GIScience & Remote Sensing

ISSN: 1548-1603 (Print) 1943-7226 (Online) Journal homepage: www.tandfonline.com/journals/tgrs20

Ground calibration of Sentinel-3 SLSTR thermal infrared bands and validation of LST algorithms

Jesús Puchades, Raquel Niclòs, Lluís Pérez-Planells, César Coll, Frank-Michael Göettsche, Jose Antonio Valiente & Enric Valor

To cite this article: Jesús Puchades, Raquel Niclòs, Lluís Pérez-Planells, César Coll, Frank-Michael Göettsche, Jose Antonio Valiente & Enric Valor (2026) Ground calibration of Sentinel-3 SLSTR thermal infrared bands and validation of LST algorithms, GIScience & Remote Sensing, 63:1, 2661533, DOI: 10.1080/15481603.2026.2661533

To link to this article: https://doi.org/10.1080/15481603.2026.2661533

Published online: 22 Apr 2026.

GISCIENCE & REMOTE SENSING
2026, VOL. 63, NO. 1, 2661533
https://doi.org/10.1080/15481603.2026.2661533

Taylor & Francis
by informa

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Ground calibration of Sentinel-3 SLSTR thermal infrared bands and validation of LST algorithms

Jesús Puchades[a], Raquel Niclòs[a], Lluís Pérez-Planells[b], César Coll[a], Frank-Michael Göettsche[c], Jose Antonio Valiente[b] and Enric Valor[a]

[a]Thermal Remote Sensing Group, Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia, Burjassot, Spain; [b]Fundación Centro de Estudios Ambientales del Mediterráneo (CEAM), Paterna, Spain; [c]IMK-ASF, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

ABSTRACT

Land surface temperature (LST) is an important variable for studying environmental changes and monitoring the Earth climate system. The SLSTR on board Sentinel-3A and 3B satellites is a key sensor that provides LST data on a global scale with daily revisit. The objectives of this paper are: (i) To check the on-board calibration of the SLSTR thermal-infrared bands calibration against ground-measured data for the first time. (ii) To validate the SLSTR LST product and an alternative emissivity-dependent split-window algorithm (SWA) for SLSTR data. To this end, we used the temperature based (t-based) and radiance based (r-based) methods. Ground LSTs measured from narrow-band, directional radiometer measurements from nine homogeneous sites were used for the calibration assessment and the t-based validation. The nine sites represent four different biomes. R-based validation was performed over the same t-based sites plus other five sites of varied biomes where ground LST measurements were not available. The analysis was performed for a dataset corresponding to the period 2020–2022. Results from the calibration assessment showed R-RMSDs (robust root mean square difference) between the simulated and satellite brightness temperatures of 0.7 K and 0.6 K for bands 8 and 9, respectively, for both sensors (3A and 3B), with small median differences. Such R-RMSDs were well within the uncertainty limits of the simulated brightness temperatures. The validation of the operational SLSTR product and the SWA yielded consistent results when comparing Sentinel-3A and 3B. For the t-based method, the SLSTR product showed an overall daytime (nighttime) R-RMSD of 1.5 K (1.7 K), with medians (satellite minus ground) varying between −2.0 K and 1.5 K and thus exceeding the recommended ±1.0 K threshold. For the SWA, R-RMSDs were 1.3 K both for daytime and the nighttime, and medians varied between −0.2 K and −0.9 K for daytime and nighttime. The r-based results confirmed the large medians in the SLSTR product, which ranged from −1.5 K to 1.7 K (from 0.1 K to 0.4 K for SWA). The SLSTR product showed R-RMSDs of 1.0 K (1.0 K) for the overall daytime (nighttime), while they were 0.6 K (0.3 K) for the SWA, which were lower than those for the t-based validation. The results presented here contributed to the global validation of LSTs derived from the SLSTR sensor onboard the Sentinel-3A and 3B satellites.

ARTICLE HISTORY
Received 04 November 2025
Accepted 14 April 2026

KEYWORDS
Land surface temperature (LST); sea and land surface temperature radiometer (SLSTR); split-window (SW); validation

1. Introduction

The World Meteorological Organization (WMO) and the Global Climate Observing System (GCOS 2022) consider land surface temperature (LST) as an essential climate variable (ECV) for Earth monitoring (Hollmann et al. 2013). LST is a key variable for estimating crop evapotranspiration and climate trends analysis (Anderson et al. 2012; Li et al. 2023). Thermal infrared (TIR) remote sensing provides a unique tool for LST estimation at a global scale. Accurate LST retrieval is operationally critical for hydrological and climatological applications, as it directly influences the soil moisture content (Carlson 2007) and surface energy fluxes (Li et al. 2013). In water resource management, precise LST data enables better prediction of drought conditions (AghaKouchak et al. 2015), while in climate studies it improves the characterization of urban heat islands, land–atmosphere interactions and regional climate patterns. However, the accuracy and

CONTACT Jesús Puchades jesus.puchades@uv.es

precision of satellite LST products must be evaluated regularly with reference ground LSTs from different surface covers and atmospheric conditions at the global scale. For satellite LST products, the GCOS recommended thresholds of ±1.0 K in accuracy (bias) and 1.0 K in precision (standard deviation) (GCOS 2022; Guillevic et al. 2018). Similarly, the on-board calibration of the TIR bands can be compared with ground-based measurements to check the accuracy of the satellite brightness temperatures, which are inputs of the LST products.

Satellite LSTs are usually validated through comparisons with ground LST measurement in thermally homogeneous sites in the so-called temperature-based (t-based) method (Coll et al. 2005; Göttsche et al. 2016; Hulley et al. 2019; Niclòs et al. 2018, 2023). Ideally, in situ LSTs should be obtained with TIR radiometers calibrated against traceable standards, taking directional measurements close to the satellite viewing angle in narrow bands similar to the on-board instruments. Test sites must be homogeneous in LST at the spatial scales of both the ground and satellite instrument. Thus, they are typically limited to single-component surfaces (water, full vegetation-cover crops and bare areas). This is due to the difference between the field-of-view ground radiometers (few m^2) and satellite sensors (typically > 1 km^2). Owing to the scarcity of such sites, ground LSTs measured in meteorological and climatological stations with infrared pyrgeometers (flux-meters) have been frequently used for satellite LST validation (Ghent et al. 2024; Krishnan et al. 2020; Li et al. 2020; Liu et al. 2022; Yang et al. 2020). However, pyrgeometers measure wide-band (4–50 µm), hemispherically integrated fluxes, which do not necessarily correspond to narrow-band, directional radiance measurements from satellites. This is especially true for daytime data over heterogeneous land surfaces (Wang and Liang 2009) and require further atmospheric correction (Ma et al. 2023, 2024). In addition to instrumental issues, weather and climatological stations are not usually located in thermally homogeneous areas. Therefore, ground measurements may be not representative at the satellite scale (Ma et al. 2021).

As an alternative validation approach, the radiance-based (r-based) method uses satellite brightness temperatures to estimate the reference LST based on prior knowledge of atmospheric temperature and humidity profiles and surface emissivity (Coll et al. 2024; Hulley and Hook 2012; Li et al. 2023; Wan and Li 2008). Thus, the r-based method allows to validate satellite LSTs over areas where in situ measurements are unavailable. It provides valuable information to extend the validation of LST products over a larger range of land covers and, therefore, consolidating the validation stage of the different LST products (Guillevic et al. 2018).

In this paper, we have focused ourselves on the sea and land surface temperature radiometer (SLSTR, Polehampton et al. (2023)) onboard the Copernicus Sentinel-3A and 3B satellites, which are in polar orbit with a 12-h delay, with descending (ascending) equatorial crossing at 10:00 (22:00) mean local solar time. The TIR bands of the SLSTR (bands 8 at 10.85 µm and 9 at 12.05 µm) are on-board calibrated with two black bodies that are viewed in each scan. Smith et al. (2021) reported uncertainties of 0.05 K in brightness temperatures for both Sentinel-3A and 3B. Like its predecessor ENVISAT's Advanced Along Track Scanner Radiometer, SLSTR has conical scanners to provide a dual-view: the so-called nadir (1420 km swath) and oblique view (750 km swath). An operational LST product is provided from SLSTR data at a 1 km resolution with global daily coverage (day and night, with Sentinel-3A and 3B combined), based on the split-window (SW) method applied to bands 8 and 9. The SLSTR product is only provided for the nadir view, however viewing zenith angles can reach 60º at the edge of the swath.

Unlike most SW algorithms for LST retrieval, the SLSTR algorithm (Ghent 2017) uses surface emissivity implicitly, i.e. it defines 17 static surface types or biomes from a predetermined global classification for which specific SW coefficients are calculated. Ghent et al. (2024) validated the SLSTR product from Sentinel-3A and 3B against ground LST measurements in seven sites of the surface radiation budget (SURFRAD) network (grassland, cropland and arid shrubland), which uses pyrgeometers for LST measurements. In addition, they considered four more validation sites with directional radiometer, namely two from the United States Climate Reference Network (USCRN) (grassland and cropland), one from the Atmospheric Research Measurement (ARM) (cropland) and one from the Copernicus LAW (Validation for Land Surface Temperature, Aerosol Optical Depth and Water Vapor Sentinel-3 Products Project) network (mixed forest). For the SURFRAD datasets, they showed considerable variability among sites and between daytime and nighttime, with mean biases within ±2.5 K in the daytime and ±1.5 K in nighttime, and standard deviations between 1 and 2 K. For the four sites equipped with radiometers, the LST results had daytime biases within

±1.5 K (except for the ARM cropland site) and nighttime biases within −0.9 and 0.0 K. Pérez-Planells et al. (2021) validated the SLSTR product with directional radiometer measurements in a rice paddy with seasonally changing conditions (water flooded, bare soil and full-cover crop). Results showed a bias (satellite minus ground) of 1.3 K and a total uncertainty of 1.8 K (Sentinel 3A and 3B, daytime and nighttime combined).

On the other hand, prior to the SLSTR launch, the SEN4LST project (Sobrino 2015) proposed an alternative, emissivity-dependent SW algorithm for SLSTR with synergistic use of OLCI in the Sentinel-2 mission. Its performance was assessed using MERIS/AATSR pairs, which are instruments with characteristics similar to those of OLCI/SLSTR. More recently, Pérez-Planells et al. (2021) also proposed an alternative SW algorithm (hereafter referred to as SWA) for SLSTR with an explicit dependence on surface emissivity in bands 8 and 9, which showed a bias of −0.4 K and a total uncertainty of 1.1 K for the same ground LST dataset. Emissivity-dependent SW algorithms are expected to improve LST retrieval by accounting for the dynamic emissivity characteristics of most natural surfaces, which significantly affect thermal infrared radiation measurements. By incorporating emissivity information, these algorithms are expected to reduce errors and provide more accurate LST estimates, especially over heterogeneous covers with distinctly different emissivity properties. Oppositely, biomes defined in the SLSTR LST product do not necessarily represent the heterogeneity of emissivity, since it depends on soil and vegetation types and structure, which may change within a given biome.

The objectives of the present paper are: (1) to check the on-boardonboard calibration of Sentinel-3A and 3B bands 8 and 9 using in situ LSTs from nine test sites in different biomes and climatic conditions, with a total of 838 matchups for Sentinel-3A and 938 for Sentinel-3B. All the sites were located in thermally homogeneous areas specifically chosen for LST validation and equipped with directional, narrow band TIR radiometers. (2) To extend the t-based validation of the SLSTR LST product and the emissivity-dependent SWA of Pérez-Planells et al. (2021) using the above mentioned ground LST database. (3) Application of the r-based method for validating both LST algorithms for the same nine t-based sites plus five sites in different biomes without ground LST measurements. The main novelties of the paper are: (a) For the first time, the onboard calibration of SLSTR TIR bands on both Sentinel-3A and 3B was checked against independent ground data. (b) The ground LST database used by Pérez-Planells et al. (2021) for the validation of the operational SLSTR product and the SWA was significantly enlarged in the number of matchups and six biome types for both Sentinel-3A and 3B to contribute to consolidate the validation stage. (c) Comparison of t-based and r-based methods for LST validation over the same sites.

Section 2 describes the study areas used for calibration and validation, as well as the satellite, atmospheric profile, and surface emissivity datasets. Section 3 shows the methodology for estimating ground-measured LSTs, the simulation of TOA brightness temperatures in SLSTR bands 8 and 9 from ground and atmospheric data, and the application of the r-based method to obtain reference LSTs. Section 3 also summarizes the operational SLSTR LST product and the SWA. The results for the SLSTR band calibration and the t-based and r-based validation of the SLSTR product and SWA are provided in Section 4. In Section 5, the results are discussed, and conclusions from this study are given in Section 6.

2. Study areas and datasets

2.1. T-based validation sites

A total of nine t-based validation sites from three different networks were used in this study. For these sites, 838 matchups for Sentinel-3A and 938 for Sentinel-3B were collected. Table 1 shows the corresponding cover type according to Corine (version 2020_20u1) (European Environment Agency 2019), latitude, longitude, elevation and biome according to SLSTR classification. Figure 1 shows Google Earth images of the nine test sites, which are described below:

– Two sites from the University of Valencia. The first one (rice paddy) is ~100 km^2 rice paddy area located near Valencia, Spain. This site has different land covers throughout the year, i.e. bare soil from February to May, water (flooded surface) in December, January and June, and full vegetation cover from July to

Table 1. Sites for t-based and r-based validation: Corine land cover class, latitude, longitude, elevation, and SLSTR biome.

Site	Biome Corine land cover/SLSTR	Latitude (°N)	Longitude (°E)	Elevation(m)	In situ LST
(a) Rice paddy	Rice fields/1 irrigated cropland	39.2729	−0.3185	1	✓
(b) Shrubland	Transitional woodland-shrub/6/13 closed broadleaved deciduous forest/shrubland	39.2242	−0.8969	800	✓
(c) Lake Constance	Water bodies/26 water	47.6050	9.4440	396	✓
(d) Gobabeb	Bare, sparse vegetation/20 bare area unknown	−23.5200	15.0832	426	✓
(e) KIT forest	Broad-leaved Forest/6 closed broadleaved deciduous forest	49.0760	8.4310	130	✓
(f) Hyytiälä (HYY)	Mixed Forest/10 mixed forest	61.8463	24.2956	170	✓
(g) Puechabon (PUE)	Transitional woodland-shrub/15 sparse vegetation	43.7369	3.6008	500	✓
(h) Svartbergest (SVA)	Coniferous Forest/9 open needleleaved forest	64.1708	19.7472	160	✓
(i) Robson_Creek (ROB)	Transitional woodland-shrub/5 broadleaved evergreen forest	−17.1175	145.6301	710	✓
(j) Soria Forest (SOF)	Broad-leaved Forest/8 closed needleleaved forest	41.5912	−2.5508	1120	–
(k) Albufera	Lagoon/26 water	39.3388	−0.3500	1	–
(l) Vineyard	Vineyards/2 rainfed cropland	39.5800	−1.2700	770	–
(m) Saline	Coastal lagoons/26 water	37.9966	−0.7262	1	–
(n) Olive Orchard	Olive grove/15 sparse vegetation	37.8251	−3.9680	481	–

mid-September. According to the SLSTR product it corresponds to the irrigated cropland biome. The second site (shrubland) is an ~200 km^2 high-plain area of shrublands located in eastern Spain and one of the most representative Mediterranean natural canopies (Data available in Niclòs et al. 2025). According to the SLSTR product, its land cover class is closed broadleaved deciduous forest and shrubland. The sites are equipped with Apogee SI-121 radiometers that measure the radiance in the 8–14 µm band, with a field of view of 22°. Several radiometers were set up at each site, at least one looking at the surface at angles close to nadir and another pointing to the sky at a zenith angle of 53° to measure the downwelling sky radiance. Measurements were taken every 4 s, and data corresponding to three minutes before and after the satellite overpass were used to obtain the average temperature and its temporal variability during the overpass. Before being installed at the stations, radiometers were calibrated against a Landcal P80P blackbody, which was in turn calibrated in the National Physical Laboratory (NPL) (Yamada et al. 2023a, 2023b). According to these experiments, the blackbody showed an error of 0.05 K against the NPL reference radiometer traceable to the SI (Système International) standards. The uncertainty of the radiometer calibration was around 0.1 K in the range of 0 °C–50 °C against the NPL reference blackbody traceable to SI units (Niclòs et al. 2015; Theocharous, Usadi, and Fox 2010), which is better than the 0.2 K value specified by the manufacturer. Apogee SI-121 radiometers have been used previously in other works (Niclòs et al. 2015, 2018; Pérez-Planells et al. 2021; Sánchez et al. 2015).

- Two sites from the Karlsruhe Institute of Technology (KIT). They are equipped with Heitronics KT15.85 IIP radiometers, which measure TIR radiance in the 9.6–11.5 µm spectral band, with a field of view of 8.5° and an uncertainty of 0.3 K according to the manufacturer (https://www.heitronics.com/). The first site is located at Lake Constance (Germany, Switzerland, Austria) and is classified as a water biome. A set of radiometers were installed on a car ferry that crosses the lake 16 times per day to measure the lake water surface temperature. The sky radiometer was set up at the complementary angle to the radiometer measuring the target because of the specular reflection of the downwelling atmospheric radiance over water bodies. Measurements were taken every 15 s and then averaged along 1 km of the ferry's track around the center of the lake (Hulley et al. 2021; Yang et al. 2020). The second site is located in Gobabeb, Namibia, on the highly homogeneous Namib gravel plains. The site has a hyper arid desert climate and is highly stable in space and time. According to the SLSTR product, the biome of the site is a "bare area, unknown". In Gobabeb, one radiometer is measuring the sky radiance at a zenith angle of 53°, and the other is looking at the surface. Both radiometers were configured to provide measurements every minute (Zhang et al. 2025).
- Five stations were set up in previously unrepresented forest biomes from the Copernicus Space Component LAW network. The stations have the same configuration as the Gobabeb station described above and are the following:
- KIT Forest, which was set up in 2020. It is on the premises of the KIT Campus North in Karlsruhe, Germany. The radiometers are mounted at the top of a 200 m meteorological mast and point directly to a closed broadleaved deciduous forest biome according to the SLSTR product.

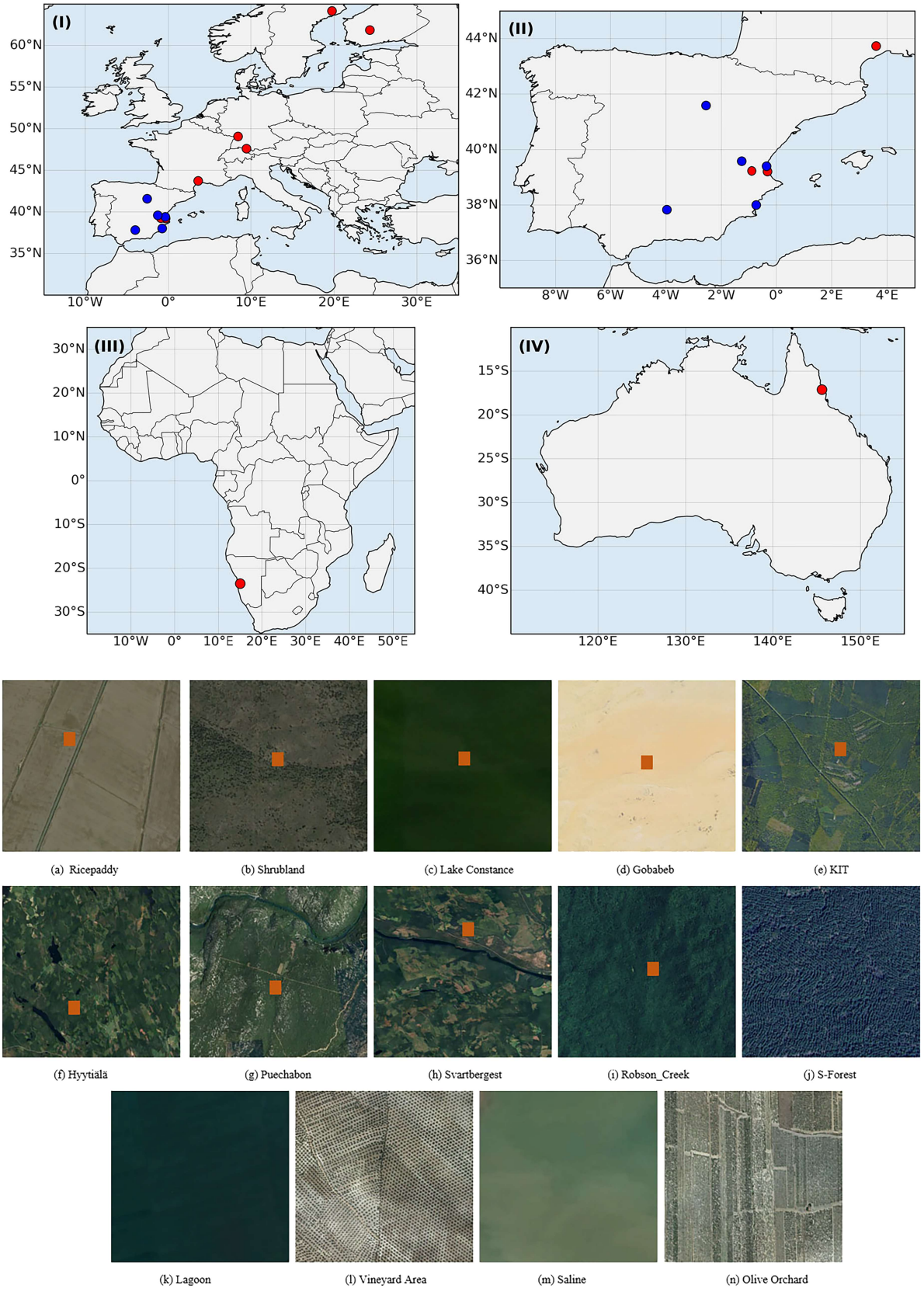

Figure 1. Top side: geographical locations of the study areas, (I) Europe, (II) Spain, (III) Africa, (IV) Australia. Red (blue) dots are t-based (r-based) study areas. Bottom side: images from Google Earth of the different test sites studied in this work labeled (a)–(n) as in Table 1. Orange squares are the location of the measuring stations. The images cover an area of 1×1 km.

The other four LAW stations have been operating since 2021 and were installed at existing CEOS land product validation supersites. They are:

- Hyytiälä (HYY), in the Juupajoki forest, Finland, classified as mixed forest.
- Puéchabon (PUE), in a forest area in South France, classified as sparse vegetation.
- Svartberget (SVA), in an experimental forest in Sweden, classified as open needleleaved forest.
- Robson Creek (ROB), in a rainforest area in Australia, classified as broadleaved evergreen forest.

The thermal homogeneity at the SLSTR scale of the t-based validation sites was checked through the standard deviation of LST for 2×2 pixels centered on each site, which ranged from 0.3 K for the Lake Constance, HYY and SVA forest sites to 0.7 K for the Gobabeb site and 0.8 K for the shrubland and PUE sites.

2.2. R-based validation sites

For the r-based validation method, we used the same sites as for the t-based method plus five additional sites without ground LST measurements representing different biomes. The additional sites are (Table 1, Figure 1):

- Soria Forest (SOF): a 200 km^2 forest located in northern Spain. Most trees are deciduous poplars, and the biome of the site is closed needleleaved forest.
- Albufera Lake: a large body of fresh water located south of Valencia, Spain. The site is classified as water.
- Vineyard: a large area of vineyards in eastern Spain. The corresponding land cover varies from bud break in April to leaf fall in mid-October, and for the rest of the year, it is leafless. The site biome is rainfed cropland.
- Saline: an area located in Spain dedicated to the production of salt by flooding plots of land with sea water. The site biome is water.
- Olive orchard: a large tract of land in southern Spain dedicated to olive production. Olives are a rainfed crop and the rows of olive trees are separated by bare soil. The biome of the olive orchard site is sparse vegetation.

SOF was grouped into the forest class, Albufera Lake and Saline were grouped into the water class, and Vineyard and Olive Orchard represented two new classes (rainfed cropland and sparse vegetation, respectively).

The study was therefore carried out with 6 classes covering a total of 11 different SLSTR biomes, so we can check the influence of the biome on the LST. All the sites have a minimum size of 3×3 Sentinel-3 SLSTR pixels (3×3 km^2).

2.3. Sentinel 3 data

TIR data from the Sentinel-3 SLSTR sensor can be obtained from the Copernicus website (https://dataspace.copernicus.eu/browser). It allows the downloading of level 1 (L1) images, which contain the brightness temperatures (BTs) of SLSTR bands 8 and 9, georeferencing data as well as additional information such as the satellite zenith angle, pixel quality data indicating cloud conditions, noise level or postprocess filling. Approximately 40% of the total available data in the study period were rejected. In order to ensure cloud-free conditions, we discarded each case with cloud values or postprocess filling in any pixel of the 2×2 pixels window selected over the site. We also used the operational cloud mask (named "cloud_in") provided by the product to check that the values were 0. In this way, any type of cloud and shadow of the cloud was removed. On the other hand, it is possible to download level 2 (L2) data that include the product LST, cloud mask, atmospheric water vapor content (WVC), pixel quality bands, fractional vegetation cover (FVC), LST uncertainty for each pixel, assigned biome and terrain elevation. Since ground LSTs were measured close to nadir, we checked that the corresponding Sentinel-3 zenith observation angle was lower than 40° in order to avoid thermal anisotropy effects between the ground and satellite data. According to Lagouarde, Kerr, and Brunet (1995), differences between nadir and off-nadir temperatures in

plant canopies and bare soils are typically negligible for zenith angles up to 30º–40º. We chose the less restrictive limit (40°) for maximizing the number of available satellite observations. Approximately 12% of the total data were discarded after angle filtering.

2.4. Atmospheric data

In order to check the calibration of the sensors and to apply the r-based method, it is necessary to know the atmospheric profiles of temperature and humidity at the time of satellite overpass to calculate the atmospheric upward radiance, downwelling radiance and transmittance. The MODTRAN 5.2 radiative transfer code was used (Berk et al. 2008) together with atmospheric profiles from the National Center of Environment Prediction (NCEP). Atmospheric data were downloaded from the website https://rda.ucar.edu/datasets/ds083.2/dataaccess/ NCEP profiles are given every 6 h with a spatial resolution of 1°, so the radiances and transmissivities were calculated for the four profiles closest to each site and for the two closest times to the satellite overpass. The profiles for each time were linearly interpolated to the overpass time, and then, we performed spatial interpolations weighted by the inverse of the squared distance to the site. Finally, the atmospheric radiances and transmittances were convolved with the spectral response function of the SLSTR bands 8 and 9. It should be noted that, due to the coarse spatial and temporal resolution of the NCEP profiles, the interpolation procedure might introduce uncertainties in the radiative transfer simulations, which are evaluated in Sections 3.2 and 3.3.

2.5. Land surface emissivity data

For each instrumented site, emissivity values for the spectral bands of the ground radiometers (either 8–14 µm for Apogee or 9.6–11.5 µm for Heitronics) are necessary to calculate ground LSTs from in situ radiance measurements. The emissivity values used in this study are given in Table 2 with the reference from which they were obtained. The emissivity values of the LAW sites are given in each database and do not vary throughout the year. On the other hand, emissivities for SLSTR bands 8 and 9 are required in all sites for simulating satellite BTs from ground data (Section 3.2), as inputs for the SWA and for r-based validation. Table 3 shows the SLSTR band emissivities for each site with reference to the method used for estimation.

Surface emissivities for the Gobabeb site were measured in situ with a field spectrometer (Göttsche et al. 2018) and convolved with the spectral response functions of the Heitronics KT15.85 and SLSTR bands 8 and 9. Water emissivities were obtained from the parametrization equation proposed by Niclòs et al. (2009), which considers the variation in emissivity as a function of the viewing zenith angle for each band. Emissivities for the forest sites, olive grove and vineyard areas were obtained using the vegetation cover method (VCM) proposed by Valor and Caselles (1996), with SLSTR level 2 FVC data and vegetation and soil emissivities obtained from Pérez-Planells et al. (2022). Salt emissivity was measured with the Box method by Rubio, Caselles, and Badenas (1997). For shrubland and rice paddy, emissivities were measured with the Box and the TES methods. Coll et al. (2016), Niclòs et al. (2011, 2018, 2023) using multiband CIMEL electronique

Table 2. Band emissivities for the spectral range of the ground radiometers used for LST measurements. References for the emissivity values are given.

Site	Reference	Spectral range	Emissivity
Rice paddy – dry bare soil	García-Santos et al. (2014)	8–14 µm	0.967
Rice paddy – wet bare soil	García-Santos et al. (2014)	8–14 µm	0.973
Rice paddy – vegetation	Pérez-Planells et al. (2021)	8–14 µm	0.983
Rice paddy – water	Pérez-Planells et al. (2021)	8–14 µm	0.986
Shrubland	Niclòs et al. (2018)	8–14 µm	0.986
Lake Constance	Thome et al. (1998)	9.6–11.5 µm	0.973
Gobabeb	Göttsche and Hulley (2012)	9.6–11.5 µm	0.940
KIT Forest	Hu et al. (2022)	9.6–11.5 µm	0.988
Hyytiälä (HYY)	Copernicus LAW LST dataset	9.6–11.5 µm	0.991
Puéchabon (PUE)	Copernicus LAW LST dataset	9.6–11.5 µm	0.989
Svartberget (SVA)	Copernicus LAW LST dataset	9.6–11.5 µm	0.991
Robson Creek (ROB)	Copernicus LAW LST dataset	9.6–11.5 µm	0.991

Table 3. Emissivities for SLSTR bands 8 and 9 used for each site of the study. Emissivity ranges are given when VCM is used (variable FVC) and for water emissivity that varies with viewing angle. In addition, the method to retrieve emissivity site is given.

Site	Method/reference	Band 8 (10.85 µm)	Band 9 (12.05 µm)
Rice paddy – dry bare soil	TES method, Niclòs et al. (2023)	0.970	0.972
Rice paddy – wet bare soil	TES method, Niclòs et al. (2023)	0.972	0.977
Rice paddy – vegetation	Box method, Niclòs et al. (2023)	0.985	0.980
Rice paddy – water	Niclòs et al. (2009)	0.988–0.992	0.985–0.988
Shrubland	VCM method, Valor and Caselles (1996)	0.980–0.985	0.983–0.987
Lake Constance	Niclòs et al. (2009)	0.991–0.992	0.986–0.988
Gobabeb	Spectrometer, Göttsche et al. (2018)	0.946	0.972
KIT Forest	VCM method, Valor and Caselles (1996)	0.980–0.986	0.982–0.986
Hyytiälä (HYY)	VCM method, Valor and Caselles (1996)	0.980–0.986	0.980–0.988
Puéchabon (PUE)	VCM method, Valor and Caselles (1996)	0.980–0.986	0.980–0.988
Svartberget (SVA)	VCM method, Valor and Caselles (1996)	0.980–0.986	0.980–0.988
Robson Creek (ROB)	VCM method, Valor and Caselles (1996)	0.980–0.986	0.980–0.988
Soria Forest (SOF)	VCM method, Valor and Caselles (1996)	0.982–0.986	0.984–0.986
Albufera	Niclòs et al. (2009)	0.990–0.992	0.986–0.988
Vineyard	VCM method, Valor and Caselles (1996)	0.969–0.976	0.961–0.972
Saline	Box method, Rubio, Caselles, and Badenas (1997)	0.991	0.990
Olive grove	VCM method, Valor and Caselles (1996)	0.972–0.976	0.966–0.973

CE-312 radiometers, with one wide band (8–13 µm) similar to Apogee and five narrow bands, two of them being similar to SLSTR bands 8 and 9 (10.9–11.7 µm and 10.2–11.0 µm, respectively).

3. Methodology

3.1. Ground LST estimation

From the at-surface radiance measured in band i of the ground radiometers, L_i^{surf}, ground LST (T_g) can be obtained from:

$$B_i(T_g) = \frac{L_i^{\text{surf}} - (1 - \varepsilon_i)\cdot L_i^{\downarrow}}{\varepsilon_i}, \tag{1}$$

where B_i is the Planck function for band i, ε_i is the surface emissivity for the radiometer band (Table 2), and $L_i^{\downarrow}$ is the sky radiance measured with a high-reflectance gold panel or at the representative zenith angle of Sentinel-3 (Guillevic et al. 2018). The Planck function integrated for band i (in $W\cdot m^{-2}\cdot\mu m^{-1}\cdot sr^{-1}$) is given by

$$B_i(T_g) = \frac{c_1}{\lambda_i^5\left[\exp\left(\frac{c_2}{\lambda_i \cdot T_g}\right) - 1\right]}, \tag{2}$$

where $c_1 = 1.1911\cdot10^8$ ($W\cdot m^{-2}\cdot\mu m^4\cdot sr^{-1}$), $c_2 = 1.4388\cdot10^4$ ($\mu m\cdot K$), and λ_i is the effective wavelength for radiometer band i (either 11.05 µm for Apogee or 10.55 µm for Heitronics). In this work, we used the effective wavelength approximation for the band-integrated Planck's function for each ground or satellite radiometer band. Then, T_g can be obtained from Equation (1) by inverting the Planck function. Using the effective wavelength approximation is accurate for narrow band radiometers such as those used here. However, it may be not accurate enough for wider band instruments.

The uncertainty associated with T_g consists of three contributions: (i) the calibration uncertainty of the instrument (0.2 K for Apogee and 0.3 K for Heitronics, here we take the largest value 0.3 K). (ii) The uncertainty in surface emissivity (0.005 according to Coll et al. 2012) which is equivalent to 0.35 K in terms of LST for both radiometers. The assumed error in emissivity is reasonable for many natural surfaces, but it could be too low for desert and saline environments, for which the uncertainty impact would be proportionally larger. (iii) The standard deviation of the different measurements taken during 6 min of measurements around the satellite overpass, which is on the order of 0.3 K. The three sources of uncertainty were added in quadrature to obtain the total uncertainty of ground LSTs, which is 0.6 K.

3.2. Simulation of satellite sensor BTs from ground data

We simulated the radiance for bands $i = 8$ and 9 (L_i^{sen}) of the satellite sensor from the in situ LSTs, emissivity data and atmospheric profiles according to:

$$L_i^{sen} = \left(\varepsilon_i \cdot B_i(T_g) + (1 - \varepsilon_i) \cdot L_i^{\downarrow}\right) \cdot \tau_i + L_i^{\uparrow}, \tag{3}$$

where ε_i is the band-effective emissivity (Table 3), $L_i^{\downarrow}$ is the sky radiance, τ_i is the atmospheric transmittance and $L_i^{\uparrow}$ is the upwelling atmospheric radiance simulated for the atmospheric profiles with MODTRAN 5.2. In Equation (3), the Planck function for the SLSTR bands is the same as in Equation (2), with effective wavelength of 10.85 µm for band 8 and 12.05 µm for band 9. BT in band i (T_{bi}) is then related to at-sensor radiance as

$$B_i(T_{bi}) = L_i^{sen}, \tag{4}$$

so that the simulated BTs for each SLSTR band can be calculated from Equation (3) by inverting the Planck function. These values are then compared with the measured satellite BTs to check the on-board calibration of the Sentinel-3 SLSTR TIR bands.

The uncertainty of the BTs simulated from ground measurements according to Equation (3) consists of three contributions. (i) The uncertainty in the atmospheric parameters, which is calculated by varying the mixing ratio at each profile level by 10% and by varying the air temperature by 1 K at each level (Hook et al., 2004). Such variations can be considered as typical uncertainties in atmospheric measurements or small spatial and temporal mismatches that impact simulations. For the AATSR sensor (with bands similar to Sentinel-3), Coll et al. (2012) estimated such impacts as 0.43 K and 0.81 K for band 8 and band 9, respectively. (ii) The uncertainty in emissivity, which is assumed to be 0.005 and corresponds to 0.35 K and 0.31 K for bands 8 and 9, respectively (Coll et al. 2012). (iii) The impact of the total uncertainty in the ground-measured LST (Section 3.1), which propagates into the satellite BT simulation (0.38 K). The three uncertainty components were added in quadrature to give the total uncertainty in the simulated BTs, which is 0.67 K and 0.94 K for bands 8 and 9, respectively.

3.3. R-based validation procedure

We applied the r-based method proposed by Wan and Li (2008) in the simplified formulation of Coll et al. (2024). Satellite BTs for bands 8 and 9 were first extracted for the site coordinates and then converted to at-sensor radiances, L_i^{sen}, via Planck's function (Equation (4)). The radiative transfer equation was applied to obtain at-ground radiance from at-sensor radiances, atmospheric parameters and emissivity

$$B_i(T_{gi}) = \frac{L_i^{sen} - L_i^{\uparrow}}{\tau_i \cdot \varepsilon_i} - \frac{(1 - \varepsilon_i) \cdot L_i^{\downarrow}}{\varepsilon_i}. \tag{5}$$

Equation (5) is applied to both channels 8 and 9 and, after the inversion of Planck's function, two different reference LST values are obtained (T_{g8} and T_{g9}). Since the LST does not depend on wavelength, T_{g8} and T_{g9} should be equal in principle. However, due to errors in the atmospheric and emissivity inputs required, there might be differences between both LSTs in such a way that the larger the difference $\delta = T_{g8} - T_{g9}$, the larger the errors on LST calculations, so δ should be close to zero in an ideal case. Following Wan and Li (2008), we choose a threshold of δ to keep small atmospheric and emissivity errors in the derivation of the r-based LST. Coll et al. (2024) showed that for δ within ±0.7 K, the error in the estimated LST is within ±1.0 K; the other cases are discarded, and r-based validation is only performed for cases meeting the δ threshold requirement. Furthermore, LSTs derived from band 8 (T_{g8}) were taken as the reference ground LSTs, since they have smaller uncertainties.

3.4. SLSTR LST level 2 operational product

The Sentinel-3 SLSTR LST level 2 operational product is a biome-dependent split-window algorithm (Chang et al. 2021). The LST is given by,

$$T_{\text{prod}} = a_{f,i,wvc} + b_{f,i}(T_{b8} - T_{b9})^{\sec(\theta/m)} + (b_{f,i} + c_{f,i})T_{b9}, \tag{6}$$

where T_{b8} and T_{b9} are the BTs for bands 8 and 9, respectively. The coefficients $a_{f,i,wvc}$, $b_{f,i}$, and $c_{f,i}$ depend on the surface biome (i), FVC (heref), and WVC. m is a parameter that depends on the satellite zenith angle θ.

Each pixel is classified as one of twenty-seven different land cover types from the Globcover biome product with a resolution of 300 m (Bicheron 2008). For each biome, coefficients $a_{f,i,wvc}$, $b_{f,i}$, and $c_{f,i}$ are determined for the vegetation and bare soil components and are linearly weighted by the FVC value of each pixel. FVC estimates are derived from SLSTR optical bands with a temporal resolution of ten days and provided with the product. In the case of permanent waters, different coefficients are given for daytime and nighttime. The SLSTR biome types assigned to the study sites are shown in Table 1.

The SLSTR product does not depend explicitly on emissivity, rather it accounts implicitly for varying surface radiative properties through static biome-dependent coefficients and dynamic FVC estimates. Therefore, it can be applied to all types of surface covers when the pixel is classified into one of the predefined biomes. However, there may be surface types with changing covers through the year, so algorithm coefficients cannot be adapted correctly.

3.5. Explicitly emissivity dependent split-window algorithm

The explicitly emissivity dependent algorithm proposed by Niclòs et al. (2011) for SEVIRI and adapted for SLSTR in Pérez-Planells et al. (2021) requires emissivities in bands 8 and 9, WVC and satellite zenith angle θ. The LST is calculated as

$$\begin{aligned} T_{\text{SWA}} = & T_{b8} + a_0 + a_1\cdot(\sec(\theta) - 1) + (a_2 + a_3\cdot(\sec(\theta) - 1))\cdot(T_{b8} - T_{b9}) + (a_4 + a_5\cdot(\sec(\theta) - 1))\cdot(T_{b8} - T_{b9})^2 \\ & + (a_6 + a_7 W + a_8 W^2)\cdot(1 - \varepsilon_{\text{mean}}) - (a_9 + a_{10} W)\cdot\varepsilon_{\text{diff}}, \end{aligned} \tag{7}$$

where a_0–a_{10} are constant algorithm coefficients, $W = \text{WVC}/\cos\theta$, $\varepsilon_{\text{mean}} = \frac{\varepsilon_8 + \varepsilon_9}{2}$ and $\varepsilon_{\text{diff}} = \varepsilon_8 - \varepsilon_9$.

Pérez-Planells et al. (2021) estimated the total LST error of the algorithm to be 1.5 K for surfaces such as wet bare soil, dry bare soil, water or full vegetation cover. They considered the error associated with the algorithm's model fitting (1.4 K) and those arising from the algorithm inputs: (i) emissivity endmembers and FVC (0.5 K), (ii) WVC (0.1 K), and (iii) BTs (0.1 K).

The algorithm requires estimations of surface emissivity in both bands for a variety of land covers at the global scale. However, this approach is feasible since emissivity in the 10–13 μm window shows relatively small variability, with $\varepsilon_{\text{mean}} > 0.97$ and $\varepsilon_{\text{diff}}$ in the ±0.01 range for most natural surfaces (water, vegetation, and soils), and can be modeled for each surface type from FVC estimates and laboratory or field emissivity measurements (Valor and Caselles 1996).

Figure 2 shows a graphical summary of the various procedures described in this section. The colors are associated with the different procedures: r-based validation in yellow, emissivity-dependent SWA in orange, ground LST derivation in green, and the LST product in gray. The blue boxes represent the atmospheric parameters, the pink one is the measured and modeled emissivities, and the red one is the BTs measured by the ground radiometers. The brown block at the bottom of Figure 2 corresponds to the calibration of the SLSTR sensor.

4. Results and discussion

In this section, we present the results for the TIR bands calibration and the t-based and r-based validation of the SLSTR product and the SWA. For each of these, we compared the satellite-measured temperatures (BTs and LSTs based on both algorithms) with ground references, which were simulated BTs and LSTs from ground measurements, as well as those extracted by r-based retrieval. The standard deviation (SD) is extremely non robust against outliers. An alternative to these results from changing the formula for the

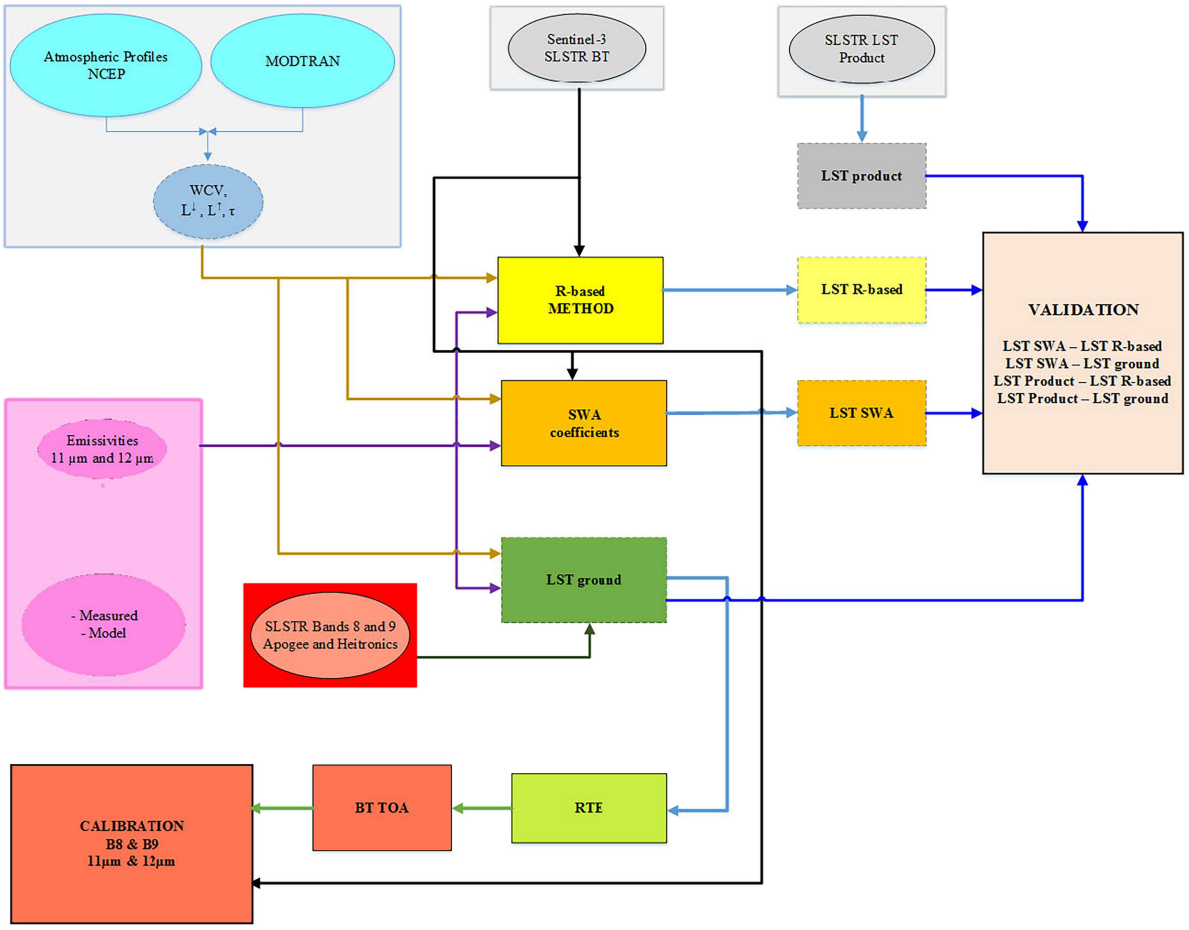

Figure 2. Flow diagram for validation and calibration.

sample standard deviation by replacing the average by the median (which is extremely robust against outliers), using absolute values of the deviations instead of squared deviations and using their median instead of their average. Robust statistics were used as a metric for the comparison, i.e. the median value and the robust standard deviation (RSD, Wilrich 2007), defined as,

$$\mathrm{RSD} = \mathrm{median}|(T_{\mathrm{sat}} - T_g)_k - \mathrm{median}(T_{\mathrm{sat}} - T_g)_k| \cdot 1.483, \tag{8}$$

where k refers to the matchup in the database. The RSD is usually a few tenths of K lower than the SD because it filters out the outliers. The robust root-mean-square difference (R-RMSD) is then calculated as the root sum of squares of the median difference and the RSD. In order to facilitate comparison with previous works, we also included the usual SD and RMSD statistics.

4.1. TIR band calibration

Figure 3 plots the satellite BTs against the simulated BTs for bands 8 (a) and 9 (b) of Sentinel-3A for daytime and nighttime data combined. Data from the different ground stations in Table 1 were grouped into four classes according to the biomes assigned by Sentinel-3: irrigated croplands (rice paddy), water (Lake Constance), desert (Gobabeb Wind), and forest (shrubland, KIT, HYY, PUE, SVA, and ROB). All the forest sites were grouped together in a single biome to facilitate the analysis. We can see that there is good agreement between the simulated and satellite BTs for both bands and all biomes for a wide range (250–310 K) of surface temperatures. The correlation coefficient (r^2) between the satellite and simulated temperatures were high, ranging from 0.98 for the forest biome to 0.9999 for the irrigated croplands in

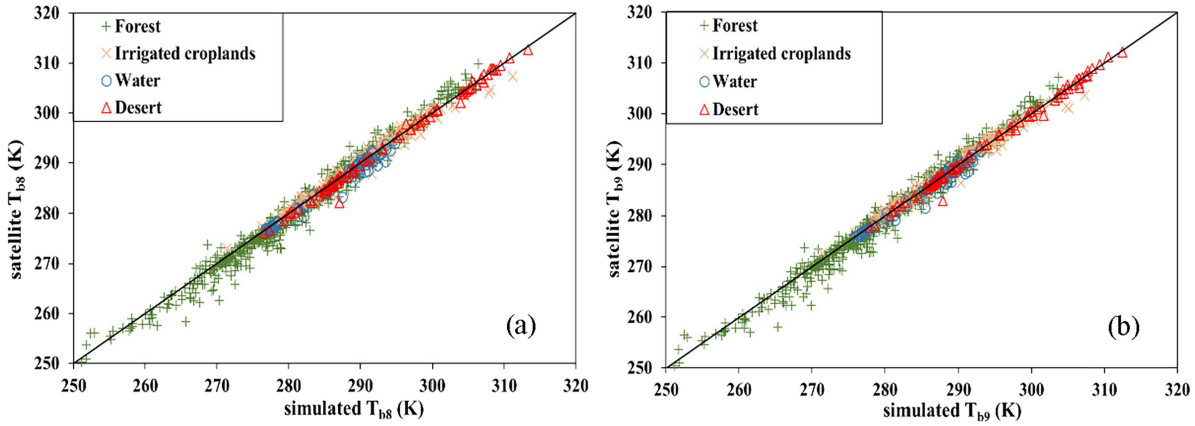

Figure 3. Satellite BTs against simulated BTs for Sentinel-3A SLSTR bands 8 (a) and 9 (b), daytime and nighttime data combined. The data are grouped by biome. The 1:1 line is shown in black.

Table 4. Median/SD and R-RMSD/RMSD values of the differences between satellite and simulated BTs in bands 8 and 9 for each ground measurement site and all sites combined. Results for Sentinel-3A and 3B are shown.

		Sentinel-3A			Sentinel-3B		
Land covers	Bands	Median/SD (K)	R-RMSD/RMSD (K)	Match-ups	Median/SD (K)	R-RMSD/RMSD (K)	Match-ups
Shrubland	Band 8	0.3/1.8	1.4/1.8	86	0.6/2.0	1.5/2.0	85
	Band 9	0.3/1.7	1.4/1.7		0.5/1.9	1.4/1.9	
Rice paddy	Band 8	0.3/1.2	0.7/1.2	136	0.3/1.2	0.7/1.2	124
	Band 9	0.2/1.1	0.7/1.4		0.3/1.1	0.7/1.1	
KIT	Band 8	−0.3/1.2	0.6/1.2	130	−0.3/0.9	0.5/0.9	158
	Band 9	−0.1/1.2	0.5/1.2		−0.0/0.9	0.5/0.9	
Gobabeb	Band 8	−0.3/0.7	0.4/0.7	87	−0.3/0.5	0.4/0.6	136
	Band 9	0.0/0.7	0.3/0.7		−0.2/0.5	0.3/0.5	
Lake Constance	Band 8	−0.4/0.9	0.6/1.1	54	−0.4/1.0	0.6/1.2	60
	Band 9	−0.4/0.8	0.6/1.0		−0.3/1.0	0.4/1.1	
HYY	Band 8	−0.9/1.2	1.2/1.6	112	−1.0/1.2	1.3/1.6	113
	Band 9	−0.7/1.1	1.0/1.4		−0.7/1.1	1.0/1.4	
PUE	Band 8	−0.2/1.3	0.7/1.3	152	−0.2/1.3	0.7/1.3	166
	Band 9	0.0/1.2	0.7/1.2		0.1/1.2	0.7/1.2	
ROB	Band 8	−0.2/1.5	0.7/1.5	49	−0.2/1.5	0.8/1.6	65
	Band 9	−0.2/1.3	0.5/1.3		−0.2/1.3	0.8/1.4	
SVA	Band 8	−0.6/1.3	1.0/1.4	118	−0.7/1.5	1.1/1.6	116
	Band 9	−0.5/1.3	0.9/1.4		−0.5/1.4	1.0/1.5	
Total	Band 8	−0.3/1.2	0.8/1.3	924	−0.3/1.2	0.8/1.3	1023
	Band 9	−0.2/1.2	0.7/1.3		−0.1/1.2	0.7/1.2	

both bands, that coefficient was included only to show linear consistencies. Moreover, we did not find any evidence of sensor drift over time nor seasonal dependence of the calibration residuals for any of the test sites during the period of time considered in the study.

The results for Sentinel-3B were similar. These data are also shown in Table 4, where median and R-RMSD values are given for the Sentinel-3A and 3B data for each ground site. The medians (R-RMSDs) for both sensors differed by 0.3 K (0.2 K) or less for all the sites in the two bands, thus showing a consistent performance.

In Table 4 median biases range from −1.0 K for HYY (forest) in band 8 and −0.7 in band 9, to around 0.3 K for the rice paddy, with KIT, Gobabeb and Lake Constance giving values close to zero. The forest sites showed certain variability in median values, PUE and ROB yielding negligible biases in both bands. Consequently, the forest R-RMSDs ranged from values around 1.0–1.3 K in HYY to 0.5–0.8 K in PUE and ROB. This fact can be observed in the larger dispersion of forest data in Figure 3. Forests are complex and varied biomes, so ground LST and emissivity measurements are a difficult task, which may explain in part such variability. Moreover, the uncertainties in emissivities could be larger than those assumed in Section 3.2 because of the forest biome complexity. The lowest R-RMSDs were obtained for the Gobabeb site (0.4 K and 0.3 K for bands 8 band 9 of both sensors, respectively). Taking all the sites together,

Table 4 shows small median biases (between −0.1 and −0.3 K) and R-RMSD of 0.8 K for band 8 and 0.7 K for band 9. The R-RMSD values of Table 4 are compatible with the uncertainty analysis of Section 3.2 for most sites in the two bands and sensors, thus suggesting a good performance of the SLSTR TIR bands calibration.

R-RMSDs results for the shrubland site are larger due to random uncertainty at daytime and nighttime for both satellites. This may indicate difficulties to obtain ground data (few meters) that are representative of the satellite data (1 km). This drawback can occur especially in forest biomes where the mixture of different components (end-members) within the view of the radiometers installed on the mast may not represent the satellite view.

4.2. T-based results

Table 5 shows the median/SD and R-RMSD/RMSD for the SLSTR product and the SWA. The biomes are the same as those in Figure 3, and the results for Sentinel-3A and 3B, daytime, nighttime and altogether are shown. In the case of water, only daytime values were available because the ferry on which the radiometers are installed does not sail at nighttime overpass times. The different forest and shrubland sites were grouped together in one single forest biome because the median and R-RMSDs were very similar in all the cases.

According to Table 5, the median biases and R-RMSDs were quite similar for the Sentinel-3A and 3B LSTs obtained from each algorithm for all the biomes, showing again good consistency between both sensors. The SLSTR product biases exceeded the ±1.0 K threshold for desert and irrigated croplands at daytime and nighttime, with negative biases for desert (−2.1 for Sentinel-3A and −2.2 K for 3B) and positive biases for irrigated croplands (1.6 and 2.0 K). For water and forest, the SLSTR product showed much smaller biases (within ±0.3 K). The alternative SWA only yields an excessive bias for desert at daytime (−1.8 K for Sentinel-3A and −1.7 for 3B), and the other biomes are generally within ±0.6 K for most cases.

The R-RMSDs in Table 5 show very similar values for the SLSTR product and the SWA over water (0.7 K for Sentinel-3A and 0.9 K for Sentinel-3B) and forest sites (lower than 1.5 K). However, the SLSTR product shows larger R-RMSDs over desert (2.2 K for Sentinel-3A and 2.3 K for 3B) and irrigated croplands (1.8 K and 2.2 K), daytime and nighttime data combined. In contrast, the alternative SWA yielded lower R-RMSDs for the desert (1.6 K and 1.3 K for Sentinel-3A and 3B, respectively) and irrigated croplands (0.9 K and 0.8 K). The SLSTR algorithm considers one single set of coefficients for the desert biome (regardless of the soil composition, texture and roughness) and one single set for the irrigated cropland biome (regardless of

Table 5. Median/SD values and R-RMSD/RMSDs (in K) of the differences between satellite LSTs and ground LSTs (t-based validation) for each biome, for the SLSTR product and the SWA.

			SLSTR product		SWA	
			Sentinel-3A	Sentinel-3B	Sentinel-3A	Sentinel-3B
Lake	Day	Median/SD	0.3/1.0	0.2/1.1	−0.5/1.1	−0.7/1.1
		R-RMSD/RMSD	0.7/1.0	0.9/1.1	0.7/1.4	0.9/1.5
	Night	Median/SD	–	–	–	–
		R-RMSD/RMSD	–	–	–	–
	All	Median/SD	–	–	–	–
		R-RMSD/RMSD	–	–	–	-
Desert	Day	Median/SD	−2.5/0.9	−2.5/0.8	−1.8/0.8	−1.7/0.7
		R-RMSD/RMSD	2.6/2.5	2.6/2.6	1.9/1.9	1.9/1.9
	Night	Median/SD	−1.9/1.0	−1.9/0.8	−0.2/0.9	0.0/0.6
		R-RMSD/RMSD	2.0/2.0	2.0/2.0	0.6/0.9	0.6/0.6
	All	Median/SD	−2.1/1.0	−2.2/0.9	−1.0/1.1	−0.6/1.1
		R-RMSD/RMSD	2.2/2.3	2.3/2.3	1.6/1.5	1.3/1.3
Irrigated cropland	Day	Median/SD	1.1/1.5	0.8/1.6	−0.2/1.6	−0.3/1.5
		R-RMSD/RMSD	1.5/1.9	1.5/1.9	1.2/1.7	1.1/1.6
	Night	Median/SD	1.8/1.2	1.9/1.4	0.3/1.1	0.2/1.2
		R-RMSD/RMSD	1.9/2.1	2.1/2.1	0.8/1.1	0.4/1.2
	All	Median/SD	1.6/1.3	2.0/1.3	0.1/1.3	0.2/1.2
		R-RMSD/RMSD	1.8/2.0	2.2/2.1	0.9/1.3	0.8/1.2
Forest	Day	Median/SD	0.2/1.8	0.3/1.7	−0.4/1.7	−0.4/1.7
		R-RMSD/RMSD	1.4/1.9	1.4/2.0	1.3/2.0	1.4/1.8
	Night	Median/SD	−0.2/1.9	−0.1/1.8	−0.9/1.8	−0.7/1.7
		R-RMSD/RMSD	1.6/2.2	1.3/2.2	1.5/2.2	1.2/1.9
	All	Median/SD	0.0/1.8	0.1/1.8	−0.6/1.8	−0.6/1.6
		R-RMSD/RMSD	1.3/2.0	1.3/2.0	1.3/2.0	1.3/1.8

the soil and vegetation type and structure), which seems not to match the particular sites studied here. The SWA, which is based on emissivity estimates from each site, shows a better agreement with the ground LSTs in both sites.

For the two biomes with the largest biases (desert and irrigated croplands), Figure 4 shows the dependence of satellite minus ground LST differences on WVC and on ground LST (Sentinel-3A only, similar results for -3B). For irrigated croplands, the SLSTR product overestimates ground LSTs independently of WVC, while it largely underestimates ground LST for lower WVCs for the desert site. We can also see that for high temperatures, the bias is close to zero for the irrigated croplands, while it is the opposite for the desert. The SWA differences are closer to zero irrespective of WVC for desert and croplands, but they show negative biases for high ground LSTs, especially for the desert. This explains the large SWA daytime bias for the desert site mentioned above.

For irrigated croplands (rice paddy site), the results can be studied more in depth by taking into account the changes of land cover at the site due to the rice phenology, with three different surface types throughout the year (wet and dry bare soil, full vegetation and flooded soil). Table 6 shows the specific results for each land cover and for the SLSTR product and the SWA and for Sentinel-3A and 3B. Cover

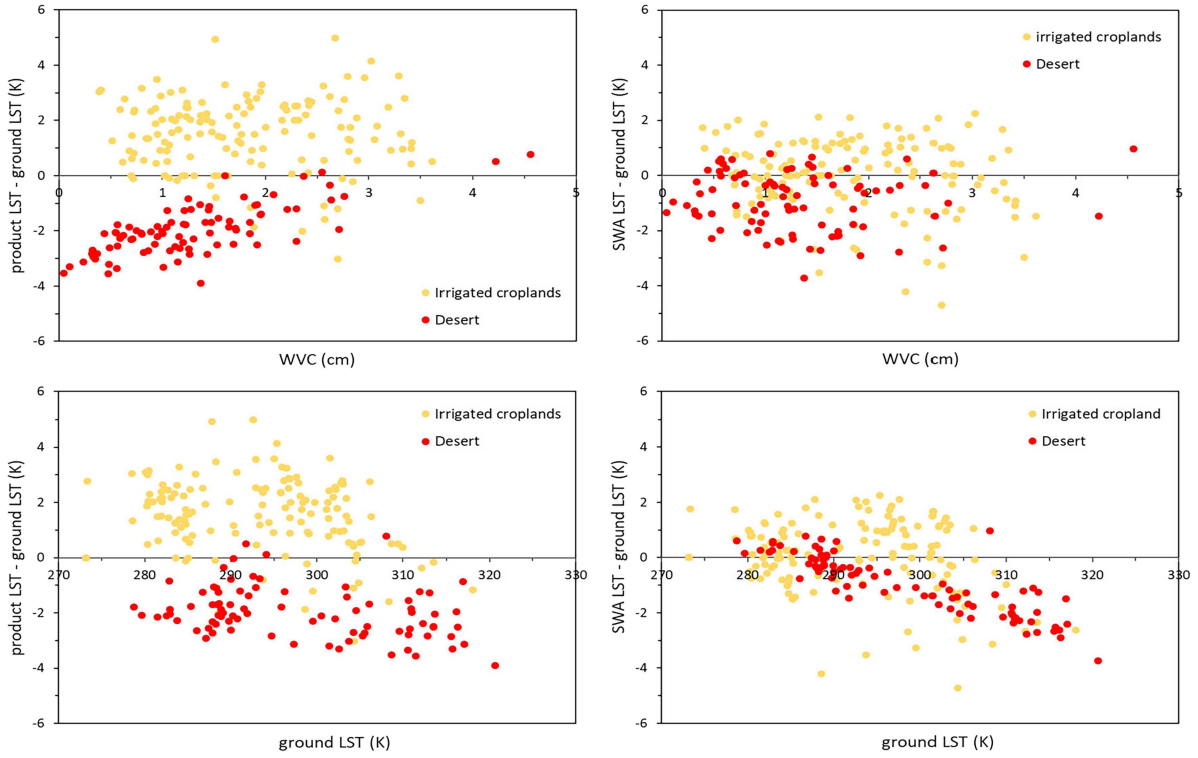

Figure 4. Satellite minus ground LST against atmospheric water vapor content (WVC) (top panels) and ground LST (bottom panels) for the SLSTR product (left panels) and the SWA (right panels) for the irrigated cropland (yellow dots) and desert biomes (red dots).

Table 6. Median/SD values and R-RMSD/RMSDs (in K) of the differences between satellite LSTs and ground LSTs (t-based validation, daytime and nighttime combined) for each land cover in the irrigated cropland site for the SLSTR product and the SWA.

		Sentinel-3A		Sentinel-3B	
		SLSTR product	SWA	SLSTR product	SWA
Full vegetation	Median/SD	1.6/1.4	0.3/1.5	1.8/1.5	0.4/1.5
	R-RMSD/RMSD	1.8/2.0	1.0/1.5	2.0/2.1	1.1/1.5
Bare soil	Median/SD	1.3/0.9	0.8/0.8	0.7/1.1	0.2/1.0
	R-RMSD/RMSD	1.5/1.6	1.0/1.1	1.0/1.4	0.7/1.0
Flooded soil	Median/SD	2.0/1.5	−0.2/1.3	2.4/1.1	0.0/0.6
	R-RMSD/RMSD	2.3/2.3	0.6/1.3	2.4/2.3	0.3/0.6

changes are not considered in the fixed biome assignment in the SLSTR product and this may be one of the reasons why the SWA, which uses dynamic emissivities depending on cover type, performed better for this particular biome for the three covers. It can be seen that for both satellites, very similar results are obtained.

Although the t-based dataset was limited to satellite viewing angles <40°, we checked for possible residual anisotropy effects by comparing the results for two subsets (0°–20° and 20°–40°) in the sparse vegetation and rainfed croplands biomes, which are prone to LST directionality. In both biomes, no significant differences were observed with respect to the total data, with maximum variations of 0.2 K. Moreover, we found no evidence of seasonal dependence of t-based validation residuals at any of the sites with the SLSTR product or SWA during the study period.

4.3. R-based results

Table 7 shows the median/SD and R-RMSD/RMSD values from the r-based validation, with two additional biomes, namely, sparse vegetation (olive orchard) and rainfed croplands (vineyard). In addition, the water biome now includes the Saline and Albufera sites, and the forest biome includes the SOF site. The total number of valid matchups is 1978 for Sentinel-3A and 2190 for Sentinel-3B. Table 7 also shows the fraction (in %) of valid cases passing the δ condition for each biome, day and night, and all together. In general, valid cases were more than 70% for all biomes. The only exceptions were the sparse vegetation and rainfed croplands sites (61% and 46%, respectively). The number of valid cases was similar between the daytime and nighttime, and for the whole LST range of each site. However, the number of valid cases decreased slightly with atmospheric humidity. These results were similar to those obtained by Coll et al. (2024). We tested different δ values to check how the fraction of valid cases would change. For $\delta = \pm 0.5$ K (± 1.0 K) it

Table 7. Median/SD values and R-RMSD/RMSDs (in K) of the differences between satellite LSTs and r-based LSTs (r-based validation) for each biome, for the SLSTR product and the SWA. The fraction (in %) of valid cases in each biome is given in parenthesis.

			SLSTR product		SWA	
			Sentinel-3A	Sentinel-3B	Sentinel-3A	Sentinel-3B
Lake	Day (71)	Median/SD	0.6/0.5	0.6/0.5	−0.2/0.4	−0.1/0.4
		R-RMSD/RMSD	0.8/0.9	0.8/0.8	0.4/0.5	0.4/0.5
	Night (68)	Median/SD	0.7/0.6	0.7/0.6	−0.1/0.4	−0.1/0.5
		R-RMSD/RMSD	0.8/0.9	0.9/0.9	0.4/0.5	0.4/0.5
	All (70)	Median/SD	0.7/0.5	0.6/0.6	−0.1/0.4	−0.1/0.4
		R-RMSD/RMSD	0.9/0.9	0.9/0.9	0.4/0.5	0.4/0.5
Desert	Day (84)	Median/SD	−2.2/0.7	−2.3/0.7	−1.6/0.6	−1.5/0.5
		R-RMSD/RMSD	2.3/2.3	2.4/2.4	1.7/1.7	1.6/1.7
	Night (88)	Median/SD	−1.3/0.7	−1.4/0.8	0.4/0.4	0.5/0.5
		R-RMSD/RMSD	1.5/1.4	1.6/1.5	0.5/0.5	0.7/0.7
	All (87)	Median/SD	−1.7/0.9	−1.7/0.9	−0.1/1.1	0.0/1.1
		R-RMSD/RMSD	1.8/1.8	1.8/1.3	1.2/1.2	1.3/1.2
Irrigated cropland	Day (67)	Median/SD	1.5/0.9	1.5/1.0	−0.2/0.7	−0.1/0.8
		R-RMSD/RMSD	1.6/1.8	1.5/1.8	0.4/0.7	0.4/0.8
	Night (70)	Median/SD	1.7/0.9	1.6/0.9	0.0/0.5	−0.1/0.4
		R-RMSD/RMSD	1.8/1.8	1.8/1.9	0.3/0.5	0.2/0.4
	All (69)	Median/SD	1.6/0.9	1.5/1.0	−0.1/0.5	−0.1/0.6
		R-RMSD/RMSD	1.7/1.8	1.7/1.9	0.3/0.5	0.3/0.6
Forest	Day (87)	Median/SD	0.2/0.8	0.4/1.0	−0.3/0.5	−0.2/0.5
		R-RMSD/RMSD	0.8/0.9	0.8/1.2	0.5/0.5	0.4/0.6
	Night (88)	Median/SD	0.5/0.7	0.7/0.8	−0.2/0.4	−0.1/0.3
		R-RMSD/RMSD	0.7/0.9	0.9/1.1	0.3/0.4	0.3/0.4
	All (87)	Median/SD	0.4/0.8	0.5/0.9	−0.2/0.4	−0.2/0.4
		R-RMSD/RMSD	0.7/0.9	0.8/1.1	0.3/0.5	0.3/0.5
Sparse vegetation	Day (53)	Median/SD	0.2/0.5	0.6/0.6	0.1/0.8	0.0/1.0
		R-RMSD/RMSD	0.5/0.6	0.8/0.8	1.0/0.9	0.9/1.0
	Night (70)	Median/SD	0.5/0.5	0.6/0.5	−0.1/0.4	−0.1/0.4
		R-RMSD/RMSD	0.6/0.7	0.8/0.8	0.3/0.4	0.3/0.4
	All (61)	Median/SD	0.3/0.5	0.6/0.5	−0.1/0.7	−0.1/0.7
		R-RMSD/RMSD	0.6/0.6	0.9/0.8	0.9/0.7	0.6/0.7
Rainfed cropland	Day (34)	Median/SD	1.1/0.5	1.2/0.7	0.0/0.6	0.3/0.8
		R-RMSD/RMSD	1.2/1.2	1.4/1.5	0.6/0.6	1.1/0.9
	Night (58)	Median/SD	0.8/0.6	0.8/0.7	−0.1/0.4	0.0/0.4
		R-RMSD/RMSD	1.0/1.0	1.2/1.1	0.3/0.4	0.3/0.4
	All (46)	Median/SD	1.0/0.6	0.9/0.7	−0.1/0.5	0.0/0.6
		R-RMSD/RMSD	1.1/1.1	1.2/1.2	0.4/0.5	0.5/0.6

would be 55% (80%), so the δ threshold used is a good compromise between sufficient accuracy and a significant number of data.

Results were similar for Sentinel-3A and 3B, except for SWA over rainfed cropland at daytime (R-RMSD of 0.6 K for 3 A and 1.1 K for 3B). In general, the mean biases and R-RMSDs were comparable to those from the t-based validation, with the SLSTR product showing mean biases around −1.5 K for desert and 1.5 K for irrigated croplands and the SWA yielding a mean bias of −1.5 K for desert at daytime. The SWA biases were close to zero for all the other cases. The low biases and R-RMSDs shown by the SWA may be due to the fact that it uses the same emissivity database (Table 3) as the r-based method. For the SLSTR product, the R-RMSDs were around 2.0 K for desert and irrigated croplands and above the 1.0 K accuracy threshold for rainfed croplands. For the SWA, only the desert at daytime (R-RMSD > 1.5 K) exceeded the accuracy requirement.

As for t-based validation, Figure 5 shows satellite minus r-based LST differences against WVC and r-based LST for desert and irrigated croplands, but also for rainfed croplands. For desert and irrigated croplands results are consistent with Figure 4. The SLSTR product shows negative biases for low WVCs and high surface temperatures for desert, and positive biases irrespective of WVC and LST for irrigated croplands. This is also observed for rainfed croplands. SWA biases only show significant correlation with LST for desert, with large negative values for high temperatures, as shown in Figure 4. In this case, we can see two distinct clusters in the desert data with respect to WVC (top-left panel of Figure 5), one with slightly overestimated LSTs corresponding to nighttime, and the other one corresponding to underestimated daytime LSTs. The dependence of the SWA bias on the LST was only observed for the desert site, which requires further research to improve the algorithm for high LSTs. For irrigated and rainfed croplands, biases are small and independent of WVC and LST.

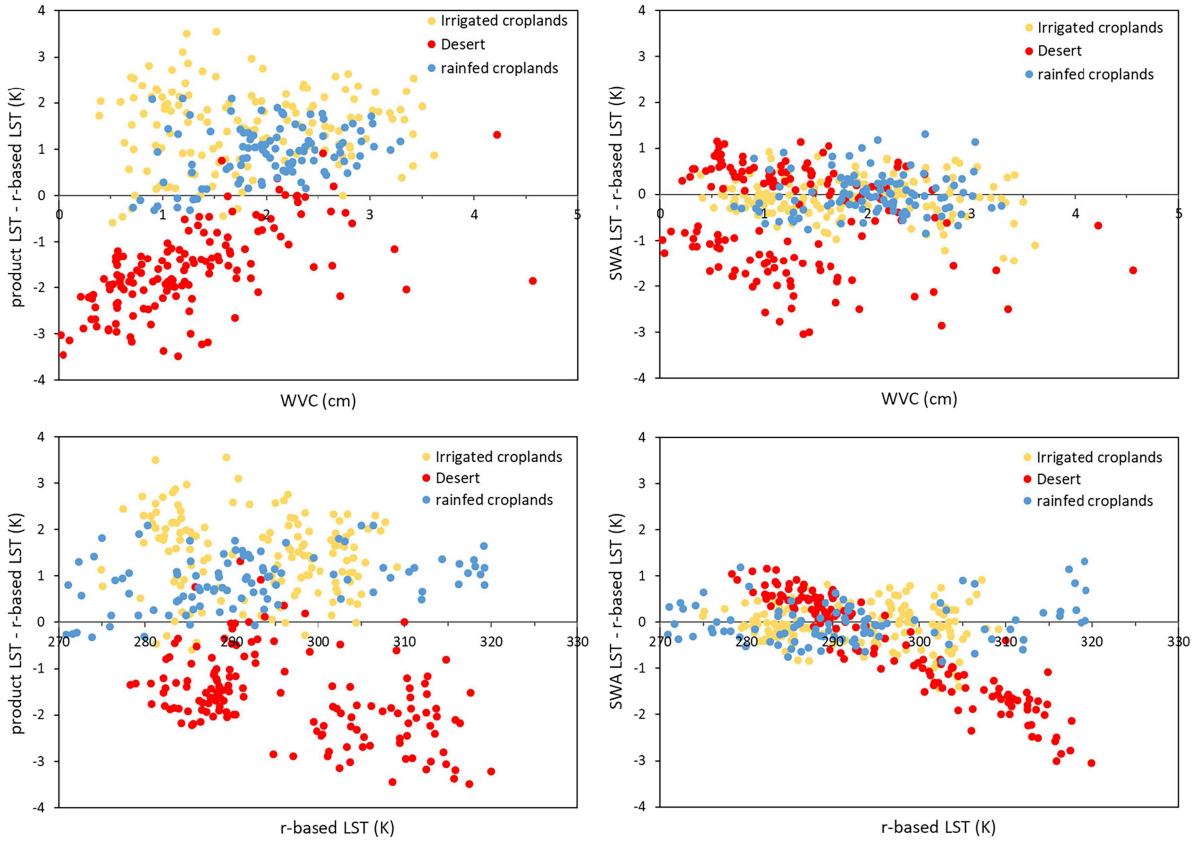

Figure 5. Satellite minus r-based LST against atmospheric water vapor content (WVC) (top panels) and ground LST (bottom panels) for the SLSTR product (left panels) and the SWA (right panels) for irrigated cropland (yellow dots), desert (red dots) and rainfed croplands (blue dots).

4.4. Summary of results per biome

We summarized the previous t-based and r-based R-RMSDs by biome in Table 8. For the lake biome, the SLSTR product showed similar results for t-based and r-based validation (0.8 K and 0.9 K, respectively). The SWA presented comparable errors for the t-based validation (0.8 K) and smaller values for the r-based validation (0.4 K).

For the forest biome, the t-based validation yielded similar R-RMSDs for both algorithms (1.3 K). The r-based validation showed uncertainties of 0.7 K for the SLSTR product and 0.3 K for SWA. For the desert biome, the SLSTR product showed R-RMSDs about 2 K for both t-based and r-based validation, while they were lower for the SWA (1.4 K). For the irrigated croplands biome, the SWA presented better results than SLSTR product for t-based validation (0.9 K and 2.0 K, respectively) and r-based validation (0.3 K and 1.7 K, respectively). Of all the biomes, the irrigated croplands showed the greatest differences between the algorithms in both validations.

For the sites with r-based validation only, the SWA and SLSTR product showed similar results (0.7 K) for the sparse vegetation biome. However, for rainfed croplands, the SLSTR product presented a larger R-RMSD (1.1 K) than the SWA (0.4 K).

5. Comparison with previous validation studies

We can compare the validation results of the present study with previously published results. Yang et al. (2020) obtained large RMSDs above 4 K for vegetated areas and 5 K for deserts. They showed t-based validations of the SLSTR product (Sentinel-3A only) in four sites equipped with pyrgeometers measuring hemispheric wide-band fluxes over meadow, cropland, desert steppe, and desert biomes. However, they also used in situ LST from Lake Constance and Gobabeb Wind, which showed RMSDs of 0.7 K and 2.4 K, respectively, also used in this study and equipped with a Heitronics KT15.85 directional radiometer. Lake Constance and Gobabeb results were similar to those obtained here for the Sentinel-3A SLSTR product, with an R-RMSD of 0.7 K and a small bias for water (daytime data only) and an R-RMSD of 2.2 K for deserts (daytime and nighttime data combined). Similarly, Li et al. (2023) performed t-based validations of Sentinel-3A LSTs over the same meadow, cropland, and desert steppe sites as Yang et al. (2020) and a mixed forest site also equipped with pyrgeometers plus two additional sites (desert steppe and wetland) with ground LSTs derived from directional, narrow band radiometers. For all the sites together, Li et al. (2023) obtained RMSDs >3 K at daytime and >2 K at night for both the SLSTR product and the SWA of Pérez-Planells et al. (2021) studied here. These values were larger than the t-based R-RMSDs obtained in the present paper for Sentinel-3A, which ranged from 2.6 K (2 K) in the daytime (nighttime) for the SLSTR product and 1.9 K (0.6 K) for the SWA in the desert site and up to 0.7 K at daytime for both algorithms at the water site. The high RMSDs reported by Yang et al. (2020) and Li et al. (2023) can largely be attributed to the use of ground LST measurements taken with pyrgeometers and the thermal heterogeneity of the sites, than to issues with the LST algorithms.

Li et al. (2023) also performed r-based validations of the SLSTR product and the SWA (Sentinel-3A) for sites representing desert, forest and inland waters: for all sites together, the daytime (nighttime) RMSD was

Table 8. T-based and r-based validation R-RMSDs (in K) for each biome, for the SLSTR product and the SWA.

		t-based	r-based
Lake	SLSTR product	0.8	0.9
	SWA	0.8	0.4
Forest	SLSTR product	1.3	0.7
	SWA	1.3	0.3
Desert	SLSTR product	2.0	2.0
	SWA	1.4	1.4
Irrigated croplands	SLSTR product	2.0	1.7
	SWA	0.9	0.3
Sparse vegetation	SLSTR product	–	0.7
	SWA	–	0.7
Rainfed croplands	SLSTR product	–	1.1
	SWA	–	0.4

1.8 K (1.4 K) for the SLSTR product and 1.6 K (0.8 K) for the SWA, which are significantly smaller than their t-based validation results. For the desert sites in both the SLSTR product and the SWA, biases were negative in Li et al. (2023) and in the present work. Li et al. (2023) reported a daytime (nighttime) RMSD of 2.1 K (1.6 K) for the SLSTR product and 1.6 K (0.7 K) for the SWA, which were in close agreement with the R-RMSDs obtained here for the same biome (see Table 7). For inland water, Li et al. (2023) obtained RMSDs of 1.2 K (1.3 K) for daytime (nighttime) for the SLSTR product and of 1.6 K (0.9 K) for SWA, which were larger than the R-RMSDs obtained here (0.9 K for the SLSTR product and 0.4 K for the SWA, nighttime and daytime).

R-based validation of the SLSTR product from Li et al. (2023) for forest showed RMSDs of 0.7 K (0.8 K) in the daytime (nighttime), which is in close agreement with Table 7. For the SWA, they report a bias of 1.1 K and R-RMSD of 1.6 K in the daytime (0.3 K and 0.8 K, respectively, in nighttime), which were significantly larger than those obtained here. The total number of matchups used by Li et al. (2023) was only 148 while we used 873. Owing to the inherent difficulty of ground LST measurements, the forest biome is usually misrepresented in t-based validation studies because directional effects may be large in such biome and radiometers are difficult to install at forest sites to collect data comparable with satellite measurements. One exception is the KIT forest site, which has been used in previous studies (Dave et al. 2025; Pérez-Planells and Göttsche 2023).

Therefore, SWA should be used at the regional scale or in areas where surface emissivities can be known with good accuracy. However, the authors recommend using the SLSTR product for global-scale LST studies, as it does not require any information on surface emissivity.

6. Conclusion

In order to check the on-board calibration of split-window bands 8 and 9 of the SLSTR on satellites Sentinel-3A and 3B, we collected a database of directional, narrow band radiometric in situ LSTs from nine sites representing different biomes (water, desert, irrigated croplands and forests). Using ground LSTs, emissivities and atmospheric data, BTs were simulated in both bands and compared with concurrent satellite BTs. We found a good agreement between the simulated and satellite BTs for both SLSTR instruments, with a mean bias (R-RMSD) of −0.3 K (0.8 K) for band 8 and −0.2 K (0.7) for band 9. Such differences are within the uncertainty limits of the simulated BTs, showing that both the bands and sensors are operating consistently in an unbiased and accurate manner.

The in situ LST database was also used for the t-based validation of the operational SLSTR product and an emissivity-dependent SWA. It should be noted that the calibration assessment shares part of the ground LST uncertainty (calibration of ground radiometers, emissivity and variability in ground temperature, Section 3.1). However, we did not apply any recalibration of the SLSTR data based on the ground LSTs, so the t-based results are not influenced by the calibration assessment, and both analyses are independent from each other.

Again, the validation results were mostly similar for Sentinel-3A and 3B. The mean biases of the SLSTR product did not meet the ±1.0 K accuracy threshold recommended by the GCOS for desert (−2 K) and irrigated croplands (1.5 K) neither at daytime nor at nighttime. In contrast, the SWA showed excessive bias (−1.7 K) only for the desert site in the daytime. For water and forest, both LST algorithms yielded mean biases within ±1.0 K and R-RMSDs of 1.5 K. For all the sites combined, the daytime (nighttime) R-RMSD of the SLSTR product was 1.5 K (1.7 K), while it was 1.3 K (1.3 K) for the SWA.

A significantly larger dataset was available for the r-based validation, which additionally to these nine t-based sites with four biomes, also included five sites with two new biomes (rainfed croplands and sparse vegetation). Results showed mean biases exceeding the GCOS threshold for the SLSTR product for desert and irrigated croplands both at daytime and nighttime. The SWA bias exceeded the threshold for the desert site at daytime, and was close to zero for all other cases. The emissivity of desert surface is typically more variable than that of vegetated areas owing to changes in composition, moisture, texture, etc., so there might be larger errors in the emissivity-dependent SWA. This is also true for the SLSTR product since only one desert class is considered. Irrigated croplands are typically dynamic and show larger variability during the growing season. As shown by Pérez-Planells et al. (2021) and the present study, emissivity modeling through FVC throughout the year fits well with the phenological variability of rice crops. Therefore, it appears that the constant coefficients assigned by the SLSTR algorithm are not correctly allocated in this case. Recently, Ghent et al. (2024) proposed to subdivide the irrigated croplands biome of the SLSTR product into three distinct biomes with different coefficients. This fact agrees with the t-based validation

results and may indicate problems with the algorithms for those particular cases rather than issues with the in situ data. The SLSTR product showed a mean bias around 1.0 K for rainfed croplands. The overall daytime (nighttime) R-RMSDs were 1.0 K (1.0 K) for the SLSTR product and of 0.6 K (0.3 K) for the SWA, which are lower than those for the t-based validation.

While the t-based method is the reference method for satellite LST validation, representative ground measurements of LST are difficult to obtain and may be affected by instrumental issues. TIR directional, narrow band radiometers are preferred over hemispheric, wide band pyrgeometers since the former represent more closely satellite measurements. Also, ground LST measurements are influenced by the inherent thermal heterogeneity of land surfaces at sub-pixel scale, and by high-frequency temporal fluctuations at the ground instrument scale due to atmospheric turbulence (Lagouarde, Irvine, and Dupont 2015). In addition, there may be directional effects between ground and satellite LSTs, especially for partially vegetated surfaces in the daytime, where angular variations of LST may be larger than 1 K (Ermida et al. 2014). These issues limit the application of the t-based method to single-component biomes in thermally homogeneous areas. The r-based method provides an alternative for biomes where representative in situ measurements are not feasible but require surface emissivity and atmospheric transmittances and radiances in the 10–12.5 µm window band.

The validation results of this study suggest that the SWA shows better agreement with the ground LSTs than the SLSTR product, provided that the emissivity in SLSTR bands 8 and 9 is known. However, this requirement is difficult to accomplish at the global scale and therefore the SLSTR product can be used in cases where emissivity is unknown. The present study contributes to the validation of LSTs derived from the SLSTR. The improvements that the work presents were the use of ground LST measured with narrow-band radiometers similar to the thermal bands used by the SLSTR sensor and the evaluation of the calibration of these bands over different types of study surfaces. Continued studies using the t-based and r-based validation methods are necessary for the global assessment of the accuracy and precision of LSTs from the Sentinel-3A and 3B satellites and to ensure the quality of the product in subsequent applications. As a future work, we consider using ERA5 atmospheric profiles, which have better spatial and temporal resolutions than NCEP and may have a positive effect on the r-based method by increasing the number of cases passing the δ condition and the accuracy of the LST calculations.

Author contributions

CRediT: **Jesús Puchades:** Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Writing – original draft; **Raquel Niclòs:** Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing – review & editing; **Lluís Pérez-Planells:** Conceptualization, Data curation, Methodology, Software, Writing – review & editing; **César Coll:** Investigation, Methodology, Supervision, Validation, Writing – review & editing; **Frank-Michael Göettsche:** Data curation, Visualization, Writing – review & editing; **Jose Antonio Valiente:** Data curation, Visualization; **Enric Valor:** Supervision, Visualization, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

We thank the European Space Agency for the project PP0097749, as part of the Sentinel-3 Validation Team (S3VT). This work was carried out within the framework of the project UV-INV_AE_4226564 of the own program of the University of Valencia Research Vice-Rectorate, Special Actions call. It was also carried out within the previos projects Tool4Extreme (PID2020-118797RB-I00), funded by the Spanish Ministerio de Ciencia e Innovación and the Agencia Estatal de Investigación (MCIN/AEI/10.13039/501100011033), and PROMETEU/2021/016, funded by the Conselleria d'Educació, Cultura, Universitats i Ocupació of the Generalitat Valenciana.

ORCID

Jesús Puchades 0000-0003-2205-6817
Raquel Niclòs 0000-0001-6630-7118
Lluís Pérez-Planells 0000-0003-1035-9757
César Coll 0000-0001-6665-3870

Frank-Michael Göettsche 0000-0001-5836-5430
Jose Antonio Valiente 0000-0003-4589-5496
Enric Valor 0000-0003-1144-1381

Data availability statement

All datasets used in this study are available from various open sources: Sentinel-3 data: https://browser.dataspace.copernicus.eu/; NCEP data: https://gdex.ucar.edu/datasets/d083002/dataaccess/; LST LAW data: https://law.acri-st.fr/data-access; LST University of Valencia data: https://doi.org/10.5281/zenodo.17465571 (Niclòs et al. 2025).

References

AghaKouchak, A., A. Farahmad, F. S. Melton, J. Teixeria, M. C. Anderson, B. D. Wardlow, and C. R. Hain. 2015. "Remote Sensing of Drought: Progress, Challenges and Opportunities." *Reviews of Geophysics* 53 (2): 452–480. https://doi.org/10.1002/2014RG000456.

Anderson, M. C., R. G. Allen, A. Morse, and W. P. Kustas. 2012. "Use of Landsat Thermal Imagery in Monitoring Evapotranspiration and Managing Water Resources." *Remote Sensing of Environment* 122: 50–65. https://doi.org/https://doi.org/10.1016/j.rse.2011.08.025.

Berk, A., G. P., Anderson, P. K., Acharya, and E. P., Shettle. 2008. MODTRAN5. 2.0. 0 user's manual. Spectral Sciences Inc., Burlington MA, Air Force Research Laboratory, Hanscom MA.

Bicheron, P. 2008. GLOBCOVER Products Description and Validation Report.

Carlson, T. 2007. "An Overview of the "triangle method" for Estimating Surface Evapotranspiration and Soil Moisture from Satellite Imagery." *Sensors* 7 (8): 1612–1629. https://doi.org/10.3390/s7081612.

Chang, S., H. Chen, B. Wu, E. Nasanbat, N. Yan, and B. Davdai. 2021. "A Practical Satellite-Derived Vegetation Drought Index for Arid and Semi-Arid Grassland Drought Monitoring." *Remote Sens (Basel)* 13: 414. https://doi.org/10.3390/rs13030414.

Coll, C., V. Caselles, E. Valor, and R. Niclòs. 2012. "Comparison between Different Sources of Atmospheric Profiles for Land Surface Temperature Retrieval from Single Channel Thermal Infrared Data." *Remote Sensing of Environment* 117: 199–210. https://doi.org/10.1016/j.rse.2011.09.018.

Coll, C., V. García-Santos, R. Niclòs, and V. Caselles. 2016. "Test of the MODIS Land Surface Temperature and Emissivity Separation Algorithm with Ground Measurements over a Rice Paddy." *IEEE Transactions on Geoscience and Remote Sensing* 54: 3061–3069. https://doi.org/10.1109/TGRS.2015.2510426.

Coll, C., R. Niclòs, J. Puchades, V. García-Santos, M. Perelló, and L. Pérez-Planells. 2024. "Demonstrating the Suitability of the Radiance- Based Method for Assessing the Accuracy of MODIS Land Surface Temperature Products." *IEEE Transactions on Geoscience and Remote Sensing* 62: 1–15. https://doi.org/10.1109/TGRS.2024.3454377.

Coll, C., V. Caselles, J. M. Galve, E. Valor, R. Niclòs, J. M. Sánchez, and R. Rivas. 2005. "Ground Measurements for the Validation of Land Surface Temperatures Derived from AATSR and MODIS Data." *Remote Sensing of Environment* 97: 288–300. https://doi.org/10.1016/j.rse.2005.05.007.

European Environment Agency. 2019. CORINE Land Cover 2018 (vector), Europe, 6-yearly - version 2020_20u1. https://doi.org/10.2909/71c95a07-e296-44fc-b22b-415f42acfdf0 https://sdi.eea.europa.eu/catalogue/copernicus/api/records/71c95a07-e296-44fc-b22b-415f42acfdf0?language=all.

Dave, J. A., V. Pandya, F.-M. Goettsche, H. K. Varchand, P. N. Parmar, D. D. Desai, D. B. Kardani, et al. 2025. "A Hybrid Physics-Based Method for Estimating Land Surface Temperature Using Radiative Transfer Simulations and Machine Learning Model from Sentinel-3A SLSTR Observations." *Earth Sci. Informatics* 18: 150. https://doi.org/10.1007/s12145-024-01530-0.

Ermida, S. L., I. F. Trigo, C. C. DaCamara, F. M. Göttsche, F. S. Olesen, and G. Hulley. 2014. "Validation of Remotely Sensed Surface Temperature over an Oak Woodland Landscape - the Problem of Viewing and Illumination Geometries." *Remote Sensing of Environment* 148: 16–27. https://doi.org/10.1016/j.rse.2014.03.016.

García-Santos, V., E. Valor, V. Caselles, C. Coll, and MÁ Burgos. 2014. "Effect of Soil Moisture on the Angular Variation of Thermal Infrared Emissivity of Inorganic Soils." *IEEE Geoscience and Remote Sensing Letters* 11: 1091–1095. https://doi.org/10.1109/LGRS.2013.2286747.

GCOS. 2022. The 2022 GCOS ECVs Requirements (GCOS 245).

Ghent, D. 2017. Title: S3 Validation Report-SLSTR.

Ghent, D., J. S. Anand, K. Veal, and J. Remedios. 2024. "The Operational and Climate Land Surface Temperature Products from the Sea and Land Surface Temperature Radiometers on Sentinel-3A and 3B." *Remote Sens (Basel)* 16: 3403. https://doi.org/10.3390/rs16183403.

Göttsche, F. M., and G. C. Hulley. 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. https://doi.org/10.1016/j.rse.2012.05.010.

Göttsche, F. M., F. S. Olesen, I. F. Trigo, A. Bork-Unkelbach, and M. A. Martin. 2016. "Long Term Validation of Land Surface Temperature Retrieved from MSG/SEVIRI with Continuous In-Situ Measurements in Africa." *Remote Sens (Basel)* 8: 410. https://doi.org/10.3390/rs8050410.

Göttsche, F.-M., F., Olesen, L., Poutier, S., Langlois, Wimmer, W., V., Garcia Santos, C., Coll, Niclos, R., Arbelo, M., and J. P., Monchau. 2018. Fiducial Reference Measurements for Validation of Surface Temperature from Satellites (FRM4STS), Report from the Field Inter-Comparison Experiment (FICE) for Land Surface Temperature (OFE-D130-LST-FICE-Report-V1-Iss-1-Ver-1).

Guillevic, P., F., Göttsche, J., Nickeson, G., Hulley, Ghent, D., Y., Yu, I., Trigo, et al. 2018. Land Surface Temperature Product Validation Best Practice Protocol. Version 1.1 58. https://doi.org/10.5067/doc/ceoswgcv/lpv/lst.001.

Hollmann, R., C. J. Merchant, R. Saunders, C. Downy, M. Buchwitz, A. Cazenave, E. Chuvieco, et al. 2013. "The ESA Climate Change Initiative: Satellite Data Records for Essential Climate Variables." *Bulletin of the American Meteorological Society* 94: 1541–1552. https://doi.org/10.1175/BAMS-D-11-00254.1.

Hook, S.J., G. Chander, J.A. Barsi, R.E. Alley, A. Abtahi, F.D. Palluconi, B.L. Markham, R.C. Richards, S.G. Schladow, and D.L. Helder. 2004. "In-flight validation and recovery of water surface temperature with Landsat-5 thermal infrared data using an automated high-altitude lake validation site at Lake Tahoe." IEEE Transactions on Geoscience and Remote Sensing 42(12): 2767–2776. 2767-2776. https://doi.org/10.1109/tgrs.2004.839092.

Hu, T., K. Mallick, G. C. Hulley, L. P. Planells, F. M. Göttsche, M. Schlerf, P. Hitzelberger, et al. 2022. "Continental-Scale Evaluation of Three ECOSTRESS Land Surface Temperature Products over Europe and Africa: Temperature-Based Validation and Cross-Satellite Comparison." *Remote Sensing of Environment* 282: 113296. https://doi.org/10.1016/j.rse.2022.113296.

Hulley, G. C., and S. J. Hook. 2012. "A Radiance-Based Method for Estimating Uncertainties in the Atmospheric Infrared Sounder (AIRS) Land Surface Temperature Product." *Journal of Geophysical Research: Atmospheres* 117(D20117). https://doi.org/10.1029/2012JD018102.

Hulley, G., S. Shivers, E. Wetherley, and R. Cudd. 2019. "New ECOSTRESS and MODIS Land Surface Temperature Data Reveal Fine-Scale Heat Vulnerability in Cities: A Case Study for Los Angeles County, California." *Remote Sens (Basel)* 11: 2136. https://doi.org/10.3390/rs11182136.

Hulley, G. C., F. M. Gottsche, G. Rivera, S. J. Hook, R. J. Freepartner, M. A. Martin, K. Cawse-Nicholson, and W. R. Johnson. 2021. "Validation and Quality Assessment of the ECOSTRESS Level-2 Land Surface Temperature and Emissivity Product." *IEEE Transactions on Geoscience and Remote Sensing : A Publication of the IEEE Geoscience and Remote Sensing Society* 60: 1–23. https://doi.org/10.1109/TGRS.2021.3079879.

Krishnan, P., T. P. Meyers, S. J. Hook, M. Heuer, D. Senn, and E. J. Dumas. 2020. "Intercomparison of in Situ Sensors for Ground-Based Land Surface Temperature Measurements." *Sensors (Switzerland)* 20: 1–26. https://doi.org/10.3390/s20185268.

Lagouarde, J. P., Y. H. Kerr, and Y. Brunet. 1995. "An Experimental Study of Angular Effects on Surface Temperature for Various Plant Canopies and Bare Soils." *Agricultural and Forest Meteorology* 77: 167–190. https://doi.org/10.1016/0168-1923(95)02260-5.

Lagouarde, J.-P., M. Irvine, and S. Dupont. 2015. "Atmospheric Turbulence Induced Errors on Measurements of Surface Temperature from Space." *Remote Sensing of Environment* 168: 40–53. https://doi.org/10.1016/j.rse.2015.06.018.

Li, Z. L., T. Bo-Hui, H. Wu, H. Ren, G. Yan, Z. Wan, I. Trigo, and J. A. Sobrino. 2013. "Satellite-Derived Land Surface Temperature: Current Status and Perspectives." *Remote Sensing of Environment* 131: 14–37. https://doi.org/10.1016/j.rse.2012.12.008.

Li, R., H. Li, L. Sun, Y. Yang, T. Hu, Z. Bian, B. Cao, Y. Du, and Q. Liu. 2020. "An Operational Split-Window Algorithm for Retrieving Land Surface Temperature from Geostationary Satellite Data: A Case Study on Himawari-8 AHI Data." *Remote Sens (Basel)* 12: 2613. https://doi.org/10.3390/rs12162613.

Li, R., H. Li, T. Hu, Z. Bian, F. Liu, B. Cao, Y. Du, L. Sun, and Q. Liu. 2023. "Land Surface Temperature Retrieval from Sentinel-3A SLSTR Data: Comparison Among Split-Window, Dual-Window, Three-Channel, and Dual-Angle Algorithms." *IEEE Transactions on Geoscience and Remote Sensing* 61: 1–14. https://doi.org/10.1109/TGRS.2023.3288584.

Liu, W., J. Shi, S. Liang, S. Zhou, and J. Cheng. 2022. "Simultaneous Retrieval of Land Surface Temperature and Emissivity from the FengYun-4A Advanced Geosynchronous Radiation Imager." *Int J Digit Earth* 15: 198–225. https://doi.org/10.1080/17538947.2021.2019844.

Ma, J., J. Zhou, T. Zhang, W. Tang, and Y. Wang. 2024. "A Comprehensive Validation Scheme for Satellite-Derived Land Surface Temperature Dataset." *IEEE Transactions on Geoscience and Remote Sensing* 62: 1–12. https://doi.org/10.1109/TGRS.2024.3488083.

Ma, J., J. Zhou, S. Liu, F. M. Göttsche, X. Zhang, S. Wang, and M. Li. 2021. "Continuous Evaluation of the Spatial Representativeness of Land Surface Temperature Validation Sites." *Remote Sensing of Environment* 265: 112669. https://doi.org/10.1016/j.rse.2021.112669.

Ma, J., J. Zhou, F.-M. Göttsche, Z. Wang, H. Wu, W. Tang, M. Li, and S. Liu. 2023. "An Atmospheric Influence Correction Method for Longwave Radiation-Based In-Situ Land Surface Temperature." *Remote Sensing of Environment* 293: 113611. https://doi.org/10.1016/j.rse.2023.113611.

Niclòs, R., J. A. Valiente, M. J. Barberà, and C. Coll. 2015. "An Autonomous System to Take Angular Thermal-Infrared Measurements for Validating Satellite Products." *Remote Sens (Basel)* 7: 15269–15294. https://doi.org/10.3390/rs71115269.

Niclòs, R., J., Puchades, S., Arribas, and J. A., Valiente. 2025. Brightness Temperature data from 2 sites (shrubland and rice paddy) during period 2021 to 2023. https://doi.org/10.5281/zenodo.17465571.

Niclòs, R., V. Caselles, E. Valor, C. Coll, and J. M. Sánchez. 2009. "A Simple Equation for Determining Sea Surface Emissivity in the 3-15 μm Region." *International Journal of Remote Sensing* 30: 1603–1619. https://doi.org/10.1080/01431160802541523.

Niclòs, R., J. M. Galve, J. A. Valiente, M. J. Estrela, and C. Coll. 2011. "Accuracy Assessment of Land Surface Temperature Retrievals from MSG2-SEVIRI Data." *Remote Sensing of Environment* 115: 2126–2140. https://doi.org/10.1016/j.rse.2011.04.017.

Niclòs, R., L. Pérez-Planells, C. Coll, J. A. Valiente, and E. Valor. 2018. "Evaluation of the S-NPP VIIRS Land Surface Temperature Product Using Ground Data Acquired By an Autonomous System at a Rice Paddy." *ISPRS Journal of Photogrammetry and Remote Sensing* 135: 1–12. https://doi.org/10.1016/j.isprsjprs.2017.10.017.

Niclòs, R., M. Perelló, J. Puchades, C. Coll, and E. Valor. 2023. "Evaluating Landsat-9 TIRS-2 Calibrations and Land Surface Temperature Retrievals against Ground Measurements Using Multi-Instrument Spatial and Temporal Sampling Along Transects." *International Journal of Applied Earth Observation and Geoinformation* 125: 103576. https://doi.org/10.1016/j.jag.2023.103576.

Pérez-Planells, L., and F.-M. Göttsche. 2023. "Combined Modelling of Annual and Diurnal Land Surface Temperature Cycles." *Remote Sensing of Environment* 299: 113892. https://doi.org/10.1016/j.rse.2023.113892.

Pérez-Planells, L., R. Niclos, E. Valor, and F.-M. Gottsche. 2022. "Retrieval of Land Surface Emissivities over Partially Vegetated Surfaces from Satellite Data Using Radiative Transfer Models." *IEEE Transactions on Geoscience and Remote Sensing* 60: 1–21. https://doi.org/10.1109/TGRS.2022.3224639.

Pérez-Planells, L., R. Niclòs, J. Puchades, C. Coll, F. M. Göttsche, J. A. Valiente, E. Valor, and J. M. Galve. 2021. "Validation of sentinel-3 Slstr Land Surface Temperature Retrieved By the Operational Product and Comparison with Explicitly Emissivity-Dependent Algorithms." *Remote Sens (Basel)* 13: 2228. https://doi.org/10.3390/rs13112228.

Polehampton, E., C., Cox, D., Smith, D., Ghent, M., Wooster, W., Xu, Bruniquel, J., C., Henocq, and S., Dransfeld. 2023. Title: Copernicus Sentinel-3 SLSTR Land User Handbook Filename Sentinel3-SLSTR-Land-Handbook-i1r3-14April2023.docx Copernicus Sentinel-3 SLSTR Land User Handbook Changes Log List of Changes Version Section Change.

Rubio, E., V. Caselles, and C. Badenas. 1997. "Emissivity Measurements of Several Soils and Vegetation Types in the 8-14 µm Wave Band: Analysis of Two Field Methods." *Remote Sensing of Environment* 59: 490–521. https://doi.org/10.1016/S0034-4257(96)00123-X.

Sánchez, J. M., R. López-Urrea, C. Doña, V. Caselles, J. González-Piqueras, and R. Niclòs. 2015. "Modeling Evapotranspiration in a Spring Wheat from Thermal Radiometry: Crop Coefficients and E/T Partitioning." *Irrigation Science* 33: 399–410. https://doi.org/10.1007/s00271-015-0476-2.

Smith, D., S. E. Hunt, M. Etxaluze, D. Peters, T. Nightingale, J. Mittaz, E. R. Woolliams, and E. Polehampton. 2021. "Traceability of the Sentinel-3 SLSTR Level-1 Infrared Radiometric Processing." *Remote Sens (Basel)* 13: 374. https://doi.org/10.3390/rs13030374.

Sobrino, J. A. 2015. SEN4LST: LST retrieval: Algorithm Theoretical Basis Document; European Space Agency, Contract 22900/09/NL/BJ; 2012; p. 25. https://doi.org/10.13140/RG.2.2.17177.83048.

Theocharous, E., E., Usadi, and N. P., Fox. 2010. CEOS comparison of IR brightness temperature measurements in support of satellite validation. Part I: Laboratory and Ocean surface temperature comparison of radiation thermometers.

Thome, K., F. Palluconi, T. Takashima, and K. Masuda. 1998. "Atmospheric Correction of Aster." *IEEE Transactions on Geoscience and Remote Sensing* 36: 1199–1211. https://doi.org/10.1109/36.701026.

Valor, E., and V. Caselles. 1996. "Mapping Land Surface Emissivity from NDVI: Application to European, African, and South American Areas." *Remote Sensing of Environment* 57: 167–184. https://doi.org/10.1016/0034-4257(96)00039-9.

Wan, Z., and Z. L. Li. 2008. "Radiance-Based Validation of the V5 MODIS Land-Surface Temperature Product." *International Journal of Remote Sensing* 29: 5373–5395. https://doi.org/10.1080/01431160802036565.

Wang, K., and S. Liang. 2009. "Global Atmospheric Downward Longwave Radiation over Land Surface under All-Sky Conditions from 1973 to 2008." *Journal of Geophysical Research: Atmospheres* 114(D19101). https://doi.org/10.1029/2009JD011800.

Wilrich, P. T. 2007. "Robust Estimates of the Theoretical Standard Deviation to Be Used in Interlaboratory Precision Experiments." *Accreditation and Quality Assurance* 12: 231–240. https://doi.org/10.1007/s00769-006-0240-7.

Yamada, Y., S., Harris, M., Hayes, R., Simpson, W., Wimmer, R., Holmes, T., Nightingale, et al. 2023a. Results from CEOS International Thermal Infrared Radiometer Inter-comparison (CRIC) Part 2 of 3: Laboratory Comparison of Radiometers.

Yamada, Y., S., Harris, M., Hayes, R., Simpson, W., Wimmer, R., Holmes, T., Nightingale, et al. 2023b. CEOS International Thermal Infrared Radiometer Comparison: Part I: Laboratory Comparison of Radiometers and Blackbodies. Submitted to Journal of Atmospheric and Oceanic Technology.

Yang, J., J. Zhou, F.-M. Göttsche, Z. Long, J. Ma, and R. Luo. 2020. "Investigation and Validation of Algorithms for Estimating Land Surface Temperature from Sentinel-3 SLSTR Data." *International Journal of Applied Earth Observation and Geoinformation* 91: 102136. https://doi.org/10.1016/j.jag.2020.102136.

Zhang, H., T. Hu, B.-H. Tang, K. Mallick, X. Zheng, M. Wang, A. Olioso, et al. 2025. "Deep Learning Coupled with Split Window and Temperature-Emissivity Separation (DL-SW-TES) Method Improves Clear-Sky High-Resolution Land Surface Temperature Estimation." *ISPRS Journal of Photogrammetry and Remote Sensing* 225: 1–18. https://doi.org/10.1016/j.isprsjprs.2025.04.016.