures in selected window regions of the spectrum. The Global Modeling and Assimilation Office (GMAO) profiles for temperature and water vapor and the Real-Time Global Sea Surface Temperature (RTGSST) are used as input to the Optimal Spectral Sampling (OSS) radiative transfer model to calculate the simulated spectra. The TES reference measurements selected cover a 4-year period of time from mid 2005 through mid 2009 with the selection criteria being; observation latitudes greater than -30° and less than 30° , over ocean, Global Survey mode (nadir view) and retrieved cloud optical depth of less than or equal to 0.01. The TES cloud optical depth retrievals are used only for screening purposes and no effects of clouds on the radiances are included in the forward model. This initial screening results in over 55 000 potential reference spectra spanning the four year period. Presented is a trend analysis of the time series of the residuals (observation minus calculations) in the TES 2B1, 1B2, 2A1, and 1A1 bands, with the standard deviation of the residuals being approximately equal to 0.6 K for bands 2B1, 1B2, 2A1, and 0.9 K for band 1A1. The analysis demonstrates that the trend in the residuals is not significantly different from zero over the 4-year period. This is one method used to demonstrate

that the relative radiometric calibration is stable over time, which is very important for any longer term analysis of TES retrieved products (L2), particularly well-mixed species such as carbon dioxide and methane.

1 Introduction

The Tropospheric Emission Spectrometer (TES) is a Fourier Transform Spectrometer on board the NASA Aura platform (Beer et al., 2001; Beer, 2006; Schoeberl et al., 2006). TES has a number of observational modes (e.g. global survey, step-and-stare, transect). The high spectral resolution of the instrument makes it very useful for identifying and quantifying trace atmospheric gases, among which are ozone, carbon monoxide, methane, and ammonia. The user community consists of researchers interested in global air quality and climate change. This study utilizes observations in global survey mode, where TES makes measurements along the satellite track with a spacing of ~ 180 km. TES nadir spectra have 0.06 cm^{-1} unapodized spectral resolution with footprints of $8 \times 5\text{ km}^2$ resulting from the averages of 16 element detector arrays where each detector has a $0.5 \times 5\text{ km}^2$ nadir footprint. The TES spectral range is covered by four filters: 2B1 ($650\text{--}900\text{ cm}^{-1}$), 1B2 ($920\text{--}1150\text{ cm}^{-1}$), 2A1 ($1100\text{--}1340\text{ cm}^{-1}$) and 1A1 ($1900\text{--}2250\text{ cm}^{-1}$). The noise characteristics for each band are taken from Worden et al. (2006), and are listed in Table 1. In general, bands 1B2 and 2A1 have the lowest Noise Equivalent Detector Temperature (NEDT) values, band 2B1 has a slightly higher value and band 1A1 is the most noisy. In an analysis such as this, with the assumption of pure white (Gaussian) noise, there should be no expectation of bias on the average results given the large number



Correspondence to: T. C. Connor
(tconnor@aer.com)

Table 1. Bulk Statistics for TES bands 2B1, 1B2, 2A1, 1A1.

Band	Frequency	Mean, K		Delta, K Night–Day	NEDT K	Standard Deviation, K	
		Day (TES-OSS), K	Night (TES-OSS), K			Day	Night
2B1	900.280–901.480 cm ⁻¹	-0.224	-0.430	-0.21	1.0	0.603	0.628
1B2	1128.10–1128.35 cm ⁻¹	-0.044	-0.403	-0.36	0.3	0.615	0.600
2A1	1128.10–1128.35 cm ⁻¹	0.363	0.011	-0.35	0.3	0.596	0.582
1A1	2133.00–2134.00 cm ⁻¹	NA	-0.360	NA	1.5	0.891	0.862

of data points. Further, in this study we average the channels that span the micro-windows listed in Table 1, thus mitigating the effects of the noise on the standard deviation of the residuals. For band 2B1 the number of channels contributing to the brightness temperature used for comparisons is 20, it is 5 channels used in bands 1B2 and 2A1, and it is 16 channels for band 1A1. A description of the absolute calibration method used for the TES instrument is described in Worden et al. (2006). Shephard et al. (2008) described significant improvements obtained from further modifications to the TES Level 1B (L1B) calibration algorithms, evaluated using on-orbit TES nadir radiance comparisons for carefully selected, nearly coincident spectral radiance measurements from the Atmospheric Infrared Sounder (AIRS) on Aqua and targeted underflights of the Scanning High-resolution Interferometer Sounder (S-HIS). The comparisons of TES with S-HIS showed mean and standard deviation differences of less than 0.3 K at warmer brightness temperatures of 290–295 K. Larger TES/S-HIS comparison differences were observed for the higher-frequency TES 1A1 filter, which has less upwelling radiance signal. The TES/AIRS comparisons showed mean differences of less than 0.3 K at 290–295 K with standard deviation less than 0.6 K for the majority of the spectral regions and brightness temperature range. In the atmospheric window regions (sensing at or near the surface) the TES comparisons with both AIRS and S-HIS showed that AIRS or S-HIS minus TES produce similar differences for the 2B1 and 1B2 filters, but the 2A1 filter brightness temperature differences were 0.2–0.3 K warmer than those for filter 2B1 and 1B2. The Shephard et al. (2008) study also contains a description of the procedure used to warm up the optical bench, which was used to adjust and maintain the alignment of the beam splitter in December 2005. The procedure increases the integrated spectral magnitude thus providing a fourfold increase in the signal-to-noise ratio at higher frequencies, particularly band 1A1.

Aumann et al. (2006) successfully utilized an approach for validating the radiometric stability of AIRS measurements in surface viewing channels using buoy based observations of sea surface temperature over tropical oceans and a radiative transfer model. Here, we use a similar approach to validate the radiometric stability of TES measurements over time. We

use the day versus night statistics to show that the method reveals the expected climatological diurnal variability in the sea surface skin temperatures, thus producing accurate residual (observations minus calculations) time series upon which robust statistical analysis may be conducted. The accuracy of any time-dependent analysis of the TES Level 2 (L2) data products is predicated on the assumption of negligible drift in radiance measurements made by the instrument over the period of record. The aim of this paper is to explore the validity of that assumption as it pertains to the TES instrument.

In the next sections, we discuss the following: in Sect. 2 the overall approach adopted to generate the calculated brightness temperatures used for comparison to the TES measurements, along with the data used in the calculations. Section 3 is a discussion concerning the results of the comparison and the statistical techniques applied in the analysis for trend. Section 4 contains a brief summary of the findings.

2 Approach

The approach used to validate the stability of the TES calibration involves an analysis of the difference between measured and simulated TES radiances in relatively transparent (“window”) spectral regions. The spectral regions chosen for each TES filter are shown in Table 1. The atmospheric contribution to the radiance is computed by running a full radiative transfer calculation for each case, which is made more computationally amenable by the development of an Optimal Spectral Sampling (OSS) (Moncet et al., 2008) based model for TES. OSS was designed specifically to address the need for highly accurate real-time monochromatic radiative transfer calculations. The OSS model requires that training be performed for the particular instrument, where the training involves generating thousands of reference high resolution spectra, upon which a search for the optimal spectral nodes (i.e. monochromatic frequency positions) is performed. The criteria for the search are that the weighted combination of the nodes for a given instrument channel reproduce the reference spectra to a user defined level of accuracy. In the particular case of the TES instrument the training was done with reference to calculations performed using the Line By Line

Radiative Transfer Model (LBLRTM), a reference model described in more detail in Clough et al. (2005). The final numerical accuracy is within 0.05 K over all parts of the spectrum with respect to the LBLRTM training calculations. The transparent (window) regions of the spectrum used in this study will show much less error relative to the reference calculation. A conservative estimate would be 0.01 K in brightness temperature. The inputs required in order to obtain a calculated value that is representative of a TES measurement are:

1. an accurate, reliable, and traceable metric for the surface contribution at the time and place of the TES measurement, and
2. a reasonable estimate of the atmospheric state coincident with the TES measurement.

For the surface contribution, one must have accurate knowledge of the emissivity and temperature. The spectral emissivity of ocean water in the thermal infrared is well known (Masuda et al., 1988; Wu and Smith, 1997) and effectively invariant for the nadir view geometry of the TES instrument in Global Survey mode. The temperature of the ocean surface is derived from the National Centers for Environmental Prediction (NCEP) Real Time Global Sea Surface Temperature (RTGSST) product (Thiebaut et al., 2003), which is produced for a 0.5° by 0.5° grid on a daily basis and represents the average temperature for that day at 1m depth.

The RTGSST was developed by NCEP and is used primarily in support of daily weather forecasting. The gridded product is derived from a two-dimensional variational interpolation of buoy and ship surface temperature measurements, and AVHRR satellite measured Sea Surface Temperature (SST) from the previous day. Following the AIRS stability study in Aumann et al. (2006), the analysis is limited to within 30° latitude of the equator. This region has a high density of buoy measurements leading to low standard deviations in the RTGSST relative to the tropical verification buoys. The standard deviation between the RTGSST and the verification buoys is typically between 0.45 K and 0.55 K for the tropical oceans, with some element of seasonality present in the difference between the verification buoys and the RTGSST (Aumann et al., 2006).

The atmospheric temperature and water vapor profiles required for the simulation are taken from the Global Modelling and Assimilation Office (GMAO) fields used as the initial guess in the TES retrievals. Profiles of ozone, carbon monoxide, methane and nitrous oxide are taken from the tropical supplement to the US Standard Atmosphere (NASA, 1966), with CO_2 scaled to a concentration of 388 ppmv. Ostensibly, the spectral regions chosen for surface viewing are relatively clear from the spectral impacts of trace species so that precise knowledge of their concentration will have little to no impact on the final analysis. An assumption implicit throughout the analysis is that any trend in the surface

or atmospheric conditions is well captured by the geophysical datasets used as input to the radiative transfer model and therefore any trend in the residuals would have to be attributed to the instrument.

For every TES Global Survey there are sixteen orbits with a frequency of about one Global Survey for every two days. The instrument has the capability of gathering thousands of nadir spectra per Global Survey, thus providing a great deal of data upon which to build robust statistical conclusions. The total number of nadir spectra taken during the four year period of analysis was just under 2.2 million, of which approximately 2.6 % met the screening criteria described below:

1. measurements between -30° and $+30^\circ$ latitude
2. over ocean scenes, and
3. “clear sky” conditions defined as TES retrieved Cloud Effective Optical Depth ≤ 0.01 .

Note that the TES cloud optical depth retrievals (described in detail by Eldering et al., 2008 and Kulawik et al., 2006) are used here purely for screening purposes and that no effects of clouds on the radiances are included in the forward model. The observations were further split into day and night categories with numbers of observations equalling $\sim 36\,200$ and $\sim 21\,200$ for day and night respectively. In reference to Fig. 1, which shows the spatial distribution of the day and night measurements, there is relatively uniform spatial distribution with areas of greater observation counts (illustrated by the “warmer” colours) along the equatorial Pacific, the southern edge of the area of analysis (i.e. around 30° S) in all oceans, and along the east coast of Africa in the Indian Ocean. There is a decrease in the relative frequency of measurements passing the screening threshold during night-time views in the Northern hemisphere, which in large part accounts for the decrease in the counts during the night versus the day.

Figure 2a illustrates a relationship between the observed minus calculated TES brightness temperature for the 900.28 cm^{-1} – 901.48 cm^{-1} window in the 2B1 band and TES retrieved Cloud Effective Optical Depth up to an optical depth of 0.05. Figure 2b is the same as Fig. 2a, but for an optical depth of up to 0.01. The figure also shows the correlation coefficient for the residuals versus the effective optical depth. The value of -0.23 for the correlation coefficient in Fig. 2b indicates that 5 % of the variance in the residuals can be explained by the cloud effective optical depth. For comparison, in Fig. 2a, the correlation coefficient is around -0.6 , thus indicating that the effective optical depth describes 36 % of the variance in the residuals and complicates the analysis. By setting the Cloud Effective Optical Depth threshold to 0.01 we leave out much of the impact of clouds on radiances, which results in a smaller standard deviation of the residuals. This threshold still allows for tens of thousands

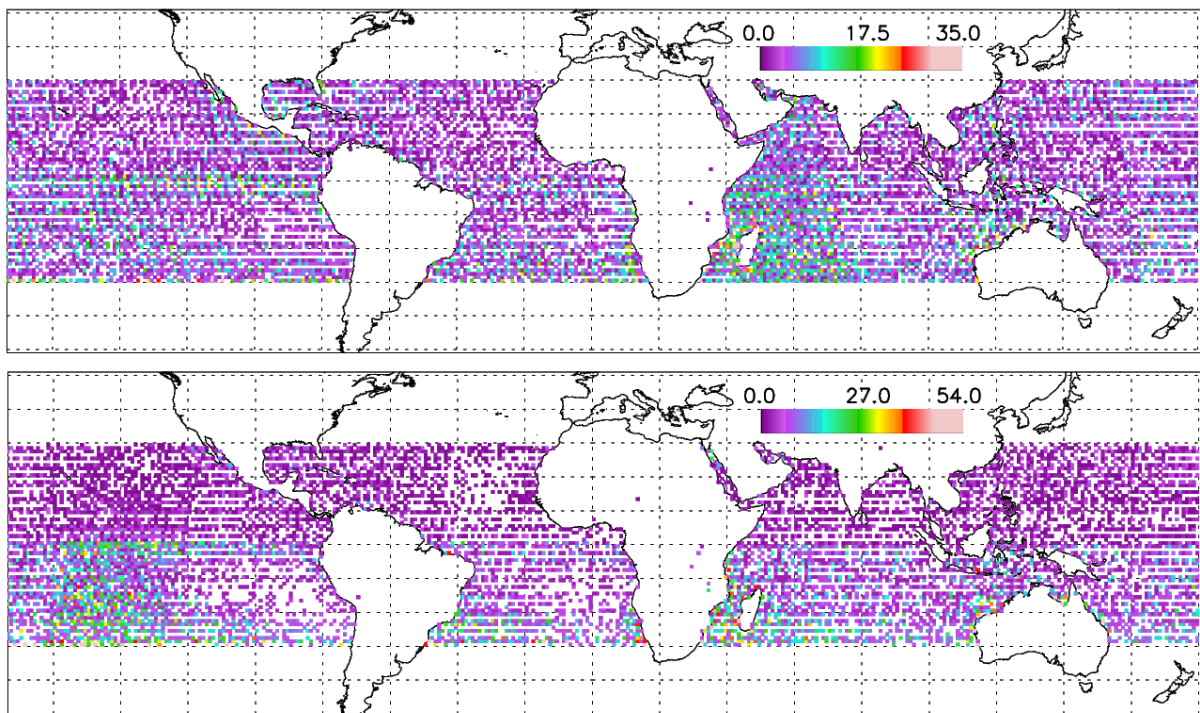


Fig. 1. Day (top panel) and night (bottom panel) number of observation counts in 1° latitude and longitude bins.

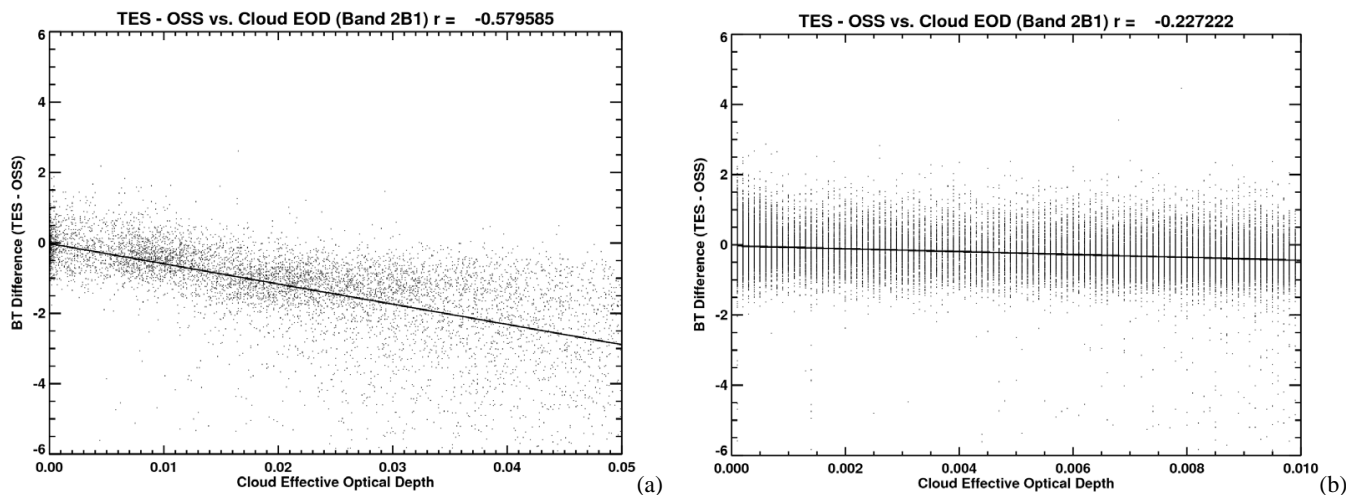


Fig. 2. (a) Relationship between Cloud Effective Optical Depth and Residuals for a Cloud Effective Optical Depth of up to 0.05. (b) Relationship between Cloud Effective Optical Depth and Residuals for a Cloud Effective Optical Depth of up to 0.01.

of spectra to be used in the analysis, resulting in robust statistical interferences. Figure 2a and b also provide evidence that the effective optical depth is providing a good screen for clear sky conditions as the calculated values in the residuals contain no treatment for cloud.

The calculated brightness temperature in window regions for each of the four TES bands is subtracted from the observed brightness temperature in these windows and a trend analysis conducted on the differences. The primary

assumption made about any time series that is implicit in analysing for trend is that it must be covariance stationary, which implies:

1. Constant and finite expected value
2. Constant and finite variance
3. Constant and finite covariance

If any of the assumptions of covariance stationary data are not met then the resulting trend and error statistics associated with the trend run the risk of being biased and incorrect.

3 Results and discussion

A representative example of a time series generated in this analysis is given in Fig. 3, which shows the OSS calculations subtracted from the TES measurements as a function of time for the 2A1 band during the descending (night) orbit. The figure demonstrates that the measurements have been consistent, which is to say that there is no obvious change in the variance or the mean of the residuals occurring through this 4-year period.

Figure 4 is the histogram of this time series for both day and night measurements along with a best-fit Gaussian distribution, showing that the residuals exhibit behaviour that looks approximately normal, but with heavier tails. The distribution of the residuals is tested for goodness-of-fit to a normal distribution later in this section. The standard deviation of ~ 0.6 K in the residuals for bands 2B1, 1B2 and 2A1 is indicative of the overall effectiveness of the technique given the NEDT metrics outlined in Table 1, the standard deviation between RTGSST and the verification buoys, and the fact that the average of multiple channels are used to derive the comparison brightness temperatures, which mitigates the effect of the NEDT. The 1A1 band is subject to a higher signal to noise ratio and therefore exhibits more spread with respect to the residuals (standard deviation of ~ 0.9 K), which is to be expected for this band. The accuracy of the method is explored in more detail in the following sections. Table 1 shows that TES bands 2B1, 1B2 and 2A1 all exhibit similar bulk statistics with respect to the analysis of residuals, while band 1A1 exhibits a much higher degree of spread. Band 1A1 will be discussed in more detail later in the paper. The means are different across each band and within the bands as a function of night and day measurement times, but the standard deviations are all very similar. Shephard et al. (2008) found a similar pattern of differences across the bands when comparing radiances from AIRS and SHIS as a reference. Specifically, they found bands 1B2 and 2B1 to have a slight negative bias (-0.02 K and -0.07 K respectively, for TES-AIRS), and band 2A1 to have a slight positive bias (0.26 K) with respect to the reference measurements. These favourable comparisons are taken as a further indication of the accuracy of the modeling approach.

3.1 Day and night differences

The day night differences are expected as a consequence of using the RTGSST, which is a daily average, as a surrogate for the instantaneous skin temperature measured by the satellite. The expected diurnal variation about the daily mean at 1 m depth is approximately 0.2 K (Kennedy et al., 2007). The

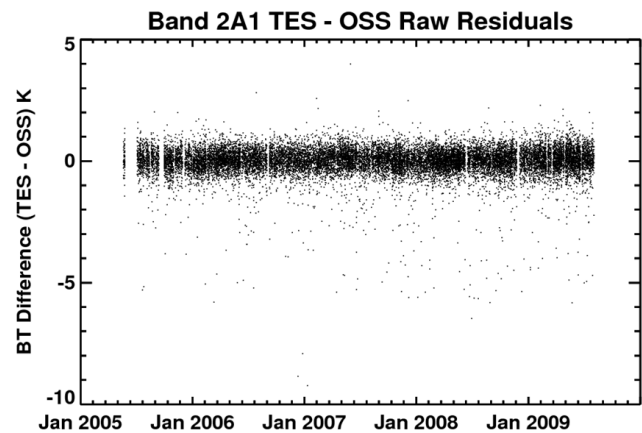


Fig. 3. Brightness temperature differences on a per-spectrum basis. No obvious violation of covariance stationary assumption.

skin temperature, which is what is measured by the TES instrument, will have a constant cool bias relative to the temperature at 1 m depth for winds that are greater than 6 m s^{-1} . This cool bias is estimated to be -0.17 K with a 0.07 K standard deviation (Donlon et al., 2002), and has effectively no relationship to time of day (at least in the presence of a wind greater than 6 m s^{-1}). The justification for basing the analysis on the assumption of winds at or above 6 m s^{-1} is that, on average, this should be the case. Donlon et al. (2002) state that 70 % of the time in the tropical oceans the winds will blow at or above 6 m s^{-1} . Kennedy et al. (2007) produced a climatology of diurnal SST variations that is based on an analysis of hourly drifting buoy observations made in the tropics within 20° of the equator. This analysis shows that the expected difference in bulk temperature for the TES overpass time of 01:45 p.m. is 0.20 K above the daily mean, and for the overpass time of 01:45 a.m. it is 0.13 K below the daily mean. Therefore, the expected daytime TES measurement, in the presence of winds greater than 6 m s^{-1} may be represented as:

$$BT_{\text{day_measures}} = BT_{\text{day_calculated}} - 0.17 \text{ K}_{\text{cool_skin_bias}} + 0.20 \text{ K}_{\text{day_warm_bias}} \quad (1)$$

where the cool bias is to account for the difference between the bulk buoy measurement and the TES skin measurement and the positive bias is to account for overpass occurring at a time when the water column is expected to be about 0.2° above the daily average. The expected nighttime TES measurement, again, in the presence of wind speeds greater than 6 m s^{-1} is represented as:

$$BT_{\text{night_measured}} = BT_{\text{night_calculated}} - 0.17 \text{ K}_{\text{cool_skin_bias}} - 0.13 \text{ K}_{\text{night_cool_bias}} \quad (2)$$

In reference to Table 1, it is clear that bias in the measured minus calculated do not exactly align with the theoretical

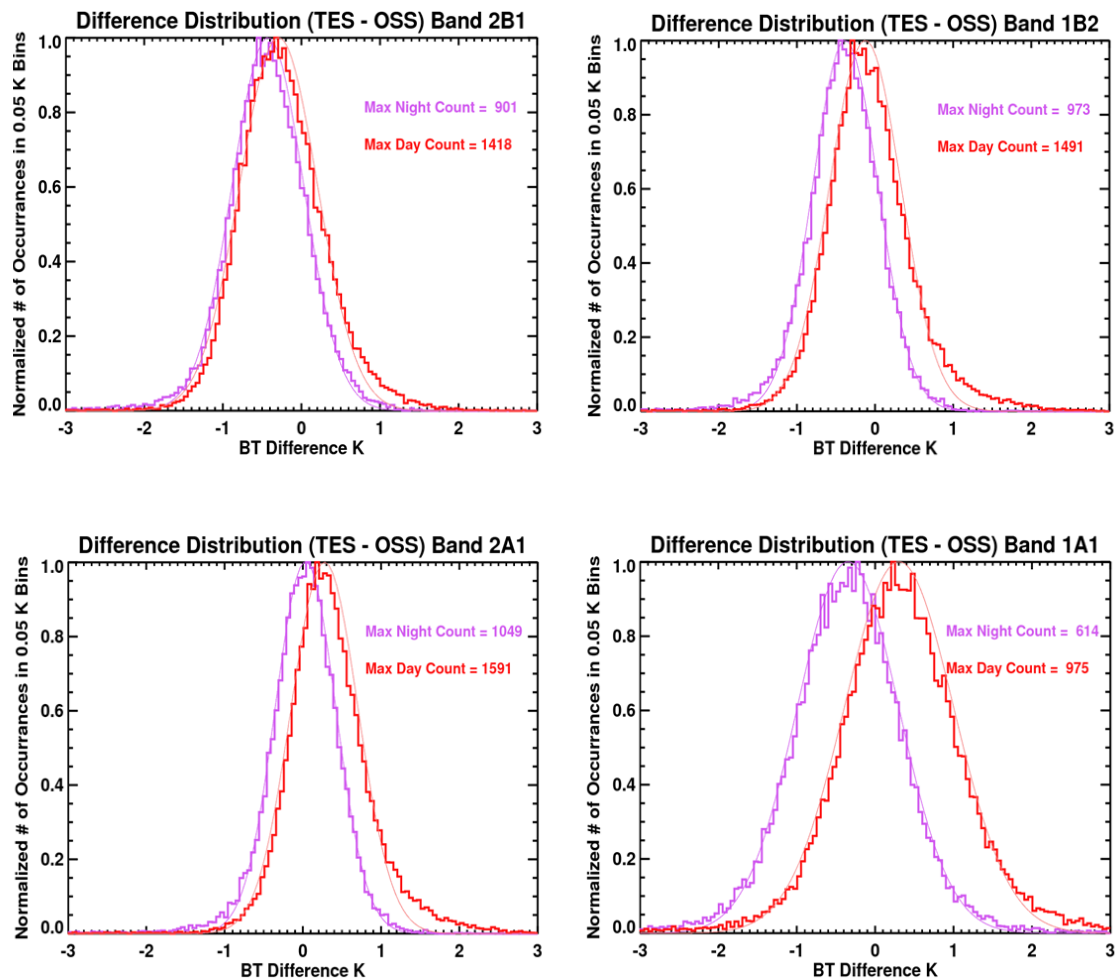


Fig. 4. Histograms of raw residual data for the four TES bands with ascending (day) orbits in red and descending (night) orbits in purple. Day-night differences are discussed in Sect. 3.1.

outlay above, however, we may effectively account for all biases and look at the ability of the instrument to resolve the diurnal variation in the skin temperature by analyzing the double differences (Aumman et al., 2006). The double differences are simply:

$$\left[\text{Mean} (BT_{\text{night_measured}}) - \text{Mean} (BT_{\text{night_calculated}}) \right] - \left[\text{Mean} (BT_{\text{day_measured}}) - \text{Mean} (BT_{\text{day_calculated}}) \right] \quad (3)$$

where the expected double difference is just equal to the diurnal difference based on the TES overpass times for day and night. This value is approximately -0.33 K and examination of Table 1 indicates that for bands 2A1 and 1B2, the differences are calculated as -0.35 K and -0.36 K respectively, which are in very good agreement with the theoretically derived value. Indeed, when compared with the AIRS finding of a diurnal difference of -0.38 ± 0.05 K (Aumann et al., 2006), where AIRS and TES have nearly coincident equator crossing times (separated by ~ 15 min), this agreement in

the average diurnal difference is expected and reassuring in that it confirms the stability of the method which implies that the statistical behaviour of the measurements themselves are being accurately captured, especially in so far as trends are concerned.

The interpretation of band 2B1 in the context of the diurnal difference analysis is more nuanced. The instrument is measuring the diurnal difference in the correct direction, but the magnitude of the difference is reduced compared to what is expected. It is apparent from Fig. 4 that the distributions of day and night residuals for band 2B1 are offset by an amount that is less than expected, but uniform through the distribution. Given the large number (tens of thousands) of measurements in this analysis, the different noise characteristics of the TES bands would not be expected to play a role in the day/night differences. The less than expected offset between the day and night for band 2B1 cannot be explained by the spectral variation in the effect of clouds on the radiances either. The clouds appear to affect the three

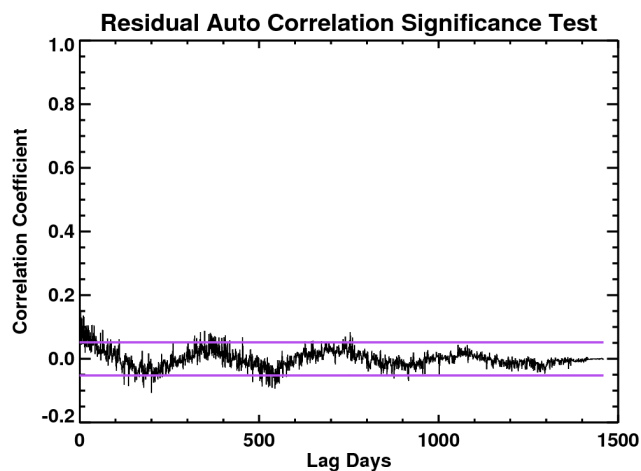


Fig. 5. Significance of auto-correlation in the residuals data – time series is not covariance stationary due to the significance of the auto correlations at ~ 180 , 360 and 540 days.

bands in an identical fashion, with correlation coefficients for day and night residuals versus cloud effective optical depth equal to -0.23 (Fig. 1) and -0.16 respectively, so while there is clearly an expected impact from clouds in an absolute sense (i.e. a negative bias) it does not lead to differences in day/night differences across bands 2B1, 1B2 and 2A1. A difference in the water depth at which the radiance is emitted for different frequencies is another factor that can impact the comparison of two differing spectral regions (McKeown et al., 1995). The 2B1 channels will receive radiance that has been emitted from a part of the skin that is closer to the actual air-sea interface (Wieliczka et al., 1989). The $\sim 1128\text{ cm}^{-1}$ channels in bands 1B2 and 2A1 will theoretically get most of their energy from a portion of the ocean skin that is about $20\text{ }\mu\text{m}$ deep, while the $\sim 900\text{ cm}^{-1}$ channels for band 2B1 will receive most of the energy from about $7\text{ }\mu\text{m}$ deep. The 900 cm^{-1} channels are emitting very close to the air-sea interface, which has the potential to confound the comparison to bands 1B2 and 2A1.

3.2 Analysis for trend

In the analysis for trend a first step of binning the residuals data into daily resolution is taken. The data are then tested for the assumption that the residuals are a stationary time series. Figure 5 is a plot of the autocorrelation coefficient versus lag days for the 2A1 night residuals, showing the 95 % significance threshold. Where values for the lagged correlation coefficients cross this threshold they are assumed to be significant at the 95 % confidence limit, and stationary assumptions made about the residuals are not valid. The plot demonstrates that there are packets of significant autocorrelations around lags at 180, 360, and 540 days thus indicates that the data is non-stationary and needs to be transformed before any regression to assure that error statistics can be

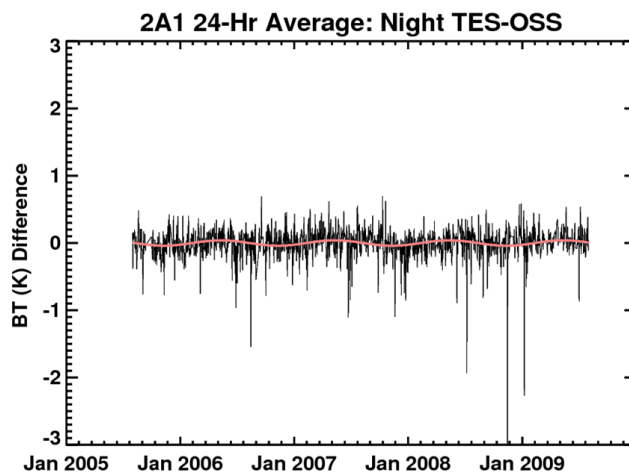


Fig. 6. Daily resolution of a time series containing a yearly harmonic.

considered robust. There are a number of techniques available to deal with the seasonality in the residuals, but in this particular case it is straightforward to simply remove them. The technique outlined by Wilks (2006) is selected, which prescribes a:

1. fourier transformation into the frequency domain
2. power spectrum analysis of the frequencies
3. identification of the frequencies with the most power and mask everything else
4. transform the masked spectra back into the time domain.

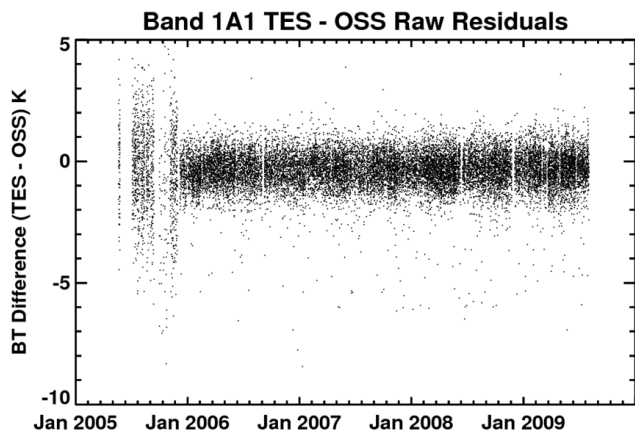
The result of the above steps can be seen in Fig. 6, which shows the times series for band 2A1 at night, binned in 24-h increments, and the seasonal cycle found through the Fourier decomposition. The seasonal cycle is simply subtracted from the time series for a set of residuals that have been transformed into a stationary time series (i.e. at this point the mean, variance and covariance are assumed constant over the period).

3.2.1 Analysis of the 1A1 band

The 1A1 band required some additional data transformation in order to get a stationary dataset. Figure 7 shows that the time series has a change in the variance in late 2005. This is the result of the optical bench warm-up, which took place in December of that year. A description of the optical bench warm-up procedure may be found in Shephard et al. (2008). The impact of this event on the measurements is evident from the dramatic change in variance. The technique used in this study to account for the non-stationary variance is to find the average variance before and after the optical bench warm-up

Table 2. Results of the χ^2 test for normal distribution.

Band	Degrees of Freedom		Calculated χ^2		χ^2 at $\alpha = 5\%$		Is The Distribution Normal?	
	Day	Night	Day	Night	Day	Night	Day	Night
2B1	21	20	205.0	54.6	32.7	31.4	No	No
1B2	21	20	161.3	68.5	32.7	31.4	No	No
2A1	24	20	117.6	42.8	36.4	31.4	No	No
1A1	23	27	110.7	57.8	35.2	40.1	No	No

**Fig. 7.** Optical bench warm-up in late 2005 causes a significant change in the standard deviation of the residuals.

and simply divide it back into the time series, thus normalizing the variance across the entire time series. The same steps described in the previous section are then applied to address the seasonality in the mean of the residuals. Finally the trend analysis is conducted in the same manner as for the other bands. Another feature of the 1A1 radiances that set them apart from the other bands is that the day time statistics for this band will be influenced in a non-negligible way by reflected sun light. This mainly manifests as a high daytime bias in the mean for the residuals, as well as a higher difference between the day time and night time residuals. However, this constant bias in measured minus modelled should not have any deleterious effects on the trend analysis.

3.2.2 Analysis for trend – test for normal distribution

The distribution of the residuals needs to be considered as it will impact the test selected for analysing the trend. A simple least-squares regression will not suffice if the underlying data comes from a distribution that is other than normal. The data in Table 2 provide the details of a χ^2 test for normal distribution of the residuals for each of the bands over day and night time. Provided in the table are the calculated χ^2 statistic, the bin size used to generate the histogram upon which the fit is

performed, the degrees of freedom resulting from the number of bins minus the number of bins with zero frequency, and the value for the 5% significance level for the given degrees of freedom. For the goodness-of-fit test the null (H_0) and alternative hypothesis (H_a) are stated as:

H_0 : The residuals are from a normal population

H_a : The residuals are not from a normal population

The results of the tests show that the null hypothesis is rejected in all cases at the 5% significance level and the conclusion is that the residuals are not from a normal population. This is especially true for the day time residuals, where the test statistic is very large compared to the reference at 5%. The assumption that the residuals come from a normal distribution is likely not a good one. As a consequence of this finding, it is necessary that a test for trend be conducted where the test is non-parametric and makes no assumption about the heredity of the distribution from which the residuals are sampled.

3.2.3 Analysis for trend – Mann-Kendall test

The non-parametric test chosen is the Mann-Kendall test for trend. The Mann-Kendall test is particularly useful for trend analysis in data derived from environmental systems for the following reasons; the test does not require the assumption of normally distributed data, and is insensitive to outliers (Lisell and Grimvall, 2002). The test is applicable in cases when the data is assumed to be from a covariance stationary time series, which is the case after the adjustments described in the previous sections are applied to the TES time series data. The Mann-Kendall considers whether the variable tends to increase or decrease with time by computing a test statistic, which is calculated as the sum of the signs of the slopes for every combination of two data points from the timeseries (Gilbert, 1987) The equations relevant to applying the Mann-Kendall test are given below:

$$S = \sum_{i=2}^n \sum_{j=1}^{i-1} \text{sign}(x_i - x_j) \quad (4)$$

Table 3. Mann-Kendall trend analysis results for each TES band.

Band	S		$SE(S)$		Z-statistic		S significant at $\alpha = 5\%$	
	Day	Night	Day	Night	Day	Night	Day	Night
2B1	14 351	13 573	18 661	18 661	0.77	0.73	No	No
1B2	21 899	5317	18 661	18 661	1.17	0.28	No	No
2A1	413	-1043	18 661	18 661	0.02	-0.06	No	No
1A1	29 391	6509	18 661	18 661	1.57	0.35	No	No

where:

n = number of observations

$\text{sign}(x_i - x_j) = -1$ for $(x_i - x_j) < 0$

$\text{sign}(x_i - x_j) = 0$ for $(x_i - x_j) = 0$

$\text{sign}(x_i - x_j) = 1$ for $(x_i - x_j) > 0$

The variance of S is given as:

$$\text{Var}(S) = n(n-1)(2n+5)/18. \quad (5)$$

The test statistic is calculated as:

$$Z = S / \sqrt{\text{Var}(S)} \quad (6)$$

and is compared against the critical value from a Standard Normal Table. The appropriate value at the 5% significance level for this test is 1.96. For the Mann-Kendall test the null hypothesis and alternative hypothesis may be stated as:

H_0 : the value of S is zero

H_a : the value of S is not zero.

Table 3 lists the results of applying the Mann-Kendall test to the residuals for the four TES bands for both day and night. The S entry in Table 3 is the result of applying the Eq. (4) to the data, and $SE(S)$ is simply the standard error of S , which is taken as the square root of the variance calculated in Eq. (5). The individual Z-scores are given in the table and demonstrate that the null hypothesis at a 5% significance level cannot be rejected. These results lead us to state that over the period of analysis in this study that there has been no discernible drift in the radiance measurements made by the TES instrument.

4 Conclusions

In this analysis $\sim 57\,400$ clear sky tropical ocean TES TOA radiances in window regions were compared with TOA simulated radiances generated from radiative transfer calculations using RTGSST and the atmospheric state between the surface and the TOA being estimated from the GMAO profiles. The resulting standard deviation of the residuals are approximately equal to 0.6 K for bands 2B1, 1B2, 2A1, and 0.9 K for band 1A1. The analysis shows that the measurements made

by the TES instrument have been very stable over the four years examined in this study. The statistical evidence from the comparison shows that there is no reason to reject the null hypothesis of no trend for all the TES bands over the course of four years, thus concluding that any trend in the instrument over this time period can be assumed to equal zero. This is a very important result when it comes to performing trend analysis using the TES level 2 products (e.g. CO₂ and CH₄), for it may be said that any trend observed in level 2 data (that was measured between August 2009 and August 2005) is not attributable to the satellite, but the environment being measured. It is important to note that as the instrument continues along its useful life, these conclusions will need to be revisited as they are only valid over the four-year period of analysis in this study.

Acknowledgements. We would like to thank Ned Snell at AER for his helpful insights and discussions, and Robert Morris at Environment Canada for his help with the Mann-Kendall trend statistical analysis. This work was supported by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Edited by: T. von Clarmann

References

- Aumann, H. H., Broberg, S., Elliott, D., Gaiser, S., and Gregorich, D.: Three years of AIRS radiometric calibration validation using sea surface temperatures, *J. Geophys. Res.*, 111, D16S90, doi:10.1029/2005JD006822, 2006.
- Beer, R.: TES on the Aura mission: Scientific objectives, measurements and analysis overview, *IEEE T. Geosci. Remote*, 44(5), 1102–1105, doi:10.1109/TGRS.2005.863716, 2006.
- Beer, R., Glavich, T. A., and Rider, D. M.: Tropospheric Emission Spectrometer for the Earth Observing Systems Aura satellite, *Appl. Optics*, 40, 2356–2367, doi:10.1364/AO.40.002356, 2001.
- Clough, S. A., Shephard, M. W., Mlawer, E. J., Delamere, J. S., Iacono, M. J., Cady-Pereira, K., Boukabara, S., and Brown, P. D.: Atmospheric radiative transfer modeling: A summary of the AER codes, *J. Quant. Spectrosc. Ra.*, 91, 233–244, doi:10.1016/j.jqsrt.2004.05.058, 2005.

- Donlon, C. J., Minett, P. J., Gentemann, C., Nightingale, T. J., Barton, I. J., Ward, B., and Murray, M. J.: Toward Improved Validation of Satellite Sea Surface Skin Temperature Measurements for Climate Research, *J. Climate*, 15, 353–369, 2002.
- Eldering, A., Kulawik, S. S., Worden, J., Bowman, K., and Osterman, G.: Implementation of cloud retrievals for TES atmospheric retrievals: 2. Characterization of cloud top pressure and effective optical depth retrievals, *J. Geophys. Res.*, 113, D16S37, doi:10.1029/2007JD008858, 2008.
- Gilbert, R. O.: Statistical methods for environmental pollution monitoring, Van Nostrand Reinhold, New York, 1987.
- Kennedy, J. J., Brohan, P., and Tett, S. F. B.: A global climatology of the diurnal variations in sea-surface temperature and implications for MSU temperature trends, *Geophys. Res. Lett.*, 34, L05712, doi:10.1029/2006GL028920, 2007.
- Kulawik, S. S., Worden, J., Eldering, A., Bowman, K., Gunson, M., Osterman, G. B., Zhang, L., Clough, S., Shephard, M. W., and Beer, R.: Implementation of cloud retrievals for Tropospheric Emission Spectrometer (TES) atmospheric retrievals: part 1, Description and characterization of errors on trace gas retrievals, *J. Geophys. Res.*, 111, D24204, doi:10.1029/2005JD006733, 2006.
- Libiseller, C. and Grimvall, A.: Performance of Partial Mann-Kendall Tests for Trend Detection in the Presence of Covariates, *Environmetrics*, 13(1), 71–84, 2002.
- Masuda, K., Takashima, T., and Takayama, Y.: Emissivity of pure and sea waters for the model sea surfaces in the infrared window regions, *Remote Sens. Environ.*, 24, 313–329, 1988.
- McKeown, W., Bretherton, F., Huang, H. L., Smith, W. L., and Revercomb, H. L.: Sounding the skin of water: Sensing air/water interface temperature gradients with interferometry, *J. Atmos. Ocean. Tech.*, 12, 1313–1327, 1995.
- Moncet, J.-L., Uymin, G., Lipton, A. E., and Snell, H. E.: Infrared Radiance Modeling by Optimal Spectral Sampling, *J. Atmos. Sci.*, 65(12), 3917–3934, doi:10.1175/2008JAS2711.1, 2008.
- NASA: U.S. Standard Atmosphere Supplements, U.S. Government Printing Office, Washington, DC, 1966.
- Schoeberl, M. R., Douglass, A. R., Hilsenrath, E., Bhartia, P. K., Beer, R., Waters, J. W., Gunson, M. R., Froidevaux, L., Gille, J. C., Barnett, J. J., Levelt, P. F., and DeCola, P.: Overview of the EOS aura mission, *IEEE T. Geosci. Remote*, 44(5), 1066–1074, doi:10.1109/TGRS.2005.861950, 2006.
- Shephard, M. W., Worden, H. M., Cady-Pereira, K. E., Lampel, M., Luo, M., Bowman, K. W., Sarkissian, E., Beer, R., Rider, D. M., Tobin, D. C., Revercomb, H. E., Fisher, B. M., Tremblay, D., Clough, S. A., Osterman, G. B., and Gunson, M.: Tropospheric Emission Spectrometer nadir spectral radiance comparisons, *J. Geophys. Res.*, 113, D15S05, doi:10.1029/2007JD008856, 2008.
- Thiebaut, J., Rogers, E., Wang, W., and Katz, B.: A new high-resolution blended real time global sea surface temperature analysis, *B. Am. Meteorol. Soc.*, 84(5), 645–656, 2003.
- Wieliczka, D. M., Weng, S., and Querry, M. R.: Optical Constants of Water in the Infrared, *Appl. Optics*, 28, 1714–1719, 1989.
- Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, 2nd Edn., International Geophysics Series, Vol. 59, Academic Press, 627 pp., 2006.
- Worden, H., Beer, R., Bowman, K., Fisher, B., Luo, M., Rider, D., Sarkissian, E., Tremblay, D., and Zong, J.: TES level1 algorithms: Interferogram processing, geolocation, radiometric and spectral calibration, *IEEE T. Geosci. Remote*, 44(5), 1288–1296, doi:10.1109/TGRS.2005.863717, 2006.
- Wu, X. and Smith, W. L.: Emissivity of rough sea surface for 8–13 μm : Modeling and verification, *Appl. Optics*, 36(12), 2609–2619, 1997.