An objective methodology for infrared land surface emissivity evaluation

Zhenglong Li,[1] Jun Li,[1] Xin Jin,[1] Timothy J. Schmit,[2] Eva E. Borbas,[1] and Mitchell D. Goldberg[2]

Received 28 March 2010; revised 20 July 2010; accepted 26 July 2010; published 24 November 2010.

[1] Land surface emissivity (LSE) in the infrared (IR) window region (8–12 μm) governs the thermal emissions from the Earth's surface. Many LSE databases, retrieved from various satellite instruments, are available for studying climate, Earth-atmosphere interaction, weather, and the environment. The precision (standard deviation) and accuracy (bias) of these databases remain unclear. In this study, we introduce an objective and efficient method for quantitatively evaluating the LSE precision using satellite radiance observations. The LSE brightness temperature (Tb) deviations, defined as the standard deviations of Tb differences between satellite observations and radiative transfer calculations, can be estimated by minimizing the impacts from land surface temperature (LST) and atmospheric profiles. This is followed by the estimation of LSE precision. This method does not need the true LSE measurements. It only needs ancillary information such as atmospheric profiles and LST, both of which do not require high accuracy and thus can be obtained from a numerical weather prediction forecast or analysis. The method is applied to six different monthly LSE databases from August 2006 and 2007, and the results are presented. The error sources affecting the method are identified and the sensitivity to these errors is studied.

Citation: Li, Z., J. Li, X. Jin, T. J. Schmit, E. E. Borbas, and M. D. Goldberg (2010), An objective methodology for infrared land surface emissivity evaluation, *J. Geophys. Res.*, *115*, D22308, doi:10.1029/2010JD014249.

1. Introduction

[2] Land surface emissivity (LSE) is the relative power of the land surface to emit energy by radiation. It is defined as the ratio of the energy radiated by the land surface to energy radiated by a blackbody at the same temperature. LSE has a spectral variation from as low as 0.6 to 1 depending on the surface materials, soil moisture, vegetation cover, and surface roughness.

[3] Accurate LSE in the infrared (IR) region, especially in the thermal infrared region (TIR; 8–12 μm), plays an important role in studying climate, Earth-atmosphere interaction, radiation budget, weather, and the environment. An accurate LSE is necessary for properly assimilating IR radiances over land. As early as 1977, *Kornfield and Susskind* [1977] pointed out that accurate LSE is necessary in order to accurately estimate land surface temperature (LST) from satellite measurements. *Li et al.* [2007] showed that changing LSE by 0.01 will result in approximately a 0.5 K brightness temperature (Tb) change in the IR window region. *Hulley et al.* [2009b] found that an emissivity error of 0.015 results in an LST error of about 1 K at 8.6 μm at a temperature of 300 K. The LSE and LST together govern the territorial emission through the Planck function, one important component of the climate system used to study surface radiation budget and energy flux. According to *Jin and Liang* [2006], the assumption of a flat constant emissivity spectrum, which is utilized by the National Oceanic and Atmospheric Administration National Centers for Environmental Prediction model and the National Center for Atmospheric Research (NCAR) Community Land Model version 2 (CLM2), induces errors in modeling the surface energy budget, especially over large arid and semiarid areas where LSE is significantly smaller than the assumed value. Using the coupled NCAR Community Atmosphere Model version 2 (CAM2)–CLM2, their study demonstrated that warm biases as large as 1.5 K for surface air temperature are found by using a constant value of 0.96 (instead of the observed value of 0.90 by Moderate-Resolution Imaging Spectroradiometer (MODIS)) for broadband LSE, and the LST is found to have warm biases of 1–1.5 K over certain desert areas. *Seemann et al.* [2008] demonstrated that accurate LSE could improve atmospheric profile sounding retrievals. By using the University of Wisconsin (UW) Baseline Fit (BF) emissivity database instead of a flat value of 0.98, the root-mean squared error (RMSE) of the retrieved total precipitable water is reduced from 3.8 to 2.5 mm compared to ground-based microwave measurements.

[1]Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin-Madison, Madison, Wisconsin, USA.
[2]Center for Satellite Applications and Research, National Environmental Satellite, Data, and Information Service, National Oceanic and Atmospheric Administration, Camp Springs, Maryland, USA.

Copyright 2010 by the American Geophysical Union.
0148-0227/10/2010JD014249

[4] LSE can be measured through laboratory analysis [*Nerry et al.*, 1990] and estimated through satellite remote sensing. LSE naturally varies both temporally and spatially. It would be expensive and impractical for laboratory analysis to capture these variations. The most practical method to measure LSE is through satellite remote sensing. Sensors measure the radiance emitted and reflected from the Earth surface and the atmosphere. LSE can be derived from these radiance measurements. Many emissivity databases are available now from different satellite instruments. The operational MODIS LSE (MOD11) products are retrieved using a physical algorithm, which takes two observations (night and daytime) and assumes LSE is invariable while LST is variable [*Wan and Li*, 1997]. A similar broadband emissivity research product from the High-Resolution Infrared Radiation Sounder is also derived [*Ruston et al.*, 2008], but with a coarser spatial resolution. *Seemann et al.* [2008] developed a global LSE database by combining the operational MODIS LSE products and hyperspectral laboratory measurements. The Atmospheric Infrared Sounder (AIRS) operational LSE products are retrieved using a multivariable linear regression method followed by a simultaneous physical algorithm [*Susskind and Blaisdell*, 2008]. *Li et al.* [2007] also showed that AIRS full spectrum LSE could be obtained through a simultaneous physical retrieval algorithm. The Infrared Atmospheric Sounding Interferometer (IASI) preliminary research test product is derived using a multistage regression technique for faster computation [*Zhou et al.*, 2010]; the preliminary IASI emissivity data set is not only a work in progress but also promising, and a final version will be produced in the near future. Therefore, the results derived from this preliminary IASI emissivity product as shown in this paper are used as a reference for testing the emissivity evaluation algorithm described.

[5] A major limitation of satellite-derived LSE is the lack of evaluation or validation. Unfortunately, direct validation of LSE is far more difficult than many other meteorological parameters. As mentioned before, an accurate LSE can be obtained using laboratory measurements. However, it is difficult to use laboratory measurements for validation. For example, the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) [*Hulley et al.*, 2009a; *Sabol et al.*, 2009; *Sobrino et al.*, 2007] and MODIS [*Wang et al.*, 2007; *Wan*, 2008] use laboratory measurements to validate their LSE retrievals. However, their comparisons are only over homogeneous areas where laboratory measurements can represent the pixels (spatial resolution of 90 m for ASTER and 1 km for MODIS). This method is hard to apply to other LSE retrievals, which typically have pixel sizes larger than a few kilometers. The increased pixel size limits the number of homogeneous sites that can be used for validation. Furthermore, the comparisons are not conclusive due to a limited sample size. Another difficulty is the temporal collocation of the laboratory measurements and the satellite retrievals. In those validation studies, the one-time laboratory measurement usually is used as true to validate the retrievals from different times at the same location. As mentioned before, LSE changes with surface conditions such as vegetation cover and soil moisture. Validation using laboratory measurements, for lack of temporal information, will be problematic if the actual LSE has a strong temporal variation.

[6] Qualitative evaluation of LSE over a large area can be indirectly achieved. For example, positive impacts are shown if the retrieved LSE improves the moisture retrievals, especially in the lower atmosphere [*Seemann et al.*, 2008; *Li et al.*, 2008]. However, such analysis does not evaluate LSE channel by channel, and the evaluation is only qualitative, not quantitative.

[7] In this paper, an objective and efficient evaluation method to quantify the LSE precision (the standard deviation of errors) is introduced. This method does not need the true LSE measurements. It uses the satellite radiance observations as references. The comparisons can be done for any clear retrievals. It only needs ancillary information such as atmospheric profiles and LST, both of which are not required to be of high accuracy and thus can be obtained from a numerical weather prediction (NWP) forecast or analysis. This method is not considered to be a validation method since no true LSE measurements are used for comparison. However, this method will provide quantitative information about the LSE precision. Section 2 provides details about all the databases and measurements used in this study, followed by the demonstration of the new evaluation method using simulation data in section 3. Application to Spinning Enhanced Visible and Infrared Imager (SEVIRI) radiance measurements is shown in section 4. Discussions are provided in section 5; results and summaries are in section 6.

[8] It is emphasized that the method presented is not able to quantify the LSE accuracy (the mean bias of errors) since there are no true LSE measurements. No attempts are made to discuss the LSE accuracy.

2. Data

2.1. SEVIRI Introduction

[9] In this study, IR radiance measurements from the SEVIRI onboard the Europe's Meteosat Second Generation (MSG) are used to evaluate LSE precision from different databases. SEVIRI is an IR imager onboard the geostationary satellite, Meteosat-8/9 of MSG [*Schmetz et al.*, 2002]. It only takes 15 min to finish a full disk scan with a resampling spatial spacing of 3 km at nadir. Among the 11 channels, the three window channels (8.7, 10.8, and 12 μm) in the TIR are used for LSE evaluation. It is noted that the method presented here is not limited to using SEVIRI; it could be performed using IR radiance measurements from any instruments with the appropriate IR channels.

2.2. Matchup Database for Simulation Study

[10] A simulation study is conducted to demonstrate the LSE precision evaluation in section 3. For the simulation, LST and the atmospheric profiles are needed for the radiative transfer calculations in addition to LSE. A matchup data set [*Li et al.*, 2009], including temperature and moisture profiles from radiosonde observations (RAOB) from the U.S. Department of Energy Atmospheric Radiation Measurement (ARM) program at the southern Great Plains site [*Miloshevich et al.*, 2006] at Lamont, OK (C1, 36°37′N, 97°30′W), the Global Forecast System (GFS) 6 h forecast, the laboratory-measured LSE spectrum from the MODIS emissivity library (http://www.icess.ucsb.edu/modis/EMIS/html/em.html) and the ASTER spectral library [*Salisbury et al.*, 1994], and the LST measured by the infrared radiometer at the ARM site [*Morris et al.*, 2006], is used to simulate the radiances at SEVIRI's three IR window bands

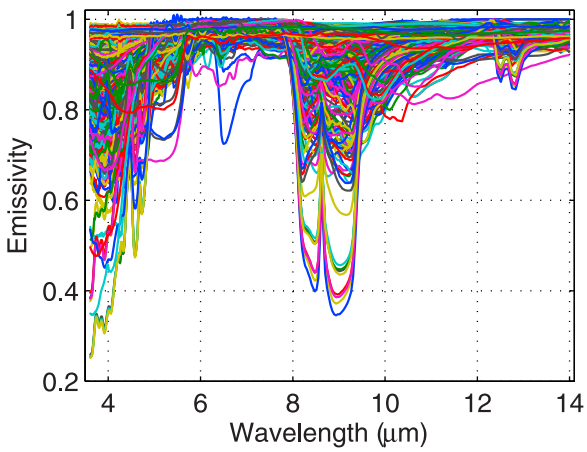

Figure 1. The land surface emissivity hyperspectra from laboratory measurements.

(8.7, 11, and 12 μm). The time coverage is from August 2006 to August 2009. The sample size for clear skies with RAOB data at the ARM site is 1718.

[11] The ARM RAOBs are preferred rather than the conventional RAOB in simulating the radiances because they are more frequent (four times a day) and have better overall quality than the conventional RAOBs [*Turner et al.*, 2003; *Li et al.*, 2009]. The sampling rate is 2 s through the flight. Figure 1 shows the LSE spectra in the infrared region from the laboratory-measured emissivity database. They are convolved to the SEVIRI spectrum using the SEVIRI's spectral response function (SRF). There are 332 samples of the emissivity spectrum. They cover most of the Earth surface conditions. Each of them is used at least five times in the data set, and some are used six times.

[12] From the matchup database, two sets of radiances are calculated: the simulated SEVIRI radiances and the calculated radiances based on the estimation of the surface and atmospheric parameters. Both sets of radiances are calculated using the Pressure-Layer Fast Algorithm for Atmospheric Transmittance (PFAAST) models [*Hannon et al.*, 1996]. PFAAST is based on the line-by-line radiative transfer model version 8.4 [*Clough and Iacono*, 1995] and the high-resolution transmission molecular absorption database-2000 [*Rothman et al.*, 1992] with updates (aer_hitran_2000_ updat_01.1). Table 1 shows the input parameters for the radiative transfer calculations.

[13] The simulated SEVIRI radiances are calculated with RAOB profiles, the radiometer-measured LST and the laboratory-measured LSE as inputs. This set of radiances is used as reference in the simulation study. The other set of radiances is calculated with the estimated atmospheric profiles and surface parameters, and it will be compared with the first set of radiances to derive LSE precisions. In the simulation study in section 3, the collocated GFS forecast is used as the estimated atmospheric parameters. Since there are no estimates of LST and LSE in the matchup database, they are generated using $x_r = x_t + E(\delta x_t)$, where x_r is the estimated parameter, x_t is the true parameter used to simulate SEVIRI radiances, and $E(\delta x_t)$ is a random number with a bias of 0 and an standard deviation of δx_t (3 K; 0.05, 0.025, and 0.015 for LST; and emissivities of 8.7, 10.8, and 12 μm, respectively).

2.3. LSE Databases

[14] In section 4, the SEVIRI radiance measurements from August 2006 are used to evaluate six monthly LSE databases over the desert areas (10°N–35°N and 15°W–55°E) in Africa and part of Asia. This area is chosen because the LSE retrieval in desert areas has been a challenge. Also, the viewing angle dependency over deserts is very small [*Labed and Stoll*, 1991], and thus can be neglected in the discussion. The six LSE databases are categorized into three groups: MODIS products, AIRS products, and IASI preliminary research test product [*Zhou et al.*, 2010].

[15] 1. The MODIS products include two operational MODIS products (MOD11 and MYD11), version 4 from Terra (MODIS-T), and Aqua (MODIS-A), respectively, and the UW High Spectral Resolution IR emissivity (UWiremis) database [*Borbas et al.*, 2007]. The operational MODIS LSE product is derived using a physics-based algorithm, which uses a pair of day/night observations in seven thermal IR bands to solve the nonlinear inverse problem [*Wan and Li*, 1997; *Wan*, 2008]. The UWiremis algorithm [*Borbas et al.*, 2007] was used to create high spectral resolution emissivity spectra (wave number resolution between 2 and 4 cm^{-1}, at 416 wave numbers) from a combination of laboratory measurements of selected materials, and the UW BF global IR land surface emissivity database [*Seemann et al.*, 2008]. The algorithm is based on a statistical regression: The first 6 eigenvectors of 123 selected high spectral resolution laboratory spectra were regressed against the 10 hinge points of the monthly UW BF emissivity data.

[16] 2. The operational AIRS LSE products are derived using a regression plus a simultaneous physical solution approach [*Susskind and Blaisdell*, 2008]. The retrieval is performed using 15 longwave channels and 10 shortwave channels [*Hulley et al.*, 2009b; *Susskind and Blaisdell*, 2008]. The monthly database is only available for four wavelengths (Table 2) with a spatial resolution of 1°. The ascending orbits (AIRS-A) are separated from the descending orbits (AIRS-D) because of their different local passing times, with ascending orbits mainly in daytime and descending orbits in nighttime over Africa.

[17] 3. The IASI LSE is a preliminary research test product [*Zhou et al.*, 2010] generated using a regression approach for a faster process. In contrast to the operational

Table 1. Inputs for Radiative Transfer Calculations for the Two Sets of SEVIRI Radiances[a]

Input Parameters	LST	LSE of 8.7 μm	LSE of 10.8 μm	LSE of 12 μm	Atmospheric Profiles
Simulated SEVIRI radiances	IR radiometer measurements		Laboratory measurements		RAOB
Calculated radiances	$x_r = x_t + E(\delta x_t)$				GFS forecast
	δx_t = 3 K	δx_t = 0.05	δx_t = 0.025	δx_t = 0.015	

[a]The LST and LSE for calculated radiances are generated using $x_r = x_t + E(\delta x_t)$, where x_r is the estimated parameter, x_t is the true parameter used to simulate SEVIRI radiances, and $E(\delta x_t)$ is a random number with a bias of 0 and an standard deviation of δx_t.

Table 2. Spatial Resolution, Time Coverage, Spectral Channels, and Version of the LSE Databases

Database	Spatial Resolution (Degrees)	Date	Available Channels	Version
MODIS-A	0.05	Aug 2006	3.75, 3.959, 4.05, 8.55, 11.03, and 12.02 μm	4
MODIS-T	0.05	Aug 2006	3.75, 3.959, 4.05, 8.55, 11.03, and 12.02 μm	4
UWiremis	0.05	Aug 2006	5 cm^{-1} between 699.3 and 2774.3 cm^{-1}	4
AIRS-A	1	Aug 2006	3.82, 8.31, 10.41, and 12.02 μm	5
AIRS-D	1	Aug 2006	3.82, 8.31, 10.41, and 12.02 μm	5
IASI	0.5	Aug 2007	0.25 cm^{-1} between 645 and 2760 cm^{-1}	N/A

AIRS LSE, the IASI monthly research product is available at its original spectral resolution of 0.25 cm^{-1}. August 2007 data are used instead of August 2006 data because IASI data became available in October 2006.

[18] Table 2 summarizes the spatial, temporal, and spectral characteristics of each database. All of them have different spectral characteristics from SEVIRI (see Figure 2, note there is an offset of 0.03 between spectra of IASI and UWiremis for illumination). They are linearly interpolated to the SEVIRI spectrum except for IASI and UWiremis, which are convolved to the SEVIRI spectrum using the SEVIRI SRF.

3. Methodology

[19] In this study, we use standard deviation to describe the Tb deviation, which is defined as

$$\delta^2 = \overline{\left[Tb_c - Tb_o - \overline{(Tb_c - Tb_o)}\right]^2}, \quad (1)$$

where δ is the Tb deviation, Tb_c and Tb_o are the calculated and observed Tb, and $\bar{}$ is the mean of the samples under consideration. Let

$$\delta Tb = Tb_c - Tb_o - \overline{(Tb_c - Tb_o)}. \quad (2)$$

Equation (1) can be written as

$$\delta^2 = \overline{\delta Tb^2}. \quad (3)$$

[20] The simplest way to evaluate LSE using satellite observations is to compare the observed radiances with the calculated radiances. If the Tb deviation is small, the emissivity retrievals are likely good. However, this comparison is not conclusive because there are three components of error sources in the radiative transfer calculation: the atmospheric profile, LST, and the surface emissivity. Without knowing the error contribution from the other two components and the error correlations among the three components, it is difficult to evaluate and quantify LSE precision with Tb deviations.

[21] To study the impact of each individual component on Tb deviation, four experiments are conducted using the matchup database introduced in section 2.2. In the first three experiments, only one component is set as an estimate while the others are set as true. Therefore, in each experiment, the calculated Tb deviation is caused by the error from the corresponding component only. In experiment 1, the profiles are estimated profiles (GFS forecast profiles); in experiment 2 (EXP2), the LST is from estimation; and in experiment 3, the LSE is from estimation. In experiment 4 (EXP4), all three components are from estimates so the calculated Tb deviation in EXP4 is the total Tb deviation. Since the errors of LST and LSE are randomly assigned, the error correlations among the three components can be ignored (more details are given later). Therefore, the total Tb deviation $\hat{\delta}$ can be expressed as

$$\hat{\delta}^2 = \delta_a^2 + \delta_\varepsilon^2 + \delta_{ts}^2, \quad (4)$$

where δ_a, δ_ε, and δ_{ts} are Tb deviations for the atmospheric profiles, LSE, and LST, respectively. Quantifying LSE precision is equivalent to solving the LSE Tb deviations δ_ε for a set of given total Tb deviations.

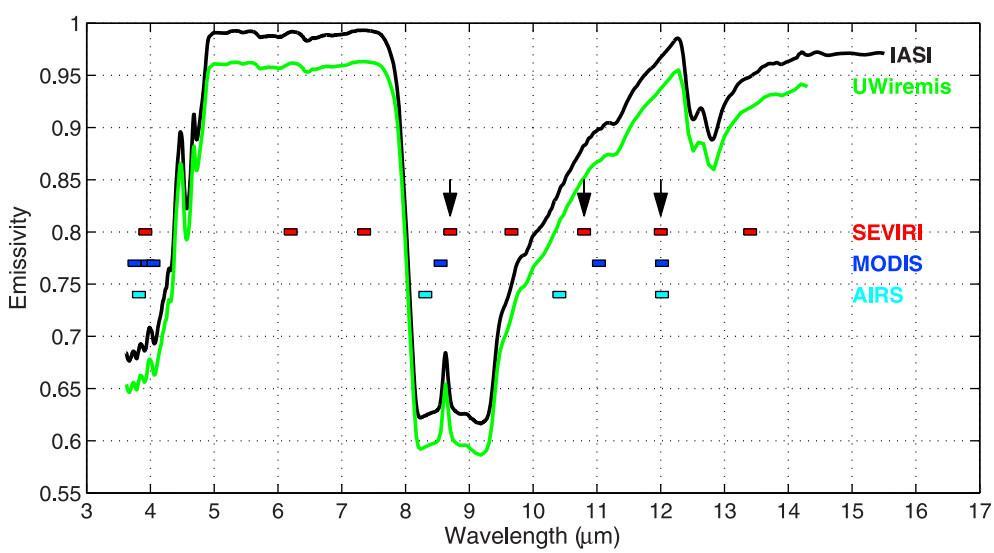

Figure 2. The spectral coverage of different LSE databases. The three SEVIRI window channels are highlighted by the black arrows to evaluate the LSE precision.

Table 3. Tb Deviations Caused by Errors of Atmospheric Profiles, LST and LSE[a]

	δ_a	δ_{ts}	δ_ε	$\hat{\delta}$[b]
8.7 μm	0.50	2.17	1.76	2.83
10.8 μm	0.54	2.32	1.14	2.63
12 μm	0.71	2.13	0.76	2.34

[a]Numbers are in Kelvin (K).
[b]The last column shows the total Tb deviations. Note that both LST and LSE have large contributions to the total Tb deviations. The values are calculated with the matchup data from southern Great Plains Atmospheric Radiation Measurement site.

[22] The results of the four experiments are shown in Table 3. The error of the profiles has a small impact on the Tb calculations in the IR window region; the Tb deviations are only 0.50, 0.54, and 0.71 K for 8.7, 10.8, and 12 μm, respectively. The 12 μm channel is more sensitive to profiles because it is more sensitive to water vapor. It should be noted that using more accurate atmospheric profiles from more sophisticated NWP models could further reduce this impact. However, less accurate atmospheric profiles would have larger impacts on the Tb calculations. The LST has much larger impacts; the Tb deviations are about three to four times as large as those caused by the atmospheric profile error. The LSE error affects different channels in different ways. The LSE of 8.7 μm was given a standard deviation error of 0.05, the largest of the three, and it has the largest Tb deviation (1.76 K) among the three channels. The LSE of the other two channels are given smaller standard deviation errors, and the Tb deviations are smaller. Especially for 12 μm, the Tb deviation is only 0.77 K, which is comparable to that caused by the atmospheric profiles.

[23] The previous analysis shows that LST has a greater impact on Tb deviations than LSE, both of which have a greater impact than the atmospheric profiles. The GFS forecast is an operational product, and its performance should be stable. So in equation (4), δ_a might be approximated using the values from the second column of Table 3 if the sample size is large enough. However, it is still impossible to solve LSE Tb deviation from equation (4) because LST contributions are also unknown. Since LST contributions are even larger than those from LSE, they cannot be ignored. In this study, the new method uses Tb differences between different channels to evaluate the LSE precision. All three SEVIRI channels are in TIR window regions and have the same LST. As will be shown later by Table 4, channel Tb differences greatly reduce the common impacts from the LST, which is very useful for LSE evaluation. As a result, channel differences are much less sensitive to LST and atmospheric profiles than LSE.

[24] To illustrate the advantage of using channel Tb differences, the same four experiments are performed, except that the three "channels" are three "channel differences," specifically "channel 1" is 10.8–8.7 μm, "channel 2" is 12–10.8 μm, and "channel 3" is 8.7–12 μm. Instead of Tb deviations, differences in brightness temperature (DTb) deviations Δ, defined as the standard deviation of differences in channel difference between calculations and observations, are used: The letters i and j denote the channel indices. Substituting equation (2) into equation (5), we have

$$\Delta_{ij}{}^2 = \overline{\left(\delta Tb_i - \delta Tb_j\right)^2}. \tag{6}$$

[25] Columns 2–5 of Table 4 show the individual and total DTb deviations ($\hat{\Delta}$) from the three components for the three "channels." Although the DTb deviation caused by atmospheric profiles (column 2) is only slightly reduced, their contribution to the total DTb deviation is very small. On the contrary, the LST impacts are significantly reduced; the DTb deviations are reduced from above 2 K (see column 3 of Table 3) to less than 0.6 K, indicating that DTb deviations are much less sensitive to LST error. At the same time, LSE impacts still remain strong and have in fact slightly increased.

[26] The atmospheric profiles are from the GFS forecast, which is a mature operational product. Statistically it is expected that the precision remains fairly consistent and stable if the sample size is large enough. Therefore, Tb and DTb deviation caused by the error of the atmospheric profiles should be consistently small. On the other hand, the LST error has a wide range depending on the surface/weather conditions and the algorithm used to estimate it. Over cold and dry areas, the precision is usually better due to less impact from moisture. Over warm and moist areas or desert areas, the precision is usually worse due to more moisture and worse LSE precision. Therefore, it is necessary to ensure that the DTb deviations remain insensitive to LST error. Figure 3 shows the impacts from EXP2. As the precision of the LST degrades from 3 to 5 K, Tb deviations increase rapidly by at least 1.4 K, while DTb deviations increase slowly by less than 0.13 K.

[27] The last column in Table 4 shows the percentage of the contribution by the LSE in the total DTb deviations using $P_\varepsilon = \frac{\Delta^2}{\hat{\Delta}^2}$. For all three channels, the LSE contribution is more than 93%. Therefore, the contribution by the atmosphere and LST can be omitted. Ignoring the error correlations among the three components, the LSE DTb deviation can be approximated using

$$\Delta_{\varepsilon,ij}^2 = \hat{\Delta}_{ij}^2, \tag{7}$$

where $\Delta_{\varepsilon,ij}$ is the LSE DTb (between the ith and jth channels) deviations, and $\hat{\Delta}_{ij}$ is the total DTb deviation. The next step is to estimate the LSE Tb deviations for given DTb deviations. According to the definition, LSE DTb deviation can be expressed as

$$\Delta_{\varepsilon,ij}^2 = \overline{\left(\delta Tb_{\varepsilon,i} - \delta Tb_{\varepsilon,j}\right)^2} = \delta_{\varepsilon,i}^2 + \delta_{\varepsilon,j}^2 - 2\overline{\delta Tb_{\varepsilon,i}\delta Tb_{\varepsilon,j}}, \tag{8}$$

where $\delta_{\varepsilon,i}$ and $\delta_{\varepsilon,j}$ are the LSE Tb deviation for the ith and jth channels, respectively, and $\overline{\delta Tb_{\varepsilon,i}\delta Tb_{\varepsilon,j}}$ is the correlative deviation between the ith and jth channels. Therefore,

$$\delta_{\varepsilon,i}^2 + \delta_{\varepsilon,j}^2 - 2\overline{\delta Tb_{\varepsilon,i}\delta Tb_{\varepsilon,j}} = \hat{\Delta}_{ij}^2. \tag{9}$$

$$\Delta_{ij}{}^2 = \overline{\left\{Tb_{c,i} - Tb_{c,j} - \left(Tb_{o,i} - Tb_{o,j}\right) - \overline{\left[Tb_{c,i} - Tb_{c,i} - \left(Tb_{o,j} - Tb_{o,j}\right)\right]}\right\}^2}. \tag{5}$$

Table 4. DTb Deviations Caused by Error of Atmospheric Profiles, LST and LSE[a]

	Δ_a	Δ_{ts}	Δ_ε	$\hat{\Delta}$[b]	P_ε[c]
10.8–8.7 μm	0.50	0.48	2.11	2.15	96.3
12–10.8 μm	0.51	0.53	1.35	1.4	93.0
8.7–12 μm	0.60	0.53	1.88	1.94	93.9

[a]Numbers are in Kelvin (K).
[b]This column shows the total DTb deviations.
[c]The percentage of the contribution by the LSE shown in the last column shows that the LSE is the most dominant contribution to the total DTb deviation. The values are calculated with the matchup data from southern Great Plains Atmospheric Radiation Measurement site.

Or, for the three channel differences, we have

$$\begin{aligned} \delta_{\varepsilon,1}^2 + \delta_{\varepsilon,2}^2 - 2\overline{\delta Tb_{\varepsilon,1}\delta Tb_{\varepsilon,2}} &= \hat{\Delta}_{12}^2 \\ \delta_{\varepsilon,2}^2 + \delta_{\varepsilon,3}^2 - 2\overline{\delta Tb_{\varepsilon,2}\delta Tb_{\varepsilon,3}} &= \hat{\Delta}_{23}^2 \\ \delta_{\varepsilon,1}^2 + \delta_{\varepsilon,3}^2 - 2\overline{\delta Tb_{\varepsilon,1}\delta Tb_{\varepsilon,3}} &= \hat{\Delta}_{13}^2. \end{aligned} \tag{10}$$

[28] Assume there are no correlative deviations between different channels. This is equivalent to assuming there are no LSE error correlations between different channels. Without a large sample size of laboratory measurements, we are not able to directly evaluate this assumption. However, the impacts of this assumption will be discussed later in sections 4 and 5. Then equation (10) can be written as

$$\begin{aligned} \delta_{\varepsilon,1}^2 + \delta_{\varepsilon,2}^2 &= \hat{\Delta}_{12}^2 \\ \delta_{\varepsilon,2}^2 + \delta_{\varepsilon,3}^2 &= \hat{\Delta}_{23}^2 \\ \delta_{\varepsilon,1}^2 + \delta_{\varepsilon,3}^2 &= \hat{\Delta}_{13}^2. \end{aligned} \tag{11}$$

[29] These three equations can be solved to obtain the LSE Tb deviation of each channel. As an example, the total DTb deviations in Table 4 (column 5) are used to solve equation (11). The new equations are

$$\begin{aligned} \delta_{\varepsilon,1}^2 + \delta_{\varepsilon,2}^2 &= 2.15^2 \\ \delta_{\varepsilon,2}^2 + \delta_{\varepsilon,3}^2 &= 1.4^2 \\ \delta_{\varepsilon,1}^2 + \delta_{\varepsilon,3}^2 &= 2.01^2. \end{aligned} \tag{12}$$

and the solutions are

$$\begin{aligned} \delta_{\varepsilon,1} &= 1.83 \text{ K} \\ \delta_{\varepsilon,2} &= 1.13 \text{ K} \\ \delta_{\varepsilon,3} &= 0.83 \text{ K}. \end{aligned} \tag{13}$$

Compared to the LSE Tb deviation in column 4 of Table 3, the solutions are very close to the actual values, and the differences are less than 0.07 K. Or in the LSE space (see equations (14) and (15) on how to estimate LSE precision from LSE Tb deviation), the differences are less than 0.0016 (0.16%). These are excellent estimations of LSE precision considering the uncertainty of laboratory measurements is 0.002 (0.2%) [*Korb et al.*, 1999].

[30] The next step is to estimate LSE precision from the estimated Tb deviations. Letting the emissivity weighting function of the ith channel be $K_{\varepsilon,i}$, the LSE precision can be estimated using

$$\sigma_{\varepsilon,i} = \sqrt{\overline{\left(\frac{\delta Tb_{\varepsilon,i}}{K_{\varepsilon,i}}\right)^2}}, \tag{14}$$

where $\sigma_{\varepsilon,i}$ is the LSE precision of the ith channel. In reality, there is no way to determine $\delta Tb_{\varepsilon,i}$, and an approximation is made

$$\sigma_{\varepsilon,i} = \delta_{\varepsilon,i}\sqrt{\overline{\left(\frac{1}{K_{\varepsilon,i}}\right)^2}}. \tag{15}$$

With the estimated LSE Tb deviation, the LST Tb deviation can be calculated using the equation

$$\delta_{ts} = \sqrt{\hat{\delta}^2 - \delta_a^2 - \delta_{\hat{\varepsilon}}^2}. \tag{16}$$

Using the data from Table 3, the calculated LST Tb deviation is

$$\begin{aligned} \delta_{ts,1} &= 2.10 \text{ K} \\ \delta_{ts,2} &= 2.31 \text{ K} \\ \delta_{ts,3} &= 2.07 \text{ K}. \end{aligned} \tag{17}$$

[31] Compared to the LST Tb deviation in column 3 of Table 3, the calculations are very close to truth, and the differences are very small (less than 0.07 K). Letting the LST weighting functions of the ith channels be $K_{ts,i}$, LST precision can be calculated using

$$\sigma_{ts,i} = \delta_{ts,i}\sqrt{\overline{\left(\frac{1}{K_{ts,i}}\right)^2}}. \tag{18}$$

The estimated LST precision will be examined as a self-consistency test to evaluate the effectiveness of the new

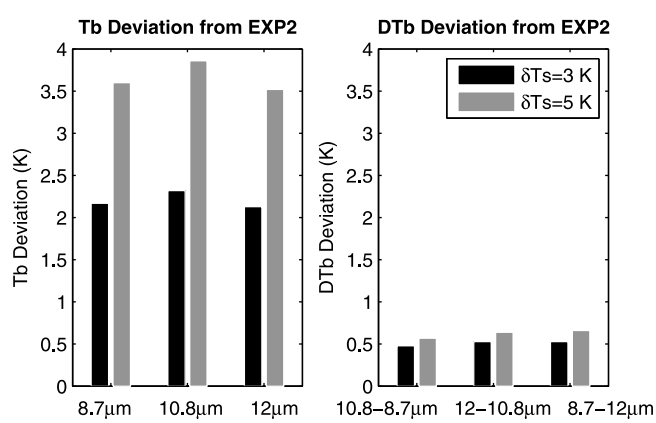

Figure 3. The impacts of LST precision on (left) Tb deviation and (right) DTb deviation from experiment 2 from the simulation study using the matchup data set from southern Great Plains ARM site. The black color represents an LST with a precision of 3 K and the gray color represents an LST with a precision of 5 K. Note the Tb deviation is very sensitive, while DTb deviation is much less sensitive to the LST precision.

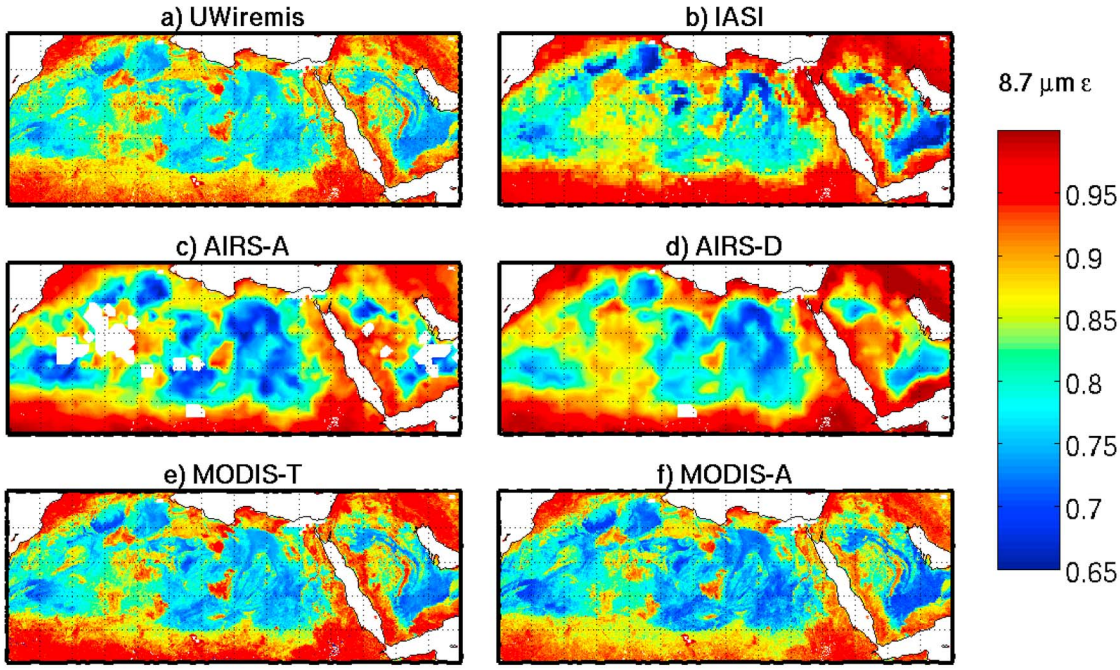

Figure 4. LSE images of 8.7 μm from (a) UWiremis, (b) IASI, (c) AIRS-A, (d) AIRS-D, (e) MODIS-T, and (f) MODIS-A. All images are monthly products for August 2006 except IASI of August 2007 over the desert areas (10°N–35°N and 15°W–55°E) in Africa and part of Asia. Blank areas are either ocean or no data.

method. It is expected that different channels and different databases have similar estimated LST precision since the same LST products are used.

4. LSE Evaluation With SEVIRI Radiance Measurements

[32] In this section, radiances are calculated with PFAAST using the estimated atmospheric profiles from GFS 6 h forecasts, the LST from the European Centre for Medium-Range Weather Forecasts (ECMWF) analysis, and the LSE from the six LSE databases. They are compared to the SEVIRI measurements at three window channels (8.7, 10.8, and 12 μm) from August 2006, which are used as references.

[33] Figures 4–6 show LSE imagery for 8.7, 10.8, and 12 μm from different databases over the desert area (10°N–35°N and 15°W–55°E) in Africa and part of Asia for August 2006. For 8.7 μm, the LSE varies geographically according to land surface condition. Over the desert areas the LSE is small (<0.8), while over most other areas the LSE is large (>0.9). All six databases show a similar geographical distribution, but with visible differences. For 10.8 and 12 μm, the LSE looks quite different. The AIRS and IASI LSE show much larger spatial variations than the MODIS products (UWiremis, MODIS-T. and MODIS-A). Some of the AIRS LSE are even larger than 1.0 (Figures 5d, 6c, and 6d). The geographical distributions are also quite different, especially for 12 μm. Quantitative comparisons (Table 5) show there are substantial differences among different LSE databases. The 8.7 μm, due to the large natural variations, has the largest differences (up to 0.035). Although 10.8 and 12 μm have much smaller natural variation, the differences could still be well above 0.02. With these substantial differences between the six databases, users will have difficulty determining which database meets their needs. Note the MODIS-A LSE is more similar to UWiremis because the latter is derived from the former. In this section, quantitative evaluation will be performed on these databases. The demonstration in the last section does not include any correlative Tb deviations. There are two types of error correlations: Type 1 is the correlation among the three components for radiative transfer calculations (atmospheric profiles, LST, and LSE), and type 2 is the LSE error correlation among three different channels. Without extensive laboratory measurements, we are not able to quantify these correlations. However, when applying this method to the real data, efforts are made to minimize the impacts of ignoring those correlations.

[34] Equations (4) and (7) assume that there are no type 1 correlations. The simulation study in the previous section ensures the noncorrelation by randomly generating LSE and LST. When applying the method to real data, it is believed that the type 1 correlations are reduced if the three components are taken from three independent data sources. In this section, the atmospheric profiles are from the GFS 6 h forecast, and the LST is from the ECMWF analysis. They are both from August 2006 and temporally interpolated to SEVIRI observation time.

[35] Ignoring type 2 correlations is the basis of transferring equation (10) to (11). However, type 2 correlations exist in almost all the databases. When these correlations are strong and ignored, equation (11) is problematic, and the solutions are no longer reliable. As an extreme case, the total DTb deviations using the MODIS-A LSE database are 0.95, 0.64, and 1.25 K for the three channel differences. Using equation (11), no realistic solutions are found, indicating there are strong LSE error correlations among the three window channels.

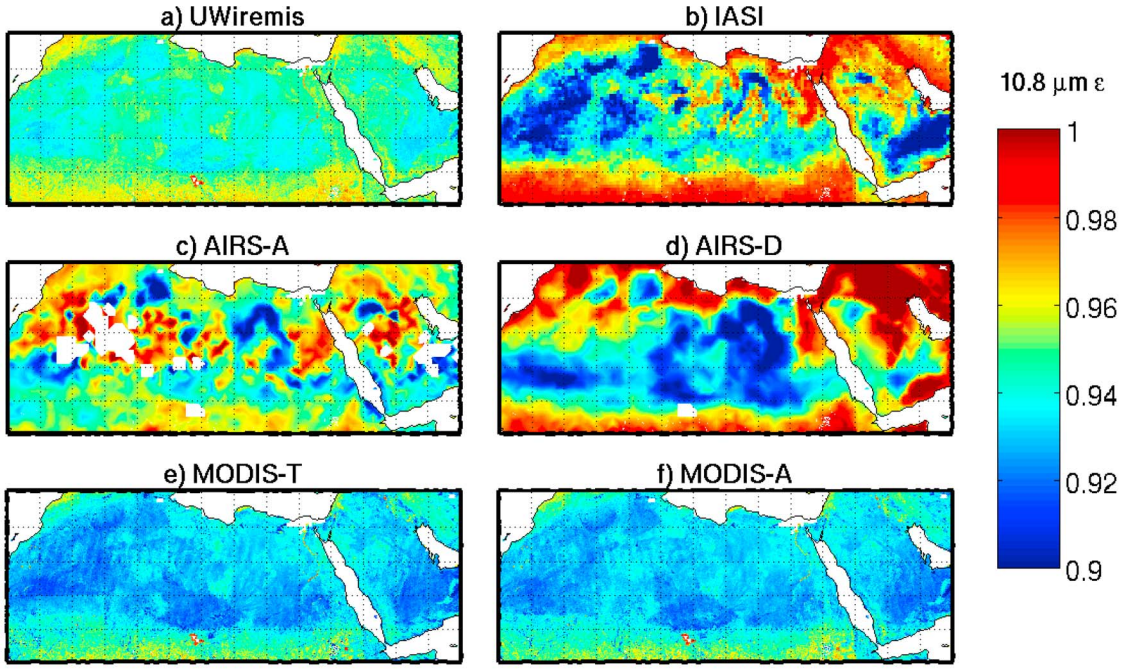

Figure 5. Same as Figure 4 except for 10.8 μm.

[36] Instead of trying to quantify type 2 correlations, efforts are made to find situations with the least correlation. For example, to estimate the UWiremis 8.7 μm LSE Tb deviation using SEVIRI radiance measurements, the 10.8 μm LSE could be chosen from six databases, as could the 12.0 μm LSE. Therefore, there are a maximum of 36 solutions for calculating the UWiremis 8.7 μm LSE Tb deviation. Some of the solutions are not realistic. Figure 7a shows all the realistic LSE Tb deviations of UWiremis 8.7 μm estimated with SEVIRI radiance measurements. Unrealistic values are defined to be imaginary numbers or real numbers less than 0. The x axis represents the LSE database used for 12 μm LSE, and y axis represents the LSE database used for 10.8 μm. Different colors represent different values of the estimated UWiremis 8.7 μm LSE Tb deviation. For example, the value at (3, 2) or (AIRS-A, IASI) is 0.99 K, meaning that the LSE of 10.8 μm is taken from IASI and the LSE of 12 μm is taken from AIRS ascending, and the estimated UWiremis 8.7 μm LSE Tb deviation is 0.99 K.

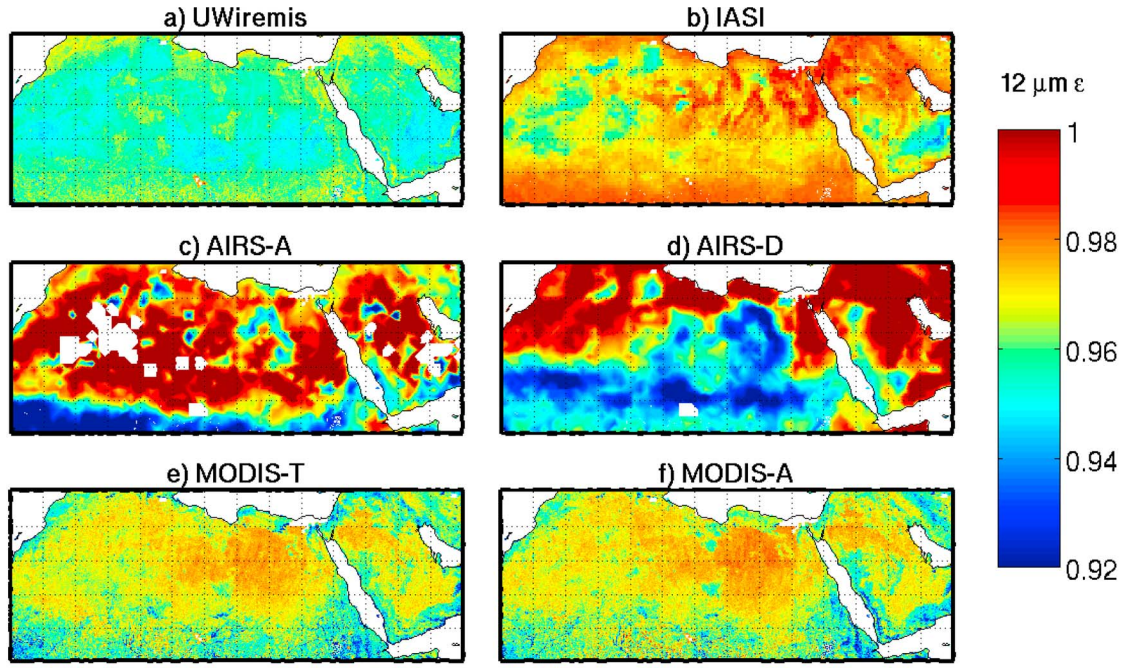

Figure 6. Same as Figure 4 except for 12 μm.

Table 5. Standard Deviations of the Differences of the LSE Products Between UWiremis and Others Over the Desert Areas (10°N–35°N and 15°W–55°E) in Africa and Part of Asia[a]

Channel	IASI	AIRS-A	AIRS-D	MODIS-T	MODIS-A
8.7 μm	0.035	0.034	0.035	0.019	0.010
10.8 μm	0.018	0.016	0.022	0.016	0.004
12 μm	0.008	0.029	0.022	0.019	0.011

[a]Note the MODIS-A LSE is more similar to UWiremis because the latter is derived from the former. The LSE databases are monthly products of August 2006 except IASI of August 2007.

[37] Figure 7a shows that different combinations of 10.8 and 12 μm LSE return different values for UWiremis 8.7 μm LSE Tb deviations. However, most solutions are clustered between 0.89 and 1.15 K, or they are very close to each other. Figures 7b–7f show the results for the other five databases. Similar to UWiremis, those results look clustered. These results indicate that the omission of correlation terms has little impact on precisely estimating 8.7 μm LSE Tb deviation. The same approach is applied to estimating the LSE Tb deviations for 10.8 and 12 μm (Figures 8 and 9). Although the estimated Tb deviations are still somewhat clustered for each database, they look more scattered than those for the 8.7 μm (Figure 7). This scattering indicates the estimation of 10.8 and 12 μm LSE is more affected by omitting the correlation terms.

[38] From Figures 7–9, each estimated Tb deviation has many realistic values. For example, there are 22 different realistic values for UWiremis 8.7 μm in Figure 7a. It is important to screen out those correlation-contaminated solutions. The detection of correlation-contaminated combinations is based on the following steps:

[39] 1. For any combination and any channel, if the solution is smaller than 0.2 K, it is considered correlation contaminated. An LSE Tb deviation of 0.2 K is equivalent to an LSE precision of 0.005, which is unlikely for the satellite-derived LSE product.

[40] 2. From Figures 8 and 9, the error correlation between 10.8 and 12 μm is much stronger than the others. Therefore, any combination that 10.8 and 12 μm come from the same database is abandoned.

[41] 3. A simple statistical method is used to detect the remaining combinations that suffer from error correlation. The solutions are assumed of the distribution of normality. For each solution, a normalized distance to the mean value is calculated using

$$D_i = \frac{|x_i - \bar{x}|}{\sigma}, \tag{19}$$

where D_i is the normalized distance from the mean, x_i is the ith solution, $\bar{x}$ is the mean of all the solution, and σ is the standard deviation of the solutions. The thresholds are 1, 1.5, and 1 for 8.7, 10.8, and 12 μm respectively. 10.8 μm is

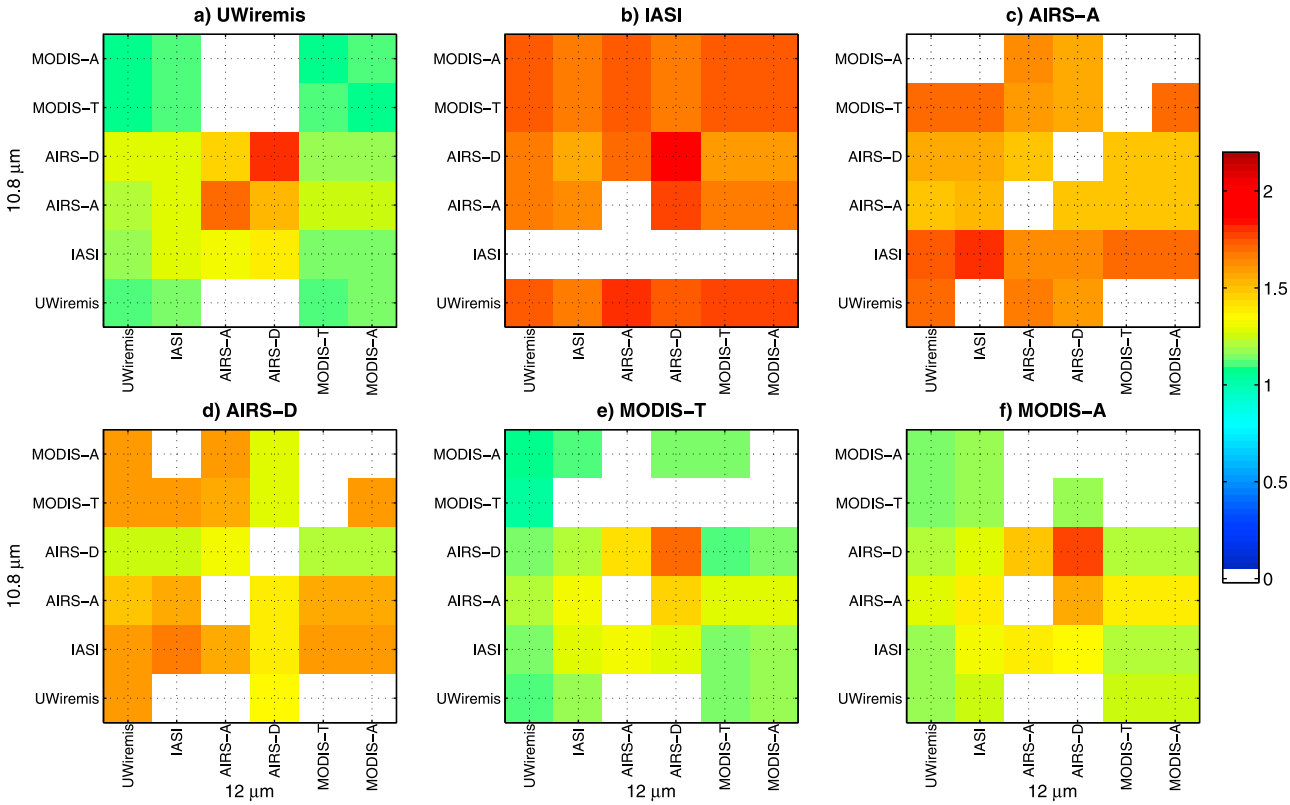

Figure 7. The estimated 8.7 μm LSE Tb deviations from different combinations of LSE at 10.8 and 12 μm. All six LSE databases are from August 2006 except IASI of August 2007 over the desert areas (10°N–35°N and 15°W–55°E) in Africa and part of Asia. The SEVIRI radiance observations from August 2006 are used as references. The x axis represents databases used for 12 μm LSE, and the y axis represent databases used for 10.8 μm LSE. The blank elements represent no realistic solutions.

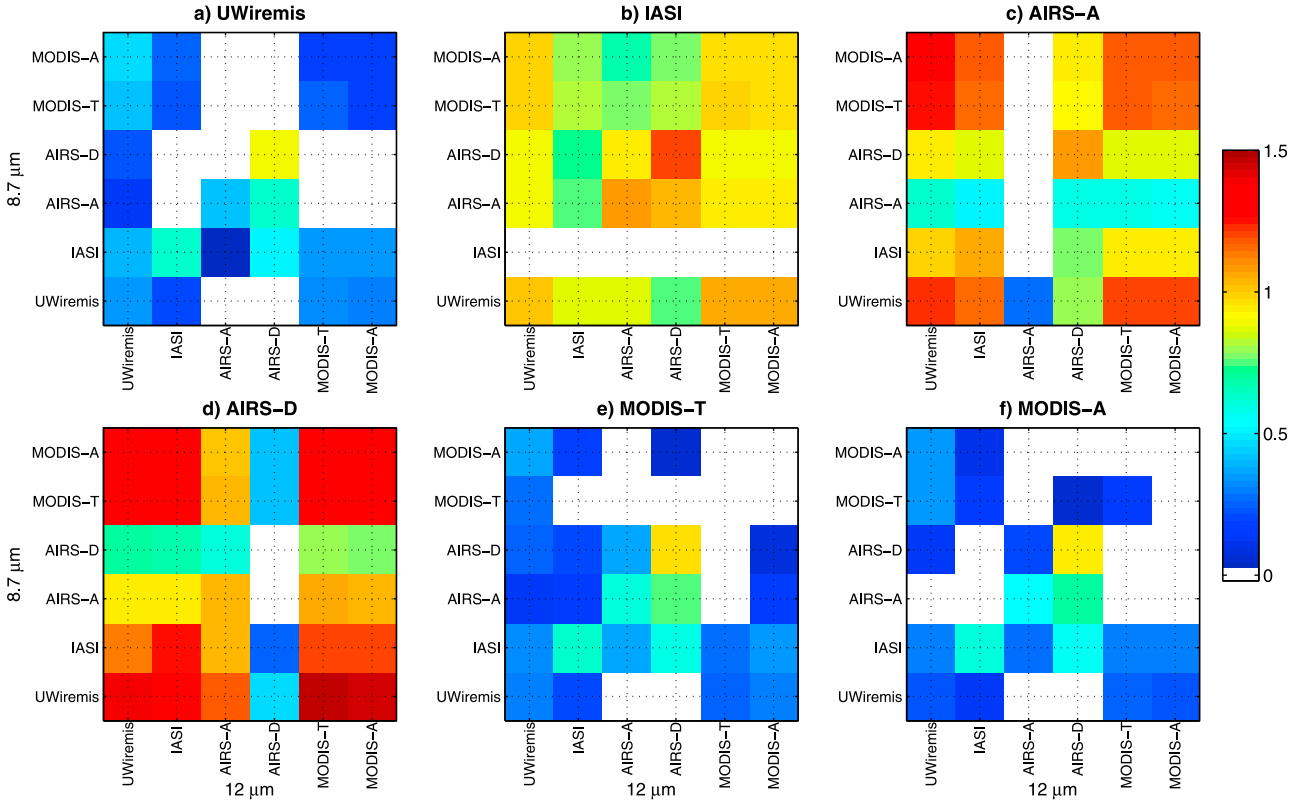

Figure 8. Same as Figure 7 except for 10.8 μm. The x axis represents databases used for 12 μm LSE, and the y axis represents databases used for 8.7 μm LSE.

given a larger threshold because the solutions shown in Figure 8 appear more scattered than in Figures 7 and 9. Any solution with a distance larger than the given threshold is abandoned.

[42] For a specific combination, if the estimated Tb deviation for any channel is considered correlation contaminated, the estimated Tb deviations of all three channels are not reliable. For example, in Figure 8b, if the 8.7 μm LSE is from MODIS-A and the 12 μm LSE is from AIRS-A, the estimated 10.8 μm IASI Tb deviation is 0.673 K, much smaller than the majority of the others. Therefore, this combination of the three channels is considered correlation contaminated. As a result, the estimated AIRS-A Tb deviation of 12 μm is significantly larger than the majority of others in Figure 9c. All three estimated Tb deviations from this combination are eliminated.

[43] After screening out all the suspicious combinations, the remaining combinations are averaged and assigned as the LSE Tb deviations. Table 6 shows the averaged Tb deviations for each database and each channel. For easier interpretation, the LSE Tb deviations are converted to LSE precisions using equation (15) (see Table 7). For 8.7 μm, the MODIS products (UWiremis, MODIS-T, and MODIS-A) show significantly better precision (around 0.030) than the IASI and AIRS (precisions larger than 0.038). However, the IASI product is a work in progress and may not be optimal yet. The AIRS-A has better precision than AIRS-D possibly because the physical retrieval is affected by improper modeling of reflected solar contamination in AIRS midwave IR region. The UWiremis improves 8.7 μm LSE precision by 0.002 because the hyperspectral LSE captures the spike (Figure 1) around 8.7 μm better than the original BF or the operational MODIS-A products. The latter two essentially give a flat value for LSE between 8.3 and 9.3 μm.

[44] Although the LSE natural variation in the 10.8 and 12 μm regions is much smaller than in the 8.7 μm region (Figure 1), physical improvement of the 10.8 and 12 μm LSE by AIRS seems to be a failure. AIRS-A has precisions of 0.0209 and 0.0377 for 10.8 and 12 μm, and AIRS-D has precisions of 0.0275 and 0.0317. All of these precisions are worse than the others in this particular case. The quality of AIRS LSE can also be seen from the LSE imagery. In Figures 5d, 6c, and 6d, the AIRS product shows many areas with LSE smaller than 0.9 (dark blue) over the desert areas, which might not be realistic. Figure 10a shows all the laboratory-measured LSE spectra with LSE less than 0.8 in the 8.7 μm region (they are considered desert samples). Among these spectra, except two samples, all of them have an LSE larger than 0.9 at 10.8 μm. For 12 μm, most of them are larger than 0.95. Figure 10b shows all the laboratory-measured LSE spectra with LSE less than 0.9 in the 10.8 μm region. These spectra are unlikely from desert because all LSE at 8.7 μm are larger than 0.85 except for two. In Figures 5d, 6c, and 6d, the AIRS product shows many areas with an LSE larger than 1.0, which is not realistic. *Hulley et al.* [2009b] also found many AIRS LSE larger than 1.0. All of these unrealistic values of AIRS contribute to less accurate LSE precision.

[45] It is interesting that MODIS products (UWiremis, MODIS-T, and MODIS-A) show surprisingly better results

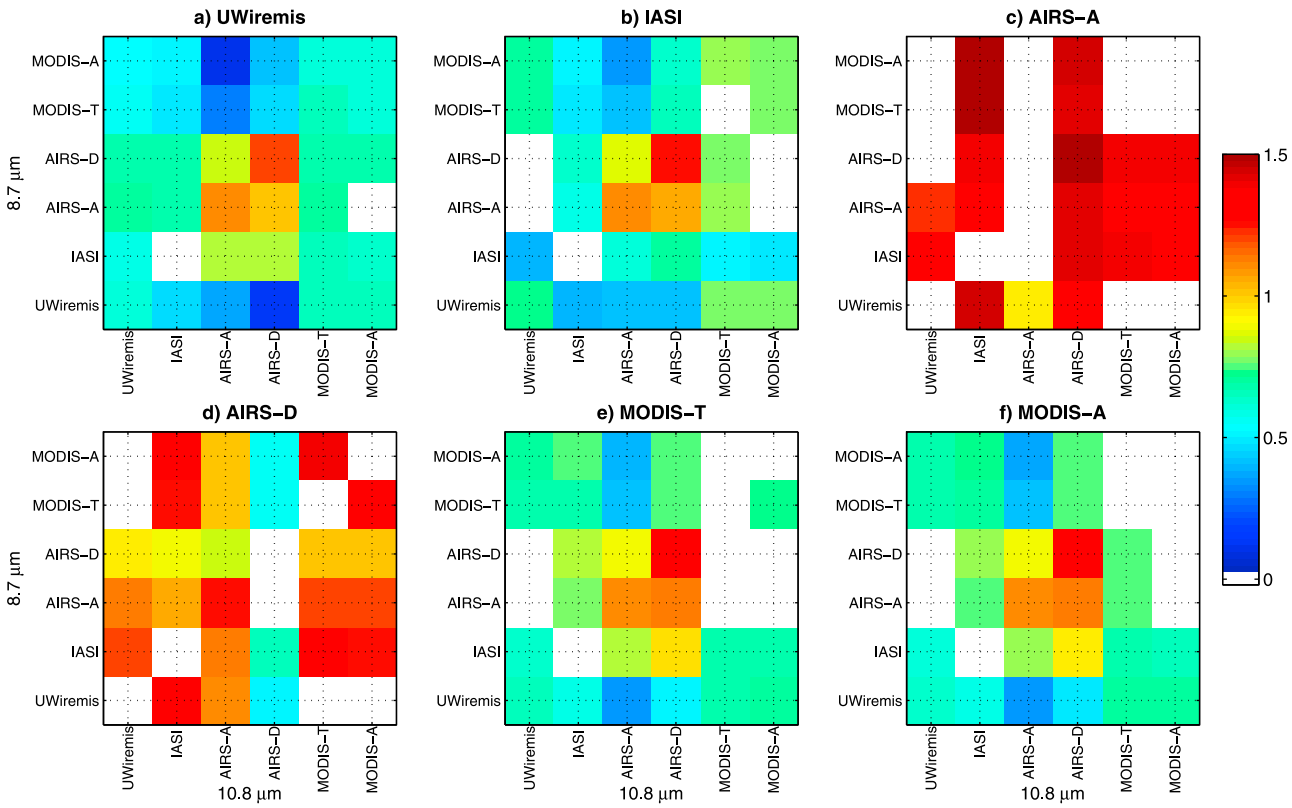

Figure 9. Same as Figure 7 except for 12 μm. The x axis represents databases used for 10.8 μm LSE, and the y axis represents databases used for 8.7 μm LSE.

for 10.8 and 12 μm. The estimated LSE precisions are all smaller than 0.019. AIRS and IASI have precisions all larger than 0.017. In the split-window algorithm, the MODIS 10.8 and 12 μm LSE are assigned based on classification-based modeling [*Wan*, 2008; *Snyder et al.*, 1998]. In the day/night algorithm, some knowledge of the classification-based emissivity is used in the assignment of the initial values of MODIS 10.8 and 12 μm LSE. The AIRS algorithm uses a similar approach for the initial guess in the regression. Our results show that the day/night algorithm is better on handling LSE at 10.8 and 12 μm than the AIRS physical algorithm, indicating that the classification-based modeling seems to provide quality initial values for physical iteration. The disadvantage of classification-based modeling is its inability to capture the temporal variation. However, the temporal variation in 10–12 μm might be very weak [*Wan*, 2008]. The reason why AIRS LSE product has worse precision than MODIS products seems to be the difficulty in the physical iteration, such as cloud contamination due to large footprint size. Table 7 shows that 12 μm has worse precision than 10.8 μm despite the fact that LSE at 12 μm has smaller natural variations than at 10.8 μm (Figure 1). The reason for this difference is still under investigation.

[46] As an indirect way to evaluate the new method, one would expect the estimated LST precisions by different databases and different channels to be similar. The LST precisions are estimated using equations (16) and (18), and results are presented in Table 8. They are all between 4.47 and 4.93 K with a mean of 4.65 K and a median of 4.63 K. Figure 11 shows the histogram of the estimated LST precisions. The small standard deviation of 0.16 K indicates the estimated LST precisions are consistent. These results add confidence to the new method.

5. Discussion

[47] Although the method presented provides quantitative evaluations of LSE precision, it cannot be used for validation because no true LSE measurements are used. There are four error sources that may potentially cause the LSE Tb deviation estimation to be inaccurate in equation (11): (1) the uncertainty in the radiative transfer calculation and satellite observation noise, (2) the uncertainty in estimating DTb deviations due to inaccurate atmospheric profiles and LST, (3) type 1 error correlations (among the three components for radiative transfer calculations), and (4) type 2 error correlations (among LSE of the three channels). Among the four

Table 6. Estimated LSE Tb Deviations Over the Desert Area (10°N–35°N and 15°W–55°E) Using SEVIRI Radiance Observations of August 2006 as References[a]

	8.7 μm	10.8 μm	12 μm
UWiremis	1.186	0.398	0.590
IASI	1.695	0.928	0.657
AIRS-A	1.567	0.952	1.402
AIRS-D	1.532	1.253	1.178
MODIS-T	1.202	0.427	0.676
MODIS-A	1.249	0.447	0.683

[a]Standard deviation, in Kelvin (K). The LSE databases are monthly products of August 2006 except IASI of August 2007.

Table 7. Estimated LSE Precision Over the Desert Area (10°N–35°N and 15°W–55°E) Using SEVIRI Radiance Observations of August 2006 as References[a]

	8.7 μm	10.8 μm	12 μm
UWiremis	0.0299	0.0087	0.0159
IASI	0.0427	0.0203	0.0177
AIRS-A	0.0395	0.0209	0.0377
AIRS-D	0.0386	0.0275	0.0317
MODIS/Terra	0.0303	0.0094	0.0182
MODIS/Aqua	0.0314	0.0098	0.0184

[a]Standard deviation, unitless. The LSE databases are monthly products of August 2006 except IASI of August 2007.

sources, the uncertainty in the radiative transfer calculation and the satellite observations altogether is estimated to be less than 0.25 K from the intercomparison of radiative transfer models and instrument characteristics. Furthermore, this uncertainty is the only one that is always positive and can be deducted from the right-hand side of equation (11). However, the deduction of $[0.25\ \mathrm{K}]^2$ is too small to have a substantial impact on the solution and thus can be neglected. The other three errors may be either positive or negative. It is also difficult to quantify them, especially the type 1 and type 2 error correlations. However, it is worthy to investigate the sensitivity of the LSE Tb deviation estimates to the error in equation (11). If a solution is insensitive to the error, it is considered a reliable estimation; if a solution is sensitive to the error, it is considered a less reliable estimation.

[48] To study that assertion, a specific combination of the three channels of LSE is chosen: 8.7 μm from IASI, 10.8 μm from MODIS-T, and 12 μm from AIRS-A. Using equation (11), the estimated LSE Tb deviations are 1.718, 0.384, and 1.398 K, which are very close to the values in Table 6 (1.695, 0.427, and 1.402 K). Error values between −0.60 and 0.60 K with an increment of 0.12 K are added to the right-hand side of equation (11). Consequently, each equation has 11 different errors. In total there are $11^3 = 1331$ error combinations. The histograms of all the realistic solutions from the 1331 combinations are shown in Figure 12. The dashed lines represent the solutions without any errors. The narrow distribution of IASI 8.7 μm and AIRS-A 12 μm indicates that they are insensitive to the error, which means the estimated IASI 8.7 μm and AIRS-A 12 μm Tb deviations are reliable. The broad histogram distribution of MODIS-T 10.8 μm, on the other hand, indicates that it is sensitive to the errors. Therefore, the estimated MODIS-T 10.8 μm Tb deviation is less reliable. Table 9 shows the confidence levels of the estimated Tb deviations at different confidence intervals. For example, the possibility of the estimated Tb deviation for IASI 8.7 μm falling between 1.695 ± 0.2 is 100%. The broader the confidence interval is, the larger the confidence level is. The confidence levels in Table 9 again show that the estimated Tb deviations for the IASI 8.7 μm and AIRS-A 12 μm are reliable, while the estimation for the MODIS-T 10.8 μm is less reliable.

[49] The previous discussion shows that a larger estimated Tb deviation is more reliable than a smaller one. This is a general conclusion, which can be drawn from mathematical analysis of equation (11). Therefore, the estimated LSE Tb deviations for 8.7 μm in Table 6 are reliable since they are all larger than 1 K. This can also be seen from Figure 7, as the estimated Tb deviations from different combinations are mostly clustered. The estimated LSE Tb deviations for the 10.8 μm of AIRS-D and the 12 μm of AIRS-A and AIRS-D are also reliable because they are larger than 1.0 K. The estimations for the 10.8 μm of IASI and AIRS-A and the 12 μm of UWiremis, IASI, MODIS-A, and MODIS-T are less reliable because their values are between 0.59 and 0.96. The remaining values are the least reliable since they are all less than 0.5 K. However, the conclusion that these LSE have better precisions than AIRS is still valid for two reasons: (1) The average of all reasonable estimations should reduce the uncertainty of the solutions and (2) if they have similar (even worse) precisions as AIRS, there is little chance for the estimates to be so small, according to Figure 12 and Table 9.

[50] It should be noted that all the analysis and discussion in this study is based on standard deviation of Tb and DTB differences. Any radiance biases from either observations or the radiative transfer model are not considered because they have no impact on the analysis.

6. Summary and Conclusions

[51] The goal of this study is to quantitatively evaluate LSE precision using satellite observations, SEVIRI IR window bands, in this study. For that purpose, a new efficient, objective, and conclusive method is introduced. To minimize the impact from LST, the new method uses

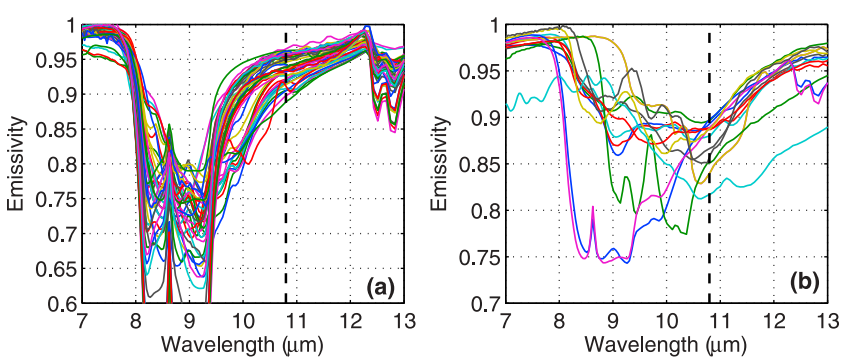

Figure 10. (a) LSE hyperspectra with 8.7 μm LSE less than 0.8 from the laboratory LSE measurements. (b) LSE hyperspectra with 10.8 μm LSE less than 0.9 from the laboratory LSE measurements. The dashed line represents the wavelength of 10.8 μm.

Table 8. Estimated Precisions for LST From European Centre for Medium-Range Weather Forecasts Analysis of August 2006 by Different Channels of Different LSE Databases Over the Desert Area (10°N–35°N and 15°W–55°E) Using SEVIRI Radiance Observations of August 2006 as References[a]

	8.7 μm	10.8 μm	12 μm
UWiremis	4.47	4.48	4.87
IASI	4.68	4.53	4.82
AIRS-A	4.80	4.57	4.77
AIRS-D	4.70	4.51	4.93
MODIS-T	4.50	4.50	4.82
MODIS-A	4.51	4.49	4.82

[a]All the values are clustered between 4.47 and 4.93 K, indirectly indicating that estimated LSE precisions are reliable.

channel DTb to derive the LSE Tb deviations, defined as the standard deviations of Tb differences between satellite observations and radiative transfer calculations. This method does not need the true LSE measurements. It only needs three components for the radiative transfer calculations: LST from the ECMWF analysis, atmospheric profiles from the GFS forecast, and LSE from satellite retrieval databases. There are no precision requirements for the first two components. It is believed that the error correlations among the three components are reduced since the three components are from three different sources. However, strong LSE error correlations could exist among the three channels, which are impossible to quantify without enough real LSE measurements. As a result, omitting the error correlation makes it difficult to accurately estimate the LSE Tb deviations, especially for 10.8 and 12 μm. Instead of trying to quantify error correlations, focus has been on identifying database combinations with fewer error correlations. Results are averaged and assigned as the estimated LSE Tb deviations, from which LSE precision is estimated.

[52] The method is applied to six LSE databases: the UWiremis database, the IASI preliminary research test product, the operational AIRS-A products, the operational AIRS-D products, the operational MODIS-A, and the operational MODIS-T products. All of them are monthly products from

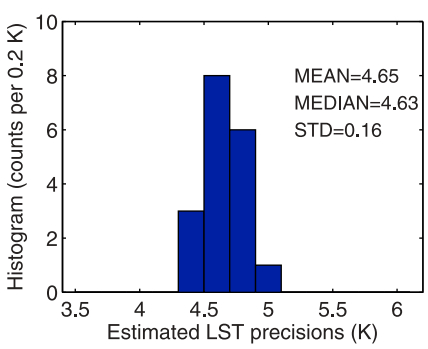

Figure 11. The histogram of the estimated precision for LST from European Centre for Medium-Range Weather Forecasts analysis of August 2006. There are 18 LST precisions from 6 LSE databases and 3 channels. The small standard deviation indicates the estimated LST precisions consistent from different channels and different LSE databases. This indirectly verifies that the estimated LSE precisions are reliable.

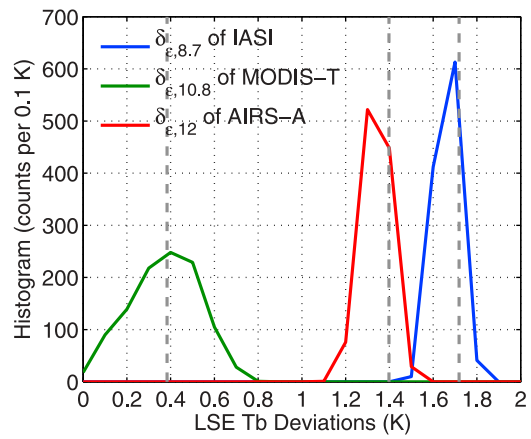

Figure 12. Histograms of the estimated LSE Tb deviations for the LSE combination of 8.7 μm from IASI, 10.8 μm from MODIS-T, and 12 μm from AIRS-A. The LSE Tb deviations are estimated with different errors added between −0.6 and 0.6 K with an increment of 0.12 K. Narrow distributions of IASI 8.7 μm and AIRS-A 12 μm indicate that they are insensitive to errors, therefore, their estimated LSE Tb deviations are more reliable than 10.8 μm of MODIS-T.

August 2006, except IASI of August 2007. All the LSE databases are spatially interpolated to SEVIRI grids. Spectrally, they are either linearly interpolated (broadband) or convolved (hyperspectral) to SEVIRI IR bands. The satellite observations are from SEVIRI at 0600 UTC on 1 August 2006. The results show that the MODIS LSE has much better precision (around 0.03) than AIRS and IASI (around 0.04) for 8.7 μm, which has the largest natural variations among the three channels. The UWiremis improves the operational MODIS-A products by 0.002 by better depicting the spectral variation around 8.7 μm. The AIRS has comparable precisions to IASI (around 0.04), although the latter could be further improved by physical iterations. However, the physical algorithm fails to improve the LSE of 10.8 and 12 μm. Both AIRS-A (0.0209 and 0.0377 for 10.8 and 12 μm) and AIRS-D (0.0275 and 0.0317 for 10.8 and 12 μm) show much worse precisions than MODIS (smaller than 0.019) and IASI (0.0203 and 0.0177 for 10.8 and 12 μm) products. The results indicate that (1) the classification-based modeling seems to provide quality initial values for physical retrieval of these two channels, and (2) the AIRS physical algorithm seems to have difficulty (such as cloud contamination) improving the first guess in 10.8 and 12 μm. The disadvantage of the classification-based modeling is its inability to capture the temporal variation.

Table 9. Confidence Levels of the Estimated Tb Deviations at Different Confidence Intervals[a]

Confidence Interval	0.1 (%)	0.2 (%)	0.3 (%)
8.7 μm of IASI	95.9	100	100
10.8 μm of MODIS-T	46.4	76.3	94.8
12 μm of AIRS-A	90.0	99.9	100

[a]Demonstrated by a specific LSE combination of 8.7 μm from IASI, 10.8 μm from MODIS-T, and 12 μm from AIRS-A. For the same confidence interval, the reliability of the three estimated LSE Tb deviations are 8.7 μm of IASI > 12 μm of AIRS-A > 10.8 μm of MODIS-T.

[53] Unlike the validations using laboratory measurements over very homogenous areas, the method presented can be performed for all clear-sky LSE retrievals. This method can be applied to any database using any satellite observations as long as the appropriate channels are available. It is important to point out that this method will only perform better with more independent LSE databases coming in the future, because more independent databases may reduce the chances of error correlations. If there are not enough independent LSE databases, lack of good knowledge in error correlations of LSE databases makes it difficult to obtain reliable precision estimates with the method presented in this paper. This method can also be applied to other window channels in the region of 8–12 μm. Also, it is important to emphasize that this method can only be applied to window channels. Channels from other spectral regions have quite different sensitivity to the LST, which makes it less effective in reducing the LST impact.

[54] The method presented is applied to estimate standard deviation of LSE Tb deviations, therefore LSE precisions. All the conclusions are made based on the analysis of LSE precisions. It is still possible that a LSE database has high precision with low accuracy (large bias error), which makes the RMSE large. An attempt is made to apply the same method to estimate LSE RMSE; the estimated RMSE Tb deviations are scattered with values in a wide range, indicating there are more LSE error correlations. The extra LSE error correlations come from the LSE bias errors, which, when significant, will increase the LSE error correlations substantially.

[55] **Acknowledgments.** This work is partly supported by the National Oceanic and Atmospheric Administration GOES-R Algorithm Working Group program NA06NES4400002. The views, opinions, and findings contained in this report are those of the author(s) and should not be construed as an official National Oceanic and Atmospheric Administration or U.S. government position, policy, or decision. The authors would like to specifically thank the MODIS and AIRS science teams for making high-quality data available to the research community. The authors would also like to thank Daniel K. Zhou of NASA Langley Research Center for providing IASI preliminary emissivity research test product for this study.

References

Borbas, E. E., R. O. Knuteson, S. W. Seemann, E. Weisz, L. Moy, and H.-L. Huang (2007), A high spectral resolution global land surface infrared emissivity database, paper presented at Joint 2007 EUMETSAT Meteorological Satellite Conference and the 15th Satellite Meteorology and Oceanography Conference, Am. Meteorol. Soc., Amsterdam, 24–28 Sept. 2007.

Clough, S. A., and M. J. Iacono (1995), Line-by-line calculations of atmospheric fluxes and cooling rates: 2. Applications to carbon dioxide, ozone, methane, nitrous oxide and the halocarbons, *J. Geophys. Res.*, *100*, 16,519–16,535, doi:10.1029/95JD01386.

Hannon, S., L. L. Strow, and W. W. McMillan (1996), Atmospheric infrared fast transmittance models: A comparison of two approaches, *Proc. SPIE Int. Soc. Opt. Eng.*, *2830*, 94–105.

Hulley, G. C., S. J. Hook, and A. M. Baldridge (2009a), Validation of the North American ASTER Land Surface Emissivity Database (NAALSED) version 2.0 using pseudo-invariant sand dune sites, *Remote Sens. Environ.*, *113*(10), 2224–2233, doi:10.1016/j.rse.2009.06.005.

Hulley, G. C., S. J. Hook, E. Manning, S.-Y. Lee, and E. Fetzer (2009b), Validation of the Atmospheric Infrared Sounder (AIRS) version 5 land surface emissivity product over the Namib and Kalahari deserts, *J. Geophys. Res.*, *114*, D19104, doi:10.1029/2009JD012351.

Jin, M., and S. Liang (2006), An improved land surface emissivity parameter for land surface models using global remote sensing observations, *J. Clim.*, *19*(12), 2867–2881, doi:10.1175/JCLI3720.1.

Korb, A. R., J. W. Salisbury, and D. M. D'Aria (1999), Thermal-infrared remote sensing and Kirchhoff's law: 2. Field measurements, *J. Geophys. Res.*, *104*, 15,339–15,350, doi:10.1029/97JB03537.

Kornfield, J., and J. Susskind (1977), On the effect of surface emissivity on temperature retrievals, *Mon. Weather Rev.*, *105*(12), 1605–1608, doi:10.1175/1520-0493(1977)105<1605:OTEOSE>2.0.CO;2.

Labed, J., and M. P. Stoll (1991), Angular variation of land surface spectral emissivity in the thermal infrared: Laboratory investigations on bare soils, *Int. J. Remote Sens.*, *12*, 2299–2310, doi:10.1080/01431169108955259.

Li, J., J. Li, E. Weisz, and D. K. Zhou (2007), Physical retrieval of surface emissivity spectrum from hyperspectral infrared radiances, *Geophys. Res. Lett.*, *34*, L16812, doi:10.1029/2007GL030543.

Li, Z., J. Li, W. P. Menzel, T. J. Schmit, J. P. Nelson III, J. Daniels, and S. A. Ackerman (2008), GOES sounding improvement and applications to severe storm nowcasting, *Geophys. Res. Lett.*, *35*, L03806, doi:10.1029/2007GL032797.

Li, Z., J. Li, W. P. Menzel, J. P. Nelson III, T. J. Schmit, E. Weisz, and S. A. Ackerman (2009), Forecasting and nowcasting improvement in cloudy regions with high temporal GOES sounder infrared radiance measurements, *J. Geophys. Res.*, *114*, D09216, doi:10.1029/2008JD010596.

Miloshevich, L. M., H. Voemel, D. N. Whiteman, B. M. Lesht, F. J. Schmidlin, and F. Russo (2006), Absolute accuracy of water vapor measurements from six operational radiosonde types launched during AWEX-G and implications for AIRS validation, *J. Geophys. Res.*, *111*, D09S10, doi:10.1029/2005JD006083.

Morris, V., C. Long, and D. Nelson (2006), Deployment of an infrared thermometer network at the Atmospheric Radiation Measurement Program Southern Great Plains Climate Research Facility, paper presented at Sixteenth Atmospheric Radiation (ARM) Science Team Meeting, U.S. Dept. of Energy, Richland, WA.

Nerry, F., J. Labed, and M. P. Stoll (1990), Spectral properties of land surfaces in the thermal infrared 1. Laborataory measurements of absolute spectral emissivity signatures, *J. Geophys. Res.*, *95*, 7027–7044, doi:10.1029/JB095iB05p07027.

Rothman, L. S., R. R. Gamache, R. H. Tipping, C. P. Rinsland, M. A. H. Smith, D. C. Benner, V. M. Devi, J.-M. Flaud, C. Camy-Peyret, and A. Perrin (1992), The HITRAN molecular database: Editions of 1991 and 1992, *J. Quant. Spectrosc. Radiat. Transf.*, *48*, 469–507, doi:10.1016/0022-4073(92)90115-K.

Ruston, B., F. Weng, and B. Yan (2008), Use of a one-dimensional variational retrieval to diagnose estimates of infrared and microwave surface emissivity over land for ATOVS sounding instruments, *IEEE Trans. Geosci. Remote Sens.*, *46*, 393–402, doi:10.1109/TGRS.2007.910219.

Sabol, D., A. Gillespie, E. Abbott, and G. Yamada (2009), Field validation of the ASTER Temperature-Emissivity Separation algorithm, *Remote Sens. Environ.*, *113*(11), 2328–2344, doi:10.1016/j.rse.2009.06.008.

Salisbury, J. W., A. Wald, and D. M. D'Aria (1994), Thermal-infrared remote sensing and Kirchhoff's law: 1. Laboratory measurements, *J. Geophys. Res.*, *99*, 11,897–11,911, doi:10.1029/93JB03600.

Schmetz, J., P. Pili, S. Tjemkes, D. Just, J. Kerkmann, S. Rota, and A. Ratier (2002), An introduction to Meteosat Second Generation (MSG), *Bull. Am. Meteorol. Soc.*, *83*(7), 977–992, doi:10.1175/1520-0477(2002)083<0977:AITMSG>2.3.CO;2.

Seemann, S. W., E. E. Borbas, R. O. Knuteson, G. R. Stephenson, and H.-L. Huang (2008), Development of a global infrared land surface emissivity database for application to clear sky sounding retrievals from multispectral satellite radiance measurements, *J. Appl. Meteorol. Climatol.*, *47*(1), 108–123, doi:10.1175/2007JAMC1590.1.

Snyder, W. C., Z. Wan, Y. Zhang, and Y. Z. Feng (1998), Classification-based emissivity for land surface temperature measurement from space, *Int. J. Remote Sens.*, *19*, 2753–2774, doi:10.1080/014311698214497.

Sobrino, J. A., J. C. Jimenez-Munoz, L. Balick, A. R. Gillespie, D. A. Sabol, and W. T. Gustafson (2007), Accuracy of ASTER Level-2 thermal-infrared Standard Products of an agricultural area in Spain, *Remote Sens. Environ.*, *106*(2), 146–153, doi:10.1016/j.rse.2006.08.010.

Susskind, J., and J. Blaisdell (2008), Improved surface parameter retrievals using AIRS/AMSU data, *Proc. SPIE*, *6966*, 10–12, doi:10.1117/1112.774759.

Turner, D. D., B. M. Lesht, S. A. Clough, J. C. Liljegren, H. E. Revercomb, and D. C. Tobin (2003), Dry bias and variability in Vaisala RS80-H radiosondes: The ARM experience, *J. Atmos. Oceanic Technol.*, *20*(1), 117–132, doi:10.1175/1520-0426(2003)020<0117:DBAVIV>2.0.CO;2.

Wan, Z. (2008), New refinements and validation of the MODIS land-surface temperature/emissivity products, *Remote Sens. Environ.*, *112*, 59–74, doi:10.1016/j.rse.2006.06.026.

Wan, Z., and Z. Li (1997), A physics-based algorithm for retrieving land-surface emissivity and temperature from EOS/MODIS data, *IEEE Trans. Geosci. Remote Sens.*, *35*(4), 980–996, doi:10.1109/36.602541.

Wang, K., Z. Wan, P. Wang, M. Sparrow, J. Liu, and S. Haginoya (2007), Evaluation and improvement of the MODIS land surface temperature/emissivity products using ground-based measurements at a semi-desert site on the western Tibetan Plateau, *Int. J. Remote Sens.*, *28*(11), 2549–2565, doi:10.1080/01431160600702665.

Zhou, D. K., A. Larar, X. Liu, S. L. William, L. L. Strow, P. Yang, P. Schlussel, and X. Calbet (2010), Global land surface emissivity retrieved from satellite ultraspectral IR measurements, *IEEE Trans. Geosci. Remote Sens.*, doi:10.1109/TGRS.2010.2051036, in press.

E. E. Borbas, X. Jin, J. Li, and Z. Li, Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin-Madison, 1225 W. Dayton St., Madison, WI 53706, USA. (zhenglong.li@ssec.wisc.edu)

M. D. Goldberg and T. J. Schmit, Center for Satellite Applications and Research, National Environmental Satellite, Data, and Information Service, National Oceanic and Atmospheric Administration, 5200 Auth Rd., NOAA Science Center, Rm. 810, Camp Springs, MD 20746, USA.