Remote Sensing of Environment 113 (2009) 1556–1565

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Evaluation of ASTER and MODIS land surface temperature and emissivity products using long-term surface longwave radiation observations at SURFRAD sites

Kaicun Wang *, Shunlin Liang

Department of Geography, University of Maryland, College Park, MD 20742, United States

ARTICLE INFO

Article history:
Received 12 August 2008
Received in revised form 9 March 2009
Accepted 14 March 2009

Keywords:
Land surface temperature
Emissivity
MODIS
ASTER
Surface radiation budget

ABSTRACT

Land surface temperature (LST) and emissivity are key parameters in estimating the land surface radiation budget, a major controlling factor of global climate and environmental change. In this study, Terra Advanced Spaceborne Thermal Emission Reflection Radiometer (ASTER) and Aqua MODerate resolution Imaging Spectroradiometer (MODIS) Collection 5 LST and emissivity products are evaluated using long-term ground-based longwave radiation observations collected at six Surface Radiation Budget Network (SURFRAD) sites from 2000 to 2007. LSTs at a spatial resolution of 90 m from 197 ASTER images during 2000–2007 are directly compared to ground observations at the six SURFRAD sites. For nighttime data, ASTER LST has an average bias of 0.1 °C and the average bias is 0.3 °C during daytime. Aqua MODIS LST at 1 km resolution during nighttime retrieved from a split-window algorithm is evaluated from 2002 to 2007. MODIS LST has an average bias of −0.2 °C. LST heterogeneity (defined as the Standard Deviation, STD, of ASTER LSTs in 1×1 km^2 region, 11×11 pixel in total) and instrument calibration error of pyrgeometer are key factors impacting the ASTER and MODIS LST evaluation using ground-based radiation measurements. The heterogeneity of nighttime ASTER LST is 1.2 °C, which accounts for 71% of the STD of the comparison, while the heterogeneity of the daytime LST is 2.4 °C, which accounts for 60% of the STD. Collection 5 broadband emissivity is 0.01 larger than that of MODIS Collection 4 products and ASTER emissivity. It is essential to filter out the abnormal low values of ASTER daily emissivity data in summer time before its application.

1. Introduction

Land surface temperature (LST) and emissivity are key parameters in calculating the land surface radiation budget, a major controlling factor of global climate and environmental change. LST and emissivity can be used directly to estimate surface upwelling longwave radiation based on the Planck function and the Stefan–Boltzmann law. LST is closely related to soil moisture (Wan et al., 2004) and canopy evapotranspiration (Wang et al., 2006, 2007b; Wang & Liang, 2008). Furthermore, LST has been assimilated into land surface models to improve land–atmosphere exchange simulation (Qin et al., 2007; Rodell et al., 2004).

As it is practically impossible to obtain such information from ground-based measurements at the regional and global scales, the use of satellites in the thermal infrared (TIR) region appears very attractive (Kerr et al., 2000). Evaluation of satellite LST and emissivity retrievals is important because their accuracy is critical to their use in a range of applications. Feedback from validation activities also helps to improve the generation of these products (Wang et al., 2007a; Wan, 2008). Sobrino et al. (2007) evaluate ASTER LST in Spain using measurements collected at a day. Coll et al. (2007) evaluated ASTER and emissivity product using ground measurements of surface temperature, emissivity, and atmospheric radiosonde profiles for low spectral contrast surface in Spain. It is very difficult to validate LST and emissivity (Wan et al., 2002; Wan, 2008; Coll et al., 2005). Because of the large spatial variation in LSTs, especially during the daytime, it is essential to carefully select validation sites (Wan et al., 2002; Wan, 2008; Coll et al., 2005). Surface-measured emissivity has been used to evaluate MODIS Collection 4 LST/emissivity products (Wang et al., 2007a; Wan, 2008).

Surface longwave radiation is related to LST and emissivity (Liang, 2004; Wang et al., 2005). Recently, high quality long-term surface longwave radiation measurements are available globally, such as Surface Radiation Budget Monitoring (SURFRAD, http://www.srrb.noaa.gov/surfrad/ and Augustine et al., 2000), FLUXNET (http://daac.ornl.gov/FLUXNET/; Baldocchi et al., 2001), Atmospheric Radiation Measurement (ARM, http://www.arm.gov/; Gautier & Landsfeld, 1997; Stokes & Schwartz, 1994) and Baseline Surface Radiation Network (BSRN, http://bsrn.ethz.ch/; Ohmura et al., 1998). It is helpful to investigate the capability of these measurements in evaluating the satellite LST and emissivity products.

Wang et al. (2008a,b) evaluated MODIS Collection 4 LST/emissivity products using longwave radiation measurements collected at FLUXNET sites. They concluded that MODIS Collection 4 LST/emissivity has an obvious negative bias, up to −3 K. Some substantial improvements have been made to MODIS Collection 5 LST/emissivity

* Corresponding author. Tel.: +1 301 405 4538.
E-mail address: kcwang@umd.edu (K. Wang).

0034-4257/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2009.03.009

Table 1
A summary of the description of sites.

Site	Lat/lon	Land cover	Elevation (km)	NDVI	Relative humidity (%)
Bondville, IL	40.05 N, 88.37 W	Cropland	0.213	0.38	66
Boulder, CO	40.13 N, 105.24 W	Grassland	1.689	0.29	45
Fort Peck, MT	48.31 N, 105.10 W	Grassland	0.634	0.23	58
Goodwin Creek, MS	34.25 N, 89.87 W	Pasture	0.098	0.54	63
Penn State, PA	40.72 N, 77.93 W	Cropland	0.376	0.48	64
Sioux Falls, SD	43.73 N, 96.62 W	Cropland	0.473	0.41	66

Multi-year (2002–2007) average of Normalized Difference vegetation Index (NDVI) from MODIS and multi-year (2002–2007) average relative humidity (%) collected at SURFRAD sites are also shown.

products (Wang et al., 2007a; Wan, 2008). In this study, we evaluate MODIS Collection 5 LST and emissivity products and ASTER standard LST and emissivity products using long-term accurate longwave radiation measurements collected at SURFRAD sites.

2. MODIS and ASTER LST/emissivity products

As part of the NASA Earth Observing System (EOS) project, two MODIS instruments were placed onboard the Terra and Aqua satellite platforms, to provide information for global studies of the atmosphere, land, and ocean processes (Salomonson et al., 1989). Aqua overpasses at around local times of 1:30 p.m. (ascending mode) and 1:30 a.m. (descending mode), while Terra overpasses at local times of 10:30 a.m. (descending mode) and 10:30 p.m. (ascending mode).

Advantages of the MODIS instruments include their global coverage, high radiometric resolution and dynamic ranges, and accurate calibration in the thermal infrared (TIR) bands (Xiong et al., 2008). Making use of the middle infrared and TIR measurements (1-km spatial resolution at nadir), MODIS land surface products can supply global coverage of temperature and emissivity at daily or quasi-daily temporal and 1-km spatial resolution retrieved by the generalized split-window algorithm (Wan & Dozier, 1996), or at 5-km spatial resolution retrieved by the MODIS day/night LST algorithm (Wan & Li, 1997). MODIS Collection 5 LST at 1-km spatial resolution from a split-window algorithm is evaluated in this paper.

Wang et al. (2007a) and Wan (2008) discussed the improvement of MODIS collection 5 LST/emissivity products. Two refinements made to the MODIS day/night LST algorithm have had obvious effects on the LST/emissivity values.

First, our previous study (Wang et al., 2007a) has shown that surface emissivity during the rainy season is less than that of the dry season. This contradicts to our expectation that surface emissivity increases with soil moisture if the surface emissivity is lower than the water surface emissivity. This temporal pattern indicates that the effect of cloud contamination plays an important role in the day/night algorithm. Therefore, the Collection 5 day/night LST algorithm tries to avoid cloud-contaminated observations in the process to select pairs of day and night observations by checking the values of brightness temperature in band 31, and always using the fixed initial emissivity values in the day/night algorithm (Wan, 2008). The regression method previously used in the development of the day/night algorithm to assign the initial values has been dropped.

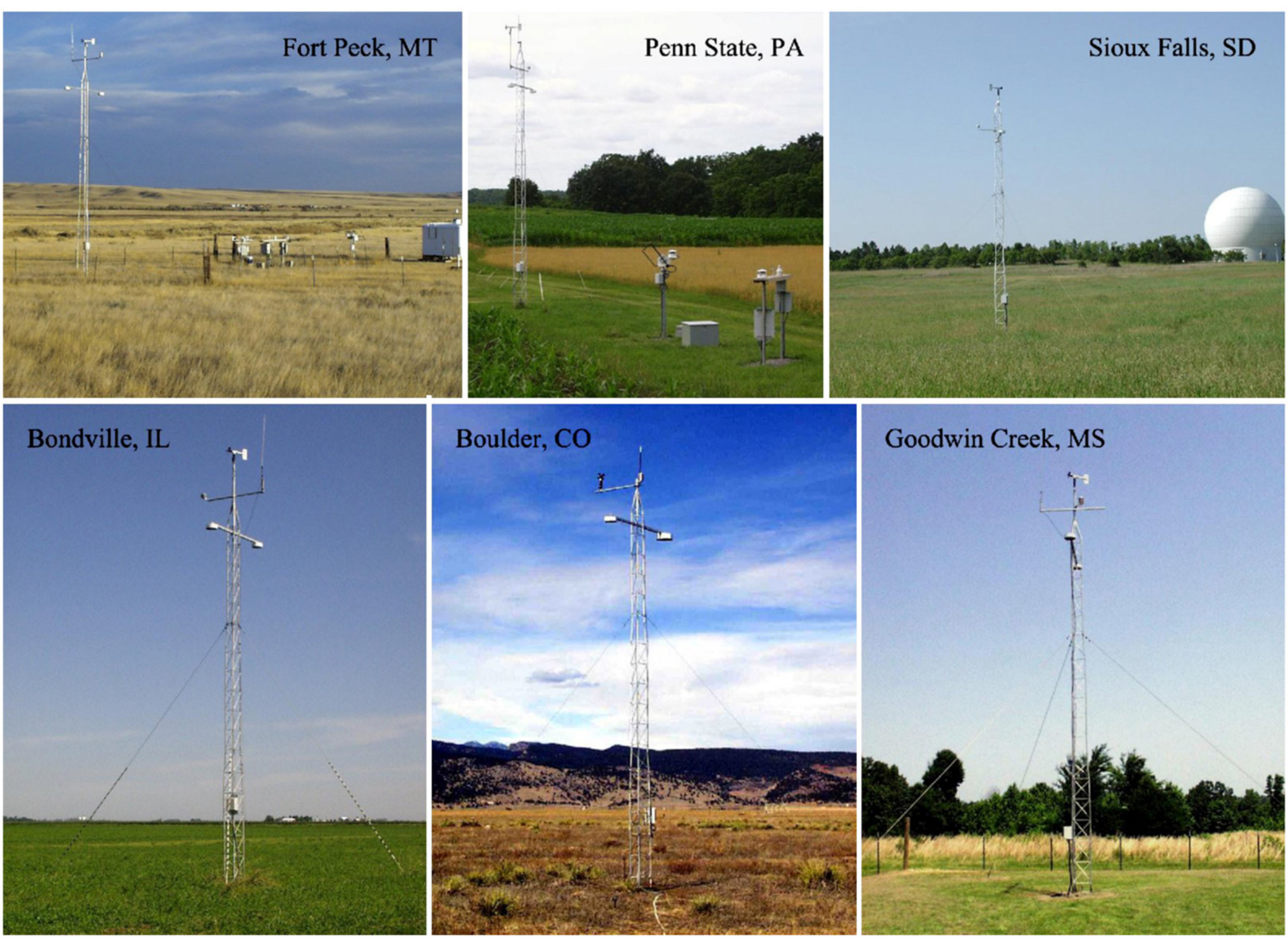

Fig. 1. Photos of the six SURFRAD sites. Pyrgeometers used to collected surface upwelling longwave radiation are deployed at 10-meter-high towers.

Second, cloud-contaminated LSTs in monthly products are removed by using constraints on the temporal variations (δT) in clear-sky LSTs over a period of 32 days. The value of δT depends on land cover types (Wan, 2008). There are three major steps in the removal scheme. Step 1 removes the worst LSTs that differ from the 32-day maximum by more than 4 times the δT value, or differ from the 16-day maximum by more than 3 times the δT value. Step 2 removes the LSTs that deviate from the 8-day maximum by more than 2 times the δT value, then calculate the 8-day average value of the remaining LSTs. Step 3 removes the LSTs that digress from the 8-day average value by more than the δT value (See Wan, 2008 for detail).

ASTER is a high spatial resolution radiometer on board the EOS Terra satellite, which consists of three separate subsystems: the visible and near infrared (VNIR), the shortwave infrared (SWIR) and the thermal infrared (TIR) (Yamaguchi et al., 1998). The TIR subsystem has five spectral channels between 8 and 12 µm with a spatial resolution of 90 m. The multispectral TIR allows the retrieval of land surface temperature (LST) and emissivity spectra at high spatial resolution. LST and spectral emissivities are retrieved from ASTER TIR data by means of the Temperature Emissivity Separation (TES) method (Gillespie et al., 1998). It is applied to at-ground TIR radiances, which have been corrected for atmospheric effects with the ASTER standard atmospheric correction algorithm (Palluconi et al., 1999), and requires knowledge of the downward sky irradiance. The ASTER TIR standard correction algorithm is based on radiative transfer calculations using the MODTRAN code (Berk et al., 1999), with input atmospheric profiles extracted from either the Global Data Assimilation System (GDAS) product or the Naval Research Laboratory (NRL) climatology model.

3. Surface measurements

Accurate and precise ground-based measurements in differing climatic regions are essential to refine and verify satellite-based estimates, as well as to support specialized research (http://www.srrb.noaa.gov/surfrad/surfpage0.html). To fill this niche, SURFRAD was established in 1993 through the support of NOAA's Office of Global Programs. Its primary objective is to support climate research with accurate, continuous, long-term measurements pertaining to the surface radiation budget over the United States.

Six SURFRAD stations are selected in this study (see Table 1). All of the six sites are located in large flat agricultural areas covered by crops and grass where surface emissivity is consistently high. Independent measures of upwelling and downward, solar and infrared radiation are the primary measurements. Data are downloaded, quality controlled, and processed into daily files that are distributed in near real time by anonymous FTP and the WWW (http://www.srrb.noaa.gov). The ground-based measurements of longwave radiation collected at SURFRAD sites have been used in our previous studies (Wang & Liang, 2009; Wang et al., 2008a,b).

A pyrgeometer is deployed at a 10-m high tower to measure upwelling longwave radiation (see Fig. 1). The effective diameter of the field-of-view of the pyrgeometer mounted on a 10-m tower is about 30–45 m. They are sensitive to the spectral range from 3.0 to 50 µm (ventilated Eppley pyrgeometer). All instruments at each station are exchanged for newly calibrated instruments on an annual basis (Augustine et al., 2000, 2005). Calibrations are performed by world-recognized organizations. SURFRAD pyrgeometers are calibrated using three standards maintained at the Surface Radiation Research Branch (SRRB) Field Test and Calibration Facility at Table Mountain near Boulder, CO. The Peak Irradiance Response (PIR) standards' calibrations are traceable to a blackbody calibration device in Davos, Switzerland, where they are calibrated annually. Finally, in order to maintain continuity between the retired instruments and their replacements, all instruments are gauged against three standard instruments before and after field deployment. All the calibration information can be found at http://www.srrb.noaa.gov/surfrad/getcals.html. SURFRAD uses quality assurance and quality control to provide the best possible data which are subjected to automated procedures as the daily files are processed. At present, datasets undergo this first-level check and a daily "eye" check before being

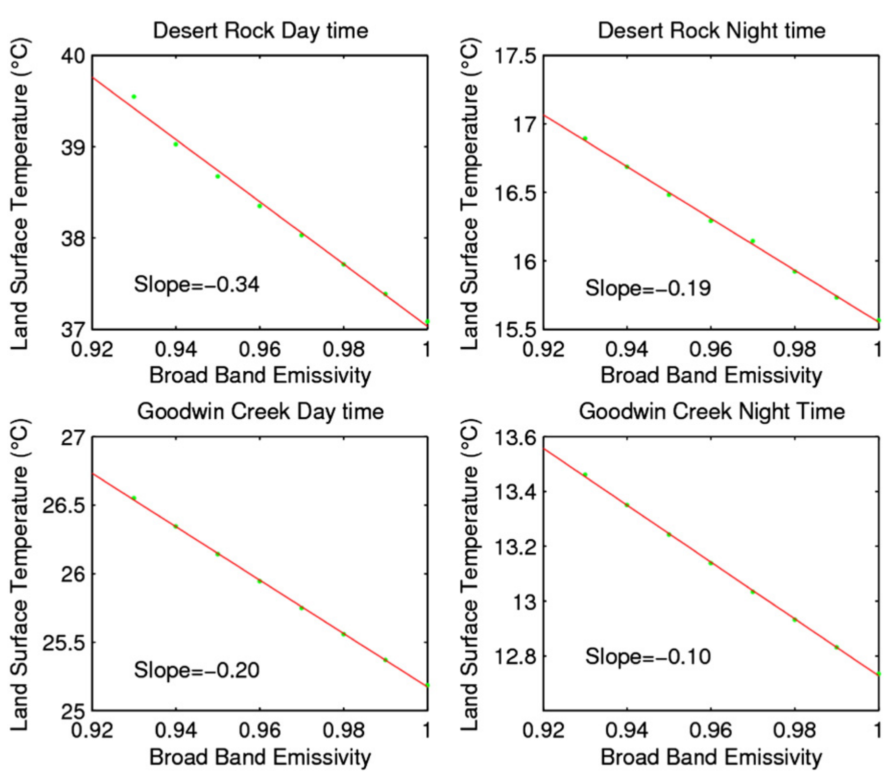

Fig. 2. The relationship between multi-year-average (2002–2007) Land Surface Temperature (LST) derived from assumed broadband emissivity using Eq. (2) at one dry and one humid site. The slope (unit: °C/0.01 broadband emissivity) demonstrates the sensitivity of LST to broadband emissivity.

released. Our recent study on downward longwave radiation is measured by the same type of pyrgeometers, documented an error of 3–5 W m^{-2} at the SURFRAD sites (Wang & Liang, in press), equivalent to an error of 0.5 °C–0.8 °C in LST.

LST (T_S) is related to surface longwave radiation by the Stefan–Boltzmann law:

$$L_{\uparrow} = \varepsilon_b \cdot \sigma \cdot T_s^4 + (1-\varepsilon_b) \cdot L_{\downarrow} \quad (1)$$

where $L_{\uparrow}$ is surface upwelling longwave radiation, ε_b is broadband emissivity over the entire infrared region, σ is the Stefan–Boltzmann's constant (5.67×10^{-8} W m^{-2} K^{-4}), and $L_{\downarrow}$ is atmospheric downwelling longwave radiation at the surface. Therefore, LST can be estimated from:

$$T_s = \left[\frac{L_{\uparrow} - (1-\varepsilon_b) \cdot L_{\downarrow}}{\varepsilon_b \cdot \sigma}\right]^{1/4} \quad (2)$$

Broadband emissivity ε_b is estimated from MODIS narrowband emissivity Collection 5 retrievals in the thermal infrared region (Wang et al., 2005):

$$\varepsilon_b = 0.2122 \cdot \varepsilon_{29} + 0.3859 \cdot \varepsilon_{31} + 0.4029 \cdot \varepsilon_{32} \quad (3)$$

where ε_{29}, ε_{31} and ε_{32} are MODIS band 29, 31 and 32 narrowband emissivities. The narrow band emissivities are derived from the MODIS day/night LST algorithm.

The accuracy of LST calculated from ground-based $L_{\uparrow}$ and $L_{\downarrow}$ measurements using Eq. (2) depends on the satellite broadband emissivity retrievals. One can infer that the sensitivity of LST derived to broadband emissivity depends on the contrast of $L_{\uparrow}$ and $L_{\downarrow}$. The greater the contrast, the larger the sensitivity. It is expected that the contrast is larger during daytime than that during nighttime, and larger in the dry areas than that in the humid areas. We calculated the sensitivity of LST to broadband emissivity at such sites: Desert Rock, NV (36.63N, 116.02W) and Goodwin Creek, MS at Terra overpasses. The broadband emissivity is assumed to vary from 0.93 to 1.0 with an interval of 0.01, and then LST is calculated from 3-minute-averages of $L\uparrow$ and $L_{\downarrow}$. We averaged the data during 2002–2007 and the results are shown in Fig. 2. The greatest sensitivity is 0.35 °C/0.01 broadband emissivity for the Desert Rock site during daytime and the least sensitivity is 0.10 °C/0.01 broadband emissivity for humid Goodwin Creek, MS during nighttime. It is shown that for vegetated surfaces, LST can be estimated from $L_{\uparrow}$ and $L_{\downarrow}$ at an accuracy of 0.2–0.5 °C given that the error of broadband emissivity is about ± 0.02.

$L_{\uparrow}$ is measured at a 10-m high tower at SURFRAD sites. A temperature inversion layer may occur at nighttime, meaning T_S is less than the air temperature at 10 m high. Water vapor in the atmosphere layer from the surface to a height of 10 m absorbs surface longwave emission and emits longwave radiance at the same time. Under this condition, the measured $L_{\uparrow}$ will be greater than the actual surface $L_{\uparrow}$. Our previous study showed that the difference may be up to 2–3 Wm^{-2} for the crop and grass surface (Wang, 2004), resulting

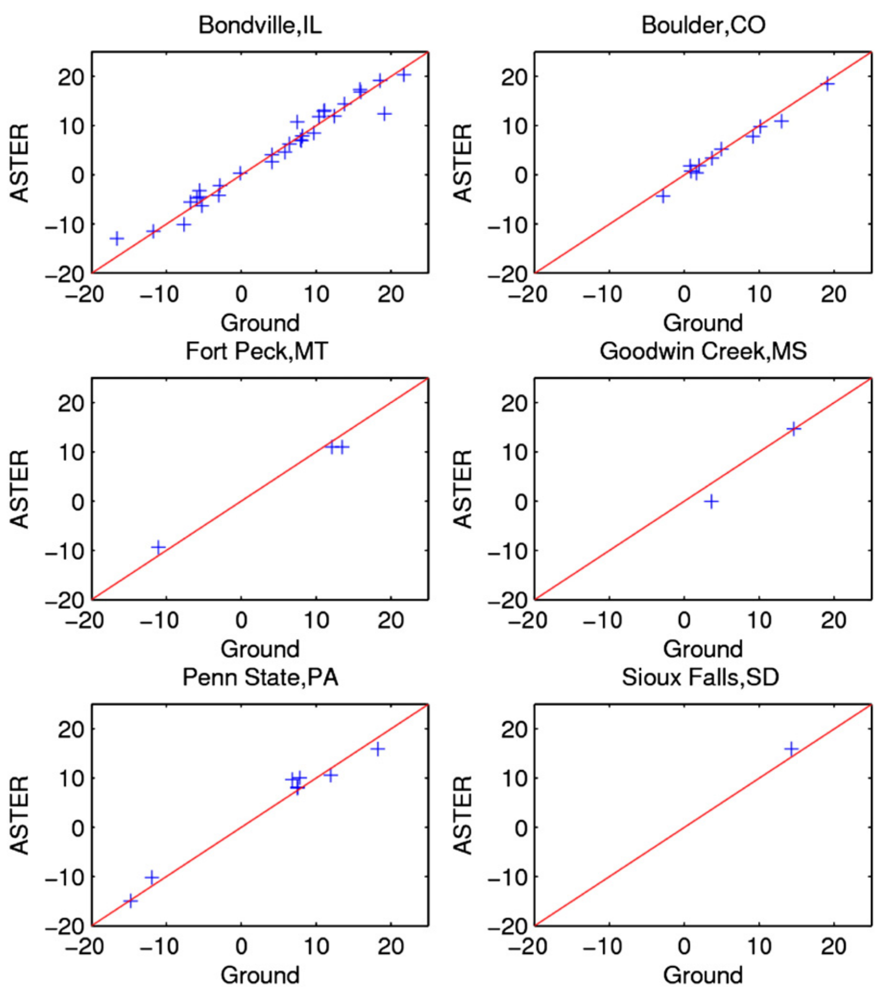

Fig. 3. Scatterplots of the comparison of LST from ASTER nighttime retrievals at 90-m resolution and ground-based measurements at six SURFRAD sites. Data used here is from 2000 to 2007.

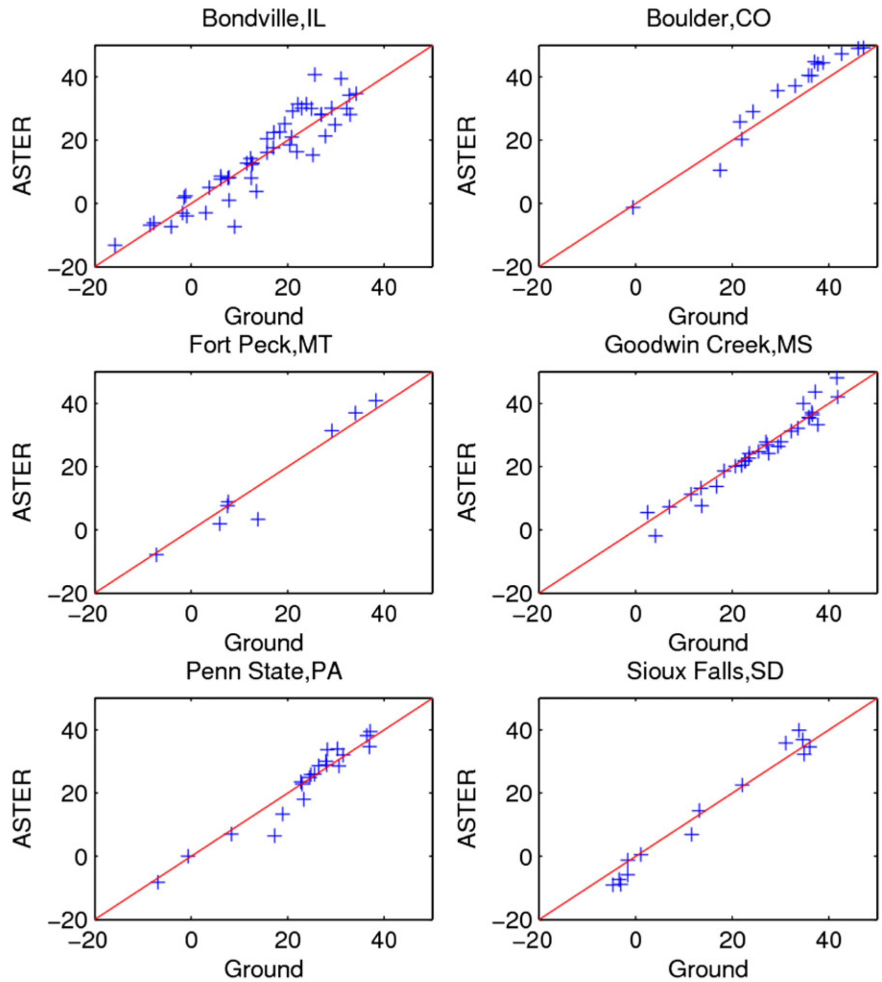

Fig. 4. Scatterplots of the comparison of LST from ASTER daytime retrievals at 90-m resolution and ground-based measurements at six SURFRAD sites. Data used here is from 2000 to 2007.

in the errors of 0–0.3 °C in T_S calculated from Eq. (2). The conditions during daytime are opposite, that is T_S may have the error of −0.3–0 °C during daytime for the SURFRAD crop and grass sites.

LST from satellite and ground measurements may differ according to their measurement methods. MODIS and ASTER use directional measurements in the atmospheric window, while ground-based longwave radiation measurements are hemispheric, wider spectrum derivations. If the surface is black body or gray body with a Lambertian assumption, the two LSTs are the same (Wang et al., 2005). However, we still cannot accurately quantify the directional emissivity of natural surfaces.

4. Results

4.1. ASTER LST

ASTER has the high spatial resolution of 90 m for thermal infrared bands, the similar size to the scale of ground-based measurements ($L_\uparrow$ is observed at a 10 m high tower). Therefore, we can compare ground-based LST with ASTER retrievals directly at the SURFRAD sites. The results are shown in Figs. 3 and 4. The ASTER standard LST products supply longitude and latitude information of 11×11 control points for a total of 700×830 LST pixels. We interpolate the longitude and latitude information of the other points using a bi-linear interpolation method. This may introduce an error in the geolocation information of the pixels extracted from ASTER LST products. Therefore, when compared with ground-based LST, the ASTER LSTs are averaged in a 3×3 window and the comparison results are affected by the heterogeneity of LST.

Table 2 summarizes the statistics of the comparisons. For the six sites at nighttime, the bias varies from −0.7 °C to 1.6 °C, with an average of 0.1 °C; STD varies from 0.9 °C to 3.0 °C, with an average of

Table 2
A summary of statistical parameters of the comparison of ASTER LST at 90-m spatial resolution and ground-based measurements collected at SURFRAD sites.

Site	Time (sample number)	Satellite average scale (270 m)				Satellite average scale (1000 m)			
		Bias	STD	R	H	Bias	STD	R	H
Bondville IL	Night (30)	0.1	1.9	0.98	0.9	0.1	1.9	0.98	0.9
	Day (48)	0.4	5.6	0.92	1.9	0.3	5.6	0.92	1.9
Boulder, CO	Night (11)	−0.6	0.9	0.99	1.0	−0.6	0.9	0.99	1.0
	Day (19)	4.0	3.8	0.98	2.7	3.8	3.7	0.98	2.7
Fort Peck, MT	Night (3)	−0.7	2.2	1.0	1.4	−0.8	1.7	1.0	1.3
	Day (8)	−0.8	4.5	0.97	1.7	−0.7	4.5	0.98	1.6
Goodwin Creek, MS	Night (3)	−0.5	3.0	0.99	0.9	−0.4	3.0	1.0	0.9
	Day (31)	−0.5	2.9	0.97	2.5	−0.4	3.0	0.97	2.5
Penn State, PA	Night (8)	0.5	1.8	0.99	1.5	0.5	1.8	0.99	1.5
	Day (21)	−0.4	3.6	0.96	2.5	−0.4	3.5	0.96	2.4
Sioux Falls, SD	Night (1)	1.6			1.4	1.4			1.0
	Day (14)	−0.9	3.7	0.99	1.9	−0.9	3.8	0.99	1.8
All	Night (56)	0.1	1.6	0.99	1.2	0.1	1.5	0.99	1.2
	Day (141)	0.3	4.0	0.96	2.4	0.3	4.0	0.96	2.4

The data used here is from 2000 to 2007. The multi-year average values of the heterogeneity (H) of the LST at 90-m resolution (defined as the standard deviation of the LST in a certain area) are shown. The unit of Bias, STD and H is °C.

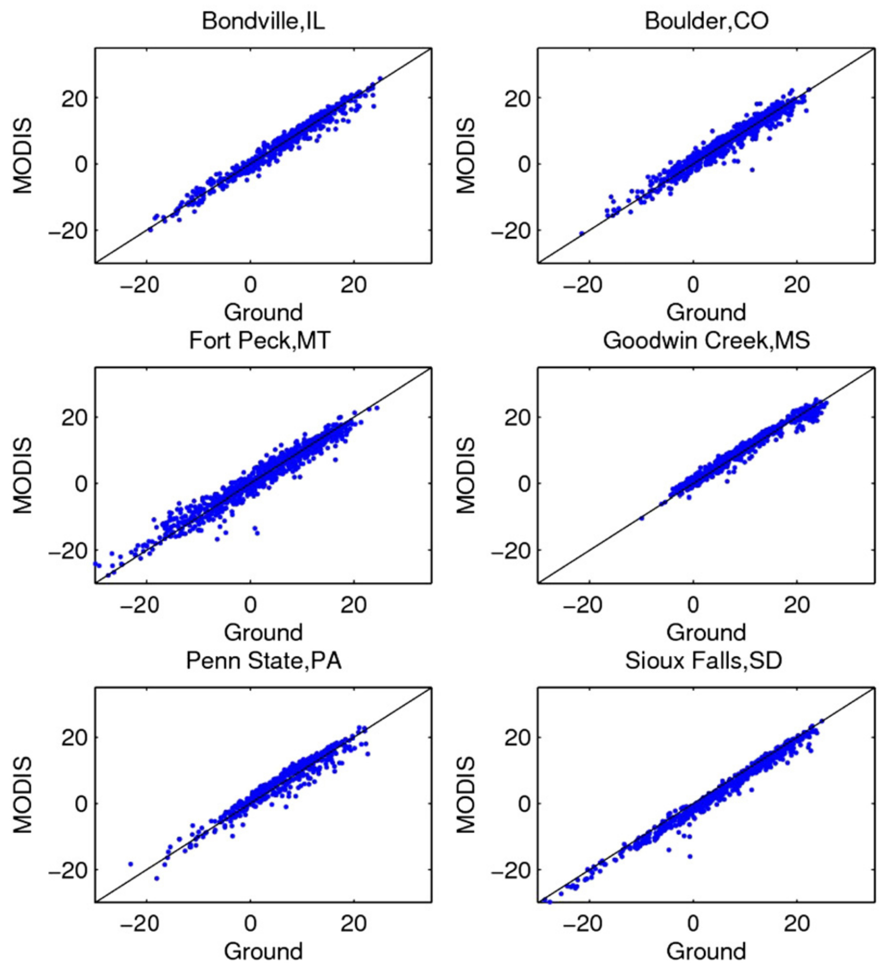

Fig. 5. Scatterplots of the comparison of LST from Aqua MODIS Collection 5 nighttime retrievals at 1-km resolution and ground-based measurements at six SURFRAD sites. Data used here is from 2002 to 2007.

1.6 °C; and the correlation coefficient varies from 0.98 to 1.0, with an average of 0.99. For the six sites at daytime, bias varies from −0.9 °C to 4.0 °C, with an average of 0.3 °C; STD varies from 2.9 °C to 5.6 °C, with an average of 4.0 °C; the correlation coefficient varies from 0.92 to 0.99, with an average of 0.96.

The comparison results for daytime are not as good as those during nighttime, especially in terms of STD. To investigate this, we calculate the degree of heterogeneity of LST, defined as the standard deviation of the ASTER LST 90-m pixels in a certain region. Table 2 shows the heterogeneity calculated over a 270×270 m^2 and 1000×1000 m^2 area. For the six sites, the heterogeneity varies from 0.9 to 1.5 during nighttime, with an average of 1.2 °C, and from 1.7 to 2.7 during daytime, with an average of 2.4 °C. The heterogeneity of LST can account for 71% of STD during nighttime and 60% of STD during daytime. Table 2 also shows that the heterogeneities are highly correlated with STDs at the six sites, which indicates that the STD of the comparison is controlled by surface heterogeneity.

With only one ground instrument per site, it is not possible to assess whether the single ground LST measurement over a small spot on the site is truly representative of the LST at the satellite pixel scale, and thus the ground LSTs may be biased for the comparison with satellite-derived LSTs. Such biases may be different depending on the time of day, season, general meteorological conditions, etc. Multiple sampling over areas comparable to the pixel size is necessary to evaluate the LST variability of the site at the ground scale and thus the quality of the ground data. To assess this effect, we compared ground measurements with averages of ASTER LST at 3×3 and 11×11 window and the results are shown in Table 2. The two comparisons have the similar biases.

4.2. MODIS LST

Given the scale difference between MODIS LST retrievals and ground-based measurements, and the large heterogeneity of the LST

Table 3
A summary of statistical parameters of the comparison of nighttime (about 1:30 a.m. local time) LST from Aqua/MODIS split-window LST at 1-km spatial resolution and ground-based measurements collected at SURFRAD sites.

Site	Bias (°C)	STD (°C)	Correlation coefficient	Heterogeneity (°C)	MODIS Collection 4	ASTER	MODIS Collection 5
Bondville, IL	−0.1	1.5	0.99	0.19	0.97	0.96	0.98
Boulder, CO	0.1	1.7	0.97	0.65	0.97	0.97	0.98
Fort Peck, MT	−0.1	2.5	0.98	0.42	0.97	0.98	0.98
Goodwin Creek, MS	0.3	1.4	0.99	0.16	0.97	0.96	0.98
Penn State, PA	0.3	2.0	0.97	0.36	0.97	0.97	0.98
Sioux Falls, SD	−1.6	1.5	0.99	0.26	0.97	0.97	0.98
All	−0.2	1.8	0.98	0.34	0.97	0.97	0.98

The data used here is from 2002 to 2007. The multi-year average values of the heterogeneity of LST at 1-km resolution (defined as the standard deviation of LST over the four pixels enclosing the site) are shown. Multi-year (2002–2006) broadband emissivity of MODIS monthly LST/emissivity products at a spatial resolution of 0.05° is also shown. The numbers of MODIS observations for each site are: 662 (Bondville), 985 (Boulder), 829 (Fort Peck), 633 (Goodwin Creek), 520 (Penn State) and 638 (Sioux Falls).

during daytime (see Section 4.1), comparison was only carried out during Aqua nighttime overpasses (about 1:30 a.m.) when spatial variation of LST is smaller compared to daytime. Available data during 2002–2007 are shown here (MODIS Collection 5 LST/emissivity products in much of 2006 were not available when this work was done).

Fig. 5 shows the comparison scatterplots of Aqua/MODIS split-window LST and ground-based measurements collected at the six SURFRAD sites. Table 3 summarizes the statistical parameters of the comparisons. The two LSTs agree well at most sites. The bias in five of the six sites is less than 0.3 °C and only one site has a −1.6 °C bias. The average of the bias over the six sites is −0.2 °C. STD varies from 1.5 °C to 2.6 °C, with an average of 1.8 °C over the six sites. The heterogeneity at 1:30 a.m. is expected to be less than that of 1.2 °C at early nighttime (22:30 p.m.).

Wang et al. (2008a,b) used the ground-based measurements from Ameriflux sites. Two sites in their report (Bondville, IL, 40.01 N, 88.29 W, cropland, and Fort Peck, MT, 48.31 N, 105.10 W, grassland) are close to the sites in this study (Bondville, IL, 40.05 N, 88.37 W, cropland, and Fort Peck, MT, 48.31 N, 105.10 W, grassland). Wang et al. (2008a,b) reported that MODIS Collection 4 LST at Bondville, IL has a negative bias of −3.1 °C, and an STD of 3.4 °C, and at Fort Peck, MT, a negative bias of −2.2 °C and a STD of 2.5 °C. The results reported in this study are much better than that in Wang et al. (2008a,b). The major reason for this is that the accuracy of longwave radiation measurements at SURFRAD sites is higher than that at corresponding Ameriflux sites. Thanks to the better sensor calibration at SURFRAD sites, the error of longwave radiation at SURFRAD Bondville and Fort Peck sites is 3–4 W m^{-2} (a standard deviation, equivalent to an error of 0.5–0.7 °C in LST) while it is 10 W m^{-2} (equivalent to an error of 1.6 °C in LST) at Ameriflux Bondville and Fort Peck sites (Wang & Liang, in press).

4.3. MODIS and ASTER emissivity

Table 3 shows the broadband emissivities from MODIS Collection 4 and Collection 5 LST and emissivity products. The emissivity is from MODIS monthly LST/emissivity products at a spatial resolution of 0.05°. The Collection 5 broadband emissivity at each site is greater than that of Collection 4 and ASTER. On average, the Collection 5 broadband emissivity is about 0.01 larger than that of MODIS Collection 4 and ASTER emissivity. Fig. 6 shows the seasonal variation of broadband emissivity over six SURFRAD sites. It is evident that the seasonal variation of the Collection 5 broadband emissivity for the SURFRAD sites is very small.

Fig. 6 also shows the broadband emissivity calculated from ASTER daily narrowband retrievals (Ogawa et al., 2003), demonstrating that the ASTER emissivity is similar to MODIS emissivity. The ASTER emissivity is lower in summer time. Surface emissivity should increase with the increased vegetation coverage and soil moisture. These abnormal low emissivity values should be filtered out for various

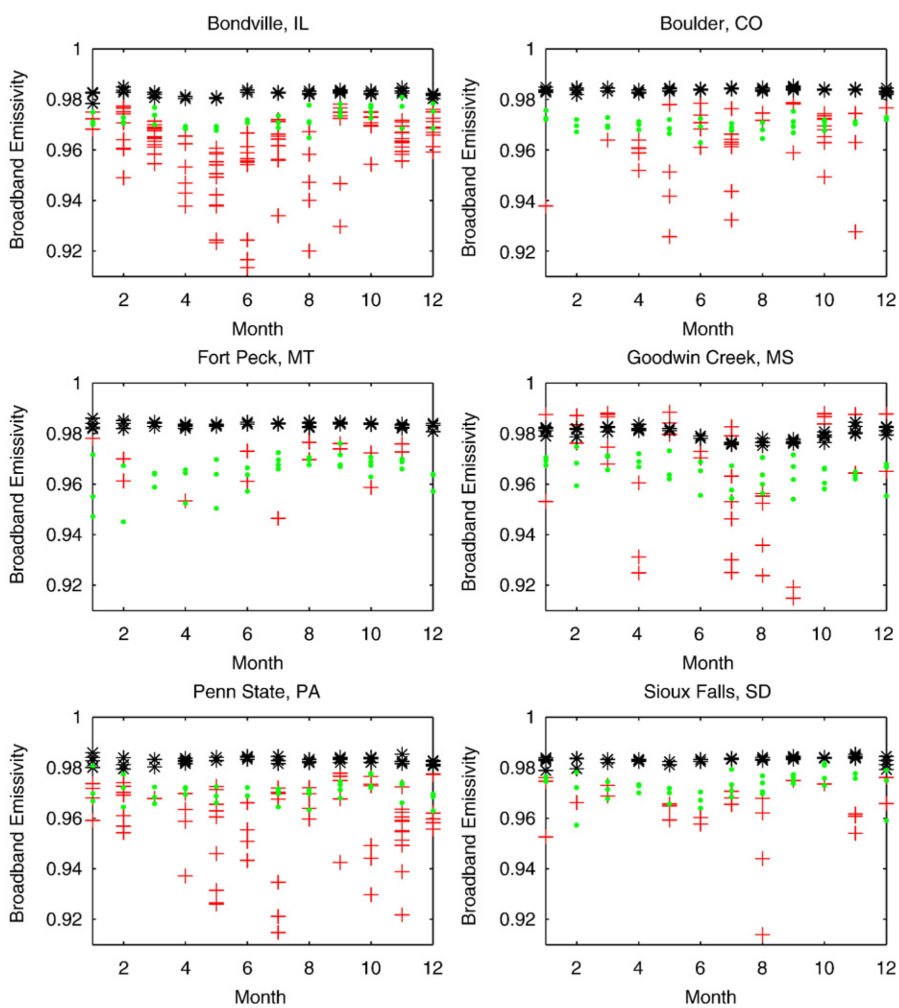

Fig. 6. Broadband emissivity calculated from MODIS Collection 4 (green dot) and Collection 5 (black star) monthly emissivity products and ASTER daily emissivity products (red plus sign) at a resolution of 0.05° at six SURFRAD sites. MODIS data used here is from 2002 to 2006 and ASTER data is from 2000 to 2007. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

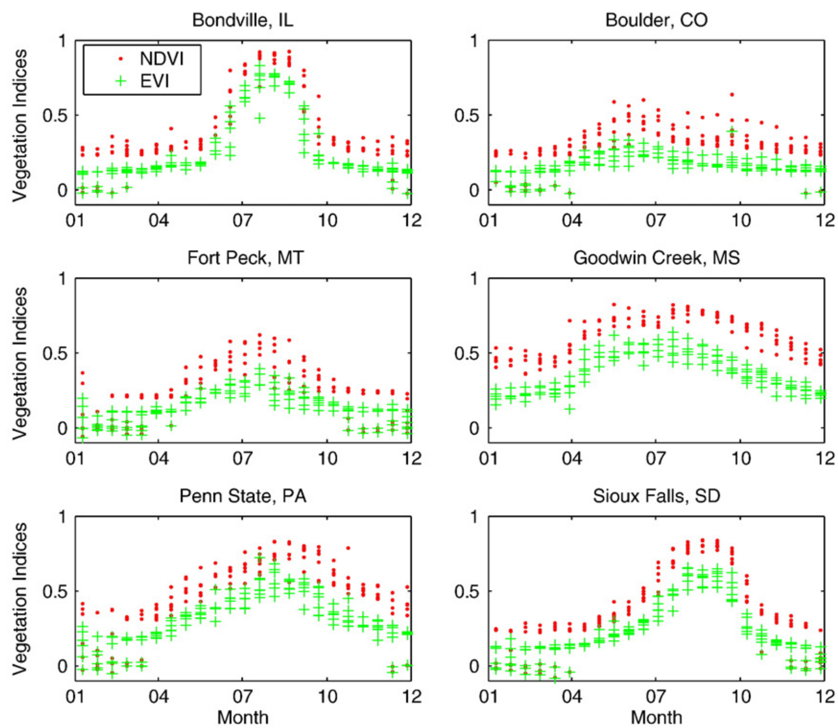

Fig. 7. Seasonal variation of MODIS 16-day Normalized Difference Vegetation Index (NDVI) and Enhanced Vegetation Index (EVI) at the six SURFRAD sites from 2001 to 2007.

applications. The multi-year averages of the ASTER and MODIS broadband emissivity are summarized in Table 3. Although ASTER emissivity differs from MODIS at each site, its average over the six sites is 0.97, very close to MODIS Collection 4 products and 0.01 less than that of MODIS Collection 5 products.

Fig. 7 shows that the MODIS vegetation indices have higher values in summer, which reflects the vegetation coverage at these sites in summer time. Fig. 8 also shows that soil moisture at the three sites also has higher values during early summer. Therefore, one can see that these lower values of ASTER broadband emissivity are inaccurate. Tonooka (2001, 2005) argued that the accuracy of ASTER standard LST/emissivity algorithm is worse when atmospheric water vapor content is higher, and that the ASTER standard algorithm tends to underestimate emissivity (Fig. 21 of Tonooka, 2005), which results large errors (both negative and positive) in LST retrieval (Fig. 8 of Tonooka, 2005). Surface water vapor pressure at the SURFRAD sites is shown in Fig. 9. The surface water vapor pressure is an indicator of atmospheric total water vapor amount, and the higher values of surface water vapor pressure indicate that atmospheric total water vapor amount is higher in the summer. Therefore, we infer that the lower values of ASTER broadband LST result from the limitations of the ASTER standard LST/emissivity algorithm when correcting atmospheric effects on thermal infrared radiance under conditions of high atmospheric water vapor content.

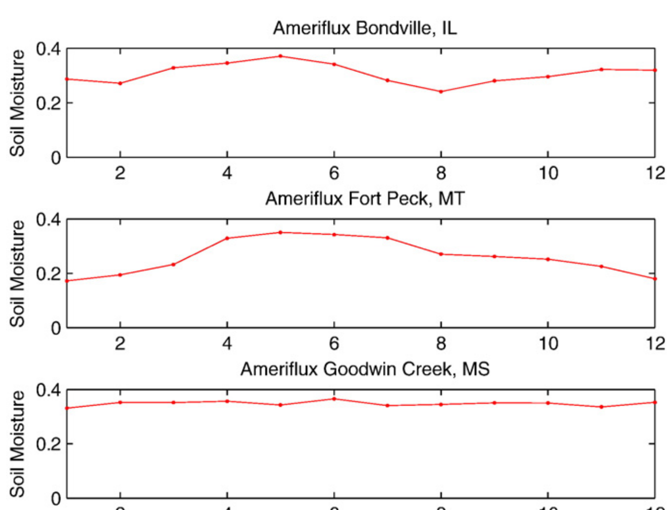

Fig. 8. Monthly soil at surface layer at three Ameriflux sites Bondville (40.01 N, 88.29 W), Fort Peck (48.31 N, 105.10 W), and Goodwin Creek (34.25 N, 89.87 W) averaged from 2000 to 2007. This can represent the soil moisture conditions at the corresponding SURFRAD sites where soil moisture is not measured.

5. Discussion and conclusions

In this study, Aqua MODIS Collection 5, and Terra ASTER LST and emissivity products are evaluated using long-term ground-based longwave radiation measurements collected at six SURFRAD sites from 2000 to 2007. LST is estimated from ground-based longwave radiation measurements collected at 10-m high towers, and from broadband emissivity from MODIS monthly LST/emissivity data (Wang et al., 2005). The sensitivity study showed that LST can be accurately estimated from ground-based longwave radiation measurements.

ASTER LST is directly compared to ground-based measurements because of its high spatial resolution (90 m). A total of 197 ASTER images were collected for six SURFRAD sites during 2000–2007. Comparison results showed that ASTER LST has a small bias. For the six sites during nighttime, the bias varies from −0.7 °C to 1.6 °C, with an average of 0.1 °C; STD varies from 0.9 °C to 3.0 °C, with an average of 1.6 °C; and the correlation coefficient varies from 0.98 to 1.0, with an average of 0.99. For the six sites during daytime, the bias varies from −0.9 °C to 4.0 °C; with an average of 0.3 °C; STD varies from 2.9 °C to 5.6 °C, with an average of 4.0 °C; and the correlation coefficient varies from 0.92 to 0.99, with an average of 0.96. The heterogeneity varies from 0.9 °C to 1.5 °C for the six sites during nighttime, with an average of 1.2 °C, while during daytime, the heterogeneity varies from 1.7 °C to 2.7 °C for the six sites, with an average of 2.4 °C. The heterogeneity of

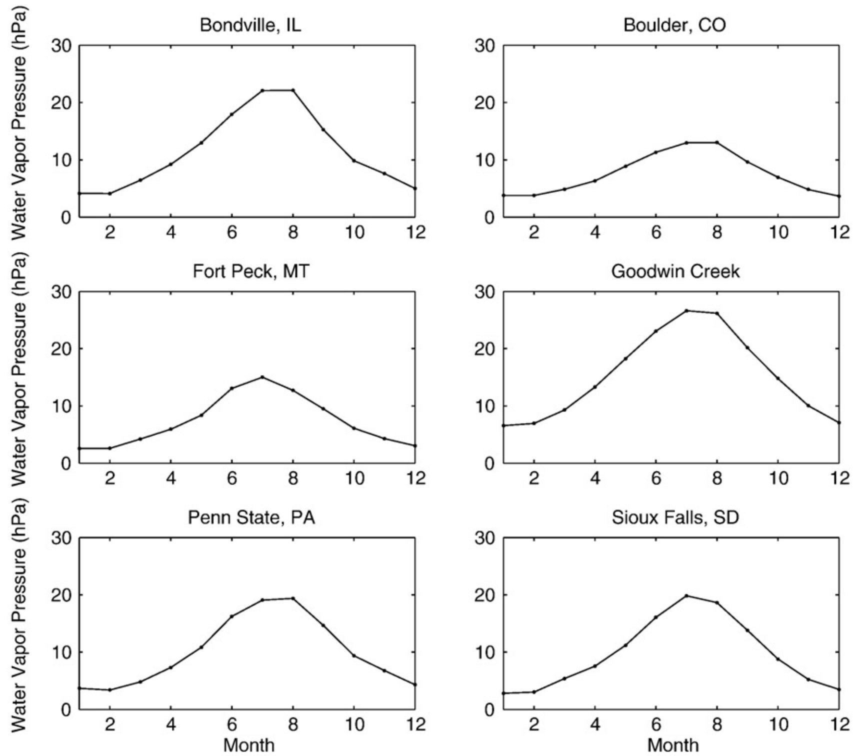

Fig. 9. The surface monthly water vapor pressure at the SURFRAD sites averaged from data collected from 2000 to 2007. The surface water vapor pressure is an indicator of atmospheric total water vapor amount.

LST at the ASTER TIR scale can account for 71% of STD during nighttime and 60% of STD during daytime at the ASTER TIR scale.

SURFRAD sites were not picked for evaluating satellite land products at a scale of 1 km or larger which is clearly seen in the pictures of Fig. 1, especially for surface characteristics that has large spatial heterogeneity, such as LST. Therefore, they are not suitable for MODIS LST evaluation. In this study, we only evaluate MODIS LST products at late night when LST heterogeneity is expected to be the least during a day (See Table 3). The Aqua MODIS LST at nighttime (about 1:30 a.m. local time) retrieved from the split-window algorithm was compared to ground-based measurements because spatial heterogeneity of LST at this time is believed to be at its lowest (without the effect of solar radiation on LST). The comparison of MODIS LST at 1-km resolution and ground-based measurements at six SURFRAD sites shows that the MODIS LST retrieval agrees well with ground-based measurements. The bias at five of the six sites is less than 0.3 °C, and −1.6 °C at the other one site, with an average of −0.2 °C. The STD varies from 1.5 °C to 2.6 °C for the six sites, with an average of 1.8 °C. The negative bias reported previously (e.g. Wang et al., 2007a, 2008a,b) in MODIS Collection 4 LST/emissivity products has been mostly removed. The large LST errors obtained by Wang et al. (2008a,b) may be due in a large part to the LST heterogeneity of the FLUXNET sites. Our study also shows that the accuracy of ground-based measurements is also a key parameter affecting the evaluations.

The major disadvantage of using ground-based longwave radiation measurements to evaluate satellite LST retrievals is that it only supplies one-point measurement at each site. In this study, we have demonstrated that the heterogeneities are highly correlated with STDs of the comparisons between ground-based LST measurements and ASTER and MODIS LST products at these six sites (Tables 2 and 3), which indicates that the STD of the comparison is controlled by surface heterogeneity. We further demonstrated that heterogeneity of LST accounts for the 60–70% of the STD of the comparisons between ground-based measurements and satellite LST retrievals. We found that LST heterogeneity for these six sites is 1.3–1.8 K at nighttime and 2–3 K at daytime at a scale of 90 m because solar radiation increases LST heterogeneity during daytime. Daytime LST heterogeneity has distinctive seasonal variations depending on surface incident solar radiation, soil moisture and land cover types around the site while nighttime LST heterogeneity is rather stable. One can infer that the real STDs of ASTER and MODIS LST retrievals are much less than those reported in this study. The STD would be less if more measurements were available at each site.

The accuracy of LST derived from ground-based longwave radiation measurements is another major factor influencing the ASTER and MODIS LST evaluation. The accuracy of longwave radiation measurements mainly depends on the sensor calibration (Wang & Liang, in press). In a recent study, we showed that thanks to the better sensor calibration at SURFRAD sites, the error of longwave radiation at SURFRAD sites is 3–4 W m^{-2} (a standard deviation, equivalent to an error of 0.6–0.7 °C in LST). The accuracy of broadband emissivity used to calculate LST from ground-based longwave radiation measurements is another parameter influencing the LST evaluation. We have shown that LST can be estimated from longwave radiation measurements at an accuracy of 0.2–0.5 °C given that the error of broadband emissivity is about ±0.02. LST from satellite and ground measurements may differ according to their measurement method. MODIS and ASTER use directional measurements in the atmospheric window, while ground-based longwave radiation measurements are hemispheric, wider spectrum derivations. However, we still cannot accurately quantify the directional emissivity of natural surfaces. Our previous study showed that the temperature gradient between surface and height (10 m at SURFRAD sites) where $L_{\uparrow}$ is measured may result in error of 0–0.3 °C in LST derived from $L_{\uparrow}$ and $L_{\downarrow}$ (Wang, 2004).

Collection 5 broadband emissivity calculated from MODIS monthly LST/emissivity is 0.01 greater than Collection 4 products and ASTER products, which is comparable to the 0.011 underestimation of Collection 4 broadband emissivity reported by our previous study (Wang et al., 2007a). There are some abnormally low ASTER emissivity retrievals during the summer because the ASTER standard LST/emissivity algorithm has difficulty in correcting atmospheric effects during the conditions of relative high atmospheric water vapor content (Tonooka, 2001, 2005). After filtering out these abnormal low values, the long-term averaged ASTER broadband emissivity is very close to that of MODIS Collection 4.

Acknowledgements

The ground-based measurements were obtained from SURFRAD (http://www.srrb.noaa.gov). MODIS and ASTER satellite data were obtained from http://www.edcdaac.usgs.gov/modis/dataprod.html with help from Dr. Wenhui Wang and graduate student Juan Du. This study is partially supported by NOAA grant NA07NES4400001 and NASA grant NNX08DC53G.

References

Augustine, J. A., DeLuisi, J. J., & Long, C. N. (2000). SURFRAD—A national surface radiation budget network for atmospheric research. *Bulletin of the American Meteorological Society*, *81*(10), 2341−2357.

Augustine, J. A., Hodges, G. B., Cornwall, C. R., Michalsky, J. J., & Medina, C. I. (2005). An Update on SURFRAD—The GCOS surface radiation budget network for the continental United States. *Journal of Atmospheric and Oceanic Technology*, *22*, 1460−1472.

Baldocchi, D., et al. (2001). FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor and energy flux densities. *Bulletin of the American Meteorological Society*, *82*, 2415−2434.

Berk, A., Anderson, G. P., Acharya, P. K., Chetwynd, J. H., Bernstein, L. S., Shettle, E. P., et al. (1999). MODTRAN 4 user's manual Hascom AFB, MA: Air Force Research Laboratory, Space Vehicles Directorate, Air Force Materiel Command 95 pp.

Coll, C., Caselles, V., Galve, J. M., Valor, E., Niclos, R., Sanches, J. M., et al. (2005). Ground measurements for the validation of land surface temperatures derived from AATSR and MODIS data. *Remote Sensing of Environment*, *97*, 288−300.

Coll, C., Caselles, V., Valor, E., Niclòs, R., Sánchez, J. M., Galve, J. M., et al. (2007). Temperature and emissivity separation from ASTER data for low spectral contrast surfaces. *Remote Sensing of Environment*, *110*, 162−175.

Gautier, C., & Landsfeld, M. (1997). Surface solar radiation flux and cloud radiative forcing for the Atmospheric Radiation Measurement (ARM) Southern Great Plains (SGP): A satellite, surface observations, and radiative transfer model study. *Journal of the Atmospheric Sciences*, *54*(10), 1289−1307.

Gillespie, A. R., Matsunaga, T., Rokugawa, S., & Hook, S. J. (1998). Temperature and emissivity separation from Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images. *IEEE Transactions on Geoscience and Remote Sensing*, *36*, 1113−1125.

Kerr, Y. H., Lagouarde, J. P., Nerry, F., & Ottle, C. (2000). Land surface temperature retrieval techniques and applications. In D. A. Quattrochi & J. C. Luvall (Eds.), *Thermal remote sensing in land surface processes* (pp. 33−109). Boca Raton, FL: CRC Press.

Liang, S. (2004). *Quantitative remote sensing of land surfaces.* John Wiley and Sons, Inc. 534 pages.

Ogawa, K., Schmugge, T., Jacob, F., & French, A. (2003). Estimation of land surface window (8–12 μm) emissivity from multi-spectral thermal infrared remote sensing — A case study in a part of Sahara Desert. *Geophysical Research Letters*, *30*(2), 1067. doi:10.1029/2002GL016354

Ohmura, A., et al. (1998). Baseline Surface Radiation Network (BSRN/WCRP): New precision radiometry for climate research. *Bulletin of the American Meteorological Society*, *79*(10), 2115−2136.

Palluconi, F., Hoover, G., Alley, R., Jentoft-Nielsen, M., & Thompson, T. (1999). *An atmospheric correction method for ASTER thermal radiometry over land.* ASTER algorithm theoretical basis document Pasadena, CA: Jet Propulsion Laboratory revision 3.

Qin, J., Liang, S., Liu, R., Zhang, H., & Hu, B. (2007). A weak-constraint based data assimilation scheme for estimating surface turbulent fluxes. *IEEE Geoscience and Remote Sensing Letters*, *4*(4), 649−653.

Rodell, M., et al. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society*, *85*, 381−394.

Salomonson, V., Barnes, W., Maymon, P., Montgomery, H., & Ostrow, H. (1989). MODIS: advanced facility instrument for studies of the Earth as a system. *IEEE Transactions on Geoscience and Remote Sensing*, *27*, 145−153.

Sobrino, J. A., Jiménez-Muñoz, J. C., Balick, L., Gillespie, A. R., Sabol, D. A., & Gustafson, W. T. (2007). Accuracy of ASTER level-2 thermal-infrared standard products of an agricultural area in Spain. *Remote Sensing of Environment*, *106*, 146−153.

Stokes, G. M., & Schwartz, S. E. (1994). The Atmospheric Radiation Measurement (ARM) program: Programmatic background and design of the cloud and radiation test bed. *Bulletin of the American Meteorological Society*, *75*(7), 1201−1221.

Tonooka, H. (2001). An atmospheric correction algorithm for thermal infrared multispectral data over land — A water–vapor scaling method. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(3), 682−692.

Tonooka, H. (2005). Accurate atmospheric correction of ASTER thermal infrared imagery using the WVS method. *IEEE Transactions on Geoscience and Remote Sensing*, *43*, 2778−2792.

Wan, Z. (2008). New refinements and validation of the MODIS land-surface temperature/emissivity products. *Remote Sensing of Environment*, *112*, 59−74.

Wan, Z., & Dozier, J. (1996). A generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Transactions on Geoscience and Remote Sensing*, *34*, 892−905.

Wan, Z., & Li, Z.-L. (1997). A physics-based algorithm for retrieving landsurface emissivity and temperature from EOS/MODIS data. *IEEE Transactions on Geoscience and Remote Sensing*, *35*, 980−996.

Wan, Z., Wang, P., & Li, X. (2004). Using MODIS land surface temperature and normalized difference vegetation index products for monitoring drought in the Southern Great Plains, USA. *International Journal of Remote Sensing*, *25*(1), 61−72.

Wan, Z., Zhang, Y., Zhang, Y. Q., & Li, Z. -L. (2002). Validation of the landsurface temperature products retrieved from Moderate Resolution Imaging Spectroradiometer data. *Remote Sensing of Environment*, *83*, 163−180.

Wang, K., (2004). A study on surface characteristics over the Tibetan Plateau using satellite remote sensed data and ground-based measurements. Dissertation of Peking University, pp144.

Wang, K., Li, Z., & Cribb, M. (2006). Estimation of evaporative fraction from a combination of day and night land surface temperature and NDVI: A new method to determine the Priestley –Taylor parameter. *Remote Sensing of Environment*, *102*, 293−305.

Wang, K., & Liang, S. (2008). An improved method for estimating global evapotranspiration based on satellite determination of surface net radiation, vegetation index, temperature, and soil moisture. *Journal of Hydrometeorology*, *9*(4), 712−727.

Wang, K., & Liang, S. (2009). Estimation of surface net radiation from solar shortwave radiation measurements. *Journal of Applied Meteorology and Climatology*, *48*, 634−643.

Wang, K., & Liang, S. (in press). Estimation of global surface downward longwave radiation under clear and cloudy conditions. *Geophysical Research Letter.*

Wang, W., Liang, S., & Meyer, T. (2008a). Validating MODIS land surface temperature products. *Remote Sensing of Environment*, *112*, 623−635.

Wang, K., Liang, S., Wang, D., & Zheng, T. (2008b). Simultaneous estimation of surface photosynthetically active radiation and albedo from GOES. *Remote Sensing of Environment,* revised.

Wang, K., Wan, Z., et al. (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. *International Journal of Remote Sensing*, *28*(11), 2549−2565.

Wang, K., Wan, Z., Wang, P., Sparrow, M., Liu, J., Zhou, X., et al. (2005). Estimation of surface long wave radiation and broadband emissivity using Moderate Resolution Imaging Spectroradiometer (MODIS) land surface temperature/emissivity products. *Journal of Geophysical Research*, *110*, D11109. doi:10.1029/2004JD005566

Wang, K., Wang, P., Li, Z., Cribb, M., & Sparrow, M. (2007). A simple method to estimate actual evapotranspiration from a combination of net radiation, vegetation index, and temperature. *Journal of Geophysical Research*, *112*, D15107. doi:10.1029/2006JD008351

Xiong, X. -X., Chiang, K. -F., Wu, A., Barnes, W. L., Guenther, B., Salomonson, V. V., et al. (2008). Multiyear on-orbit calibration and performance of Terra MODIS thermalemissive bands. *IEEE Transactions on Geoscience and Remote Sensing*, *46*, 1790−1803.

Yamaguchi, Y., Kahle, A. B., Tsu, H., Kawakami, T., & Pniel, M. (1998). Overview of Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). *IEEE Transactions on Geoscience and Remote Sensing*, *36*, 1062−1071.