Remote Sensing of Environment 154 (2014) 19–37

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Validation of Land Surface Temperature products derived from the Visible Infrared Imaging Radiometer Suite (VIIRS) using ground-based and heritage satellite measurements

Pierre C. Guillevic [a,*], James C. Biard [b,c], Glynn C. Hulley [a], Jeffrey L. Privette [c], Simon J. Hook [a], Albert Olioso [d], Frank M. Göttsche [e], Robert Radocinski [a], Miguel O. Román [f], Yunyue Yu [g], Ivan Csiszar [g]

[a] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[b] Cooperative Institute for Climate and Satellites, North Carolina State University, Asheville, NC, USA
[c] NOAA's National Climatic Data Center, Asheville, NC, USA
[d] Institut National de la Recherche Agronomique (INRA), EMMAH, Avignon, France
[e] Karlsruhe Institute of Technology, Karlsruhe, Germany
[f] Terrestrial Information Systems Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA
[g] NOAA's Center for Satellite Application and Research, College Park, MD 20740, USA

ARTICLE INFO

Article history:
Received 9 April 2014
Received in revised form 22 July 2014
Accepted 5 August 2014
Available online xxxx

Keywords:
Land Surface Temperature
VIIRS
Validation
MODIS
Ground-based measurements

ABSTRACT

Thermal infrared satellite observations of the Earth's surface are widely used to retrieve Land Surface Temperature (LST) and monitor LST changes around the world. Since January 2012, the Visible Infrared Imaging Radiometer Suite (VIIRS) onboard the Suomi National Polar-Orbiting Partnership (S-NPP) has provided daily observations of LST with a spatial resolution of 750 m at nadir. Comparison of the standard VIIRS LST product with the equivalent daily standard product from the Moderate Resolution Imaging Spectroradiometer (MODIS) collection-5 and with ground-based measurements over vegetated and inland water surfaces showed good agreement. Analysis indicated the accuracy and precision of the VIIRS product over these cover types was 0.2 K and 2.0 K respectively provided the analyses included appropriate compensation for any spatial heterogeneity in LST within the validation site. However, comparisons between in situ LST and the VIIRS and MODIS LST over arid and semi-arid regions indicate both satellite products significantly underestimate the LST, and the VIIRS algorithm can have large errors in the retrieved LST over areas of high atmospheric water vapor. Errors of up to 4 K were observed over semi-arid and arid areas due to incorrect characterization of emissivity, and differences of up to 15 K were observed over areas with high atmospheric water content between the VIIRS LST and matching MODIS LST.

1. Introduction

Land Surface Temperature (LST) is a key variable for surface water and energy budget calculations that can be obtained globally and operationally from satellite observations. The Visible Infrared Imaging Radiometer Suite (VIIRS) instrument was launched in October 2011 on the Suomi National Polar-Orbiting Partnership (S-NPP) satellite. VIIRS was designed to improve upon the capabilities of the Advanced Very High Resolution Radiometer (AVHRR) onboard NOAA's operational polar-orbiting satellites and provide observational overlap and continuity with both the AVHRR and the Moderate Resolution Imaging Spectroradiometer (MODIS) instruments on the NASA Terra and Aqua platforms of the NASA Earth Observing System (Justice et al., 2013). High temporal and spatial resolution LST products known as Environmental Data Records (EDR) have been derived from VIIRS data since processing began January 18th, 2012. These products provide a new source of LST for many applications, including weather forecasting (Meng, Li, Zhan, Shi, & Liu, 2009; Zheng et al., 2012), short-term climate prediction (Reichle et al., 2009; Reichle, Kumar, Mahanama, Koster, & Liu, 2010), extreme weather monitoring (Anderson, Hain, Wardlow, Mecikalski, & Kustas, 2011), and irrigation and water resource management including agricultural drought forecasting (Anderson, Allen, Morse, & Kustas, 2012; Kerr, Lagouarde, Nerry, & Ottlé, 2004). LST is particularly useful for agricultural drought forecasting since it is very sensitive to plant water stress and a strong indicator of changes in root zone soil moisture (Anderson et al., 1997, 2012; Anderson, Norman, Diak, Kustas, & Mecikalski, 1997; Guillevic & Koster, 2002; Guillevic et al., 2002; Moran et al., 2009).

The VIIRS thermal bands measure the spectral radiance emitted by the land surface and the atmosphere. The surface-emitted radiance is

* Corresponding author at: Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, USA. Tel.: +1 818 354 5034.
E-mail address: pierre.c.guillevic@jpl.nasa.gov (P.C. Guillevic).

http://dx.doi.org/10.1016/j.rse.2014.08.013
0034-4257/

attenuated by atmospheric constituents such as clouds, haze and absorbing gases (mainly water vapor, carbon dioxide, ozone and methane). Under clear sky conditions, the spectral radiance at the top of the atmosphere is the sum of three components: (1) the radiance emitted by the land surface and attenuated by the atmosphere, (2) the atmospheric radiance reflected by the land surface and attenuated by the atmosphere, and (3) the radiance emitted by the atmospheric constituents in the direction of the sensor. In order to retrieve LST from the thermal infrared radiance measured by VIIRS, the effect of the atmosphere must be removed and surface radiance has to be corrected for emissivity effects, which can otherwise introduce large uncertainties especially for split-window based algorithms (Jacob et al., 2008, Chap. 10; Kerr et al., 2004). Preliminary VIIRS LST EDR datasets became available publicly on October 22, 2012, and are currently being evaluated by scientists from NASA and NOAA, among others.

This study presents validation results for the VIIRS LST EDR obtained from comparisons with ground-based measurements and operational LST products from Aqua MODIS (1:30 am/pm satellite orbit). The limitations of the VIIRS LST algorithm are discussed, and guidance on methodologies and recommended good practice for validating moderate resolution satellite-based LST products is provided. Section 2 presents a short review of the different methods commonly used to validate satellite LST products and enumerates the sources of errors associated with each method. Section 3 describes the specific protocols used to evaluate the VIIRS LST EDR against ground based measurements and Aqua MODIS LST products, and discusses the challenges in retrieving the LST from satellite measurements. The different satellite products and the associated retrieval algorithms used in the study are presented in Section 3, and the in situ reference datasets are presented in Section 4. Section 5 presents the VIIRS LST validation results, while Section 6 suggests future algorithm refinements and provides general guidance on retrieving LST from remotely sensed data.

2. Satellite LST validation approaches

In order to maximize the usefulness of LST for research and studies it is necessary to know the uncertainty in the LST measurement. The VIIRS LST was designed to meet the quality specifications of operational users, such as Numerical Weather Prediction modelers. Multiple validation methods and activities are necessary to assess LST compliance with the specifications. A detailed presentation of previous satellite-based LST validation efforts is available in review studies by Li et al. (2013), Merchant et al. (2013) and Schneider, Ghent, Corlett, Prata, and Remedios (2012). Four different methods have been widely used to validate and determine the uncertainties in LST products derived from satellite measurements:

- Temperature based validation. This approach involves comparisons with ground-based measurements of LST, and has been frequently used to validate LST products retrieved from MODIS (Bosilovich, 2006; Coll et al., 2005, Coll, Galve, Sanchez, & Caselles, 2010; Guillevic et al., 2012, 2013; Hook, Vaughan, Tonooka, & Schladow, 2007; Trigo, Monteiro, Olesen, & Kabsch, 2008; Wan, 2008; Wang & Liang, 2009; Wang, Liang, & Meyers, 2008), from the Spinning Enhanced Visible and Infrared Imager onboard Meteosat Second Generation (MSG/SEVIRI) (Göttsche et al., 2013; Kabsch, Olesen, & Prata, 2008; Trigo et al., 2008), from AVHRR (Prata, 1994), from the Advance Spaceborne Thermal Emission and Reflection (ASTER) radiometer onboard Terra (Sobrino et al., 2007), from the Along Track Scanning Radiometer (ATSR) (Prata, 1994), or from VIIRS (Li et al., 2014). This approach allows the uncertainties in LST products to be determined, however, a large number of in situ measurements are needed if the validation site is spatially heterogeneous in order to characterize it correctly (Guillevic et al., 2012). Furthermore, most field radiometers collect observations at nadir angles, whereas wide field-of-view satellite scanners like VIIRS collect most observations off-nadir. These limitations provide significant uncertainty that is very difficult to eliminate. Therefore, the method is particularly suited for studies over inland water bodies which provide large spatially homogenous temperature targets and can be used to both validate and refine the retrieval algorithm (Coll, Hook, & Galve, 2009; Hook et al., 2007; Hulley, Hook, & Schneider, 2011). Unfortunately, validation over water does not assess the LST algorithm correction for surface emissivity.
- Scene-based comparisons. This approach involves comparing a new satellite LST product with a heritage LST product (Guillevic et al., 2013; Hulley & Hook, 2009a; Jacob et al., 2004; Trigo et al., 2008). The method can be particularly valuable for finding spatial disagreements between LST products for a wide range in cover types. However, this is not an absolute validation and satellite LST intercomparisons alone do not provide an independent validation measurement unless one of the satellite products has been independently validated. Different retrieval algorithms based on similar assumptions and formulations (e.g. split-window) can be highly consistent with each other but biased when compared to ground reference measurements. Also, this approach requires accounting for differences in spatial resolution, view angle and overpass time between the two different satellite datasets.
- Radiance-based validation (Coll, Wan, & Galve, 2009; Hulley & Hook, 2012; Niclòs, Galve, Valiente, Estrela, & Coll, 2011; Wan, 2014; Wan & Li, 2008). This approach requires precise estimates of channel specific surface emissivity values and atmospheric temperature, and water vapor profiles coincident with the satellite overpass. LST values are then derived by inverting a radiative transfer model. Radiance-based validation has the advantage that temperature measurements are not required at the time of the overpass. Instead emissivity measurements made at a different time can be used with model-based atmospheric information. The method is best for large-scale validation efforts on a global scale or for products with coarse spatial resolution.
- Time series comparisons (Hook et al., 2007; Merchant et al., 2013). This method is used to detect problems that can occur during the instrument's life, e.g. calibration drift (Hook et al., 2007), or unrealistic outliers due to cloud coverage (Schneider et al., 2012). However, the approach requires relatively long time series of observations over very stable targets over time. The VIIRS standard LST has only been available for 1.5 year and a longer record is required before the data lend themselves to this approach.

The four different approaches are complementary and provide different levels of information about the quality of the retrieved LST. These four methods are part of the validation plan for the ATSR LST products (Schneider et al., 2012), for example, and are usually required to achieve Stage-3 validation status as defined by the MODIS land validation protocol (http://landval.gsfc.nasa.gov).

3. Validation methodology for VIIRS LST

Under clear sky conditions, the top of atmosphere radiance measured by a spaceborne sensor ($L_{sat,\lambda}$) includes contributions from the surface emission, the atmospheric upwelling radiance ($L^{\uparrow}_{sky,\lambda}$) and atmospheric downwelling radiance ($L^{\downarrow}_{sky,\lambda}$) reflected by the Earth's surface and attenuated by the atmosphere (Eq. 1). Retrieval algorithms rely on one or more top-of-atmosphere spectral measurements to account for atmospheric effects and estimate LST.

$$L_{sat,\lambda} = \left[\varepsilon_\lambda B_\lambda(LST) + (1-\varepsilon_\lambda) L^{\downarrow}_{sky,\lambda}\right] \tau_\lambda + L^{\uparrow}_{sky,\lambda} \quad (1)$$

where ε_λ is the spectral emissivity at wavelength λ or associated with a specific (relatively narrow) domain $[\lambda_1, \lambda_2]$ centered on wavelength λ, $B_\lambda(T)$ is the Planck function describing the radiance of a black body at temperature T, and τ_λ is the atmospheric attenuation.

The approaches presented for validating VIIRS LST EDR are based on comparisons with ground-based measurements and LST products from other instruments, e.g. Aqua MODIS LST products. The study is designed to quantify the spatial variability and atmospheric effects on LST derived from space-borne thermal infrared instruments. This section presents the validation approaches, their relative complexity, and the accuracy that can be achieved by them.

3.1. Validation with ground-based measurements

Validation with ground-based measurements involves a comparison of satellite-derived LST with collocated and simultaneously acquired LST retrievals from in situ radiometers, which also require accurate knowledge of surface emissivity, ε_λ, over their smaller surface targets. Ground-based LST (LST_{ground}) is retrieved from in situ measurements of radiance emitted from the surface ($L_{ground,\lambda}$) and downwelling radiance from the sky ($L^{\downarrow}_{sky,\lambda}$) (Eq. 2).

$$LST_{ground} = B_\lambda^{-1}\left[\frac{1}{\varepsilon_\lambda}\left(L_{ground,\lambda} - (1-\varepsilon_\lambda)L^{\downarrow}_{sky,\lambda}\right)\right] \quad (2)$$

where all remaining symbols have the same meaning as in Eq. (1).

The primary uncertainties in ground-based LST retrieval are associated with the accuracy of surface emissivity and down-welling radiance (Hook et al., 2007). Generally, neither spectral nor directional measurement of downwelling thermal radiation from the atmosphere is routinely sampled in the field. When derived from models (Brutsaert, 1975; Idso, 1981; MODTRAN) or directional measurements (Kondratyev, 1969), estimates of atmospheric radiation can have significant uncertainties, especially for warm and humid atmospheres.

3.1.1. Spatial representativeness of ground-based LST

Depending on the experimental design, i.e. the sensor's field of view, and the height and angle at which the sensor is mounted, the footprint of a ground-based infrared instrument is typically from 1 to 10 m on a side for spectral radiometers and 10 to 100 m for pyrgeometers. For example, the spatial representativeness of ground-based LST derived from SURFRAD's pyrgeometer measurements is around 70 m × 70 m. Most vegetated landscapes contain various land cover types or soils, and therefore, the LST measured by a station at one specific location usually does not represent the surrounding area that is included in the lower resolution satellite sensor footprint (e.g. around 1 km).

To address this issue an up-scaling model can be used to help interpret validation results over heterogeneous land surfaces (Guillevic et al., 2012). A summary description is repeated here for completeness. The approach uses a physically based land surface model driven by in situ atmospheric forcing measurements and high-resolution imagery to describe satellite LST footprints over ground stations. To represent the spatial variability within a satellite pixel, it is assumed that the sub-pixel temperature variability is mainly due to land cover heterogeneity and variability in surface biophysical parameters, such as vegetation density, emissivity or albedo. Because the canopy evapotranspiration is usually greater than the bare soil evaporation, the temperature of vegetated areas under low soil water stress conditions is typically lower than the temperature of barren surfaces. The method estimates the temperature of each land cover class inside a mixed pixel using a land surface model driven by the measured atmospheric forcing and observed surface biophysical properties from a nearby tower/site. The atmospheric forcing is assumed to be uniform over the satellite footprint, which is approximately 1 km × 1 km for VIIRS. Depending on the canopy structure, sun illumination and viewing directional configurations, satellites measure different surface radiometric temperatures, particularly over sparsely vegetated regions and open canopies (Guillevic et al., 2013; Lagouarde, Ballans, Moreau, Guyon, & Coraboeuf, 2000). The selected validation sites represent short vegetation areas (i.e. grassland, cropland), inland waters or bare soils associated with small directional effects that have not been represented in the approach.

3.1.2. The up-scaling model

The SEtHyS land surface model (Coudert, Ottlé, Boudevillain, Demarty, & Guillevic, 2006) was used to quantify differences in subpixel temperature between classes of surface biophysical properties (i.e. different surface types or different vegetation densities) with respect to the LST measured by the station. This reduces the impact of model systematic errors and uncertainties in the atmospheric forcing on the assessment of the satellite pixel LST. The SEtHyS land surface model is a two-source energy balance model that simulates the energy and water transfer between the surface and the atmosphere, and describes the evolution of surface state variables such as LST and soil moisture. In addition to LST estimates, the selected ground stations also provide accurate measurements of local environmental information, in particular the atmospheric forcing required by the SEtHyS model. These measurements include air temperature, relative humidity and wind speed at the surface, and incoming shortwave and longwave radiation.

The steps of the scaling methodology are as follows:

1. Calibration of the SEtHyS model using ground observations. This task determines the optimal set of internal model parameters that allows the model to describe the observed in situ LST (see Guillevic et al., 2012, for a detailed description of the model calibration method).
2. Characterization of the satellite footprint, which depends on pixel geolocation, viewing zenith angle, satellite altitude (824 km for S-NPP, and 705 km for Terra and Aqua satellites) and the instantaneous field of view (911 μrad for VIIRS, and 1315 μrad for MODIS at moderate resolution).
3. Representation of the LST of each surface end-member using the SEtHyS model forced by the observed biophysical properties at high-resolution. We used the leaf area index (LAI) derived from MODIS Normalized Difference Vegetation Index (NDVI) standard products at 250 m spatial resolution to describe the spatial variability of vegetation density around the station (see Section 4.3: MODIS ancillary products).
4. Calculation of LST at satellite resolution from a weighted mean of n radiative contributions from each land cover class (Guillevic et al., 2012) (Eq. 3).

$$LST = \left[\frac{1}{\varepsilon}\sum_{i=1}^{n} f_i\,\varepsilon_i\,T_i^4\right]^{\frac{1}{4}} \quad with \quad \varepsilon = \sum_{i=1}^{n} f_i\,\varepsilon_i \quad and \quad \sum_{i=1}^{n} f_i = 1 \quad (3)$$

where f_i is the cover fraction of land cover class i at temperature T_i and with broadband emissivity ε_i. ε is the broadband surface r-emissivity as defined by Norman and Becker (1995).

The scaling methodology requires high-resolution information about vegetation density. The NDVI is an indicator of green biomass, and an indicator of the photosynthetic efficiency of the plants. Greener areas are usually associated with higher evapotranspiration and, consequently, characterized by lower LST. We use a logarithmic relationship between NDVI and LAI to estimate the LAI at 250 m resolution (Guillevic et al., 2012). The relationship is calibrated using NDVI and LAI values derived from MODIS data at 1 km. Then, LAI at 250 m resolution is estimated by applying the determined relationship to MODIS NDVI at 250 m spatial resolution.

3.2. Comparison with heritage satellite LST data

The method provides useful quality information with respect to spatial patterns in LST product differences. However, product intercomparisons do not represent a comprehensive validation and cannot substitute validation efforts with ground-based reference data. For example, two different products can be highly consistent with each

other while at the same time significant discrepancies are observed between the products and the ground reference data. This is especially true if the LST retrieval algorithms are based on similar assumptions, e.g. VIIRS and MODIS standard LST products are both based on the split window technique. When comparing multiple satellite LST products, the three main sources of discrepancies that can affect the quality of algorithm performance evaluations are differences in spatial resolution, overpass times, view angles, and cloud cover. The impact of directional effects on satellite LST products has been described by Guillevic, Gastellu-Etchegorry, Demarty, and Prevot (2003, 2013), Lagouarde et al. (2000), Pinheiro, Privette, and Guillevic (2006), Sobrino, Jiménez-Munoz, and Verhoef (2005) and Trigo et al. (2008). VIIRS and MODIS observations are never strictly simultaneous and differences in sensor footprint increase with differences in view zenith angle. To reduce the effect of cloud contamination on validation results, only cloud free data are used during different periods of the year to analyze possible seasonal effects on the observed discrepancies. However, the decision on whether a pixel is cloud-free is typically based on characteristics of the data themselves and there may be some ambiguity in the cloud assessment.

3.2.1. Match-up tool

In this study we used a match-up tool to select coincident VIIRS and MODIS granules with respect to the satellite overpass times and view angles. Given a date range, satellite angular separation limit and satellite time separation limit, the application retrieves time spans in which both MODIS and VIIRS on their two different platforms observe the same place on the Earth within those limits. The application uses orbit histories for the two satellites to select time spans within the date range where the position of one satellite passes within the angular separation limit of a position held by the other satellite at a previous time that is within the time separation limit. These “near miss” time spans are referred to as Simultaneous Nadir Overpasses (SNOs). Once the list of SNOs has been produced, the times are used to search for processed LST products from the two satellites. The results can be filtered by date/time, geolocation, angular separation distance, or time separation interval. For the purposes of this study, the maximum time separation and angular separation between the two satellites are set to 10 minutes and 2°, respectively. The results were also filtered to select granules with ground tracks intersecting the boundaries of the contiguous United States.

4. Satellite data

4.1. VIIRS LST EDR

Since August 11, 2012, the VIIRS EDR has been operationally produced using a single split window algorithm that is effectively insensitive to solar radiation (Yu, Privette, & Pinheiro, 2005). The algorithm uses brightness temperatures measured in channel M15 (T_{15}) and channel M16 (T_{16}) centered on 10.76 μm and 12.01 μm, respectively (Eq. 4).

$$LST = a_0 + a_1 T_{15} + a_2 (T_{15} - T_{16}) + a_3 (sec\theta_v - 1) + a_4 (T_{15} - T_{16})^2 \quad (4)$$

where a_k (with $k = 0$ to 4) are the algorithm coefficients and θ_v is the sensor zenith angle. Daytime and nighttime sets of coefficients were derived for 17 different surface types from regression analysis of MODTRAN radiative transfer simulations for globally representative atmospheric and surface conditions. The International Geosphere-Biosphere Programme (IGBP) global classification map is used to identify the surface type associated with each pixel. The algorithm regression coefficients were generated from an ensemble of MODTRAN radiative transfer simulations using a comprehensive set of geophysical parameters (VIIRS LST ATBD, 2011). Surface temperatures and coherent atmospheric temperature and water vapor profiles were derived from National Center for Environmental Prediction (NCEP) global simulations at 2.5° × 2.5° spatial resolution. LST values were sampled from 196 K to 327 K. Distribution of band-averaged spectral emissivity values for each surface type (Table 1) was derived from the MOSART database (VIIRS LST ATBD, 2011). These were used to produce a total of 268,128 samples representing 12 days and nights (1 day and night per month) over a global grid, and provided an ensemble of training data covering global, diurnal and seasonal features (VIIRS LST ATBD, 2011). The regression coefficients used in the algorithm since August 11, 2012 (Interface Data Processing Segment (IDPS) version Mx7.3) were trained for sensor zenith angles less than 40°, and are presented in Tables 2 and 3.

The spatial resolution of VIIRS raw radiometric measurements at moderate resolution is around 750 m at nadir and around 1.5 km at the edge of the swath. VIIRS detectors are rectangular, with the smaller dimension projecting along the scan. At nadir, three detector footprints are aggregated to form a single VIIRS pixel. Moving along the scan away from nadir, the detector footprints become larger both along track and along scan, due to geometric effects and the curvature of the Earth (Wolfe et al., 2013). The pixel aggregation scheme is changed from three to two detectors at a scan angle of around 32°, and from two to one detector at around 48°, which provides a more uniform pixel size over the scan.

The accuracy and precision requirements specified by NOAA's Joint Polar Satellite System (JPSS) program for the VIIRS LST EDR are 1.5 K and 2.5 K, respectively, for clear conditions. These requirements were driven primarily by the Numerical Weather Prediction community. VIIRS granules are archived and distributed by NOAA's National data centers (http://www.class.ncdc.noaa.gov). The NASA's Land Product Evaluation and Analysis Tool Element (LPEATE) routinely provides 50 km × 50 km subsets of VIIRS granules centered on selected validation sites for calibration and validation purposes (http://viirsland.gsfc.nasa.gov/Subsets.html). We used data from both CLASS and LPEATE in this study.

4.2. MODIS LST product

The daily daytime and nighttime Collection-V5 MODIS LST level-2 products are derived from the MODIS instrument onboard the Terra and Aqua satellites: the two products are referred to as MOD11_L2 and MYD11_L2, respectively. Satellite overpass times at the equator are around 10:30 am/pm (solar local time) for Terra and 1:30 am/pm for Aqua, and the spatial resolution of MODIS at nadir is 927 m. Along each scan, MODIS's off-nadir scan angle increases to values up to 65°, which causes the sensor's spatial resolution to degrade to about 6 km in the along-scan direction. The generalized split-window algorithm (Wan & Dozier, 1996) is used to derive LST values from brightness temperature measurements in MODIS band 31 (T_{31}) and band 32 (T_{32}) centered on 11.03 μm and 12.02 μm, respectively (Eq. 5).

$$LST = b_0 + \left(b_1 + b_2 \frac{1-\varepsilon}{\varepsilon} + b_3 \frac{\Delta\varepsilon}{\varepsilon^2}\right) \frac{T_{31} + T_{32}}{2} + \left(b_4 + b_5 \frac{1-\varepsilon}{\varepsilon} + b_6 \frac{\Delta\varepsilon}{\varepsilon^2}\right) \frac{T_{31} - T_{32}}{2} \quad (5)$$

where ε and Δε are the mean and the difference of the emissivities in bands 31 and 32. The algorithm coefficients b_k (with $k = 0$–6) depend on viewing zenith angle, surface air temperature (T_{air}) and atmospheric water vapor content. The coefficients were derived for daytime and nighttime from regression analysis of radiative transfer simulation data for a comprehensive set of LST values varying from $T_{air} - 16$ K to $T_{air} + 16$ K (Wan & Dozier, 1996). In the standard LST product, information about surface air temperature and total column water vapor is taken from the MODIS atmospheric profile product (MOD07)

Table 1
Spectral emissivity values associated with the 17 IGBP surface types used by the VIIRS and MODIS LST algorithms, respectively. For each surface type, the VIIRS algorithm uses distributions of spectral emissivity, while the MODIS algorithm (MYD11) estimates the surface emissivity values using prescribed green and senescent components (dry/fine or coarse components for snow).

Id	Surface type	VIIRS emissivity				MODIS emissivity			
		10.76 μm (M15)		12.01 μm (M16)		11.03 μm (band 31)		12.02 μm (band 32)	
		Mean	STD	Mean	STD	Green	Senescent	Green	Senescent
1	Evergreen needleleaf forest	0.985	1.5e−3	0.986	9.5e−4	0.987	0.987	0.989	0.989
2	Evergreen broadleaf forest	0.982	1.1e−3	0.985	9.6e−4	0.981	0.981	0.984	0.984
3	Deciduous needleleaf forest	0.981	3.1e−3	0.982	2.6e−3	0.987	0.986	0.989	0.988
4	Deciduous broadleaf forest	0.974	1.4e−5	0.975	4.9e−4	0.981	0.968	0.984	0.971
5	Mixed forest	0.979	9.4e−4	0.980	2.4e−4	0.981	0.968	0.984	0.971
6	Closed shrublands	0.974	4.0e−3	0.977	3.2e−3	0.983	0.973	0.987	0.975
7	Open shrublands	0.974	2.2e−3	0.977	1.4e−3	0.972	0.970	0.973	0.975
8	Woody savannas	0.981	3.8e−3	0.982	2.1e−3	0.982	0.975	0.985	0.978
9	Savannas	0.977	2.3e−3	0.980	7.3e−4	0.983	0.973	0.987	0.975
10	Grasslands	0.970	6.3e−3	0.974	6.9e−3	0.983	0.973	0.987	0.975
11	Permanent wetlands	0.984	3.5e−3	0.979	2.8e−3	0.992	0.992	0.988	0.988
12	Croplands	0.978	5.0e−3	0.982	3.5e−3	0.983	0.977	0.987	0.982
13	Urban and built-up	0.972	7.2e−3	0.977	7.6e−3	0.970	0.966	0.976	0.972
14	Cropland/natural vegetation	0.976	3.0e−3	0.980	2.8e−3	0.983	0.973	0.987	0.975
15	Snow and ice	0.994	1.6e−3	0.987	2.1e−3	0.993	0.984	0.990	0.971
16	Barren/sparsely vegetated	0.973	7.7e−3	0.972	6.0e−3	0.965	0.965	0.972	0.972
17	Inland water	0.990	6.5e−4	0.987	1.9e−3	0.992	0.992	0.988	0.988

(Wan, 2008). For each surface type, the spectral emissivity values in bands 31 and 32 (Table 1) are defined as a combination of green and senescent components (Snyder, Wan, Zhang, & Feng, 1998).

4.3. MODIS ancillary products

Additional products derived from the MODIS Terra alone (with the name prefix "MOD"), the MODIS Aqua instrument alone (prefix "MYD") or from both instruments (prefix "MCD") were used to describe the spatial variability of surface biophysical parameters around the field stations. These were the vegetation index at 250 m (MOD13Q1, and MYD13Q1) and 1 km (MOD13A2) spatial resolution and vegetation density, i.e. leaf area index, at 1 km spatial resolution (MCD15A2). The temporal resolution of NDVI products derived from both MODIS instruments at 250 m is 8 days. A full description of the products is given by Justice et al. (1998), and updated information is available at the NASA Land Processes Distributed Active Archive Center (LPDAAC) website (https://lpdaac.usgs.gov/lpdaac/products/modis_products_table/).

4.4. ASTER ancillary data

The ASTER sensor was launched together with the MODIS sensor on the Terra satellite in December 1999. ASTER has five channels in the thermal infrared domain (between 8.1 μm and 11.6 μm) used to derive LST and emissivity products at a spatial resolution of 90 m at nadir and at a temporal resolution of 16 days. View zenith angles are less than 9° for normal ASTER viewing conditions. ASTER LST and emissivity products (referred to as AST08 and AST05 respectively) are generated in a swath format using the Temperature Emissivity Separation (TES) algorithm (Gillespie et al., 1998) during daytime and nighttime. ASTER data were used in this study to evaluate the spatial variability of LST around field stations and to evaluate the spatial representativeness of ground-based measurements.

5. Validation sites

Nine in situ validation sites were selected to evaluate the VIIRS LST EDR over various land cover types: cropland, grassland, sparsely vegetated arid area, desert, and inland waters (Table 4). Two of the sites,

Table 2
Daytime regression coefficients (a_i, with i from 0 to 4) used in the VIIRS LST split window algorithm (Eq. 4) since August 11, 2012 (IDPS version Mx7.3). The description of the 17 IGBP surface types is provided in Table 1.

Surface type	Daytime algorithm coefficients				
	a_0	a_1	a_2	a_3	a_4
1	−6.33485	1.028104	1.310552	1.063013	0.441287
2	−5.47409	1.024861	1.660752	2.42386	0.327702
3	−4.58919	1.022091	1.103522	0.813863	0.571229
4	−5.45372	1.033022	1.811434	−2.70106	0.298936
5	−7.51475	1.033524	1.201031	1.246776	0.438322
6	−2.44143	1.016145	1.566102	0.370878	0.193816
7	−7.09271	1.033233	1.350841	1.213785	0.367221
8	−9.81976	1.041334	0.970194	1.372316	0.55345
9	−10.6068	1.04433	1.158848	1.021086	0.463121
10	−6.44958	1.031742	1.303886	0.059388	0.394892
11	−7.78559	1.033159	0.558588	1.036486	0.740771
12	−11.9967	1.049311	1.160366	2.728394	0.434421
13	−7.32977	1.034073	1.576136	0.978909	0.268421
14	−9.31956	1.04016	1.069135	2.379238	0.469663
15	−4.9299	1.01913	1.683574	0.352144	−0.26357
16	−12.7833	1.052898	0.944545	0.889798	0.506456
17	−8.92885	1.033913	1.870167	1.479963	0.354069

Table 3
Nighttime regression coefficients (a_i, with i from 0 to 4) used in the VIIRS LST split window algorithm (Eq. 4) since August 11, 2012 (IDPS version Mx7.3). The description of the 17 IGBP surface types is provided in Table 1.

Surface type	Nighttime algorithm coefficients				
	a_0	a_1	a_2	a_3	a_4
1	−2.44023	1.013721	1.597063	0.397226	0.243329
2	−10.9737	1.043302	1.337757	1.192763	0.433421
3	−2.81076	1.015627	1.253511	0.782135	0.474349
4	−0.67262	1.008506	1.782233	1.031163	0.193119
5	−1.58225	1.011321	1.569283	0.874003	0.341845
6	−2.86866	1.017388	1.169604	0.40632	0.470555
7	−3.67031	1.020234	1.367489	0.974629	0.383254
8	−6.1826	1.027303	1.131303	0.819621	0.519747
9	−7.93398	1.034157	1.219383	1.250769	0.450993
10	−2.19848	1.015395	1.473563	1.304318	0.286378
11	−4.76334	1.021443	1.198395	0.313569	0.606909
12	−0.98175	1.010598	1.322288	−0.39396	0.397286
13	0.269089	1.006037	1.40562	0.363574	0.370285
14	−3.08412	1.016865	1.563887	0.810411	0.296177
15	−3.29337	1.013452	1.323036	0.251886	−0.22787
16	−8.63783	1.037961	1.034632	0.83134	0.478393
17	−8.89917	1.033886	1.848356	1.511793	0.354354

Table 4
List of validation sites including geolocation, elevation, surface emissivity and basic description of the surface type at station location and around the station within VIIRS footprints.

Site location	Id	Latitude	Longitude	Elevation	Surface type at station	Surface type around station	Surface emissivity
Lake Tahoe, CA/NV	LTO	39.153°N	120.000°W	1897 m	Inland water	Inland water	0.990
Table Mountain, CO	TBL	40.126°N	105.238°W	1692 m	Sparse grassland	Grassland/crop	0.973
Bondville, IL	BON	40.051°N	88.373°W	213 m	Grassland	Cropland	0.976
Goodwin Creek, MS	GWN	34.255°N	89.873°W	96 m	Grassland	Grassland	0.975
Fort Peck, MT	FPK	48.308°N	105.102°W	636 m	Grassland	Grassland	0.979
Desert Rock, NV	DRA	36.623°N	116.020°W	1004 m	Arid shrubland	Arid shrubland	0.966
Penn State U., PA	PSU	40.720°N	77.931°W	373 m	Cropland	Cropland/forest	0.972
Sioux Falls, SD	SXF	43.734°N	96.623°W	483 m	Grassland	Grassland/urban	0.978
Gobabeb, Namibia	GOB	23.551° S	15.051° E	425 m	Desert area	Desert area	0.950

located at Lake Tahoe, CA/NV, and Gobabeb, Namibia have been operated continuously for several years for LST validation. The others are part of NOAA's Surface Radiation (SURFRAD) operational network initially developed to characterize surface radiation. Around half of the sites are representative of homogeneous areas at scales appropriate for validating VIIRS products, i.e. a large lake, desert or grassland. The remaining sites are located over relatively heterogeneous areas in terms of land cover types and surface biophysical parameters.

5.1. Inland water site: Lake Tahoe, CA/NV

Since 1999, NASA's Jet Propulsion Laboratory (JPL) has been maintaining four continuous LST monitoring stations on Lake Tahoe, CA/NV, a 35 km long and 15 km wide lake on the California–Nevada border. Each station has a JPL-built self-calibrating thermal infrared radiometer that measures surface brightness temperature in the 8–14 μm atmospheric window from a height of 1 m and several bulk temperature sensors, placed ~2 cm beneath the surface (Fig. 1). The temporal resolution of the in situ measurements is 2 minutes (one measurement of the target's radiance and self-calibration process every 2 minutes). The radiometers are typically exchanged at 6-month intervals for maintenance. Validation at JPL's NIST-traceable calibration facility indicates that changes during deployments are small and less than 0.05 K (Hook et al., 2003). A full meteorological station (wind speed, wind direction, air temperature, relative humidity and net radiation) is also deployed at each station. Channel-specific (8–14 μm) incoming atmospheric radiation required for atmospheric correction (Eq. 2) are derived from MODTRAN 5.2 simulations using atmospheric profiles obtained from local sounding balloon launches and model data generated by the National Centers for Environmental Prediction (NCEP). NCEP produces global model values on a $1° \times 1°$ grid at 6 h intervals. The atmospheric profiles at the satellite overpass times are derived from NCEP data using linear interpolation between the two successive intervals bracketing the observation time. Lake Tahoe is on a model grid point and no spatial interpolation is required. More information on the measurements is available at http://laketahoe.jpl.nasa.gov.

5.2. Desert site: Gobabeb, Namibia

Gobabeb is located in the Namib Desert in Namibia, and is one of four permanent validation stations operated by Karlsruhe Institute of Technology (KIT) within the framework of the Satellite Application Facility on Land Surface Analysis (LSA SAF, http://landsaf.meteo.pt/). The main purpose of the station is to validate and monitor LST products retrieved from thermal infrared satellite measurements, and it is supported by the European Organization for the Exploitation of Meteorological Satellites

Fig. 1. Images of four different field stations used in the study: JPL's instrumented buoys over Lake Tahoe (upper left), KIT's 30 m tower located at Gobabeb, Namibia (upper right), and two stations from NOAA's SURFRAD network—Bondville, IL (bottom left) and Desert Rock, NV (bottom right).

(EUMETSAT) (Göttsche et al., 2013). The Gobabeb station was specifically developed to validate LST products derived from the Spinning Enhanced Visible and Infrared Imager onboard Meteosat Second Generation (MSG/SEVIRI). The site represents a large and homogeneous gravel plain sparsely covered by dry grass. The core instruments are KT-15.85 IIP infrared self-calibrating radiometers, commercialized by Heitronics Infrarot Messtechnik GmbH in Wiesbaden, Germany. The KT-15.85 IIP measures thermal infrared radiance in the 9.6–11.5 μm domain and obtains brightness temperatures with an absolute accuracy of ±0.3 K (Theocharous, Usadi, & Fox, 2010). Two KT-15 sensors with a field of view (FOV) of 8.5° are mounted next to each other at 25 m height (Fig. 1) and observe an area of about 14 m^2 each. An additional KT-15 faces the sky at 53° with respect to zenith and measures the spectral downwelling longwave radiance, which is used to correct for the reflected atmospheric component (Eq. 2). All station measurements are collected once per minute. The surface emissivity of the gravel plain is considered constant and is assumed by LSA SAF to be 0.95 for MSG/SEVIRI at 10.8 μm, which is in close agreement with ASTER and MOD21 emissivities; this emissivity value has been validated with additional field measurements (Göttsche & Hulley, 2012).

5.3. The Surface Radiation Budget Network (SURFRAD)

The Surface Radiation Budget Network (SURFRAD) was established in 1993 with a primary objective of supporting climate research with accurate, continuous, long-term measurements of the surface radiation budget over the United States in support of the global Baseline Surface Radiation Network (BSRN) (Augustine, DeLuisi, & Long, 2000, Augustine, Hodges, Cornwall, Michalsky, & Medina, 2005). The seven SURFRAD stations used in the study operate in climatologically diverse regions and are representative of various land cover types (Table 4). Quality-controlled measurements of all relevant radiative components (upwelling and downwelling, solar and infrared, solar direct and diffuse, photosynthetically active radiation, solar ultraviolet-B radiation), and meteorological parameters are measured. SURFRAD stations are programmed to sample at 15-second intervals and to provide 1-minute averages of each parameter. SURFRAD instruments are meticulously maintained, and all instruments are replaced on an annual basis with freshly calibrated instruments. The primary measurements used in this study to derive ground-based LST are the upwelling and downwelling thermal infrared radiances, which are measured by two pyrgeometers (Eppley Precision Infrared Radiometer, spectral range 3.5–50.0 μm). The accuracy of the Eppley pyrgeometer is about 4.2 W m^{-2}, and the precision of the instrument is less than 1 W m^{-2} for nighttime measurements and around 2 W m^{-2} for daytime measurements (Philipona et al., 2001). The spatial representativeness of the pyrgeometer measurements is around 70 × 70 m^2. The surface emissivity of each site used to derive in situ LST (Eq. 2) is estimated from a spectral-to-broadband relationship (Eq. 6) from Ogawa, Schmugge, and Rokugawa (2008) using ASTER spectral emissivity products (Hulley & Hook, 2009b) (Table 4):

$$\varepsilon_{8-13.5\mu m} = 0.026\varepsilon_{11} + 0.269\varepsilon_{12} + 0.357\varepsilon_{13} + 0.359 \quad (6)$$

where ε_{11}, ε_{12}, ε_{13} are the ASTER-derived spectral emissivity values in bands 11, 12 and 13 centered on 8.65 μm, 9.1 μm and 10.6 μm. The emissivity integrated between 8 and 13.5 μm from Eq. (7) is assumed to be the best estimate of the broadband emissivity (Cheng, Liang, Yao, & Zhang, 2013), and is used here to represent the pyrgeometer spectral domain. The instrumental error alone gives rise to an uncertainty in retrieved LST of less than 1 K (Guillevic et al., 2012). Measurements from SURFRAD have already been used by Guillevic et al. (2012), Heidinger, Laszlo, Molling, and Dan Tarpley (2013), Wang et al. (2008) and Wang and Liang (2009) for evaluating ASTER, GOES and MODIS LST products, for example.

6. Results

6.1. Comparisons with satellite heritage data

To evaluate the relative agreement between VIIRS and MODIS LST products, the matchup tool described in Section 2 was used to select 44 Simultaneous Nadir Overpasses (SNOs) (37 daytime and 7 nighttime SNOs) of VIIRS and Aqua MODIS granules acquired from August 2012 to July 2013, which is within the VIIRS LST beta version release period, over the contiguous United States. Based on similar assumptions to describe the atmospheric effects, i.e. split window methods, VIIRS and MODIS (MYD11) LST products are generally in relatively good agreement, especially at nighttime (Fig. 2). The observed bias, standard deviation (STD) and root mean square (RMS) of the differences between the two LST products (VIIRS minus MYD11) are around 2.2 K, 2.3 K and 3.2 K respectively for daytime observations and 0.3 K, 1.1 K and 1.1 K for nighttime observations when accounting for all SNOs identified in the selected period of time (see Figs. 3 and 4, and Table 5 for detailed statistics depending on surface type). Regardless of surface type, better agreement between the satellite products is seen at night with the bias and STD being lower than 0.8 K and 1.7 K, respectively. During the daytime, significant discrepancies are observed over sparsely vegetated areas (open shrublands, savannah) and barren soils, for which the RMS of the difference is up to 5 K. Large biases of around 3.2 K are observed for broadleaf deciduous forest where the vegetation density varies with seasons. However, with non-systematic occurrence, differences in LST products are significantly high, up to 15 K, over barren surfaces or open shrublands for specific dates. This effect is clearly shown when analyzing geographical features in maps of differences (VIIRS minus MODIS LST) calculated for four dates associated with different atmospheric conditions over the southwestern USA and northwestern Mexico: hot and wet in August 2012, cool and dry in October 2012, hot and dry in June 2013 and cool and wet at nighttime in August 2012 (Fig. 2). On August 11, 2012, differences in LST up to 15 K are observed over barren surface and open shrublands (Figs. 2 and 3).

Results obtained for barren surfaces show evidence of two distinct trends between the VIIRS and MODIS LST and the agreement between the products is strongly degraded for LST values higher than 330 K. However, over the same region but for different dates, discrepancies in the products are significantly reduced and negligible at nighttime (Fig. 4). The observed differences were not present in the VIIRS Sensor Data Record (SDR) (equivalent to MODIS level-1b products). On the August 11, 2012, the date on which maximum discrepancies were observed, the at-sensor radiances measured in the 11 μm and 12 μm bands of VIIRS and MODIS (VIIRS bands M15 and M16 and MODIS bands 31 and 32) were in good agreement (Fig. 5). The bias and standard deviation calculated between VIIRS and MODIS at-sensor brightness temperatures differ by 0.45 K and 1.8 K and cannot explain the differences observed when comparing the level-2 (swath-based) LST products. The higher bias obtained between VIIRS channel M15 and MODIS band 31 may be partly explained by slight differences in spectral domains. Since our use of SNOs minimizes the impact of directional effects and spatial resolution on the differences between VIIRS and MODIS at-sensor radiances, the discrepancies in LST are mainly introduced by the retrieval algorithms. As such, either one or both split-window algorithms do not accurately represent the wide range of natural surface variability and atmospheric conditions.

The effect of emissivity is also not predominant, or it would have been present in all granules since both algorithms are using classification of constant emissivity values for each surface type. However, we note on August 11, 2012 that differences in LST products are well correlated with the spatial distribution of atmospheric total column water vapor contents derived from MODIS (MYD07) (Fig. 6, upper panel). Differences in LST higher than 10 K and up to 15 K were observed over the Sonoran Desert, which covers large parts of the southwestern United States and of northwestern Mexico, where the MOD07 water

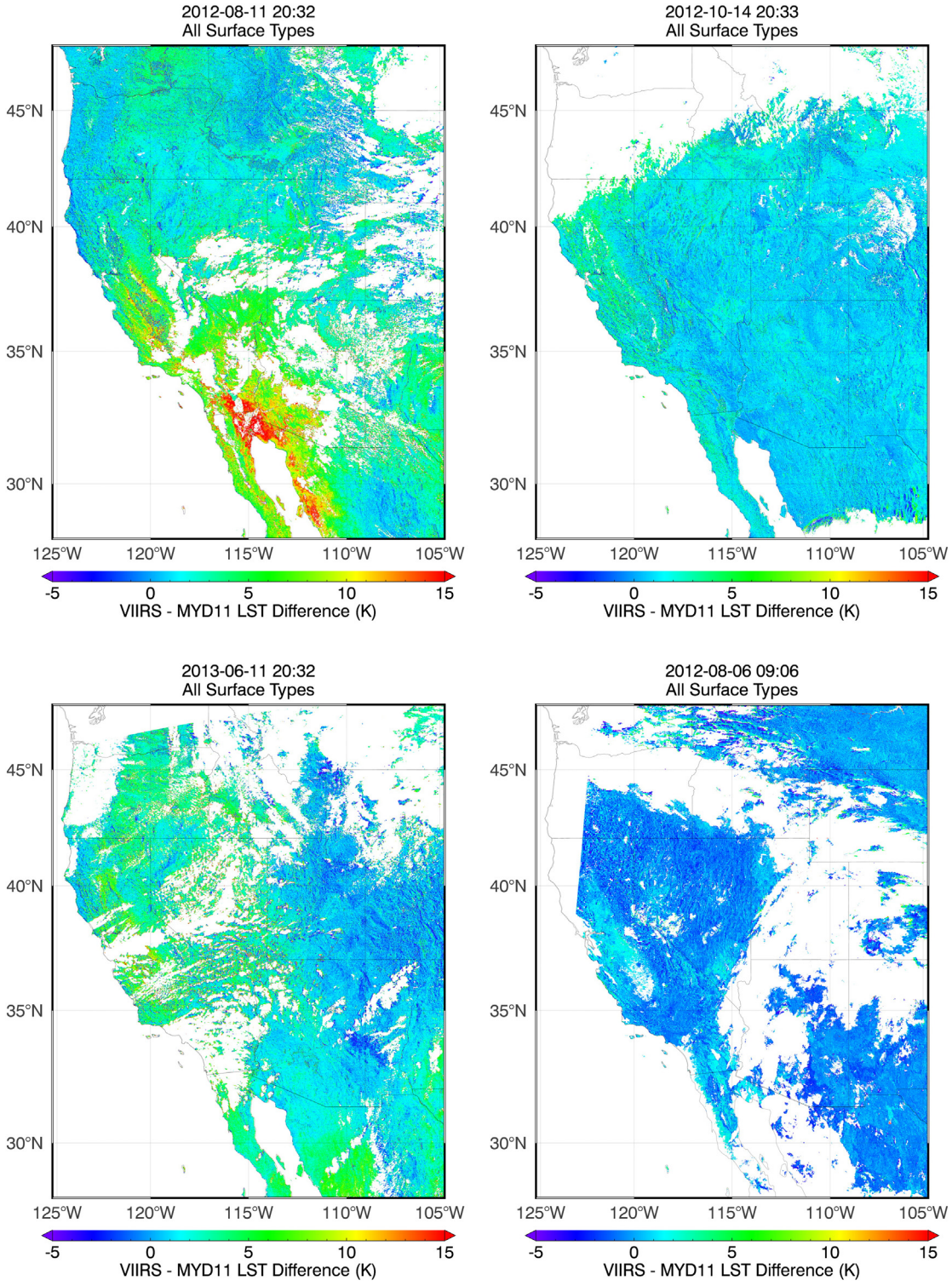

Fig. 2. Differences between VIIRS and MODIS (MYD11) LST products observed over the western USA on four different dates associated with different atmospheric conditions: hot and wet in August 11, 2012, cool and dry in October 14, 2012, hot and dry in June 11, 2013 and cool and wet at nighttime in August 6, 2012. The white areas over land are regions where good-quality retrievals were not available (clouds, etc.).

vapor content was higher than 4 cm and the LST was higher than 340 K (Fig. 6, bottom panel). As expected, the difference between VIIRS brightness temperatures (M15 minus M16) increases when the atmospheric water content increases (Fig. 7) due to higher absorption in band M16 than M15, and the split window methods exploit such a relationship to remove the atmospheric effects from at-sensor radiances and retrieve the LST. Ideally, the atmosphere should have a very small influence on validation results. However, on August 11, 2012, the differences between VIIRS and MODIS LST strongly increased from 5 to 15 K when the atmospheric total column water vapor content increased from 3 to 5 cm over regions with very warm surface temperatures (Fig. 8).

For very humid atmospheric conditions, the differences between the VIIRS brightness temperatures were around 6 K (Fig. 7). Such values strongly enhance the contribution of the quadratic term in the VIIRS

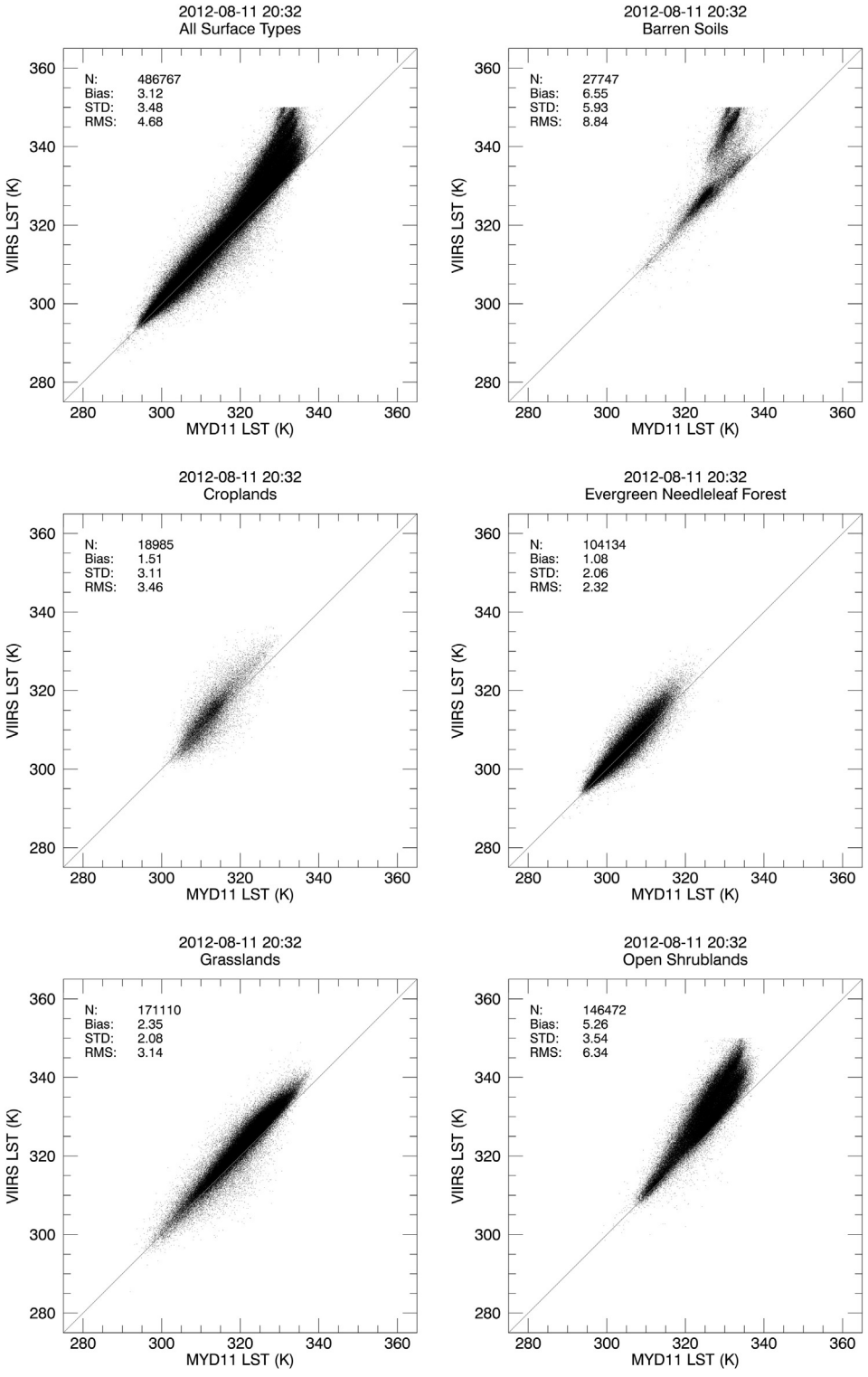

Fig. 3. VIIRS vs. MODIS (MYD11) LST products observed over the western USA on August 12, 2014 at 20:32 UTC (daytime). Results obtained for all 17 IGBP surface types (upper left) and selected surface types are represented. Large differences are observed over barren soils.

algorithm formulation (Eq. 4). The above results provide evidence that the coefficient a_4 in the split window formulation used for VIIRS must be carefully defined. In the current VIIRS algorithm this is attempted with a single set of coefficients which, therefore, must incorporate the algorithm's entire knowledge about all globally possible daytime atmospheric conditions. This study indicates that a single set of coefficients is insufficient to achieve the required product accuracy. Moreover, the current set of VIIRS coefficients is derived from regressions over an ensemble of radiative transfer simulations that accounts for LST values up to 327 K maximum, but simulations associated with LST values above 300 K are not statistically represented. Results suggest that the performance of the VIIRS algorithm is significantly degraded for very humid atmospheres, and over warm arid regions for which surface air temperature and LST may differ considerably, e.g. by more than 20 K.

The poor performance of VIIRS LST for arid surface conditions and high water vapor atmospheres has been verified with a third

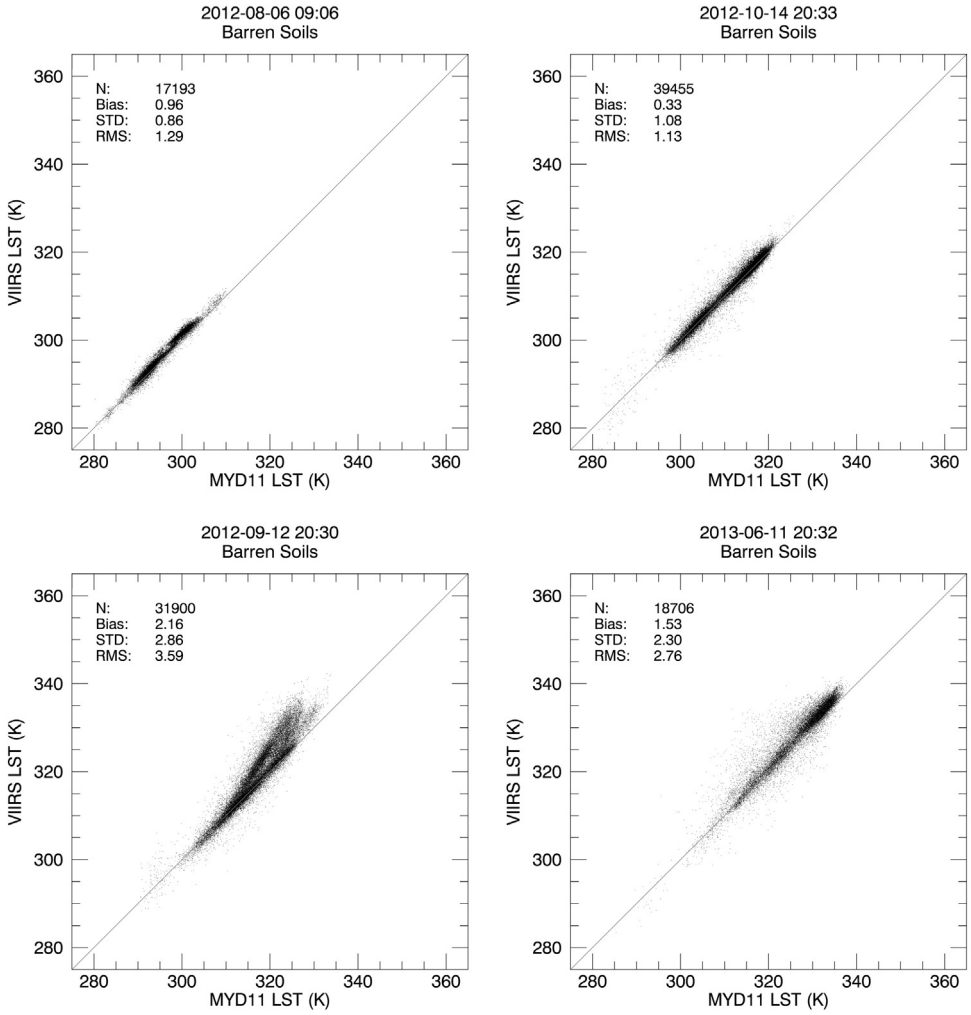

Fig. 4. VIIRS vs. MODIS (MYD11) LST products observed over barren areas in the western USA for different dates at 9:06 UTC (night) and around 20:30 UTC (daytime) associated with different atmospheric water vapor content.

Table 5
Comparisons of coincident VIIRS and MODIS LST products acquired from August 2012 to July 2013 over contiguous United States during daytime (37 VIIRS granules) and nighttime (7 VIIRS granules). Bias, standard deviation (STD) and root mean square (RMS) of the differences (VIIRS minus MODIS LST) are represented. No statistics were calculated for number of coincident pixels (N) lower than 100. The description of the 17 IGBP surface types is provided in Table 1.

Surface type	Daytime				Nighttime			
	N	Bias	STD	RMS	N	Bias	STD	RMS
1	1207066	1.23	2.12	2.45	106924	0.07	1.31	1.31
2	636	1.31	1.49	1.98	–	–	–	–
3	–	–	–	–	–	–	–	–
4	260829	3.23	1.08	3.41	1523	0.50	0.88	1.01
5	422628	1.28	1.33	1.84	41372	0.09	1.12	1.12
6	32933	1.06	2.01	2.27	5637	0.22	0.86	0.89
7	3072468	2.64	2.74	3.80	268565	−0.08	1.10	1.10
8	235621	2.16	2.31	3.17	11416	0.01	0.94	0.94
9	21248	4.54	3.02	5.45	–	–	–	–
10	4789463	2.57	2.11	3.33	905205	0.39	0.95	1.03
11	871	1.28	2.35	2.68	–	–	–	–
12	2033586	1.97	2.12	2.89	148459	0.37	0.90	0.97
13	126656	2.07	2.02	2.89	4771	−0.35	1.12	1.18
14	542517	1.49	1.59	2.18	9377	0.04	1.07	1.07
15	–	–	–	–	–	–	–	–
16	446044	3.06	3.83	4.90	24071	0.72	0.95	1.19
17	417660	0.56	1.05	1.19	81904	0.76	1.65	1.82
All	13610262	2.25	2.36	3.26	1609259	0.30	1.07	1.12

independent dataset: the MYD21 LST products (Hulley & Hook, 2010) that will be released by NASA in 2014. MYD21 LST products are derived using the Temperature and Emissivity Separation (TES) algorithm (Gillespie et al., 1998) modified for MODIS bands 29 (8.55 μm), 31 (11 μm), and 32 (12 μm). Initially developed for ASTER, the TES algorithm is now used to derive LST from MODIS (Hulley & Hook, 2010), and SEVIRI (Jimenez-Munoz, Sobrino, Mattar, Hulley, & Gottsche, 2014). TES is a physically based algorithm and retrieves coherent LST and spectral surface emissivity products using observations of atmospheric profiles, i.e. total water vapor content and temperature, and a full radiative transfer model (MODTRAN). Both VIIRS and MYD11 LST for August and October 2012 were compared with MYD21 LST (Fig. 9). The split window methods, using fixed emissivity values, generally underestimate the LST over very warm arid regions. The observed discrepancies significantly increase under warm conditions when the difference between the thermal infrared radiance from the surface and the atmosphere increases. This behavior is clearly observed between MYD11 and MYD21 products for most atmospheric conditions. However, it is not observed for VIIRS LST under humid atmospheric conditions where the effect of the atmosphere is predominant. When the difference between the thermal infrared radiance from the Earth's surface and from the atmosphere is small (e.g., at nighttime and during cool weather), the emissivity effect is partially compensated by the reflected downward atmospheric radiation. Comparisons between VIIRS and

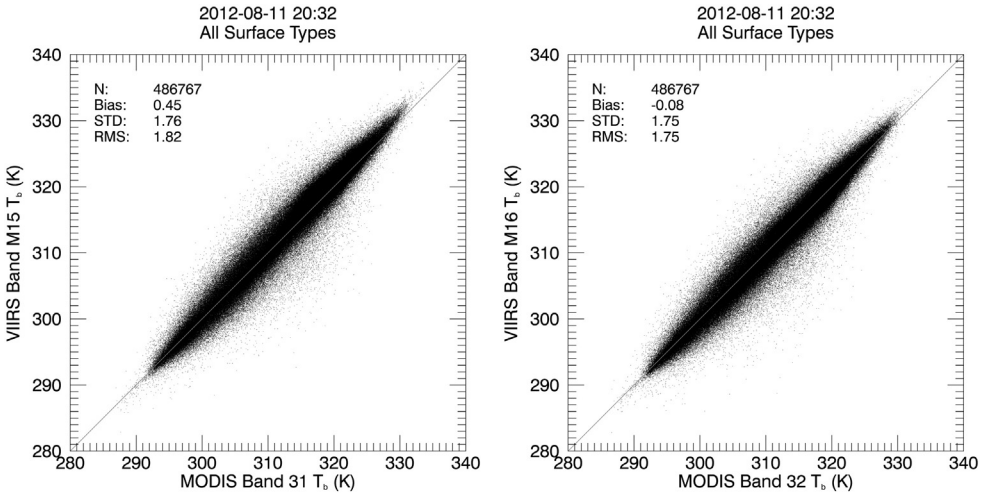

Fig. 5. VIIRS vs. MODIS brightness temperatures (T_b) centered around 11 μm (left: VIIRS band M15 vs. MODIS band 31) and 12 μm (right: VIIRS band M16 vs. MODIS band 32) measured over the western USA on August 11, 2012.

MODIS products showed that the VIIRS LST is systematically underestimated over barren surfaces and significantly overestimated for warm surfaces and wet atmospheres.

6.2. Comparisons with ground-based measurements

6.2.1. Characterization of the spatial representativeness of ground-based LST using ASTER

LST and emissivity products from ASTER with a spatial resolution of 90 m were used to quantify the spatial variability of LST around the SURFRAD stations. The stations located at Lake Tahoe (inland water) and Gobabeb, Namibia (desert) are very homogeneous in terms of land cover type and surface biophysical properties and are, therefore, not considered here. The coarser spatial resolution of the VIIRS and MODIS pixels was estimated by aggregating 90 m ASTER pixels to form 1 km pixels centered on each SURFRAD station. We used LST and surface emissivity standard deviation calculated for subsets of 11×11 ASTER pixels to assess the spatial representativeness of the station measurements (Table 6). All available clear-sky ASTER granules were selected to calculate the statistics. With the exception of the Sioux Falls site, which represents a patchwork of land cover types (grassland, lake, urban), the sites are characterized by relatively small emissivity differences (standard deviation around 0.01). Two sites, Fort Peck, MT and Desert Rock, NV, appeared more spatially homogeneous than the other sites, with median and maximum values of LST STD around 1.3 K and 2.3 K respectively. The sites located near Sioux Falls, SD, Table Mountain, CO and Bondville, IL, are more heterogeneous and exhibit a higher spatial variability of both LST and surface emissivity. Daytime and nighttime ASTER data were used, with a higher variability usually measured during daytime. These results clearly suggest that over heterogeneous areas the quality of validation results based on comparisons with ground-based LST strongly depends on the geolocation accuracy, the satellite sensor's footprint, and point spread function. A very limited number of ASTER cloud free scenes were available for each site after August 2012, and ASTER data were not directly used in the validation process. The surface emissivity used in the upscaling model was derived from model optimization to account for seasonal variations, as described by Guillevic et al. (2012).

6.2.2. VIIRS LST EDR vs. ground-based LST

To evaluate the performance of the VIIRS algorithm, the LST standard EDR, retrieved with the single split window method (the baseline algorithm for VIIRS since August 11, 2012), was compared with ground-based LST obtained at nine different stations. If necessary, the ground-based LST was scaled up to account for LST spatial variability around the stations. Since the sites located in Gobabeb, Namibia and Lake Tahoe, California are homogeneous at the scale of MODIS and VIIRS, the ground-based LST did not need to be adjusted for site inhomogeneity. All data used in the analysis are publicly available and include quality assurance metrics.

The comparisons of LST data were undertaken from August 2012 to October 2013 using VIIRS, NOAA's SURFRAD and JPL network ground stations, with the exception of the Gobabeb station for which only 5 months of data were available. Satellite overpasses with no obvious clouds and quality flags provided with the VIIRS LST EDR were used to select matchups during clear days. The upscaling model was specifically developed for VIIRS LST validation by Guillevic et al. (2012). The first step of the scaling represents the calibration/optimization of the SEtHyS land surface model using in situ measurements. This was performed for each site over 10-day periods in spring or summer with very low cloud coverage. The scaling is performed by simulating the LST spatial distribution around the station due to surface spatial variability. The calibration methodology and the relationship between MODIS-based NDVI and LAI at 250 m spatial resolution around the stations are also described in Guillevic et al. (2012). The scaling model accounts for the view angle dependency of the VIIRS footprint, which varies discontinuously from 0.75 km at nadir to around 1.5 km at the edge of the scan.

Validation results with in situ measurements show that the VIIRS algorithm performs well over most vegetated and inland water surfaces (Fig. 10 and Table 7). The accuracy (bias between satellite and ground-based LST) and precision (standard deviation of the differences) of the VIIRS LST EDR required by the JPSS program are 1.5 K and 2.5 K, respectively (Justice et al., 2013). With the exception of the desert site in Gobabeb, where VIIRS does not meet the requirements, the absolute value of the bias between VIIRS LST and in situ LST calculated for daytime and nighttime data varies from 0.16 K to 0.61 K when accounting for scaling effects. The standard deviation of the differences between VIIRS and in situ LST varies from 1.80 K to 2.65 K over land sites, with values slightly above the requirements for Bondville, IL (where the STD of the differences is 2.65 K). Over the vegetated SURFRAD sites, VIIRS systematically overestimated the LST.

Differences between VIIRS and ground-based LST products can mainly be attributed to low spatial representativeness of in situ measurements due to surface heterogeneities, insufficient stratification in VIIRS algorithm coefficients as needed to represent a wide range of global surface conditions, large errors in surface emissivity values for bands M15 and M16 estimated from land-cover types, or due to the presence of clouds. Maximizing the data quality by using strong cloud filtering significantly reduces the number of observations available but minimizes the effect of cloud contamination in validation results.

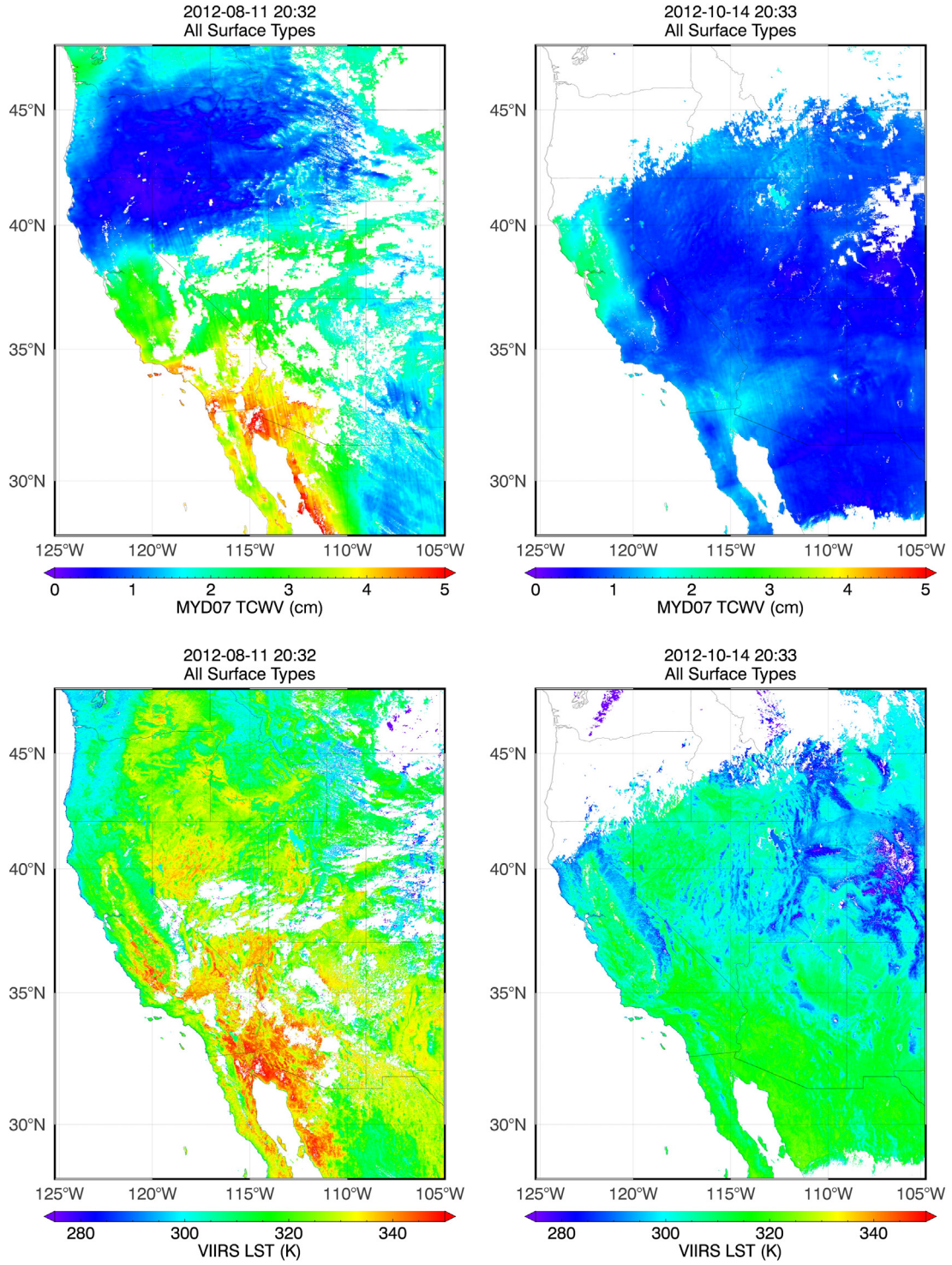

Fig. 6. Total column water vapor content (TCWV) products derived from MODIS (MYD07) (upper), and VIIRS LST products (bottom) on August 11, 2012 (left) and October 14, 2012 (right).

The effect of spatial variability is usually more significant during daytime than nighttime (Table 7), since effects of structural shading, evaporative cooling and surface-air temperature differences are greater during the day than at night. For example, over vegetated areas not subject to water stress, the transpiration of plants amplifies thermal gradients between areas with different vegetation types and densities. Over croplands near Bondville, IL or sparsely distributed grassland near Table Mountain, CO, the scaling model is able to account for the high bias observed between VIIRS and in situ LST at daytime—a bias of around 2.7 K at Table Mountain and 1.9 K at Bondville without scaling (see Table 6 for the results achieved with and without scaling). The Bondville station is located on a 250 m × 250 m patch of grass surrounded by crops. The effect of the crops on validation results directly depends on the season, soil wetness (and resulting transpiration) and the maturity of the crops, which explains why the VIIRS LST is much lower than the station LST when the plants (corn and soybeans) are well developed and why the VIIRS LST is significantly higher after the harvest. This scaling trend has been shown by Wang et al. (2014) when validating other surface radiation products. Even if the Bondville site is well instrumented and maintained, the measurement is only

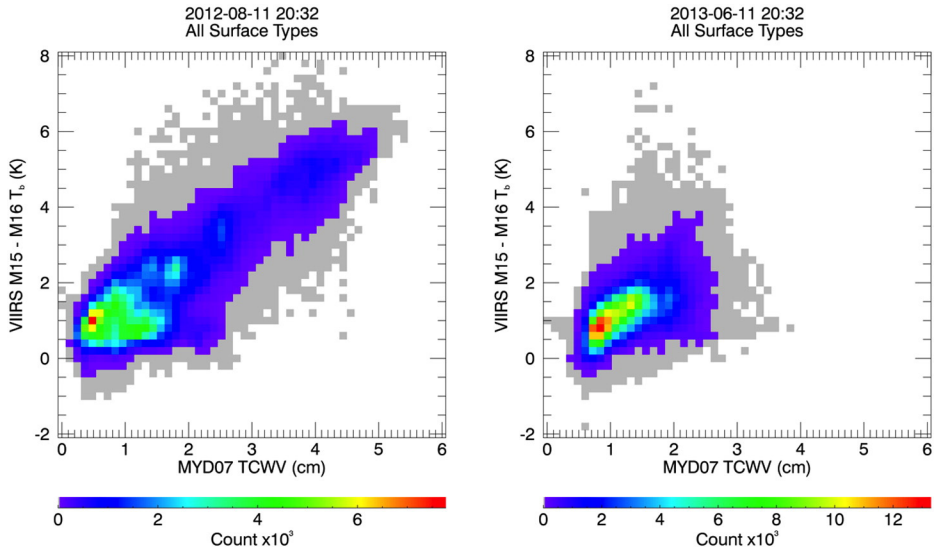

Fig. 7. Difference between VIIRS band M15 and M16 brightness temperatures (T_b) vs. MODIS-derived total column water vapor content (TCWV) observed on different dates associated with different atmospheric conditions: hot and wet in August 11, 2012 (left), and hot and dry in June 11, 2013 (right).

representative of about one tenth of the VIIRS footprint, which makes it not well suited for quantitatively estimating the quality of satellite products at moderate resolution due to seasonal spatial variability effects.

The best validation results were at Lake Tahoe (Table 7) where the bias and STD of the differences between VIIRS LST and LST measured over Lake Tahoe are lower than 0.15 K and 0.4 K, respectively. The validation was performed using day and night in situ data collected by four different buoy stations over 15 months. The results are highly consistent both spatially (between the stations) and temporally (seasons, day or night) and clearly suggest that in situ measurements over water bodies provide the level of accuracy needed for sensor calibration and temporal drift detection. However, validation over water cannot estimate the errors in the surface emissivity assumptions or modeling inherent in an LST algorithm.

In the previous section, we showed that VIIRS and MODIS MYD11 LST products are quite consistent over contiguous United States for relatively cool and low atmospheric water vapor content conditions. Similar results have been found at Gobabeb in Namibia where the bias between the VIIRS and MODIS LST is lower than 0.5 K (Fig. 11). However, both VIIRS and MODIS products significantly underestimate the LST of the Namibian desert by more than 4 K when compared with ground-based reference measurements (Fig. 11 and Table 7). Previous studies over arid and semi-arid validation sites (Justice et al., 2013; Li et al., 2014) have also shown a systematic underestimation of VIIRS LST by up to 5 K due to overestimation of spectral emissivity values used in the VIIRS algorithm for barren surfaces. At Gobabeb, VIIRS and MYD11 emissivities are up to 3% higher than in situ measurements (Göttsche & Hulley, 2012), which already explains a large part of the observed biases. The LST underestimation by MODIS (MOD11 and MYD11 products) was first described by Wan (2008) and Hulley and Hook (2010), and the MODIS algorithm in collection 6 has been modified to reduce this effect (Wan, 2014). In contrast, the VIIRS LST errors under such conditions have not yet been resolved. These results illustrate that validation should not be based on comparisons with existing satellite data alone.

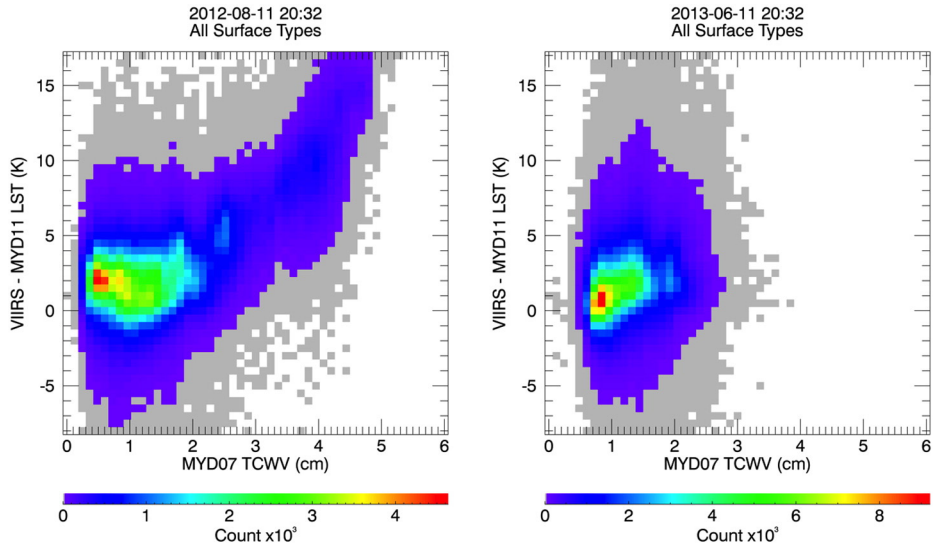

Fig. 8. Differences between VIIRS and MODIS (MYD11) LST products vs. MODIS-derived total column water vapor content (TCWV) observed on different dates associated with different atmospheric conditions: hot and wet in August 11, 2012 (left), and hot and dry in June 11, 2013 (right) at 20:32 UTC.

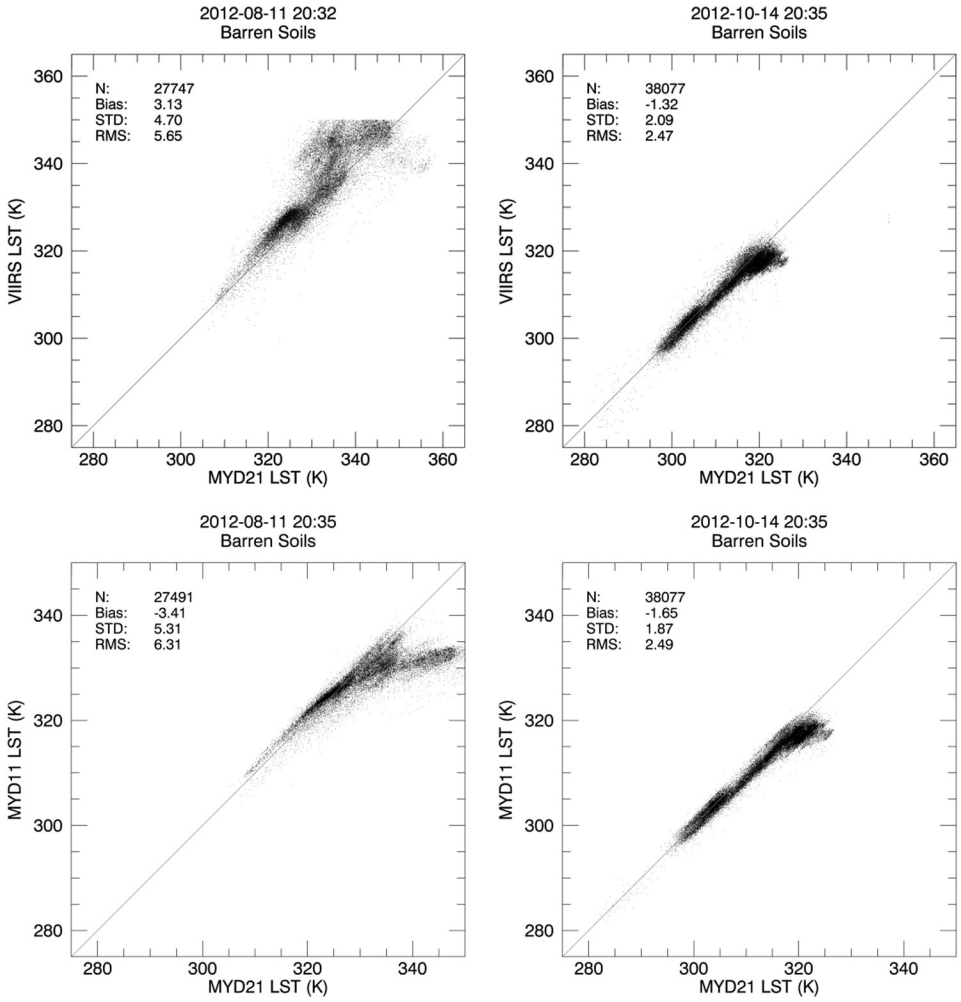

Fig. 9. VIIRS vs. MODIS (MYD21) LST products (upper) and MYD11 vs. MYD21 LST products (bottom) observed over barren areas on August 11, 2012 (left) and October 14, 2012 (right).

7. Discussion

Results have shown that the two validation methods used in the study, namely the comparisons with ground-based and between different satellite products complement each other and are both useful for characterizing the performance and limitations of the VIIRS LST retrieval algorithm. Comparisons with ground-based LST measurements are the most traditional and well-used approaches, and in most cases they provide suitable validation results for well-defined and dedicated sites. However, because of the limited number of high quality sites with spatially homogeneous surface temperatures, additional methods must be employed to characterize algorithm performance over the full range of surface types and conditions. In our study, for example, in situ validation was not able to detect the strong degradation of the VIIRS algorithm performance under high atmospheric water vapor content.

Such discrepancies were only successfully diagnosed by comparing VIIRS and MODIS swaths. However, the characterization of uncertainties on existing LST satellite products would not have been possible without independent and traceable sites, or from rigorous simulations (Hulley, Hughes, & Hook, 2012). Satellite product inter-comparisons provide useful information about spatial patterns in LST deviations and can be used to monitor the performance of a product and detect significant quality issues. However, they cannot provide rigorous and quantitative validation results. The use of the MYD21 LST product, which is based on other assumptions than the split window methods, showed that VIIRS was underperforming with large cold biases over arid regions far exceeded VIIRS accuracy requirements. Moreover, comparisons with in situ measurements at Gobabeb showed that coincident VIIRS and MODIS LST split-window products can be very consistent in a relative manner but nonetheless may both be biased with respect to an independent ground reference data.

Accurate in situ observations of LST at dedicated validation sites provide the most reliable validation results. However, due to their high cost and logistical barriers, there are only a few sites around the world that are dedicated to LST validation. Two of these sites were used in this study—one located in a desert area in Namibia operated by KIT and

Table 6
Spatial representativeness of SURFRAD ground-based measurements for comparisons with VIIRS LST products at moderate resolution. Subsets of 11 × 11 ASTER pixels at 90 m centered on station locations are used to calculate the standard deviations (STD) of LST and surface emissivity around the stations. N represents the number of clear-sky ASTER granules used to assess the median, minimum (Min) and maximum (Max) values of STD.

Site location	N	STD of LST (K)			STD of emissivity (−)		
		Min	Median	Max	Min	Median	Max
Table Mountain, CO	28	0.43	2.46	4.81	0.006	0.010	0.016
Bondville, IL	52	0.28	0.76	4.02	0.006	0.013	0.023
Goodwin Creek, MS	26	0.62	2.06	4.27	0.007	0.009	0.016
Fort Peck, MT	9	0.40	1.26	2.18	0.007	0.010	0.014
Desert Rock, NV	69	0.36	1.01	2.38	0.004	0.005	0.012
Penn State U., PA	23	0.49	1.38	2.57	0.007	0.010	0.018
Sioux Falls, SD	9	0.92	2.90	4.82	0.014	0.018	0.024

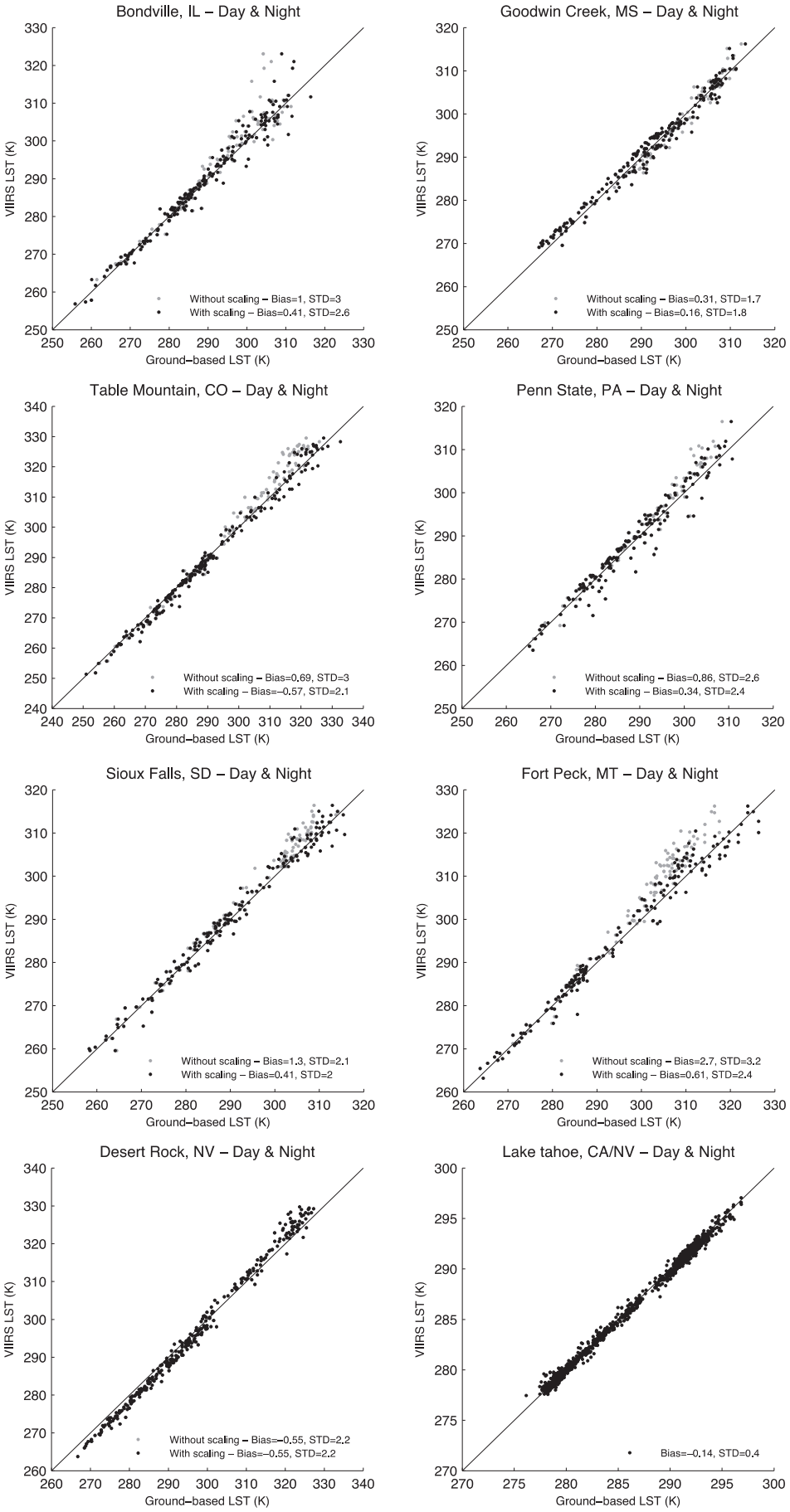

Fig. 10. VIIRS LST EDR vs. ground-based LST measurements at eight validation sites representative of various land surface types over the contiguous United States.

Table 7
Bias, standard deviation (STD) and root mean square error (RMSE) of the differences calculated between VIIRS LST EDR and ground-based LST with and without scaling. Results over the validation datatset, daytime periods and nighttime periods are represented. Statistics account for all clear-sky data from August 11, 2012 to October 16, 2013. No scaling was carried out over Lake Tahoe and Gobabeb sites.

Period		N	Satellite vs. non scaled LST			Satellite vs. scaled-up LST		
			Bias	STD	RMSE	Bias	STD	RMSE
Lake Tahoe	All	1073	−0.14	0.40	0.42	–	–	–
	Daytime	524	−0.12	0.43	0.45	–	–	–
	Nighttime	549	−0.16	0.36	0.39	–	–	–
Table Mountain	All	210	0.59	2.99	3.04	−0.57	2.16	2.23
	Daytime	89	2.69	3.34	4.27	−0.10	2.81	2.80
	Nighttime	121	−0.99	1.23	1.58	−0.92	1.42	1.69
Bondville	All	195	1.00	3.03	3.18	0.41	2.65	2.67
	Daytime	96	1.87	3.93	4.34	0.60	3.52	3.55
	Nighttime	99	0.16	1.30	1.31	0.22	1.33	1.34
Goodwin Creek	All	199	0.31	1.17	1.72	0.16	1.80	1.80
	Daytime	98	−0.40	1.76	1.80	−0.69	1.84	1.96
	Nighttime	101	1.00	1.30	1.64	1.00	1.30	1.65
Fort Peck	All	174	2.65	3.19	4.14	0.61	2.44	2.51
	Daytime	94	4.66	2.86	5.46	0.97	3.00	3.13
	Nighttime	80	0.29	1.50	1.52	0.19	1.48	1.48
Desert Rock	All	309	−0.55	2.21	2.27	−0.55	2.21	2.27
	Daytime	134	1.41	1.89	2.35	1.41	1.89	2.35
	Nighttime	175	−2.05	0.81	2.20	−2.05	0.81	2.20
Penn State U.	All	162	0.86	2.55	2.69	0.34	2.41	2.43
	Daytime	62	2.06	2.78	3.45	0.90	2.77	2.89
	Nighttime	100	0.11	2.09	2.08	0.00	2.10	2.08
Sioux Falls	All	172	1.28	2.06	2.41	0.41	1.96	1.99
	Daytime	81	2.27	2.02	3.03	0.50	2.31	2.35
	Nighttime	91	0.41	1.66	1.70	0.31	1.59	1.61
Gobabeb	All	252	−4.56	2.54	5.21	–	–	–
	Daytime	124	−5.53	2.93	6.25	–	–	–
	Nighttime	128	−3.62	1.63	3.96	–	–	–

one on Lake Tahoe operated by JPL. Both sites have provided the most reliable quantitative validation results. The Gobabeb desert site, in particular, was useful for assessing the negative bias of VIIRS LST over arid areas due to an overestimation of emissivity. Comparisons at the inland water sites also provide evidence of good overall quality of the VIIRS radiometric measurements. Validation results at Lake Tahoe are considerably less impacted by the cumulated and diffused effects of spatial variability, atmospheric effects, sun illumination, viewing angles and surface emissivity uncertainties. In situ LST retrieved over water is associated with low experimental uncertainties, and is very stable over time. Inland water measurements provide the necessary quality to perform long-term LST product validation and radiometric calibration drift monitoring.

Although the SURFRAD network was not initially designed for LST validation, SURFRAD measurements are useful for validating satellite LST products (Guillevic et al., 2012; Heidinger et al., 2013; Wang & Liang, 2009; Wang et al., 2008). According to results obtained for VIIRS in this study, or for MODIS by Guillevic et al. (2012), we strongly recommend that ground-based LST validation only be performed over sites that are highly homogeneous in surface temperature. High-resolution LST or NDVI datasets can be used to select appropriate validation sites. In this study, we used simple statistics to analyze the spatial representativeness of in situ LST measurements, but more sophisticated approaches exist. For example, Roman et al. (2009) has introduced the use of geostatistical metrics (e.g., empirical variograms) to estimate the spatial variability of satellite-derived surface albedo around locations of interest. Note that currently no operational satellite LST algorithms correct for view angle (e.g. normalization to nadir angle). Therefore, some systematic errors are typically unavoidable when comparing satellite data to nadir-pointing field instruments.

Validation based on comparisons with existing satellite products requires careful identification of appropriate matchups between the different datasets. Previous multi-sensor comparison studies (Guillevic et al., 2012; Trigo et al., 2008) found differences up to 12 K between MODIS and SEVIRI-derived LST over sparsely vegetated woodlands due to directional effects. Pinheiro et al. (2006) found similar differences in directional field radiometer measurements over African woodlands. Appropriate matchups significantly reduce the discrepancies induced by directional effects such as shadows and variable footprint size. However, the impact of differences in spatial resolutions, spatial weighting and spectral responses on validation results is difficult to assess and cannot be completely reduced with matchups. For example, a systematically higher VIIRS LST product (relative to MODIS LST) may reflect a possible issue in VIIRS, MODIS or both retrieval algorithms. However, the global RMSE of the differences is around 3.2 K during daytime and 1.1 K during nighttime, and may reflect intrinsic differences in sensor characteristics and experimental/matchup design.

8. Conclusions

Through comparisons with Aqua/MODIS LST products and ground-based measurements, two problems were identified in the VIIRS LST EDR:

A systematic underestimation of LST over barren surfaces due to inaccurate spectral emissivity values used in the VIIRS algorithm. Surface emissivity ancillary data used in the algorithm are derived from fixed

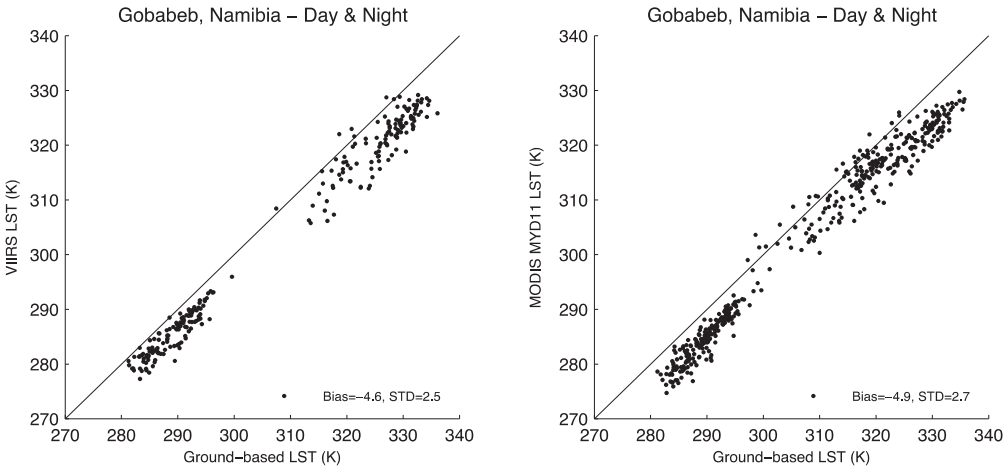

Fig. 11. VIIRS (left) and MODIS MYD11 (right) LST products vs. ground-based LST measurements at Gobabeb, Namibia. Due to an overestimation of surface emissivity values used in the algorithms, both VIIRS and MODIS products significantly underestimate the LST of the Namibian desert by more than 4 K on average.

values depending on a limited number of surface types and do not fully encompass the natural variation in surface emissivity. This represents a well-known intrinsic problem associated with split window techniques over land that cannot be easily corrected for the current LST VIIRS algorithm. A possible solution for split window methods would be to explicitly account for surface emissivity values as parameters of the split window formulation (Wan and Dozier, 1997), and to use maps of retrieved dynamic surface emissivities (like those from ASTER-TES, MODIS MOD11B1 or MOD21 products) as algorithm inputs.

A strong overestimation of LST over arid and semi-arid regions under hot and very humid conditions due to the use of a non-representative sampling of environmental conditions, i.e., atmospheric conditions and surface temperatures, when determining the VIIRS algorithm coefficients. A better representation of extreme atmospheric water vapor conditions in the training dataset should improve the algorithm performance, especially over arid regions (bare surfaces and open shrubland according to the IGBP classification) when surface air temperature and LST have very large differences. Both effects are much greater in day than at night.

Results have shown that comparisons with existing satellite data and ground-based measurements are complementary for identifying and characterizing the limitations of the VIIRS—or any other—LST algorithm. For some of the cases investigated in this study, two different satellite LST products can be in very good agreement since they use a similar algorithm, however they may differ considerably from the corresponding ground-based reference measurements. Furthermore, it was shown that except for slight discrepancies due to differences in sensor-specific spectral functions, there is excellent agreement between at-sensor radiances measured by VIIRS (channels M15 and M16) and MODIS (bands 31 and 32), which demonstrates the high quality of the VIIRS thermal infrared level-1 product.

Because spatial representativeness and directional effects are quite difficult to compensate for, quantitative assessment of algorithm uncertainties requires dedicated and high quality in situ LST measurements over sites that are homogeneous at the spatial scale of the satellite observing system. Ground-based LST measurements over water bodies provide the most stable and reliable validation dataset for monitoring calibration drift. However, in order to evaluate LST algorithms over a representative range of land surface temperatures (e.g. from 223 K to 343 K) and for situations with strong surface overheating, validation over sites representing solid land surface types is indispensable.

Long-term and routine validation efforts require additional comparisons with other satellite-derived products, such as the physically-based MODIS MYD21, which provides consistent LST and surface emissivity (LST&E) accuracy over all land cover types. Continued validation of VIIRS LST EDRs using the radiance-based methods following the methodology defined by Wan and Li (2008) will provide validation over a more diverse set of conditions on a global scale. This is because in general it is easier to find sites that are homogeneous in emissivity than homogeneous in temperature at km-scale resolutions. Emissivities at these sites can be defined using either in situ measurements or from the ASTER Global Emissivity Database (ASTER-GED) developed by Hulley and Hook (2009b) for example. The VIIRS algorithm is currently unable to accommodate spatial and temporally-adjusted emissivity values since it does not produce a dynamically retrieved land surface emissivity product. Consequently, the performance of the algorithm is strongly reduced over arid and sparsely vegetated regions. However, the generation of a dynamic emissivity product would ensure continuity with existing LST&E products such as from the current MODIS MOD11B1 and MOD21 products, and would enable the development of merged products using both split-window and dynamic emissivity retrieval such as from the TES algorithm.

Acknowledgments

The research described in this paper was carried out at (1) the Cooperative Institute for Climate and Satellites–North Carolina under Cooperative Agreement NA09NES4400006 supported by the Joint Polar Satellite System program, and by the National Oceanic and Atmospheric Administration (NOAA)'s Climate Data Record project, and at (2) the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, under contract with the National Aeronautic and Space Administration (NASA). The VIIRS data are distributed by the NOAA's Comprehensive Large Array-Data Stewardship System (CLASS, http://www.class.ncdc.noaa.gov). The MODIS data used in the study are distributed by the NASA Earth Observing System Data and Information System (http://earthdata.nasa.gov). The in situ data from Gobabeb, Namibia, were collected within the context of the LSA-SAF (http://landsaf.ipma.pt), a project funded by the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT). The authors would like to thank Dr. Zhengming Wan from University of Santa Barbara, CA, USA, Dr. John Augustine from NOAA's Earth System Research Laboratory, Boulder, CO, USA and Dr. Alain Sei from Northrop Grumman Aerospace Systems, Redondo Beach, CA, USA for their scientific support regarding MODIS, SURFRAD and VIIRS products.

References

Anderson, M. C., Norman, J. M., Diak, G. R., Kustas, W. P., & Mecikalski, J. R. (1997). A two-source time-integrated model for estimating surface fluxes using thermal infrared remote sensing. *Remote Sensing and Environment*, *60*, 195–216.

Anderson, M. C., Hain, C. R., Wardlow, B., Mecikalski, J. R., & Kustas, W. P. (2011). Evaluation of a drought index based on thermal remote sensing of evapotranspiration over the continental U.S. *Journal of Climate*, *24*, 2025–2044.

Anderson, M. C., Allen, R. G., Morse, A., & Kustas, W. P. (2012). Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources. *Remote Sensing of Environment*, *122*, 50–65, http://dx.doi.org/10.1016/j.rse.2011.08.025.

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*, 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.

Bosilovich, M. G. (2006). A comparison of MODIS land surface temperature with *in situ* observations. *Geophysical Research Letters*, *33*, L20112, http://dx.doi.org/10.1029/2006GL027519.

Brutsaert, W. (1975). On a derivable formula for long-wave radiation from clear skies. *Water Resources Research*, *11*(5), 742–744.

Cheng, J., Liang, S., Yao, Y., & Zhang, X. (2013). Estimating the optimal broadband emissivity spectral range for calculating surface longwave net radiation. *IEEE Geoscience and Remote Sensing Letters*, *10*, 401–405.

Coll, C., Caselles, V., Galve, J. M., Valor, E., Niclòs, R., Sanchez, 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., Hook, S. J., & Galve, J. M. (2009). Land surface temperature from the advanced Along-Track Scanning Radiometer: Validation over inland waters and vegetated surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, *47*, 350–360.

Coll, C., Wan, Z., & Galve, J. M. (2009b). Temperature-based and radiance-based validations of the V5 MODIS land surface temperature product. *Journal of Geophysical Research, [Atmospheres]*, *114*, D20102, http://dx.doi.org/10.1029/2009JD012038.

Coll, C., Galve, J. M., Sanchez, J. M., & Caselles, V. (2010). Validation of Landsat-7/ETM+ thermal-band calibration and atmospheric correction with ground-based measurements. *IEEE Transactions on Geoscience and Remote Sensing*, *48*, 547–555.

Coudert, B., Ottlé, C., Boudevillain, B., Demarty, J., & Guillevic, P. (2006). Contribution of thermal infrared remote sensing data in multiobjective calibration of a dual-source SVAT model. *Journal of Hydrometeorology*, *7*(3), 404–420, http://dx.doi.org/10.1175/JHM503.1.

Gillespie, A., Rokugawa, S., Matsunaga, T., Cothern, J. S., Hook, S., & Kahle, A. B. (1998). A temperature and emissivity separation algorithm for Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images. *IEEE Transactions on Geoscience and Remote Sensing*, *36*(4), 1113–1126.

Göttsche, F. -M., & Hulley, G. C. (2012). Validation of six satellite-retrieved land surface emissivity products over two land cover types in a hyper-arid region. *Remote Sensing of Environment*, *124*, 149–158.

Göttsche, F. -M., Olesen, F. -S., & Bork-Unkelbach, A. (2013). Validation of land surface temperature derived from MSG/SEVIRI with in situ measurements at Gobabeb, Namibia. *International Journal of Remote Sensing*, *34*(9–10), 3069–3083.

Guillevic, P., Koster, R., Suarez, M., Bounoua, L., Collatz, G., Los, S., et al. (2002). Influence of the interannual variability of vegetation on the surface energy balance - A global sensitivity study. *Journal of Hydrometeorology*, *617–629*.

Guillevic, P., & Koster, R. D. (2002). Role of interannual vegetation variability in climate. *Bulletin of the American Meteorological Society*, *83*(12), 1753–1754.

Guillevic, P., Gastellu-Etchegorry, J. P., Demarty, J., & Prevot, L. (2003). Thermal infrared radiative transfer within three-dimensional vegetation covers. *Journal of Geophysical Research, [Atmospheres]*, *108*, 4248.

Guillevic, P. C., Privette, J. L., Coudert, B., Palecki, M.A., Demarty, J., Ottlé, C., et al. (2012). Land Surface Temperature product validation using NOAA's surface climate

observation networks - Scaling methodology for the Visible Infrared Imager Radiometer Suite (VIIRS). *Remote Sensing of Environment*, *124*, 282–298, http://dx.doi.org/10.1016/j.rse.2012.05.004.

Guillevic, P. C., Bork-Unkelbach, A., Göttsche, F. M., Hulley, G. C., Gastellu-Etchegorry, J. -P., Olesen, F. S., et al. (2013). Directional viewing effects on satellite Land Surface Temperature products over sparse vegetation canopies—A multisensor analysis. *IEEE Geoscience and Remote Sensing Letters*, *99*, 1–5, http://dx.doi.org/10.1109/LGRS.2013.2260319.

Heidinger, A. K., Laszlo, I., Molling, C. C., & Dan Tarpley, D. (2013). Using SURFRAD to Verify the NOAA single-channel Land Surface Temperature algorithm. *Journal of Atmospheric and Oceanic Technology*, *30*, 2868–2884, http://dx.doi.org/10.1175/JTECH-D-13-00051.1.

Hook, S. J., Prata, A. J., Alley, R. E., Abtahi, A., Richards, R. C., Schladow, S. G., et al. (2003). Retrieval of lake bulk and skin temperatures using Along-Track Scanning Radiometer (ATSR-2) data: A case study using Lake Tahoe, California. *Journal of Atmospheric and Oceanic Technology*, *20*(4), 534–548.

Hook, S. J., Vaughan, R. G., Tonooka, H., & Schladow, S. G. (2007). Absolute radiometric in-flight validation of mid infrared and thermal infrared data from ASTER and MODIS on the Terra spacecraft using the Lake Tahoe, CA/NV, USA, automated validation site. *IEEE Transactions on Geoscience and Remote Sensing*, *45*, 1798–1807.

Hulley, G. C., & Hook, S. J. (2009a). Intercomparison of versions 4, 4.1 and 5 of the MODIS Land Surface Temperature and emissivity products and validation with laboratory measurements of sand samples from the Namib Desert, Namibia. *Remote Sensing of Environment*, *133*, 1313–1318.

Hulley, G. C., & Hook, S. J. (2009b). The North American ASTER Land Surface Emissivity Database (NAALSED) Version 2.0. *Remote Sensing of Environment*, 1967–1975.

Hulley, G. C., & Hook, S. J. (2010). Generating consistent Land Surface Temperature and emissivity products between ASTER and MODIS data for earth science research. *IEEE Transactions on Geoscience and Remote Sensing*, 1304–1315, http://dx.doi.org/10.1109/TGRS.2010.2063034.

Hulley, G. C., Hook, S. J., & Schneider, P. (2011). Optimized split-window coefficients for deriving surface temperatures from inland water bodies. *Remote Sensing of Environment*, *115*, 3758–3769.

Hulley, G. C., & Hook, S. J. (2012). A radiance-based method for estimating uncertainties in the Atmospheric Infrared Sounder (AIRS) land surface temperature product. *Journal of Geophysical Research Letters*, *117*, D20117, http://dx.doi.org/10.1029/2012JD019102.

Hulley, G. C., Hughes, C. G., & Hook, S. J. (2012). Quantifying uncertainties in land surface temperature and emissivity retrievals from ASTER and MODIS thermal infrared data. *Journal of Geophysical Research, [Atmospheres]*, *117*, D23113, http://dx.doi.org/10.1029/2012JD018506.

Idso, S. B. (1981). A Set of equations for full spectrum and 8- to 14-μm and 10.5- to 12.5 μm thermal radiation from cloudless skies. *Water Resources Research*, *17*(2), 295–304.

Jacob, F., Petitcolin, F., Schmugge, T., Vermote, E., French, A., & Ogawa, K. (2004). Comparison of land surface emissivity and radiometric temperature derived from MODIS and ASTER sensors. *Remote Sensing of Environment*, *90*, 137–152.

Jacob, F., Schmugge, T., Olioso, A., French, A. N., Ogawa, K., Petitcolin, F., et al. (2008). Potential of thermal infrared remote sensing for the monitoring of land surfaces. Book Chapter. In S. Liang (Ed.), *Advances in Land Remote Sensing: System, Modeling, Inversion and Application* (pp. 243–269) (Chapter 10).

Jimenez-Munoz, J. C., Sobrino, J. A., Mattar, C., Hulley, G., & Gottsche, F. -M. (2014). Temperature and emissivity separation from MSG/SEVIRI Data. *IEEE Transactions on Geoscience and Remote Sensing*(99), 1–15, http://dx.doi.org/10.1109/TGRS.2013.2293791.

Justice, C. O., Vermote, E., Townshend, J. R. G., DeFries, R., Roy, D. R., Hall, D. K., et al. (1998). The Moderate resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing*, *36*(4), 1228–1249.

Justice, C. O., Román, M.O., Csiszar, I., Vermote, E. F., Wolfe, R., Hook, S. J., et al. (2013). Land and cryosphere products from Suomi NPP VIIRS: Overview and status. *Journal of Geophysical Research, [Atmospheres]*, *118*, 9753–9765, http://dx.doi.org/10.1002/jgrd.50771.

Kabsch, E., Olesen, F. S., & Prata, F. (2008). Initial results of the land surface temperature (LST) validation with the Evora, Portugal ground-truth station measurements. *International Journal of Remote Sensing*, *29*(17–18), 5329–5345.

Kerr, Y. H., Lagouarde, J. P., Nerry, F., & Ottlé, C. (2004). Land surface temperature retrieval: Techniques and applications: Case of the AVHRR. In D. A. Quattrochi, & J. C. Luwall (Eds.), *Thermal remote sensing in land surface processes* (pp. 33–109). Boca Raton Fl: CRC Press.

Kondratyev, K. Y. (1969). Radiation in the Atmosphere. *International geophysics series*, *vol. 12*, New York, NY, London: Academic Press.

Lagouarde, J. -P., Ballans, H., Moreau, P., Guyon, D., & Coraboeuf, D. (2000). Experimental study of brightness surface temperature angular variations of Maritime Pine (Pinus pinaster) stands. *Remote Sensing of Environment*, *72*, 17–34.

Li, Z. -L., Tang, B. H., Wu, H., Ren, H. Z., Yan, G. J., Wan, Z. M., et al. (2013). Satellite-derived land surface temperature: Current status and perspectives. *Remote Sensing of Environment*, *131*, 14–37, http://dx.doi.org/10.1016/j.rse.2012.12.008.

Li, H., Sun, D., Yu, Y., Wang, H., Liu, Y., Liu, Q., et al. (2014). Evaluation of the VIIRS and MODIS LST products in an arid area of Northwest China. *Remote Sensing of Environment*, *142*, 111–121, http://dx.doi.org/10.1016/j.rse.2013.11.014.

Meng, C. L., Li, Z. -L., Zhan, X., Shi, J. C., & Liu, C. Y. (2009). Land surface temperature data assimilation and its impact on evapotranspiration estimates from the Common Land Model. *Water Resources Research*, *45*, W02421, http://dx.doi.org/10.1029/2008WR006971.

Merchant, C. J., Matthiesen, S., Rayner, N. A., Remedios, J. J., Jones, P. D., Olesen, F., et al. (2013). The surface temperatures of Earth: Steps towards integrated understanding of variability and change. *Geoscientific Instrumentation Methods and Data Systems*, *2*(305–321), 2013, http://dx.doi.org/10.5194/gi-2-305-2013.

Moran, M. S., Scott, R. L., Keefer, T. O., Hernandez, M., Nearing, G. S., Emmerich, W. E., et al. (2009). Partitioning evapotranspiration in semiarid grassland and shrubland ecosystems using time series of soil surface temperature. *Agricultural and Forest Meteorology*, *149*, 59–72, http://dx.doi.org/10.1016/j.agrformet.2008.07.004.

Niclòs, R., Galve, J. M., Valiente, J. A., Estrela, M. J., & Coll, C. (2011). Accuracy assessment of land surface temperature retrievals from MSG2-SEVIRI data. *Remote Sensing of Environment*, *115*, 2126–2140.

Norman, J. M., & Becker, F. (1995). Terminology in thermal infrared remote sensing of natural surfaces. *Agricultural and Forest Meteorology*, *77*, 153–166.

Ogawa, K., Schmugge, T., & Rokugawa, S. (2008). Estimating broadband emissivity of arid regions and its seasonal variations using thermal infrared remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, *46*(2), 334–343.

Philipona, R., Dutton, E., Stoffel, T., Michalsky, J., Reda, I., Stifter, A., et al. (2001). Atmospheric longwave irradiance uncertainty: Pyrgeometers compared to an absolute sky-scanning radiometer, atmospheric emitted radiance interferometer, and radiative transfer model calculations. *Journal of Geophysical Research, [Atmospheres]*, *106*, 28,129–28,141.

Pinheiro, A.C. T., Privette, J. L., & Guillevic, P. (2006). Modeling the observed angular anisotropy of land surface temperature in a savanna. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(4), 1036–1047 (April 2006).

Prata, A. J. (1994). Land surface temperatures derived from the advanced very high resolution radiometer and the along-track scanning radiometer: 2. Experimental results and validation of AVHRR algorithms. *Journal of Geophysical Research: Atmospheres*, *99*(D6), 13025–13058.

Reichle, R. H., Bosilovich, M. G., Crow, W. T., Koster, R. D., Kumar, S. V., Mahanama, S. P., et al. (2009). Recent advances in land data assimilation at the NASA Global Modeling and Assimilation Office. In S. K. Pard (Ed.), *Data assimilation for atmospheric, oceanic and hydrologic applications* (pp. 407–428). London, United Kingdom: Springer-Verlag.

Reichle, R. H., Kumar, S. V., Mahanama, S. P. P., Koster, R. D., & Liu, Q. (2010). Assimilation of satellite-derived skin temperature observations into land surface models. *Journal of Hydrometeorology*, *11*, 1103–1122, http://dx.doi.org/10.1175/2010JHM1262.1.

Roman, M.O., Schaaf, C. B., Woodcock, C. E., Strahler, A. H., Yang, X., Braswell, R. H., et al. (2009). The MODIS (Collection V005) BRDF/albedo product: Assessment of spatial representativeness over forested landscapes. *Remote Sensing of Environment*, *113*, 2476–2498, http://dx.doi.org/10.1016/j.rse.2009.07.009.

Schneider, P., Ghent, D., Corlett, G., Prata, F., & Remedios, J. (2012). AATSR validation: LST validation protocol. *Internal publication*, UL-NILU-ESA-LST-LVP Issue 1 Revision 0, page 1–39. http://lst.nilu.no/Portals/73/Docs/Reports/UL-NILU-ESA-LST-LVP-Issue1-Rev0-1604212.pdf

Sobrino, J. A., Jiménez-Munoz, J. C., & Verhoef, W. (2005). Canopy directional emissivity: Comparison between models. *Remote Sensing of Environment*, *99*, 304–314.

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.

Snyder, W. C., Wan, Z., Zhang, Y., & Feng, Y. -Z. (1998). Classification-based emissivity for land surface temperature measurement from space. *International Journal of Remote Sensing*, *19*(14), 2753–2774.

Theocharous, E., Usadi, E., & Fox, N.P. (2010). *CEOS comparison of IR brightness temperature measurements in support of satellite validation.* Part I: Laboratory and ocean surface temperature comparison of radiation thermometers. Teddington, UK: National Physical Laboratory (NPL REPORT OP 3, ISSN: 1754–2944).

Trigo, I. F., Monteiro, I. T., Olesen, F., & Kabsch, E. (2008). An assessment of remotely sensed land surface temperature. *Journal of Geophysical Research, [Atmospheres]*, *113*, http://dx.doi.org/10.1029/2008JD010035 D17108.

VIIRS LST ATBD (2011). Joint Polar Satellite System (JPSS) VIIRS Land Surface Temperature Algorithm Theoretical Basis Document (ATBD). http://npp.gsfc.nasa.gov/sciencedocuments/ATBD_122011/474-00051_LandSurfTemp_Rev-_20110422.pdf

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. (2008). Radiance-based validation of the V5 MODIS land-surface temperature product. *International Journal of Remote Sensing*, *29*, 5373–5393.

Wan, Z. (2014). New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. *Remote Sensing of Environment*, *140*, 36–45, http://dx.doi.org/10.1016/j.rse.2013.08.027.

Wang, W., Liang, S., & Meyers, T. P. (2008). Validating MODIS land surface temperature products using long-term nighttime ground measurements. *Remote Sensing of Environment*, *112*, 623–635.

Wang, K., & Liang, S. (2009). Evaluation of ASTER and MODIS land surface temperature and emissivity products using long-term surface longwave radiation observations at SURFRAD sites. *Remote Sensing of Environment*, *113*(7), 1556–1565.

Wang, Z., Schaaf, C. B., Strahler, A. H., Chopping, M. J., Román, M.O., Shuai, Y., et al. (2014). Evaluation of MODIS albedo product (MCD43A) over grassland, agriculture and forest surface types during dormant and snow-covered periods. *Remote Sensing of Environment*, *140*, 60–77.

Wolfe, R. E., Lin, G., Nishihama, M., Tewari, K. P., Tilton, J. C., & Isaacman, A.R. (2013). Suomi NPP VIIRS prelaunch and on-orbit geometric calibration and characterization.

Journal of Geophysical Research, [Atmospheres], http://dx.doi.org/10.1002/jgrd.50873 (2013JD020508).

Yu, Y., Privette, J. L., & Pinheiro, A. (2005). Analysis of the NPOESS VIIRS Land Surface Temperature algorithm using MODIS data. *IEEE Transactions on Geoscience and Remote Sensing, 43*(10), 2340–2350.

Zheng, W., Wei, H., Wang, Z., Zeng, X., Meng, J., Ek, M., et al. (2012). Improvement of daytime land surface skin temperature over arid regions in the NCEP GFS model and its impact on satellite data assimilation. *Journal of Geophysical Research, [Atmospheres], 117*, D06117, http://dx.doi.org/10.1029/2011JD015901.