Evaluation of globally gridded SST products from NOAA, CMC and UKMeto using AIRS and CrIS SST measurements.

H. H. Aumann, E. M. Manning, R. C. Wilson and J. Vasquez

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

Abstract

We compare the daily Sea Surface skin Temperature (SST) derived from calibrated radiances from AIRS (Atmospheric Infrared Sounder) on EOS Aqua and CrIS (Crosstrack Interferometer Sounder) on SNPP between 2013 and 2019 to each other and to three major gridded daily Sea Surface Temperature (SST) products: The NOAA Real Time Global (RTG) SST, the Canadian Meteorological Centre (CMC) SST, and the OSTIA SST from the UK MetOffice. The mean and standard deviation of SST models relative to independent observation with the SST derived from AIRS or CrIS under clear conditions are an effective way of evaluating the fidelity and trends of SST models. AIRS and CrIS show very similar patterns relative to the models, dominated by difference between the skin and the buoy temperatures, but CrIS is on average 0.2 K warmer than AIRS. The daily standard deviation of the difference between the AIRS, the CMC and the OSTIA is 0.4 K, a level previously only seen in SST models relative to the Argo buoys. Between 2013 and 2019, all three models show artifacts related to version changes, which limit the interpretation of zonal trend from the 2013-2019 time period at the 3 mK/yr level. The 5 year anomaly trend in (AIRS-CrIS) double difference is relatively latitude zone and day/night consistent at between +2 mK/yr and -1 mK/yr, although this results can't be assumed to be valid for all channels and the full dynamic range of the scenes encountered by AIRS and CrIS.

1. Introduction

The Sea Surface Temperature (SST) is a key component of climate research, and a potential resource for the evaluation of the accuracy and stability of the SST derived from space-borne sensors. We focus on three daily gridded SST products: The NOAA Real Time Global (RTG) SST (Thiebault et al. 2003), the Canadian Meteorological Centre (CMC) SST (Brasnett 2008)

and the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) SST (Donlon et al. 2012). Between 2002 and 2019 the RTG production used AVHRR data from NOAA 16-19, and more recently AVHRR equivalent VIIRS data from NOAA 20. The CMC SST has been produced on a 0.2 degree grid since 1991 (v2.0), since 2016 (v3.0) it has been available on a 0.1 degree grid. Sensors used in the production of the CMC SST include the AVHRR from NOAA-18, 19, the European Meteorological Operational-A (METOP-A) and Operational-B (METOP-B), and microwave data from the Advanced Microwave Scanning Radiometer 2 (AMSR2) onboard the GCOM-W satellite. The version 2.0 of the OSTIA used in our analysis has been produced on a 5 km grid by the UK Meteorological Office since 2013, using AVHRR, the Spinning Enhanced Visible and Infrared Imager (SEVIRI), the Geostationary Operational Environmental Satellite (GOES) imager, the Infrared Atmospheric Sounding Interferometer (IASI, Blumstein et al. 2008), the Tropical Rainfall Measuring Mission Microwave Imager (TMI). All three products also use in situ observations of the SST from drifting and moored buoys and ships. Common to these models is that they refer to a daily mean buoy equivalent temperature at 2 meter below the surface. All data used in the production of the SST models have to be converted to buoy equivalent temperatures using a local time, latitude, season and data source dependent model. The resulting model temperatures thus become sensitive to changes in the input data, potentially creating trend artifacts.

Finding globally distributed and independent truth data to evaluate the stability of these SST models for the global open oceans is difficult. Fiedler et al. 2019 (Table 4) stated a trend of -1.9+/-1.0 mK/yr (2 sigma) for the CMC, +10.6+/-3.5 mK/yr (2 sigma) for the OSTIA v1.0 SST relative to the Global Tropical Moored Buoy Array (GTMBA) for the August 2002 - December 2007 period. This trend is valid only within about 6 degrees of the equator. The Argo buoys are more uniformly distributed with latitude, but their number increased between 2002 and 2009 by a factor of six, making trends relative to SST models difficult to interpret. This absence of a stable and independent reference provides an opportunity for the use of Atmospheric Infrared Sounder (AIRS, Aumann et al. 2003) and Crosstrack Interferometer Sounder (CrIS) data, since neither has been used in the RTG, the CMC and the OSTIA production. The objective of our paper is to use AIRS and CrIS radiance derived SSTs in the evaluation of these three SST products and to each other.

2. Data

For AIRS we used the near-real-time L1b v5 calibrated data available from the GSFC/DISC since September 2002. For CrIS SNPP we used the retrospective calibration based on algorithms developed in 2018 at the UW-Madison Space Science and Engineering Center: Hank Revercomb, and at the UMBC Atmospheric Spectroscopy Laboratory: Larrabee Strow (2018), Suomi NPP CrIS Level 1B Normal Spectral Resolution V2, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), for data between August 2012 and March 2019. The AIRS and CrIS data used the clear footprints saved as part of the ACDS (AIRS Calibration Data Subset) and the equivalent CCDS (CrIS Calibration Data Subset). AIRS and CrIS used the same clear filter concept based on a 3x3 footprint Spatial Coherence Test (SCT, Aumann et al. 2019). Samples where the SCT is less than a 0.5K threshold are identified as SCT clear, typically more than 30,000 each day. While AIRS and CrIS both are in 1:30AM ascending node orbits, their altitudes are 703 km and 825 km, respectively. As a result different locations are identified as clear by AIRS and CrIS. Between May 2013 and March 2019 data from AIRS and CrIS and the three previously identified SST models were available.

We evaluated the SST models using the atmospheric window channels at 1231.3 cm^{-1} for AIRS, with brightness temperature bt1231, and at 1232.5 cm^{-1} for CrIS, bt1232. The SST derived from bt1231 and bt1232 are referred to as sst1231 and sst1232, or AIRS SST and CrIS SST, depending on the context. The difference between the model SST and the AIRS measured bt1231 is typically 1K related to surface emissivity plus 2.5K due to water vapor absorption. For the CrIS bt1232 the difference is 1K plus 3.5K. Both sst1231 and sst1232 use a split window type water vapor correction based on the brightness temperature difference between the 1231.3 cm^{-1} and 1227.7 cm^{-1} channels for AIRS, 1232.5 cm^{-1} and 1227.5 cm^{-1} for CrIS. The water vapor correction was regression trained on 1403 open ocean profiles from ECMWF (Aumann et al. 2019). For the calculations of the clear AIRS spectra we used SARTA. SARTA is the Radiative Transfer Model (RTM) without clouds developed for AIRS (Strow et al. 2006). SARTA is based on HITRAN2008, updated with the current version of the CKHD3.2 water continuum. For CrIS we used the Community Radiative Transfer Model (CRTM) 2.1.3. CRTM is based on LBLRTM

v12.2, with the water continuum absorption based on MT_CKD, version 2.4.

For each day we calculated the sst1231 and sst1232 for the clear footprints and matched them by their longitude and latitude to the nearest grid point of the models. We then separated the data into eight groups, day and night data, and four latitude bands: 30N-50N, 30N to 0, 0 to 30S, and 30S-50S. For each group we calculated the daily mean $=\sum(\text{SST-model})/N$, and standard deviation $\text{stddev}=\sqrt{(\sum(\text{SST-mean-model})^2/N)}$, where N is the daily number of the SST matchups.

3. Results

We first discuss the results in terms of the comparisons of (AIRS-model) and (CrIS-model) and then as the direct difference (AIRS-CrIS).

1. (AIRS-model) and (CrIS-model)

Figure 1 illustrates results in terms of the time series of the daily mean(sst1231-model) for AIRS and CrIS for the night overpasses of the combined 30S-0, 0-30N latitude zone of the oceans, using on average 8700 clear footprint each day from AIRS, 7500 clear footprints per day from CrIS. The patterns in the time series with AIRS and CrIS are very similar, dominated by a peak-to-peak 0.25K seasonal modulation. While AIRS is 0.61K colder than the CMC, CrIS is 0.43 K colder, i.e. CrIS is 0.18K warmer than AIRS in the night 30S-30N zone. This difference will be discussed later.

AIRS and CrIS SST are expected to be colder at night than the models: The skin is on average 0.2K colder than the buoys (Donlon et al. 2002) with a small wind speed dependence (Gentemann & Minnett 2008, Figure 4). In addition, the models refer to the daily mean temperature at about 2 meter depth, while the AIRS SST refer to the skin temperature at the overpass time. On average, the daily mean buoy temperature is 0.21K warmer for the 1:30 PM, 0.11K colder for the 1:30 AM overpasses of the 30S-30N zone (Kennedy et al. 2007), but there is a considerable seasonal and zonal variability. It is this seasonal variability which dominates the difference between the AIRS and CrIS surface SST and the SST at 2 meter below the surface in

the models.

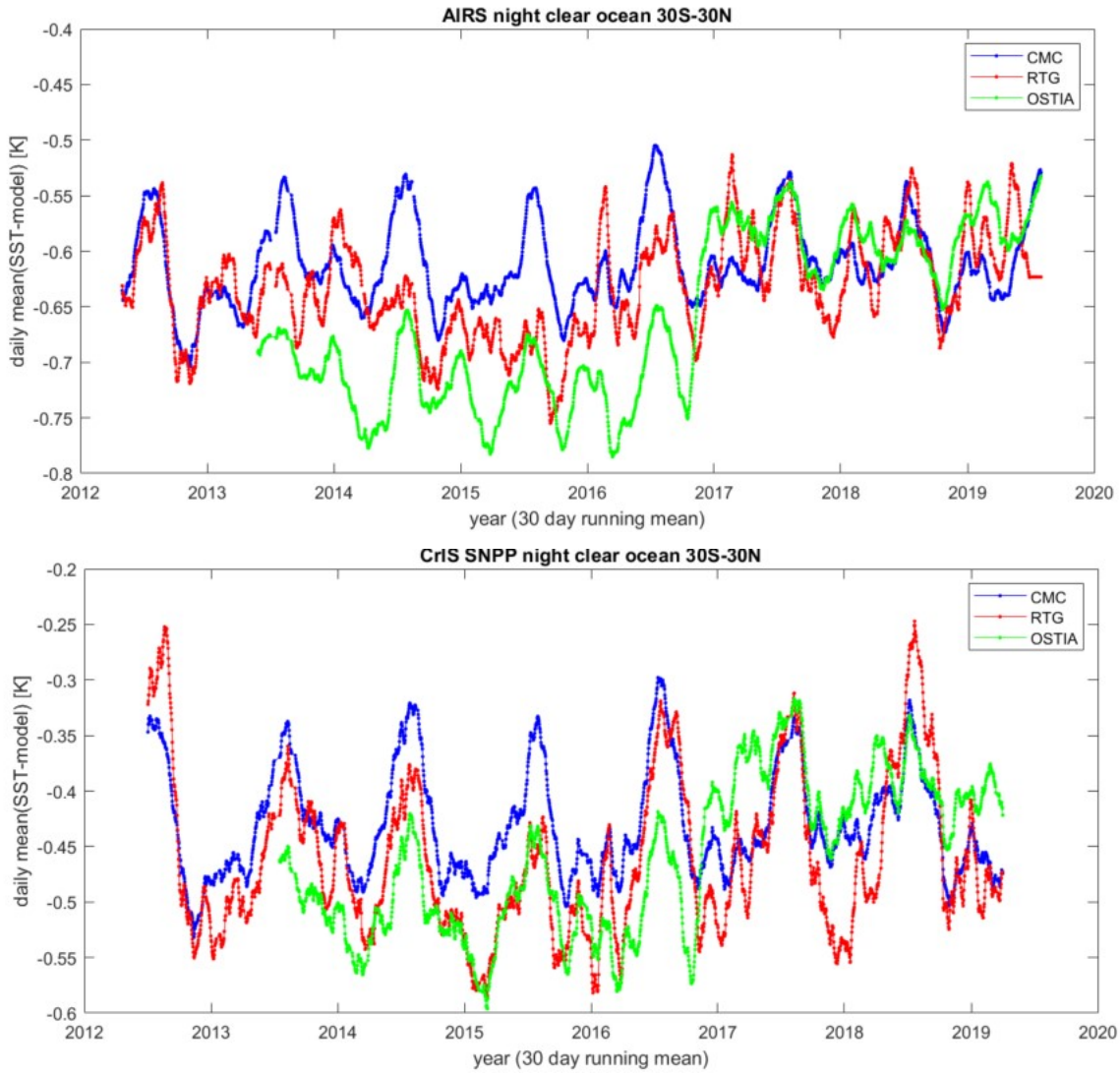

Figure 1. Time series of the daily mean(sst1231-model) for AIRS (top) and CrIS (bottom) for the night overpasses of the 30S to 30N latitude zone of the oceans.

However, the amplitude of the modulation changed between the model and over the years due to version changes related to the changes in the input data and other changes to the algorithms. Between 2013 and 2019 the peak-to-peak (p-p) amplitude relative to AIRS was about 0.1K for the OSTIA, 0.15K p-p for the CMC, decreasing to 0.1K in 2016. The RTG modulation was less regular. The seasonal modulation decreased in the CMC (blue) in 2017, with a 15 mK shift in the bias, likely associated with a version change. For the RTG (red) the seasonal modulation doubled starting in 2016, with little change in the mean bias. For the OSTIA (green) the bias after 2017

shifted by about 100 mK with little change in the seasonal modulation.

Figure 2 (top) shows the time series of stddev(AIRS-model) for the night overpasses of the 30S to 30N latitude zone of the oceans, smoothed with a 30 day running mean. The stddev (AIRS-OSTIA) (green) and stddev (AIRS-CMC) are very similar, with neither showing the prominent seasonal pattern in the mean. Figure 2 (bottom) shows the time series of stddev(SST-model) for the CrIS night overpasses of the 30S to 30N latitude zone of the oceans. The stddev(CrIS-CMC), averaged for the 5 years, is 0.51K, compared to 0.44K for AIRS.

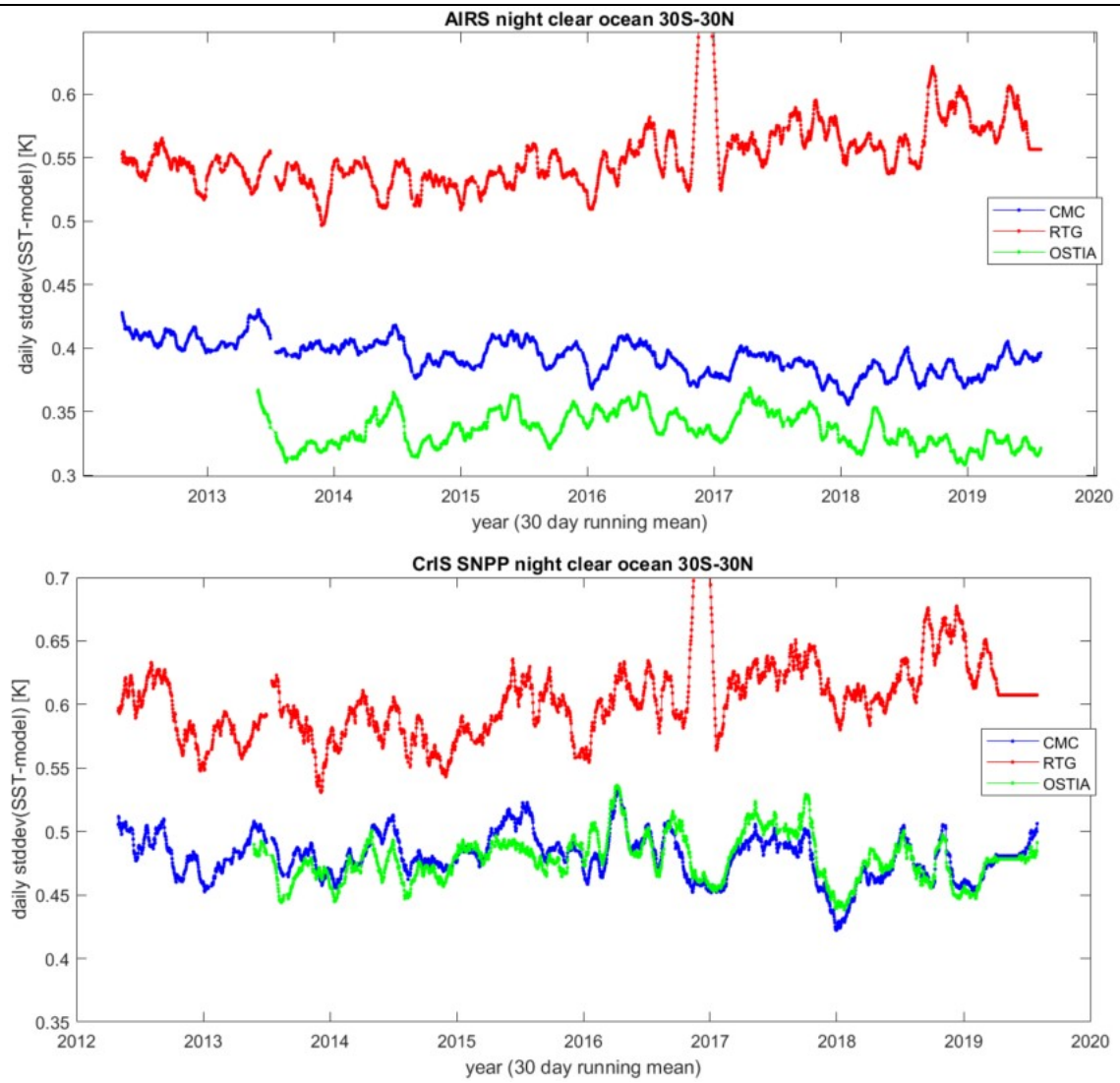

Figure 2 Time series of stddev(SST-model) for the AIRS and CrIS night overpasses of the 30S to 30N latitude zone of the oceans. Note the difference in the y-scale for AIRS and CrIS.

The general character, differences and evidence of version changes discussed for the night overpasses of the 30N-30S oceans are seen for the day and night overpasses of other latitude bands. With increasing distance from the equator the seasonal modulation of the diurnal cycle of the difference between the skin and the 2 meter temperatures becomes increasingly more pronounced. Due to the similarity of the results from AIRS and CrIS, we show only the AIRS time series. All results are summarized in Table 1. Figures 3 and 4 shows the time series for the night overpasses of latitudes between 30N-50N and between 30S-50S.

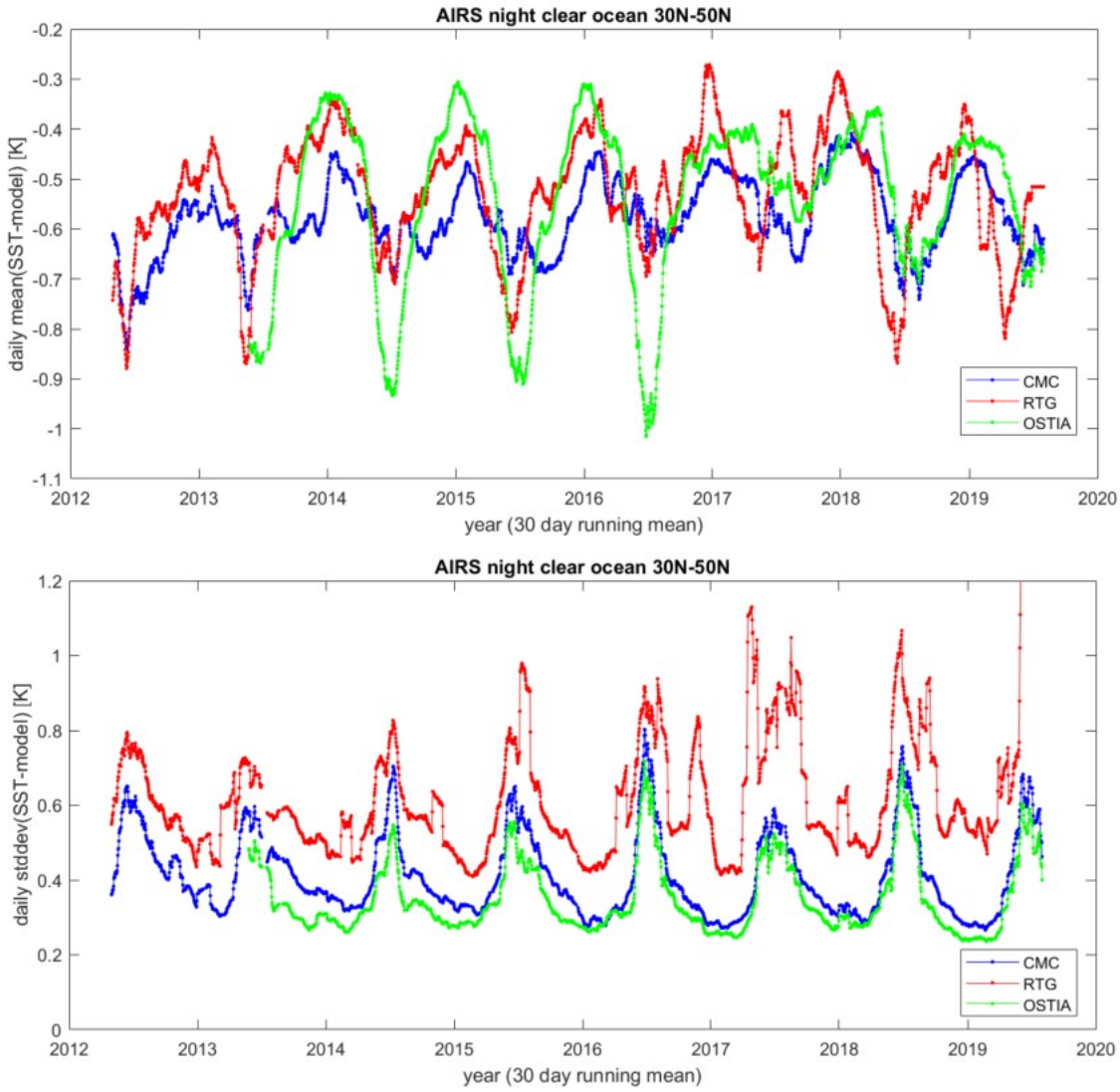

Figure 3. Time series of the daily mean and stddev of (sst1231-model) for AIRS for the night overpasses of the latitude zones 30N-50N.

Version changes make the interpretation of the anomaly trends relative to the RTG or the OSTIA even more limited. For this reason Table 1 a. and b. summarizes anomaly trends and bias only for AIRS-CMC, and CrIS-CMC as function of the latitude bands and day and night. AIRS and CrIS both shows a pattern of a positive trend relative to the CMC above the equator, negative below the equator.

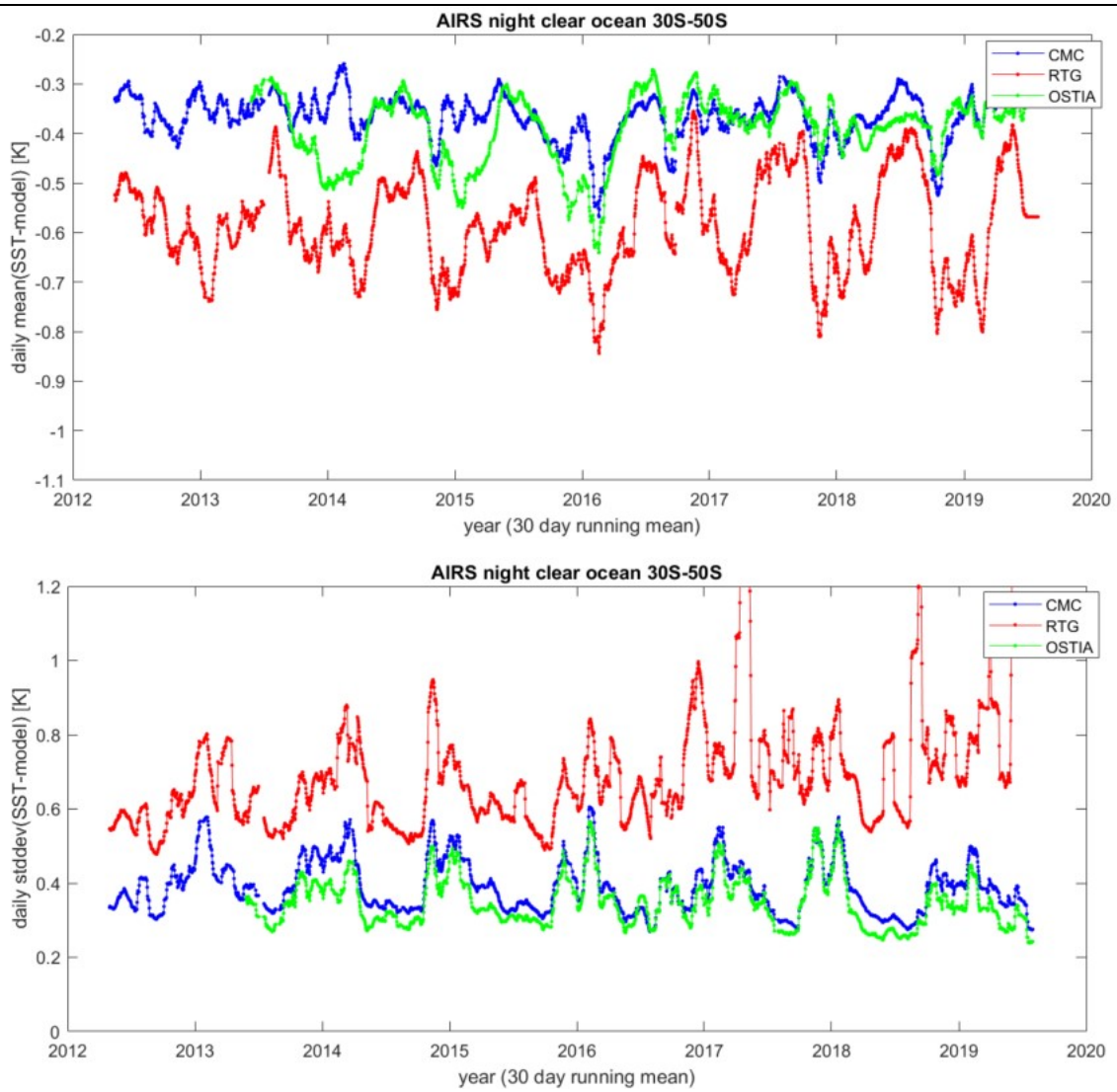

Figure 4. Time series of the daily mean and stddev of (sst1231-model) for AIRS for the night overpasses of the latitude zones between 30S-50S.

AIRS-CMC.airs	Day trend [mK/yr]	Day bias [K]	Day count	Night trend [mK/yr]	Night bias [K]	Night count
30-50N	+ 8.4 ± 1.6	-0.23	2154	+10.5 ± 1.5	-0.56	1848
0-30N	+ 6.9 ± 1.1	-0.21	6935	+ 6.3 ± 0.7	-0.62	4623
30S-0	+ 1.5 ± 0.9	-0.21	8187	+ 0.6 ± 0.8	-0.60	4177
50S-30S	-4.6 ± 1.2	-0.07	2773	-3.2 ± 0.9	-0.34	1892

a)

CrIS-CMC.cris	Day trend [mK/yr]	Day bias [K]	Day count	Night trend [mK/yr]	Night bias [K]	Night count
30-50N	+ 7.5 ± 1.7	+ 0.04	1771	+ 7.1 ± 2.0	-0.28	1497
0-30N	+ 4.2 ± 1.5	-0.03	6153	+ 3.6 ± 1.4	-0.42	3920
30S-0	+ 1.0 ± 1.2	-0.08	7274	+ 0.3 ± 1.3	-0.15	2231
50S-30S	-5.1 ± 143	+0.22	2200	-2.2 ± 1.4	-0.06	1514

b)

Table 1. Anomaly trends and mean in a) (AIRS-CMC.airs) and b) (CrIS-CMC.cris) , and daily mean sample count. The number following ± is the 1 sigma confidence uncertainty of the trend.

2. Direct AIRS/CrIS comparison.

Between May 2012 and March 2019 AIRS and CrIS data were both available. This allow a direct comparison of AIRS and CrIS derived SST. We already noted that the typically 40,000 clear footprints identified each day are from different locations. In the analysis of daily mean of the double difference (AIRS-CMC.airs) minus (CrIS-CMC.cris), where CMC.airs and CMC.cris are the CMC SST values closest to the AIRS and CrIS locations, respectively, the location differences and model artifacts cancel. However, the time series of the double difference for the 30S-30N latitude zones, day (red), night (blue), Figure 5, still shows a seasonal pattern, more pronounced at night than during the day. This requires further analysis. Table 2 shows the anomaly trend and bias for the AIRS-CrIS double difference.

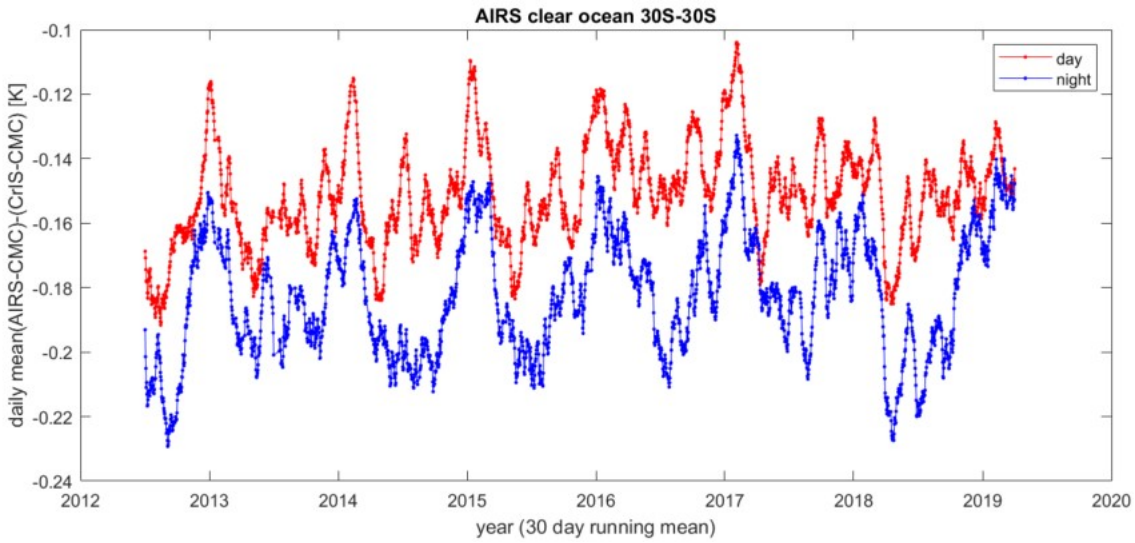

Figure 5. (AIRS-CMC.airs)- (CrIS-CMS.cris) double difference

(AIRS-CMC.airs) –(CrIS-CMC.cris)	Day trend [mK/yr]	Mean Day bias [K]	Mean Day count	Night trend [mK/yr]	Mean Night bias [K]	Mean Night count
30-50N	+ 0.9 ± 1.2	-0.27	1771	+ 3.1 ± 1.4	-0.28	1497
0-30N	+ 3.0 ± 0.9	-0.17	6153	+ 2.3 ± 1.1	-0.20	3920
30S-0	+ 0.7 ± 0.7	-0.14	7274	+ 0.0 ± 1.0	-0.17	2231
50S-30S	+ 0.4 ± 0.9	-0.29	2200	-1.4 ± 1.0	-0.29	1514

Table 2. (AIRS-CMC.airs) –(CrIS-CMC.cris) double difference mean and trends. The number following ± is the 1 sigma confidence uncertainty of the trend.

4. Discussion

The daily stddev(AIRS-CMC) and (AIRS-OSTIA), 0.4 K average between 2013 and 2019, indicate a level of agreement between the models and independent SST observations seen previously only relative to the Argo buoys in the 2000-2011 period (Fiedler et al. 2019). The AIRS stddev relative to OSTIA and CMC shows a small but steady decrease, presumably related to improvements in the models. A degradation of the RTG is seen as the steadily increasing stddev starting in 2016, until the RTG production was ended in 2019.

The evaluation of trends from the data between 2013 and 2019 is impacted by version changes in the CMC, start-up problem and version changes in the OSTIA, and the steady degradation of the RTG throughout our analysis period. Even the small, previously noted 15 mK version change related shift in the CMC would create a 3 mK/yr trend artifact between 2013 and 2019. Table 1 shows that the one sigma bias anomaly trend uncertainties for AIRS and CrIS relative to the CMC range from +8 mK/yr to -5 mK/yr depending on latitude. For a longer time series, i.e. AIRS-CMC between 2002 and 2020, the 15 mK shift due to a version change in 2017 would only create a 1 mK/yr trend artifact, i.e. trends larger than 1 mK/yr would be significant.

Trends in the AIRS-CMC and CrIS-CMC are dominated by CMC version changes. In trends based on the double difference, version changes in the CMC cancel, and the AIRS-CrIS trends range from +3 and -1 mK/yr, relatively latitude zone and day/night consistent. These trends are very low, but pertain only to the channels used for the SST derivation. There is no assurance that they are valid for all channels and the full dynamic range of the scenes encountered by AIRS and CrIS.

CrIS is typically about 200 mK warmer than AIRS at all latitudes. This difference could be related to the combination of a bias in the absolute calibration of AIRS and CrIS and a difference in the RTM used for the water vapor correction of the SST. The 200 mK difference between AIRS and CrIS appears to be dominated by an absolute calibration difference, consistent in magnitude with expectations.

Calibration bias: AIRS is a grating array spectrometer. As such the SST is calculated from the signal measured by a single detector at 1231.3 cm^{-1} (8 microns). Pagano et al. (2018, Figure 4) show a 2-sigma absolute radiometric uncertainty for the channels near 8 microns of 80 mK for scene temperatures in the 280-300K range for the L1b V5 calibration. CrIS is an interferometer with a 3x3 focal plane array, where 9 independent detectors measure the signal for a wide wavelength range. In the previous figures we showed the daily mean from CrIS averaged over all nine detectors. If we instead were to calculate the daily mean for each detector, we find a 140 mK day/night independent difference between the warmest and coldest detector. The difference between AIRS and CrIS seen in our data could thus be simply related to the absolute calibration

difference.

Water vapor correction: We used SARTA with the latest CKHD 3.2 water continuum. For CrIS we used the CRTM 2.1.3 with a 2008 vintage MT_CK 2.4. The effect of the water vapor absorption on for the CrIS bt1232 is typically 3.5K. The water vapor continuum in MTCK 2.4 was decreased in more recent editions. CRTM would have to be 6% too strong to create a warm bias in sst1232 of 200 mK (3.5*0.06=0.2K). A 2% overcorrection is non unlikely, which would explain 70 mK of the difference between AIRS and CrIS. A CRTM overcorrection of the water vapor continuum may explain the consistently slightly larger stddev of CRIS relative to the CMC and OSTIA models.

4. Summary

The mean and standard deviation of SST models relative to independent observation with the SST derived from AIRS or CrIS under clear conditions are an effective way of evaluating the fidelity and trends of SST models. AIRS and CrIS show very similar patterns relative to the models. CrIS being 200 mK warmer than AIRS appears to be consistent with the combination of expected water vapor continuum correction and absolute calibration differences. The CMC and OSTIA are major improvements over the RTG. The daily standard deviation of the difference between the AIRS, the CMC and the OSTIA of 0.4 K is a level previously only seen in SST models relative to the Argo buoys. Between 2013 and 2019, all three models show artifacts related to version changes, which limit the interpretation of zonal trend from the 2013-2019 time period. Only the CMC is available for the 2013 – 2019 AIRS and CrIS availability. The 5 year anomaly trends in (AIRS-CrIS) double difference avoid the effect of version changes. There we find trends between at +3 and -1 mK/yr. This requires additional analysis.

Acknowledgments

This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA. Sergio DeSouza-Machado, U. Maryland Baltimore County, provided support with the SARTA RTM. We used the prototype of the daily AIRS Calibration Data Subset (ACDS v7), which will be available free of charge from the GSFC DISC. Currently the

ACDS v5 is available at https://disc.gsfc.nasa.gov/datasets/AIRXBCAL_005/summary?keywords=AIRXBCAL. An equivalent CrIS Calibration data subset, conceptually equivalent to the ACDS v7, will be available from the GSFC/DISC soon.

Between 2001 and 2016, the RTGSST was freely available on a 0.5 degree grid via FTP at ftp://polar.ncep.noaa.gov/pub/history/sst/rtg_low_res/, for example, rtg_sst_grb_0.5.20020831. More recently it was available from https://polar.ncep.noaa.gov/sst/rtg_high_res/. As of 2019 its production has been discontinued. The CMC and OSTIA data are freely available from https://podaac.jpl.nasa.gov/dataset/CMC0.2deg-CMC-L4-GLOB-v2.0?ids=&values=&search=CMC and https://podaac.jpl.nasa.gov/dataset/OSTIA-UKMO-L4-GLOB-v2.0?ids=&values=&search=%22OSTIA-UKMO-L4-GLOB-v2.0%22

References

Aumann, H.H. , M.T. Chahine, C. Gautier, M. Goldberg, E. Kalnay, L. McMillin, H. Revercomb , P.W. Rosenkranz , W. L. Smith , D. H. Staelin, L. Strow and J. Susskind, (2003), “AIRS/AMSU/HSB on the Aqua Mission: Design, Science Objectives, Data Products and Processing Systems”, IEEE Trans. Geosci. Remote Sens., Vol.41, 253-264.

Aumann, H. H., Broberg, S., Manning, E., & Pagano, T. (2019). Radiometric Stability Validation of 17 Years of AIRS Data Using Sea Surface Temperatures. *Geophysical Research Letters*, 46. https://doi.org/10.1029/2019GL085098

Brasnett B., 2008. The impact of satellite retrievals in a global sea-surface-temperature analysis. Q.J.R. Meteorol. Soc., 134, 1745-1760. doi: 10.1002/qj.319.

Blumstein, D., Chalon, G., Carlier, T., Buil, C., Hebert, Ph., Maciaszek, T., Ponce, G., Phulpin, T., Tournier, B.,Simeoni, D., Astruc, P., Clauss, A., Kayal, G., and Jegou, R., "IASI instrument: Technical overview and measured performances," Infrared Spaceborne Remote Sensing XII, M.Strojnik Editor, Proc SPIE **5543**, 196-207 (2008).

Donlon, C.J., P.J. Minnett, C. Gentemann, T.J. Nightingale, I.J. Barton, B. Ward and M.J. Murray (2002), ”Toward Improved Validation of Satellite Sea Surface Skin Temperature Measurements for Climate Research’, J. Climate, 15, 353-369.

Fiedler, E.K., A. McLaren, V. Banzon, B. Brasnett, S. Ishizaki, J. Kennedy, N. Rayner, J. Roberts_Jones, G. Corlett, C. Merchant and C. Donlon (2019) “Intercomparison of long-term sea surface temperature analysis using the GHRSST Multi-Product Ensemble (GMPE) system”, Remote Sensing of Environment 222, 18-23 https://doi.org/10.1016/j.se.2019.12.015

Gentemann, C. L., and P. J. Minnett (2008), “Radiometric measurements of ocean surface thermal variability”, J. Geophys. Res., 113, C08017, doi:10.1029/2007JC004540.

Kennedy, J.J., P. Brohan and S. F. B. Tett (2007) “A global climatology of the diurnal variations in sea-surface temperature and implications for MSU temperature trends”, GRL, VOL. 34, L05712, doi:10.1029/2006GL028920.

Pagano, T.S., H.H. Aumann, S. Broberg, E. Mannng , K. Overoye, and M. Weiler (2018)”Update to the Absolute Radiometric Accuracy of AIRS on Aqua”, Asia-Pacific Remote Sensing 2018.

Strow, L. L., S.E. Hannon, S. DeSouza-Machado, H. E. Mottler and D. C. Tobin (2006), Validation of the atmospheric infrared sounder radiative transfer algorithm, JGR 111 D09S06 doi: 10.1029/2005JD006146.

Thiebaux, Jean, B. Katz, and Wanqui Wang (2002) “New sea-surface temperature analysis implemented at NCEP”, Ocean Marine Branch contribution NCEP/OMB No. 197.