

Article

Quality Assessment of S-NPP VIIRS Land Surface Temperature Product

Yuling Liu ^{1,*}, Yunyue Yu ², Peng Yu ¹, Frank M. Göttsche ³ and Isabel F. Trigo ⁴

¹ Earth System Science Interdisciplinary Center at University of Maryland, College Park, MD 20740, USA; E-Mail: peng.yu@noaa.gov

² Center for Satellite Applications and Research, NOAA/NESDIS, College Park, MD 20740, USA; E-Mail: yunyue.yu@noaa.gov

³ Karlsruhe Institute of Technology (KIT), Eggenstein-Leopoldshafen 76344, Germany; E-Mail: frank.goettsche@kit.edu

⁴ IPMA, Instituto Português do Mar e da Atmosfera, Lisbon 1749-077, Portugal; E-Mail: isabel.trigo@ipma.pt

* Author to whom correspondence should be addressed; E-Mail: yuling.liu@noaa.gov; Tel.: +1-301-683-2573.

Academic Editors: Changyong Cao, Richard Müller and Prasad S. Thenkabail

Received: 24 July 2015 / Accepted: 14 September 2015 / Published: 21 September 2015

Abstract: The VIIRS Land Surface Temperature (LST) Environmental Data Record (EDR) has reached validated (V1 stage) maturity in December 2014. This study compares VIIRS v1 LST with the ground in situ observations and with heritage LST product from MODIS Aqua and AATSR. Comparisons against U.S. SURFRAD ground observations indicate a similar accuracy among VIIRS, MODIS and AATSR LST, in which VIIRS LST presents an overall accuracy of -0.41 K and precision of 2.35 K. The result over arid regions in Africa suggests that VIIRS and MODIS underestimate the LST about 1.57 K and 2.97 K, respectively. The cross comparison indicates an overall close LST estimation between VIIRS and MODIS. In addition, a statistical method is used to quantify the VIIRS LST retrieval uncertainty taking into account the uncertainty from the surface type input. Some issues have been found as follows: (1) Cloud contamination, particularly the cloud detection error over a snow/ice surface, shows significant impacts on LST validation; (2) Performance of the VIIRS LST algorithm is strongly dependent on a correct classification of the surface type; (3) The VIIRS LST quality can be degraded when significant brightness temperature

difference between the two split window channels is observed; (4) Surface type dependent algorithm exhibits deficiency in correcting the large emissivity variations within a surface type.

Keywords: VIIRS LST EDR; split window algorithm; surface type dependency; LST uncertainty

1. Introduction

Land surface temperature (LST) is a critical parameter in the weather and climate system controlling surface heat and water exchange with the atmosphere [1]. It has been used in many applications, including weather forecasting [2,3], irrigation and water resource management particularly agricultural drought forecasting [4,5], and urban heat island monitoring [6]. Remote sensing in the thermal infrared (TIR) provides a unique resource of obtaining LST information at the regional and global scales [7]. Satellite LSTs have been routinely produced from geostationary and polar-orbiting satellites. As of 2014, more than 30 years of global satellite LST data had been accumulated, which is an important component of the Climate Data Records (CDRs).

Many algorithms have been developed for LST retrieval including single and multi-channel algorithms, e.g., [8–15]. Among these, regression algorithms based on the split window (SW) technique have been the most widely used due to their simplicity, effectiveness and robustness. The SW algorithm utilizes the differential atmospheric absorption in two adjacent channels within the thermal infrared atmospheric windows, generally centered at about 11 μm and 12 μm . Many efforts have been made since the late 1980s to extend the SW method, initially developed for sea surface temperature estimates, to retrieve the LST. With modifications to treat the spatio-temporal and spectral variations of the Land Surface Emissivity (LSE), the large difference between the LST and the air temperature, the total column water vapor (WV) in the atmosphere, and the viewing zenith angle (VZA), a variety of SW algorithms for LST retrieval have been developed [7,16]. Many operational LST products have been generated using SW algorithms such as Advanced Very High Resolution Radiometer (AVHRR), Advanced Along-Track Scanning Radiometer (AATSR) [17], Moderate Resolution Imaging Spectroradiometer (MODIS) [11], and Spinning Enhanced Visible and Infrared Imager (SEVIRI) [18,19].

The SW approach has been applied to the Visible Infrared Imaging Radiometer Suite (VIIRS) instrument, a primary sensor onboard the Suomi National Polar-orbiting Partnership (S-NPP) satellite for measuring earth surface parameters. VIIRS was designed to improve upon the capabilities of the operational AVHRR and provide observation continuity with MODIS [20]. LST is one of Environmental Data Records (EDRs) measured by VIIRS, with a moderate spatial resolution of 750 m at nadir during the satellite over-pass times at both day and night. Unlike the SW algorithm for MODIS LST production, coefficients of the SW LST algorithm for VIIRS LST production are surface type dependent, *i.e.*, they depend on the 17 International Geosphere-Biosphere Programme (IGBP) types. Validation of such surface type dependent LST algorithm is particularly needed.

As a quality control (QC) procedure, the Joint Polar Satellite System (JPSS) products including the EDRs are managed through a series of maturity status review processes (*i.e.*, beta, provisional validated V1, validated V2, and validated V3 stages). It is also designed to ensure that scientists, meteorologists

and other specialists are prepared and able to utilize data through the JPSS program. The VIIRS LST product completed the validated V1 stage maturity review in December 2014. The V1 stage is a critical milestone in the JPSS EDRs production; it is defined as “using a limited set of samples, the algorithm output is shown to meet the threshold performance attributes identified in the JPSS level 1 requirements”.

Some researchers have evaluated the VIIRS LST product during its beta maturity stage [7,21,22]. In this study, we present evaluations of the VIIRS LST data since its provisional stage from April 2014 when a newly calibrated algorithm coefficients set was implemented. The evaluations are based on comparisons of the VIIRS LST data with the ground station observations and with the heritage LST products from MODIS Aqua collection 5 and AATSR. Furthermore, a method is presented to quantify the uncertainty of LST derived from the surface type dependent algorithm. In addition, a proxy method, consisting of the use of the VIIRS LST algorithm to MODIS observations, is adopted for the interpretation of the cross comparison between the VIIRS and MODIS LSTs. The outline of the paper is as follows. Section 2 gives a detailed description of the data sets including the satellite data and ground in situ data as well as the data quality control procedures. Section 3 provides a description of the methodology used in this study. The validation results are presented in Section 4. The limitation of the VIIRS LST algorithm and the uncertainty caused by the surface type input are analyzed and discussed in Section 5. Finally, the concluding remarks are provided in Section 6.

2. Data

Multiple data sets were used in this study: VIIRS LST data, which are to be assessed; ground observations from the SURFACE RADIATION budget observing network (SURFRAD) and from the validation station “Gobabeb” (Namibia) operated by Karlsruhe Institute of Technology (KIT) are used as reference data for temperature based (T-based) validation; MODIS LST and AATSR LST are reference data used in cross validation. In addition, the radiative transfer simulation database is used for the theoretical analysis of the LST uncertainty caused by the surface type input.

2.1. VIIRS LST EDR

Two algorithms coexist in the VIIRS LST software package: SW and dual split window algorithms (DSW). The VIIRS EDR has been operationally generated using SW algorithm since 11 August 2012. The VIIRS moderate resolution channels M15 and M16 centered at 10.76 μm and 12.01 μm , respectively, are utilized in the LST algorithm.

$$LST_i = a_0(i) + a_1(i)T_{15} + a_2(i)(T_{15} - T_{16}) + a_3(i)(\sec \theta - 1) + a_4(i)(T_{15} - T_{16})^2 \quad (1)$$

Where $a_j(i)$ are algorithm coefficients derived from regression analyses; index i denotes 17 land surface types defined by the International Geosphere-Biosphere Programme (IGBP); T_{15} and T_{16} are corresponding brightness temperatures of M15 and M16 bands; θ is satellite viewing zenith angle. Physically, brightness temperature T_{15} in the SW algorithm is utilized as primary estimate of the land surface temperature; the brightness temperature difference ($T_{15} - T_{16}$), with its first order and second order, are applied for atmospheric corrections; and the satellite viewing zenith angle is applied further for atmospheric path correction. Algorithm coefficients $\{a_j\}$ are clarified by daytime and nighttime atmospheric conditions.

VIIRS LST data is archived and distributed by NOAA Comprehensive Large Array-data Stewardship System (CLASS) [23]. Although VIIRS cryoradiator opened doors on 19 January 2012, 1 February 2012 is chosen as the start date in this study considering the signal at very early stage might be unstable. For T-based validation, VIIRS LST data is from the CLASS subset data provided by NASA's Land Product Evaluation and Analysis Tool Element (LPEATE); for cross-comparison, the granule data is from CLASS. All VIIRS LST is reprocessed by using the up to date LUT. The local calculation has been verified with operational product and the difference is found within the floating calculation error.

2.2. Reference Data

2.2.1. MODIS LST Product

The time for VIIRS overpass equator is at local time 1:30 am/pm which is same as MODIS Aqua so the MODIS Aqua LST product MYD11_L2 (collection 5), equivalent to VIIRS LST EDR, is selected as a reference to evaluate the performance of the VIIRS LST. The generalized split window algorithm [11] is used to derive LST value from brightness temperature measurements in MODIS band 31 and band 32, centered at 11.02 μm and 12.03 μm , respectively.

$$LST = A_0 + (A_1 + A_2 \frac{1-\varepsilon}{\varepsilon} + A_3 \frac{\Delta\varepsilon}{\varepsilon^2}) \frac{(T_{11} + T_{12})}{2} + (A_4 + A_5 \frac{1-\varepsilon}{\varepsilon} + A_6 \frac{\Delta\varepsilon}{\varepsilon^2}) \frac{(T_{11} - T_{12})}{2} \quad (2)$$

Where A_i ($i = 0-6$) are algorithm coefficients depending on viewing zenith angle, surface air temperature and water vapor content. ε and $\Delta\varepsilon$ are the mean and difference of the surface emissivity in band 31 and band 32. The coefficients are derived from regression analysis for a LST value ranging from $T_{air} - 16$ K to $T_{air} + 16$ K for C5 LST product [24].

2.2.2. AATSR LST Product

The basic form of the Advanced Along Track Scanning Radiometer (AATSR) LST algorithm expresses the LST as a linear combination of the two brightness temperatures T_{11} and T_{12} . The weak non-linearity by permitting the temperature difference to vary with a power n is introduced into the algorithm [25].

$$LST = a_{f,i,pw} + b_{f,i} (T_{11} - T_{12})^n + (b_{f,i} + c_{f,i}) T_{12} \quad (3)$$

Where $n = 1/\cos(\theta/m)$, θ is satellite zenith angle and m is a variable parameter controlling the dependence on view angle. $a_{f,i,pw} = d(\sec(\theta) - 1)pw + fa_{v,i} + (1 - f)a_{s,i}$, f is vegetation fraction and pw is total precipitable water $b_{f,i} = fb_{v,i} + (1 - f)b_{s,i}$, $c_{f,i} = fc_{v,i} + (1 - f)c_{s,i}$.

These coefficients are determined by regression over simulation dataset for 14 land cover classes ($I = 1-14$). The parameters d and m are empirically determined using the radiative transfer simulations for regions where validation data are available [25]. $m = 5$ and $d = 0.4$ are used in [26] for LST derivation. The available AATSR LST data covers the time period from 2002 to 2012. Therefore the data from 1 February 2012 to 8 April 2012 are used in this study.

2.2.3. SURFRAD Ground Observations

The SURFRAD stations provide high quality *in situ* measurements of surface upwelling and downwelling long wave radiations along with other meteorological parameters [27–29]. The pyranometer is usually mounted on a 10 meter high tower in each SURFRAD site, facing downward to measure the surface upwelling radiation. The spatial representativeness is about 70 m × 70 m [21]. Observations from SURFRAD stations have been widely used for evaluating satellite-based estimates of surface radiation, for validating hydrology, weather prediction, climate models and satellite LST products from ASTER, GOES and MODIS [1,15,30–33]. In this study, SURFRAD observations from February 2012 to April 2015 over 7 sites as shown in Table 1 are used for validation of the VIIRS LST retrieval.

Table 1. Geo-Location and surface type of the seven SURFRAD stations.

No.	Site Location	Station	Lat(N)/Lon(W)	Surface Type
1	Bondville, IL	BON	40.05/88.37	Crop Land
2	Fort Peck, MT	FPT	48.31/105.10	Grass Land
3	Goodwin Creek, MS	GWN	34.25/89.87	Grassland
4	Table Mountain, CO	TBL	40.13/105.24	Grass/Crop Land
5	Desert Rock, NV	DRA	36.63/116.02	Shrub Land
6	Pennsylvania State University, PA	PSU	40.72/77.93	Mixed Forest
7	Sioux Falls, SD	SFX	43.73/97.49	Cropland

The *in situ* surface skin temperature, T_s , is estimated using the following equation

$$T_s = ((R^\uparrow - (1-\varepsilon)R^\downarrow) / \sigma\varepsilon)^{1/4} \quad (4)$$

where R^\uparrow and R^\downarrow are upwelling and downwelling long wave fluxes respectively, ε is the surface broadband emissivity, and σ is Stefan-Boltzmann constant *i.e.*, $5.67051 \times 10^{-8} \text{ W} \cdot \text{m}^{-2} \cdot \text{K}^{-4}$. ε is estimated from a spectral to broadband relationship [34] as shown in Equation (5).

$$\varepsilon = 0.2122 \times \varepsilon_{29} + 0.3859 \times \varepsilon_{31} + 0.4029 \times \varepsilon_{32} \quad (5)$$

where, ε_{31} and ε_{32} are narrowband emissivity of the MODIS bands 29, 31, and 32 centered at 8.52 μm , 11.03 μm and 12.02 μm . Instead of using the fixed emissivity value for each site, the monthly emissivity from the Global Infrared Land Surface Emissivity database [35] is used for broadband emissivity calculation to better characterize the emissivity change over sites.

The accuracy of the LST estimated by Equation (4) depends on the accuracy of the upwelling and downwelling radiation and the broadband emissivity. The accuracy of the pyrgeometer is claimed to be about $9 \text{ W} \cdot \text{m}^{-2}$ [27]. The broadband emissivity matches well the ground measurements, with a standard deviation of 0.0085 and a bias of 0.0015 [33].

2.2.4. Ground Observation at Gobabeb, Namibia

KIt isIt is LST validation station at Gobabeb (latitude 23.55°S, longitude 15.05°E, 406 m a.s.l.) is located in the hyper-arid climate of the Namibia Desert on large (several thousand km^2) and highly homogeneous gravel plains (Figure 1). The gravel plain consists mainly of gravel and sand (about 75%)

and spatially well distributed desiccated grass (about 25%), but there are also some smaller wadis and rock outcrops. This site is well characterized and is dedicated to LST validation [18,36,37].

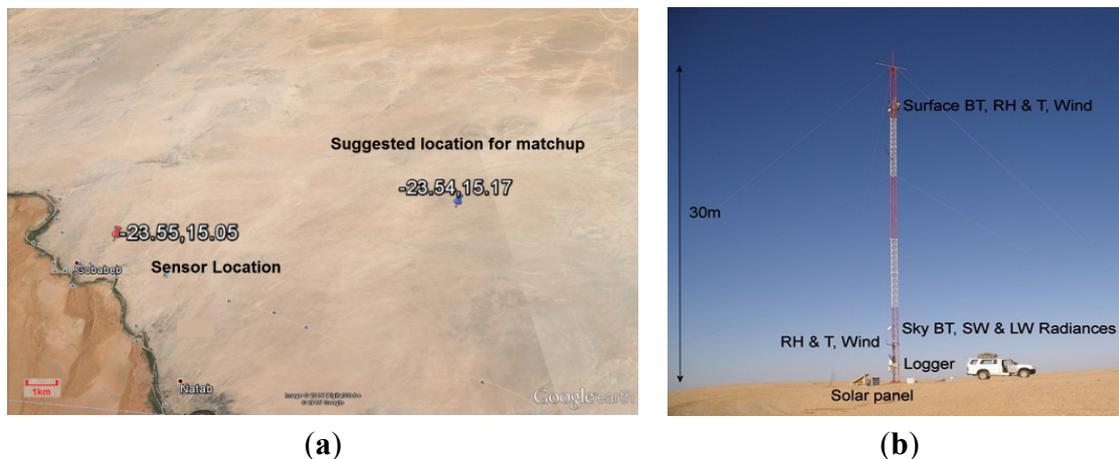


Figure 1. (a) is Geographic landscape in Gobabeb station in Namibia and (b) is the instrumentation for LST measurement: two radiometers measure the surface-leaving radiance ($9.6\text{--}11.5\ \mu\text{m}$) from the gravel plain, which is highly homogenous over at least $2500\ \text{km}^2$. A third radiometer measures sky radiance.

The station instruments are mounted at several heights of 30 m high wind profiling tower. The main instrument for the *in situ* determination of LST at KI isIt is validation stations is the precision radiometer “KT15.85 IIP” produced by Heitronics GmbH, Wiesbaden, Germany. KT15.85 IIP radiometers measure thermal infra-red radiance between $9.6\ \mu\text{m}$ and $11.5\ \mu\text{m}$, have a temperature resolution of 0.03 K and an accuracy of $\pm 0.3\ \text{K}$ over the relevant temperature range [36]. The KT15.85 IIP has a drift of less than 0.01% per month: the high stability is achieved by linking the radiance measurements via beam-chopping (a differential method) to internal reference temperature measurements and was confirmed by a long-term parallel run with the self-calibrating radiometer “RotRad” from CSIRO, which is continuously stabilized with 2 blackbodies [38]. Due to the KT-15 IIP’s narrow spectral response function and the small distance between the radiometers and the surface, atmospheric attenuation of the surface-leaving TIR radiation is negligible. However, the measurements of the surface-observing KT-15 IIPs contain radiance emitted by the surface (*i.e.*, the target signal) as well as reflected downward IR radiance from the atmosphere, which needs to be corrected for [36]. Therefore, at each station an additional KT-15.85 IIP measures downward longwave IR radiance from the atmosphere at 53° VZA: measurements under that specific zenith angle are directly related to downward hemispherical radiance [39] so that no ancillary data for deriving ground truth LST are needed.

Brightness temperatures from the surface pointing radiometers are converted to radiances; these are corrected for reflected downwelling radiance using Satellite Application Facility on Land Surface Analysis [36] emissivity and measured downwelling radiance. LST is obtained from these corrected surface leaving radiances. Ground data, corresponding to observations taken every minute throughout the whole year of 2012, are used for VIIRS LST validation.

2.3. Quality Control Procedures

Quality control of the both ground data and satellite data is critical to reliable results. The following procedures are performed in this study.

(1) Ground data quality control

Two procedures are used for ground data quality control: the first takes into account the quality flag (QF) included in the data, e.g., in SURFRAD data set, a QF of zero indicates that the corresponding data point is good, having passed all QC checks; the other is the temporal variation test by checking the Standard Deviation (STD) of ground observations in 30 min temporal interval centered at the observation time. In our practice, it is observed that LST can be varied significantly in 30 min time scale. We therefore have tested using different threshold values to filter out noise data and found that 1.5 STD is a reasonable threshold to remove noisy data and to still keeping reasonable time-varying LST data.

(2) Satellite data quality control

The main purpose for satellite data quality control is to reduce the impact from the cloud contamination and suboptimal atmospheric conditions. Therefore, the QF for cloud condition in LST product is used to constrain the cloud to be confidently clear. In addition, the spatial variation test, *i.e.*, the STD of the 3 by 3 pixel brightness temperature of the channel at 11 μm centered at the matchup pixel, is applied as an additional cloud filter [15]. The STD should be small over thermal homogeneous surfaces unless there is cirrus or cloud cover. The spatial variation test is widely used in LST validation studies. For example, Li *et al.* [30] used the neighboring 5×5 box for MODIS LST validation using 10 year SURFRAD data. In this study, the threshold is set as 1.5 K, although it may be slightly higher (e.g., 1.75 K) for sites like Boulder [33]. We intend to include all angle measurements in the validation so the LST data quality flag is not applied as it includes the viewing zenith angle restriction.

(3) Match up process

For T-based validation, the spatially closest pixel to the site is used for the matchup and temporally the ground site observations match with the granule start time. Because the subset VIIRS LST data provided by LPEATE is aggregated to ~ 5 min, which is same to MODIS v5 LST data, the temporal difference between satellite observations and ground measurements is up to 5 min. Different from MODIS and VIIRS, the overpass time of AATSR LST is calculated at pixel level so that the AATSR LST data is concurrent to the ground measurements. To evaluate the relative agreement between VIIRS and MODIS LST products, Simultaneous Nadir Overpasses (SNOs) tool [40] is used to search for the “just-miss” scenes for both satellite. The SNO tool provides the service for three regions: polar area (north and south), low latitude area and continental US area. The SNO criteria are set as 10 min temporal difference and 250 km nadir distance. More than 100 SNOs are chosen in the cross comparison of VIIRS and MODIS Aqua granules acquired from 2012 to early 2014.

3. LST Assessment Methodology

There are many challenges in evaluation of the satellite LST products. Among these, one should bear in mind that: *in situ* LST measurements are usually performed within a very small area that may not

represent well the measurement from satellite sensor, that is of the much larger pixel footprint; temporal variability of LST, particularly during the daytime, requires a stable *in situ* measurement and tolerance of very short time difference to the satellite observation; high quality *in situ* measurements are extremely limited, reducing the statistical significance of the results, and limiting the seasonal and global representativeness.

Three approaches are widely used for LST products validation: T-based method, radiance based (R-based) method and the cross-satellite comparison method. Obtaining reference (or “truth”) LST value is the key. In the T-based and the cross-satellite approach, the reference LST is the ground measurement and the cross-satellite measurement, respectively. While in the R-base method, numerical radiative transfer model is applied to calculate the reference LST using the satellite sensed top of atmosphere brightness temperatures. This method, however, requires accurate atmospheric profile and surface emissivity information which are hard to obtain. In this study, we used the T-based method and cross-satellite comparison methods to assess the VIIRS LST quality.

3.1. T-Based Validation Method

The T-based method is a direct comparison analysis of ground measurements of LST and the corresponding satellite estimates. It is based on the assumption that the ground LST measurements would represent fairly well the satellite LSTs. Obviously, such assumption may be problematic in some ground sites where thermal homogeneity is a serious issue. Some field campaigns have been conducted to validate LST products over surfaces such as lake, grassland and rice fields [24,26,41–45]. However, the *in situ* data collected from field campaigns are very limited and costly. Researchers studied various methodologies to characterize/calibrate traditional ground station data for the T-based LST validation [46–48]. In this study, we use ground observations of surface leaving longwave radiation to estimate *in situ* LST such as from the SURFRAD [1,31,33,49]. The T-based method is limited by the spatial variability of LSTs, especially during the daytime [50].

3.2. Cross Satellite Comparison Method

The cross comparison with heritage satellite LST product is widely used for satellite LST evaluation [29,51,52]. However, this is not an absolute validation unless one of the satellite products has been independently validated [21]. As the VIIRS is expected to replace MODIS in the future, the cross comparison to the MODIS LST will provide the evaluation of the VIIRS LST retrieval performance with respect to characterization of the differences, *i.e.*, spatial pattern, systematic error budget, which may reflect the algorithm difference, limitations and error sources. Since the launch of MODIS Terra in December 1999, the MODIS LST data has been widely used and evaluated by many individual and institutional investigators [40,43,53].

LST product from AATSR is also used in this study. For the nadir view of AATSR, the instantaneous field of view (IFOV) is 1 km × 1 km at the center of swath, which is also the case for MODIS observations close to nadir. One of the special features of the AATSR instrument is its use of a conical scan to give a dual-view of the Earth's surface. AATSR LST product has also been available over a decade and many studies have been conducted to assess and validate this product [9,26,54]. The validation result indicates that AATSR LST data has a fairly good accuracy, *e.g.*, an average error of −0.9 K

with a STD of 0.9 K was reported by Coll *et al.* [26] using ground measurements from experimental site in an area of rice crops.

3.3. VIIRS LST Uncertainty to Input Imprecision

The need for a correct characterization of the uncertainty associated with satellite retrievals is becoming increasingly recognized. The VIIRS LST algorithm (see Section 2.1 for further details) uses as explicit inputs the TOA brightness temperatures (split-window channels M15 and M16) and the pixel view zenith angle. Other variables, such as the pixel land cover classification and solar zenith angle are implicit inputs, since they are used to select the correct set of coefficients to be used in the retrieval (see Equation (1)). The surface type misclassification implies the use of inaccurate emissivity for the split window bands, which are built up according to emissivity and land cover type data to represent different IGBP surfaces in the regression process therefore affecting the algorithm coefficients retrieval. The overall typing accuracy for the 17 land cover types is expected to be 70 percent at moderate spatial resolution [55], which means that about 30% of pixels might be misclassified as other surface types. Several studies have pointed surface emissivity as one of the most relevant sources of LST, and here we will consider in particular the expected uncertainty in VIIRS LST, which can be attributed to land cover/emissivity misclassification, as summarized by Equations (6) and (7).

$$S_i^2 = \sum_{j=1}^{17} (P_{ij} \times \sigma^2(\epsilon_{ij})) \quad (6)$$

$$S_{sf}^2 = \sum_{i=1}^{17} \frac{S_i^2 \times N_i}{N_{total}} \quad (7)$$

Where p_{ij} is the probability of (mis-)classification of surface type i ($I = 1, 2 \dots 17$) to be j ($j = 1, 2 \dots 17$).

Table 2. Surface type distribution over land.

No.	IGBP Land Surface Type	Percentage
1	Evergreen Needle Leaf Forests	1.91
2	Evergreen Broadleaf Forests	9.25
3	Deciduous Needle leaf Forests	1.17
4	Deciduous Broadleaf Forests	0.78
5	Mixed Forests	5.85
6	Closed Shrub Lands	0.06
7	Open Shrub Lands	15.18
8	Woody Savannahs	8.16
9	Savannahs	8.13
10	Grasslands	8.45
11	Permanent Wetlands	0.87
12	Croplands	7.25
13	Urban build-up	0.39
14	Croplands/Natural Vegetation Mosaics	4.11
15	Snow ice	10.46
16	Barren	12.54
17	Water Bodies	5.45

ε_{ij} is the LST difference between LST calculated with the equation for surface type i and with the equation for surface type j for each pixel with i surface type; $\sigma^2(\varepsilon_{ij})$ is the respective error variance. S_i^2 represents the error variance associated to the pixel land cover classification, for each IGBP type i under either day or night condition. S_{sf}^2 represents the error variance for all IGBP types and all day/night conditions. N_i represents the number of samples for surface type IGBP i . N_{total} represents the total number of samples for all cases. P_{ij} is obtained from the class composition of commission errors for VIIRS surface type quality assessment (Damien Sulla-Menashe, VIIRS ST V1 Quality Assessment 2 April 2014). The type of water bodies is not included in above commission error; therefore we exclude water bodies in the uncertainty analysis. N_i/N_{total} actually represents the proportion of each surface cover over land. The global surface type distribution over land (Table 2) based on statistical analysis of a whole year of surface type input is used in the theoretical analysis.

4. Results

4.1. Comparison with SURFRAD Data

Figure 2 shows overall comparisons of the VIIRS LSTs (a) and the MODIS LSTs (b) against the SURFRAD LSTs. The number of VIIRS matchups is twice that of MODIS matchups: On the one hand, this is due to a better coverage of VIIRS compared to MODIS; on the other hand, it is a result of different cloud flag definition. In the MODIS LST product, the cloud free pixels affected by nearby clouds are excluded, which is not the case for the VIIRS LST product. It presents that accuracy and precision of the VIIRS LSTs are -0.41 K and 2.35 K, respectively, which is better than that of the MODIS LSTs (*i.e.*, -1.36 K and 2.50 K, respectively). Note that a better accuracy/precision of the VIIRS LSTs are at nighttime (-0.24 K/ 1.97 K) compared to that at daytime (-0.71 K/ 2.86 K). The better nighttime performance is expected because the thermal heterogeneity is usually higher during daytime and the atmospheric water vapor is less and the land surface behaves almost homogeneously at night [7]. This result is consistent with the results of other studies [21,33,47,56].

Note also that overall the VIIRS LSTs are 0.95 K warmer than the MODIS LSTs.

A similar comparison of the VIIRS LSTs and the AATSR LSTs against the SURFRAD LSTs is shown in Figure 3. In which, bias and of the VIIRS LSTs are -0.78 K and 2.34 K, respectively, which is comparable to that of the AATSR LSTs (*i.e.*, -0.20 K and 2.42 K, respectively). Note that VIIRS LST is on average 0.5 K colder than AATSR LST.

Note that circled matchups of the VIIRS LSTs in Figures 2 and 3 are significantly lower than the ground measurements. These are suspicious cloud contaminated data since temperature of the cloud top is mostly lower than the land surface during the time. Although the VIIRS QF from the cloud mask product and additional cloud filtering have been utilized, it appears insufficient for the validation purpose. It is found that all matchups in the circle are with snow/ice cover, which suggests a degradation of the cloud detection over bright surfaces. Four types of misclassification have been found for snow/ice identification with the cloud mask including the multi-layered cloud misclassified as snow [57]. Cloud leakage has been reported by EDR groups and snow/ice/cloud differentiation has been listed as the major issue for further improvement [58,59]. Besides, the snow/ice EDR only provides temporal snow for daytime, which leads to the incorrect surface type used in the VIIRS LST retrieval at night. Therefore,

the VIIRS nighttime LST is likely degraded by misuse of the surface cover information. In order to solve this problem, the nighttime snow/ice detection was introduced to the operational product on 22 May 2014. However, it cannot help the analysis of past data prior to that date.

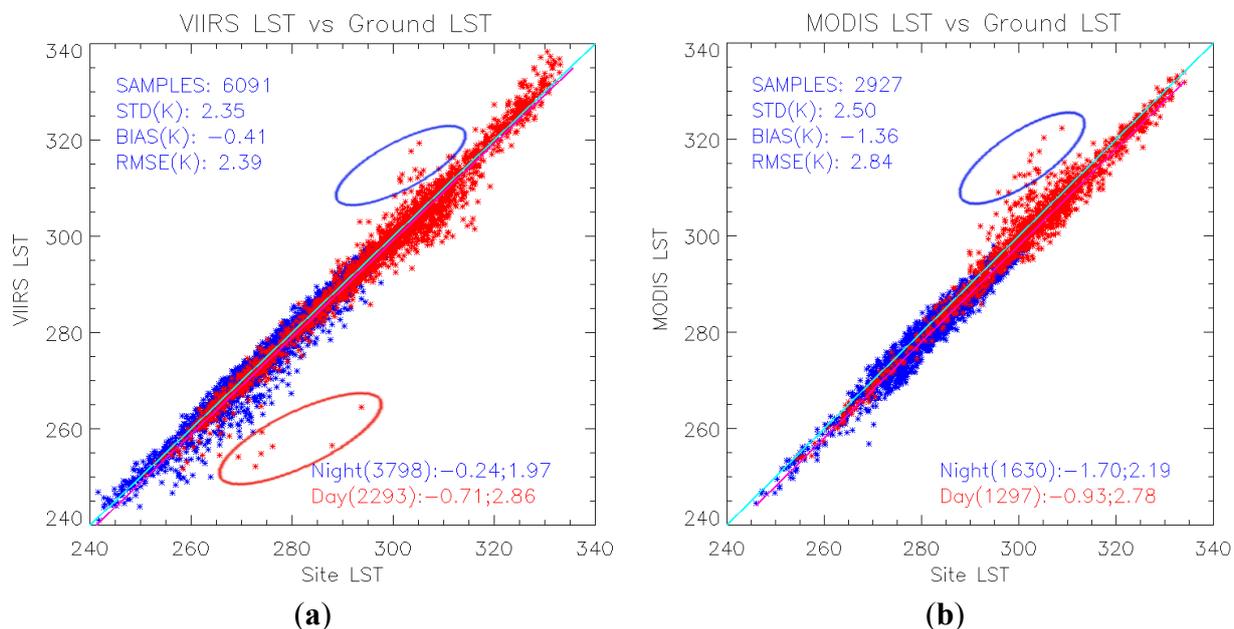


Figure 2. Scatter plots of the VIIRS LSTs (a) and MODIS LSTs (b) against the SURFRAD LSTs compared in the period from February 2012 to April 2015. Overall accuracy and precision of the satellite LSTs referring the SURFRAD LSTs are noted, as well as the daytime and nighttime cases. Some VIIRS LST plots are circled as suspicious cloud contaminated plots (red).

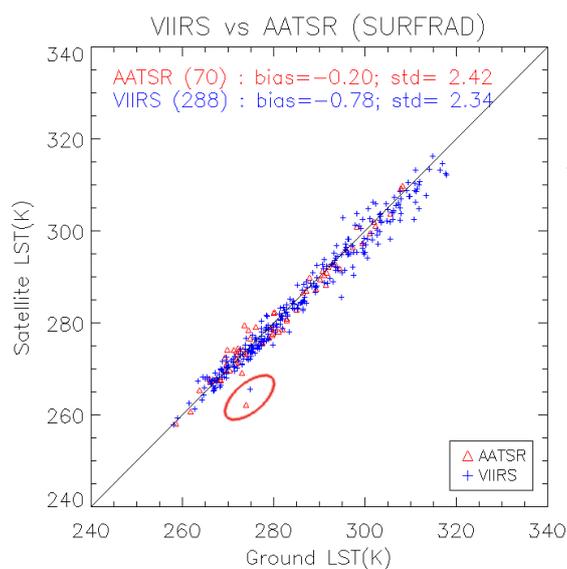


Figure 3. Scatter plots of the VIIRS LSTs (blue) and the AATSR LSTs (red) against the SURFRAD LSTs compared in the period from 1 February 2012 to 8 April 2012. Overall accuracy and precision of the satellite LSTs referring the SURFRAD LSTs are noted. Some VIIRS LST plots are circled as suspicious cloud contaminated plots.

Difference between the VIIRS LSTs and SURFRAD LSTs also demonstrates a strong seasonal variation. As shown in Table 3, the best agreement occurs in fall with a bias of -0.23 K and STD as of 1.82 K and the worst agreement in spring with a bias of -0.57 K and STD as of 2.56 K; the seasonal pattern is more significant at daytime than at nighttime. The seasonal variation is also reported in MODIS LST validation [30]. Two most relevant error sources in LST retrieval are atmospheric water vapor absorption and the surface emissivity uncertainty. Significant decrease of the atmospheric transmittance at 11 and 12 μm with increase of the water vapor introduces significant error in the split window algorithm when the surface temperature is high [60], which is the fundamental reason of the worsen performance in spring and summer compared to that in fall and winter. Besides, the large discrepancies in spring are also attributed to the cloud contamination over snow/ice surface (those matchups in the red circle of Figure 2 happened in spring) and considerably warm LST retrievals about 6 – 10 K greater than ground observations over Bondville station in late spring to early summer. As shown in blue circle of Figure 2, the feature is found in both VIIRS LST and MODIS LST validation results against SURFRAD observations. In their MODIS LST validation study, Li *et al.* [30] compared 10 years 16-day average NDVI and daily emissivity datasets from the MODIS observations and found that this feature might be caused by anomalous NDVI-emissivity relationship, *i.e.*, emissivity does not change accordingly with the NDVI change during the time period. Guillevic *et al.* [21] mentioned that validation results obtained for stations surrounded by croplands present strong seasonal dependency: station observations may be closer/deviate more from the temperature of surrounding fields, according to crop maturity. As such, VIIRS LST tends to be much lower than that of Bondville station when plants in surrounding fields (corn and soybeans) are well developed in summer and significantly higher after the harvest.

Table 3. Seasonal variation from the validation using SURFRAD measurements.

Season	Samples	Overall		Day		Night	
		Bias	STD	Bias	STD	Bias	STD
Spring	1549	-0.57	2.55	-0.58	3.16	-0.56	2.13
Summer	1433	-0.12	2.46	-0.90	3.70	0.26	1.40
Fall	1734	-0.23	1.82	-0.46	1.97	-0.07	1.70
Winter	1372	-0.72	2.21	-0.85	1.80	-0.63	2.44

In our dataset, however, we mainly observed higher LST at early growth stage from May to June but not after the harvest; and this feature is not obvious from other sites with cropland cover such as over the SXF site. Therefore, other impacts should be investigated. Some matchups with higher LST estimation are listed in Table 4. It mainly happens at local time 1–3 p.m. When crops are short, they will have little shade on ground. From the angles shown in columns of STZ (Satellite Zenith Angle), STAZ (Satellite Azimuth Angle), SOZ (Solar Zenith Angle) and SOAZ (Solar Azimuth Angle), the sun illuminates crops with a certain angle from 17 deg to 36 deg centering at about 20 deg, and the satellite views the soil surface and vegetation stem rather than the canopy.

Radiometer of the station, however, is always looking down at nadir measuring the surface upwelling radiance from the vegetation canopy. Such observation difference might be another reason leading to lower ground observations than satellite LST estimation.

Table 4. Details about the match ups over Bondville site with higher satellite LST retrievals than ground measurements.

Viirs_lst	Surfrad_lst	BT15	BT16	Date	Time	STZ	STAZ	SOZ	SOAZ
312.19	304.77	306.05	302.51	2013140	1905	33.39	-98.15	26.34	-133.20
313.65	305.98	308.04	304.88	2013163	1835	14.57	78.10	19.43	-146.77
311.39	301.43	306.27	303.76	2013164	1815	39.15	74.33	17.73	-159.42
313.91	304.41	307.21	303.14	2013169	1825	32.07	75.08	17.95	-155.56
310.84	303.62	304.40	300.88	2013170	1805	50.70	72.09	16.85	-169.68
313.14	305.44	308.69	306.60	2014128	1855	6.24	-100.44	27.15	-142.34
312.03	301.27	308.06	306.73	2014130	1815	44.90	73.11	23.26	-161.37
308.33	300.28	302.04	299.51	2014130	2000	67.60	-90.80	35.98	-118.65
309.05	302.28	305.02	303.27	2014144	1855	6.25	-97.24	24.05	-137.94
311.75	303.35	307.29	305.41	2014151	1825	38.72	74.25	19.61	-154.49
309.26	303.50	304.43	302.02	2014154	1910	25.26	-99.96	24.39	-130.55
306.46	299.39	303.54	302.82	2014165	1900	15.91	-99.07	22.50	-132.85

Table 5. Discrepancies between VIIRS LST and Ground LST over site and associated surface types.

Site	Surface Types	Samples Number	Overall			Nighttime			Daytime		
			Bias	Std	Rmse	Bias	Std	Rmse	Bias	Std	Rmse
BON	Cropland	768	-0.42	2.92	2.95	-0.48	2.05	2.10	-0.27	4.33	4.33
BON	Snow/ice	39	0.12	1.34	1.33	0.12	0.83	0.80	0.12	1.50	1.48
DRA	Closed Shrublands	97	-0.96	1.42	1.71	-1.32	0.84	1.56	-0.45	1.88	1.91
DRA	open shrublands	1128	-0.18	1.57	1.58	-0.58	0.88	1.05	0.26	2.00	2.01
DRA	Barren	149	-0.23	1.55	1.56	-1.04	0.75	1.28	0.87	1.67	1.88
FPK	Grass	491	-0.19	1.84	1.85	0.07	1.63	1.63	-0.70	2.12	2.23
FPK	Crop/vegetation Mosaic	90	-1.13	2.61	2.83	-1.70	2.86	3.31	-0.08	1.69	1.67
FPK	Snow/ice	56	-3.16	5.57	6.36	-	-	-	-3.16	5.57	6.36
GWN	Woody Savannahs	390	0.06	2.69	2.69	1.39	1.75	2.23	-2.10	2.56	3.30
GWN	Crop/vegetation Mosaic	487	-0.18	2.52	2.52	1.28	1.61	2.06	-2.20	2.11	3.05
PSU	Deciduous broadleaf forests	21	-0.85	2.52	2.60	-0.48	2.55	2.51	-1.77	2.39	2.80
PSU	Grass	157	-0.28	1.85	1.86	-0.21	1.93	1.93	-0.37	1.75	1.77
PSU	Cropland	35	-1.16	2.20	2.46	-1.21	2.38	2.63	-0.91	1.04	1.31
PSU	Crop/vegetation Mosaic	406	-0.15	2.51	2.51	-0.19	2.56	2.56	0.00	2.34	2.32
PSU	Snow/ice	105	-1.30	3.10	3.35	-2.29	3.67	4.29	-0.72	2.56	2.64
SXF	Cropland	762	-0.44	2.33	2.37	-0.13	2.07	2.07	-1.08	2.69	2.90
SXF	Snow/ice	119	-1.91	3.64	4.10	-1.94	1.94	2.72	-1.90	4.10	4.50
TBL	Grass	749	-0.68	1.81	1.94	-0.70	1.59	1.74	-0.63	2.35	2.43
TBL	Snow/ice	41	-1.36	1.80	2.24	-2.43	0.80	2.54	-1.06	1.90	2.14

To characterize the spatial representativeness of the ground site LST, ASTER LST product is used by aggregating 90 m ASTER pixels to form 1 km pixels centered at each station [7,21]. In this study, google earth image is used for visual check of the surface heterogeneity. The SURFRAD sites DRA and FPK appear more homogeneous than other sites. The quality of validation results over relatively heterogeneous sites depend on the satellite footprint, geolocation accuracy, surface type accuracy as well

as the emissivity settings of ground LST calculation. The discrepancies between VIIRS LST and ground measurements are analyzed over each site and associated surface types (Table 5). All sites except DRA and GWN present seasonal snow cover. The validation results are strongly impacted by cloud contamination in FPK and SXF sites for snow cover. The analysis result suggests land cover discrepancies between sites and satellite footprints. For example, FPK site and surroundings are located within grassland areas; however 90 out of 637 matchup pixels are classified as crop/vegetation mosaic, which results in a relatively large error. Similarly PSU site is with cropland cover on site; however, 17 matchups pixels are classified as deciduous broadleaf forests, which also causes a significant error. For the DRA site, although there were 149 matchups misclassified as barren surface, a good agreement between the satellite observations and ground measurements is obtained. This is possibly because that the emissivity setting for barren surface (emissivity pair of 0.965 and 0.97 for VIIRS band 15 and 16, respectively) is close to the bushy surface of the site. Furthermore, there would also be considerable bush shading at the site so that no obvious underestimation observed. It is also noted that the remaining 97 matchups are classified as closed shrubland at DRA site. According to the IGBP surface type definitions, shrub canopy cover is greater than 60% and 10%–60% for closed shrubland type and open shrubland type, respectively. Therefore surface type over the DRA site might change depending on the green and dry season.

4.2. Comparison with Data from Gobabeb, Namibia

The same QC control procedure as for the SURFRAD sites is implemented and the validation results are shown in Figure 4. Gobabeb *in situ* LST are also used to validate MODIS Aqua LST (collection 5), which are used as a reference.

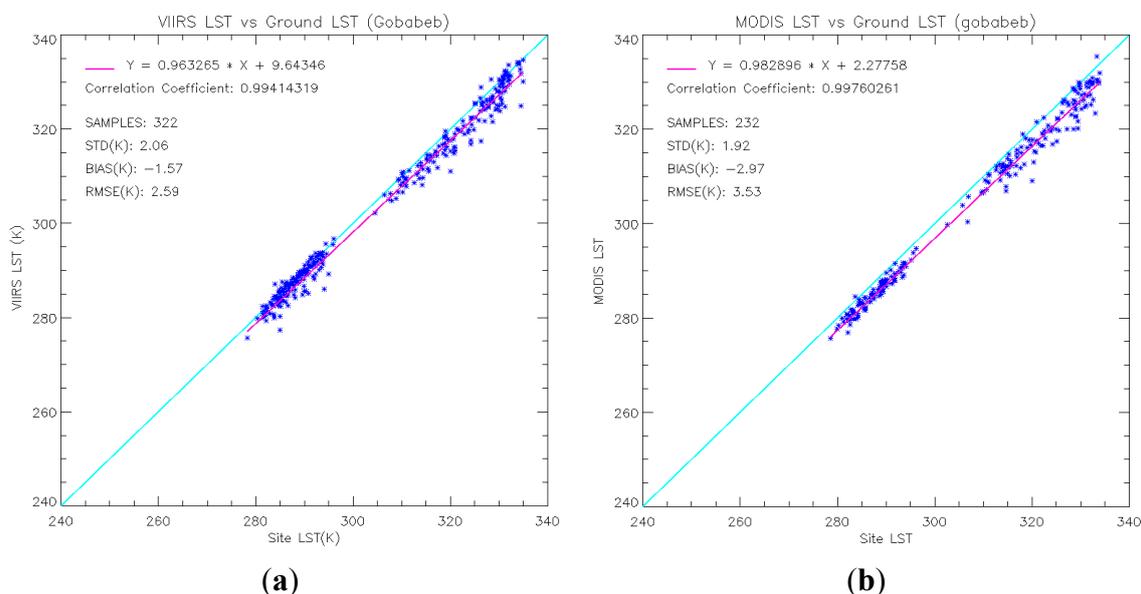


Figure 4. Validation result against the data in Gobabeb, Namibia in 2012: VIIRS LST (a) and MODIS LST V5 (b).

The results show that the VIIRS and MODIS algorithms underestimate *in situ* LST with a bias of 1.57 K and 2.97 K, respectively, whereas they achieve similar precisions of 2.06 K for VIIRS and 1.92 K for

MODIS. As shown in Figure 1, the location used for the comparison is about 13 km east of the validation station, where the gravel plain is highly homogeneous over large areas [37]. Using additional *in situ* measurements across the gravel plains, Göttsche *et al.* [36] demonstrated that the surface conditions at Gobabeb station are highly representative of the gravel plains and that there is an excellent match between the operational SEVIRI LST retrieved by EUMETSAT's Land Surface Analysis-Satellite Application Facility (LSA-SAF) and Gobabeb station LST [18], with typical monthly biases of less than 1.0 K and rms errors of about 1.0 K to 1.5 K.

Wan [54] clearly described the underestimation of MODIS v5 LST over bare soil sites. Three possible sources are considered for the large LST error: (1) The original split window algorithm does not well cover the wide range of LSTs; (2) the large errors in surface emissivity values in bands 31 and 32 estimated from land-cover types; (3) effect of dust aerosols that has not been considered in the R-based validation. Emissivity adjustment model for bare soil pixel as well as a new set of split window algorithm coefficients is incorporated into the day/night algorithm, resulting C6 level 2 LST products. Above reasons (1) and (2) are also applicable for the large error in VIIRS LST. Reason 3 is not investigated in this study.

Besides, the misclassification of surface type over Gobabeb site is observed, with 4 matchup pixels being classified as evergreen broadleaf forest. These lead to large LST errors and were therefore removed from the validation results. Again this demonstrates the impact of surface type misclassification on LST estimates. It is necessary to understand the LST uncertainty associated with the surface type input. Section 5 presents this analysis in detail.

4.3. Cross Comparison with MODIS Aqua LST

The cross comparison of the VIIRS and MODIS LST product is conducted at granule level using the SNO service. As described in the quality control section, the temporal difference is restricted to 10 min. Over 100 scenes are chosen covering each month of one year in continental US, low latitude and polar area representing low, middle to high latitudes climate. The overall comparison results as shown in Figure 5, indicate that MODIS LST and VIIRS LST produce a consistent measurement with a bias of 0.77 K and STD of 1.97 K (VIIRS minus MODIS).

Considering that the cloud residue and surface cover difference within two satellite footprints have strong impact on the cross comparison, the spatial variation test is applied to both MODIS and VIIRS LSTs, which results in the exclusion of two third of match-up pixels as shown in Figure 5b. The viewing angle difference screening is further applied based on Figure 5b. Therefore Figure 5c, representing the cross comparison results with cloud and VZA screening, shows a bias of 0.7 K and STD of 1.13 K. In order to check if the discrepancy is due to the LST retrieval algorithm, a proxy like method is used for VIIRS LST calculation, *i.e.*, using MODIS sensor data records (BT and geometry information) as input for VIIRS LST retrieval. This way the impact of difference in the sensor data records can be excluded. The result shows a bias of 0.5K and STD of 0.7 K (Figure 5d).

We also apply the above proxy like procedure to generate multiple daily global data, which leads to similar results. For example, a bias of 0.13 K and STD of 0.72 K is obtained for the global proxy comparison on 22 April 2014; a bias of 0.5 K and STD of 0.55 K is obtained on 19 December 2014. This exercise demonstrates that the algorithm difference is rather small in terms of uncertainty; more significant uncertainties are due to the sensor characterization, thermal heterogeneity of the land surface,

temporal difference, and the angular anisotropy of land surface emissivity and temperature. A positive bias of the order of 0.5 K is found between VIIRS and MODIS LST. The comparison of VIIRS and MODIS LST over surface types is summarized in Table 6. It is noted that the comparison result does not cover all surface types, e.g., snow/ice in nighttime and closed shrubland in daytime, and the number of match-ups in each surface type varies significantly from 2 to 552,550. The surface types with less than 500 samples are excluded from the following discussion due to lack of statistical representativeness. In the daytime, the LST difference ranges from 0.2 K to 3.3 K for absolute bias and from 0.57 K to 1.49 K for STD; in the nighttime, the LST difference ranges from 0.06 K to 2.11 K for absolute bias and from 0.93 K to 1.53 K for STD. Large discrepancies are found over open shrubland, savannahs and barren soil, for which the difference is up to 3.3 K. The comparison results are restrained by availability of SNOs and accuracy of the VIIRS LST surface type information.

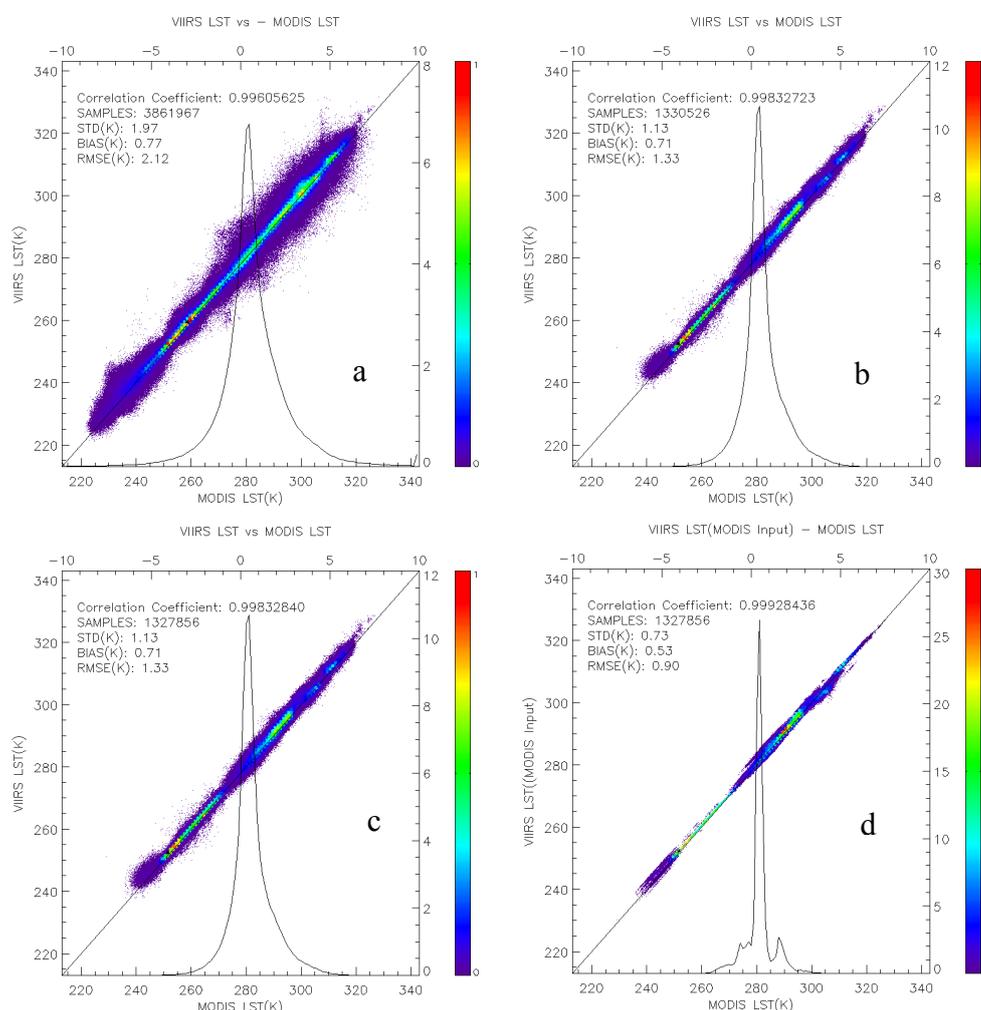


Figure 5. Cross-comparison results between VIIRS and AQUA for the whole period and area under analysis. **(a)** all comparison results under cloud clear condition ; **(b)** based on a, spatial variation tests are added ; **(c)** based on b, angle difference is added ; **(d)** based on c, VIIRS LST is calculated using MODIS data as input and then compare to MODIS LST.

However, we also observed some very large discrepancies in the cross satellite comparison of VIIRS and MODIS LST products. For example, difference over 10 K was observed on 28 December 2013 over

Australia, where VIIRS overpasses at UTC 04:17 and MODIS Aqua overpasses at UTC 04:45. Therefore, the temporal separation is about 30 min. The corresponding granules are used as a case study to investigate the problem.

Table 6. Cross comparison of VIIRS LST and MODIS LST over surface type. This result is corresponding to Figure 5c. *i.e.*, the data has been filtered to include only LST with the angle difference within 10 degrees and possible cloud contamination excluded. The overpass includes areas in low latitude, high latitude and US.

Surface Type	Night			Day		
	Bias	STD	Samples	Bias	STD	Samples
Evergreen Needleleaf Forest	−0.36	1.15	31	0.24	1.01	12
Evergreen broadleaf Forest	−0.06	0.93	11110	0.20	0.92	40085
Deciduous Needleleaf Forest	1.70	2.09	104	−1.41	0.28	2
Deciduous Broadleaf Forest	0.50	0.93	1947	−0.46	0.97	1871
Mixed Forest	−0.10	1.28	5666	−0.72	1.19	218
Closed Shrublands	1.60	0.97	858	-	-	-
Open Shrublands	2.11	1.34	166680	−0.37	1.43	1097
Woody Savannahs	0.15	1.13	124278	0.46	1.36	7728
Savannahs	0.76	1.03	145338	3.34	1.01	505
Grasslands	0.46	1.24	51831	−0.33	1.35	259
Wetlands	0.61	1.34	4371	1.72	1.13	340
Croplands	0.21	1.23	26030	−0.32	1.06	11583
Urban	0.60	1.22	769	0.38	1.40	52
Natural Vegetation Mosaics	0.44	1.04	53593	1.14	1.49	2551
Snow/ice	-	-	-	0.47	0.57	552550
Barren	2.04	1.24	1222	1.08	1.12	111549
Water	1.40	1.53	3073	−0.19	1.00	553

The overall granule comparison of VIIRS and MODIS LST (Figure 6a) shows that VIIRS LST is statistically 2 K warmer than MODIS LST and the maximum difference is over 10 K. To examine whether these highest discrepancies are caused by the sensor input, we examine the brightness temperature at 11 μm and the BT difference of the two split windows; the results are displayed in Figure 6b,c, respectively.

It is found that the VIIRS BT₁₅ is on average 1 K colder than MODIS BT₃₁ (Figure 6b) but the VIIRS BT difference between two split windows is statistically 1 K higher than that in MODIS (Figure 6c), which is considered as the main cause for the large LST discrepancy as the impact of BT difference on VIIRS LST is not linear but quadratic growth. For verification, we calculate VIIRS LST using the same proxy method in generating Figure 5d and then compare with MODIS LST. The LST discrepancy becomes much smaller as shown in (Figure 6d) with a bias of 0.01 K. Therefore, it once again suggests that the algorithm difference is not the main cause for the large discrepancies. However, VIIRS LST might overcorrect the atmospheric absorption under very high BT difference condition, which results in the LST degradation for the particular case.

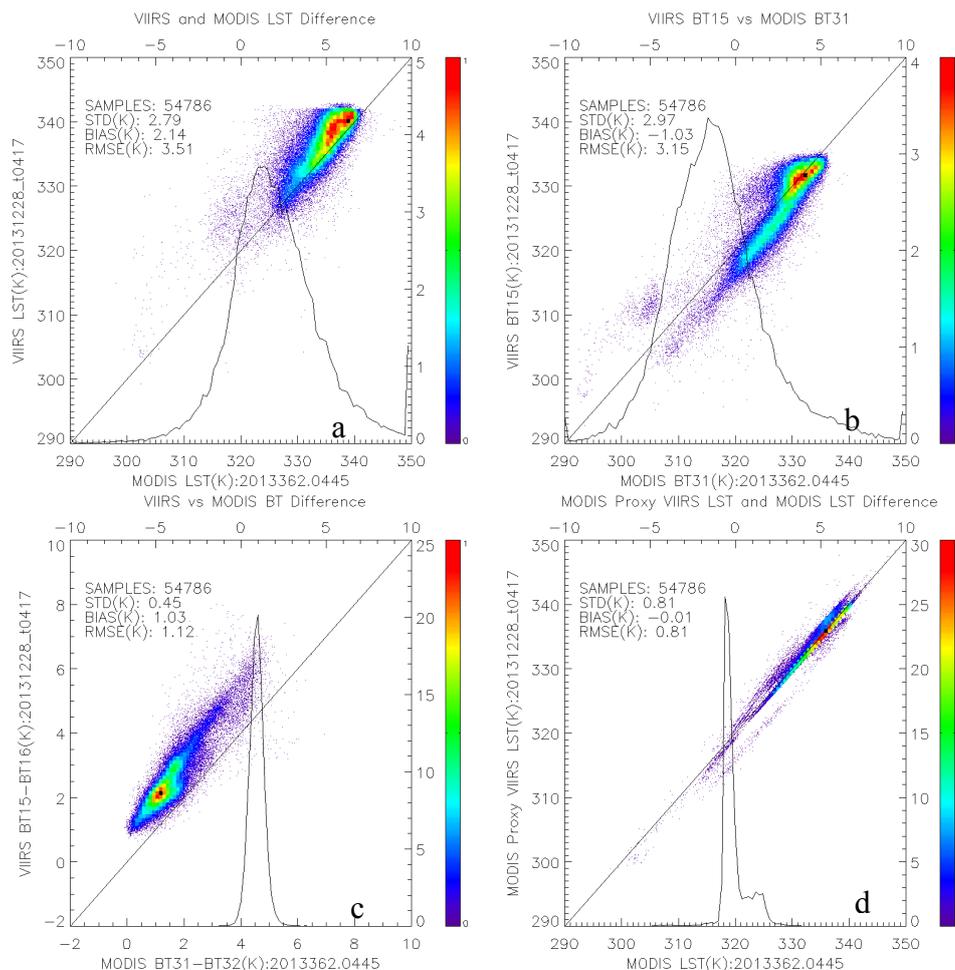


Figure 6. Cross-comparison results between VIIRS and AQUA of the case study on 28 December 2013. **(a)** Overall comparison results under cloud clear condition; **(b)** Brightness temperature comparison of VIIRS band 15 and MODIS Aqua band; **(c)** the BT difference comparison between VIIRS (BT15-BT16) and MODIS (BT31-BT32); **(d)** 31 based on a, VIIRS LST is calculated using MODIS data as input and then compare to MODIS LST

5. Discussion

5.1. Impact from the Non-Linear Term

Note that split window algorithm applied for the VIIRS LST retrieval is a heritage of the sea surface temperature (SST) retrieval; it is a linearization approach of the radiative transfer equation. Over water, surface emissivity difference between the two split window channels (centered at about 11 and 12 μm wavelengths, respectively) is ignorable so that brightness temperature (BT) difference between the split window channels represents well the atmospheric absorption. Over land, however, the BT difference includes the emissivity difference as well as the atmospheric absorption. In particularly, the VIIRS LST algorithm applies a linear and a quadratic term (Equation (1)), $(T_{11} - T_{12})$ and $(T_{11} - T_{12})^2$, for correcting atmospheric absorption. Emissivity difference between the two channels may introduce significant error. The non-linear SW algorithm is proposed primarily to improve the accuracy of LST retrieval [61]. Many similar forms of nonlinear SW algorithms have been developed in the literature.

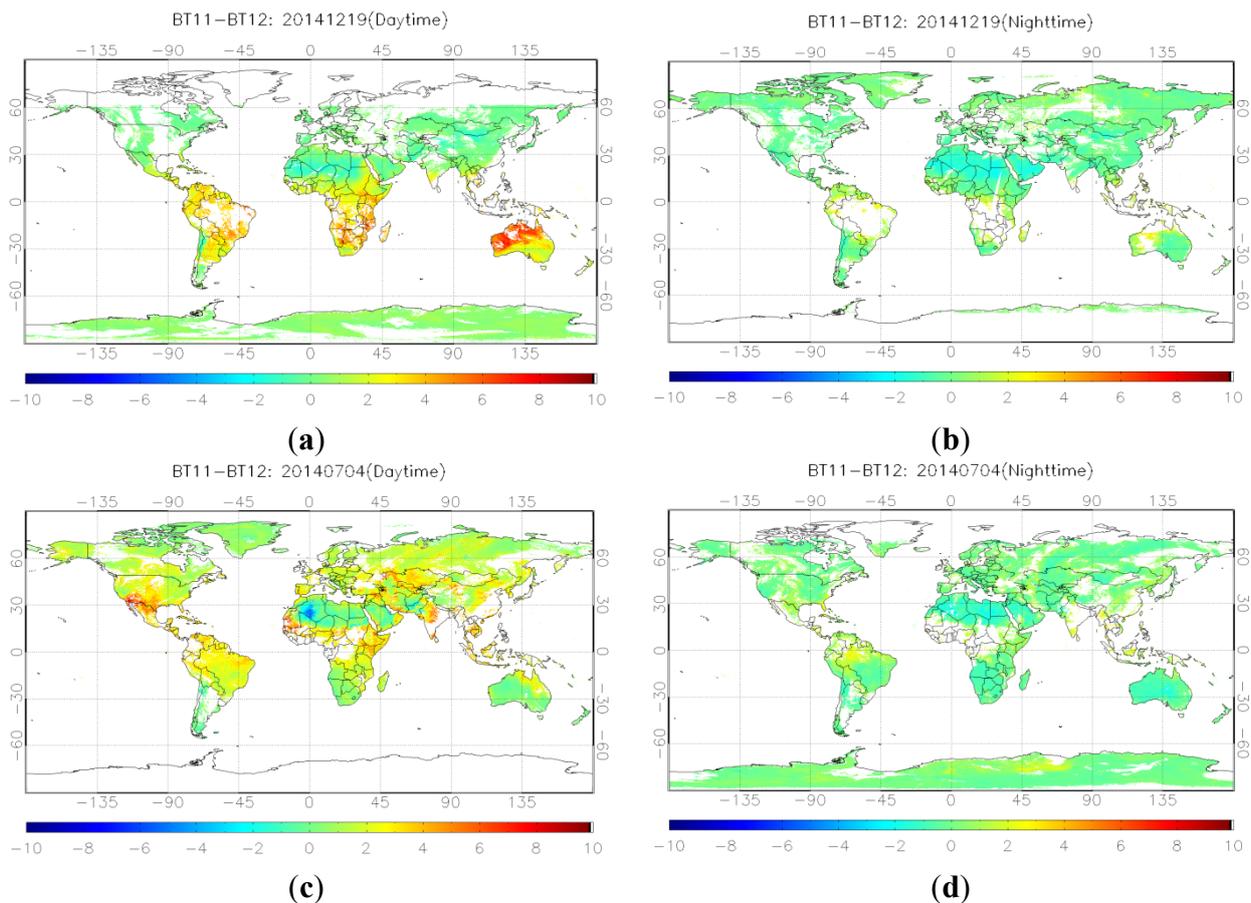


Figure 7. Global BT difference distribution map for 19 December 2014 at daytime (a) and nighttime (b); 4 July 2014 at daytime (c) and nighttime (d).

In our simulation database, the BT difference is within -1 K to 3 K for most cases but the real orbit data may go far beyond this range. Figure 7 shows the daily global BT distribution on 19 December 2014 and 4 July 2014, representing the hemisphere summer and winter. The global BT difference distribution presents an obvious regional feature, high BT difference centers in the low latitude area and low BT difference centers in middle and high latitude area with a seasonal variation. The BT difference distribution also varies under day and night condition, more homogeneous at nighttime compared to those at daytime. It is mostly less than 4 K at nighttime even in low latitude areas but close or even higher than 10 K at daytime particularly over Australia and South Africa. Refining the algorithm by extending the representativeness of the simulation database used for algorithm regression does not provide a promising improvement; it may improve over some areas but likely degrade over some other areas due to large variation at global scale. A specific coefficient set is recommended in this case to counter for the large variation within a surface type.

5.2. Uncertainty due to Error of Surface Type

Note also that the VIIRS LST algorithm is surface type dependent and therefore the impact of error in surface type input is a concern. Theoretical analysis of the algorithm uncertainty due to the surface type has been described in Section 3.3. Figure 8 shows that the overall LST uncertainty caused by surface type misclassification is 0.73 K, more specifically, 0.83 K for nighttime and 0.61 K for daytime. The impact

varies significantly with different surface types and day/night conditions. The most significant impact is found over closed shrubland in which an LST uncertainty of 1 K is found at daytime and 1.9 K at nighttime. Though the accuracy over snow/ice surface type is as high as 91%, the 9% misclassification causes 1 K LST uncertainty. Therefore, the impact of surface type accuracy on LST uncertainty is closely related to the emission characteristics differences between the correct surface type and misclassified surface types. It is noted that the daytime uncertainty is smaller than nighttime uncertainty from the theoretical analysis. The combination of atmospheric conditions and SW coefficients turn night-time LST estimates more sensitive to errors in surface type classification than daytime retrievals.

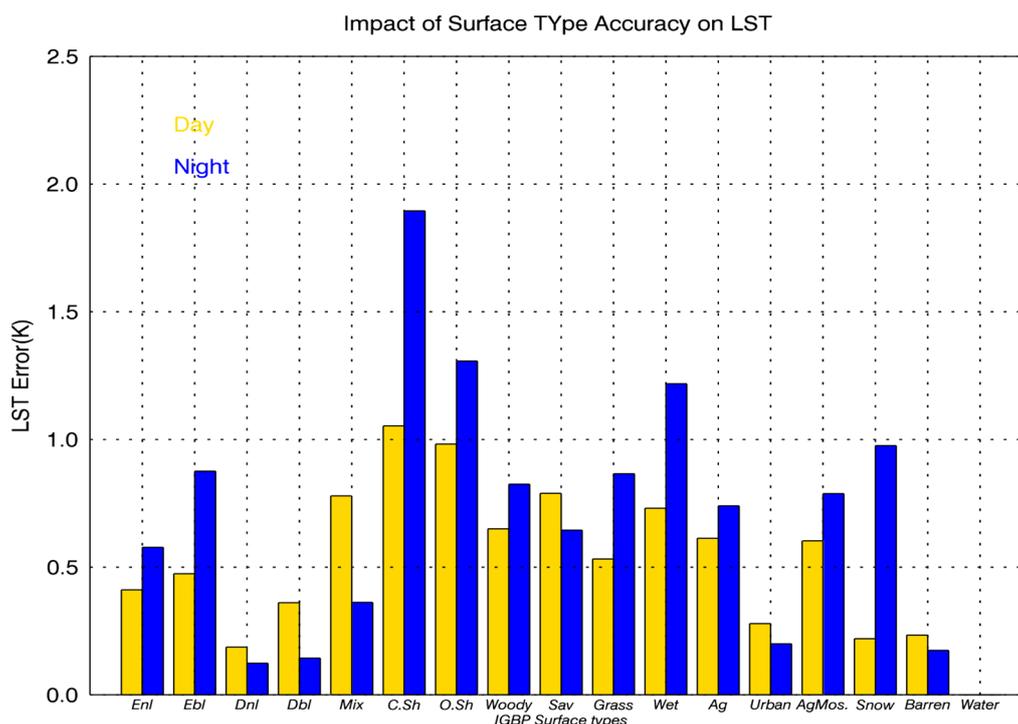
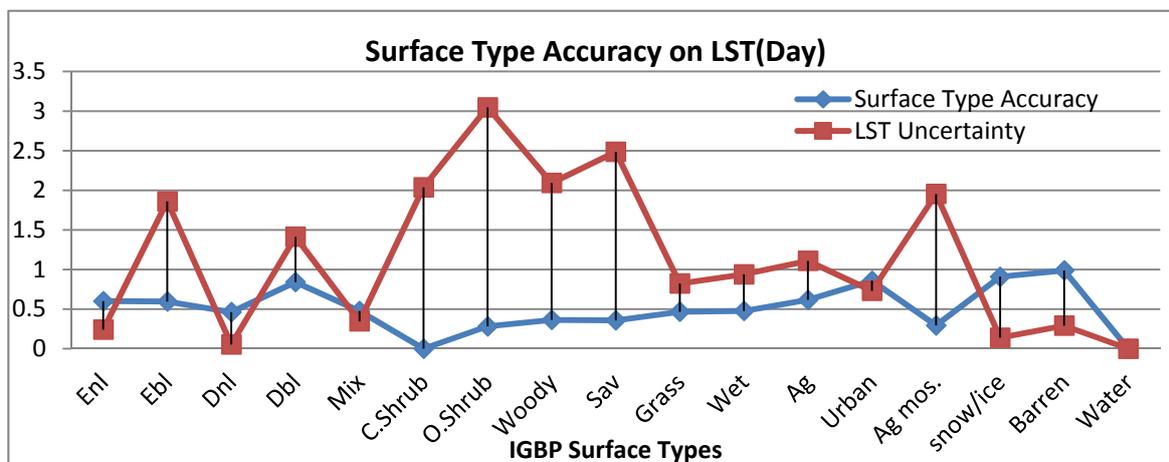


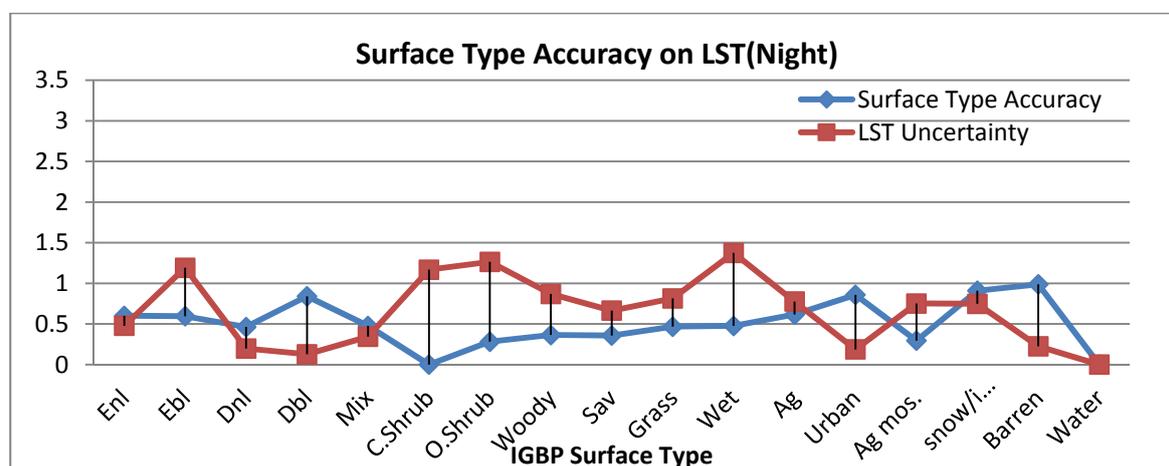
Figure 8. LST uncertainty associated with the uncertainty in surface type classification. These values are estimated using the simulation dataset for all surface types and day/night conditions.

In addition to the theoretical estimation, a set of global daily VIIRS data on 22 October 2014 for both daytime and nighttime are used as a real data case study. The blue line in Figure 9 represents the surface type accuracy from 0 (zero percent) to 1 (100 percent). The red line represents the LST uncertainty introduced by the surface type misclassification. As expected, the higher the surface type accuracy is, the smaller the uncertainty of the LST retrieval. Impact at daytime is more significant than that at nighttime. The surface type misclassification causes overall LST errors of 1.2 K, with 1.5 K and 0.8 K for daytime and nighttime respectively, which is larger than the theoretical results based on the simulation data. The maximum impact of 3 K is found at daytime over open shrubland and 1.4 K at nighttime over permanent wetland. LST performance is severely affected over closed shrub land, open shrub land, woody savannahs and croplands/Natural vegetation mosaics at daytime. Night-time LST uncertainties are generally smaller than those obtained for daytime, and tend to be higher over permanent wetland, close shrubland, open shrubland and evergreen broadleaf forest. Furthermore, some problems are observed from the surface type accuracy table, for instance closed shrub land, all classified as

evergreen needle leaf forests which is quite distinct to shrub land in terms of surface emission characteristics, causing large LST errors in this case.



(a)



(b)

Figure 9. Impact of surface type accuracy (blue line, ranging from 0 to 1) on LST uncertainty (red line, in K) for daytime (a) and nighttime (b).

6. Conclusions

Two validation approaches are used in this study, namely the T-based method to compare the VIIRS LST retrieval with ground LST measurements and the comparison with different satellite products, e.g., MODIS Aqua LST product and AATSR LST. These two methods are complementary to each other and both are useful to quantify and characterize the accuracy of VIIRS LST product and help refine the LST retrieval algorithms. The comparisons with ground measurements in SURFRAD indicate that the VIIRS LST EDR yields a reasonable accuracy with an average bias of -0.41 K and RMSE of 2.38 K for the seven sites. The accuracy at nighttime is better than that in daytime. In addition, the VIIRS LST quality shows significant seasonality with a better performance in fall and winter than that in spring and summer. The comparisons with ground measurements in Gobabeb, Namibia, indicate that VIIRS and MODIS underestimate the LST over desert area by 1.57 K and 2.97 K, respectively.

The cloud contamination and surface type misclassification are found to have a great impact on the validation result. Additional cloud screening is strongly recommended in the LST validation and applications. The surface type misclassification under given accuracy introduces a 0.7 K uncertainty in LST when using the MODTRAN simulation database, lower than that from the real case (1.2 K or so). Currently, VIIRS LST is the only product using surface type dependent algorithm among the existing operational LST products. It strongly depends on the accurate classification of surface cover at satellite footprint. The consistency of the surface type product, surface type mixed pixel and misuse of the surface type information will definitely affect and limit the performance of the VIIRS LST retrieval. During the time of this study, surface type EDR is updated quarterly but now has changed to an annual product, which will potentially have a negative impact on surface type dependent LST retrieval. As shown in Table 4, there is more than one surface classification for matchups over DRA, FPK, GWN, and PSU sites in addition to the temporal snow/ice cover. It might be either misclassification or the natural variation of the surface cover along the year. The impact on LST validation varies depending on how distinct/ close the emission characteristics are among those surface types. With the annual surface type product, the LST algorithm should be able to account for significant emissivity variability which might go beyond the range for a typical IGBP surface type. It is certainly a big challenge for the current VIIRS LST algorithm. A possible solution will be to explicitly employ emissivity values as the parameters of the split window algorithm. Many algorithms of this type have been proposed, for instance, the LST algorithms for MODIS [11], for GOES [47] as well as for the future GOES-R [1]. Additional refinement is needed to improve the emissivity explicit algorithm performance over arid and semi-arid regions, e.g., dynamic surface emissivity setting [21] and separate coefficient set for this surface type [54].

The cross comparison with existing satellite data at granule level, e.g., MODIS LST product, indicates that VIIRS LST and MODIS LST are in very good agreement with over 100 SNO scenes. However, large discrepancies are found when the BT difference between the two split window channels is very large. This usually indicates a hot and humid weather condition and a coupled effect of significant spectral emissivity difference between split window channels. This high BT difference is mostly found at daytime and spatially distributed in low latitude area. This problem is not observed in the ground validation. The analysis results using MODIS sensor data as input for VIIRS LST retrieval clearly show that the large difference between VIIRS LST and MODIS LST for this particular situation is attributed to the sensor data difference, particularly the split window BT difference. Special attention needs to be paid to the usage of the quadratic term in LST retrieval, e.g., the contribution of this term has to be restricted so that it won't cause the overcorrection of the atmospheric effect, thus cause the overestimate of the LST. Possible solutions include the stratification of the retrieval by water vapor or a separate coefficient set particularly developed for very high BT difference situation.

Acknowledgments

This study was supported by NOAA grant NA09NES4400006 (Cooperative Institute for Climate and Satellites-CICS) at the University of Maryland/ESSIC. The manuscript contents are solely the opinions of the authors and do not constitute a statement of policy, decision, or position on behalf of NOAA or the U. S. Government.

Author Contributions

Yuling Liu designed the research, collected satellite data and implemented the data analysis as well as wrote the manuscript. Yunyue Yu guided the whole study including research contents, methodology *etc.* and has the greatest contribution on the revisions of the manuscript. Peng Yu provided some tools and ideas for uncertainty and data result analysis. Frank-M. Göttsche provided the *in situ* LST measurements over Africa and revised the manuscript with a focus on the part of temperature based validation. Isabel F. Trigo contributed to the interpretation of the results. All the authors worked on the revisions of the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Yu, Y.; Tarpley, D.; Privette, J.L.; Flynn, L.E.; Xu, H.; Chen, M.; Vinnikov, K.Y.; Sun, D.; Tian, Y. Validation of GOES-R satellite land surface temperature algorithm using SURFRAD ground measurements and statistical estimates of error properties. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 704–713.
2. Meng, C.L.; Li, Z.-L.; Zhan, X.; Shi, J.C.; Liu, C.Y. Land surface temperature data assimilation and its impact on evapotranspiration estimates from the common land model. *Water Resour. Res.* **2009**, *45*, W02421.
3. Zheng, W.; Wei, H.; Wang, Z.; Zeng, X.; Meng, J.; Ek, M.; Mitchell, K.; Derber, J. Improvement of daytime land surface skin temperature over arid regions in the NCEP GFS model and its impact on satellite data assimilation. *J. Geophys. Res. Atmos.* **2012**, *117*, D06117.
4. Anderson, M.C.; Kustas, W.P.; Norman, J.M.; Hain, C.R.; Mecikalski, J.R.; Schultz, L.; Gonzalez-Dugo, M.P.; Cammalleri, C.; d’Urso, G.; Pimstein, A.; *et al.* Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. *Hydro. Earth Sys. Sci.* **2011**, *15*, 223–239.
5. Anderson, M.C.; Allen, R.G.; Morse, A.; Kustas, W.P. Use of Landsat thermal imagery in monitoring evapotranspiration and managing water resources. *Remote Sens. Environ.* **2012**, *122*, 50–65.
6. Rajasekar, U.; Weng, Q. Urban heat island monitoring and analysis by data mining of MODIS imageries. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 86–96.
7. Li, H.; Sun, D.; Yu, Y.; Wang, H.; Liu, Y.; Liu, Q.; Du, Y.; Wang, H.; Cao, B. Evaluation of the VIIRS and MODIS LST products in an arid area of Northwest China. *Remote Sens. Environ.* **2014**, *142*, 111–121.
8. Jiménez-Muñoz, J.C.; Sobrino, J.A. A generalized single-channel method for retrieving land surface temperature from remote sensing data. *J. Geophys. Res.* **2003**, *108*, 4688–4695.
9. Prata, A.J. Land surface temperatures derived from the AVHRR and the ATSR. 1, Theory. *J. Geophys. Res.* **1993**, *98*, 689–16702.
10. Coll, C.; Caselles, V.; Sobrino, J.A.; Valor, E. On the atmospheric dependence of the split-window equation for land surface temperature. *Int. J. Remote Sens.* **1994**, *15*, 105–122.

11. Wan, Z.; Dozier, J. A generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Trans. Geosci. Remote Sens.* **1996**, *34*, 892–905.
12. Gillespie, A.R.; Rokugawa, S.; Hook, S.J.; Matsunaga, T.; Kahle, A.B. *Temperature/Emissivity Separation Algorithm Theoretical Basis Document*, version 2.4; NASA/GSFC: Greenbelt, MD, USA, 1996.
13. Yu, Y.; Tarpley, D.; Privette, J.; Goldberg, M.; Raja, M.; Vinnikov, K.; Xu, H. Developing algorithm for operational GOES-R land surface temperature product. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 936–951.
14. Sun, D.; Pinker, R.T. Estimation of land surface temperature from a Geostationary Operational Environmental Satellite (GOES-8). *J. Geophys. Res.* **2003**, *108*, doi:10.1029/2002JD002422.
15. Gillespie, A.R.; Rokugawa, S.; Matsunaga, T.; Cothorn, J.S.; Hook, S.; Kahle, A.B. A temperature and emissivity separation algorithm for Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images. *IEEE Trans. Geosci. Remote Sens.* **1998**, *36*, 1113–1126.
16. Dash, P.; Göttsche, F.-M.; Olesen, F.-S.; Fischer, H. Land surface temperature and emissivity estimation from passive sensor data: Theory and practice—Current trends. *Int. J. Remote Sens.* **2002**, *23*, 2563–2594.
17. Coll, C.; Valor, E.; Galve, J.M.; Mira, M.; Bisquert, M.; García-Santos, V.; Caselles, E.; Caselles, V. Long-term accuracy assessment of land surface temperatures derived from the advanced along-track scanning radiometer. *Remote Sens. Environ.* **2012**, *116*, 211–225.
18. Freitas, S.C.; Trigo, I.F.; Bioucas-Dias, J.M.; Goettsche, F.-M. Quantifying the uncertainty of land surface temperature retrievals from SEVIRI/Meteosat. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 523–534.
19. Niclòs, R.; Galve, J.M.; Valiente, J.A.; Estrela, M.J.; Coll, C. Accuracy assessment of land surface temperature retrievals from MSG2-SEVIRI data. *Remote Sens. Environ.* **2011**, *2011*, 2126–2140.
20. Hulley, G.C.; Hook, S.J. 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 Sens. Environ.* **2009**, *133*, 1313–1318.
21. Guillevic, P.C.; Biard, C.J.; Hulley, G.C.; Privette, J.L.; Hook, S.J.; Olioso, A.; Göttsche, F.M.; Radocinski, R.; Román, M.O.; Yu, Y.; Csiszar, I. Validation of land surface temperature products derived from the Visible Infrared Imaging Radiometer Suite (VIIRS) using ground-based and heritage satellite measurements. *Remote Sens. Environ.* **2014**, *154*, 19–37.
22. Justice, C.O.; Román, M.O.; Csiszar, I.; Vermote, E.F.; Wolfe, R.; Hook, S.J.; Friedl, M.; Wang, Z.; Schaaf, C.; Miura, T.; *et al.* Land and cryosphere products from Suomi NPP VIIRS: Overview and status. *J. Geophys. Res.* **2013**, *118*, 9753–9765.
23. CLASS. Available online: <http://www.nsof.class.noaa.gov> (accessed on 14 September 2015).
24. Wan, Z. New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product. *Remote Sens. Environ.* **2014**, *140*, 36–45.
25. Prata, A.J. Land Surface Temperature Measurement from Space: AATSR Algorithm Theoretical Basis Document. Available online: https://earth.esa.int/c/document_library/et_file?folderId=13019&name=DLFE-660.pdf (accessed on 15 December 2014).

26. Coll, C.; Caselles, V.; Galve, J.M.; Valor, E.; Niclòs, R.; Sánchez, J.M.; Rivas, R. Ground measurements for the validation of land surface temperatures derived from AATSR and MODIS data. *Remote Sens. Environ.* **2005**, *97*, 288–300.
27. Augustine, J.A.; DeLuisi, J.J.; Long, C.N. SURFRAD—A national surface radiation budget network for atmospheric research. *Bull. Am. Meteorol. Soc.* **2000**, *81*, 2341–2357.
28. Augustine, J.A.; Hodges, G.B.; Cornwall, C.R.; Michalsky, J.J.; Medina, C.I. An update on SURFRAD—The GCOS surface radiation budget network for the continental United States. *J. Atmos. Ocean. Technol.* **2005**, *22*, 1460–1472.
29. Hulley, G.C.; Simon, J.H. The North American ASTER land surface emissivity database (NAALSED), version 2.0. *Remote Sens. Environ.* **2009**, *113*, 1967–1975.
30. Li, S.; Yu, Y.; Sun, D.; Tarpley, D.; Zhan, X.; Chiu, L. Evaluation of 10 year AQUA/MODIS land surface temperature with SURFRAD observations. *Int. J. Remote Sens.* **2014**, *35*, 830–856.
31. Liu, Y.; Yu, Y.; Tarpley, D.; Wang, X.; Wang, Z. Initial assessment of Suomi NPP VIIRS Land Surface Temperature (LST) algorithms ballroom G. Presented at the AMS 93rd Annual Meeting, Austin, TX, USA, 5–10 January 2013.
32. Salisbury, J.W.; D’Aria, D.M. Emissivity of Terrestrial Materials in the 8–14 μm atmospheric window. *Remote Sens. Environ.* **1992**, *42*, 83–106.
33. Wang, K.; Liang, S. Evaluation of ASTER and MODIS land surface temperature and emissivity products using long-term surface longwave radiation observations at SURFRAD sites. *Remote Sens. Environ.* **2009**, *113*, 1556–1565.
34. Wang, K.; Wan, Z.; Wang, P.; Sparrow, M.; Liu, J.; Zhou, X.; Haginoya, S. Estimation of surface long wave radiation and broadband emissivity using MODIS land surface temperature/emissivity product. *J. Geophys. Res.* **2005**, *110*, D11109.
35. Seemann, S.W.; Borbas, E.E.; Knuteson, R.O.; Stephenson, G.R.; Huang, H.-L. Development of a global infrared land surface emissivity database for application to clear sky sounding retrievals from multi-spectral satellite radiance measurements. *J. Appl. Meteor. Climatol.* **2008**, *47*, 108–123.
36. Göttsche, F.-M.; Olesen, F.-S.; Bork-Unkelbach, A. Validation of land surface temperature derived from MSG/SEVIRI with *in situ* measurements at Gobabeb, Namibia. *Int. J. Remote Sens.* **2013**, *34*, 3069–3083.
37. Göttsche, F.-M.; Hulley, G.C. Validation of six satellite-retrieved land surface emissivity products over two land cover types in a hyper-arid region. *Remote Sens. Environ.* **2012**, *124*, 149–158.
38. Kabsch, E.; Olesen, F.; Prata, F. Initial results of the land surface temperature (LST) validation with the Evora, Portugal ground-truth station measurements. *Int. J. Remote Sens.* **2008**, *29*, 5329–5345.
39. Kondratyev, K.Y. *Radiation in the Atmosphere*; Academic Press: New York, NY, USA, 1969.
40. Simultaneous Nadir Overpasses (SNOs) Tool. Available online: <http://ncc.nesdis.noaa.gov/SNOPredictions> (accessed on 14 September 2015).
41. Theocharous, E.; Usadi, E.; Fox, N.P. *CEOS Comparison of IR Brightness Temperature Measurements in Support of Satellite Validation. Part I: Laboratory and Ocean Surface Temperature Comparison of Radiation Thermometers*; National Physical Laboratory: Teddington, UK, 2010.
42. Wan, Z.; Zhang, Y.; Zhang, Q.; Li, Z. Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data. *Remote Sens. Environ.* **2002**, *83*, 163–180.

43. Wan, Z.; Zhang, Y.; Zhang, Q.; Li, Z. Quality assessment and validation of the MODIS global land surface temperature. *Int. J. Remote Sens.* **2004**, *25*, 261–274.
44. Coll, C.; Wan, Z.; Galve, J.M. Temperature-based and radiance-based validations of the V5 MODIS land surface temperature product. *J. Geophys. Res.* **2009**, *114*, D20102.
45. Hook, S.J.; Vaughan, R.G.; Tonooka, H.; Schladow, S.G. 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 Trans. Geosci. Remote Sens.* **2007**, *45*, 1798–1807.
46. Guillevic, P.; Privette, J.; Coudert, B.; Palecki, M.A.; Demarty, J.; Ottlé, C.; Augustine, J.A. Land Surface Temperature product validation using NOAA's surface climate observation networks-scaling methodology for the Visible Infrared Imager Radiometer Suite (VIIRS). *Remote Sens. Environ.* **2012**, *124*, 282–298.
47. Hale, R.C.; Gallo, K.P.; Tarpley, D.; Yu, Y. Characterization of *in situ* locations for calibration/validation of satellite-derived land surface temperature data. *Remote Sens. Lett.* **2011**, *2*, 41–50.
48. Ermida, S.L.; Trigo, I.F.; DaCamara, C.C.; Göttsche, F.M.; Olesen, F.S.; Hulley, G. Validation of remotely sensed surface temperature over an oakwood landscape-The problem of viewing and illumination geometries. *Remote Sens. Environ.* **2014**, *148*, 16–27.
49. Sun, D.; Yu, Y.; Fang, L.; Liu, Y. Toward an operational land surface temperature algorithm for GOES. *J. Appl. Meteor. Clim.* **2013**, *52*, 1974–1986.
50. Wan, Z.; Li, Z.-L. Radiance-based validation of the V5 MODIS land-surface temperature product. *Int. J. Remote Sens.* **2008**, *29*, 5373–5395.
51. Jacob, F.; Petitcolin, F.; Schmugge, T.; Vermote, E.; French, A.; Ogawa, K. Comparison of land surface emissivity and radiometric temperature derived from MODIS and ASTER sensors. *Remote Sens. Environ.* **2004**, *90*, 137–152.
52. Trigo, I.F.; Monteiro, I.T.; Olesen, F.; Kabsch, E. An assessment of remotely sensed land surface temperature. *J. Geophys. Res.* **2008**, *113*, D17108.
53. Wan, Z.; Li, Z. A physics-based algorithm for land-surface emissivity and temperature from EOS/MODIS data. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 980–996.
54. Soliman, A.; Duguay, C.; Saunders, W.; Hachem, S.P.-A. Land surface temperature from MODIS and AATSR: Product development and intercomparison. *Remote Sens.* **2012**, *4*, 3833–3856.
55. VIIRS Surface Type EDR ATBD. Available online: http://www.star.nesdis.noaa.gov/jpss/documents/ATBD/D0001-M01-S01-024_JPSS_ATBD_VIIRS-Surface-Type_A.pdf (accessed on 15 September 2015).
56. Yu, Y.; Privette, J.L.; Pinheiro, A.C. Analysis of the NPOESS VIIRS land surface temperature algorithm using MODIS data. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 2340–2350.
57. Hutchison, K.D.; Iisager, B.D.; Mahoney, R.L. Enhanced snow and ice identification with the VIIRS cloud mask algorithm. *Remote Sens. Lett.* **2013**, *4*, 929–936.
58. Kopp, T.; Heidinger, A.; Thomas, W. VIIRS Cloud Mask (VCM) Provisional Status. Available online: http://www.star.nesdis.noaa.gov/jpss/documents/meetings/2013/EDRProvReview/VCM_provisional_brief_Jan2013.pdf (accessed on 10 June 2015).

59. Kopp, T.J.; Thomas, W.; Heidinger, A.K.; Botambekov, D.; Frey, R.A.; Hutchison, K.D.; Iisager, B.D.; Brueske, K.; Reed, B. The VIIRS Cloud Mask: Progress in the first year of S-NPP toward a common cloud detection scheme. *J. Geophys. Res. Atmos.* **2014**, *119*, 2441–2456.
60. VIIRS LST ATBD. Available online: http://www.star.nesdis.noaa.gov/jpss/documents/ATBD/D0001-M01-S01-022_JPSS_ATBD_VIIRS-LST_A.pdf (accessed on 15 September 2015).
61. Li, Z.; Tang, B.; Wu, H.; Ren, H.; Yan, G.; Wan, Z.; Trigo, I.F.; Sobrino, J.A. Satellite-derived land surface temperature: current status and perspectives. *Remote Sens. Environ.* **2013**, *131*, 14–37.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).