Atmos. Meas. Tech., 10, 351–371, 2017
www.atmos-meas-tech.net/10/351/2017/
doi:10.5194/amt-10-351-2017

Atmospheric Measurement Techniques
Open Access
EGU

Global clear-sky surface skin temperature from multiple satellites using a single-channel algorithm with angular anisotropy corrections

Benjamin R. Scarino[1], **Patrick Minnis**[2], **Thad Chee**[1], **Kristopher M. Bedka**[2], **Christopher R. Yost**[1], **and Rabindra Palikonda**[1]

[1]Science Systems and Applications, Inc., 1 Enterprise Parkway, Suite 200, Hampton, VA 23666, USA
[2]NASA Langley Research Center, 21 Langley Blvd MS 420, Hampton, VA 23681-2199, USA

Correspondence to: Benjamin R. Scarino (benjamin.r.scarino@nasa.gov)

Received: 10 March 2016 – Published in Atmos. Meas. Tech. Discuss.: 5 April 2016
Revised: 22 November 2016 – Accepted: 19 December 2016 – Published: 27 January 2017

Abstract. Surface skin temperature (T_s) is an important parameter for characterizing the energy exchange at the ground/water–atmosphere interface. The Satellite ClOud and Radiation Property retrieval System (SatCORPS) employs a single-channel thermal-infrared (TIR) method to retrieve T_s over clear-sky land and ocean surfaces from data taken by geostationary Earth orbit (GEO) and low Earth orbit (LEO) satellite imagers. GEO satellites can provide somewhat continuous estimates of T_s over the diurnal cycle in non-polar regions, while polar T_s retrievals from LEO imagers, such as the Advanced Very High Resolution Radiometer (AVHRR), can complement the GEO measurements. The combined global coverage of remotely sensed T_s, along with accompanying cloud and surface radiation parameters, produced in near-realtime and from historical satellite data, should be beneficial for both weather and climate applications. For example, near-realtime hourly T_s observations can be assimilated in high-temporal-resolution numerical weather prediction models and historical observations can be used for validation or assimilation of climate models. Key drawbacks to the utility of TIR-derived T_s data include the limitation to clear-sky conditions, the reliance on a particular set of analyses/reanalyses necessary for atmospheric corrections, and the dependence on viewing and illumination angles. Therefore, T_s validation with established references is essential, as is proper evaluation of T_s sensitivity to atmospheric correction source.

This article presents improvements on the NASA Langley GEO satellite and AVHRR TIR-based T_s product that is derived using a single-channel technique. The resulting clear-sky skin temperature values are validated with surface references and independent satellite products. Furthermore, an empirically adjusted theoretical model of satellite land surface temperature (LST) angular anisotropy is tested to improve satellite LST retrievals. Application of the anisotropic correction yields reduced mean bias and improved precision of GOES-13 LST relative to independent Moderate-resolution Imaging Spectroradiometer (MYD11_L2) LST and Atmospheric Radiation Measurement Program ground station measurements. It also significantly reduces inter-satellite differences between LSTs retrieved simultaneously from two different imagers. The implementation of these universal corrections into the SatCORPS product can yield significant improvement in near-global-scale, near-realtime, satellite-based LST measurements. The immediate availability and broad coverage of these skin temperature observations should prove valuable to modelers and climate researchers looking for improved forecasts and better understanding of the global climate model.

1 Introduction

Surface skin temperature (T_s) is a critical quantity for characterizing the exchange of energy between the Earth's surface and the atmosphere. Consistent land and ocean measurements of T_s are essential for regional and global climate assessment and weather model data assimilation. Surface en-

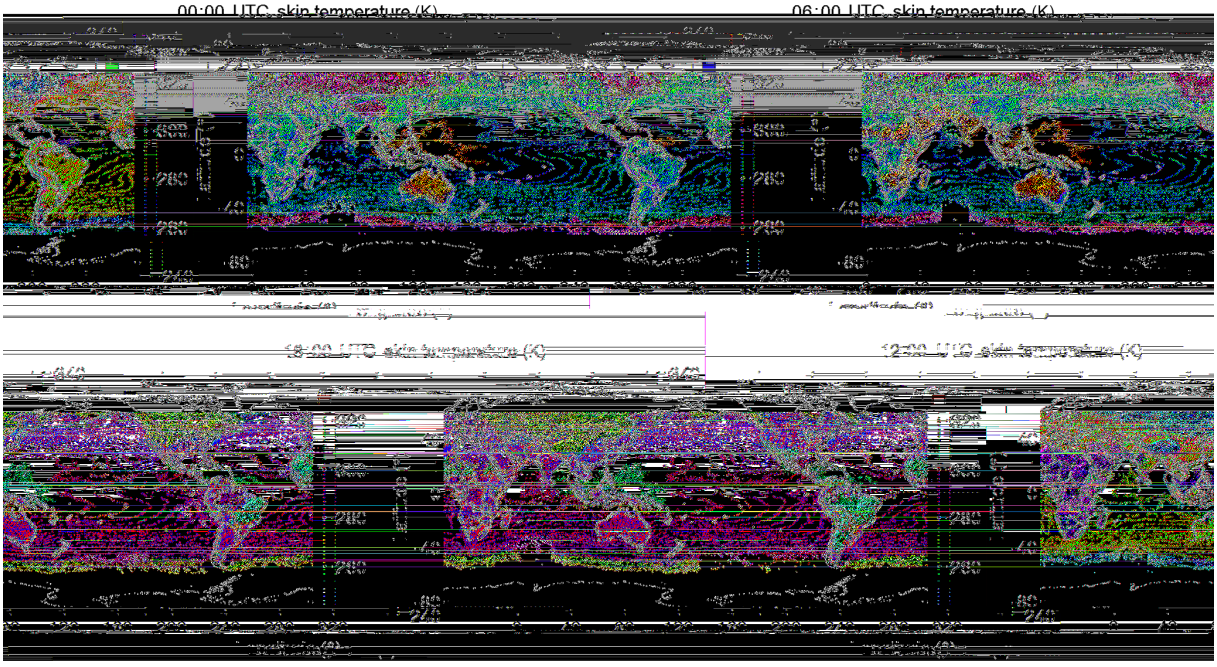

Figure 1. Mean merged, clear-sky surface skin temperature values from GOES-East, GOES-West, Meteosat-9, MTSAT-2, and INSAT-3D, October 2015.

ergy balance and top-of-atmosphere (TOA) radiative budget calculations rely on the accuracy of these surface parameters (Bodas-Salcedo et al., 2008). In addition to surface flux analyses, T_s retrievals are used to minimize model prediction uncertainty by updating model state values with observations at regular time steps – an important consideration for climate and numerical weather prediction (NWP) models (Garand, 2003 Tsuang et al., 2008 Reichle et al., 2010 Ghent et al., 2010 Guillevic et al., 2012 Draper et al., 2015). The modeling community could benefit significantly from the provision of frequent, spatially contiguous, global land and ocean T_s data (Rodell et al., 2004 Bosilovich et al., 2007). Many other uses of T_s as well as the status and future of T_s retrievals are summarized by Li et al. (2013). It is clear that the need is growing for higher accuracy, global coverage, and greater temporal and spatial resolution of T_s retrievals from satellite imager data.

Satellite-based T_s retrieval, validation, and modeling studies originate from a variety of sources, e.g., the National Environmental Satellite, Data, and Information Service (NESDIS) or the National Oceanic and Atmospheric Administration (NOAA) via the Advanced Very High Resolution Radiometer (AVHRR) series and the Geostationary Operational Environmental Satellite (GOES) sensors (Prata, 1993, 1994 Coll and Caselles, 1997 Sobrino and Raissouni, 2000 Kerr et al., 2004 Sobrino et al., 2004 Yu et al., 2009, 2012a, b Sun et al., 2012). Specifically, using two different single-channel land surface temperature (LST) algorithms, Heidinger et al. (2013) and Minnis et al. (2016) found good agreement with the NOAA ESRL Surface Radiation (SURFRAD) network in verification studies using LST retrievals from GOES and AVHRR alone, respectively. Furthermore, near-realtime (NRT) LST is produced operationally from Meteosat Spinning Enhanced Visible and Infrared Imager (SEVIRI) data, which offer continuous coverage of Europe and Africa, and served as the focus of several LST validation studies (Sobrino and Romaguera, 2004 DaCamara, 2006 Kabsch et al., 2008 Trigo et al., 2008 Göttsche et al., 2013). Retrievals using radiances from the Moderate Resolution Imaging Spectroradiometer (MODIS) have been both the target and standard for a number of LST verification studies (Wan et al., 2002, 2004, 2008 Coll et al., 2009 Jiménez et al., 2012). Duan et al. (2014) used four daily observations from Terra and Aqua MODIS to capture the diurnal cycle of LST, which is critical for full characterization of the climate system. Wang et al. (2014) conducted a three-way T_s comparison using MODIS, in situ ground observations, and model simulations. They note the high importance of accurate cloud-clearing and the inherit difficulties of resolution scaling when comparisons are conducted between satellite data and point references – conclusions supported in a similar MODIS daytime LST verification study conducted by Williamson et al. (2013).

With more reliable calibrations, operational geostationary Earth orbiting (GEO) and low Earth orbiting (LEO) satellite imagers are being used to derive cloud and radiation properties in NRT, e.g., Minnis et al. (2008a). The combination of GOES-East (GOES-13), GOES-West (GOES-15), Me-

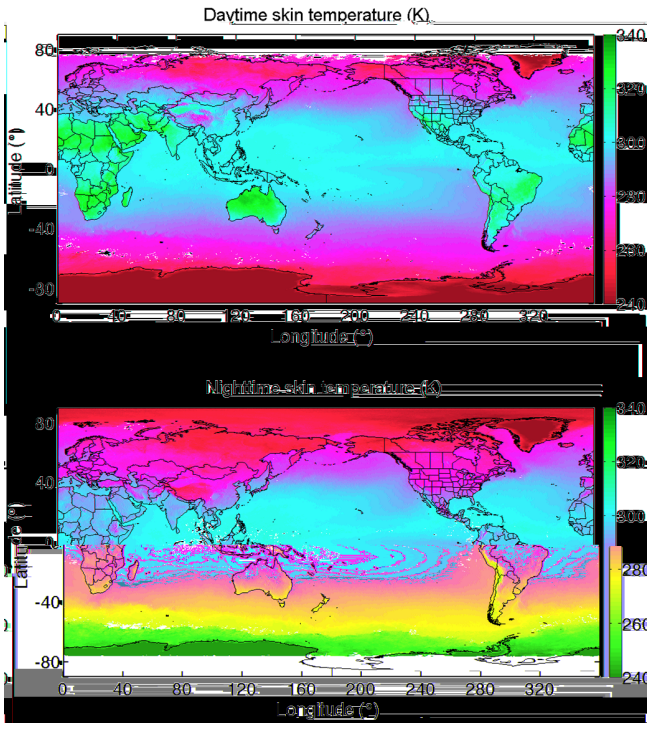

Figure 2. Average surface skin temperature from NOAA-19 AVHRR, October 2013.

teosat Second Generation (MSG Meteosat-9 or Meteosat-10), MTSAT-2 (recently replaced by Himawari-8), and the Indian Space Research Organization INSAT-3D provides high-temporal-resolution (1 h nominal) quasi-global T_s data produced in NRT, with a shared single-channel retrieval algorithm (e.g., Fig. 1). The methodology (Sect. 3) is flexible and easily transportable to other GEO and LEO imagers, including the current AVHRR instruments on the NOAA and EUMETSAT MetOp platforms. AVHRR T_s retrievals supplement the GEO data and fill in missing measurements over polar regions (e.g., Fig. 2). This same method is being applied to historic and current imager datasets, particularly as part of the Satellite ClOud and Radiative Property retrieval System (SatCORPS) analyses of AVHRR data for provision of a NOAA Climate Data Record (Minnis et al., 2016), and for MODIS, GEO, and Suomi National Polar-orbiting Partnership (S-NPP) Visible Infrared Imaging Radiometer Suite data as part of the Clouds and Earth's Radiant Energy System (CERES) project (e.g., Minnis et al., 2010).

This article highlights recent improvements made to the SatCORPS NRT satellite T_s product (Scarino et al., 2013), via comparisons of GOES and AVHRR T_s retrievals with established sea surface temperature (SST) and LST reference datasets. The influence of NWP source on retrieved T_s values is also examined. The main improvements over the earlier version are enhanced pixel-level resolution output and hourly GEO retrieval time steps. The SatCORPS T_s retrieved from GOES and AVHRR data is evaluated by comparing with reference datasets based on in situ and satellite measurements. Section 2 provides an overview of these product and reference datasets, as well descriptions of ancillary validation sets and how the reanalysis input is configured. Explanation of the single-channel T_s retrieval algorithm is provided in Sect. 3.

The results and discussion are presented together in two sections: one for SST and the other for LST. Section 4 focuses only on SatCORPS SST results because the validation and angular corrections differ from those used for LST. Validation of the LST and application of the theoretical model developed by Vinnikov et al. (2012) comprise the main topics of Sect. 5. Coefficients for the Vinnikov et al. (2012) model were determined empirically from near-simultaneous measurements covering a limited viewing range from a small collection of sites. Because those sites represent varied surface and climate types, the resulting coefficients could potentially serve as an effective initial step in the process of correcting LST for angular dependency. The included subsections then highlight our independent assessment of the Vinnikov et al. (2012) model over a large area, starting with details on the origin and use of the correction model (Sect. 5.1), testing of its broad influence (Sect. 5.2) with validation results/discussion relative to satellite (Sect. 5.3) and ground references (Sect. 5.4), and an uncertainty discussion related to the spatial homogeneity surrounding the ground reference (Sect. 5.5). Finally, Sect. 6 summarizes the main conclusions. The combined GEO and AVHRR retrievals allow for high-resolution temporal monitoring of the T_s diurnal cycle, an essential state variable for numerical weather model data assimilation and climate studies (e.g., Draper et al., 2015). The T_s products and uncertainties described here should be valuable for improving surface energy flux analyses and numerical weather prediction due to their NRT global availability over land and ocean.

2 Data

2.1 Satellite data for surface skin temperature retrieval

Clear-sky surface skin temperature is retrieved from channel 4 (11 µm) radiances taken by the NOAA-18 or NOAA-19 AVHRR for the period January–December 2008 or 2013, respectively, in the Global Area Coverage (GAC) format. The nominal satellite equatorial crossing time is 13:30 LT. A GAC pixel radiance is formed by averaging the radiances of four consecutive raw 1 km AVHRR pixels along the scan direction. The process is repeated after skipping the fifth pixel and so on to produce consecutive GAC pixels along the scan line. Two scan lines are then skipped and the pixel averaging is applied again to the third scan line. Thus, a GAC pixel nominally covers a 1 km × 4 km area (a 4 km^2 pixel) but, because of sampling, represents a 3 km × 5 km area that yields

an effective resolution of ∼ 4 km. The AVHRR data were analyzed with the SatCORPS-A1 methodology (Minnis et al., 2016) to retrieve cloud properties, TOA broadband fluxes, and clear-sky surface skin temperature. Clear pixels are determined from the SatCORPS cloud mask. Details of the skin temperature retrieval process are given in Sect. 3.

Hourly channel-4 (10.8 µm) data from GOES-13 (GOES-East) and GOES-15 (GOES-West) taken during January, April, July, and October (hereafter, JAJO) 2013 are used to retrieve T_s for validation with surface and other satellite surface skin temperature datasets. Furthermore, GOES-13 and GOES-15 data are employed to test the angular anisotropy parameterization. The nominal GOES imager resolution is 4 km. The pixels are sub-sampled, however, to an effective resolution of 8 km during full disk and hourly hemispheric scans. That is, every other pixel is skipped during realtime full disk and hemispheric processing to improve computational speed and produce more manageable output file size. The actual pixel measurements and geolocation attributes, however, are still representative of a $4 \times 4\,km^2$ area. These data were analyzed with a version of SatCORPS-A1 adapted to the GOES channels as described by Minnis et al. (2008a).

2.2 Validation data

For validation comparisons, this study employs surface and satellite-based references. The SatCORPS AVHRR SST values are compared to the daily high-resolution blended SST analysis described by Reynolds et al. (2007). It comprises the NOAA "Optimum Interpolation" SST (OI SST) version 2 high-resolution dataset, which consists of a global $0.25° \times 0.25°$ grid of blended satellite (AVHRR two- and three-channel algorithms and Advanced Microwave Sounding Radiometer (until 2011) data) and in situ measurements of daily SST. It covers the period from January 1981 to the present.

The version 5 Aqua MODIS LST/Emissivity product (MYD11_L2 hereafter, MYD11), which is derived from the generalized split-window algorithm (Wan and Dozier, 1996 Wan and Li, 1997 Snyder and Wan, 1998), is used to validate the SatCORPS AVHRR and GOES LST values. The dataset includes values of LST retrieved from clear-sky 1 km MODIS pixels and surface spectral emissivity values. Because MYD11 is derived from different data using a different type of algorithm, and is accurate to ±1 K or less (Wan, 2008 Wan et al., 2002, 2004), it serves well as an independent reference for comparing with the GOES retrievals.

Surface radiometer measurements from the Atmospheric Radiation Measurement (ARM) Southern Great Plains (SGP) Central Facility (36.3° N, 97.5° W) 11 µm upwelling/downwelling infrared thermometer (IRT Morris, 2006) serve as another LST validation source. The ARM IRT ground-based radiation pyrometers provide measurements of the equivalent blackbody brightness temperature for the 9.6–11.5 µm spectral band every 60 s. From a 10 m height with 30.5° FOV, the upwelling IRT, with a specified accuracy of ±0.5 K, measures the effective ground radiating temperature, i.e., the temperature equivalent of the ground infrared radiant energy assuming the surface emissivity (ε_s) is equal to 1.0 (Morris, 2006). A true skin temperature T_s can, therefore, be determined as

$$T_s = B^{-1}\left\{\frac{\left[B\left(T_o\right) - \left(1-\varepsilon_s\right) \times B\left(T_{o\downarrow}\right)\right]}{\varepsilon_s}\right\}, \quad (1)$$

where ε_s is from the CERES 11 µm database (e.g., Chen et al., 2004) and the spectral downwelling narrowband brightness temperature ($T_{o\downarrow}$) is measured by a 2 m height uplooking IRT, which is oriented so that the zenith view of the sky is reflected into the lens by a gold mirror, and has a narrow 2.64° FOV (Morris, 2006). The Planck function for the particular waveband is $B(T)$, and T_o is temperature equivalent to the surface-leaving blackbody radiance. Note that the ARM downwelling IRT at the Lamont, OK, Central Facility was no longer operating in 2013, and therefore $T_{o\downarrow}$ was acquired from the nearby Lamont, OK, Extended Facility downwelling IRT, which operates in unison with the Central Facility instrument. It is expected that there is negligible variation in $T_{o\downarrow}$ over the ∼ 200 m distance between the two sites.

2.3 Supplementary ASTER data

Following the studies of Wang and Liang (2009) and Guillevic et al. (2014), high-resolution Terra Advance Spaceborne Thermal Emission and Reflection (ASTER) LST and emissivity (AST_08 and AST_05, respectively) product data from 2001 through 2015 (available complete years) are used to measure the spatial homogeneity of LST in both a 4×4 and $8 \times 8\,km^2$ area centered on the ARM SGP ground station. The ASTER LST product has a 90 m spatial resolution at nadir, derived from five infrared channels using the temperature–emissivity separation (TES) method (Yamaguchi et al., 1998 Gillespie et al., 1998). Each ASTER granule consists of 700 × 830 LST pixels, which can be referenced to two 11 × 11 matrices of geocentric latitude and geodetic longitude. Bilinear interpolation is used to estimate latitude and longitude for each LST pixel. These LST values are subsetted into blocks of 45 × 45 pixels to simulate the spatial extent of a 4 km GOES pixel centered on the Central Facility and to assess the spatial representativeness of the site relative to the surrounding region. A subset block of 89 × 89 pixels is meant to represent the worst possible disparity between the satellite measurement and the ground station based on the most extreme pixel-to-point matches possible, i.e., the ARM site being situated in any corner of the GOES pixel.

2.4 Reanalysis input

Numerical weather model output parameters are used as input to compute TOA brightness temperatures (T_{TOA}). These

include the model surface air (T_a') and skin (T_s') temperatures, and vertical temperature and humidity profiles. The realtime GEO retrievals employ National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS EMC, 2003) model forecasts accessed from the Man-computer Interactive Data Analysis System (McIDAS Lazzara et al., 1999). Non-realtime GEO studies utilize either GFS or Modern-Era Retrospective Analysis for Research and Applications (MERRA Rienecker et al., 2011) reanalyses. The impacts of using one reanalysis or the other are examined by analyzing the same satellite data using each of the two reanalyses during the T_s retrieval.

MERRA data have a spatial resolution of 0.5° latitude × 0.66° longitude over the globe. The surface skin temperature is available hourly, while the temperature and humidity profiles are provided every 6 h. A total of 43 atmospheric layers are used. The version of GFS used here has a 1.25° horizontal resolution and up to 11 levels in the vertical, and it provides data every 6 h. No model values of T_s' are available in the GFS version over land, so T_s' is estimated from T_a' as a function of local time and season.

3 Single-channel skin temperature retrieval

The method for calculating T_s from 11 µm T_{TOA} observations is an updated, higher-resolution version of that described by Scarino et al. (2013). Because some imagers (e.g., AVHRR-1, GOES-13) lack split-window capabilities, the single-channel method best allows historical consistency in application amongst many distinct sensors (Sun and Pinker, 2003 Jiménez-Muñoz and Sobrino, 2010 Heidinger et al., 2013). The process to determine T_s first employs the cloud mask algorithm developed for CERES to classify pixels as cloudy or clear on a chosen grid (Minnis et al., 2008b). The algorithm relies on comparisons of observations with estimates of the clear-sky T_{TOA} or reflectance at 0.65, 3.8, and 10.8 µm. Those estimates are made using the CERES 10'-regional clear-sky albedo and land surface emissivity databases (Chen et al., 2004, 2010), along with the appropriate bidirectional and directional reflectance models, angularly dependent sea surface emissivity models, predicted skin temperature, and corrections for atmospheric absorption and emission (Minnis et al., 2011). The emissivity for water surfaces is estimated using a wind-speed-dependent model developed from theoretical calculations using the approach of Jin et al. (2006). A constant wind speed of 5 knots is assumed for all sea surface pixels.

The observed or modeled radiance at the TOA can be represented as

$$B\left(T_{\mathrm{TOA}}\right) = \prod_{i=n}^{1} t_i \left[B\left(T_o\right)\right] + \left(1 - t_1\right) B\left(T_1\right) + \sum_{i=n}^{2} \left(1 - t_i\right) B\left(T_i\right) \prod_{j=i}^{1} t_j, \quad (2)$$

where T_o is the surface-leaving radiant energy equivalent brightness temperature, which comes from T_s based on the following relationship using the narrowband surface emissivity:

$$T_s = B^{-1}\left\{\frac{\left[B\left(T_o\right) - \left(1 - \varepsilon_s\right) \times L_{\downarrow}\right]}{\varepsilon_s}\right\}. \quad (3)$$

$L_{\downarrow}$ is the downwelling radiant energy at the surface:

$$L_{\downarrow} = \left(1 - t_n\right) B\left(T_n\right) + \sum_{i=1}^{n-1} \left(1 - t_i\right) B\left(T_i\right) \prod_{j=n}^{i+1} t_j. \quad (4)$$

The subscripts i and j denote an atmospheric layer, where 1 and n refer to the layers at the TOA and just above the surface, respectively (e.g., $B(T_1) \equiv B(T_{TOA})$). The atmospheric layer temperature is T_i, and B is evaluated at the central wavelength of the 11 µm band. B^{-1} is the inverse Planck function. The layer transmissivity (t_i) is determined using the correlated k distribution technique, which accounts for gaseous absorption within the spectral band of a given channel. This technique is described in detail by Goody et al. (1989) and Kratz (1995), who depict the discrete version of the spectral-mean transmission $t_{\Delta\omega}(u, p, \Theta)$ as

$$t_{\Delta\omega}\left(u, p, \Theta\right) \cong \sum_{i=1}^{n} w_i \exp\left[-k_i\left(p, \Theta\right) u\right], \quad (5)$$

where $k_i(p, \Theta)$ is an absorption coefficient as a function of pressure p and temperature Θ for a particular wavenumber ω, u is a pathlength, and w_i is a weighting factor for which the summation over n calculations must equal 1. Although the technique does not explicitly account for the details of the spectral response function, the transmissivity from the surface to the TOA is the same with and without the details of the spectral response function for the 11 µm band included in the calculations (Kratz, 1995).

The surface temperatures and atmospheric profiles are linearly interpolated temporally to the satellite image time and spatially to the center of each 0.5° × 0.5° AVHRR or 1.0° × 1.0° GEO grid box. These grid boxes of interpolated NWP sounding data are called regions. For AVHRR retrievals, the regions have resolutions up to 1.5° × 1.5° near the poles but are nominally 0.5° × 0.5° everywhere else. The same T_s retrieval methodology is used for all resolutions. The specific logic of the cloud mask algorithm can be found in Minnis et al. (2008a, 2010, 2016) and Trepte

et al. (2010), which describe cloud tests for different scenarios (e.g., scenes over snow or desert, sun-glint-influenced ocean, scenes with smoke or thin cirrus). It is important to note that although the NWP skin temperature T_s' is used as a seed value in the initial application of the cloud mask, decisions based solely on the difference between 11 µm observations and model values occur for only 2.3 % (5.3 %) of the pixels over land during the day (night). Therefore, the initial influence T_s' is significantly diminished.

After the cloud mask is applied, the mean 0.65 µm reflectance and mean 3.8 and 10.8 µm T_{TOA} (i.e., $< T_{TOA} >$) values are computed from the observed values for the clear and cloudy pixels for each region. The data are then analyzed in pixel groupings called tiles. For AVHRR, the tiles are 8×12 pixels in area, and for GEO the tiles are the same resolution as the gridded region, or $1.0° \times 1.0°$. The different tile sizes are employed to facilitate optimal processing speed. The GEO data are analyzed in NRT, while the AVHRR data have been used for climate studies, which do not have the same time constraints as NRT applications. If at least 20 % of the pixels within the tile are considered clear, the mean observed clear-sky temperature replaces the original NWP-based clear-sky temperature for the region and the cloud mask is repeated using the observed clear-sky mean brightness temperature. The 20 % criterion is used to minimize the influence of cloudy pixels on the final temperature value while still allowing sufficient sample size. If fewer than 20 % of the pixels are clear, then the original clear-sky estimate T_s' and cloud mask are retained and no value T_s is retrieved.

For those tiles satisfying the 20 % criterion, a value of T_s for each pixel is determined using a two-step process. First, the tile mean value T_s (i.e., $< T_s >$) is determined by solving Eq. (2) from the inverse of Eq. (3) (i.e, T_o' solved from T_s'). Then, the mean observed 11 µm clear-sky $< T_{TOA} >$ is used to adjust T_s' based on the difference between the $< T_{TOA} >$ and the modeled T_{TOA}' for each tile. That is, a correction is applied to the model T_s' and temperature/humidity profiles such that T_{TOA}' computed with Eqs. (2) and (5) equals $< T_{TOA} >$, thereby yielding $< T_s >$. For the AVHRR retrievals, the tile average $< T_{TOA} >$ represents an area that is smaller than the area represented by T_{TOA}', which is a regional value originating from the region-scale MERRA T_s'. Thus, all AVHRR tiles with their center within a given MERRA region use the same model profiles and T_s'. For the GEO T_s retrieval, both the observed $< T_{TOA} >$ and the modeled T_{TOA}' correspond to a $1.0° \times 1.0°$ region, because tile area matches region area for GEO imagers.

To save computational time, a value of T_s is estimated for each pixel in the tile as

$$T_s = B^{-1}\left[R_T\, B\left(T_{TOA}\right)\right], \tag{6}$$

where R_T is the ratio

$$R_T = \frac{B\left(\langle T_s \rangle\right)}{B\left(\langle T_{TOA} \rangle\right)}, \tag{7}$$

and T_{TOA} is the observed clear-sky brightness temperature for the pixel. This approach assumes that the atmospheric attenuation and contribution to the exiting radiance is proportionally the same throughout the region. It yields T_s pixel values that differ by -0.04 ± 0.20 K from T_s computed using Eqs. (2) and (5) for each pixel.

4 Sea surface temperature validation

Sea surface temperatures were retrieved as described above for the 2008 AVHRR dataset and are compared with the OI SST values. The AVHRR SST pixel data were first gridded to match the NOAA OI SST 0.25° resolution. Only those pixels classified as clear, with 100 % water fraction (based on a $1.0° \times 1.0°$ land mask) and 0 % sea ice fraction outside of sun-glint conditions, were used to compute the daily grid averages. Additionally, each pixel must be assigned a quality assurance flag of 1, indicating that there are no adjacent cloudy pixels or nearby thin cirrus (within two pixels).

Figure 3 maps the July 2008 SST means from AVHRR (Fig. 3a) and NOAA OI SST (Fig. 3b) as well as their differences (Fig. 3c), which qualitatively reveal very good agreement between the two products. The Fig. 3d scatter density plot reveals a more quantitative analysis of the ~ 3 million daily cell-to-cell comparisons. The bias and standard deviation of the difference (SDD) of the AVHRR SST relative to OI SST for July 2008 are -0.06 and 0.62 K, respectively. A high associated coefficient of determination ($R^2 > 0.99$ not shown) indicates low variance, despite apparent outliers. Disagreements over open ocean, such as those in the tropical western Pacific and northern Pacific Ocean, can be attributed to cloud-clearing differences between the two products or to the fact that the OI satellite SST is supplemented by in situ measurements from buoys and ships that are free of cloud consideration. Nevertheless, despite localized coastal differences and cloud influences, the AVHRR SST is largely consistent with the NOAA OI SST product.

Sea surface temperatures from JAJO 2013 GOES-13 are compared to NOAA OI SSTs under the same gridding, filtering, and quality assurance criteria used for the AVHRR comparisons. Whereas the AVHRR SST retrievals always utilize atmospheric corrections based on MERRA reanalysis, the GEO SST retrievals utilize either GFS or MERRA for the atmospheric corrections. The NRT GEO retrievals currently rely on GFS forecasts, whereas the MERRA reanalysis is suitable for historical GEO and AVHRR retrievals. Therefore, it is important to quantify the influence of the particular NWP reanalysis on satellite-based SST retrieval. Figure 4 compares the July 2013 GOES-13 SSTs retrieved using the GFS-based atmospheric corrections. The GOES SSTs are rather poor in both accuracy and precision relative to the reference – an absolute bias approaching -0.7 K with SDD = 1.02 K. These significant increases in bias and SDD can be attributed to the difference in NWP source, as

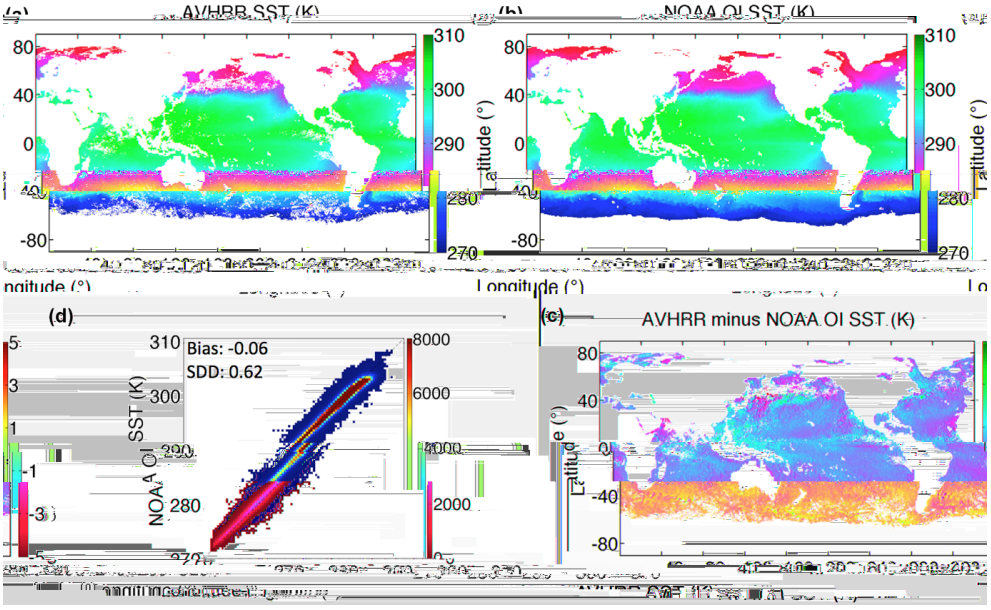

Figure 3. July 2008 **(a)** AVHRR SST, **(b)** NOAA OI SST, **(c)** SST difference, and **(d)** scatter density analysis of ∼ 3 million daily matched grid cells.

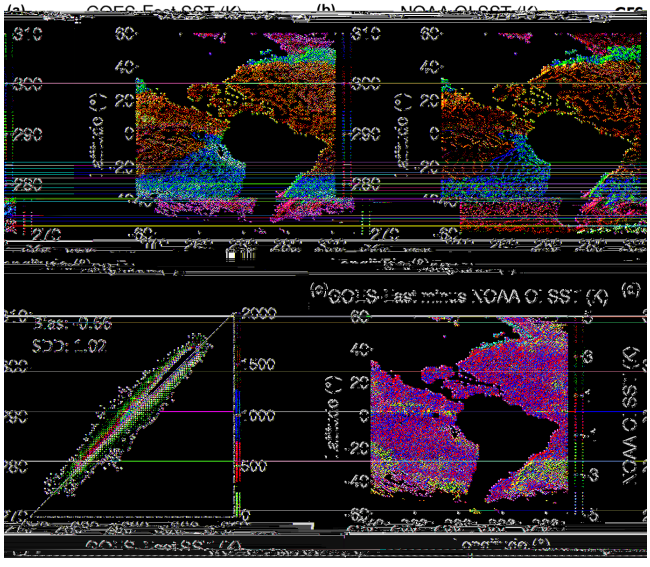

Figure 4. July 2013 **(a)** GOES-13 SST derived, in part, from GFS-based atmospheric corrections, **(b)** NOAA OI SST, **(c)** SST difference, and **(d)** scatter density analysis of ∼ 1 million daily matched grid cells.

is evident from Fig. 5. Figure 5 shows the same comparison as Fig. 4, except that MERRA profiles were used for the atmospheric corrections. Similar then to the AVHRR retrievals, MERRA-derived GOES-13 SSTs exhibit a near-zero bias and an SDD of only 0.60 K relative to the NOAA OI SST reference.

The accuracy and precision of the GFS- and MERRA-derived GOES-13 SST values for the remaining seasonal months of 2013 are illustrated in Fig. 6 along with their AVHRR counterparts for all 12 months of 2008. Mean AVHRR SST is consistently 0.1 K, or less, colder than the NOAA OI SST reference throughout the year. The AVHRR SST monthly SDD is steady near 0.6 K. The MERRA-based JAJO GOES-13 SST SDD is also steady near 0.6 K and the bias is consistently close to zero. The differences between the AVHRR and GOES biases are likely due to uncertainties in the infrared calibrations. The GFS-derived GOES-13 mean SST is consistently ∼ 0.6 K colder than the NOAA reference, with an SDD in excess of 1.0 K for the JAJO seasonal months. This discrepancy with the MERRA-based results suggests that the GFS model profiles are drier than MERRA and/or have insufficient vertical resolution to properly account for the changes in water vapor that are used to compute the atmospheric attenuation of the infrared radiation. It is unlikely that the GFS humidity is too low since it appears to have a wet bias (Yoo, 2012). An explanation for the differences in the model fields is beyond the scope of this paper. However, it is clear that the single-channel retrieval method is sensitive to the source of temperature and humidity profiles. Hereafter, the MERRA data are used for all analyses, unless indicated otherwise.

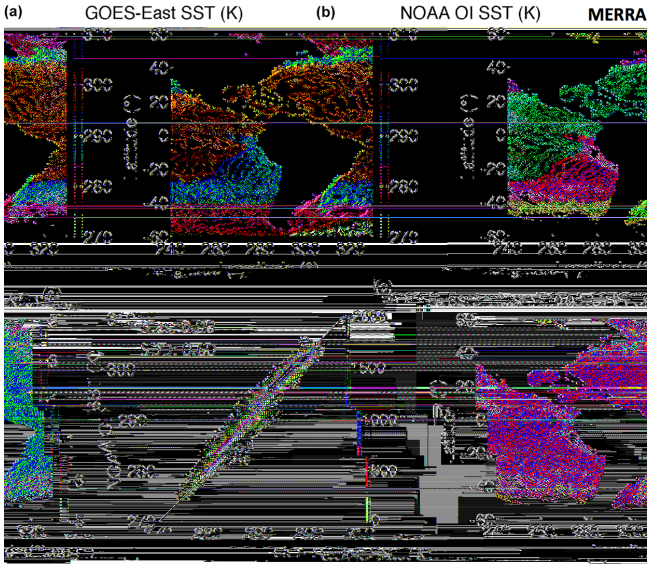

Figure 5. July 2013 **(a)** GOES-13 SST derived, in part, from MERRA-based atmospheric corrections, **(b)** NOAA OI SST, **(c)** SST difference, and **(d)** scatter density analysis of ∼ 1 million daily matched grid cells.

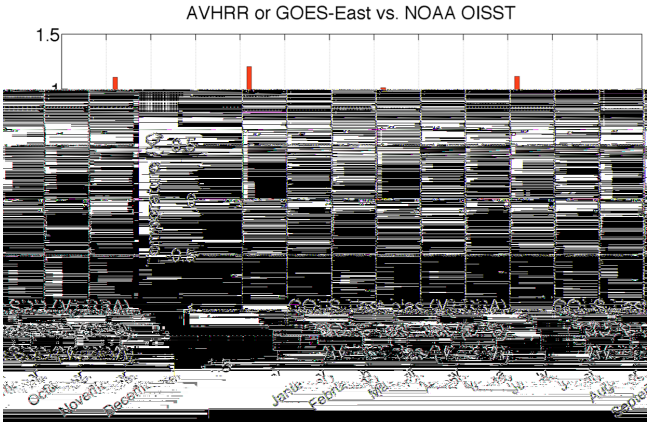

Figure 6. AVHRR (2008) and GOES-13 (2013) SST accuracy and precision relative to NOAA OI SST. For the GEO retrievals, the atmospheric correction is based on either GFS or MERRA reanalysis. Atmospheric corrections for AVHRR retrievals are strictly based on MERRA.

5 Land surface temperature angular anisotropy correction

Satellite-observed LST depends on the viewing and illumination conditions because shading, vegetation conditions, soil type, and topography affect the radiance exiting the scene (Lagouarde et al., 1995 Minnis and Khaiyer, 2000 Minnis et al., 2004). This thermal radiation anisotropy can result in the retrieved LST varying by 6 K or more for some areas (Rasmussen et al., 2010, 2011 Guillevic et al., 2013). From experimental measurements, Sobrino and Cuenca (1999) and Cuenca and Sobrino (2004) found a viewing zenith angle (VZA) dependence of LST that depends on soil type. Pinheiro et al. (2006) developed a physical model to estimate the variation of LST as a function of canopy coverage, solar zenith angle (SZA), VZA, and relative azimuth angle (RAA) for a savanna. Rasmussen et al. (2010, 2011) developed and applied a similar model to predict the LST that would be retrieved by Meteosat over Africa. Vinnikov et al. (2012) constructed a generalized model to convert satellite-measured, VZA-, SZA-, and RAA-dependent LST into a direction-independent equivalent physical temperature, which for general application, requires many sets of matched measurements from different angle sets to construct coefficients for the necessary kernels. Addressing the anisotropic effects and thereby reducing the T_s uncertainties could improve climate monitoring and be of significant benefit to data assimilation and numerical weather prediction needs (Reichle et al., 2010 Guillevic et al., 2013 Draper et al., 2015).

Accounting for 3-D radiance anisotropy for a global retrieval methodology will require the development of regional and seasonal kernels for a universal model (e.g., Vinnikov et al., 2012) or developing canopy configurations globally for physical models (e.g., Rasmussen et al., 2010). Such endeavors require many different matched datasets for a sufficiently large configuration of viewing/illumination angle combinations across many scene types and all seasons. Therefore, at present, we choose to employ the Vinnikov et al. (2012) universal empirical model for angular anisotropy correction. The model, built from varied surface and climate property observations, can serve as a baseline for angular anisotropy correction despite its development from a small number of ground sites across a limited viewing range. Therefore, the goal of this section is to independently test the efficacy of the model through use of large-area satellite and independent ground site LST comparisons. If effective, a universal anisotropic correction model such as this is certainly beneficial to NRT global retrieval of satellite-based LST. As with Vinnikov et al. (2012), our initial assessment will start regionally, i.e., within the GOES-East and GOES-West satellite domains.

5.1 Nadir-normalization model for LST retrieval anisotropy

Vinnikov et al. (2012) formulate the skin temperature at a given set of viewing and illumination angles as

$$T_s(\theta, \theta_o, \varphi) = T_n[1 + a(1-\mu) + b\psi(\theta, \theta_o, \varphi)], \quad (8)$$

where θ, θ_o, and φ correspond to VZA, SZA, and RAA, respectively. The LST at nadir is T_n, and $\mu = \cos(\theta)$. In sequence, the three terms within the brackets are referred to as the isotropic, emissivity, and solar kernels. Vinnikov et

al. (2012) chose the functional form of the solar kernel as

$$\psi(\theta, \theta_o, \varphi) = \sin(\theta)\cos(\theta_o)\sin(\theta_o) \cos(\theta_o - \theta)\cos(\varphi), \quad (9)$$

while coefficients a and b are determined empirically. Physically, the solar kernel attempts to account for the impact of solar intensity and shadowing, as well as the hotspot effect. At night, the solar kernel is defined as 0, i.e., $\psi(\theta, \theta_o \geq 90°, \varphi) \equiv 0$. The emissivity kernel accounts for the VZA dependence of the effective emissivity, which can be due to the VZA variation of the emissivity of a pure surface, the changing combination of scene components (e.g., grass, rocks, tree canopy, mountain slopes, valleys) and their respective temperatures as VZA changes, or a combination of the two.

Vinnikov et al. (2012) first estimated the emissivity kernel by determining the coefficient a in Eq. (8) at night by matching nearly simultaneous GOES-E and GOES-W T_s measurements with five ground site measurements of LST. The differences in the VZAs for the two satellites, covering a range of 43 to 66°, provided the variation in μ needed to perform the regression fit. The solar kernel coefficient for each site was then determined in the same manner using the daytime GOES measurements with the assumption that the emissivity kernel is the same for any hour of the day. This follows if one considers that when the solar kernel is integrated over the entire range of RAA (0–360°), it reduces to zero. Thus, the solar kernel, in effect, represents deviations from the emissivity kernel. Thus, features such as the hotspot, which occurs in a solar backscatter position when $\theta_o = \theta$, are compensated by lower values at a different RAA, typically in shadow, or over a range of RAAs at the same value of θ (e.g., Minnis et al., 2004). It is possible, therefore, that the emissivity kernel, or VZA correction model, could be determined during the daytime by taking measurements over a sufficient range of SZAs and RAAs at a given VZA. Doing so, however, would require a significantly large number of matched datasets in order to achieve sufficient sample size for every interdependent VZA, SZA, and RAA configuration across different surface types and seasons, and thus is much simpler to accomplish at night.

5.2 GOES-East/West LST comparison

To test the efficacy of the Vinnikov et al. (2012) three-kernel anisotropic correction, differences between the hourly GOES-East (GE) and GOES-West (GW) LST retrievals from July 2013 were computed before and after applying the angular adjustments. Prior to differencing, the 15 min discrepancy in the image retrieval at the 3 h synoptic times (00:00, 03:00, …, 21:00 UTC) was mitigated by adjusting the GE LST, which is based on images beginning 15 min before the UTC hour, to that UTC hour when the GW image scan began. This approach accounts for the specific GE and GW scanline time discrepancies. The GE data were linearly interpolated to the GW time using the nearest surrounding synoptic hours. When those surrounding hours crossed the sunrise terminator, no correction was applied because of the day-night discontinuity in LST that occurs shortly after sunrise. Data taken near the terminator (SZA between 80 and 100°) were not used. The image times at the non-synoptic hours are nearly identical, so no temporal normalization was required. To minimize calibration differences, the average nocturnal LST difference, 0.08 K, between GOES-13 and 15 within 0.5° longitude of 105° W, which is the bisector of the two views, was computed and added to all GOES-15 values.

Figure 7 plots the VZAs for GW (Fig. 7a), GE (Fig. 7b), and the GE–GW VZA differences (Fig. 7c). Although the differences are generally less than ±30°, the largest VZAs are up to 70° or more. At night, the emissivity kernel from Eq. (8) would suggest large LST differences for pairs matched at the higher VZAs in this domain. All the retrieved values of normalized LST for both satellites were adjusted to nadir to account for the anisotropic dependence.

The mean regional differences, i.e., $\mathrm{DT}_s = \mathrm{LST(GE)} - \mathrm{LST(GW)}$, are shown in Fig. 8 for the matched July 2013 data. During daytime, DT_s for the unadjusted values (Fig. 8a) is mostly positive east of 105° W and negative to the west. Notable exceptions include the positive values in the west corresponding the highest mountain ranges in Colorado, Utah, Mexico, Washington, Wyoming, Idaho, and New Mexico. After adjusting to nadir (Fig. 8b), the same patterns remain, but they are mitigated considerably with DT_s values closer to zero. Also, the corrected differences for some of the regions at extreme VZAs in the far northeast remain relatively large, perhaps because the viewing dependence increases nonlinearly for large VZA. At night, the unadjusted differences (Fig. 8c) are relatively small, i.e., $|\mathrm{DT}_s| < 2$, in most regions. The positive differences are no longer evident over the high mountains. Applying the anisotropic correction further reduces $|\mathrm{DT}_s|$ to values less than 1.0 K in nearly all cases (Fig. 8d).

Table 1 summarizes the GE–GW results. Over the eastern and western halves of the domain, $|\mathrm{DT}_s|$ drops by 0.99 and 0.54 K, respectively, during the day with the application of the anisotropy adjustment. The mean regional differences are much smaller than before correction, especially for the western region where the difference is near zero. Similarly at night, the corresponding regional differences decrease by comparable amounts and are much closer to zero than without the corrections. Furthermore, the mean absolute biases for both day and night, which are determined by the east–west sample-weighted region differences (not shown), are much closer after correction – reduced by a factor of 2 or more. Overall, the mean bias for the entire domain after correction over all non-terminator hours is 0.59 K.

Although it significantly reduces the GE–GW differences, the anisotropic correction does not eliminate all of the disagreement between the two satellite retrievals. Also, although sign difference between the means over the eastern and western domains essentially disappears for both day and night with the correction, the remaining east–west difference

Table 1. July 2013 day, night, and combined matched GOES-East minus GOES-West mean clear-sky surface skin temperature difference (K) for regions east and west of 105° W, without and with anisotropic correction. The sample-weighted average bias is shown in the bottom row.

Longitude	Combined		Day		Night	
	Without correction	With anisotropic correction	Without correction	With anisotropic correction	Without correction	With anisotropic correction
< 105° W	2.07	1.10	2.38	1.39	1.18	0.35
> 105° W	−0.53	0.01	−0.51	0.04	−0.55	−0.03
All	0.85	0.59	0.98	0.73	0.43	0.18

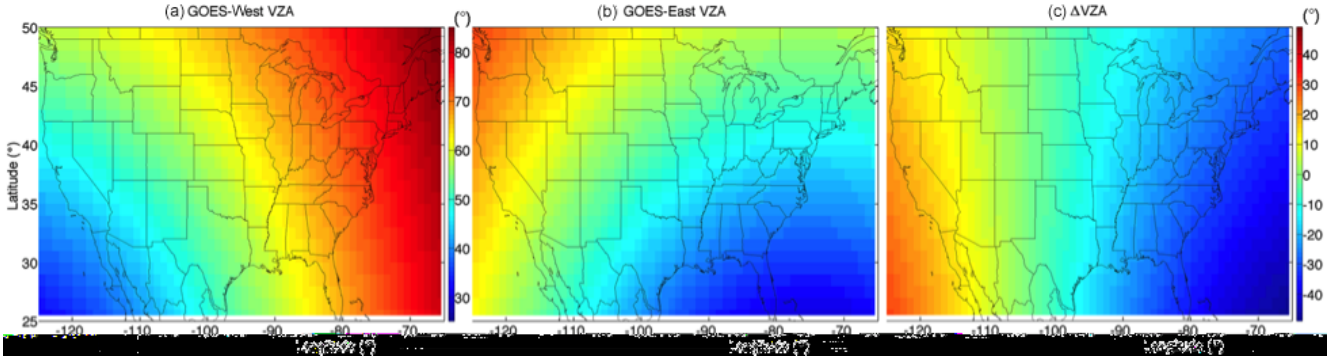

Figure 7. Viewing zenith angles for **(a)** GOES-West and **(b)** GOES-East, and **(c)** their differences over the matching domain.

suggests other factors aside from angular anisotropy affect the observed temperatures. It is possible that the solar azimuthal dependence seen in earlier studies (e.g., Minnis et al., 2004 Vinnikov et al., 2012) is not balanced out for the configurations seen here. The azimuthal dependence includes effects from both the relative solar azimuth angle and the azimuthal orientation of the terrain and vegetation. Moreover, the heating/cooling rates probably differ between the eastern and western domains because of humidity and altitude differences. Downwelling longwave radiation might play a greater role in the diurnal cycle of T_s in the eastern domain, perhaps diminishing the solar-induced anisotropy. It would be instructive to derive a daytime-specific emissivity kernel over the entire range of RAA in order to test these theories, but as alluded to in the previous subsection, such an endeavor is beyond the purview of this paper, which merely is meant to assess the current anisotropy model, as given by Vinnikov et al. (2012), applied to a large satellite dataset.

We can, however, explore how DT_s changes over the course of a day and how much the anisotropic correction diminishes those differences. To that end, the differences were averaged for each UTC for each of the four months and are plotted in Fig. 9. The July results corresponding to Fig. 8 are plotted in Fig. 9e and f as lines connecting the means at each hour. Over the western domain (red line), the uncorrected DT_s (Fig. 9e) gradually approaches zero at 09:00 UTC from near −1 K after 03:00 UTC, when the sun has set over the entire domain. At 12:00 UTC, it rises rapidly to a peak of 2.5 K near 16:00 UTC and drops precipitously after 17:00 UTC to −3 K at 22:00 UTC before increasing until 03:00 UTC. In the east (blue line), DT_s drops slowly toward zero after 01:00 UTC but only reaches 0.4 K at 06:00 UTC before increasing again. It only increases significantly after 12:00 UTC, maximizing at 3.5 K (17:00 UTC) before decreasing to 1.3 K at 21:00 UTC, when it levels off. The behavior is rather different for the corrected values (Fig. 9f), with the two curves being much closer together between 03:00 and 16:00 UTC, while also being much closer to zero overall than without the angular correction. The corrected western domain DT_s rises to near 1.0 K from 09:00 to 11:00 UTC and then drops slightly to about 0.3 K at 12:00 UTC before gradually rising to 1.0 K again by 17:00 UTC. At 17:00 UTC and after, the curves diverge significantly with the eastern data varying more extremely (rapid and continuous increase from 13:00 through 21:00 UTC) than their western counterparts, suggesting different heating/cooling rates. The bias for the entire domain (black line) shows definitively that the afternoon points are mainly responsible for the overall positive bias in Table 1. The results for the other months show that the model generally reduces the mean absolute DT_s at most hours. Some exceptions are seen at night.

Even with different heating/cooling rates, it is expected that DT_s would approach zero after correction for anisotropic

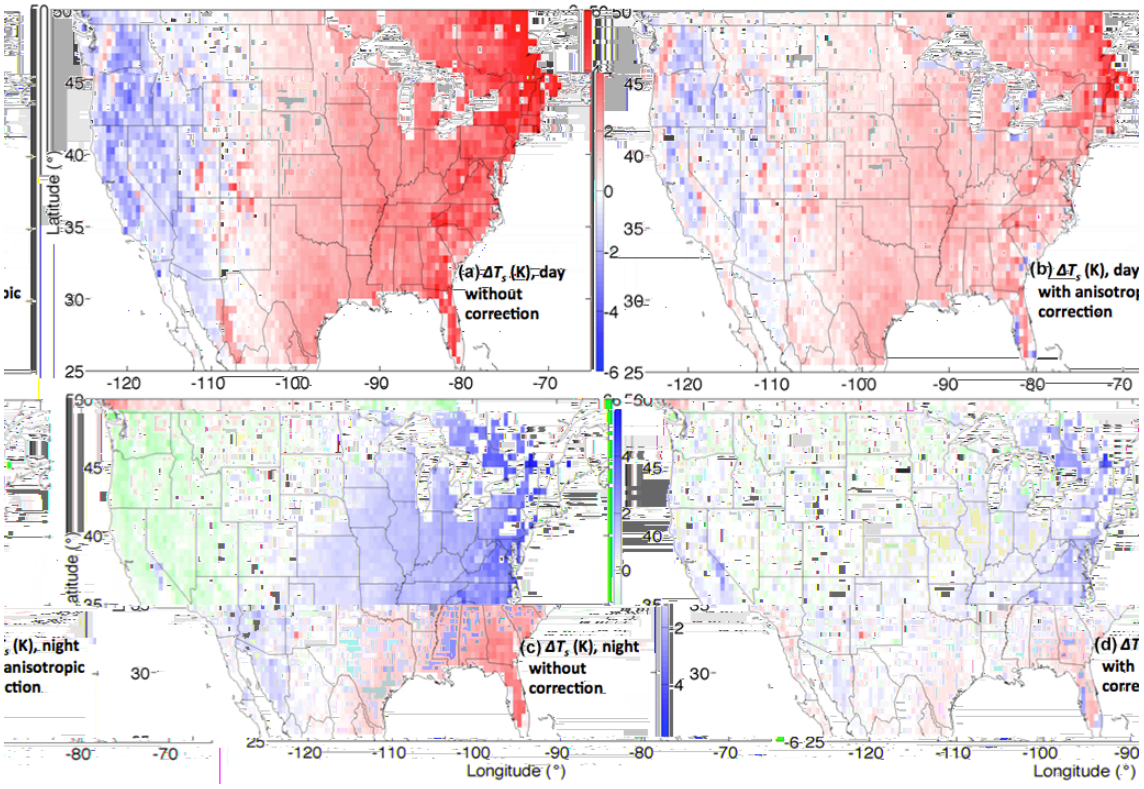

Figure 8. Mean regional GOES-East–GOES-West LST differences for July 2013. **(a)** Day without correction, **(b)** day with anisotropic correction, **(c)** night without correction, and **(d)** night with anisotropic correction.

Table 2. Seasonal and diurnal calibration gains for 2013 based on a satellite ray-matching calibration technique described by Minnis et al. (2002). The gain coefficients derive from the ratio of mean GOES-13 and Aqua MODIS time-, space-, and angle-matched radiances, later converted to brightness temperature (BT). A MODIS-consistent GOES-13 BT is attained by multiplying the appropriate gain with GOES-13 BT values. The gains include adjustment for the spectral band difference between the GOES and MODIS channels following the technique described by Scarino et al. (2016). Those spectral band adjustment factor (SBAF) slope and offset values were applied to MODIS BT values (converted from radiance) during the cross-calibration in order to yield MODIS BTs that were spectrally consistent with GOES-13.

	January		April		July		October	
	Day	Night	Day	Night	Day	Night	Day	Night
Calibration gain	0.9998	1.0018	1.0003	1.0027	1.0001	1.0030	1.0007	1.0015
SBAF slope	1.004	1.004	1.006	1.006	1.005	1.005	1.006	1.005
SBAF offset	−0.694	−0.708	−1.012	−0.994	−0.867	−0.811	−1.029	−0.952

effects as the surface air and skin temperature equilibrate. Instead of going to zero after 03:00 UTC, DT_s drops to roughly −0.8 K for the entire domain by 04:00 UTC until about 06:00 UTC, before rising more rapidly to about 0.8 K at 09:00 UTC and remaining relatively flat until 12:00 UTC. This odd behavior is likely an artifact of the sun–satellite configuration, which causes a change in the infrared channel calibrations at satellite midnight and for 3–4 h afterward. Yu et al. (2013) found that the GOES-11 and GOES-12 10.7 µm (channel 4) brightness temperatures were biased by −0.5 K relative to their daytime calibrations for 3–4 h after satellite midnight, even after an operational correction for the midnight effect had been applied. A smaller bias was evident for a couple of hours prior to midnight. This residual bias could explain the unexpected variation in DT_s seen between 03:00 and 12:00 UTC, if GOES-13 and 15 suffer from a similar UTC bias. Assuming then that the calibration biases are −0.40 and −0.80 K 2 h before and for 4 h after midnight, respectively, for GE, and the same for GW, then DT_s would almost follow the black curve in Fig. 9f exactly (assuming that $DT_s = 0$ in a perfectly calibrated system). By 06:00 UTC, DT_s would still be near −0.4 to −0.8 K because only GE is

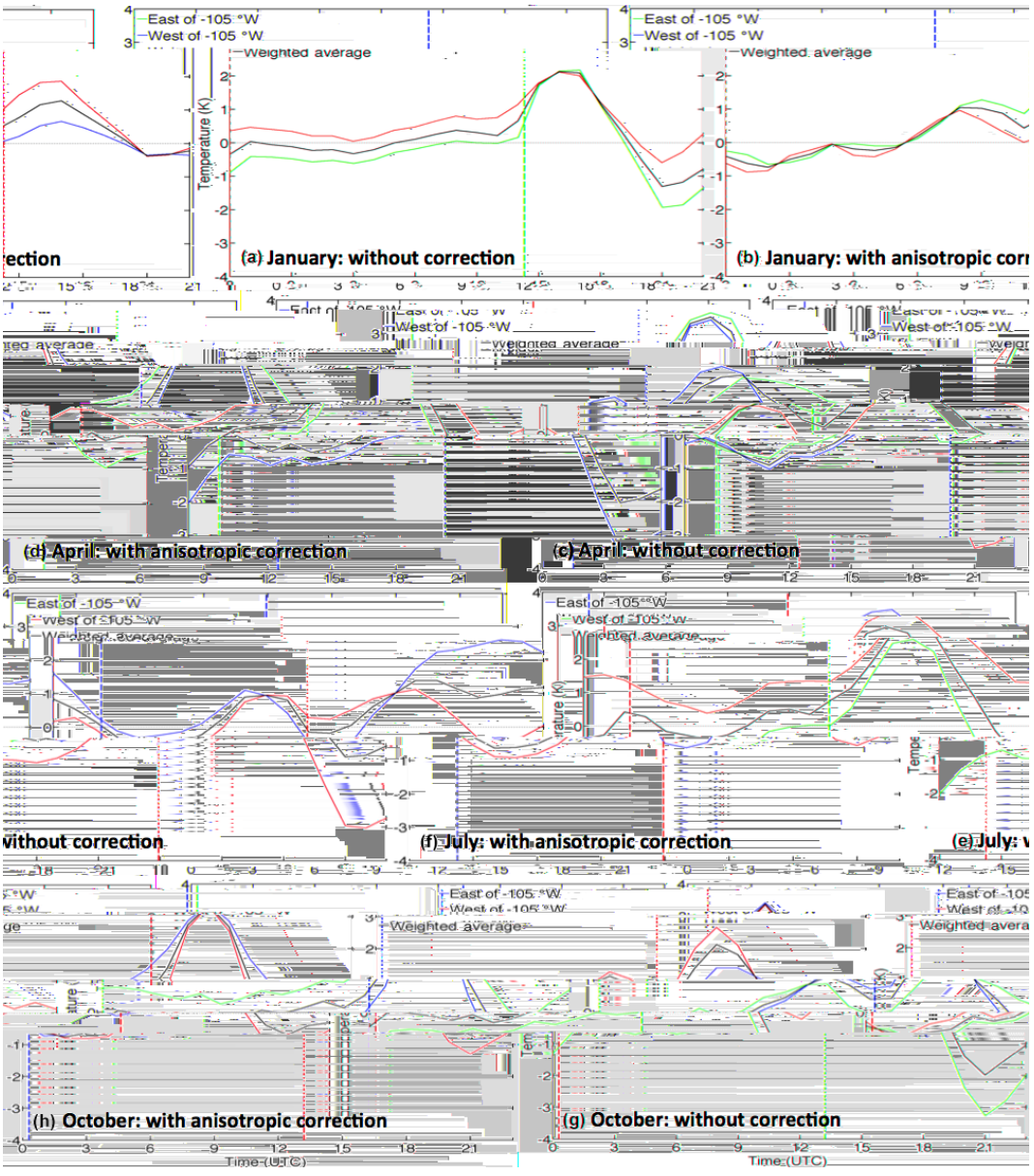

Figure 9. Mean hourly, regional GOES-East–GOES-West LST differences for January, April, July, and October 2013, (left) without and (right) with anisotropic correction. The vertical dashed lines indicate the terminator transitions to night (blue) and day (red) at 37.5° N, 105° W.

influenced by the midnight effect. By 07:00 UTC, the smaller GW pre-midnight bias would partially offset the GE bias causing DT_s to rise until 09:00 UTC, when only GW is affected. After 12:00 UTC, the daylight in the eastern half of the domain would overwhelm any remaining bias. Of course the results discussed here only represent one domain during one month, although DT_s diurnal cycles are shown for other seasonal months in Fig. 9. The midnight calibrations and the viewing/illumination angles vary with time of year. Thus it is clear that a much more comprehensive study would be needed to fully assess the angular anisotropy dependence of the retrieved T_s values in this context. Overall, however, application of the three-kernel model nets meaningful reduction of $|DT_s|$ and can perhaps be improved further by incorporating terrain considerations to account for differential heating/cooling rates.

5.3 Validation with independent MODIS LST, MYD11

The JAJO 2013 GOES-13 LST values are compared with the independent MYD11 product between 60° N and 60° S to determine if the Vinnikov et al. (2012) angular parameterization improves the consistency of the two products. The GOES-

Table 3. Bias and SDD values (K) for the GOES-13 and the MYD11 Aqua MODIS product comparison for day, night, and all times combined separated by 2013 seasonal month, without and with the Vinnikov et al. (2012) three-kernel anisotropic correction applied. The numbers in parentheses indicate the sample size for that month.

		January (920)		April (1992)		July (2401)		October (1615)	
		Without correction	With correction	Without correction	With correction	Without correction	With correction	Without correction	With correction
Bias	Combined	0.88	0.44	0.85	0.24	1.25	0.25	0.84	0.38
	Day	1.34	0.72	1.30	0.49	1.89	0.52	1.69	0.98
	Night	0.25	0.08	0.29	−0.07	0.24	−0.13	−0.03	−0.27
SDD	Combined	1.79	1.49	1.76	1.28	1.42	1.11	1.90	1.46
	Day	2.05	1.71	2.05	1.37	1.07	1.06	2.19	1.45
	Night	1.11	1.06	1.05	1.08	1.01	1.03	1.08	1.13

Figure 10. Probability distributions of LST differences from GOES-13 and the MYD11 Aqua MODIS product for day, night, and all times (combined) **(a)** without and **(b)** with the Vinnikov et al. (2012) three-kernel anisotropic correction applied, for 2013.

13 10.8 µm channel was first cross-calibrated as in Minnis et al. (2002) against its Aqua MODIS counterpart, band 31, for day and night and the JAJO seasonal months in order to minimize any calibration differences. As part of the cross-calibration, spectral differences were accounted for via convolution of Infrared Atmospheric Sounding Interferometer (IASI) hyperspectral brightness temperature measurements over the GOES-13 and Aqua MODIS 11 µm channel spectral response functions, as thoroughly detailed by Scarino et al. (2016). The diurnal/seasonal calibration coefficients and spectral band adjustment factors (SBAF) are provided in Table 2. For each Aqua overpass, the MYD11 pixel LST values are converted to pixel radiance and are averaged on the 1° × 1° GOES-East domain. The mean radiance values are then converted to mean LST and matched to within 15 min of the GOES-13 hourly scans, provided there are at least 150 valid MODIS and GOES-13 pixels per grid cell. To eliminate any differences due to surface emissivity discrepancies, the GOES-13 LST was retrieved using the MYD11 band 31 emissivity values. To effect the comparisons, both the GOES-13 and MYD11 LST values were normalized to the nadir view using Eq. (8).

Figure 10 shows histograms of the differences between the GOES-13 and MYD11 LSTs without (Fig. 10a) and with (Fig. 10b) the Vinnikov et al. (2012) anisotropic correction. Without correction, the GOES LSTs tend to be greater than their MYD11 counterparts, especially during the day. The SDD is greatest during the day at 1.82 K, and the day and night GOES biases are 1.62 and 0.19 K, respectively, resulting in a combined (both day and night) 0.98 K bias. After applying the anisotropic correction, the daytime SDD drops to 1.33 K, with the bias decreasing by 1.0 down to 0.62 K. The nocturnal bias drops to almost −0.12 K, while its SDD increases slightly from 1.07 to 1.09 K. Although the nocturnal bias changes sign, the magnitude is less than that prior to the correction.

The anisotropic correction serves the comparison well, for the most part, across all seasons. Table 3 summarizes the bias and SDD adjustments for day, night, and combined times across the JAJO seasonal months. For all months, the day-

time and combined bias and SDD values reduce substantially following nadir normalization. At night, however, the SDD changes rather subtly and sometimes does not constitute an improvement. In October, for example, the SDD increases from 1.08 to 1.13 K at night, and this is also the only month when the absolute bias increases (−0.03 to −0.27 K). Overall, the angular anisotropy adjustments reduce the bias by ∼ 0.7 K, a value resulting from a 1.0 K reduction during the day and ∼ 0.1 K (0.3 K) absolute (total) reduction at night. The overall SDD dropped by 23 % comprised of a 27 % daytime reduction and a 2 % increase at night. Thus although the emissivity kernel is based on nighttime data, these results indicate that its use to derive the daytime adjustment coefficients is built on a sound assumption of diurnal applicability. The inconsistency at night, especially in October, may be a result of the limited range of viewing angles used to construct the emissivity kernel, although it is perhaps more likely due to nighttime calibration artifacts (e.g., discussion of Fig. 9) that have not yet been fully resolved (Yu et al., 2013).

Similar results (not shown) were found for the GOES-13 LST values retrieved using GFS instead of MERRA. Unlike the SST comparisons (Figs. 4 and 5), the GFS-derived GOES LST bias and SDD values are comparable to those based on the MERRA profiles. Without applying the anisotropic correction, the daytime and nocturnal biases for GOES / GFS retrievals relative to MYD11 are 1.69 ± 1.99 and 0.22 ± 1.12 K, respectively, which are higher, but not significantly worse than the corresponding MERRA values. After applying the VZA adjustment, the day and night biases are 0.76 ± 1.43 and -0.11 ± 1.13 K, respectively. Although the GFS results over land, compared to those over ocean, are much closer to those from MERRA, the MERRA-based results are slightly more accurate, relative to MYD11, than their GFS counterparts.

5.4 Validation with ARM SGP infrared thermometer

To obtain estimates of the LST bias and SDD relative to ARM IRT measurements, only confidently clear pixels that include the ARM SGP site were selected from the 2013 GOES-13, GOES-15, and NOAA-19 AVHRR retrievals. In order to minimize the chance of cloud mask errors and edge effects, all adjacent pixels were required to be clear. Figure 11 shows the scatterplots of LST retrieved from the ARM SGP IRT and from matched GOES and AVHRR data. The IRT is a down-looking instrument that measures the upwelling radiating temperature of the ground surface, so it is considered to have a nadir view for this comparison. The points (Fig. 11a) tend to parallel the line of agreement but are mostly above it. The IRT values are 1.07 K greater than their satellite counterparts, on average, with an SDD of 1.94 K. The points corrected for anisotropy (Fig. 11b) are scattered about the line of agreement and the average difference is 0.11 K with SDD = 1.91 K. Given that agreement improves comparably for day as well as night further supports the assumption that the night-based emissivity kernel is valid during all hours of the day.

This improvement for both halves of the diurnal cycle is easier to see in Fig. 12, which plots histograms of the differences, SatCORPS–IRT, before (Fig. 12a) and after (Fig. 12b) anisotropic correction. The daytime bias approaches zero, moving from -0.94 ± 2.41 to 0.09 ± 2.38 K, while the nocturnal bias changes from -1.18 ± 1.37 to 0.13 ± 1.34 K. When only the GOES data are considered, the corrected data yield 0.16 ± 2.24 and 0.15 ± 1.18 K for day and night, respectively, and when only the AVHRR results are considered, the respective day and night corrected data yield -0.13 ± 2.79 and 0.06 ± 1.75 K. The GOES data were also analyzed using the GFS atmospheric profiles and yielded smaller absolute biases compared to their MERRA counterparts (−0.85 and −1.10 K for day and night, respectively) for no anisotropic correction. With the correction, the day and night measurements exhibit biases of 0.24 and 0.39 K, respectively, and when combined yield an overestimate of 0.32 ± 1.75 K. This bias is larger than the MERRA-based GOES Only retrievals, but the SDD is reduced by 2.2 %. Thus, the overall accuracy is similar for the two vertical profile sources for this location. Because the MERRA and GFS biases differ significantly for SSTs (Fig. 6), use of the MERRA profiles for retrieving T_s with a single IR channel is preferable. See Table 4 for a summary of all results discussed in this section. For any post-adjustment value in Table 4, with the exception of daytime ALL and AVHRR Only SDD, the individual day and night MERRA- or GFS-sourced bias and precision values are within the GOES-R specifications of 2.5 and 2.3 K, respectively (Yu et al., 2012b).

It is worth noting that Heidinger et al. (2013) reported very small changes in LST relative to SURFRAD measurements as a function of VZA and concluded that they are not a major concern. The VZA corrections developed by Vinnikov et al. (2012) and employed here, however, improve the overall absolute bias in all cases, and the SDD in most, relative to ARM. Exceptions for SDD include the combined and daytime AVHRR Only analyses. The anisotropic correction, on average, has a total bias influence close to 1.0 K, which amounts to a significant net reduction of about 0.8 K, yielding an average absolute bias of ∼ 0.1 K relative to the ARM surface site values. The average SDD reduction after correction is only 1 %, resulting in an average precision of about 1.9 K relative to ARM. This is a significantly smaller change compared to the 23 % reduction in SDD seen in the MYD11 validation, which may be due to the smaller sample size for the SGP analysis. Nevertheless, these small improvements, together with the better satellite-to-satellite normalization in Figs. 10 and 8, demonstrate that large-scale application of the three-kernel LST adjustment for anisotropic dependencies will result in a more accurate and uniform product.

Table 4. Mean bias and SDD values (K) based on results relative to the ARM IRT before and after anisotropic correction using only GOES data, only 2013 AVHRR data, and using combined 2013 GOES and AVHRR results (All). GOES SatCORPS retrievals are based on MERRA (top) and GFS (bottom) input.

	GOES		GOES corrected		AVHRR		AVHRR corrected		All		All corrected	
MERRA	Bias	SDD	Bias	SDD	Bias	SDD	Bias	SDD	Bias	SDD	Bias	SDD
Combined	−1.13	1.84	0.16	1.79	−0.80	2.20	−0.02	2.23	−1.07	1.94	0.11	1.91
Day	−0.93	2.30	0.16	2.24	−0.88	2.75	−0.13	2.79	−0.94	2.41	0.09	2.38
Night	−1.32	1.19	0.15	1.18	−0.74	1.75	0.06	1.75	−1.18	1.37	0.13	1.34
GFS	Bias	SDD	Bias	SDD								
Combined	−0.98	1.76	0.32	1.75								
Day	−0.85	2.19	0.24	2.17								
Night	−1.10	1.29	0.39	1.29								

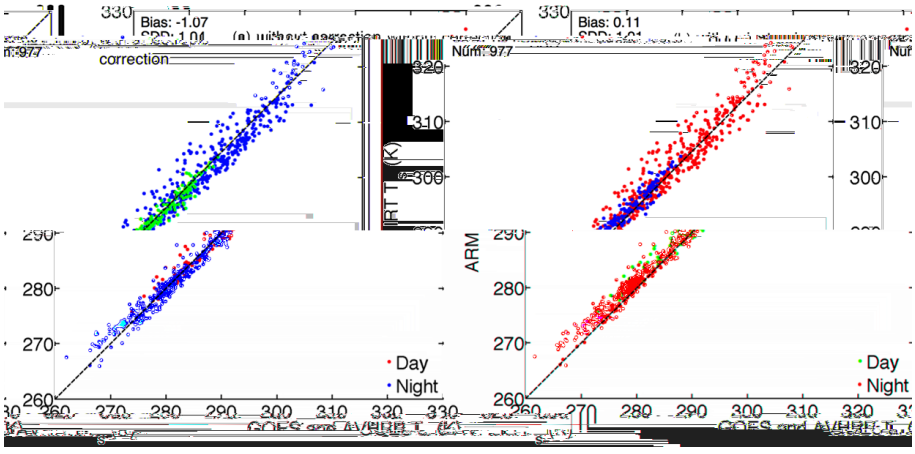

Figure 11. Scatterplots of clear-sky surface skin temperatures from 2013 GOES-13, GOES-15, and AVHRR imagery, matched with ARM SGP IRT temperatures **(a)** without and **(b)** with anisotropic correction.

5.5 Spatial homogeneity of LST

The biases in the results can be due to many factors, including errors in the assumed surface emissivities, the atmospheric profiles, and the surface observations themselves. The representativeness of the site for the much larger area is also potentially a large source of bias. This issue, sometimes called the scaling, or up-scaling, problem (Wang and Liang, 2009 Guillevic et al., 2012, 2014 Li et al., 2013), is a concern for any ground-based satellite LST validation effort, but no attempt is made here to up-scale the ground station point observations to fully characterize the relatively large pixel area of the satellite product. The potential impact of the large scale is important to mention, however. For example, Wang and Liang (2009) conclude that it is not possible to compare satellite-derived LST relative to a single ground LST measurement without introducing bias. Guillevic et al. (2014) discuss that although ground-based LST measurements are in most cases suitable for "well-defined and dedicated sites", investigative procedures into measurement performance should be employed when spanning the full range of surface types and conditions surrounding the site. In both studies, high-resolution ASTER data were used to assess the degree of heterogeneity of LST around field stations and evaluate the spatial representativeness for ground-based measurements. Heidinger et al. (2013), who forewent up-scaling, bring attention to potentially underestimated errors caused by the scaling uncertainty. Furthermore, Fang et al. (2014), in their non-scaled validation study with SURFRAD and ARM, acknowledge the need to better characterize the uncertainties of comparing point measurements with pixel observations. This need was also recognized by Pinker et al. (2009) and led Fang et al. (2014) to suggest using ASTER or MODIS. Therefore, as similarly cautioned by Heidinger et al. (2013), we advise users to be mindful of the scaling-based uncertainties of these non-scaled LSTs. To that end, the remainder of this section aims to quantify the spatial variability of LST surrounding the ARM site.

Inspired by the techniques of Wang and Liang (2009) and Guillevic et al. (2014), we use high-resolution ASTER data

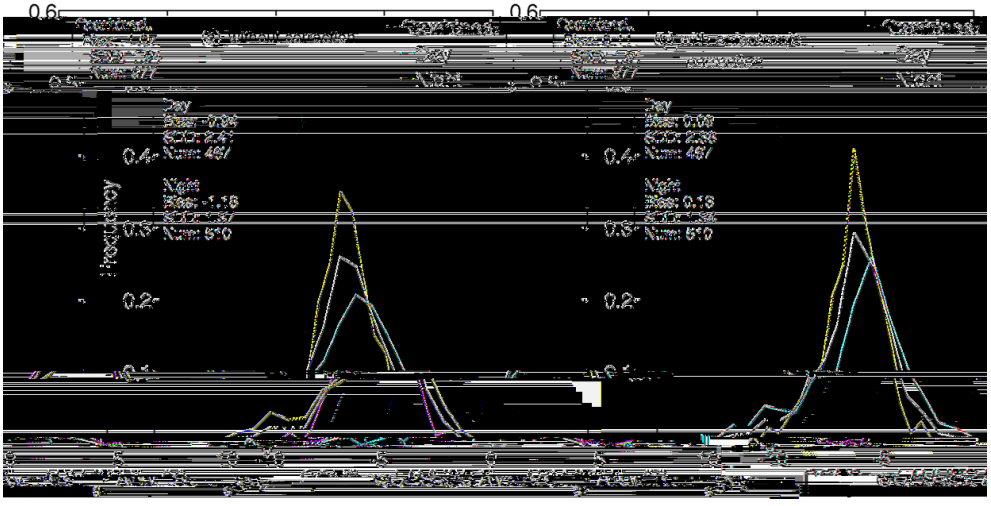

Figure 12. Probability distributions of LST differences between satellite (GOES-13, GOES-15, and AVHRR) and ARM IRT **(a)** without and **(b)** with anisotropic correction.

Table 5. Spatial homogeneity assessment of LST (K) using aggregates of 45×45 (4 km) and 89×89 (8 km) ASTER pixels at 90 m resolution centered on the ARM SGP Central Facility for day, night, and combined times, using data from 2001 through 2015. The bias and SDD indicate the mean difference and standard deviation of the difference, respectively, of the ASTER central pixel relative to the 4.05×4.05 or $8.01 \times 8.01\,km^2$ average. Bias and SDD are based on granules N, which indicates the number of aggregates where at least 95 % of the 90 m pixels signified clear-sky conditions with no known defects and also represents the number of ASTER granules used to determine the region's minimum, median, and maximum values of LST standard deviation (SD).

	Scale (km)	N	Bias	SDD	Min SD	Median SD	Max SD
Combined	4	92	−0.09	1.12	0.20	0.99	5.81
	8	83	−0.17	1.12	0.22	1.05	5.45
Day	4	32	−0.12	1.69	0.20	2.24	5.81
	8	27	−0.24	1.71	0.22	2.27	5.45
Night	4	60	−0.08	0.67	0.28	0.68	4.14
	8	56	−0.14	0.69	0.30	0.79	5.16

to measure the spatial variance of LST in both a 4×4 and $8 \times 8\,km^2$ area centered on the ARM SGP Central Facility. Spatial homogeneity is evaluated in two ways. First, the 4 or 8 km (actually 4.05 and 8.01 km) LST mean (computed from mean radiance and AST_05 surface emissivity) is compared to the central pixel. Second, the heterogeneity is assessed using the minimum, median, and maximum standard deviations (SD) of the 45×45 or 89×89 pixel area, an approach similar to the method Guillevic et al. (2014) employed for evaluating a 1 km region centered on various SURFRAD sites. Note that it is not our aim to assess the accuracy of the ASTER product but instead only the consistency of the large-area and central pixel measurements centered on the ARM site. For such an evaluation, therefore, measurements from ARM are irrelevant. The similarities of the values from the various scales determine the scaling consistency relative to a given ARM SGP measurement.

Results of the ARM spatial homogeneity analysis are presented in Table 5, where N indicates the number of ASTER granules for which at least 95 % of the pixels signified clear-sky conditions with no known defects. Whether day, night, or combined, the magnitude of the mean difference between the 4 km mean LST and central-pixel LST is near 0.1 K (0.12, 0.08 and 0.09 K, respectively), with the central pixel being slightly cooler in all cases. Although it is true that daytime LST exhibits higher uncertainty than that of nighttime (as also observed by Wang and Liang, 2009, and Guillevic et al., 2014), the SDDs of 1.69, 0.67, and 1.12 K are within the respective GOES Only day, night, and combined post-adjustment precision values of 2.24, 1.18, and 1.79 K. The same can be said for the median SDs, although the daytime value matches the 2.24 K precision exactly. It is therefore concluded that our GOES LST retrievals are too warm relative to the ARM measurement by about 0.1 K on average, which is a minor bias adjustment, with an associated uncertainty not in excess of the stated precision values. These values are consistent to the results when considering the 100 % clear granules, in which case $N = 9$, and the absolute mean bias and uncertainty are 0.12 ± 0.75 K, and median SD is 0.57 K in a range from 0.20 to 2.40 K (not shown). Interestingly, although the error is not significantly large, reducing the anisotropy-corrected GOES LST by 0.1 K would bring GOES more in line with the measurements from ARM with both day and night mean biases closer to zero.

The 8 km spatial homogeneity analysis is meant to demonstrate the most extreme cases of pixel-to-ARM matching dis-

parity, e.g., theoretical situations where the ground site is situated near the corners of the containing GOES pixel. The result is a near doubling of the bias to −0.17 K for the combined case, which includes a −0.24 K daytime contribution and a −0.14 K nighttime contribution. The SDDs are comparable to those of the 4 km assessment, which suggests that increasing the aggregation area surrounding the ARM site from 4×4 to $8 \times 8\,km^2$ does not significantly influence LST uncertainty. Although the biases are larger in magnitude than those from the 4 km analysis, if they are representative of the expected scaling error between pixel area and the point measurement, then, as before, they serve well to adjust GOES LST to match the cooler, on average, ARM measurements. Compared to the 4 km analysis, the median SD values increased only slightly to 2.27, 0.79, and 1.05 K for the day, night, and combined cases, respectively. Therefore, as expected, only the daytime SD exceeds the uncertainty determined from the ARM validation results. These results are not meant to suggest that a 4×4 or $8 \times 8\,km^2$ area LST is generally representative of any given point measurement within that region. In fact, such cases are certainly unlikely as concluded by Wang and Liang (2009) and Guillevic et al. (2014), even for well-known validation sources.

6 Summary and conclusions

Accurate assessment of global climate and improvement of climate models, as well as numerical weather forecasts, rely on consistent land and ocean T_s measurements, among others. Atmospheric flux calculations depend on the robustness of such surface variables, and NWP analyses are driven by reliable and frequent state variable updates over large spatial domains. Despite key downsides, satellite data are ideal sources of T_s given their model-ready retrieval schedule and broad continuous areal coverage. Thermal-infrared-derived T_s relies on accurate cloud clearing, atmospheric adjustment, and angular anisotropy consideration. Therefore, validation of satellite T_s relative to known standards is of critical importance.

The SatCORPS provides a T_s product retrieved from GEO and AVHRR sources using the same single-channel algorithm. The benefit of the single-channel approach is that this method is more universally applicable to historic and future satellite instruments compared to the split-window technique. Having GEO and AVHRR T_s values derived from the same algorithm reduces relative uncertainty and, hence, are better able to supplement one another. Validation of SST retrieved from both satellites demonstrates consistent accuracy and precision results of less than 0.1 and 0.6 K relative to NOAA OI SST, respectively, for atmospheric corrections based on MERRA profiles. If GFS temperature and humidity profiles are used to account for atmospheric attenuation, however, the accuracy and precision values for the GEO SST exceed 0.6 and 1.0 K, respectively. The larger negative bias and precision relative to the MERRA-based results suggests that the GFS atmosphere is drier than MERRA over the oceans, on average. This result is surprising in that satellite (Tian et al., 2013) and radiosonde (Kennedy et al., 2011) comparisons indicate that MERRA is too dry at altitudes below 500 hPa.

Daytime LST retrievals can be significantly influenced by satellite and solar viewing geometry. One must therefore account for this 3-D radiance anisotropy dependence on a global scale in order to create an accurate and uniform product. Creating a universal model such as this, however, will require the development of regional and seasonal kernels, which requires many different matched datasets for a sufficiently large configuration of viewing/illumination angle combinations across many scene types and all seasons. Such an endeavor is left for future work. Here, we have employed the Vinnikov et al. (2012) universal empirical model for angular anisotropy correction. It was developed and tested using a very limited set of measurements taken at only five sites over the United States but had not been exercised over a larger scale prior to this study. This article has highlighted independent tests of model effectiveness via large-area satellite LST comparisons and ground site validation, which effectively demonstrate the benefit of applying this anisotropic correction to LST retrievals over much of North American in all seasons.

Land surface temperatures retrieved from July 2013 matched GOES-East and GOES-West data over North America showed distinct VZA-dependent differences. Normalization of the daytime LSTs to the nadir view using the Vinnikov et al. (2012) anisotropic correction model reduced the absolute bias by a factor of 2. The remaining daytime differences are likely due to differential heating/cooling rates and topographical orientations, which can be potentially mitigated in the future by implementing terrain and vegetation considerations into the correction model. The GE–GW average nocturnal absolute LST difference is ~ 0.4 K. Applying the anisotropic correction reduces the mean absolute bias to ~ 0.2 K. Overall, application of the three-kernel model nets meaningful improvement despite a need for better terrain handling and more comprehensive study of near-midnight calibration effects for GOES satellites.

The SatCORPS retrievals from GOES-13 were compared to the Collection-5 Aqua MODIS LST product, a well-validated dataset. Normalization of both the GOES and MODIS LSTs to the nadir view reduced the daytime bias and SDD by 1.0 K and 23 %, respectively. The daytime angle-corrected GOES data were, on average, 0.62 ± 1.33 K greater than their MODIS counterpart. For nighttime, the anisotropic correction had a smaller absolute effect on the bias and resulted in a slightly increased SDD of 1.09 K. When combined, the mean GOES and MODIS difference is 0.29 ± 1.28 K, reduced from 0.98 ± 1.67 K prior to anisotropic correction. Use of the GFS profiles in place of their MERRA counterparts slightly degraded the combined

bias and precision to 0.34 and 1.36 K, respectively. Comparisons with LSTs from the ARM IRT ground station provide further evidence of the validity of the SatCORPS retrieval approach and the application of the anisotropic correction, both for day and night. On average, MERRA-based atmospheric corrections seem to perform slightly better than GFS-based attenuation for LST retrievals compared to surface and other satellite LSTs. This finding, however, should not restrict use of GFS for LST retrievals, as the differences are rather small and not strictly better/worse in all scenarios. For SST validation, the MERRA atmosphere is clearly preferred. The small improvements in bias and SDD relative to both the ARM and MYD11 validation efforts (1) demonstrate that large-scale application of the three-kernel LST adjustment for anisotropic dependencies will yield a more accurate and consistent product and (2) support the assumption of diurnal efficacy of the night-based emissivity kernel.

Further investigation is necessary for the ARM ground-site validation approach, particularly in terms of the up-scaling problem. However, a spatial homogeneity analysis using ASTER data at 4 and 8 km scales demonstrated that the average scaling error is small. Also, SD of LST surrounding the ARM measurement site only exceeded determined precision values for daytime granules, which highlights the importance of robust solar angle considerations in the satellite retrievals. Regardless, disparity between pixel- and ground-station-observed surface conditions and model sounding deficiencies are the likely contributors to the surface–satellite differences. Beyond the outlier cases, however, the corrected SatCORPS GEO and AVHRR T_s exhibit minimal mean bias along with high precision. The anisotropic correction, with an adjustment magnitude of ~ 1.0 K, affords reductions of 0.8 K and 1 % in absolute LST bias and SDD, respectively. These small reductions yield mean bias and precision values of 0.1 and 1.9 K, respectively, compared to the ground site reference.

This study has examined data from only one small part of the Earth using a single anisotropic model developed using a limited range of viewing angles. It appears to work quite well for the larger domain (central North America), which included the sites used in its development, but there remain several areas for future testing and improvement. The impact of such corrections should be tested over other areas of the globe having different vegetation and terrain. The simple linear emissivity kernel under-corrects at higher VZAs, indicating that a higher-order formulation may be needed. Biases in mountainous areas stand out even after correction, suggesting that terrain orientation and morphology may introduce additional complexity in the anisotropy. Regional determination of the Vinnikov et al. (2012) model coefficients may be ideal, but deriving those coefficient values would require many matched datasets to achieve sufficient sampling at a large variety of VZA, SZA, and RAA combinations across all seasons, a task that is left for future work. Because land areas are viewed at fixed VZAs by GEO imagers, the LST retrievals will suffer from VZA biases and, at a given local hour, solar illumination biases. Removal of those biases will improve the quality of LST monitoring and enhance the utility of these datasets for assimilation into numerical weather models. Therefore, incorporating these anisotropic corrections for LST into the near-global NRT retrievals, for overlapping GEO and LEO imagers with robust cloud screening algorithms, will benefit the data assimilation and climate research communities and hopefully lead to improved forecasts and better understanding of the global climate system.

7 Data availability

The GEO satellite skin temperature data can be accessed at http://satcorps.larc.nasa.gov/, whereas the AVHRR data are available at https://gis.ncdc.noaa.gov/all-records/catalog/search/resource/details.page?id=gov.noaa.ncdc:C00876. The angular corrections reported here have not been applied to the current versions of those datasets.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This research was supported by the NASA Modeling, Analysis, and Prediction Program and the NOAA CDR Program. Computing was supported by the NASA High End Computing Program. The authors would like to thank Sarah Bedka and Doug Spangenberg for their generous assistance with SatCORPS processing.

Edited by: M. Portabella
Reviewed by: I. Trigo and two anonymous referees

References

Bodas-Salcedo, A., Ringer, M., and Jones, A.: Evaluation of the surface radiation budget in the atmospheric component of the Hadley Centre Global Environmental Model (HadGEM1), J. Climate, 17, 4723–4748, 2008.

Bosilovich, M., Radakovich, J., Silva, A. D., Todling, R, and Verter, F.: Skin temperature analysis and bias correction in a coupled land-atmosphere data assimilation system, J. Meteorol. Soc. Jpn., 85, 205–228, 2007.

Chen, Y., Sun-Mack, S., Minnis, P., Young, D. F., and Smith Jr., W. L.: Seasonal surface spectral emissivity derived from Terra MODIS data, Proc. 13th AMS Conf. Satellite Oceanogr. and Meteorol., Norfolk, VA, 20–24 September, CD-ROM, P2.4, 2004.

Chen, Y., Minnis, P., Sun-Mack, S., Arduini, R. F., and Trepte, Q. Z.: Clear-sky and surface narrowband albedo datasets derived from MODIS data, Proc. AMS 13th Conf. Atmos. Rad. and Cloud Phys., Portland, OR, June 27–July 2, JP1.2., 2010.

Coll, C. and Caselles, V.: A split-window algorithm for land surface temperature from advanced very high resolution radiometer data: Validation and algorithm comparison, J. Geophys. Res., 102, 16697–16713, 1997.

Coll, C., Wan, Z., and Galve, J. M.: Temperature-based and radiance-based validations of the V5 MODIS land surface temperature product, J. Geophys. Res., 114, D20102, doi:10.1029/2009JD012038, 2009.

Cuenca, J. and Sobrino, J.: Experimental measurements for studying angular and spectral variation of thermal infrared emissivity, Appl. Opt., 43, 4598–4602, doi:10.1364/AO.43.004598, 2004.

DaCamara, C. C.: The Land Surface Analysis SAF: One year of pre-operational activity, in: Proc. 2006 EUMETSAT Meteorol. Satellite Conf., Helsinki, Finland, 2006.

Draper, C., Reichle, R., De Lannoy, G., and Scarino, B.: A dynamic approach to addressing observation-minus-forecast mean differences in a land surface skin temperature data assimilation system, J. Hydrometeorol., 16, 449–464, 2015.

Environmental Modeling Center: The GFS Atmospheric Model, NCEP Office Note 442, Global Climate and Weather Modeling Branch, EMC, Camp Springs, Maryland, 2003.

Duan, S.-B., Li, Z.-L., Tang, B.-H., Wu, H., Tang, R., Bi, Y., and Zhou, G.: Estimation of diurnal cycle of land surface temperature at high temporal and spatial resolution from clear-sky MODIS data, Remote Sens., 6, 3247–3262, 2014.

Fang, L., Yu, Y., Xu, H., and Sun, D.: New retrieval algorithm of deriving land surface temperature from geostationary orbiting satellite observations, IEEE T. Geosci. Remote, 52, 819–828, 2014.

Garand, L.: Toward an integrated land-ocean surface skin temperature analysis from the variational assimilation of infrared radiances, J. Appl. Meteorol., 42, 570–583, 2003.

Ghent, D., Kaduk, J., Remedios, J., Ardö, J., and Balzter, H.: Assimilation of land surface temperature into the land surface model JULES with and ensemble Kalman filter, J. Geophys. Res., 115, D19112, doi:10.1029/2010JD014392, 2010.

Göttsche, F. M., Olesen, F. S., and Bork-Unkelbach, A.: Validation of land surface temperature derived from MSG/SEVIRI with in-situ measurements at Gobabeb, Namibia, Int. J. Remote Sens., 34, 3069–3083, 2013.

Gillespie, A., Rokugawa, S., Matsunaga, T., Cothern, J. S., Hook, S., and Kahle, A. B.: A temperature and emissivity separation algorithm for Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images, IEEE T. Geosci. Remote, 36, 1113–1126, 1998.

Goody, R., West, R., Chen, L., and Crisp, D.: The correlated-k method for radiation calculations in nonhomogeneous atmospheres, J. Quant. Spectrosc. Ra., 42, 539–550, 1989.

Guillevic, P. C., Privette, J. L., Coudert, B., Palecki, M. A., Demarty, J., Ottlé, C., and Augustine, J. A.: Land surface temperature product validation using NOAA's surface climate observation networks – Scaling methodology for the Visible Infrared Imager Radiometer Suite (VIIRS), Remote Sens. Environ., 124, 282–298, 2012.

Guillevic, P. C., Bork-Unkelbach, A., Göttsche, F. M., Hulley, G., Gastellu-Etchegorry, J-.P., Olesen, F. S., and Privette, J. L.: Directional viewing effects on satellite land surface temperature products over sparse vegetation canopies – A multisensor analysis, IEEE T. Geosci. Remote, 10, 1464–1468, 2013.

Guillevic, P. C., Biard, J. C., Hulley, G. C., Privette, J. L., Hook, S. J., Olioso, A., Göttsche, F. M., Radocinski, R., Romàn, M. O., Yu, Y., and Csiszar, I.: Validation of land surface temperature products derived from the Visible Infrared Radiometer Suite (VIIRS) using ground-based and heritage satellite measurements, Remote Sens. Environ., 154, 19–37, 2014.

Heidinger, A. K., Laszlo, I., Molling, C. C., and Tarpley, D.: Using SURFRAD to verify the NOAA single-channel land surface temperature algorithm, J. Atmos. Ocean. Technol., 30, 2868–2884, 2013.

Jiménez-Muñoz, J. C., and Sobrino, J. A.: A single-channel algorithm for land-surface temperature retrieval from ASTER data, IEEE Geosci. Remote S., 7, 176–179, 2010.

Jiménez, C., Prigent, C., Catherinot, J., Rossow, W., and Liang, P. A.: Comparison of ISCCP land surface temperature with other satellite and in situ observations, J. Geophys. Res., 117, D08111, doi:10.1029/2011JD017058, 2012.

Jin, Z., Charlock, T. P., Rutledge, K., Stamnes, K., and Wang, Y.: Analytical solution of radiative transfer in the coupled atmosphere-ocean system with a rough surface, Appl. Opt., 45, 7443–7455, 2006.

Kabsch, E., Olesen, F. S., and Prata, F.: Initial results of the land surface temperature (LST) validation with the Evora, Portugal ground-truth station measurements, Int. J. Remote Sens., 29, 5329–5345, 2008.

Kennedy, A. D., Dong, X., Xi, B., Xie, S., Zhang, Y., and Chen, J.: A comparison of MERRA and NARR reanalyses with the DOE ARM SGP data, J. Climate, 24, 4541–4557, 2011.

Kerr, Y. H., Lagouarde, J. P., Nerry, F., and Ottlé, C.: Land surface temperature retrieval: Techniques and applications: Case of the AVHRR, in: Thermal Remote Sensing in Land Surface Processes, edited by: Quattrochi, D. A. and Luvall, J. C., CRC Press, 2004.

Kratz, D. P.: The correlated k-distribution technique as applied to the AVHRR channels, J. Quant. Spectrosc. Ra., 53, 501–507, 1995.

Lagouarde, J. P., Kerr, Y. H., and Brunt, Y.: An experimental study of angular effect on surface temperature for various plant canopies and bare soils, Agr. Forest Meteorol., 77, 167–190, doi:10.1016/0168-1923(95)02260-5, 1995.

Lazzara, M. A., Benson, J. M., Fox, R. J., Laitsch, D. J., Rueden, J. P., Santek, D. A., Wade, D. M., Whittaker, T. M., and Young: J. T.: The Man computer Interactive Data Access System: 25 years of interactive processing, B. Am. Meteorol. Soc., 80, 271–284, 1999.

Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F., and Sobrino, J. A.: Satellite-derived land surface temperature: Current status and perspectives, Remote Sens. Environ., 131, 14–37, 2013.

Minnis, P. and Khaiyer, M. M.: Anisotropy of land surface skin temperature derived from satellite data, J. Appl. Meteorol., 39, 1117–1129, 2000.

Minnis, P., Nguyen, L., Doelling, D. R., Young, D. F., Miller, W. F., and Kratz, D. P.: Rapid calibration of operational and research meteorological satellite imagers, Part II: Comparison of infrared channels, J. Atmos. Ocean. Technol., 19, 1250–1266, 2002.

Minnis, P., Gambheer, A. V., and Doelling, D. R.: Azimuthal anisotropy of longwave and infrared window radiances from CERES TRMM and Terra data, J. Geophys. Res., 109, D08202, doi:10.1029/2003JD004471, 2004.

Minnis, P., Nguyen, L., Palikonda, R., Heck, P. W., Spangenberg, D. A., Doelling, D. R., Ayers, J. K., Smith, Jr., W. L., Khaiyer, M. M., Trepte, Q. Z., Avey, L. A., Chang, F.-L., Yost, C. R., Chee, T.

L., and Sun-Mack, S.: Near-real time cloud retrievals from operational and research meteorological satellites, Proc. SPIE Europe Remote Sens. 2008, Cardiff, Wales, UK, 15–18 September, 8 pp., 2008a.

Minnis, P., Trepte, Q. Z., Sun-Mack, S., Chen, Y., Doelling, D. R., Young, D. F., Spangenberg, D. A., Miller, W. F., Wielicki, B. A., Brown, R. R., Gibson, S. C., and Geier, E. B.: Cloud detection in nonpolar regions for CERES using TRMM VIRS and Terra and Aqua MODIS data, IEEE T. Geosci. Remote, 46, 3857–3884, 2008b.

Minnis, P., Sun-Mack, S., Trepte, Q. Z., Chang, F.-L., Heck, P. W., Chen, Y., Yi, Y., Arduini, R. F., Ayers, J. K., Bedka, K., Bedka, S., Brown, R., Gibson, S., Heckert, E., Hong, G., Jin, Z., Palikonda, R., Smith, R., Smith, Jr., W. L., Spangenberg, D. A., Yang, P., Yost, C. R., and Xie, Y.: CERES Edition 3 cloud retrievals, AMS 13th Conf. Atmos. Rad., Portland, OR, 27 June–2 July, 5.4, 2010.

Minnis, P., Sun-Mack, S., D. F. Young, D. F., Heck, P. W., Garber, D. P., Chen, Y., Spangenberg, D. A., Arduini, R. F., Trepte, Q. Z., Smith, Jr., W. L., Ayers, J. K., Gibson, S. C., Miller, W. F., Chakrapani, V., Takano, Y., Liou, K.-N., Xie, Y., and Yang, P.: CERES Edition-2 cloud property retrievals using TRMM VIRS and Terra and Aqua MODIS data, Part I: Algorithms, IEEE T. Geosci. Remote, 49, 4374–4400, 2011.

Minnis, P., Bedka, K., Trepte, Q., Yost, C. R., Bedka, S. T., Scarino, B., Khlopenkov, K., and Khaiyer, M. M.: A consistent long-term cloud and clear-sky radiation property dataset from the Advanced Very High Resolution Radiometer (AVHRR), Climate Algorithm Theoretical Basis Document (C-ATBD), CDRP-ATBD-0826 Rev 1 AVHRR Cloud Properties – NASA, NOAA CDR Program, 19 September, 159 pp., 2016

Morris, V. R.: Infrared Thermometer (IRT) Handbook, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, ARM TR-015, available at: https://www.arm.gov/publications/tech_reports/handbooks/irt_handbook.pdf (last access: 6 January 2017), 2006.

Pinheiro, A. C. T., Privette, J. L., and Guillevic, P.: Modeling the observed angular anisotropy of land surface temperature in a Savanna, IEEE T. Geosci. Remote, 44, 1036–1047, 2006.

Pinker, R. T., Sun, D., Hung, M.-P., Li, C., and Basara, J. B.: Evaluation of satellite estimates of land surface temperature from GOES over the United States, J. Appl. Meteorol. Climatol., 48, 167–180, 2009.

Prata, A. J.: Surface temperatures derived from the advanced very high resolution radiometer and the along track scanning radiometer, 1. Theory, J. Geophys. Res., 98, 16689–16702, 1993.

Prata, A. J.: Land surface temperature determination from satellites, Adv. Space Res., 14, 15–26, 1994.

Rasmussen, M. O., Pinheiro, A. C., Proud, S. R., and Sandholt, I.: Modeling angular dependences in land surface temperatures from the SEVIRI instrument onboard the geostationary Meteosat Second Generation satellites, IEEE T. Geosci. Remote, 48, 3123–3133, 2010.

Rasmussen, M. O., Göttsche, F.-M., Olesen, F.-S., and Sandholt, I.: Directional effects on land surface temperature estimation from Meteosat Second Generation for savanna landscapes, IEEE T. Geosci. Remote, 49, 4458–4468, doi:10.1109/TGRS.2011.2144604, 2011.

Reichle, R., Kumar, S. V., Mahanama, S. P. P., Koster, R. D., and Liu, Q.: Assimilation of satellite-derived skin temperature observations into land surface models, J. Hydrometeorol., 11, 1103–1122, 2010.

Reynolds, R. W., Smith, T. M., Liu,. C., Chelton, D. B., Casey, K. S., and Schlax, M. G.: Daily high-resolution-blended analyses for sea surface temperature, J. Climate, 20, 5473–5496, 2007.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G. K., Bloom, S., Chen, J., Collins, D., Conaty, Al., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, J. Climate, 24, 3624–3648, 2011.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The global land data assimilation system, B. Am. Meteorol. Soc., 85, 381–394, 2004.

Scarino, B. R., Doelling, D. R., Minnis, P., Gopalan, A., Chee, T., Bhatt, R., Lukashin, C., and Haney, C. O.: A Web-based Tool for Calculating Spectral Band Difference Adjustment Factors Derived from SCIAMACHY Hyper-spectral Data, IEEE T. Geosci. Remote, 54, 2529–2542, 2016.

Scarino, B., Minnis, P., Palikonda, R., Reichle, R. H., Morstad, D., Yost, C., Shan, B., and Liu, Q.: Retrieving clear- sky surface skin temperature for numerical weather prediction applications from geostationary satellite data, Remote Sens., 5, 342–366, 2013.

Snyder, W. and Wan, Z.: BRDF models to predict spectral reflectance and emissivity in the thermal infrared, IEEE T. Geosci. Remote, 36, 214–225, 1998.

Sobrino, J. A. and Cuenca, J.: Angular variation of thermal infrared emissivity for some natural surfaces from experimental measurements, Appl. Opt., 38, 3931–3936, doi:10.1364/AO.38.003931, 1999.

Sobrino, J. A. and Raissouni, N.: Toward remote sensing methods for land cover dynamic monitoring: Application to Morocco, Int. J. Remote Sens., 21, 353–366, 2000.

Sobrino, J. A. and Romaguera, M.: Land surface temperature retrieval from MSG1-SEVIRI data, Remote Sens. Environ., 92, 247–254, 2004.

Sobrino, J. A. Sòria, G., and Prata, A. J.: Surface temperature retrieval from Along Track Scanning Radiometer 2 data: Algorithms and validation, J. Geophys. Res., 109, D11101, doi:10.1029/2003JD004212, 2004.

Sun, D. and Pinker, R. T.: Estimation of land surface temperature from a Geostationary Operational Environmental Satellite (GOES-8), J. Geophys. Res., 108, 4326, doi:10.1029/2002JD002422, 2003.

Sun, D., Fang, L., and Yu, Y.: GOES LST Algorithm Theoretical Basis Document, Version 3.0, NOAA NESDIS Center for Satellite Applications and Research, Camp Springs, MD, USA, 81 pp., 2012.

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

Trepte, Q. Z., Minnis, P., Trepte, C. R., Sun-Mack, S., and Brown, R.: Improved cloud detection in CERES Edition 3 algorithm and comparison with the CALIPSO Vertical Feature Mask, Proc.

AMS 13th Conf. Atmos. Rad. and Cloud Phys., Portland, OR, 27 June–2 July, JP1.32, 2010.

Trigo, I., Monteiro, I. T., Olesen, F., and Kabsch, E.: An assessment of remotely sensed land surface temperature, J. Geophys. Res., 113, D17108, doi:10.1029/2008JD010035, 2008.

Tsuang, B., Chou, M., Zhang, Y., Roesch, A., and Yang, K.: Evaluations of land ocean skin temperatures of the ISCCP satellite retrievals and the NCEP and ERA reanalyses, J. Climate, 21, 308–330, 2008.

Vinnikov, K. Y., Yu, Y., Goldberg, M. D., Tarpley, D., Ramanov, P., Istvan, L., and Chen, M.: Angular anisotropy of satellite observations of land surface temperature, Geophys. Res. Lett., 39, L23802, doi:10.1029/2012GL054059, 2012.

Wan, Z.: New refinements and validation of the MODIS land-surface temperature/emissivity products, Remote Sens. Environ., 112, 59–74, 2008.

Wan, Z. and Dozier, J.: A generalized split-window algorithm for retrieving land-surface temperature measurement from space, IEEE T. Geosci. Remote, 34, 892–905, 1996.

Wan, Z. and Li, Z.-L.: A physics-based algorithm for retrieving land-surface emissivity and temperature from EOS/MODIS data, IEEE T. Geosci. Remote, 35, 980–996, 1997.

Wan, Z., Zhang, Y., Zhang, Q., and Li, Z.: Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data, Remote Sens. Environ., 83, 163–180, 2002.

Wan, Z., Zhang, Y., Zhang, Q., and Li, Z.-L.: Quality assessment and validation of the MODIS global land surface temperature, Int. J. Remote Sens., 25, 59–74, 2004.

Wan, Z., Zhang, Y., Zhang, Q., and Li, Z.: Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data, Remote Sens. Environ., 83, 163–180, 2008.

Wang, A., Barlage, M., Zeng, X., and Draper, C. S.: Comparison of land skin temperature from a land model, remote sensing, and in situ measurement, J. Geophys. Res.-Atmos., 19, 3093–3106, 2014.

Wang, K. and Liang, S.: Evaluation of ASTER and MODIS land surface temperature and emissivity products using long-term surface longway radiation observations at SURFRAD sites, Remote Sens. Environ., 113, 1556–1565, 2009.

Williamson, S. N., Hik, D. S., Gamon, J. A., Kavanaugh, J. L., and Koh, S.: Evaluating cloud contamination in clear-sky MODIS *Terra* daytime land surface temperatures using ground-based meteorology station observations, J. Climate, 26, 1551–1560, 2013.

Yamaguchi, Y., Kahle, A. B., Tsu, H., Kawakami, T., and Pniel, M.: Overview of Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), IEEE T. Geosci. Remote, 36, 1062–1071, 1998.

Yoo, H.: Evaluation of NCEP GFS cloud properties using satellite retrievals and ground-based measurements. Ph.D. Dissertation, Univ. Maryland, College Park, MD, 159 pp., doi:10.13016/M2W90T, 2012.

Yu, Y., Tarpley, D., Privette, J. L., Goldberg, M. D., Rama Varma Raja, M. K., Vinnikov, K. L., and Xu, H.: Developing algorithm for operational GOES-R land surface, temperature product, IEEE T. Geosci. Remote, 47, 936–951, 2009.

Yu, Y., Tarpley, D., Privette, J. L., Flynn, L. E., Xu, H., Chen, M., Vinnikov, K. Y., Sun, D., and Tian, Y.: Validation of GOES-R satellite land surface temperature algorithm using SURFRAD ground measurements and statistical estimates of error properties, IEEE T. Geosci. Remote, 50, 704–713, doi:10.1109/TGRS.2011.2162338, 2012a.

Yu, Y., Tarpley, D., Xu, H., and Chen, M.: GOES-R Advanced Baseline Imager (ABI) Algorithm Theoretical Basis Document for Land Surface Temperature, Version 2.5, NOAA NESDIS Center for Satellite Applications and Research, Camp Springs, MD, USA, 93 pp., 2012b.

Yu, F., Wu, X., Rama Varma Raja, M. K., Li, Y., Wang, L., and Goldberg, M.: Diurnal and scan angle variations in the calibration of GOES imager infrared channels, IEEE T. Geosci. Remote, 51, 671–683, doi:10.1109/TGRS.2012.2197627, 2013.