



LSA SAF Land Surface Temperature: Evolution and Validation of Version 2 suite of Products

Sofia L. Ermida^{1,2}, Emanuel Dutra^{1,2}, Francisco M. Lopes^{1,2}, Sandra Gomes¹, Frank-M. Göttsche³, Isabel F. Trigo^{1,2}

¹Instituto Português do Mar e da Atmosfera (IPMA), Lisbon, Portugal

²Instituto Dom Luiz (IDL), Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal

³Karlsruhe Institute of Technology, IMKASF, Eggenstein-Leopoldshafen, Germany

Correspondence to: Sofia L. Ermida (sofia.ermida@ipma.pt)

Abstract. The main objective of the Satellite Application Facility on Land Surface Analysis (LSA-SAF) is to provide Earth Observation data to support land, land-atmosphere and biosphere applications, with commitments to maintain operational near-real time products that are consistent with corresponding Climate Data Records (CDR) produced offline. The LSA-SAF has been providing estimates of Land Surface Temperature (LST) for over two decades, based on observations from EUMETSAT satellites. This article describes the new developments introduced in the generation of LST from EUMETSAT satellites by the LSA-SAF. The so-called version 2 suite of products includes LST from SEVIRI on-board the MSG series, for both the 0-degree (MLSTv2) and IODC (IODC MLSTv2) coverage, and from AVHRR on-board Metop satellites (EDLSTv2). It also includes an all-sky product (MLST-ASv2) that combines clear-sky LST from MLSTv2 with cloudy skin temperature from a surface energy balance model developed within the LSA-SAF (METv3). The main algorithm developments with respect to the older operational version include a revision of the calibration database used to train the LST retrieval algorithm and of the land surface emissivity algorithm. The sensors calibration was also revised to solve discontinuity issues. The overall accuracy of the LSA-SAF datasets, when compared to in-situ data, is 0.20 K, 0.21 K, -0.22 K and 0.24 K for MLSTv2, IODC MLSTv2, EDLSTv2 and MLST-ASv2, respectively. The LSA-SAF also aims to provide LST datasets that meet the GCOS requirements for climate applications. Analyses of the products' stability shows that all datasets meet the GCOS requirements, with exception of the IODC MLSTv2. The LST algorithms presented here have also been adapted and implemented for the generation of LST from the new Meteosat Third Generation (MTG) and Metop Second Generation (Metop-SG), ensuring continuity of the Meteosat and Metop services. Monthly averages from each product for the respective data record period on a 0.25°x0.25° regular grid are available from <https://doi.org/10.5281/zenodo.20705177> (Ermida et al., 2026). The full resolution data is available from the LSA-SAF data server at <https://datasasaf.lsasvcs.ipma.pt/>.

1. Introduction

The Satellite Application Facility on Land Surface Analysis (LSA-SAF) is part of the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) Application Ground Segment. Its main objective is to provide Earth Observation data to support land, land-atmosphere and biosphere applications, with commitments to maintain operational near-real time products that are consistent with corresponding Climate Data Records (CDR) produced offline (Trigo et al., 2011).



35 For over two decades, the LSA-SAF has been providing estimates of Land Surface Temperature (LST) based on
 observations from EUMETSAT satellites. LST corresponds to the radiative temperature of the surface, which controls the
 outgoing flux of longwave radiation and, therefore, is critical to study Earth's energy budget. LST can be directly estimated
 using passive sensors operating in the thermal infrared (TIR) region (Li et al., 2013) and most LST products rely on
 observations within the so-called TIR atmospheric window, where atmospheric transmittance is high. Nevertheless, LST
 40 algorithms must disentangle surface upwelling radiation from atmospheric contributions in the measured Top-Of-
 Atmosphere (TOA) signal. This requires accounting for effects of Land Surface Emissivity (LSE) on surface radiation and
 for the atmospheric absorbed as well as reflected and emitted components.

The LSA-SAF currently provides LST products derived from the Spinning Enhanced Visible and Infrared Imager (SEVIRI)
 on-board Meteosat Second Generation (MSG) and the Advanced Very High Resolution Radiometer (AVHRR) onboard
 45 Metop (EUMETSAT polar-orbiting satellites). In both cases, LST primarily relies on dual-channel TIR measurements to
 better estimate atmospheric contributions, while LSE is prescribed based on land cover and surface information retrieved
 from visible channels (e.g., vegetation and snow cover). The new version 2 suite of products relies on the same algorithm,
 namely the Generalized Split-Window (GSW) algorithm (Trigo et al., 2008; Wan and Dozier, 1996), but the calibration
 dataset used to train the algorithm has been updated to a more comprehensive description of surface and atmospheric
 50 conditions (Ermida and Trigo, 2022), which also ensures a good algorithm performance in more extreme conditions.
 Additionally, the LSE algorithm was revised to improve the representation of bare surfaces. For that purpose, the Vegetation
 Cover and Temperature-Emissivity Separation (V-TES) algorithm proposed by (Ermida et al., 2023) has been implemented,
 which combines two different retrieval methods, i.e., direct LSE retrieval and based on land cover classification and
 vegetation information, and maximizes the quality of the retrievals.

55 Following its recognition as an Essential Climate Variable (ECV)(Bojinski et al., 2014) by the Global Climate Observing
 System (GCOS) program, LST has been increasingly used in climate monitoring (Agathangelidis et al., 2022; de Almeida et
 al., 2021; Cheda et al., 2026; Liu et al., 2025; Martins et al., 2024). Satellite LST is particularly useful in complementing
 near surface air temperature from stations, since it can also be used to quantify Earth's warming while being available with
 much wider spatial coverage. One of the biggest challenge in using satellite data for climate analyses is that, to obtain the
 60 required long time-series, CDRs generally rely on a combination of multiple platforms. As such, when deriving stable
 datasets from satellite observations, it is necessary to address calibration issues throughout each sensor's life-cycle and for
 the transitions between the different satellites in the mission. If multiple missions need to be used, ensuring stability can be
 further complicated by changes in sensor configuration (e.g., spatial, temporal and spectral resolutions, available spectral
 bands, scan configuration). For the LSA-SAF v2 suite, CDRs are derived only with data from a single mission (i.e., either
 65 MSG/SEVIRI or Metop/AVHRR), which reduces the impact of satellite transitions. Nevertheless, it is important to ensure
 good calibration between the sensors to avoid either drifts, e.g., due to sensor degradation, or discontinuities between
 observations provided by different satellites of the same series. Here, the latest radiance calibration coefficients provided by
 EUMETSAT were used to minimize such issues.



In this paper we describe the most recent version 2 of the LSA-SAF LST CDRs: MLSTv2, based on MSG/SEVIRI; and EDLSTv2, based on Metop/AVHRR. Additionally, an all-sky LST product is derived from SEVIRI/MSG (MLST-ASv2): it combines clear-sky MLSTv2 data with cloudy-sky surface temperature estimates generated by a surface energy balance (SEB) model, which is closely forced with SEVIRI/MSG down-welling radiation, albedo and vegetation products. We provide a comprehensive validation of the different products using in-situ data, while assessing the products' consistency and stability through the comparisons with other LST CDRs.

2. Methods and Data

2.1. LSA-SAF LST Data Records

The LSA-SAF derives LST from observations provided by two main sensors: SEVIRI on board MSG satellites and AVHRR on board Metop satellites. Figure 1 shows the times of operation of each satellite used in the production of LSA-SAF LST products. The MSG series includes four geostationary satellites, which over the last two decades provided the so-called nominal service at 0° longitude covering Africa, Europe and part of South America. Starting operations in January 2004, MSG1 (also called Meteosat-8) was the primary satellite until April 2007, when MSG2 (Meteosat-9) took over the service until January 2013; then MSG3 (Meteosat-10) replaced the latter until February 2018, being then followed by MSG4 (Meteosat-11) until March 2023, which was then swapped back to MSG3. EUMETSAT use older satellites as backup in case the primary fails and MSG 1 and 2 have been allocated to provide the Indian Ocean Data Coverage (IODC), with MSG1 being reallocated to 41.5°E and providing data since 2018 and MSG2 at 45.5°E since July 2022. The EUMETSAT Polar System (EPS) Metop series is comprised of three satellites in sun-synchronous orbit with equatorial crossing times (ECT) at 9h30 AM/PM. Within LSA-SAF, LST production is based on observations from the prime satellite: Metop-A was the primary operational satellite from 2007 until 2015, when it was replaced by Metop-B. Although Metop-C is available since 2019, the team decided to maintain operational production from Metop-B due to calibration issues in the Metop-C AVHRR infrared channels (He et al., 2016). Starting in 2024, Metop-B is slowly deorbited to conserve fuel; this service will eventually be replaced by Metop - Second Generation (Metop-SG), with the first satellite of this series already in orbit since August 2025.

The sensor on the geostationary orbit, SEVIRI, provides data at a spatial resolution of 3 km at sub-satellite point and a temporal sampling frequency of 15 minutes, while AVHRR on the polar orbit satellites provides roughly 1-km spatial resolution and two overpasses per day (although higher temporal sampling is attained in polar regions). Table 1 presents the characteristics of the different LST products generated within LSA-SAF. The SEVIRI LST data is generated at full temporal resolution on the native geostationary grid. The AVHRR LST data are produced for each swath, which are then projected onto a 0.01° sinusoidal grid and aggregated into day and nighttime datasets. The datasets are also disseminated on regular latitude-longitude grids of 0.05° and 0.01°, respectively.

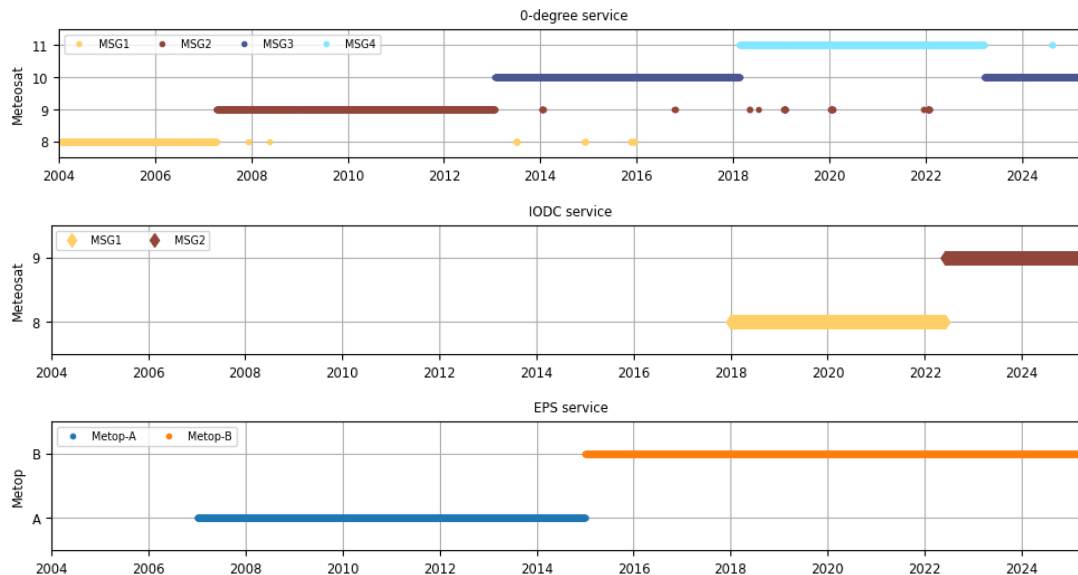


Figure 1 – Timeline of the use of each Meteosat Second Generation (0-degree and IODC service) and each Metop satellite in the production of the LSA-SAF LST products.

All LSA-SAF LST products are based on observations performed in the TIR atmospheric window, with the SEVIRI and AVHRR sensors both providing channels centered at 10.8 and 12 μm . To reduce discontinuities introduced by the different satellites in the series, prior to the generation of LSTs all radiances were re-calibrated. The SEVIRI radiance calibration coefficients were provided by EUMETSAT following the methodology of John et al. (2019). The respective AVHRR coefficients were obtained from (Coppens and Ackermann, 2014).

Table 1 – Characteristics of the LSA-SAF LST products.

Product acronym	Sensor/ Platforms	Spatial resolution and coverage	Temporal resolution and coverage	Other details
<i>MLSTv2</i>	SEVIRI/ MSG1-4	3 km; 0-degree disk (Europe, Africa, South America); native grid	15 min 2004 – 2024 NRT updates	Clear-sky only
<i>MLSTv2 [IODC]</i>	SEVIRI/ MSG1-2	3 km; IODC disk (Europe, Africa, Middle East); native grid	15 min 2018 – 2024 NRT updates	Clear-sky only
<i>MLST-ASv2</i>	SEVIRI/ MSG1-4	3 km; 0-degree disk (Europe, Africa, South America); native grid	30 min 2004 – 2024 NRT updates	All-sky: cloudy estimates from surface energy balance model
<i>EDLSTv2</i>	AVHRR/ Metop A-B	1 km; Global; sinusoidal grid	Twice-daily 2007 – 2024 NRT updates	Clear-sky only; day/night observations close to 9h30AM/PM local time



2.2. Auxiliary Datasets

The clear-sky LST retrieval algorithm used by the LSA-SAF requires estimates of atmospheric total column water vapour (TCWV), which was obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 reanalysis (Hersbach et al., 2020). ERA5 fields were projected onto each satellite's swath using bilinear interpolation. For each pixel, the estimates are matched to the satellite observation time through interpolation of the two adjacent (hourly) estimates.

Daily estimates of surface emissivity require information on fraction of vegetation cover (FVC) and snow cover (SC). Emissivity estimates for MLSTv2 and EDLSTv2 make use of the MSG Daily FVC (MDFVC) and EPS 10-days FVC (ETFVC) products, respectively, generated within the LSA-SAF (García-Haro et al., 2019, 2021). Snow cover for the 0-degree MSG disk is obtained from the SEVIRI Snow mask (SE-D-SEVIRI, H34) product generated by the Hydrology SAF (H-SAF). For Metop and IODC, we use the Snow Cover from the Interactive Multisensor Snow and Ice Mapping System (IMS) product generated by United States National Ice Center (USNIC) (U.S. National Ice Center, 2008). The IMS SC data was projected from the original 1km and 4km polar stereographic grids to the Metop and IODC swaths using a nearest neighbor approach.

Emissivity estimates also rely on information on land cover type, which is obtained from the Copernicus Global Land Service (CGLS) Land Cover product (Buchhorn et al., 2020), and on soil type, which is obtained from the United Nations Food and Agriculture Organization (FAO) and Educational, Scientific and Cultural Organization (UNESCO) Soil Map of the World (FAO-UNESCO, 2007).

Furthermore, when deriving emissivities of bare soil over deserts from direct retrievals (described in section 2.3.3), the Dust Aerosol Optical Depth (DuAOD) at 550 nm from the Copernicus Atmosphere Monitoring Service (CAMS) global reanalysis (EAC4)(Inness et al., 2019) is used to mask out retrievals with high aerosol concentrations.

Under clear-sky, the MLST-ASv2 LST is provided by the MLSTv2, while for cloudy conditions it uses surface temperature estimates provided by an Energy Balance Model under cloudy conditions (section 2.4) (Barrios et al., 2024; Martins et al., 2019). The model uses information on downwelling solar radiation (MDSSF) (Carrer et al., 2019), surface albedo (MDAL) (Carrer et al., 2018), FVC (MDFVC) and leaf area index (LAI; MDLAI) (García-Haro et al., 2019), which are all generated by the LSA-SAF. The model also ingests surface and root zone soil moisture (SSM and RZSM, respectively) provided by the H-SAF CDR H141/H142 (H SAF, 2020), and meteorological fields from ERA5. From 2021 (end of the CDR) onwards, the soil moisture information is obtained from the Near Real Time (NRT) H-SAF operational H26 product and meteorological fields from ECMWF operational short-range weather forecasts.



2.3. Clear-sky LST retrieval

2.3.1. The Generalized Split Window Algorithm

The LSA-SAF LST products rely on retrievals performed using the GSW algorithm (Trigo et al., 2008; Wan and Dozier, 1996):

$$LST = \left(A_1 + A_2 \frac{1-\bar{\varepsilon}}{\bar{\varepsilon}} + A_3 \frac{\Delta\varepsilon}{\bar{\varepsilon}^2} \right) \frac{BT_{10.8} + BT_{12.0}}{2} + \left(B_1 + B_2 \frac{1-\bar{\varepsilon}}{\bar{\varepsilon}} + B_3 \frac{\Delta\varepsilon}{\bar{\varepsilon}^2} \right) \frac{BT_{10.8} - BT_{12.0}}{2} + C, \quad (1)$$

where $BT_{10.8}$ and $BT_{12.0}$ are the brightness temperatures of the 10.8 μm and 12.0 μm channels, respectively. $\bar{\varepsilon}$ is the channels' average emissivity ($\bar{\varepsilon} = (\varepsilon_{10.8} + \varepsilon_{12.0})/2$) and $\Delta\varepsilon$ is the emissivity difference between the channels ($\Delta\varepsilon = \varepsilon_{10.8} - \varepsilon_{12.0}$). A_j , B_j , ($j = 1, 2, 3$) and C are algorithm coefficients that are calibrated for classes of atmospheric total column water vapour (TCWV; every 5 cm) and satellite view zenith angle (VZA; every 5°). The GSW algorithm is only applied to cloud-free pixels, with cloud detection being performed with software provided by the Nowcasting (NWC) SAF (<https://nwc-saf.eumetsat.int/>).

Additionally, LST uncertainties are generated on a pixel-by-pixel basis by propagating input uncertainties and considering the impact of retrieval conditions on product quality; the approach follows the same methodology of previous LST product versions, thoroughly described in Freitas et al. (2010) and corresponding Algorithm Theoretical Basis Documents (ATBD) available at the LSA-SAF website (<https://lsa-saf.eumetsat.int/en/data/products/land-surface-temperature-and-emissivity/>).

2.3.2. Algorithm Calibration Database

The algorithm coefficients were calibrated with simulations of TOA brightness temperatures as would be measured by each satellite. These were generated using the state-of-the-art MODTRAN 6 radiative transfer model (Berk et al., 2014). A newly developed database of atmospheric profiles and surface properties (Ermida and Trigo, 2022) was used as input to the model. The database employs dissimilarity criteria, as defined by Chevallier et al. (2000), and applies these to atmospheric profiles from the ERA5 reanalysis (Hersbach et al., 2020). This ensures that a wide range of atmospheric conditions are included, allowing a good representativeness of the algorithm globally. For each atmospheric profile, multiple surface conditions are selected taking into account the distribution of errors between satellite LST and ERA5 skin temperature, and the variability of emissivity values observed in different emissivity products. By varying the surface temperature and emissivities for each profile, we ensure that possible biases in ERA5 are not translated into the calibration database. The database was split into calibration and validation subsets, with 1/3 of the data being used for calibration and the remaining for obtaining the algorithm error. Figure 2 shows an example of the algorithm coefficients and root mean squared error (RMSE) as obtained for the SEVIRI onboard MSG3, highlighting the expected decrease in product accuracy with increasing view angle and atmospheric water vapour. More details on the compilation of the calibration database can be found in Ermida and Trigo (2022).

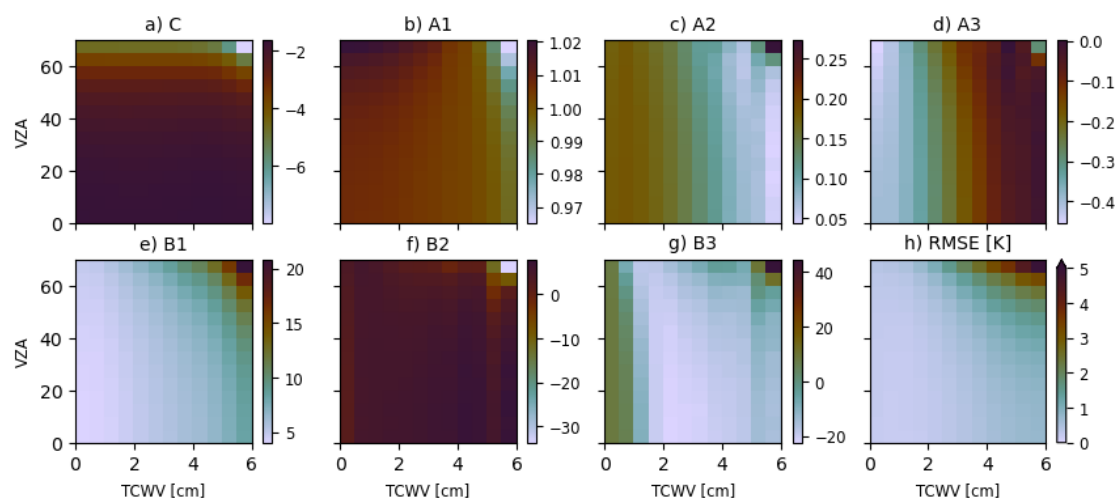


Figure 2 – Coefficients (a-g) and Root Mean Squared Error (RMSE in Kelvin; h) of the Generalized Split-window Algorithm (equation (1)), as function of view zenith angle (VZA; degrees) and total column water vapour (TCWV; cm), as obtained for the SEVIRI onboard the MSG3 satellite.

2.3.3. Land Surface Emissivity

Emissivity estimates are derived using the V-TES methodology introduced by Ermida et al. (2023), which combines the Vegetation Cover Method (VCM)(Caselles et al., 2012) with the Temperature-Emissivity Separation (TES) algorithm (Gillespie et al., 1998).

The VCM is used over vegetated surfaces, where the generally small variation in spectral emissivity limits the quality of TES retrievals. The VCM relies on information on land cover to attribute vegetation-type emissivities, and on soil type to attribute bare ground emissivities. Additionally, snow and water emissivities are attributed based on snow and water masks. The respective vegetation-type, bare ground, water and snow emissivities are obtained from spectral reflectance data from the ECOSTRESS Spectral Library version 1.0 (Meerdink et al., 2019). Each plant sample in the spectral library was attributed to a land cover based on species, and each rock/soil type was attributed to a FAO classification based on taxonomy. All used samples and their attributed classes are described in Tables A1 to A3 of Ermida et al. (2023).

Bare ground emissivities over bare and sparsely vegetated areas were then complemented with information obtained from a TES algorithm, modified to account for directional effects on emissivity (Ermida et al., 2024). Since TES requires a large number of radiative transfer simulations (per pixel and per channel), we employed the Radiative Transfer for TOVS (RTTOV) version 13 model, which is significantly faster than MODTRAN6. For the MLSTv2 products, the TES algorithm was applied to a full year of hourly data, while for the EDLSTv2 product a three-year period was used to compensate for the lower daily frequency of observations. In both cases, TES retrievals are limited to snow-free surfaces with a FVC below 20%, TCWV below 2 cm and dust AOD below 0.3, in order to obtain estimates with the best possible quality. The retrievals are then averaged over the respective period using the atmospheric transmittance as weight (Ermida et al., 2023). Global



195 TES retrievals for the EDLSTv2 product were derived from Visible Infrared Imaging Radiometer Suite (VIIRS) TOA measurements, since AVHRR does not provide the three TIR channels necessary for applying the TES. Band correction coefficients were then derived from the spectral library data to account for differences in the channel response functions. Likewise, band correction coefficients are also estimated to convert between different MSG sensors. Table 2 summarizes the data used in the TES retrievals for each LST product.

200

Table 2 – Details of the data used in TES retrievals for estimating the emissivities of each LSA-SAF LST product.

Product acronym	Sensor/ Platforms	Temporal coverage and resolution	Projection and spatial resolution	Auxiliary data
MLSTv2	SEVIRI/ MSG3	Mar 2023 – Feb 2024 1-hourly	MSG disk Native resolution	MDFVC; snow cover; CAMS Du-AOD; ERA5
MLSTv2 [IODC]	SEVIRI/ MSG1	Jan-Dec 2021 1-hourly	IODC disk Native resolution	MDFVC; snow cover; CAMS Du-AOD; ERA5
	SEVIRI/ MSG2	Mar 2023 – Feb 2024 1-hourly	IODC disk Native resolution	MDFVC; snow cover; CAMS Du-AOD; ERA5
EDLSTv2	VIIRS/ SNPP	Jan 2019 – Dec 2021 Twice-daily	Sinusoidal grid 1 km	ETFVC; snow cover; CAMS Du-AOD; ERA5

Finally, the effective emissivity of a pixel for a given channel, k , is computed as the weighted average of the vegetated and bare ground emissivities:

$$205 \quad \varepsilon_{k,eff} = \varepsilon_{k,veg} FVC + \varepsilon_{k,bg} (1 - FVC), \quad (2)$$

where $\varepsilon_{k,veg}$ and $\varepsilon_{k,bg}$ are the emissivities of the vegetated and bare ground components, FVC is the fraction of vegetation cover. If a pixel includes snow or water, then effective emissivity is obtained by averaging with the area fraction covered by snow or water:

$$\varepsilon_{k,eff} = \varepsilon_{k,eff} (1 - f_{snow}) + \varepsilon_{k,snow} f_{snow}, \quad (3)$$

$$210 \quad \varepsilon_{k,eff} = \varepsilon_{k,eff} (1 - f_{water}) + \varepsilon_{k,water} f_{water}, \quad (4)$$

where $\varepsilon_{k,snow}$ and $\varepsilon_{k,water}$ are the emissivities of snow and water, and f_{snow} and f_{water} are the respective fractions. An example of emissivities derived for each product on the 15th of March of 2024 is shown in Figure 3. More details on channel emissivity, including estimation of uncertainties, are available in Ermida et al. (2023).

2.4. All-sky LST estimates

215 The LSA-SAF MLST-ASv2 product relies on the combination of LST provided by the MLSTv2 product together with skin temperature (T_{skin}) estimated via a SEB model used within the LSA SAF to obtain the MSG Evapotranspiration (METv3) product (Barrios et al., 2024). Cloud identification relies on the software developed by the NWC-SAF. For cloudy, cloud contaminated or cloud edge pixels, LST results from the SEB based on the Tiled ECMWF Scheme for Surface Exchanges



over Land with revised hydrology (H-TESSSEL)(Albergel et al., 2012; Balsamo et al., 2009; Viterbo and Beljaars, 1995), which also provides turbulent energy fluxes (and therefore evapotranspiration, METv3) at the surface. While being based on a model scheme, the estimated skin temperature and energy fluxes are strongly constrained by remote sensing data, namely shortwave and longwave downwelling fluxes at the surface, albedo and LAI (all of these derived within the LSA SAF from MSG/SEVIRI observations) (Carrer et al., 2018, 2019; García-Haro et al., 2019) and root zone soil moisture (H-SAF product based on ASCAT data)(H-SAF, 2020).

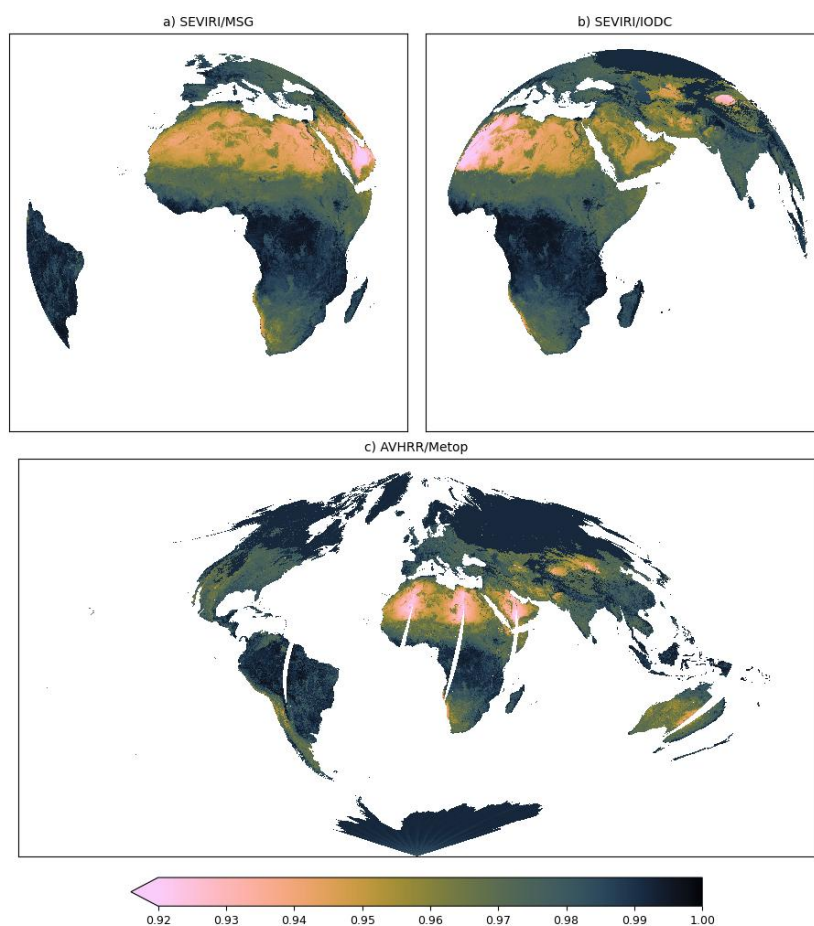


Figure 3 – Channel 10.8 μm emissivity estimates obtained with the V-TES methodology for a) SEVIRI on board 0-degree MSG, b) SEVIRI on-board IODC MSG and c) AVHRR on-board Metop, on the 15th of March 2024.

The METv3 algorithm assumes that each SEVIRI pixel is composed by multiple tiles corresponding to different land cover types: bare soil, high and low vegetation, water, snow, rock, urban. Then, for each tile i , the SEB is solved using:

$$Rn_i = H_i + LE_i + G_i, \quad (5)$$



where Rn_i is the net radiation, H_i and LE_i are the sensible and latent heat fluxes, respectively, and G_i is the ground heat flux. Net radiation, Rn_i , can also be defined as the combination of net shortwave and longwave fluxes at the surface:

$$R_{n,i} = (1 - \alpha)S_{\downarrow} + \varepsilon(L_{\uparrow} - \sigma T_{skin_i}^4), \quad (6)$$

235 where $S_{\downarrow}$ and $L_{\uparrow}$ represent the downwelling shortwave and longwave fluxes, α is the albedo, ε is the emissivity and σ the Stefan-Boltzmann constant. The latent and sensible heat fluxes are modulated by the canopy and aerodynamical resistances and the ground heat flux is a function of net radiation and LAI. More information on the parameterization of H_i , LE_i and G_i can be found in Barrios et al. (2024). The algorithm iteratively solves the closure of the SEB, resulting in an estimation of the three fluxes and the T_{skin} , from which the evapotranspiration is derived.

240 By solving the T_{skin} instead of ingesting LST retrieved from satellite TIR measurements, it is ensured that METv3 estimates can be obtained under cloudy conditions. This SKT estimate can then be used to complement the MLSTv2 and generate the MLST-ASv2 product. The use of satellite data from the SEVIRI sensor to force the SEB model results in more realistic LST estimates than those provided by schemes that rely on an atmospheric model. Nevertheless, the METv3 algorithm still relies on parameterizations that have been derived from in-situ flux measurements, which introduces additional uncertainty, and
 245 are likely to misrepresent some surface conditions.

2.5. Validation against In-situ LST estimates

Quality assessment of the LSA-SAF LST Data Records is performed through comparison against LST estimates from in-situ stations, following the recommendations by the Committee on Earth Observation Satellites Working Group (CEOS) on Calibration and Validation - Land Product Validation (LPV) Subgroup for validation of Land Surface Temperature Products
 250 (Guillevic et al., 2018). As per the agreed protocol, the LST comparisons are summarized using robust statistics, namely by computing median error (μ) as an indicator of accuracy:

$$\mu = median(LST_{sat} - LST_{insitu}), \quad (7)$$

where LST_{sat} stands for satellite LST and LST_{insitu} for in-situ LST. Precision is measured by the median of the absolute residuals (σ):

$$255 \quad \sigma = 1.4826 * median(|(LST_{sat} - LST_{insitu}) - \mu|), \quad (8)$$

where the value 1.4826 is used to convert the median of residuals into an unbiased measure of standard deviation (Göttsche et al., 2013). Total product uncertainty is here described by the RMSE:

$$RMSE = \sqrt{\frac{\sum (LST_{sat} - LST_{insitu})^2}{N}}, \quad (9)$$

where N is the sample size.

260 In-situ and satellite observations are matched by choosing the closest pixel/grid point to the station and averaging in-situ observations in a ± 3 -minute window centered at the satellite observation time (except for inland water measurements for



which a 1-hour window is considered). Only satellite observations with valid values within 3x3 pixels surrounding the station are used to avoid cloud contamination.

2.5.1. In-situ LST data

265 In-situ LST is obtained from four main databases: the Karlsruhe Institute of Technology (KIT)(Ermida et al., 2014; Göttsche et al., 2016) sites, the Baseline Surface Radiation Network (BSRN)(Driemel et al., 2018), the Surface Radiation Budget Network (SURFRAD)(Augustine et al., 2005), and the Copernicus Ground Based Observation for Validation (GBOV)(Anand et al., 2023), as detailed in Table A1.

270 Stations of the KIT network were specifically designed for LST validation, being therefore located in areas where the surface is homogeneous at the scale of a few kms. Measurements are provided by narrow band radiometers and, at the EVO and KAL sites (Table A1), multiple radiometers measure different components of the surface which allow upscaling to the satellite pixel scale (Ermida et al., 2014; Göttsche et al., 2016). For the KIT stations, LST is obtained by inverting Planck's law (Trigo et al., 2021) and emissivities are obtained either from in-situ measurements or from the MLST (operational) product (see details in Table A1).

275 For the remaining networks, radiances are measured by pyrogeometers and LST is estimated by inverting Stefan-Boltzmann's law (Trigo et al., 2021). For the BSRN and SURFRAD networks, emissivities are either derived from spectral library data or monthly values are obtained using the code provided by Ermida et al. (2020) on Google Earth Engine (GEE)(Gorelick et al., 2017). GBOV uses estimates from an emissivity database by Seemann et al. (2008), as described in the respective product documentation (Anand et al., 2023).

280 For each site, spatial heterogeneity was assessed by computing the LST standard deviation (STD) in the surrounding area, using high resolution LST derived from Landsat 8 and 9 (Ermida et al., 2020). STD was computed considering the field-of-view (FOV) of each LST product, namely an area of 3x3 grid points in the case of the EDLSTv2 (corresponding to about 3x3 km² centered in the station) and an area of 1.5x1.5 pixels in the case of the MLSTv2 and MLST-ASv2 products (corresponding to about 5x5 km²). The larger number of pixels in the case of the Metop LST aimed at accounting for the
 285 uncertainties associated to projecting the original swath LST retrievals onto the sinusoidal grid (Table 1). Stations with STD above 2K were classified as located within an "heterogeneous surface" and removed from the analysis. When Landsat data is not available (e.g. at the poles), the criteria to select the stations is the closeness to the coast, to avoid using water contaminated pixels. Table A1 indicates which stations were used to validate each LST product. The GEE code used to derived the LST STD is available at https://code.earthengine.google.com/?accept_repo=users/sofiaermida/LST_validation.

290 2.6. Stability assessment

The LSA-SAF datasets were further assessed for stability to ensure that the data are free from non-climatic discontinuities and artificial trends. Such analysis is carried out following the methodology proposed by Good et al. (2022) where each dataset's anomalies are compared against a reference dataset that we assume to be free of non-climatic effects. In Good et



al.'s work, the used reference was a set of station measurements of 2-meter air temperature located over Europe. Due to the
 295 limited spatial coverage of meteorological stations, here we consider three reference datasets with global coverage:

Air temperature (2-meter) anomalies from the Climate Research Unit temperature version 5 (CRUTEM5)(Osborn et al.,
 2021), which is available with a spatial resolution of 5°x5° latitude-longitude and a monthly-mean time resolution. For each
 grid box, the anomaly value is the average of the anomalies of all the stations inside that grid box. For this reason, the
 number of available grid boxes can vary in time depending on station availability. The anomalies are computed with respect
 300 to the 1961-1990 normal.

Skin temperature (T_{skin}) from the ERA5 reanalysis 15, which result from the assimilation of a large amount of observations.
 Despite the variability of assimilated observations over time, T_{skin} results from the SEB scheme used within ERA5, which,
 should not be much affected by changes in the assimilated data;

The MODerate Resolution Imaging Spectroradiometers (MODIS) on board the Aqua LST dataset generated by the European
 305 Space Agency's (ESA) Climate Change Initiative (CCI) for LST project (Ghent et al., 2025), which was analyzed and
 considered free from non-climatic discontinuities by (Good et al., 2022).

For the comparison with CRUTEM5, the LSA-SAF products are projected onto the 5°x5° grid by averaging pixels within the
 coarser grid. For comparisons against ERA5/satellite, LST datasets were projected onto the 0.25°x0.25° regular latitude-
 longitude grid of ERA5, also by simple average. To reduce the impact of cloud contamination, when projecting satellite
 310 products, grid-points with a number of valid pixels below 30% are discarded. ERA5 estimates where the total cloud cover is
 above 30% over clear-sky pixels (i.e. where LST is valid) are also removed.

To maximize the temporal consistency between LSA SAF LST products and ERA5, or among satellite datasets, the
 maximum time differences allowed are 7.5 minutes in the case of satellite products' match-ups, and up 30 minutes in the
 case of satellite – ERA5 match-ups. Monthly mean values of LST/T_{skin} are then derived for each dataset on a pixel-by-pixel
 315 basis.

For a given year, y , and month, m , the anomaly difference between a satellite product, SAT, and a reference dataset, REF, is
 the spatial average of the anomaly difference at each grid point, p :

$$\Delta T_{\text{anom}}(y, m) = \frac{1}{N_p} \sum_p (T_{\text{anomSAT}_p}(y, m) - T_{\text{anomREF}_p}(y, m)), \quad (10)$$

with

$$320 \quad T_{\text{anom}_p}(y, m) = T_p(y, m) - T_{\text{normal},p}(m), \quad (11)$$

where T represents the LST or T_{skin} monthly composites for each dataset at each grid point, N_p is the number of valid grid
 points, and the normal is the average of all years considered for the normal. Only grid points where both satellite and
 reference have valid values are considered in the averaging. The period used to compute the normal depends on the data
 availability of each product, as described in Table 3. In the case of the CRUTEM5, anomalies are precomputed for w.r.t. the
 325 1961-1990 period. To adjust for the datasets' normal period, for each month we compute the average CRUTEM5 anomaly in
 all years of the dataset and remove it from the original anomaly.



To compare with CRUTEM5, we consider the average of all available time steps. For ERA5 and MSG-based datasets, comparisons are performed at the hour of the maximum and minimum LST, which are identified for each month from the mean diurnal cycle, and for the average of the full diurnal cycle. For the Metop and CCI MODIS products, the daytime and nighttime overpasses are considered, as well as their average. Given the difference in observation time between the two, the EDLSTv2 was not compared against CCI MODIS. Due to the limited overlap of the time periods, IODC MLSTv2 was also not compared against CCI MODIS. To complement the comparisons against ERA5, EDLSTv2 stability was also assessed with respect to MLSTv2 (taking into account the stability established with ERA5 and MODIS). Likewise, IODC MLSTv2 is also assessed using EDLSTv2.

335

Table 3 – Reference datasets, normal period, and variables analyzed in the stability assessment of each LSA-SAF LST product. The LST_{MEAN} corresponds to the average of the full diurnal cycle or the average between the day and night overpasses.

<i>Reference</i>	MLSTv2	IODC MLSTv2	MLST-ASv2	EDLSTv2
<i>Station</i>	CRUTEM5 (all-sky) 2004-2024 LST_{MEAN}	CRUTEM5 (all-sky) 2018-2024 LST_{MEAN}	CRUTEM5 (all-sky) 2004-2024 LST_{MEAN}	CRUTEM5 (all-sky) 2007-2024 LST_{MEAN}
<i>Reanalysis</i>	ERA5 (clear-sky) 2004-2024 LST_{MAX} , LST_{MIN} , LST_{MEAN}	ERA5 (clear-sky) 2018-2024 LST_{MAX} , LST_{MIN} , LST_{MEAN}	ERA5 (all-sky) 2004-2024 LST_{MAX} , LST_{MIN} , LST_{MEAN}	ERA5 (clear-sky) 2007-2024 LST_{DAY} , LST_{NIGHT} , LST_{MEAN}
<i>Satellite</i>	MODIS (clear-sky) 2004-2021 LST_{DAY} , LST_{NIGHT} , LST_{MEAN}	EDLSTv2 (clear-sky) 2018-2024 LST_{DAY} , LST_{NIGHT} , LST_{MEAN}	-	MLSTv2 (clear-sky) 2007-2024 LST_{DAY} , LST_{NIGHT} , LST_{MEAN}

For each ΔT_{anom} combination, the respective trend (t) is estimated using the methodology by Sen(Sen, 1968). The obtained trends are evaluated taking into account the defined stability requirements provide by the GCOS for LST, namely where the goal is established at 0.1 K/decade, the baseline at 0.2 K/decade and the threshold is at 0.3 K/decade (World Meteorological Organization et al., 2025). The agreement in variability of the different time-series is also analyzed through the Pearson correlation coefficient (R).

340



3. Validation and Assessment

3.1. Validation against station data

Validation of the LSA-SAF products is performed through comparison against in-situ data. Stations of the KIT network were planned to ensure spatial representativeness at satellite pixel scale, with upscaling available for stations with sparse vegetation. For this reason, we consider these to provide a validation reference with the highest confidence. Although BSRN, SURFRAD and GBOV measurements are provided by similar sensors, we consider that the use of Landsat-based emissivities, that have higher resolution, provides higher confidence in the comparisons with BSRN/SURFRAD data. Table 4 shows the statistics, as defined in section 2.5, for the four considered networks. BSRN and SURFRAD are grouped due to the similar data processing and since SURFRAD data is only used in the validation of EDLSTv2. Figure 4 to Figure 7 show the scatterplots of satellite versus in-situ LSTs for each product and each network, showing also the statistics for day and night separately.

Table 4 - Validation statistics for the LSA-SAF LST products for each validation dataset, namely accuracy (μ ; K), precision (σ ; K), RMSE (K) and number of observations (N), for all times-of-day. For the MLST-ASv2 product, the first row corresponds to the statistics for all-sky conditions and the second row, in brackets, for cloudy-sky conditions.

	KIT				BSRN/SURFRAD				GBOV			
	μ	σ	RMSE	N	μ	σ	RMSE	N	μ	σ	RMSE	N
MLSTv2	0.28	1.33	1.83	528789	-0.90	2.05	2.32	222068	1.94	2.39	3.36	86354
IODC MLSTv2	0.13	1.28	1.55	141412	0.61	2.51	3.06	53592	1.36	2.48	3.55	36087
MLST-ASv2	0.19 (-0.14)	1.54 (2.65)	2.39 (3.46)	299756 (88913)	0.26 (0.34)	1.90 (1.85)	2.34 (2.34)	488324 (385705)	2.63 (2.71)	1.87 (2.79)	3.81 (3.95)	125407 (90327)
EDLSTv2	0.22	1.16	1.72	11968	-0.33	2.11	2.65	41611	0.79	2.61	3.26	37110

MLSTv2

The MLSTv2 (SEVIRI/MSG 0-degree) product shows very good agreement with the KIT and BSRN LSTs, with accuracy and precision values of 0.28K (-0.9K) and 1.33K (2.05K), respectively, for the KIT (BSRN) dataset (Table 4). RMSE values also indicate a good agreement, with values of 1.83K and 2.32K for KIT and BSRN, respectively. For the GBOV network, statistics indicate a slightly reduced agreement, with accuracy, precision and RMSE values of 1.94K, 2.39K and 3.36K, respectively, but these are still reasonable considering the increased uncertainty due to station representativeness. Statistics clearly reflect this issue of matching station point measurements to the satellite FOV scale, with the KIT showing overall



better agreement, followed by BSRN and GBOV. The respective scatter plots also reflect this effect (Figure 4), with MLSTv2 showing a very good agreement with KIT LSTs for both day and night-time (Figure 4a and b, respectively) for the full range of LSTs. For BSRN and GBOV data, it is noticeable an increase in discrepancies between satellite and in-situ, particularly for the higher values of daytime LST (Figure 4c and e, respectively). LST contrasts between the different land covers within a pixel tend to be more pronounced during the daytime, and the higher the LSTs the higher these contrasts are likely to be (Ermida et al., 2014). As a result, a satellite pixel containing a large variability of surfaces types is likely to show higher discrepancies with respect to the station value (which typically only encompasses one surface type). It should be noted that, although a high resolution LST product was used for a pre-selection of the more homogeneous stations, the data used corresponds to the average LST over several years with an overpass time around local 10AM. Therefore, the pronounced LST contrasts that occur at peak LST values during the warm season were smoothed and likely not fully reflected in those averages. MLSTv2 daytime accuracy (precision) values are -0.14K (1.22K), -1.02K (1.72K) and 2.67K (2.21K) for KIT, BSRN and GBOV, respectively, while the corresponding RMSE values are 1.82K, 2.48K and 3.73K (Figure 4a, c and e, respectively). The nighttime accuracy (precision) values are 0.69K (1.20K), 0.93K (1.73K) and 1.21K (2.45K) for KIT, BSRN and GBOV, respectively, and values of RMSE are 1.83K, 2.15K and 3.01K (Figure 4b, d and f, respectively).

IODC MLSTv2

The IODC MLSTv2 product also shows, overall, very good agreement with in-situ observations. Accuracy (precision) values are 0.13K (1.28K), 0.61K (2.51k) and 1.36K (2.48K) for the KIT, BSRN and GBOV networks, respectively (Table 4), and RMSE values are 1.55K, 3.06K and 3.55K. The poorer agreement for BSRN and GBOV is consistent to that seen for the 0-degree MLSTv2 product and can be attributed to station representativeness. Likewise, the scatter plots (Figure 5) also indicate a very good agreement with KIT data (Figure 5a,b), and reasonably good agreement with BSRN and GBOV (Figure 5c-f), with discrepancies being higher at daytime for the higher ranges of LST. IODC MLSTv2 daytime accuracy (precision) values are -0.29K (1.22K), -1.11K (3.17K) and 2.38K (2.09K) for KIT, BSRN and GBOV, respectively, and values of RMSE are 1.60K, 3.57K and 3.59K (Figure 5a, c and e, respectively). The nighttime accuracy (precision) values are 0.52K (1.13K), 1.52K (1.68K) and 0.33K (2.62K) for KIT, BSRN and GBOV, respectively, and values of RMSE are 1.49K, 2.48K and 3.50K (Figure 5b, d and f, respectively).

MLST-ASv2

For the all-sky product, MLST-ASv2, statistics are presented for all conditions on the first row and for cloudy-sky only in the second row (Table 4). The clear-sky statistics are not shown since these correspond to the MLSTv2 product. The MLST-ASv2 product has lower frequency (30-min) than the MLSTv2 (15-min), which is reflected in the sample sizes. The MLST-ASv2 product also shows a good agreement with in-situ, for both all-sky and cloudy-sky conditions. The all-sky accuracy (precision) values are 0.19K (1.54K), 0.26K (1.90K) and 2.63K (1.87K) for the KIT, BSRN and GBOV networks,



respectively (Table 4), and RMSE values are 2.39K, 2.34K and 3.81K. The cloudy-sky component of the product shows accuracy values that are slightly better than clear-sky counterpart (MLSTv2) for the KIT and BSRN stations, with values of -0.14K and 0.34K, but poorer in the case of GBOV, with a value of 2.71K (Table 4). The data scatter described by precision tends to be larger under cloudy than clear-sky conditions, with exception of the BSRN stations, with values of 2.65K, 1.85K and 2.79K for KIT, BSRN and GBOV, respectively. A higher variance of the satellite-station LST differences is to be expected, not only due to the increased uncertainties associated to the SEB model, but also because sub-pixel scale clouds can create large discrepancies due to the different proportions of shaded surface included in the satellite and station's FOV (Jiménez et al., 2017). This is also reflected in the product's RMSE, resulting in values of 3.46K, 2.34K and 3.95K.

The scatter plots of MLST-ASv2 versus in-situ KIT data for cloudy-sky (Figure 6) indicate a good accuracy of daytime cloudy-sky LSTs, with a value of 0.25K, but reduced precision and total uncertainty, with values of 3.41K and 4.21K (Figure 6a). The dispersion of points, particularly at warmer LSTs, highlights the increasing LST uncertainty under cloudy conditions, when estimates are less directly linked to TOA observations. In these cases, LST will be influenced by errors in the SBE, input variables, as well as cloud types (broken clouds may be particularly problematic). At nighttime for the KIT stations, there is a better agreement with in-situ, with an accuracy value of -0.40K, precision of 1.97K and RMSE of 2.34K (Figure 6b). The product shows a good agreement with BSRN data, with daytime (nighttime) accuracy, precision and RMSE values of 0.72K (0.09K), 2.16K (1.48K) and 2.64K (1.98K), respectively. Consistent with the clear-sky MLSTv2 product, larger discrepancies were found when comparing to GBOV data: the day and nighttime accuracy values are 3.62K and 2.14K, respectively, and precision (RMSE) values are 2.22K (4.73K) and 1.22K (2.88K).

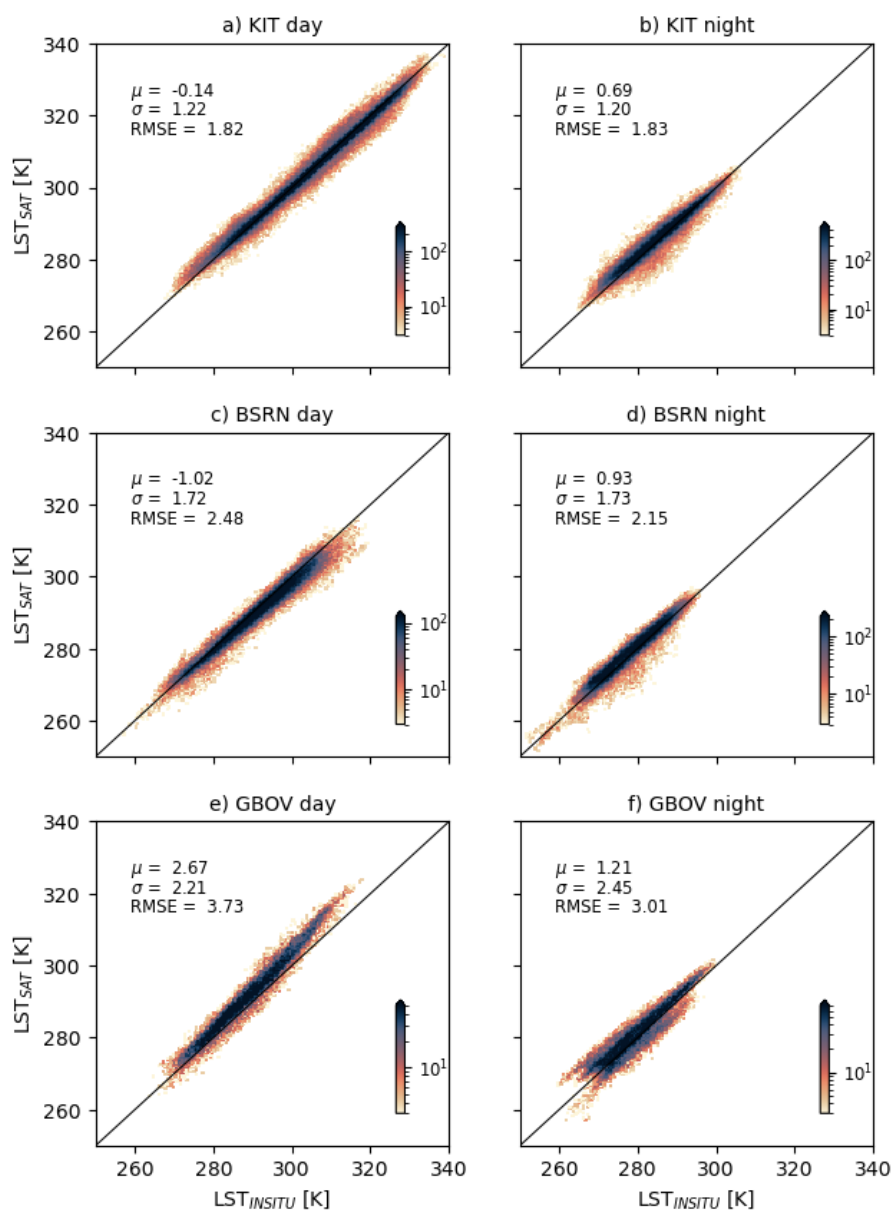
EDLSTv2

In the case of the EDLSTv2 product, statistics are shown to be more consistent across all station networks. The accuracy when compared to the KIT, BSRN/SURFRAD and GBOV stations is 0.22K, -0.33K and 0.79K, respectively. Precision and RMSE values are slightly higher for the BSRN/SURFRAD and GBOV networks, but discrepancies between the different networks are not as high as in the case of the MSG-based products. Given the higher resolution of the EDLSTv2 product, this supports the hypothesis that large discrepancies between satellite and in-situ for these stations is likely due to surface heterogeneity around the station. On the other hand, EDLSTv2 observation times around 10:30 am/pm local time (Metop local overpass) avoid the peak temperatures, when spatial heterogeneity also tends to be higher. The precision (RMSE) values are 1.16K (1.72K), 2.11K (2.65K) and 2.61K (3.26K) for comparisons with KIT, BSRN/SURFRAD and GBOV stations, respectively.

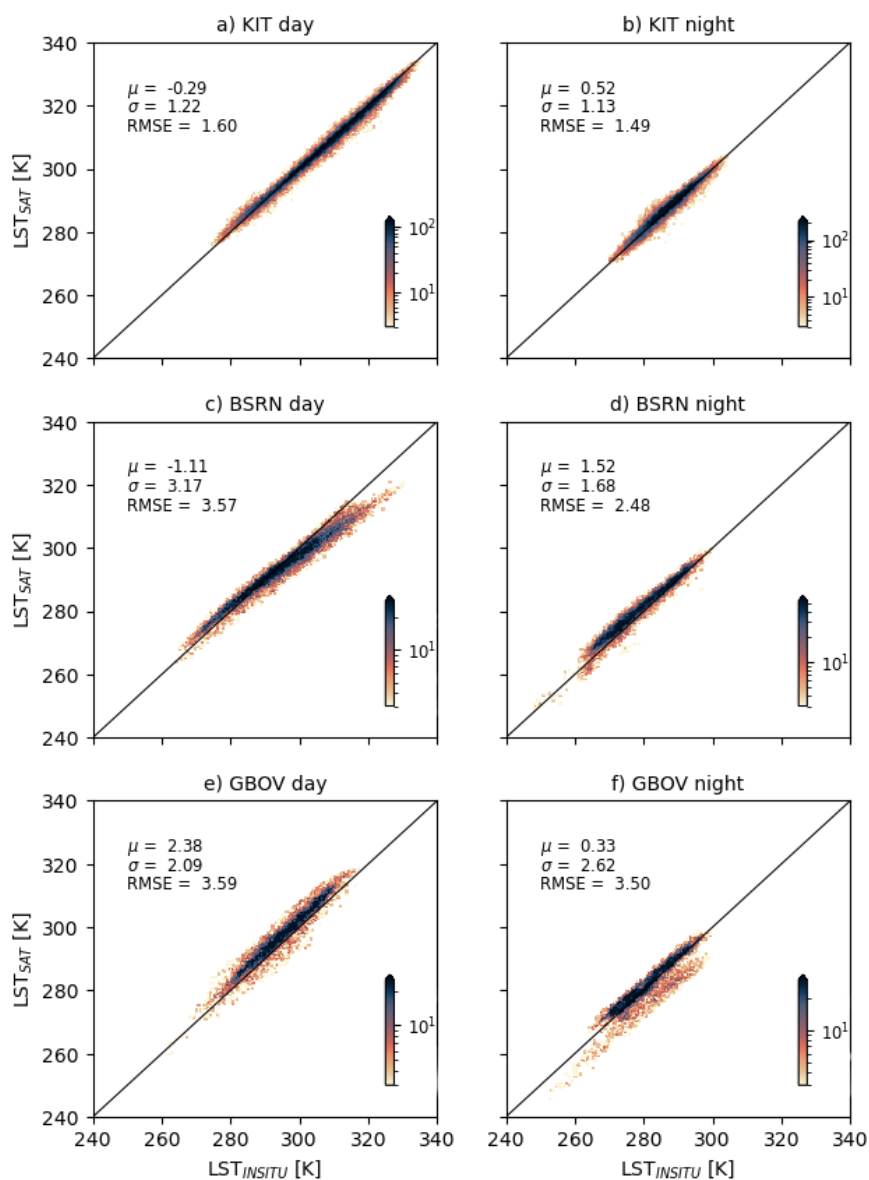
The scatter plots (Figure 7) indicate compatible performance of the product between daytime and nighttime. The daytime (nighttime) accuracy is 0.20K (0.24K), -0.82K (0.08K) and 0.81K (0.78K) for comparisons with KIT, BSRN/SURFRAD and GBOV LSTs. Precision and RMSE are all below 2K for the KIT stations, with respective values of 1.26K (1.05K) and 1.94K (1.45K) for daytime (nighttime) (Figure 7a and b). The precision and RMSE values are slightly higher for the BSRN/SURFRAD stations, likely due to increased heterogeneity, with values of precision (RMSE) of 2.28K (2.97K) and



435 2.00K (2.31K) and day and night, respectively (Figure 7c and d). The BSRN network also includes stations over Antarctica
(DOM, GVN, SPO; Table A1) that allow a validation of the product under extreme low temperatures. The scatter plots
(Figure 7c and d) suggest a good agreement at low temperatures, with slightly increased dispersion at nighttime. This is
likely related to cloud contamination, since cloud detection over very cold surfaces is particularly difficult at night. At
GBOV sites, the increased daytime values of precision (3.26K) and RMSE (3.64K) suggest larger discrepancies than in the
440 other stations, while nighttime values (2.11K and 2.81K, respectively) are more in line with the other networks. This again
suggests representativeness issues of the GBOV stations of the satellite pixel scale.



445 **Figure 4- MLSTv2 (LST_{SAT}; K) against in-situ LST (LST_{INSITU}; K) for each validation network considered, separated into daytime (left panels) and nighttime (right panels). The respective statistics, namely accuracy (μ ; K), precision (σ ; K) and RMSE (K) are also shown.**



450 Figure 5- As in Figure 4 but for IODC MLSTv2.

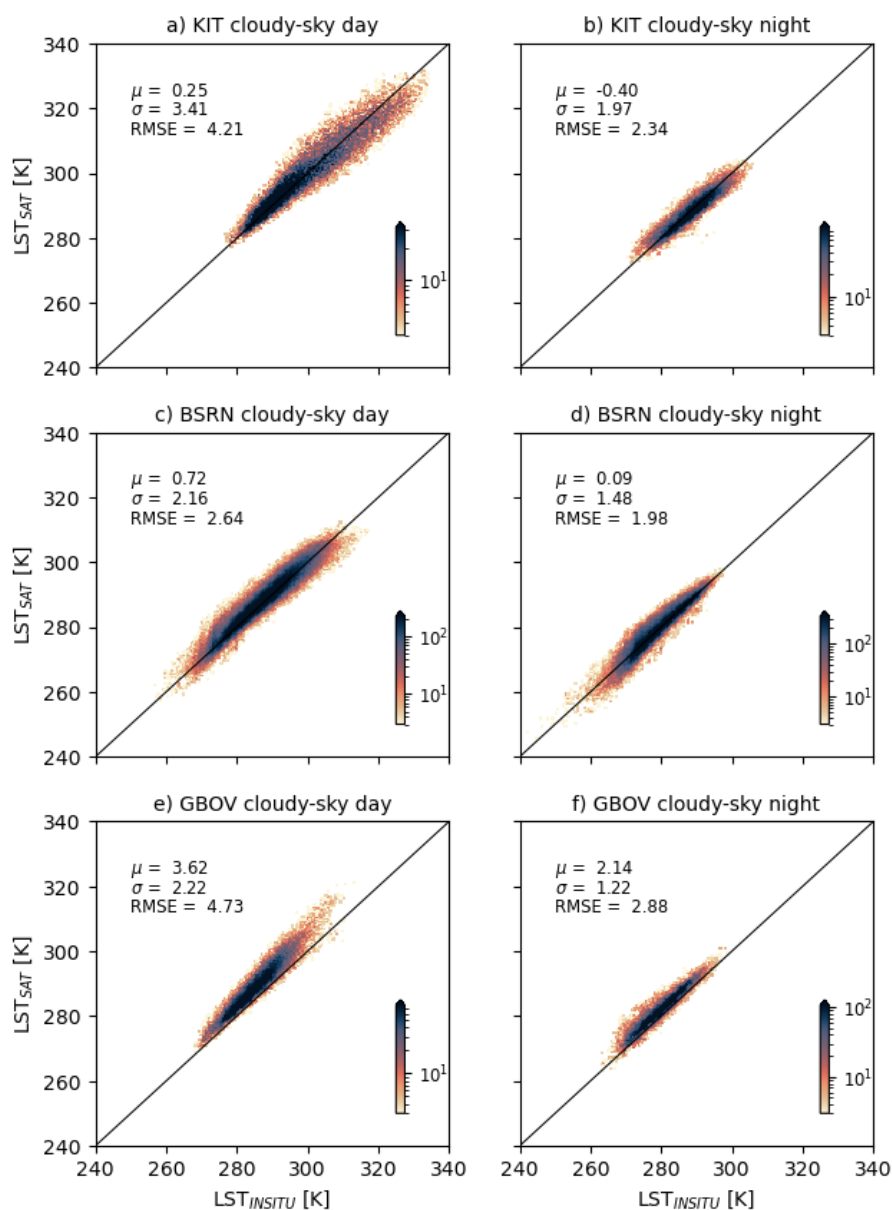


Figure 6 – As in Figure 4 but for MLST-ASv2 and for cloudy-sky conditions only.

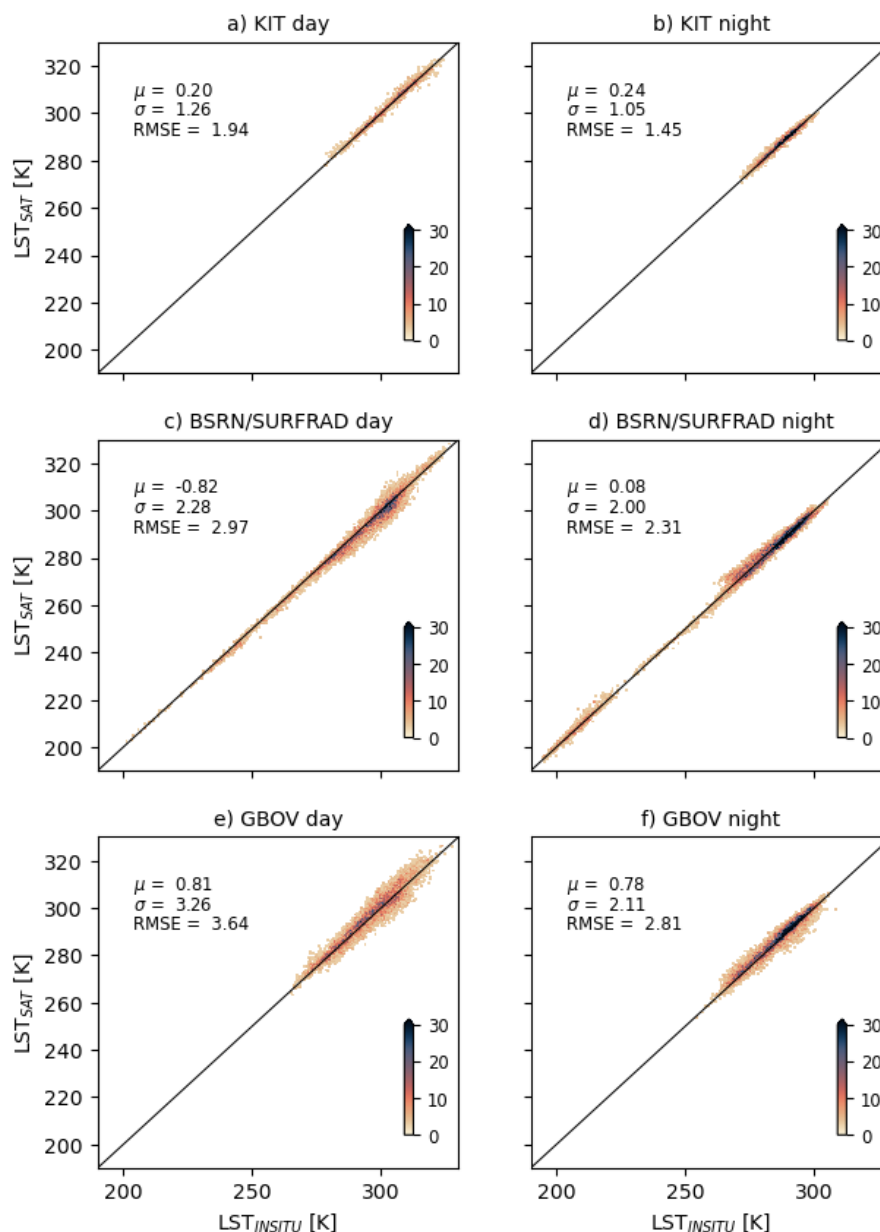


Figure 7 – As in Figure 4 but for EDLSTv2.

3.2. Product inter-comparison

In this section, we assess the consistency between the MLSTv2 and EDLSTv2 products. For that purpose, the EDLSTv2 product is projected onto the MLSTv2 grid by simple averaging and matching the LST estimates with time differences of up to 7.5 minutes. Figure 8 shows maps of the average and standard deviation of the differences between EDLSTv2 and MLSTv2 for the seasons of July to August (JJA) and December to February (DJF), for daytime and nighttime. The products



show an overall good agreement. During daytime DJF (JJA), 67.5% (59.4%) of pixels have a mean difference between $\pm 1K$ and 92.5% (84.1%) within $\pm 2K$. The daytime STD is below 2K in 72.3% (65.5%) of pixels for the DJF (JJA) season. At nighttime, differences are significantly reduced, with 90.3% (87.4%) of pixels in DJF (JJA) having mean differences within $\pm 1K$ and 99% (96.3%) having STD below 2K.

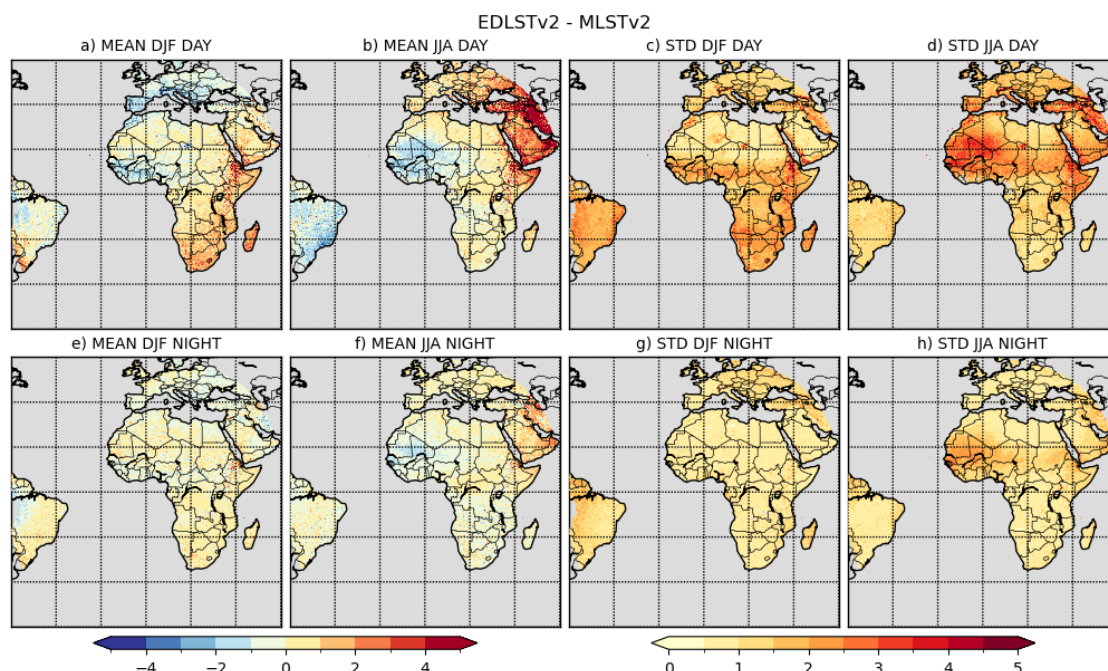


Figure 8 – Statistics of the differences between MLSTv2 and EDLSTv2 (as EDLSTv2 minus MLSTv2; K), namely the mean (a,b,e,f) and the standard deviation (STD; c,d,g,h), for the months December-January-February (DJF; a,c,e,g) and June-July-August (JJA; b,d,f,h) of 2007 to 2024, for daytime (a-d) and nighttime (e-h).

Higher discrepancies can be found in some regions, for example: 1) negative mean differences and higher STD values are found over the western part of the Saharan desert during daytime in JJA; 2) strong positive mean differences at the edge of the MSG disk, over eastern Africa/western Asia, particularly during daytime in JJA; 3) negative mean differences over South America during daytime in JJA. To further investigate these discrepancies, we analyze LST differences at various regions of interest taking into account the view angle of the Metop/AVHRR. Figure 9 shows a map of surface elevation overlapped with the regions of interest, as well as the mean and standard deviation of LST differences for each region plotted for different ranges of VZA, for daytime and nighttime, respectively.

An area over the Zagros Mountain ranges of Iran was selected for the particularly high mean differences found at daytime JJA (Figure 8b). The analysis of the VZA dependence of the differences over this region (Figure 9a) shows high discrepancies during daytime for AVHRR views opposing the MSG/SEVIRI view; however, the mean difference is close to 1K for AVHRR VZAs approximating those of SEVIRI ($>60^\circ$). This mountain range shows a strong north-south orientation, which enhances LST directional effects: during the morning, an east facing view favors the observation of shaded surfaces,

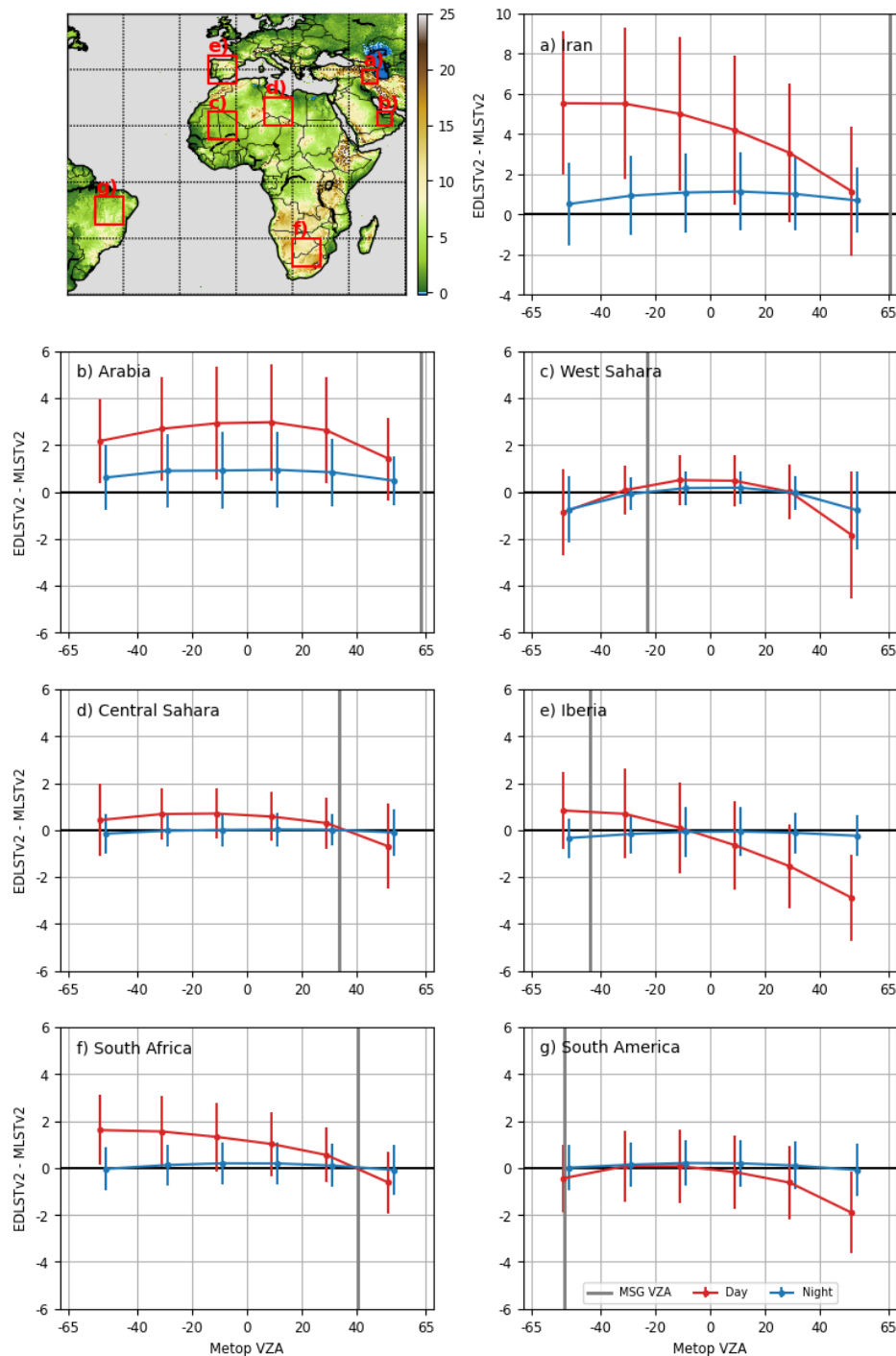


while a west facing view favors the observation of sunlit slopes. During nighttime, i.e., in the absence of shading effects, differences are much smaller and show very limited VZA dependence. A similar effect occurs over the East African Rift System, e.g. in Ethiopia, which is also roughly a north-south oriented mountain range (not shown). Over the Arabian Peninsula (Figure 9b), the daytime mean difference curve is more symmetrical, suggesting a mix of emissivity and shading effects on LST. Although the emissivity retrieval includes a VZA dependence, the very high SEVIRI VZA can result in larger surface emissivity uncertainties. There are also some mountain ranges and high sand dunes near the Persian Gulf that contribute to LST directional effects. Over these two areas (Figure 9a and b), the mean difference is around 1K for nighttime or daytime with similar view geometry. These larger discrepancies could be related to the higher uncertainties of the GSW algorithm at very high VZAs ($\sim 65^\circ$).

We also analyzed two regions in the Saharan desert with contrasting behavior of the LST differences (Figure 9c and d). Over the western region of the Sahara (Figure 9c), there is a persistent occurrence of high dust aerosol concentrations during the JJA season. Dust aerosols have significant absorption around the 11 μm spectral region, resulting in LST underestimation under high concentrations (Stante et al., 2023). This effect is higher during daytime (with warmer brightness temperatures) and the errors increase with VZA, i.e., with increasing optical path. This results in an underestimation of AVHRR LSTs when the VZA is larger than SEVIRI, as seen in Figure 9c. Over central Africa (Figure 9d), aerosol loads are less significant (e.g. Stante et al. (2023)) and thus aerosol effects are reduced, leading to much lower discrepancies. The dependence on VZA is also reduced, suggesting a good representation of emissivity directionality. The daytime VZA dependence is asymmetrical, suggesting that some shading effects occur in the area.

The regions of Iberia, southern Africa and southern America (Figure 9e-g) are all characterized by a predominance of sparse-to-moderate vegetation cover. Both show strongly asymmetrical VZA dependence during daytime. For morning overpasses, a west (east) facing view, i.e. for $\text{VZA} < 0$ ($\text{VZA} > 0$), favors the observation of sunlit (shaded) surfaces resulting in higher (lower) LSTs. The LST differences tend to be lower when SEVIRI and AVHRR views are more aligned; in the case of Iberia and Southern Africa the SEVIRI view is closer to a north-south orientation, contrasting with the AVHRR east-west orientation and leading to higher LST discrepancies. At night, VZA dependence is smaller and symmetrical and mainly reflects the proportion of vegetated versus bare ground components within the satellite FOV (Ermida et al., 2014).

For the IODC MLSTv2 product, comparisons with the EDLSTv2 yield similar results to the 0-degree MLSTv2 product (Figure A1 of supplementary material). During daytime DJF (JJA), 58.2% (47.8%) of pixels have a mean difference between $\pm 1\text{K}$ and 84.4% (76.4%) within $\pm 2\text{K}$. The daytime STD is below 2K in 72.8% (58.5%) of pixels for the DJF (JJA) season. At nighttime, 83.5% (76.9%) of pixels in DJF (JJA) have mean differences within $\pm 1\text{K}$ and 97.6% (92.7%) have STD below 2K. As in the case of the MLSTv2, the larger discrepancies can mostly be attributed to directional effects or aerosol contamination.



515 **Figure 9 – Elevation (km) and selected regions of interest (top left panel) and respective average (markers) and standard deviation (bars) of the differences between the EDLSTv2 and MLSTv2 products at each region, grouped by Metop view zenith angle (VZA; degrees), for daytime (red) and nighttime (blue). Negative (positive) VZA values indicate the scene is viewed from east (west). The average MSG view zenith angle is also indicated by the gray vertical line.**



3.3. Stability assessment

Stability of the LSA-SAF LST products is assessed by comparing LST anomalies against independent reference datasets, preferably widely used in climate studies. Here we have considered ERA5 reanalysis (Hersbach et al., 2020), one satellite-based product not part of the LSA SAF portfolio (ESA CCI MODIS/Aqua; Ghent et al., 2025), and a gridded 2m air temperature product derived from in situ stations (CRUTEM5; Osborn et al., 2021). The MLSTv2 anomalies were compared against CCI MODIS/Aqua anomalies, which was one of the sensors considered to be most stable in the study by Good et al. (2022). Good et al., reported non-significant trends for the minimum, maximum and mean temperature for the beta version of the CCI MODIS/Aqua product. The corresponding trends for the used version 4 are currently not available. Only the MLSTv2 product is compared against CCI MODIS data since its high temporal frequency allows close match-ups with the MODIS overpasses. We found that anomaly comparisons without first performing temporal matching tend to lead to lower agreement since the days considered in calculating monthly means may significantly vary due to cloud cover (not shown). For this reason, for the stability assessment time match-ups are always performed prior to the anomaly computation. The only exception is the CRUTEM5 dataset, since these data are only available only as monthly values. Table 5 presents the trend values of the LST anomalies of each LST product as obtained when compared against station, reanalysis and satellite reference.

Table 5 – Trend of the anomaly differences (ΔT_{anom} in eq. (10; K/decade) for each combination of LSA-SAF product and reference dataset described in Table 3, for the maximum/daytime ($LST_{\text{MAX/DAY}}$), the minimum/nighttime ($LST_{\text{MIN/NIGHT}}$) and mean LST (LST_{MEAN}). Values marked with * correspond to non-significant trends at the 95% confidence level.

Product		Reference				
		CRUTEM5	ERA5	CCI MODIS	MLSTv2	EDLSTv2
MLSTv2	$LST_{\text{MAX/DAY}}$	-	0.034*	-0.032	-	-0.086
	$LST_{\text{MIN/NIGHT}}$	-	-0.044	-0.058	-	-0.043
	LST_{MEAN}	0.152	-0.024*	-0.043	-	-0.061
EDLSTv2	$LST_{\text{MAX/DAY}}$	-	-0.068	-	0.086	-
	$LST_{\text{MIN/NIGHT}}$	-	-0.028	-	0.043	-
	LST_{MEAN}	-0.153	-0.055	-	0.061	-
IODC MLSTv2	$LST_{\text{MAX/DAY}}$	-	-0.784	-	-	-0.375
	$LST_{\text{MIN/NIGHT}}$	-	-0.287	-	-	-0.447
	LST_{MEAN}	0.755	-0.412	-	-	-0.378
MLST-ASv2	$LST_{\text{MAX/DAY}}$	-	0.132	-	-	-
	$LST_{\text{MIN/NIGHT}}$	-	-0.126	-	-	-
	LST_{MEAN}	0.008*	-0.059	-	-	-



MLSTv2

The anomaly difference time-series for the MLSTv2 product are also shown in Figure 10 (against ERA-5 skin temperature and CCI MODIS LST product, respectively) and Figure 11 (against CRUTEM5 dataset). When compared against ERA5, the obtained anomaly difference trends are 0.034, -0.044 and -0.024 K/decade for daily maximum (LST_{MAX}), daily minimum (LST_{MIN}) and mean LST (LST_{MEAN}), respectively, with the values of the LST_{MAX} and LST_{MEAN} being non-significant at 5% p-level (Table 5). When compared against CCI MODIS, the corresponding trends are -0.032, -0.058 and -0.043 K/decade for the LST_{MAX} , LST_{MIN} and LST_{MEAN} , respectively. All these trends in the anomalies' differences meet the goal requirement of 0.1 K/decade defined by the GCOS (World Meteorological Organization et al., 2025). Correlations of the MLSTv2 anomalies with the reference datasets are all 0.98 when compared to CCI MODIS, but are slightly lower when compared against ERA5, with values of 0.89, 0.97 and 0.95 for LST_{MAX} , LST_{MIN} and LST_{MEAN} , respectively (Figure 10), with some significant discrepancies being noticeable, for example in 2004 and 2010. Compared to CRUTEM5, the MLSTv2 LST_{MEAN} ΔT_{anom} trend is slightly higher, with a value of 0.152 K/decade, which still meets the GCOS baseline requirement (0.2 K/decade). Correlation is also lower with a value of 0.77. This is likely due to discrepancies related to the difference between air and surface temperature, to the comparison of all-sky versus clear-sky and also to the sampling frequency of each dataset: 15-min in the case of MLSTv2, while the temporal sampling between stations can differ considerably (e.g., 6-hourly or only minimum and maximum temperature) (Osborn et al., 2021).

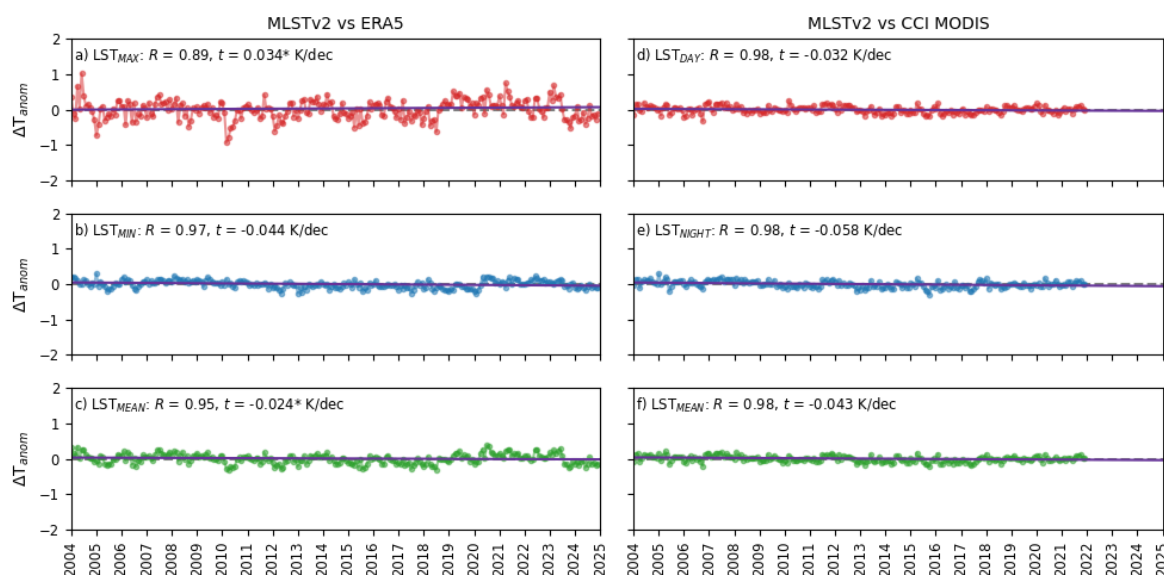


Figure 10 – Monthly anomaly difference (ΔT_{anom} in eq. (10) time series as obtained for the MLSTv2 product when compared to ERA5 (a,b,c) and CCI MODIS (d,e,f), for the maximum/daytime ($LST_{MAX/DAY}$; a,d), the minimum/nighttime ($LST_{MIN/NIGHT}$; b,e) and mean LST (LST_{MEAN} ; c,f). The trend of each series is represented by the purple line. The respective values of the trend (t ; K/decade) and correlation coefficient (R) are also shown.



We further analyze the impact of satellite transitions on the LST anomalies. For that purpose, we have computed the average and standard deviation of the anomaly differences, ΔT_{anom} , for the operation period of each platform (Table A2, available in Appendix A). When compared to ERA5, the average ΔT_{anom} of the MSG4 LST_{MAX} shows a slightly larger value of 0.15K, while the remaining MSGs have values of 0.00K, -0.06K and -0.06K, for MSG1, MSG2 and MSG3, respectively. These discrepancies are reduced when the anomalies are compared against CCI MODIS, with average ΔT_{anom} values of the LST_{MAX} of 0.01K, 0.04K, -0.06K and 0.01K for MSGs 1-4, respectively. However, it should be noted that the CCI MODIS product ends in 2021 and, therefore, it is not possible to analyze the full MLSTv2 data record. Shortly after MSG4 became operational, with the generation of NRT LST by LSA-SAF from SEVIRI on-board MSG4, it was clear that this sensor's calibration needed to be revisited. This has prompted a deeper adjustment of the previous version of the MLST product. For comparison, we performed the same analysis of the anomaly differences for the MLST product (Figure A2 and Table A6). In this analysis, it is possible to identify a significant discontinuity caused by the operation period of the MSG4, with an ERA5 average ΔT_{anom} for LST_{MAX} of 0.49K. The average ΔT_{anom} for the remaining sensors is -0.14K, -0.06K and -0.24K for MSG1, MSG2 and MSG3, respectively, which shows that the discontinuities due to sensor change have been significantly improved from the previous version.

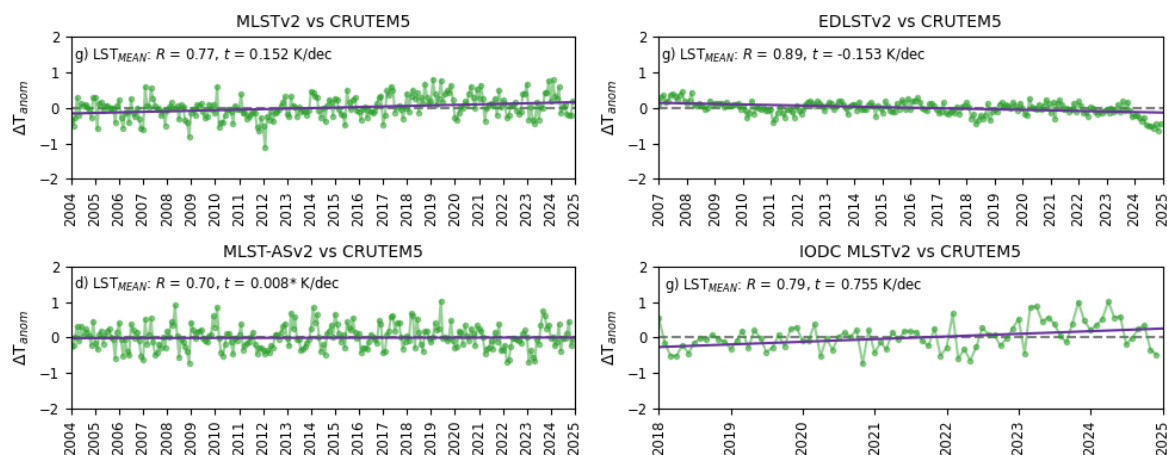


Figure 11 – As in Figure 10 but for a) MLSTv2, b) EDLSTv2, c) MLST-ASv2 and c) IODC MLSTv2 products when compared to CRUTEM5 data for the mean LST (LST_{MEAN}).

EDLSTv2

After establishing the stability of the MLSTv2, this product is used as reference for the stability assessment of the EDLSTv2. Figure 12 shows the time series of anomaly differences when comparing EDLSTv2 to ERA5 and MLSTv2. When compared against ERA5, the obtained anomaly difference trends of EDLSTv2 are -0.068, -0.028 and -0.055 K/decade for the LST_{DAY}, LST_{NIGHT} and LST_{MEAN}, respectively (Table 5). Compared to MLSTv2 the difference trends are 0.086, 0.043 and 0.061 K/decade for LST_{DAY}, LST_{NIGHT} and LST_{MEAN}, respectively. All these trends meet the goal requirement of 0.1 K/decade defined by GCOS 54. Correlations with MLSTv2/ERA5 are reasonably good, showing a better agreement at nighttime, with



a value of 0.99 (Figure 12b,e), than at daytime (0.97/0.93; Figure 12a,d), while daily (nighttime + daytime) LST anomalies present an intermediate value (0.98/0.97; Figure 12c,f). During daytime uncertainties in the LST are amplified, particularly when considering the impact of directional effects (as discussed in section 3.2), which explains the increased discrepancies. Compared to CRUTEM5, the $LST_{MEAN} \Delta T_{anom}$ trend is -0.153 K/decade, meeting the baseline requirement of GCOS. The correlation is better than for MLSTv2, with a value of 0.89 (Figure 11b). This could be related to the overpass time of Metop, for which LST values tend to be closer to air temperature. There is a conspicuous drop in the series starting in 2024 (Figure 11b), most likely due to a slight de-orbiting of Metop-B following the decision to preserve fuel, as the end of the satellite's life cycle is approaching.

An analysis of the average ΔT_{anom} in the Metop series suggests a slight sign change from Metop-A to Metop-B (Table A3), the largest discrepancies being found for LST_{DAY} when compared to MLSTv2, with ΔT_{anom} values of -0.05K and 0.04K for Metop-A and -B, respectively.

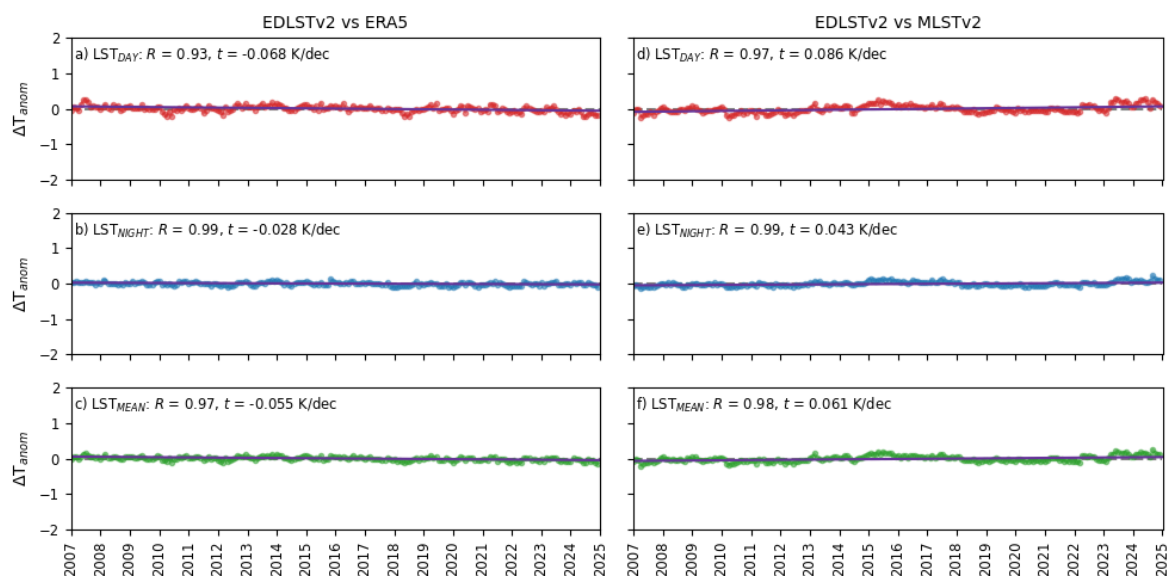


Figure 12 - As in Figure 10 but for the EDLSTv2 product when compared to ERA5 (a,b,c) and MLSTv2 (d,e,f).

IODC MLSTv2

The IODC MLSTv2 product data record is too short for climate purposes, but analyzing its stability can still provide useful information to users. For that purpose, we compared IODC MLSTv2 anomalies against ERA5 and the EDLSTv2. The MODIS dataset was not used here due to the short overlap period (4 years). Figure 13 shows the time series of the difference between LST anomalies for the IODC MLSTv2 and the ERA5/EDLSTv2 datasets. Strongly decreasing trends were found for all variables, with values of -0.784 (-0.375), -0.287 (-0.447) and -0.412 (-0.378) K/decade for $LST_{MAX/DAY}$, $LST_{MIN/NIGHT}$ and LST_{MEAN} , respectively, when compared against ERA5 (EDLSTv2). Correlation with EDLSTv2 is high, with values



around 0.96-0.97, and slightly reduced for ERA5, with values between 0.79 (LST_{MAX}) and 0.94 (LST_{MIN}). Compared to CRUTEM5, there is also a significant trend of 0.755 K/decade, the correlation being similar to ERA5 (0.79).

610 The trends described above seem to be associated to a significant shift in ΔT_{anom} values occurring around 2022, which corresponds to the transition from MSG1 to MSG2 (Figure 1). Indeed, the average ΔT_{anom} w.r.t to EDLSTv2 is between 0.06K and 0.07K for MSG1, and between -0.13K and -0.11K for MSG2 (Table A4). Compared to ERA5 this transition is particularly significant for LST_{MAX} , with values of mean ΔT_{anom} of 0.11K and -0.20K for MSG1 and MSG2, respectively.

615 For the IODC coverage, the transition from MSG1 to MSG2 was accompanied by a repositioning of the satellite from 41.5°E to 45.5°E. Therefore, it is difficult to ascertain whether these discrepancies are related to the sensor's calibration or a change in FOV (e.g., changing directional effects). Calibration coefficients are currently not available for the IODC coverage but could be made available by EUMETSAT in the near future.

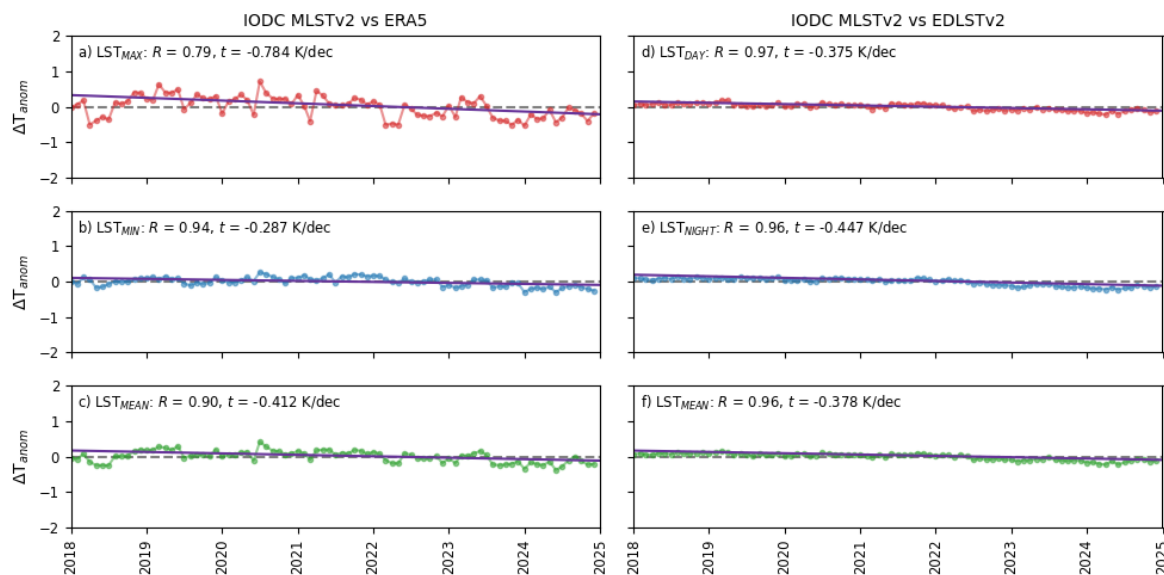


Figure 13 – As in Figure 10 but for the IODC MLSTv2 product when compared to ERA5 (a,b,c) and EDLSTv2 (d,e,f).

620 MLST-ASv2

Finally, the MLST-ASv2 anomalies were compared against ERA5 ones (Figure 14) and to CRUTEM5 (Figure 11), since to our knowledge, there are no all-sky satellite LST products that currently meet GCOS requirements and that could, therefore, be used as reference for this comparison. The obtained trends of ΔT_{anom} between MLST-ASv2 and ERA-5 (all sky) are 0.132, -0.126 and -0.059 K/decade for LST_{MAX} , LST_{MIN} and LST_{MEAN} , respectively (Table 5; Figure 14). These are slightly

625 larger than those obtained for MLSTv2, but still meeting the GCOS baseline requirement of 0.2 K/decade. Correlations are also higher, with values of 0.94, 0.95 and 0.97 for LST_{MAX} , LST_{MIN} and LST_{MEAN} , respectively (Figure 14). Compared to the clear-sky counterpart, MLSTv2, correlations are significantly improved during the day (from 0.89 to 0.94), likely because the cloudy-sky LSTs are more comparable to ERA5 values. It should be noted that, unlike for MLSTv2 where reanalysis and satellite products are carefully matched up in time, while ensuring that both correspond to clear-sky estimates, the MLST-



ASv2 data are matched independent of cloud conditions. Therefore, differences in cloud coverage between the MLST-ASv2 and ERA5 can lead to higher discrepancies. This could explain why nighttime correlations are higher under clear-sky (0.97; Figure 10b) than all-sky (0.95; Figure 14b). When compared against CRUTEM5, the ΔT_{anom} trend is actually improved, being non-significant with a value of 0.008 K/decade and meeting the GCOS target requirement. Since the MLST-ASv2 incorporates the MLSTv2 product, these results are indicative of the importance of comparing clear-sky to all-sky values. Indeed, trends in cloud cover are likely to have a different impact on an all-sky and clear-sky LST trends (Anon, 2013; Dong et al., 2023; Ermida et al., 2019; Norris et al., 2016; Pfeifroth et al., 2018b, a).

In Figure 14 there is a small but noticeable increase (decrease) of LST_{MAX} (LST_{MIN}) ΔT_{anom} around 2021-2022. The MLST-ASv2 cloudy-sky LST is produced by the METv3 algorithm. For the generation of the MLST-ASv2 product, we included data from both the CDR (2004-2020) and NRT production (2021 onwards). This seems to introduce a slight discontinuity in the dataset, with an average ΔT_{anom} of the MLST-ASv2 during the METv3 CDR (NRT) period of -0.05K (0.20K), 0.04K (-0.15K) and 0.01K (-0.05K) for LST_{MAX} , LST_{MIN} and LST_{MEAN} , respectively (Table A5).

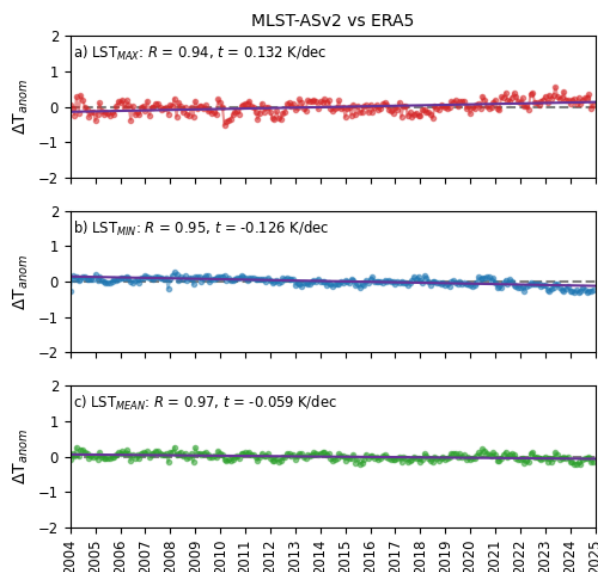


Figure 14 - As in Figure 10 but for the MLST-ASv2 product when compared to ERA5.

4. Summary and Recommendations

This work describes the new developments introduced in the generation of LST from EUMETSAT satellite by the LSA-SAF. The so-called version 2 suite of products includes LST from SEVIRI on-board the MSG series, for both the 0-degree (MLSTv2) and IODC (IODC MLSTv2) coverage, and from AVHRR on-board Metop satellites (EDLSTv2). It also includes an all-sky product (MLST-ASv2) that combines clear-sky LST from MLSTv2 with cloudy skin temperature from a surface energy balance model developed within the LSA-SAF (METv3). The main algorithm developments with respect to the older operational version include:



1. Re-calibration of TOA radiances to ensure consistency across sensors in each series;
2. Revision of the calibration database used to train the LST retrieval algorithm. The new version uses a more comprehensive database that ensures a better representation of extreme conditions. Furthermore, the new database uses profile data from a more recent reanalysis, thereby improving the vertical description of the atmosphere;
- 655 3. The land surface emissivity algorithm was revised to improve the representation of spatial variability in bare and sparsely vegetated surfaces. Additionally, the algorithm uses the algorithm also uses more recent versions of spectral library data and land cover and soil type classifications.

The generated LST datasets were validated against in-situ data from four different station networks: KIT, BSRN, SURFRAD and GBOV. However, our analysis seems to indicate that GBOV data might not be representative of satellite spatial
 660 resolutions, especially in the case of the coarser geostationary data. The overall accuracy of the LSA-SAF datasets, considering KIT, BSRN and SURFRAD data, is 0.20K, 0.21K, -0.22K and 0.24K for MLSTv2, IODC MLSTv2, EDLSTv2 and MLST-ASv2, respectively. The overall precision is 1.51K, 1.53K, 1.91K and 1.75K for MLSTv2, IODC MLSTv2, EDLSTv2 and MLST-ASv2, respectively. The total uncertainty of the products is 1.98K, 2.07K, 2.40K and 2.35K. The products also show consistent performance for day/nighttime and for very high and very low LST values.

665 The products were also analyzed for their consistency, showing good agreement when disregarding directional effects. During daytime, at least 84% of pixels have a mean difference within ± 2 K. However, directional effects are an important caveat when using satellite LST products: for specific surface conditions, LST discrepancies due to satellite view angle differences can be significant (Duffour et al., 2016; Ermida et al., 2017; Lagouarde et al., 2000; Pinheiro et al., 2004; Rasmussen et al., 2011). The LSA-SAF currently provides an estimate of the deviation from nadir view for the MLST
 670 product 68; this data layer will also be included in the new MLSTv2. While there are currently no new developments, the LSA-SAF will continue to investigate methodologies to account for directional effects on its LST products.

The LSA-SAF also aims to provide LST datasets that meet the GCOS requirements for climate applications. Therefore, we analyze the products' stability by comparing LST anomalies with three reference datasets, namely the CRUTEM5, ERA5 skin temperature and MODIS/Aqua from the ESA CCI LST project. We found that the MLSTv2 and EDLSTv2 products
 675 meet the GCOS baseline requirement of 0.2 K/decade when compared to CRUTEM5 and the target 0.1 K/decade requirement when compared to ERA5/MODIS. Conversely, the MLST-ASv2 product meets the target requirement when compared to CRUTEM5, while meeting only the baseline when compared to ERA5. This is likely due to the use of clear-sky LST instead of all-sky air temperature, which can introduce discrepancies with a climatic signal (e.g., from cloud cover). The IODC MLSTv2 product did not meet the GCOS requirements. However, this is a short dataset (starting in 2018) and trends
 680 are most likely associated to the re-positioning of the satellite position when transitioning from MSG1 to MSG2. Nevertheless, we expect new radiance calibration coefficients to be available in the near future and will then make an assessment for the need to release a new version of this dataset.

Although significant efforts have been made to obtain stable recalibrated radiances for generating LSA-SAF LST datasets, there are still some small discontinuities. Our analysis suggests that MSG3 is slightly colder, while MSG4 is slightly warmer



than the other MSGs, although the largest differences are within 0.15K. For the EDLSTv2 product, we identify a slight shift towards colder LST from 2024 onwards, associated with the orbital drift that was initiated to conserve fuel. We expect that data from the Metop-SG will be available soon, which will allow replacing any data from EPS data does not meet requirements. We also identified a small discontinuity in the MLST-ASv2 product around 2022, which we assign to the transition from METv3 CDR to METv3 NRT data, with changes of about 0.10-0.15K.

The algorithms presented here have also been adapted and implemented for retrieving LST from the new Meteosat Third Generation (MTG) and Metop-SG, thereby ensuring continuity of the Meteosat and Metop services. The overlapping period between legacy and follow-on missions will be crucial to further ensure the consistency of LSA-SAF LST datasets derived from geostationary and polar-orbiting EUMETSAT satellites.

Appendix A

Table A1 – Coordinates, used time-period, land cover, source of emissivity and LSA-SAF products validated by each in-situ station.

ID	Coordinates	Time period	Land Cover	Emissivity	Products validated
KIT Network					
EVO	38.540 °N 8.003 °W	2010-2024	Savanna	LSA-SAF MLST	MLSTv2 (IODC); MLST-ASv2; EDLSTv2
GBB_P	23.520°S 15.083°E	2013-2024	Desert	In-situ 40,49	MLSTv2 (IODC); MLST-ASv2; EDLSTv2
KAL	22.933°S 17.992 °E	2011-2017	Open Shrubland	LSA-SAF MLST	MLSTv2; MLST-ASv2; EDLSTv2
BOD	47.605 °N 9.444 °E	2023-2024	Inland water	0.986 (Masuda et al., 1988)	MLSTv2; EDLSTv2
KIT_F	49.091 °N 8.425 °E	2020-2024	Mixed Forest	From spectral library: 0.988	EDLSTv2
BSRN Network					
CAB	51.971 °N 4.927 °E	2014-2024	Cropland Mosaics	Monthly mean from Ermida et al. (2020)	MLSTv2 (IODC); MLST-ASv2; EDLSTv2
BUD	47.4291 °N 19.1822 °E	2019-2024	Cropland Mosaics	Monthly mean from Ermida et al. (2020)	MLSTv2 IODC
DOM	75.1000 °S 123.3830 °E	2010-2022	Snow and Ice	From spectral library: 0.982	EDLSTv2
GVN	70.6500 °S	2007-2022	Snow and Ice	From spectral library:	EDLSTv2



	8.2500 °W			0.982	
SPO	89.9830 °S 24.7990 °W	2007-2021	Snow and Ice	From spectral library: 0.982	EDLSTv2
SEL	15.7840 °N 91.9902 °W	2020-2024	Savanna	Monthly mean from Ermida et al. (2020)	EDLSTv2
PAY	46.8123 °N 6.9422 °E	2004-2023	Evergreen Needleleaf Forest	Monthly mean from Ermida et al. (2020)	MLSTv2; MLST-ASv2
TAT	36.0581 °N 140.1258 °E	2007-2024	Urban and Built- up	Monthly mean from Ermida et al. (2020)	EDLSTv2
TOR	58.2540 °N 26.4620 °E	2004-2020	Cropland Mosaics	Monthly mean from Ermida et al. (2020)	MLSTv2 (IODC); MLST-ASv2; EDLSTv2
SURFRAD Network					
GWN	34.255 °N 89.873 °W	2007-2024	Deciduous Broadleaf Forest	Monthly mean from Ermida et al. (2020)	EDLSTv2
BND	40.051°N 88.373°W	2007-2024	Croplands		EDLSTv2
SXF	43.734 °N 96.623 °W	2007-2024	Croplands		EDLSTv2
FPK	48.308 °N 105.102 °W	2007-2024	Grasslands		EDLSTv2
DRA	36.623 °N 116.020 °W	2007-2024	Open Shrublands		EDLSTv2
GBOV Database					
ALICE	22.2832°S 133.2511°E	2016-2022	Open Shrublands	(Seemann et al., 2008)	EDLSTv2
BE-VI	50.3050°N 5.9981°E	2021-2024	Mixed Forest		EDLSTv2
BRASS	51.3075°N 4.5199°E	2018-2024	Mixed Forest		MLSTv2; MLST-ASv2; EDLSTv2
CALPE	34.0027°S 140.5877°E	2018-2024	Closed Shrublands		EDLSTv2
DALYU	14.1592°S 131.3881°E	2018-2024	Woody Savannas		EDLSTV2



DES_M	41.5563°N 93.2855°W	2013-2024	Croplands		EDLSTV2
DRYRI	15.2590°S 132.3710°E	2018-2024	Woody Savannas		EDLSTV2
DUKEF	35.9712°N 79.0934°W	2016-2024	Mixed Forest		EDLSTV2
FR-BI	44.4937°N 0.9561°W	2020-2024	Evergreen Needleleaf Forest		MLSTv2; MLST-ASv2; EDLSTv2
FYODO	56.4476°N 32.9019°E	2018-2023	Croplands		MLSTv2 (IODC); MLST-ASv2; EDLSTv2
GINGI	31.3764°S 115.7139°E	2018-2024	Woody Savannas		EDLSTV2
GREAT	30.1920°S 120.6540°E	2018-2023	Open Shrublands		EDLSTV2
GRIGN	48.8442°N 1.9519°E	2022-2024	Croplands		MLSTv2; MLST-ASv2; EDLSTv2
HYYTI	61.8474°N 24.2948°E	2022-2024	Evergreen Needleleaf		MLSTv2; MLST-ASv2; EDLSTv2
IT-SR	43.7320°N 10.2909°E	2020-2023	Evergreen Needleleaf Forest		MLSTv2; MLST-ASv2; EDLSTv2
KANPU	26.5332°N 80.1683°E	2019-2021	Croplands		EDLSTV2
LOCHR	51.1122°N 3.8504°E	2019-2022	Deciduous Broadleaf Forest		MLSTv2 (IODC); MLST-ASv2; EDLSTv2
MANHA	39.1027°N 96.6097°W	2014-2024	Grasslands		EDLSTV2
PUECH	43.7400°N 3.6000°E	2022-2024	Evergreen Needleleaf		MLSTv2 (IODC); MLST-ASv2; EDLSTv2
ROBSO	17.1175°S 145.6301°E	2018-2024	Mixed Forest		EDLSTV2
SE-HT	56.0976°N 13.4190°E	2019-2024	Evergreen Needleleaf Forest		MLSTv2; MLST-ASv2; EDLSTv2
SKUKU	25.0197°S	2022-2022	Savannas		MLSTv2; MLST-ASv2;



	31.4969°E			EDLSTv2
SOUTH	36.6058°N 97.4888°W	2013-2023	Croplands	EDLSTV2
STURT	17.1507°S 133.3502°E	2020-2024	Savannas	EDLSTV2
SVART	64.2600°N 19.7700°E	2022-2023	Evergreen Needleleaf	EDLSTV2
THARA	50.9626°N 13.5651°E	2021-2024	Evergreen Needleleaf Forest	EDLSTV2
TUMBA	35.6565°S 148.1516°E	2019-2019	Evergreen Broadleaf	EDLSTV2
VALEN	39.5707°N 1.2882°W	2020-2024	Cropland Mosaics	MLSTv2 (IODC); MLST-ASv2; EDLSTv2
WARRA	43.0950°S 146.6545°E	2018-2021	Evergreen Broadleaf	EDLSTV2
WHROO	36.6732°S 145.0294°E	2020-2022	Croplands	EDLSTV2
WICKE	52.3094°N 0.2797°E	2017-2021	Croplands	MLSTv2 (IODC); MLST-ASv2; EDLSTv2
WILLI	35.7550°N 112.3374°W	2013-2024	Open Shrublands	EDLSTV2
WUSTE	50.5049°N 6.3310°E	2016-2022	Evergreen Needleleaf Forest	MLSTv2; MLST-ASv2; EDLSTv2
WYTHA	51.7772°N 1.3383°W	2021-2022	Croplands Mosaics	MLSTv2 (IODC); MLST-ASv2; EDLSTv2
YANCO	35.0053°S 146.3099°E	2018-2024	Grasslands	EDLSTV2
YOSEM	37.7589°N 119.8211°W	2016-2024	Evergreen Needleleaf	EDLSTV2



Table A2 - Average and standard deviation of the anomaly differences (ΔT_{anom} in eq. 10; K) for each satellite in the MLSTv2 and MLST-ASv2 time series.

		MLSTv2 vs ERA5	MLSTv2 vs MODIS	MLST-ASv2 vs ERA5
LST _{MAX/DAY}	MSG1	0.00 ± 0.31	0.01 ± 0.12	-0.06 ± 0.18
	MSG2	-0.06 ± 0.27	0.04 ± 0.09	-0.08 ± 0.17
	MSG3	-0.06 ± 0.23	-0.06 ± 0.09	0.04 ± 0.18
	MSG4	0.15 ± 0.26	0.01 ± 0.08	0.09 ± 0.16
LST _{MIN/NIGHT}	MSG1	0.03 ± 0.11	0.05 ± 0.10	0.06 ± 0.08
	MSG2	0.04 ± 0.09	0.02 ± 0.09	0.08 ± 0.07
	MSG3	-0.05 ± 0.08	-0.05 ± 0.10	-0.06 ± 0.11
	MSG4	0.01 ± 0.13	-0.01 ± 0.06	-0.05 ± 0.09
LST _{MEAN}	MSG1	0.04 ± 0.14	0.04 ± 0.09	0.04 ± 0.09
	MSG2	0.00 ± 0.14	0.02 ± 0.08	0.03 ± 0.09
	MSG3	-0.06 ± 0.09	-0.05 ± 0.08	-0.04 ± 0.09
	MSG4	0.07 ± 0.13	0.00 ± 0.05	-0.01 ± 0.08

705 **Table A3 - Average and standard deviation of the anomaly differences (ΔT_{anom} in eq. 10; K) for each satellite in the EDLSTv2 time series.**

		EDLSTv2 vs ERA5	EDLSTv2 vs MLSTv2
LST _{DAY}	Metop-A	0.03 ± 0.08	-0.05 ± 0.09
	Metop-B	-0.02 ± 0.08	0.04 ± 0.11
LST _{NIGHT}	Metop-A	0.01 ± 0.09	-0.03 ± 0.04
	Metop-B	-0.01 ± 0.08	0.02 ± 0.07
LST _{MEAN}	Metop-A	0.03 ± 0.05	-0.04 ± 0.06
	Metop-B	-0.02 ± 0.05	0.03 ± 0.09

Table A4 - Average and standard deviation of the anomaly differences (ΔT_{anom} in eq. 10(10; K) for each satellite in the IODC MLSTv2 time series.

		IODC MLSTv2 vs ERA5	IODC MLSTv2 vs EDLSTv2
LST _{MAX/DAY}	MSG1	0.11 ± 0.29	0.06 ± 0.05
	MSG2	-0.20 ± 0.21	-0.11 ± 0.05
LST _{MIN/NIGHT}	MSG1	0.06 ± 0.10	0.07 ± 0.03
	MSG2	-0.11 ± 0.10	-0.13 ± 0.05
LST _{MEAN}	MSG1	0.07 ± 0.14	0.06 ± 0.03
	MSG2	-0.13 ± 0.13	-0.12 ± 0.05



Table A5 - Average and standard deviation of the anomaly differences (ΔT_{anom} in eq. 10; K) of the MLST-ASv2 time series, for the periods of the Climate Data Record (CDR) and Near Real Time (NRT) productions of the METv3 product.

		MLST-ASv2 vs ERA5
LST _{MAX}	MET CDR	-0.05 ± 0.16
	MET NRT	0.20 ± 0.13
LST _{MIN}	MET CDR	0.04 ± 0.08
	MET NRT	-0.15 ± 0.09
LST _{MEAN}	MET CDR	0.01 ± 0.09
	MET NRT	-0.05 ± 0.09

715 **Table A6 - Average and standard deviation of the anomaly differences (ΔT_{anom} in eq. 10; K) for each satellite in the MLST (previous version) time series.**

		MLST vs ERA5	MLST vs MODIS
LST _{MAX/DAY}	MSG1	-0.14 ± 0.25	-0.12 ± 0.10
	MSG2	-0.06 ± 0.24	0.04 ± 0.10
	MSG3	-0.24 ± 0.23	-0.23 ± 0.09
	MSG4	0.49 ± 0.26	0.34 ± 0.08
LST _{MIN/NIGHT}	MSG1	-0.08 ± 0.10	-0.06 ± 0.09
	MSG2	-0.05 ± 0.09	-0.03 ± 0.09
	MSG3	-0.10 ± 0.11	-0.12 ± 0.11
	MSG4	0.24 ± 0.14	0.26 ± 0.07
LST _{MEAN}	MSG1	-0.09 ± 0.12	-0.08 ± 0.08
	MSG2	-0.06 ± 0.12	0.00 ± 0.08
	MSG3	-0.16 ± 0.13	-0.17 ± 0.09
	MSG4	0.34 ± 0.16	0.30 ± 0.06

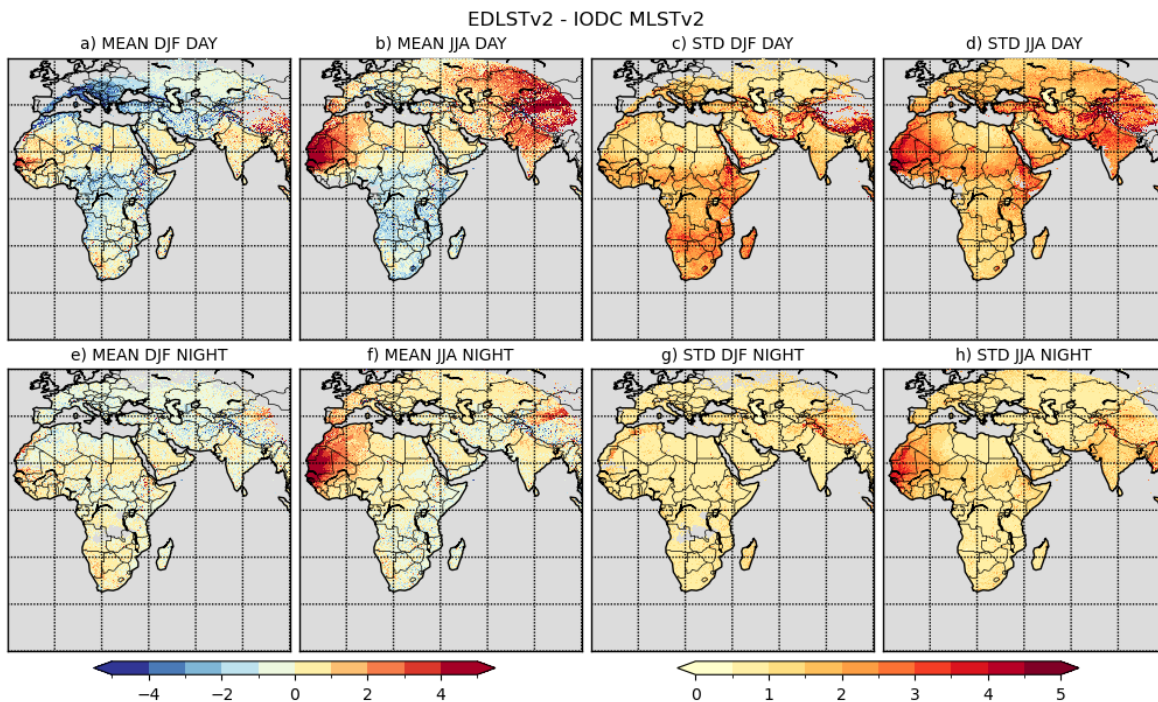


Figure A1 - Statistics of the differences between IODC MLSTv2 and EDLSTv2 (as EDLSTv2 minus IODC MLSTv2; K), namely the mean (a,b,e,f) and the standard deviation (STD; c,d,g,h), for the months December-January-February (DJF; a,c,e,g) and June-July- August (JJA;b,d,f,h) of 1980 to 2024.

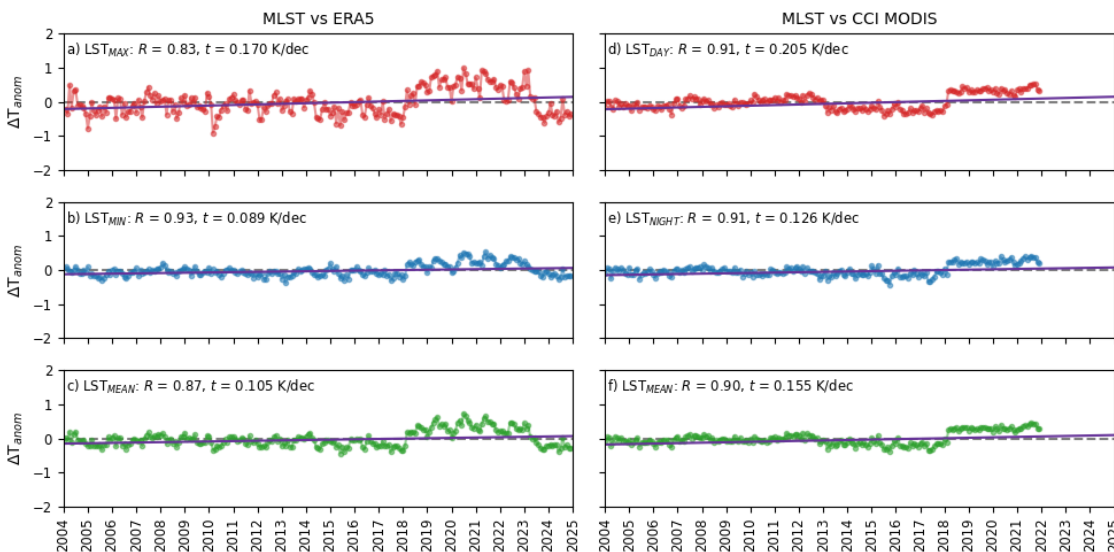


Figure A2 - Monthly anomaly difference (ΔT_{anom} in eq. 10(10) time series as obtained for the MLST (previous version) product when compared to ERA5 (a,b,c) and MODIS (d,e,f), for the maximum/daytime (LST_{MAX/DAY}; a,d), the minimum/nighttime (LST_{MIN/NIGHT}; b,e) and mean LST (LST_{MEAN}; c,f). The trend of each series is represented by the purple line. The respective values of the trend (t ; K/decade) and correlation coefficient (R^2) are also shown.



Data availability

- 730 Monthly averages from each product for the respective data record period on a 0.25°x0.25° regular grid are available from <https://doi.org/10.5281/zenodo.20705177> (Ermida et al., 2026)

Products can be accessed through the LSA-SAF data server:

- 735 <https://data.sasaf.lsasvcs.ipma.pt/PRODUCTS/MSG/MLSTv2/>
<https://data.sasaf.lsasvcs.ipma.pt/PRODUCTS/MSG-IODC/MLSTv2/>
<https://data.sasaf.lsasvcs.ipma.pt/PRODUCTS/MSG/MLST-ASv2/>
<https://data.sasaf.lsasvcs.ipma.pt/PRODUCTS/EPS/EDLSTv2/>

Information on all LSA-SAF products is available at: <https://lsa-saf.eumetsat.int/en/>

740 Code Availability

The Google Earth engine code used to select the validation stations is available at:

https://code.earthengine.google.com/?accept_repo=users/sofiaermida/LST_validation

Author contributions

- 745 Conceptualization: SE, ED, IT; Methodology: SE, ED, IT; Investigation: SE; Resources: SE, ED, SG; Software: SE, ED, SG; Data Curation: SE, ED, SG, FG; Formal Analysis: SE, FL; Validation: SE, IT; Visualization: SE; Writing – Original Draft: SE; Writing – Review and Editing: ED, FL, SG, FG, IT; Supervision: IT; Project Administration: IT; Funding Acquisition: IT. All authors have read and agreed to the published version of the manuscript.

References

- 750 Agathangelidis, I., Cartalis, C., Polydoros, A., Mavroukou, T., and Philippopoulos, K.: Can Satellite-Based Thermal Anomalies Be Indicative of Heatwaves? An Investigation for MODIS Land Surface Temperatures in the Mediterranean Region, *Remote Sensing* 2022, Vol. 14, Page 3139, 14, 3139, <https://doi.org/10.3390/RS14133139>, 2022.
- Albergel, C., Balsamo, G., de Rosnay, P., Muñoz-Sabater, J., and Boussetta, S.: A bare ground evaporation revision in the ECMWF land-surface scheme: evaluation of its impact using ground soil moisture and satellite microwave data, *Hydrol. Earth Syst. Sci.*, 16, 3607–3620, <https://doi.org/10.5194/hess-16-3607-2012>, 2012.
- 755 de Almeida, C. R., Teodoro, A. C., and Gonçalves, A.: Study of the Urban Heat Island (UHI) Using Remote Sensing Data/Techniques: A Systematic Review, *Environments* 2021, Vol. 8, Page 105, 8, 105, <https://doi.org/10.3390/ENVIRONMENTS8100105>, 2021.



- Anand, J., Dodd, E., and Nativel, S.: Algorithm Theoretical Basis Document for Land Surface Temperature products: RM2 (Direct and diffuse thermal radiation), RM8 (Emissivity) and RM9 (Surface temperature), Ground-Based Observations for Validation (GBOV) of Copernicus Global Land Products, version 1.6, 2023.
- Anon: Clouds and aerosols, Climate Change 2013 the Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, 9781107057999, 571–658, <https://doi.org/10.1017/CBO9781107415324.016>, 2013.
- Augustine, J. A., Hodges, G. B., Cornwall, C. R., Michalsky, J. J., Medina, C. I., Augustine, J. A., Hodges, G. B., Cornwall, C. R., Michalsky, J. J., and Medina, C. I.: An Update on SURFRAD—The GCOS Surface Radiation Budget Network for the Continental United States, *J. Atmos. Ocean. Technol.*, 22, 1460–1472, <https://doi.org/10.1175/JTECH1806.1>, 2005.
- Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M., and Betts, A. K.: A Revised Hydrology for the ECMWF Model: Verification from Field Site to Terrestrial Water Storage and Impact in the Integrated Forecast System, *J. Hydrometeorol.*, 10, 623–643, <https://doi.org/10.1175/2008JHM1068.1>, 2009.
- Barrios, J. M., Arboleda, A., Dutra, E., Trigo, I., and Gellens-Meulenberghs, F.: Evapotranspiration and surface energy fluxes across Europe, Africa and Eastern South America throughout the operational life of the Meteosat second generation satellite, *Geosci. Data J.*, <https://doi.org/10.1002/gdj3.235>, 2024.
- Berk, A., Conforti, P., Kennett, R., Perkins, T., Hawes, F., and van den Bosch, J.: MODTRAN6: a major upgrade of the MODTRAN radiative transfer code, 90880H, <https://doi.org/10.1117/12.2050433>, 2014.
- Bojinski, S., Verstraete, M., Peterson, T. C., Richter, C., Simmons, A., and Zemp, M.: The Concept of Essential Climate Variables in Support of Climate Research, Applications, and Policy, *Bull. Am. Meteorol. Soc.*, 95, 1431–1443, <https://doi.org/10.1175/BAMS-D-13-00047.1>, 2014.
- Buchhorn, M., Smets, B., Bertels, L., Roo, B. De, Lesiv, M., Tsendbazar, N.-E., Herold, M., and Fritz, S.: Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2019: Globe, <https://doi.org/10.5281/ZENODO.3939050>, 2020.
- Carrer, D., Moparthy, S., Lellouch, G., Ceamanos, X., Pinault, F., Freitas, S., and Trigo, I.: Land Surface Albedo Derived on a Ten Daily Basis from Meteosat Second Generation Observations: The NRT and Climate Data Record Collections from the EUMETSAT LSA SAF, *Remote Sens. (Basel)*, 10, 1262, <https://doi.org/10.3390/rs10081262>, 2018.
- Carrer, D., Ceamanos, X., Moparthy, S., Vincent, C., C. Freitas, S., and Trigo, I. F.: Satellite Retrieval of Downwelling Shortwave Surface Flux and Diffuse Fraction under All Sky Conditions in the Framework of the LSA SAF Program (Part 1: Methodology), *Remote Sens. (Basel)*, 11, 2532, <https://doi.org/10.3390/rs11212532>, 2019.
- Caselles, E., Valor, E., Abad, F., and Caselles, V.: Automatic classification-based generation of thermal infrared land surface emissivity maps using AATSR data over Europe, *Remote Sens. Environ.*, 124, 321–333, <https://doi.org/10.1016/J.RSE.2012.05.024>, 2012.
- Cheda, A., Tetzlaff, A., Blannin, J., Good, E., Sharma, V., Trigo, I., Schwaab, J., Riihelä, A., Grams, C. M., and Schröder, M.: Heat Indices for Europe Derived from Satellite Data: A Proof of Concept, *Remote Sensing* 2026, Vol. 18, Page 589, 18, 589, <https://doi.org/10.3390/RS18040589>, 2026.



- Chevallier, F., Chédin, A., Cheruy, F., and Morcrette, J.-J.: TIGR-like atmospheric-profile databases for accurate radiative-flux computation, *Quarterly Journal of the Royal Meteorological Society*, 126, 777–785, <https://doi.org/10.1002/qj.49712656319>, 2000.
- 795 Coppens, D. and Ackermann, J.: Quantification of Potential Metop-A and Metop-B AVHRR/3 IR Channel Radiance Biases Using Co-located IASI Spectra, Internal Report, 2014.
- Dong, B., Sutton, R. T., and Wilcox, L. J.: Decadal trends in surface solar radiation and cloud cover over the North Atlantic sector during the last four decades: drivers and physical processes, *Clim. Dyn.*, 60, 2533–2546, <https://doi.org/10.1007/S00382-022-06438-3/FIGURES/11>, 2023.
- 800 Driemel, A., Augustine, J., Behrens, K., Colle, S., Cox, C., Cuevas-Agulló, E., Denn, F. M., Duprat, T., Fukuda, M., Grobe, H., Haeffelin, M., Hodges, G., Hyett, N., Ijima, O., Kallis, A., Knap, W., Kustov, V., Long, C. N., Longenecker, D., Lupi, A., Maturilli, M., Mimouni, M., Ntsangwane, L., Ogihara, H., Olano, X., Olefs, M., Omori, M., Passamani, L., Pereira, E. B., Schmithüsen, H., Schumacher, S., Sieger, R., Tamlyn, J., Vogt, R., Vuilleumier, L., Xia, X., Ohmura, A., and König-Langlo, G.: Baseline Surface Radiation Network (BSRN): structure and data description (1992–2017), *Earth Syst. Sci. Data*, 10, 1491–1501, <https://doi.org/10.5194/essd-10-1491-2018>, 2018.
- 805 Duffour, C., Lagouarde, J. P., Olioso, A., Demarty, J., and Roujean, J. L.: Driving factors of the directional variability of thermal infrared signal in temperate regions, *Remote Sens. Environ.*, 177, <https://doi.org/10.1016/j.rse.2016.02.024>, 2016.
- Ermida, S. L. and Trigo, I. F.: A Comprehensive Clear-Sky Database for the Development of Land Surface Temperature Algorithms, *Remote Sens. (Basel)*, 14, 2329, <https://doi.org/10.3390/rs14102329>, 2022.
- 810 Ermida, S. L., Trigo, I. F., DaCamara, C. C., Göttsche, F. M., Olesen, F. S., and Hulley, G.: Validation of remotely sensed surface temperature over an oak woodland landscape — The problem of viewing and illumination geometries, *Remote Sens. Environ.*, 148, 16–27, 2014.
- Ermida, S. L., DaCamara, C. C., Trigo, I. F., Pires, A. C., Ghent, D., and Remedios, J.: Modelling directional effects on remotely sensed land surface temperature, *Remote Sens. Environ.*, 190, 56–69, <https://doi.org/10.1016/j.rse.2016.12.008>, 2017.
- 815 Ermida, S. L., Trigo, I. F., DaCamara, C. C., and Pires, A. C.: A Methodology to Simulate LST Directional Effects Based on Parametric Models and Landscape Properties, *Remote Sens. (Basel)*, 10, 1114, <https://doi.org/10.3390/rs10071114>, 2018.
- Ermida, S. L., Trigo, I. F., DaCamara, C. C., Jiménez, C., and Prigent, C.: Quantifying the Clear-Sky Bias of Satellite Land Surface Temperature Using Microwave-Based Estimates, *Journal of Geophysical Research: Atmospheres*, 124, <https://doi.org/10.1029/2018JD029354>, 2019.
- 820 Ermida, S. L., Soares, P., Mantas, V., Göttsche, F.-M., and Trigo, I. F.: Google Earth Engine Open-Source Code for Land Surface Temperature Estimation from the Landsat Series, *Remote Sens. (Basel)*, 12, 1471, <https://doi.org/10.3390/rs12091471>, 2020.



- Ermida, S. L., Hulley, G., Goettsche, F. M., and Trigo, I. F.: A combined Vegetation cover and Temperature-Emissivity Separation (V-TES) method to estimate land surface emissivity, *IEEE Transactions on Geoscience and Remote Sensing*, 1–1, <https://doi.org/10.1109/TGRS.2023.3301615>, 2023.
- Ermida, S. L., Hulley, G., and Trigo, I. F.: Introducing emissivity directionality to the temperature-emissivity separation algorithm, *Remote Sens. Environ.*, 311, 114280, <https://doi.org/10.1016/J.RSE.2024.114280>, 2024.
- Ermida, S. L., Dutra, E., Gomes, S., and Isabel, F. T.: LSA-SAF Monthly SEVIRI and AVHRR Land Surface Temperature from version 2 suite at 0.25°x0.25° (Version v2) [Data set], Zenodo, <https://doi.org/10.5281/ZENODO.20705177>, 2026.
- FAO-UNESCO: Digital Soil Map of the World, <https://data.apps.fao.org/map/catalog/srv/eng/catalog.search#/metadata/446ed430-8383-11db-b9b2-000d939bc5d8>, 2007.
- Freitas, S. C., Trigo, I. F., Bioucas-dias, J. M., and Götsche, F.: Quantifying the Uncertainty of Land Surface Temperature Retrievals From SEVIRI / Meteosat, *IEEE Transactions on Geoscience and Remote Sensing*, 48, 523–534, 2010.
- García-Haro, F. J., Camacho, F., Martínez, B., Campos-Taberner, M., Fuster, B., Sánchez-Zapero, J., and Gilabert, M. A.: Climate Data Records of Vegetation Variables from Geostationary SEVIRI/MSG Data: Products, Algorithms and Applications, *Remote Sensing* 2019, Vol. 11, Page 2103, 11, 2103, <https://doi.org/10.3390/RS11182103>, 2019.
- García-Haro, F. J., Campos-Taberner, M., Martínez, B., Sánchez-Ruiz, S., Camacho, F., and Gilabert, M. A.: LSA SAF: a long-term service of vegetation variables for modelling terrestrial ecosystems at regional and global scales, <https://doi.org/10.1117/12.2600318>, 11858, 305–312, <https://doi.org/10.1117/12.2600318>, 2021.
- Ghent, D., Veal, K., and Perry, M.: ESA Land Surface Temperature Climate Change Initiative (LST_cci): Daily land surface temperature from MODIS (Moderate resolution Infra-red Spectroradiometer) on Aqua, level 3 collated (L3C) global product (2002–2021), version 4.00, NERC EDS Centre for Environmental Data Analysis, <https://doi.org/doi:10.5285/a92c495827ef4c81a901f878adb6ef70>, 2025.
- Gillespie, A., Rokugawa, S., Matsunaga, T., Steven Cothorn, J., Hook, S., and Kahle, A. B.: A temperature and emissivity separation algorithm for advanced spaceborne thermal emission and reflection radiometer (ASTER) images, *IEEE Transactions on Geoscience and Remote Sensing*, 36, 1113–1126, <https://doi.org/10.1109/36.700995>, 1998.
- Good, E. J., Aldred, F. M., Ghent, D. J., Veal, K. L., and Jimenez, C.: An Analysis of the Stability and Trends in the LST_cci Land Surface Temperature Datasets Over Europe, *Earth and Space Science*, 9, e2022EA002317, <https://doi.org/10.1029/2022EA002317>, 2022.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R.: Google Earth Engine: Planetary-scale geospatial analysis for everyone, *Remote Sens. Environ.*, 202, 18–27, <https://doi.org/10.1016/J.RSE.2017.06.031>, 2017.
- Götsche, F.-M. and 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.*, 124, 149–158, 2012.
- Götsche, F.-M., Olesen, F.-S., and Bork-Unkelbach, A.: Validation of land surface temperature derived from MSG/SEVIRI with in situ measurements at Gobabeb, Namibia, *Int. J. Remote Sens.*, 34, 3069–3083, <https://doi.org/10.1080/01431161.2012.716539>, 2013.



- Göttsche, F.-M., Olesen, F.-S., Trigo, I. F., Bork-Unkelbach, A., and Martin, M. A.: Long term validation of land surface temperature retrieved from MSG/SEVIRI with continuous in-situ measurements in Africa, *Remote Sens. (Basel)*, 8, 410, 2016.
- Guillevic, P., Göttsche, F., Nickeson, J., Hulley, G., Ghent, D., Yu, Y., Trigo, I., Hook, S., Sobrino, J. A., Remedios, J., Román, M., and Camacho, F.: Land Surface Temperature Product Validation Best Practice Protocol. Version 1.1., In P. Guillevic, F. Göttsche, J. Nickeson & M. Román (Eds.), *Best Practice for Satellite-Derived Land Product Validation* (p. 58): Land Product Validation Subgroup (WGCV/CEOS), <https://doi.org/10.5067/doc/ceoswgcvt/lpv/lst.001>, 2018.
- He, K., Ignatov, A., Kihai, Y., Cao, C., Stroup, J., Editors, A., Zhao, X., Li, X., and Thenkabail, P. S.: Sensor Stability for SST (3S): Toward Improved Long-Term Characterization of AVHRR Thermal Bands, *Remote Sensing* 2016, Vol. 8, Page 346, 8, 346, <https://doi.org/10.3390/RS8040346>, 2016.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.: The ERA5 global reanalysis, *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049, <https://doi.org/10.1002/qj.3803>, 2020.
- H-SAF: Scatterometer root zone soil moisture (RZSM) data record 10km resolution – multimission, Technical Report, EUMETSAT SAF on Support to Operational Hydrology and Water Management, 2020.
- Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, *Atmos. Chem. Phys.*, 19, 3515–3556, <https://doi.org/10.5194/acp-19-3515-2019>, 2019.
- Jiménez, C., Prigent, C., Ermida, S. L., and Moncet, J.-L.: Inversion of AMSR-E observations for land surface temperature estimation: 1. Methodology and evaluation with station temperature, *Journal of Geophysical Research: Atmospheres*, 122, 3330–3347, <https://doi.org/10.1002/2016JD026144>, 2017.
- John, V. O., Tabata, T., Rüthrich, F., Roebeling, R., Hewison, T., Stöckli, R., and Schulz, J.: On the Methods for Recalibrating Geostationary Longwave Channels Using Polar Orbiting Infrared Sounders, *Remote Sensing* 2019, Vol. 11, Page 1171, 11, 1171, <https://doi.org/10.3390/RS11101171>, 2019.
- Lagouarde, J. P., Ballans, H., Moreau, P., Guyon, D., and Coraboeuf, D.: Experimental study of brightness surface temperature angular variations of maritime pine (*Pinus pinaster*) stands, *Remote Sens. Environ.*, 72, 17–34, [https://doi.org/10.1016/S0034-4257\(99\)00085-1](https://doi.org/10.1016/S0034-4257(99)00085-1), 2000.
- Liu, Z., Ma, J., Li, K., Pérez-Planells, L., Karinou, F., Agathangelidis, I., and Cartalis, C.: Assessing the Combined Impact of Land Surface Temperature and Droughts to Heatwaves over Europe Between 2003 and 2023, *Remote Sensing* 2025, Vol. 17, Page 1655, 17, 1655, <https://doi.org/10.3390/RS17091655>, 2025.



- Li, Z. L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F., and Sobrino, J. a: Satellite-derived land surface temperature: Current status and perspectives, *Remote Sens. Environ.*, 131, 14–37, <https://doi.org/10.1016/j.rse.2012.12.008>, 2013.
- 895 Martins, J. P. A., Trigo, I. F., Ghilain, N., Jimenez, C., Göttsche, F.-M., Ermida, S. L., Olesen, F.-S., Gellens-Meulenberghs, F., and Arboleda, A.: An all-weather land surface temperature product based on MSG/SEVIRI observations, *Remote Sens. (Basel)*, 11, <https://doi.org/10.3390/rs11243044>, 2019.
- Martins, J. P. A., Caetano, S., Pereira, C., Dutra, E., and Cardoso, R. M.: A satellite view of the exceptionally warm summer of 2022 over Europe, *Natural Hazards and Earth System Sciences*, 24, 1501–1520, [https://doi.org/10.5194/NHESS-24-1501-](https://doi.org/10.5194/NHESS-24-1501-2024)
900 2024, 2024.
- Masuda, K., Takashima, T., and Takayama, Y.: Emissivity of pure and sea waters for the model sea surface in the infrared window regions, *Remote Sens. Environ.*, 24, 313–329, [https://doi.org/10.1016/0034-4257\(88\)90032-6](https://doi.org/10.1016/0034-4257(88)90032-6), 1988.
- Meerdink, S. K., Hook, S. J., Roberts, D. A., and Abbott, E. A.: The ECOSTRESS spectral library version 1.0, *Remote Sens. Environ.*, 230, 111196, <https://doi.org/10.1016/J.RSE.2019.05.015>, 2019.
- 905 Norris, J. R., Allen, R. J., Evan, A. T., Zelinka, M. D., O’Dell, C. W., and Klein, S. A.: Evidence for climate change in the satellite cloud record, *Nature*, 536, 72–75, <https://doi.org/10.1038/nature18273>, 2016.
- Osborn, T. J., Jones, P. D., Lister, D. H., Morice, C. P., Simpson, I. R., Winn, J. P., Hogan, E., and Harris, I. C.: Land Surface Air Temperature Variations Across the Globe Updated to 2019: The CRUTEM5 Data Set, *Journal of Geophysical Research: Atmospheres*, 126, e2019JD032352, <https://doi.org/10.1029/2019JD032352>, 2021.
- 910 Pfeifroth, U., Bojanowski, J. S., Clerbaux, N., Manara, V., Sanchez-Lorenzo, A., Trentmann, J., Walawender, J. P., and Hollmann, R.: Satellite-based trends of solar radiation and cloud parameters in Europe, *Adv Sci Res*, 15, 31–37, <https://doi.org/10.5194/asr-15-31-2018>, 2018a.
- Pfeifroth, U., Sanchez-Lorenzo, A., Manara, V., Trentmann, J., and Hollmann, R.: Trends and variability of surface solar radiation in Europe based on surface- and satellite-based data records, *J Geophys Res Atmos*, 123, 1735–1754, <https://doi.org/10.1002/2017jd027418>, 2018b.
- 915 Pinheiro, a. C. T., Privette, J. L., Mahoney, R., and Tucker, C. J.: Directional effects in a daily AVHRR land surface temperature dataset over Africa, *IEEE Transactions on Geoscience and Remote Sensing*, 42, 1941–1954, <https://doi.org/10.1109/TGRS.2004.831886>, 2004.
- Rasmussen, M. O., Göttsche, F., Olesen, F., and Sandholt, I.: Directional Effects on Land Surface Temperature Estimation From Meteosat Second Generation for Savanna Landscapes, *IEEE Transactions on Geoscience and Remote Sensing*, 49, 4458–4468, 2011.
- 920 Seemann, S. W., Borbas, E. E., Knuteson, R. O., Stephenson, G. R., and Huang, H.-L.: Development of a global infrared land surface emissivity database for application to clear sky sounding retrievals from multispectral satellite radiance measurements, *J. Appl. Meteorol. Climatol.*, 47, 108–123, 2008.



- 925 Sen, P. K.: Estimates of the Regression Coefficient Based on Kendall's Tau, *J. Am. Stat. Assoc.*, 63, 1379–1389, <https://doi.org/10.1080/01621459.1968.10480934>, 1968.
- Stante, F., Ermida, S. L., DaCamara, C. C., Göttsche, F.-M., and Trigo, I. F.: Impact of High Concentrations of Saharan Dust Aerosols on Infrared-Based Land Surface Temperature Products, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 16, 4064–4079, <https://doi.org/10.1109/JSTARS.2023.3263374>, 2023.
- 930 Trigo, I. F., Monteiro, I. T., Olesen, F., and Kabsch, E.: An assessment of remotely sensed land surface temperature, *J. Geophys. Res.*, 113, 1–12, <https://doi.org/10.1029/2008JD010035>, 2008.
- Trigo, I. F., Dacamara, C. C., Viterbo, P., Roujean, J.-L., Olesen, F., Barroso, C., Camacho-de-Coca, F., Carrer, D., Freitas, S. C., García-Haro, J., Geiger, B., Gellens-Meulenberghs, F., Ghilain, N., Meliá, J., Pessanha, L., Siljamo, N., and Arboleda, A.: The Satellite Application Facility for Land Surface Analysis, *Int. J. Remote Sens.*, 32, 2725–2744, <https://doi.org/10.1080/01431161003743199>, 2011.
- 935 Trigo, I. F., Ermida, S. L., Martins, J. P. A., Gouveia, C. M., Göttsche, F.-M., and Freitas, S. C.: Validation and consistency assessment of land surface temperature from geostationary and polar orbit platforms: SEVIRI/MSG and AVHRR/Metop, *ISPRS Journal of Photogrammetry and Remote Sensing*, 175, 282–297, <https://doi.org/10.1016/J.ISPRSJPRS.2021.03.013>, 2021.
- 940 U.S. National Ice Center: IMS Daily Northern Hemisphere Snow and Ice Analysis at 1 km, 4 km, and 24 km Resolutions. (G02156, Version 1) [Data set], Boulder, Colorado USA. National Snow and Ice Data Center, <https://doi.org/https://doi.org/10.7265/N52R3PMC>, 2008.
- Viterbo, P. and Beljaars, A. C. M.: An Improved Land Surface Parameterization Scheme in the ECMWF Model and Its Validation, *J. Clim.*, 8, 2716–2748, [https://doi.org/10.1175/1520-0442\(1995\)008<2716:AILSPS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1995)008<2716:AILSPS>2.0.CO;2), 1995.
- 945 Wan, Z. and Dozier, J.: A generalized split-window algorithm for retrieving land-surface temperature from space, *Geoscience and Remote Sensing, IEEE Transactions on*, 34, 892–905, 1996.
- World Meteorological Organization, Intergovernmental Oceanographic Commission of the United Nations Educational, S. and C. O., International Science Council, United Nations Environment Programme, and Copernicus Climate Change Service: GCOS-245: The 2022 GCOS ECVs Requirements, 2022 edition - Updated in 2025, Geneva, 2025.

950