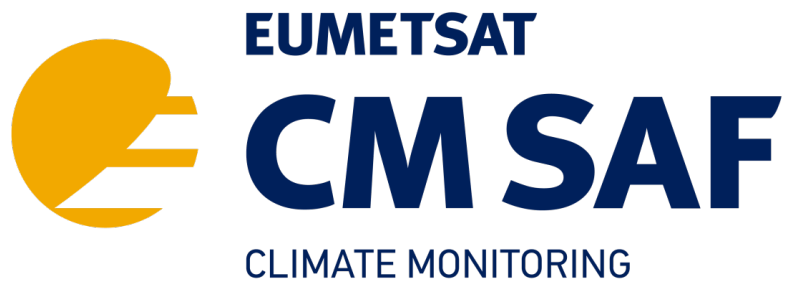


EUMETSAT Satellite Application Facility on Climate Monitoring



Validation Report

Land Surface Temperature (LST)

Edition 2

DOI: [10.5676/EUM_SAF_CM/LST_METEOSAT/V002](https://doi.org/10.5676/EUM_SAF_CM/LST_METEOSAT/V002)

Land Surface Temperature

CM-23922

Reference Number:


SAF/CM/MeteoSwiss/VAL/MET/LST/2.1

Issue/Revision Index:

2.1

Date:

07.09.2023

	Validation Report Land Surface Temperature (LST) Edition 2	Doc.: SAF/CM/MeteoSwiss/VAL/MET/LST/2.0 Issue: 2.1 Date: 07.09.2023
---	---	---

Document Signature Table

	Name	Function	Signature	Date
Authors	Anke Tetzlaff	CM SAF Scientist (MeteoSwiss)		07/09/2023
	Quentin Bourgeois	CM SAF Scientist (MeteoSwiss)		
	Jedrzej Bojanowski	CM SAF Scientist (MeteoSwiss)		
	Frank Götsche	LSA SAF Scientist (KIT)		
	Isabel Trigo	LSA SAF Science Coordinator (IPMA)		
Editor	Marc Schröder	Science Coordinator		07/09/2023
Approval	CM SAF Steering Group			
Release	Rainer Hollmann	Project Manager		


Distribution List

Internal Distribution	
Name	No. Copies
DWD Archive	1
CM SAF Team	1

External Distribution		
Company	Name	No. Copies
PUBLIC		1

Document Change Record

Issue/Revision	Date	DCN No.	Changed Pages/Paragraphs
1.0	15/08/2016	SAF/CM/MeteoSwiss/VAL/MET/LST/1	First version submitted for DRR 2.8
1.1	12/05/2017	SAF/CM/MeteoSwiss/VAL/MET/LST/1	Included changes from DRR review
2.0	21/04/2023	SAF/CM/MeteoSwiss/VAL/MET/LST/2	Second version submitted for DRR 3.9
2.1	07/09/2023	SAF/CM/MeteoSwiss/VAL/MET/LST/2.1	Reviewer comments from DRR 3.9 implemented

	Validation Report Land Surface Temperature (LST) Edition 2	Doc.: SAF/CM/MeteoSwiss/VAL/MET/LST/2.0 Issue: 2.1 Date: 07.09.2023
---	---	---

Applicable Documents

Reference	Title	Code
AD 1	CM SAF Product Requirements Document	SAF/CM/DWD/PRD/4.1

Reference Documents

Reference	Title	Code
RD 1	Product User Manual Meteosat Land Surface Temperature (LST) Edition 2	SAF/CM/MeteoSwiss/PUM/MET/LST/2.0
RD 2	Algorithm Theoretical Basis Document Meteosat Land Surface Temperature (LST) Edition 2	SAF/CM/MeteoSwiss/ATBD/MET/LST/2.1

Table of Contents

Executive Summary	8
1 The EUMETSAT SAF on Climate Monitoring.....	10
2 CM SAF Meteosat Land Surface Temperature CDR.....	12
3 Validation Strategy.....	14
4 Data sets for comparison with Meteosat LST.....	16
4.1 In-situ Observations	16
4.1.1 KIT station-based Land Surface Temperature measurements	16
4.1.2 EUSTACE 2m air temperature measurements	17
4.2 Satellite Data	18
4.2.1 ESA CCI MODIS AQUA LST CDR	18
4.2.2 ESA CCI LEO and GEO LST CDR	18
4.2.3 LSA SAF Meteosat LST CDR	19
5 Validation Results	20
5.1 Precision and Accuracy	20
5.1.1 Gobabeb Station	20
5.1.2 RMZ Station	22
5.1.3 Dahra Station	24
5.1.4 Evora Station	25
5.1.5 Summary Precision and Accuracy	27
5.2 Temporal Stability	28
5.2.1 Stability against EUSTACE 2m air temperature	28
5.2.2 Stability against ESA CCI MODIS AQUA LST	31
5.2.3 Summary Stability	34
5.3 Comparison against other LST CDRs	34
6 Conclusions.....	38
7 References.....	40
8 Glossary.....	43

List of Figures

Figure 0-1: Monthly mean time series of the CM SAF LST (red) as compared to homogenized T2m air temperature measurements (blue) at 466 stations over Europe. Top: monthly mean air temperature, bottom: monthly mean air temperature anomaly (seasonal corrected).....	8
Figure 2-1: Availability of MFG and MSG measurements.	12
Figure 4-1: Locations of the Karlsruhe Institute of Technology’s (KIT) validation stations on the SEVIRI disk.....	16
Figure 5-1: Comparison between in-situ LST and CM SAF and LSA SAF LST for different TCWV classes at Gobabeb for 2010. The boxplots show the median, the first and third quartile with whiskers at the 95th and 5th percentiles. Red: LSA SAF LST, dark blue: CM SAF LST.....	21
Figure 5-2: Comparison between in situ and CM SAF LST at Gobabeb for 2010. Left: Monthly CM SAF LST. Right: Hourly CM SAF LST.	22
Figure 5-3: Comparison between in-situ LST and CM SAF and LSA SAF LST for different TCWV classes at RMZ for 2010. The boxplots show the median, the first and third quartile with whiskers at the 95th and 5th percentiles. Red: LSA SAF LST, dark blue: CM SAF LST.	23
Figure 5-4: Comparison between in situ and CM SAF LST at RMZ for 2010. Left: Monthly CM SAF LST. Right: Hourly CM SAF LST.	23
Figure 5-5: Comparison between in-situ LST and CM SAF and LSA SAF LST for different TCWV classes at Dahra for 2010. The boxplots show the median, the first and third quartile with whiskers at the 95th and 5th percentiles. Red: LSA SAF LST, dark blue: CM SAF LST.	24
Figure 5-6: Comparison between in situ and CM SAF LST at Dahra for 2010. Left: Monthly CM SAF LST. Right: Hourly CM SAF LST.	25
Figure 5-7: Comparison between in-situ LST and CM SAF and LSA SAF LST for different TCWV classes at Evora for 2010. The boxplots show the median, the first and third quartile with whiskers at the 95th and 5th percentiles. Red: LSA SAF LST, dark blue: CM SAF LST.	26
Figure 5-8: Comparison between in situ and CM SAF LST at Evora for 2010. Left: Monthly CM SAF LST. Right: Hourly CM SAF LST.	26
Figure 5-9: Monthly daily mean time series of the monthly CM SAF LST (red) as compared to EUSTACE T2m air temperature measurements (blue) at 466 stations over Europe. The air temperature measurements, which originate from the European Climate Assessment and Dataset (ECA&D), were homogenized in the EUSTACE project. The black dashed line represents the Theil-Sen linear trend. a) Monthly means, b) Temperature anomalies (seasonal corrected) and c) CM SAF minus T2m temperature anomalies. The green dashed lines show major satellite changes with the transition MVIRI/SEVIRI in 2005.	29

Figure 5-10: Monthly daily mean anomaly time series of the CM SAF LST (red) as compared ESA CCI MODIS Aqua (blue) in Europe (upper), the Tropics (second), Northern Africa (third) and Southern Africa (fourth). The analysis is performed at night-time (01:30 am local time). The dashed green lines show major satellite transitions with the change MVIRI/SEVIRI in 2005.....32

Figure 5-11: Bias of the monthly daily mean anomaly time series of the CM SAF minus ESA CCI MODIS Aqua LST in Europe, the Tropics, Northern Africa and Southern Africa. The dashed green lines show major satellite transitions with the change MVIRI/SEVIRI in 2005.....33

Figure 5-12: Satellite-based comparison for January, March, June and September. Left: CM SAF minus ESA CCI MODIS LST (1:30 am local time, 2003-2018), middle: CM SAF minus ESA CCI LEO & GEO LST (0 am, 2009-2020) and right: CM SAF LST minus LSA SAF LST (0 am, 2004-2019).....36

List of Tables

Table 0-1: Summary of CM SAF LST accuracy (mean bias error) and precision (bias corrected root mean square error), as evaluated at the four Karlsruhe Institute of Technology's (KIT) validation stations, and compared to target requirements (threshold, target, optimal).....8

Table 3-1: Requirements on the CM LST product for hourly data (from AD-1).....14

Table 3-2: Requirements on the CM SAF LST product for monthly data (from AD-1).....15

Table 4-1: Overview of KITs validation stations.....16

Table 5-1: Statistics for the comparison of CM SAF LST and LSA SAF LST with in situ LST at Gobabeb Station in 2010. Compliance with the requirements is presented in colours (threshold, target, optimal).....20

Table 5-2: Statistics for the comparison of CM SAF LST and LSA SAF LST with in situ LST at RMZ Station in 2010. Compliance with the requirements is presented in colours (threshold, target, optimal).....22

Table 5-3: Statistics for the comparison of CM SAF LST and LSA SAF LST with in situ LST at Dahra Station in 2010. Compliance with the requirements is presented in colours (threshold, target, optimal, not fulfilled).....24

Table 5-4: Statistics for the comparison of CM SAF LST and LSA SAF LST with in situ LST at Evora Station in 2010. Compliance with the requirements is presented in colours (threshold, target, optimal, not fulfilled).....25

Table 5-5: Overall CM SAF LST precision and accuracy for KIT validation stations compared to target requirements. Accuracies are given in terms of bias and of bias-corrected RMS. Compliance with the requirements is presented in colours (threshold, target, optimal).....27

Table 5-6: Slope (trend), bias and bc-RMS for the CM SAF LST minus T2m monthly anomalies averaged spatially for different regions. The trend ($t_{CM\ SAF-T2m\ anom}$) is shown for the period 1983-1998 (left), 1999-2019 (middle) and 1983-2019 (right). In

addition, we show trends calculated for T2m (t_{T2m}) and CM SAF LST (t_{CMSAF}) anomalies separately for each period. Statically significant trends ($p < 0.05$) are marked with a *. The last five rows relate to the average over all stations. Compliance with the requirements is presented in colours (threshold, target, optimal).....30

Table 5-7: Slope (trend), bias and bc-RMS for the CM SAF LST minus ESA CCI MODIS monthly anomalies. The trend is calculated for the period 2004 to 2018. Statistically significant trends ($p < 0.05$) are marked with a *. Compliance with the requirements is presented in colours (threshold, target, optimal).....33

Table 5-8: Overall CM SAF LST stability compared to EUSTACE T2m and ESA CCI MODIS Aqua. The stability is given in terms of decadal trend in bias of the CM SAF minus reference data anomaly. Compliance with the requirements is presented in colours (threshold, target, optimal).....34

Table 5-9: Statistics CM SAF – ESA CCI LEO & GEO LST (fulldisk, 2009 to 2020).....35

Table 5-10: Statistics CM SAF – ESA MODIS AQUA LST (fulldisk, 2003 to 2018).35

Table 5-11: Statistics CM SAF – LSA SAF LST (fulldisk, 2004 to 2019).....35

Table 6-1: Summary of CM SAF LST precision (bias-corrected RMS error) compared to target requirements (threshold, target, optimal) evaluated at KITs validation stations.38

Table 6-2: Summary of CM SAF LST accuracy (mean bias error) compared to target requirements (threshold, target, optimal) evaluated at KITs validation stations.38

Table 6-3: Summary of CM SAF LST stability (decadal trend in bias) compared to target requirements (threshold, target, optimal) when evaluated over the various sites and regions.....38

Executive Summary

This CM SAF Product Validation Report provides information on the quality of the CM SAF Meteosat Land Surface Temperature (LST) Climate Data Record (CDR) derived from the Meteosat Visible and InfraRed Imager (MVIS) on board the Meteosat First Generation (MFG) and the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) observations on board the Meteosat Second Generation (MSG) satellites. The covered time of the presented CDR ranges from January 1983 to December of 2020.

The reference datasets used to evaluate the CM SAF LST data precision and accuracy were taken from four ground-based observation. The validation sites are located in different climate zones and include a wide range of atmospheric conditions for different land surfaces. The evaluation scores and their compliance with the target requirements of accuracy and precision are:

Table 0-1: Summary of CM SAF LST accuracy (mean bias error) and precision (bias corrected root mean square error), as evaluated at the four Karlsruhe Institute of Technology's (KIT) validation stations, and compared to target requirements (threshold, target, optimal)

	hourly	monthly
Accuracy	0.6 K	0.4 K
Precision	1.9 K	1.0 K

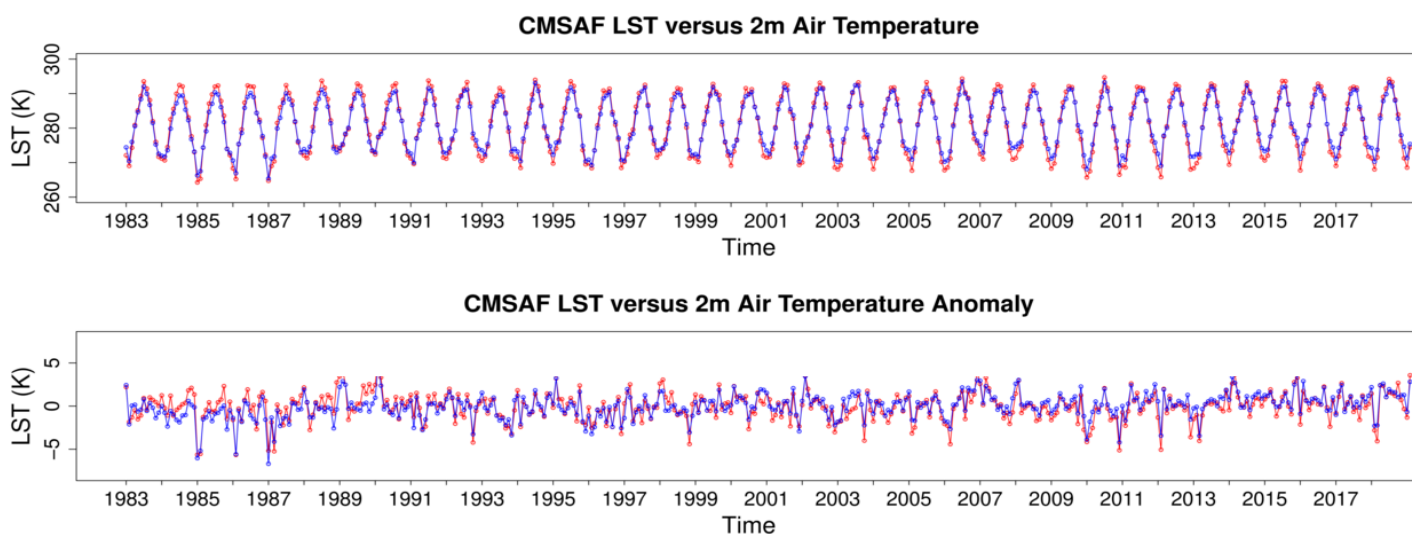




Figure 0-1: Monthly mean time series of the CM SAF LST (red) as compared to homogenized T2m air temperature measurements (blue) at 466 stations over Europe. Top: monthly mean air temperature, bottom: monthly mean air temperature anomaly (seasonal corrected).

The CM SAF LSTs have an excellent agreement with homogenized station based air temperature measurements in Europe (Figure 0-1). Good et al. (2022) have outlined that satellite-based LST data are highly correlated with air temperatures in anomaly space and can be used to assess warming trends over land if the required homogeneity is assured. We observe a decadal trend in bias between the CM SAF minus EUSTACE T2m air temperature anomalies of -0.1K/decade for the period 1999 to 2019, which reflects the optimal stability requirement. For Europe (1999 to 2019) significant trends in CM SAF LST data of 0.37K/decade are obtained, which match the station-based T2m trends of 0.34K/decade . For the period 1983 to 1998 we observe some instability in the order of -0.5K/decade . A comprehensive evaluation against ESAs Land Surface Temperature Climate Change Initiative (CCI) Moderate Resolution Imaging Spectroradiometer (MODIS) LSTs (2003-2018) shows that

	<p align="center">Validation Report Land Surface Temperature (LST) Edition 2</p>	<p>Doc.: SAF/CM/MeteoSwiss/VAL/MET/LST/2.0 Issue: 2.1 Date: 07.09.2023</p>
---	---	--

instabilities are in the order of -0.1 K (optimal requirement) to 0.8 K (threshold requirement) for regions outside of Europe.

This study suggests that CM SAF LST data can be used to monitor temperature anomalies for the new WMO 1991-2020 norm period. The study also suggests that CM SAF LST data can be used to assess warming trends from 1999 onward. Trend analysis before 1999 or trend analysis in desert regions cannot be recommended as the required homogeneity is not assured.

	<p style="text-align: center;">Validation Report Land Surface Temperature (LST) Edition 2</p>	<p>Doc.: SAF/CM/MeteoSwiss/VAL/MET/LST/2.0 Issue: 2.1 Date: 07.09.2023</p>
---	--	--

1 The EUMETSAT SAF on Climate Monitoring

The importance of climate monitoring with satellites was recognized in 2000 by EUMETSAT Member States when they amended the EUMETSAT Convention to affirm that the EUMETSAT mandate is also to “contribute to the operational monitoring of the climate and the detection of global climatic changes”. Following this, EUMETSAT established within its Satellite Application Facility (SAF) network a dedicated centre, the SAF on Climate Monitoring (CM SAF, <http://www.cmsaf.eu>).


The consortium of the CM SAF currently comprises the Deutscher Wetterdienst (DWD) as host institute, and the partners from the Royal Meteorological Institute of Belgium (RMIB), the Finnish Meteorological Institute (FMI), the Royal Meteorological Institute of the Netherlands (KNMI), the Swedish Meteorological and Hydrological Institute (SMHI), the Meteorological Service of Switzerland (MeteoSwiss), and the Meteorological Office of the United Kingdom (UK Met Office). Since the beginning in 1999, the EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF) has developed and will continue to develop capabilities for a sustained generation and provision of Climate Data Records (CDR's) derived from operational meteorological satellites.

In particular, the generation of long-term data records is pursued. The ultimate aim is to make the resulting data records suitable for the analysis of climate variability and potentially the detection of climate trends. The CM SAF works in close collaboration with the EUMETSAT Central Facility and liaises with other satellite operators to advance the availability, quality and usability of Fundamental Climate Data Records (FCDRs) as defined by the Global Climate Observing System (GCOS). As a major task, the CM-SAF utilizes FCDRs to produce records of Essential Climate Variables (ECVs) as defined by GCOS. Thematically, the focus of the CM SAF is on ECVs associated with the global energy and water cycle.

Another essential task of the CM SAF is to produce data records that can serve applications related to the new Global Framework of Climate Services initiated by the World Meteorological Organisation (WMO) World Climate Conference-3 in 2009. The CM SAF is supporting climate services at national meteorological and hydrological services (NMHSs) with long-term data records, but also with data records produced close to real time that can be used to prepare monthly/annual updates of the state of the climate. Both types of products together allow for a consistent description of mean values, anomalies, variability and potential trends for the chosen ECVs. The CM SAF ECV data records also serve the improvement of climate models both at global and regional scales.

As an essential partner in the related international frameworks, in particular WMO SCOPE-CM (Sustained COordinated Processing of Environmental satellite data for Climate Monitoring), the CM SAF - together with the EUMETSAT Central Facility, assumes the role as main implementer of EUMETSAT's commitments in support to global climate monitoring. This is achieved through:

- Application of the highest standards and guidelines as outlined by GCOS for satellite data processing,
- Processing of satellite data within a true international collaboration benefiting from developments at international level and pollinating the partnership with its own ideas and standards,
- Intensive validation and improvement of the CM SAF climate data records,

	<p style="text-align: center;">Validation Report Land Surface Temperature (LST) Edition 2</p>	<p>Doc.: SAF/CM/MeteoSwiss/VAL/MET/LST/2.0 Issue: 2.1 Date: 07.09.2023</p>
---	--	--

- Taking a major role in data record assessments performed by research organisations such as the World Climate Research Programme (WCRP). This role provides the CM SAF with strong contacts to research organizations that form a substantial user group for the CM SAF CDRs,
- Maintaining and providing an operational and sustained infrastructure that can serve the community within the transition of mature CDR products from the research community into operational environments.

A catalogue of all available CM SAF products is accessible via the CM SAF webpage, www.cmsaf.eu/. Here, detailed information about product ordering, add-on tools, sample programs and documentation is provided.

2 CM SAF Meteosat Land Surface Temperature CDR

The CM SAF LST CDR is based on 38 years of Meteosat measurements. MVIRI and SEVIRI are optical imaging radiometers mounted on the geostationary Meteosat First Generation (MFG) satellites 1 to 7 and Meteosat Second Generation (MSG) satellites 1 to 4, respectively. Meteosat satellites in operational mode are centred near 0°/0° latitude/longitude and acquire an image of a full earth disk including Europe, Africa, the Middle East and the Atlantic Ocean. MVIRI scans the full earth disk every 30 minutes with 5 x 5 km² spatial resolution at nadir in the thermal channel. SEVIRI images the full disk every 15 minutes with a horizontal resolution of 3 x 3 km² at nadir in the thermal channel. Both sensors are passive imagers with spectral bands in the visible and thermal infrared. MVIRI has three bands: a broad visible channel, a water vapour channel and a single infrared channel. SEVIRI has 12 spectral channels between 0.6 µm and 13.4 µm, which include two thermal infrared ‘split-window’ channels. In order to ensure the highest possible consistency for the LST CDR, the retrieval algorithms only use channels that are available or can be simulated from both sensors, i.e. the broadband visible channel and the 10.8 µm infrared channel.

The presented LST CDR spans the years 1983-2020, and is based on measurements of MFG-2, MFG-3, MFG-4, MFG-5, MFG-6, MFG-7, MSG-1, MSG-2, MSG-3 and MSG-4 (Figure 2-1). Gaps in the prime satellite were filled by a back-up satellite. Both CDRs were derived from Level 1.5 MVIRI and SEVIRI data provided by EUMETSAT.

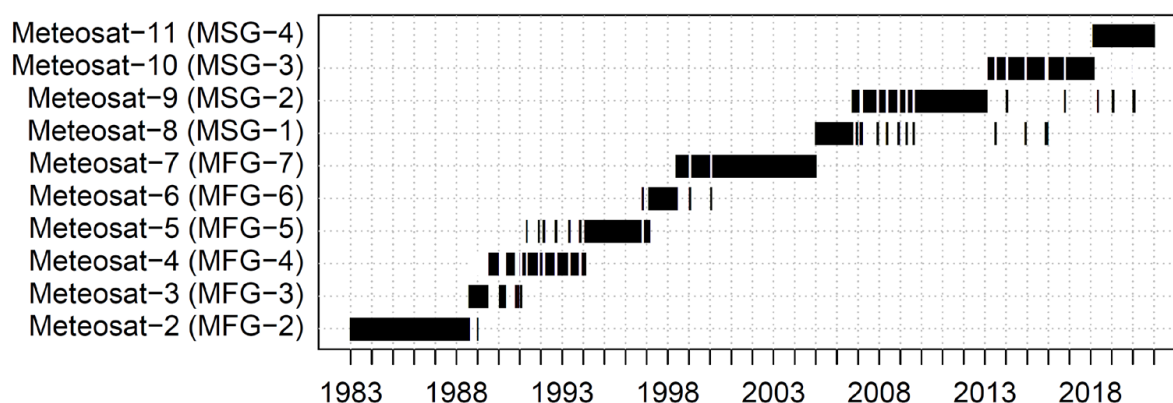



Figure 2-1: Availability of MFG and MSG measurements.

For MVIRI and SEVIRI dedicated infrared channel inter-calibration factors were provided by EUMETSAT, which are based on daily inter-calibrations of MVIRI and SEVIRI with the High Resolution Infrared Radiation Sounder (HIRS) and Infrared Atmospheric Sounding Interferometer (IASI) instrument on board the EUMETSAT Meteorological Operational satellite (MetOp) series of polar orbiting platforms. For the processing we used the calibration coefficients from release 1.0 of the Meteosat FCDR. The inter-calibration is carried out using the spectral response function of the respective MVIRI and SEVIRI sensor. For a more detailed instrument specification and a description of the calibration, the reader is referred to John et al. (2019).

The CM SAF LSTs are retrieved with the GeoSatClim LST algorithms v2022. Since the MVIRI instrument on-board Meteosat 2–7 is equipped with a single thermal infrared channel, single-channel LST retrieval algorithms are used to ensure consistency across Meteosat satellites.

The CM SAF LST data is retrieved with the Physical Mono-Window (PMW) LST model (details provided in the CM SAF LST ATBD; RD 2). Single-channel LST models strongly depend on

	Validation Report Land Surface Temperature (LST) Edition 2	Doc.: SAF/CM/MeteoSwiss/VAL/MET/LST/2.0 Issue: 2.1 Date: 07.09.2023
---	---	---

ancillary data to estimate the atmospheric state and surface properties. The CM SAF single channel model is based on radiative transfer modelling and requires ERA5 atmospheric profiles and an external monthly emissivity climatology to calculate LST.

The CM SAF LST data are provided as hourly and monthly diurnal cycle data for the period 1983-2020. The provided CM SAF LST data are hourly samples, i.e. they are instantaneous values at the full hour. The CM SAF LST product specifications are given in the corresponding Product User Manual (RD 1).

3 Validation Strategy

The purpose of the validation effort is to characterize the CM SAF LST dataset in terms of accuracy, precision and stability. Furthermore, the products are evaluated w.r.t. the requirements stated in AD 1, which are summarised by Table 3-1 and Table 3-2, and their compliance is reported.

The following statistical parameters to characterize the quality of the CM SAF LST data sets in terms of precision, accuracy and stability were selected:

- bias or mean error or MBE (**accuracy**)
- bias corrected root mean square (bc-RMS) difference (**precision**)
- decadal trend in bias as compared to a reference data set (**stability**)

with

$$\text{MBE} = \frac{1}{n} \sum_{k=1}^n (E_k - M_k)$$

$$\text{bc-RMS} = \sqrt{\frac{1}{n} \sum_{k=1}^n (E_k - M_k - \text{MBE})^2}$$

with n=number of observations, E_k =single reference observation, M_k =satellite LST. The rationale for the chosen statistical parameters is that the overall SAF Product Requirements Table should include measures for both accuracy (i.e., how close to the reference is our estimation?), precision (i.e., what is the random spread of our estimation?) and temporal stability (i.e., is the bias constant over time or is there a trend?).

The CM SAF has defined the following product requirements:

- Threshold requirement (minimum standard possible for the product release)
- Target requirement (target for the product release)
- Optimal requirement (optimal requirement from the user community) which could be retrieved with an optimal observing system.

The product requirements were defined after taking into account requirements from different users and user groups. The most established recommendations in this respect were issued by the Global Climate Observation System (GCOS) community. However, the values in Table 3-1 and Table 3-2 were also influenced by requirements from users working on regional climate monitoring and modelling applications, which often have even stricter requirements than GCOS.

Table 3-1: Requirements on the CM LST product for hourly data (from AD-1).

	Product Requirements hourly LST (K)		
	Threshold	Target	Optimal
Accuracy (bias)	1.5	1	0.5
Precision (bc-RMS)	2.5	1.5	1
Stability (decadal)	1	0.3	0.1

Table 3-2: Requirements on the CM SAF LST product for monthly data (from AD-1).

	Product Requirements		
	monthly LST (K)		
	Threshold	Target	Optimal
Accuracy (bias)	1.5	1	0.5
Precision (bc-RMS)	1.5	1	0.5
Stability (decadal)	1	0.3	0.1

To ensure that the CM SAF LST products meet their target accuracy and precision, CM SAF LSTs are validated against in-situ LST obtained at Karlsruhe Institute of Technology's (KIT) validation stations. The stations are located in homogeneous areas in Africa and Europe and cover a wide range of atmospheric conditions on the Meteosat disk. LST is a variable, which varies strongly in space and time. For the Meteosat series of satellites, only KIT's stations are located in sufficiently large and highly homogeneous areas to provide meaningful accuracy and precision estimates for satellite-based LST retrievals (Göttsche et al. 2016).

Another focus of this evaluation is on the temporal stability of the presented CM SAF LST data. Most climate applications can deal with a substantial systematic bias, since observational data are usually bias-corrected during comparisons with models or work is done in anomaly space. Furthermore, a fairly substantial random error is not necessarily a problem for detecting climate signals. However, sufficiently long-term in-situ LST measurements over homogeneous areas are not available from the LSA SAF sites or any other location within the Meteosat field of view. Therefore, the decadal product stability is evaluated against homogenised EUSTACE T2m ground air temperature measurements from 1983 to 2019 and ESA CCIs Moderate Resolution Imaging Spectroradiometer (MODIS) Aqua LSTs from 2003 to 2018. Good et al. (2022) has demonstrated that those data are stable and can be used as reference to test the homogeneity of LST CDRs.

Validation statistics are provided for hourly and monthly data, as the CM SAF Meteosat LST products are provided as hourly samples (i.e. instantaneous data at the full hour and monthly data).

4 Data sets for comparison with Meteosat LST

4.1 In-situ Observations

4.1.1 KIT station-based Land Surface Temperature measurements

LST observations from Karlsruhe Institute of Technology’s (KIT) four permanent LST validation stations are used to evaluate the accuracy and precision of the CM SAF LST products. The stations were set up specifically for validating satellite-retrieved LST products and are part of LSA SAF’s validation effort supported by EUMETSAT. They are located in large homogenous areas within Meteosat’s field of view and lie in different climate zones, which provides a broad range of atmospheric conditions for product validation (Göttsche et al. 2016). The locations of the four validation stations on the Meteosat Earth disk are indicated in Figure 4-1. An overview of the KITs validation sites is provided in Table 4-1.

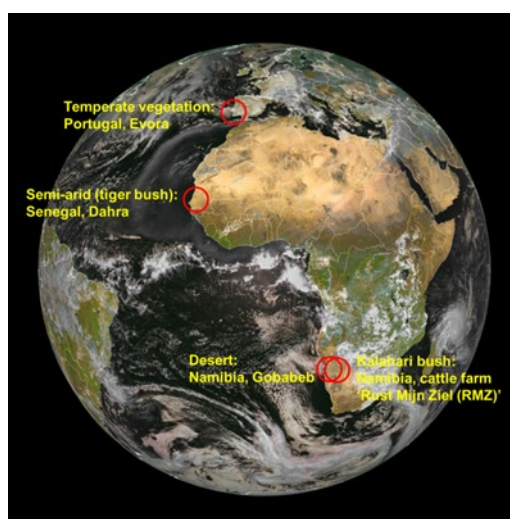



Figure 4-1: Locations of the Karlsruhe Institute of Technology’s (KIT) validation stations on the SEVIRI disk.

Table 4-1: Overview of KITs validation stations.

	Dahra	RMZ	Gobabeb	Evora
Location	Senegal Lat: 15.402336 Lon: -15.432744	Namibia Lat: -23.010532 Lon: 18.352897	Namibia Lat: -23.550956 Lon: 15.05138	Portugal Lat: 38.540244 Lon: -8.003368
Elevation	90 m	1450 m	406 m	230 m
Climate Zone	Tropical Wet-Dry	Steppe	Desert	Mediterranean
Vegetation	Grassland; 96% grass, 4% tree	Savanna; 85% grass/soil, 15% tree	Barren; 32% tree, 68% grass	Woody savanna with isolated groups of evergreen oak trees

The main instrument for the in situ determination of LST at KIT’s validation stations is the precision radiometer “KT15.85 IIP” produced by Heitronics GmbH, Wiesbaden, Germany. KT15.85 IIP radiometers measure thermal infra-red radiance between 9.6 μm and 11.5 μm ,

	Validation Report Land Surface Temperature (LST) Edition 2	Doc.: SAF/CM/MeteoSwiss/VAL/MET/LST/2.0 Issue: 2.1 Date: 07.09.2023
---	---	---

have a temperature resolution of 0.03 K and an accuracy of ± 0.3 K over the relevant temperature range (Theocharous et al. 2008). The KT15.85 IIP has a drift of less than 0.01% per month: the high stability is achieved by linking the radiance measurements via beam-chopping (a differential method) to internal reference temperature measurements and was confirmed by a long-term parallel run with the self-calibrating radiometer “RotRad” from Commonwealth Scientific and Industrial Research Organisation (CSIRO), which is continuously stabilized with 2 blackbodies (Kabsch et al. 2008). The parallel run at the Evora site started in April 2005; a year later, the agreement between the instruments was still excellent (correlation 0.99). Due to the KT-15.85 IIP’s narrow spectral response function and the small distance between the radiometers and the surface atmospheric attenuation of the surface-leaving, thermal infrared radiation is negligible. However, the measurements of the surface-observing KT-15.85 IIPs contain radiance emitted by the surface (i.e., the target signal), as well as reflected downward IR radiance from the atmosphere, which needs to be corrected for. Therefore, at each station, an additional KT-15.85 IIP measures downward longwave IR radiance from the atmosphere at 53° VZA: measurements under that specific zenith angle are directly related to downward hemispherical radiance (Kondratyev 1969), so that no ancillary data for deriving ground truth LST are needed (Göttsche et al. 2016).


Accurate estimations of land surface emissivity are essential for obtaining satellite LST products, but also for limiting the uncertainty of ground-based LST estimates. Since for vegetated sites, LSE is a dynamic quantity, in-situ LSTs at Dahra, Rust Mijn Ziel (RMZ) and Evora are derived using the LSA SAF dynamic emissivity product. In situ LST at the desert site Gobabeb is derived using a static emissivity obtained from in situ measurements (Göttsche et al. 2016).

Göttsche et al. (2016) and Ermida et al. (2014) have estimated that the uncertainty of in situ LST observations lies between 0.5 and 1K. The error of ground LST is dominated by uncertainty in local emissivity and noise in the measurements. In the case of Evora site, the upscaling process to estimate pixel-size LST (Ermida et al. 2014) constitutes a further source of uncertainty, particularly during summer months when local temperature heterogeneity is high.

4.1.2 EUSTACE 2m air temperature measurements

The CM SAF LST CDR is compared with daily mean near-surface air temperature (T2m) measurement data from the EU Surface Temperature for All Corners of Earth (EUSTACE) project (Rayner et al. 2020) for meteorological stations within Europe (Squintu et al. 2019a). The data are the only daily T2m station dataset available for a larger region, which is homogenized. Daily data are required to obtain the best possible temporal match between CM SAF LSTs and T2m. In addition, we require homogenized data to assess the temporal stability of the CM SAF LSTs.

The temperatures series, which originate from the European Climate Assessment and Dataset (ECA&D, Klok and Klein Tank, 2008), were homogenised using the quantile matching approach, which corrects for changes in station location or measurement equipment (Squintu et al. 2019b). An automated method locates breaks in the series based on a comparison with surrounding series and applies adjustments, which are estimated using homogeneous segments of surrounding series as a reference. In total, around 2100 series have been adjusted (Good et al. 2022). Data are provided across Europe from 1983 up to 2019.

	Validation Report Land Surface Temperature (LST) Edition 2	Doc.: SAF/CM/MeteoSwiss/VAL/MET/LST/2.0 Issue: 2.1 Date: 07.09.2023
---	---	---

4.2 Satellite Data

4.2.1 ESA CCI MODIS AQUA LST CDR

We used Land surface temperature from MODerate resolution Imaging Spectroradiometer (MODIS on Aqua), level 3 collated (L3C) global product version 3.00 for this validation. The dataset coverage is global over the land surface. The CDR is produced by ESAs Land Surface Temperature Climate Change Initiative (CCI). LSTs are provided on a global grid at a resolution of 0.01° longitude and 0.01° latitude.

MODIS is an advanced imaging instrument on board NASA's Terra and Aqua polar satellites. Aqua crosses the equator at 01:30 am and 01:30 pm local solar time. Aqua MODIS views Earth's entire surface every 1 to 2 days, acquiring data in 36 spectral bands. MODIS achieves full Earth coverage nearly twice per day so the daily files have small gaps primarily close to the equator where the surface is not covered by the satellite swath on that day. MODIS is one of the most advanced passive imagers in space and has proven to be very stable over time (Kilpatrick et al. 2015).

The ESA CCI (MODIS) Aqua LST CDR spans the period 2003 to 2018. Among all ESA CCI LST data sets, only the MODIS Aqua and the Advanced Along-Track Scanning Radiometer (AATSR) LST CDR appear stable (Good et al. 2022). The MODIS Aqua CDR has a longer temporal coverage compared to AATSR and is therefore chosen for this validation. It shows no climatic discontinuities associated with sensor drift over time (Good et al. 2022). For the period 2002-2018 significant trends in LST of 0.64-0.66 K/decade are obtained, which compare well with the equivalent station-based T2m trends of 0.52-0.59 K/decade (Good et al. 2022).


MODIS Aqua ESA CCI LST has been retrieved using observations from the split window channels at 11 and 12 µm. Data were processed in the University of Leicester processing chain. The Generalised Split Window (GSW) retrieval coefficients are defined for different land cover classes and diurnal condition (Dodd et al. 2019). Cloud screening is performed using the University of Leicester cloud retrieval system. The surface emissivity is taken from the Combined ASTER and MODIS Emissivity for Land (CAMEL) database, which is a global monthly mean emissivity dataset spanning the years 2000 – 2016. Further details on the retrieval algorithm can be found in the ATBD (Dodd et al. 2019) and the PUM (Dood et al. 2021).

The ESA CCI MODIS Aqua LST CDR is available at <https://climate.esa.int/en/odp/#/project/land-surface-temperature> (last visited on 14/04/23).

4.2.2 ESA CCI LEO and GEO LST CDR

In addition to MODIS, we used ESA CCIs Monthly multisensor Infra-Red (IR) Low Earth Orbit (LEO) and Geostationary Earth Orbit (GEO) land surface temperature (LST) level 3 supercollated (L3S) global product version 1.00 to compare to the CM SAF CDR. LST fields are provided at 3 hourly intervals each day at the full UTC hour (e.g. 00:00 UTC, 03:00 UTC) for the period 2009 to 2020 at 0.05 latitude/longitude grids.

The product is based on merging of available GEO data and infilling with available LEO data outside of the GEO discs (Veal et al. 2022). As data towards the edge of the GEO disc is known to have large uncertainty, data with satellite zenith angle of more than 60 degrees is discarded. For the area covered by the CM SAF LST CDR Spinning Enhanced Visible Infra-

	<p style="text-align: center;">Validation Report Land Surface Temperature (LST) Edition 2</p>	<p>Doc.: SAF/CM/MeteoSwiss/VAL/MET/LST/2.0 Issue: 2.1 Date: 07.09.2023</p>
---	--	--

Red Imager (SEVIRI) on Meteosat Second Generation is used for VZA of less than 60 degrees. For larger satellite zenith angles, LEO data (AATSR or MODIS) is combined depending on mission start and end dates and instrument downtimes. Hence, for viewing angles below 60 degrees the LST CDR is produced using the same Meteosat input data as the CM SAF LST CDR.

LSTs are retrieved using a Generalised Split Window algorithm from all instruments (Veal et al. 2022). As the CM SAF employs a single channel LST retrieval algorithm, a different retrieval algorithm is employed for similar input data. The Geostationary data were produced by the Instituto Português do Mar e da Atmosfera (IPMA) before being merged into the final dataset.

The ESA CCI LEO and GEO LST CDR is available at <https://climate.esa.int/en/odp/#/project/land-surface-temperature> (last visited on 14/04/23).

4.2.3 LSA SAF Meteosat LST CDR

LSA SAF generates and disseminates an operational Meteosat SEVIRI LST product and has recently also issued a LST CDR spanning the years 2004 to 2015. The LSA SAF LST dataset is then continued in near realtime. In contrast to CM SAF LST and similar to ESA CCI LST products, LSA SAF LSTs are obtained with a two-channel Generalised Split-Window (GSW) retrieval scheme. The GSW performs corrections for atmospheric effects based on the differential absorption in adjacent IR bands and requires EM as input data. Land Surface Emissivity is independently estimated as a function of Fraction of Vegetation Cover (FVC) and land cover classification (Trigo et al. 2018). The LSA SAF team has implemented a correction for LST angular effects (Ermida et al. 2018a), the so called kernel-hotspot model. The kernel-hotspot model is calibrated taking into account the surface characteristics (vegetation cover and structure, topography) and provides estimates of LST dependence on viewing illumination geometries (Ermida et al. 2018b).

We have used the LSA SAF monthly LST 2004-2019 climatology. LST is averaged over the 2004-2019 period, per month and hourly time slot. This dataset was derived by joining the LST CDR (MLST-R, LSA-050) for 2004-2015 and the operational NRT product (MLST, LSA-001) for 2016-2019. It is available on a 0.05 latitude and longitude grid at <https://landsaf.ipma.pt/en/products/land-surface-temperature/lst/> for registered users (last visited on 14/04/23).

A detailed description of the LSA SAF algorithm can be found in the corresponding Algorithm Technical Base Document (Trigo et al. 2018).

5 Validation Results

5.1 Precision and Accuracy

In-situ observations for 2010 for KIT's validation stations were used to evaluate the precision and accuracy of the CM SAF Meteosat LST data. Currently, these validation stations are the only ones located in highly homogeneous areas within Meteosat's field of view and suitable for evaluating Meteosat-derived LST (Göttsche et al. 2016). A description of the validation stations and in-situ LST derivation is given in Section 4.1.1.

Precision and accuracy for hourly and monthly CM SAF LST data is provided. Monthly means were calculated from match-up's with in-situ LST obtained from 1-minute averages of brightness temperature measurements. The delay between actual satellite acquisition time and nominal product time was accounted for and in-situ and CM SAF LSTs were collocated to better than a minute. CM SAF LST estimates which had corresponding in-situ observations were collocated and aggregated to monthly means.

In order to also demonstrate the high quality of the CM SAF cloud mask, no additional cloud filter was applied, something that is common for comparisons between satellite LST and in-situ LST (e.g. Göttsche et al. 2016 and Freitas et al. 2010). The performance of the CM SAF LST is evaluated by calculating mean bias error (accuracy) and bc-RMS (precision) for each station and comparing them to the target requirements specified in Table 3-1 and Table 3-2 (details in AD 1).

5.1.1 Gobabeb Station

Aggregated to monthly statistics, the two CM SAF LST products are within the target accuracy and accuracy at Gobabeb (Table 5-1, Figure 5-1 and Figure 5-2).

Table 5-1: Statistics for the comparison of CM SAF LST and LSA SAF LST with in situ LST at Gobabeb Station in 2010. Compliance with the requirements is presented in colours (threshold, target, optimal).

	CM SAF LST			LSA SAF LST		
	n	bias	bc-RMS	n	bias	bc-RMS
hourly	6108	0.81 K	1.69 K	6108	0.39 K	1.39 K
monthly	12	0.75 K	0.76 K	12	0.40 K	0.56 K

Due to the exceptionally wet January/February and October/November 2010, the presented analysis included LST observations for a wide range of atmospheric conditions, including some rather moist atmospheres (Figure 5-1).

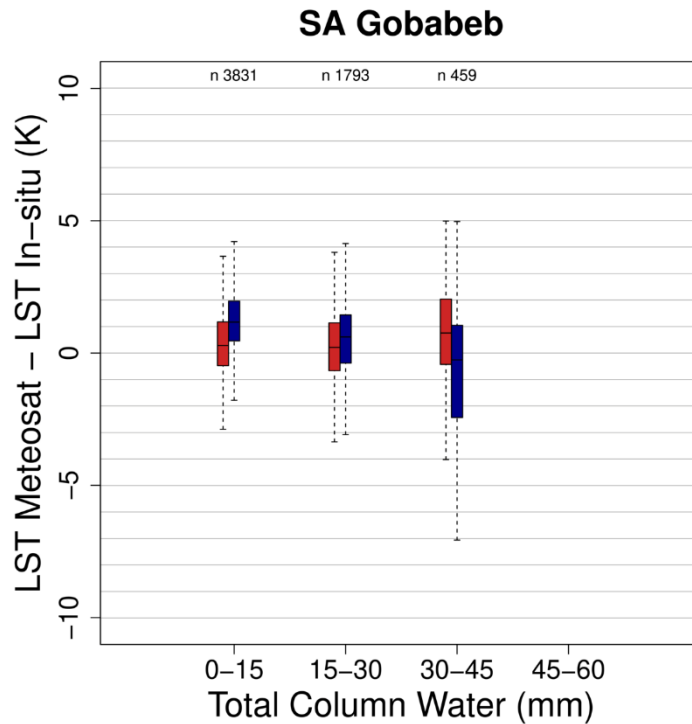


Figure 5-1: Comparison between in-situ LST and CM SAF and LSA SAF LST for different TCWV classes at Gobabeb for 2010. The boxplots show the median, the first and third quartile with whiskers at the 95th and 5th percentiles. Red: LSA SAF LST, dark blue: CM SAF LST.

For dry atmospheres, CM SAF LST has a 1 K positive bias compared to in-situ LST as well as to LSA SAF LST (Figure 5-1). Emissivity differences cannot explain this positive bias, as CM SAF emissivity in SEVIRI channel 10.8 (= University of Wisconsin Baseline Fit emissivity) is near-constant at 0.955 and, therefore, higher than the constant emissivity of 0.940 used for obtaining the in-situ LST and the near-constant 0.949 LSA SAF emissivity (Göttsche and Hulley 2012).

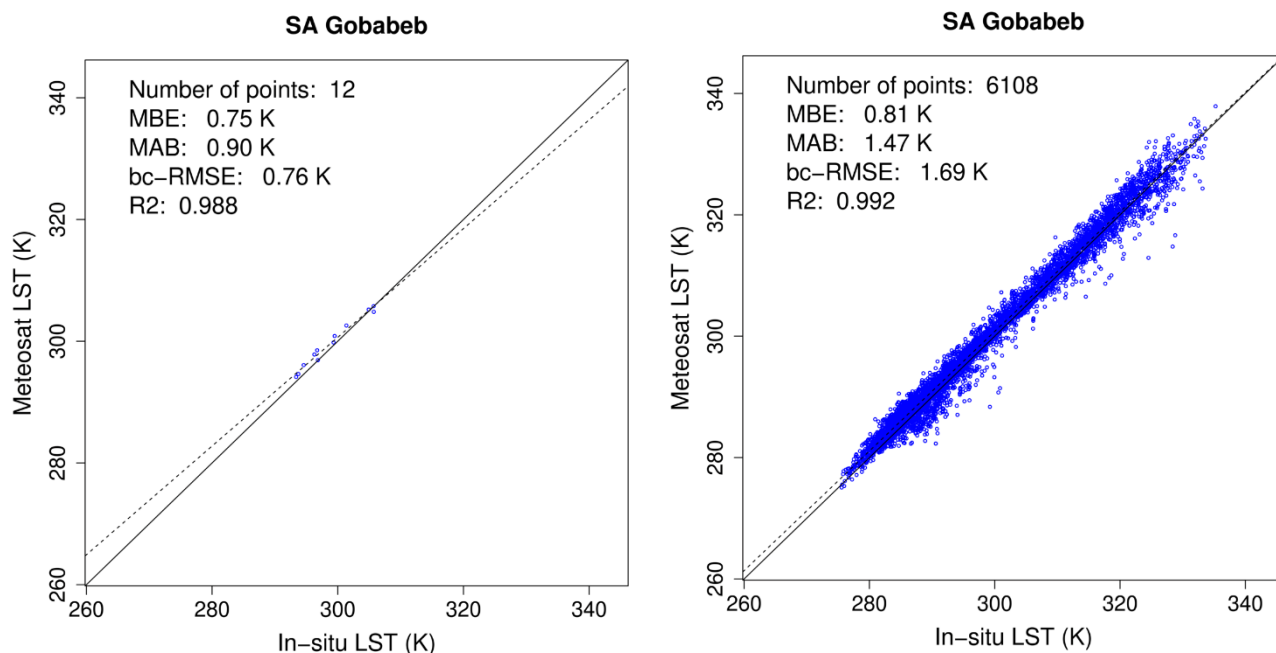


Figure 5-2: Comparison between in situ and CM SAF LST at Gobabeb for 2010. Left: Monthly CM SAF LST. Right: Hourly CM SAF LST.

5.1.2 RMZ Station

For the Savanna station RMZ the CM SAF LST products show an excellent accuracy with a bias within the optimal 0.5 K accuracy requirement (Figure 5-3, Figure 5-4). The precision (bc-RMS) of the CM SAF LST data is within the 1.5 K (hourly) and 1 K (monthly) target precision. In 2010 the RMZ site had dry atmospheres (TCWV < 30 mm) and, thus, it is concluded that the CM SAF LST retrieval algorithm performs well for dry atmospheres (Figure 5-4).

Table 5-2: Statistics for the comparison of CM SAF LST and LSA SAF LST with in situ LST at RMZ Station in 2010. Compliance with the requirements is presented in colours (threshold, target, optimal).

	CM SAF LST			LSA SAF LST		
	n	bias	bc-RMS	n	bias	bc-RMS
hourly	4815	0.08 K	1.56 K	4815	-0.59 K	1.80 K
monthly	12	-0.12 K	0.66 K	12	-0.57 K	0.25 K

For the CM SAF LST, we observe a positive bias for temperatures above 310 K (Figure 5-2). The distinct bias at high temperatures likely demonstrates some problems of the radiative transfer model to account for the considerable strong land surface overheating (up to 340 K) compared to air temperatures (around 290 K) in desert regions (Göttsche et al. 2013; Göttsche et al. 2016).

SA RMZ

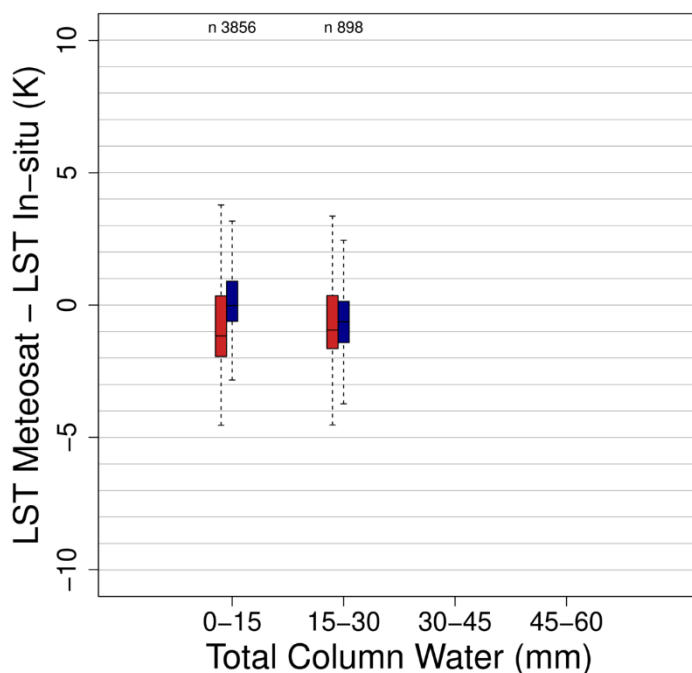


Figure 5-3: Comparison between in-situ LST and CM SAF and LSA SAF LST for different TCWV classes at RMZ for 2010. The boxplots show the median, the first and third quartile with whiskers at the 95th and 5th percentiles. Red: LSA SAF LST, dark blue: CM SAF LST.

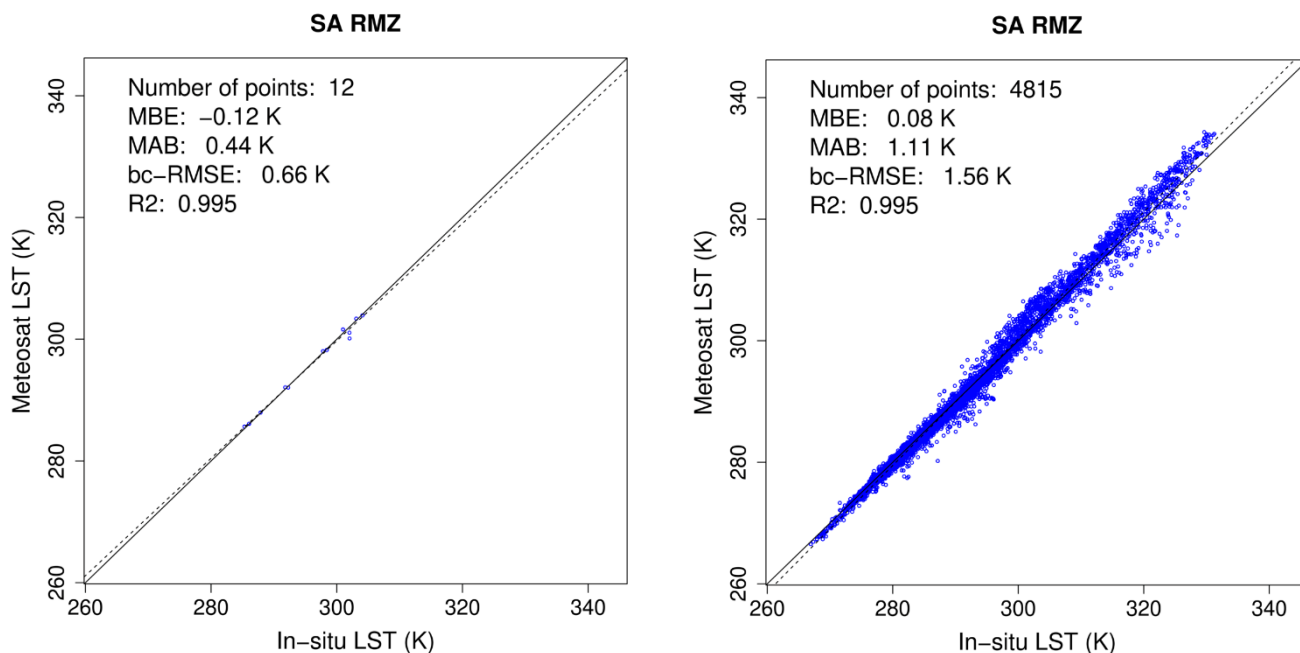


Figure 5-4: Comparison between in situ and CM SAF LST at RMZ for 2010. Left: Monthly CM SAF LST. Right: Hourly CM SAF LST.

5.1.3 Dahra Station

For the tropical site Dahra, the observed bias meets the required target accuracy. TCWVs can be up to 60 mm during the rainy season in Dahra (Göttsche et al. 2016). Instantaneous CM SAF LST products reach the target precision for TCWV up to 30 mm (Figure 5-5). For higher TCWVs, a distinct loss in precision (Figure 5-5) is observed, which is expected from single-channel models due to their sensitivity to NWP errors as detailed in Duguay-Tetzlaff et al. (2015).

Bc-RMS for CM SAF LST products are up to 3.0 K for very moist atmospheres (TCWV > 45 mm) and exceed the 2.0 K threshold precision (Table 5-3, Figure 5-6). The strong negative bias of the LSA SAF seen in Dahra under 30-45 mm WV range relates to errors in the atmospheric correction under very high aerosol loads, which in this location coincide with high moisture content (Stante et al., 2023).

Table 5-3: Statistics for the comparison of CM SAF LST and LSA SAF LST with in situ LST at Dahra Station in 2010. Compliance with the requirements is presented in colours (threshold, target, optimal, not fulfilled)

	CM SAF LST			LSA SAF LST		
	n	bias	bc-RMS	n	bias	bc-RMS
hourly	1602	-0.30 K	2.88 K	1602	-0.99 K	2.06 K
monthly	7	-0.7 K	1.72 K	7	-1.37 K	1.44 K

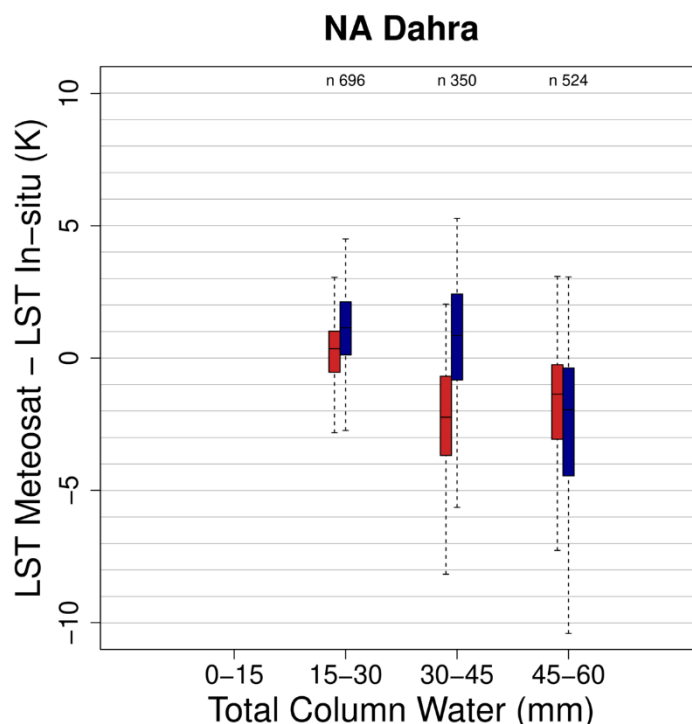


Figure 5-5: Comparison between in-situ LST and CM SAF and LSA SAF LST for different TCWV classes at Dahra for 2010. The boxplots show the median, the first and third quartile with whiskers at the 95th and 5th percentiles. Red: LSA SAF LST, dark blue: CM SAF LST.

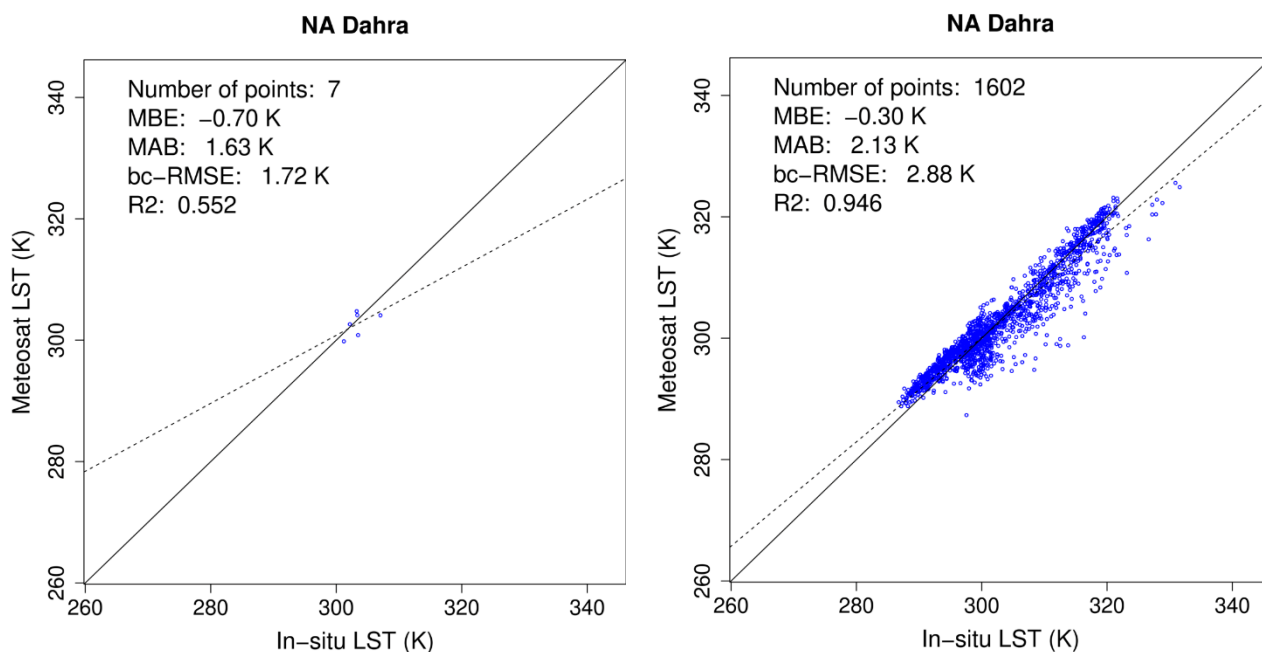


Figure 5-6: Comparison between in situ and CM SAF LST at Dahra for 2010. Left: Monthly CM SAF LST. Right: Hourly CM SAF LST.

5.1.4 Evora Station

At station Evora (Mediterranean climate zone) the precision of the CM SAF LST is within the target (monthly) and threshold (hourly) precision. The CM SAF LST show a distinct positive bias of about 1.3 K to 1.7 K at the different aggregation steps (Table 5-4, Figure 5-7). The bias is more or less constant through the entire temperature range (Figure 5-8).

Table 5-4: Statistics for the comparison of CM SAF LST and LSA SAF LST with in situ LST at Evora Station in 2010. Compliance with the requirements is presented in colours (threshold, target, optimal, not fulfilled).

	CM SAF LST			LSA SAF LST		
	n	bias	bc-RMS	n	bias	bc-RMS
hourly	3407	1.74 K	1.84 K	3407	1.24 K	1.54 K
monthly	12	1.38 K	0.78 K	12	1.03 K	0.59 K

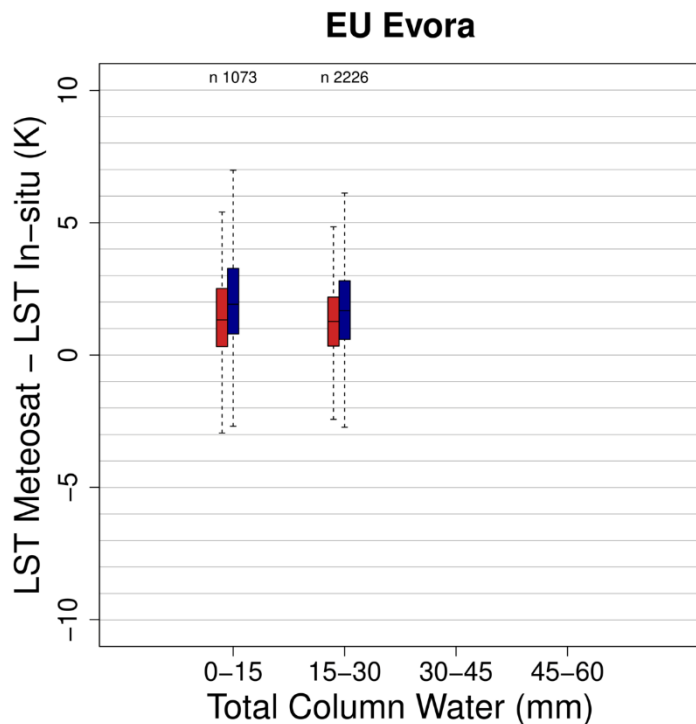


Figure 5-7: Comparison between in-situ LST and CM SAF and LSA SAF LST for different TCWV classes at Evora for 2010. The boxplots show the median, the first and third quartile with whiskers at the 95th and 5th percentiles. Red: LSA SAF LST, dark blue: CM SAF LST.

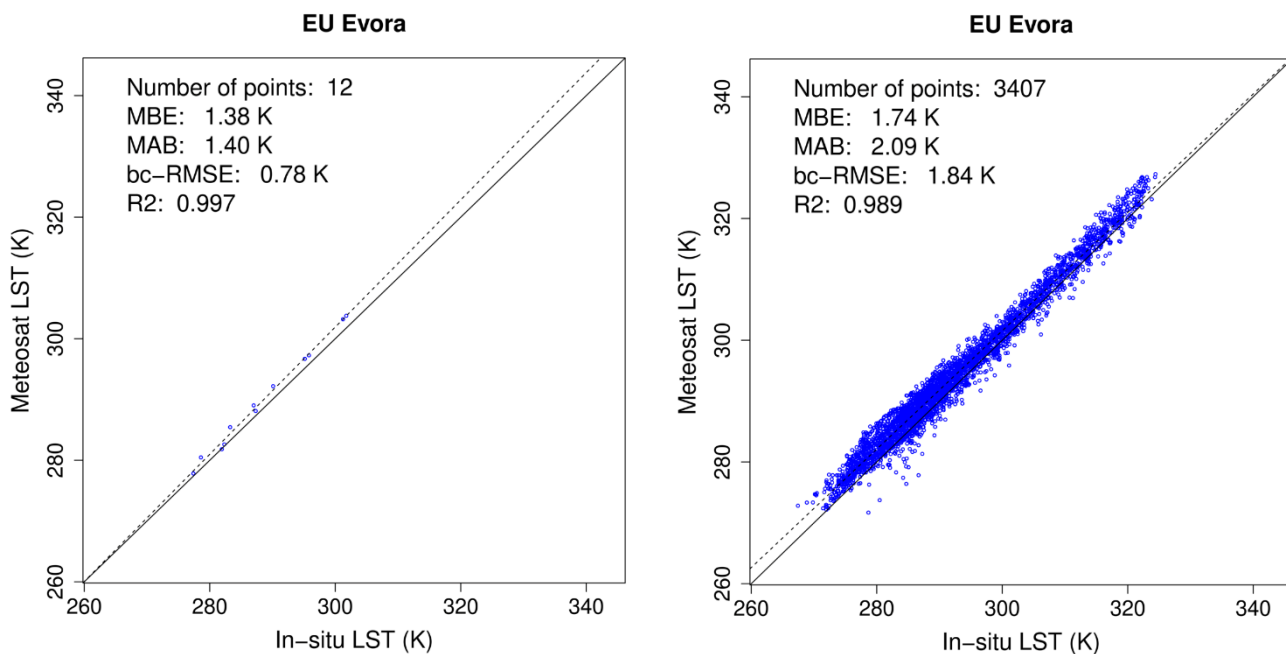


Figure 5-8: Comparison between in situ and CM SAF LST at Evora for 2010. Left: Monthly CM SAF LST. Right: Hourly CM SAF LST.


Figure 5-7 outlines that this positive bias is also present for the LSA SAF LST product at Evora. The positive LSA SAF LST bias is known from previous validation studies (Göttsche et al. 2011). This bias partially reflects the achievable accuracy with in-situ LST measurements (Göttsche et al. 2016). These have to represent large scale satellite footprints covering several square kilometres. Although the land cover at Evora is spatially quite homogeneous, it represents a mixture of grass, background soil, and trees, which cause shadows and complicate the ground-based LST determination (Ermida et al. 2014; Guillevic et al. 2013). The bias for the CM SAF LST is about 0.5 K higher compared to LSA SAF LST bias in Evora (Figure 5-7), which partially reflects emissivity differences: CM SAF emissivity is on average 0.003 lower than the respective LSA SAF emissivity.

5.1.5 Summary Precision and Accuracy

The overall precision and accuracy of the CM SAF LST data, as evaluated at the four KIT validation stations, is well within the target requirements as outlined in Table 5-5. For the monthly bias we reach the optimal requirement. The hourly bc-RMS is within the threshold when including also very moist atmospheres.

Table 5-5: Overall CM SAF LST precision and accuracy for KIT validation stations compared to target requirements. Accuracies are given in terms of bias and of bias-corrected RMS. Compliance with the requirements is presented in colours (threshold, target, optimal).

	Product Requirements CM SAF LST			Achieved Accuracy & Precision CM SAF LST	
	Threshold	Target	Optimal	Hourly	Monthly*
Bias	1.5 K	1.0 K	0.5 K	0.58	0.36
Bc-RMS	2.5 (1.5* K)	1.5 (1.0* K)	1.0 (0.5* K)	1.9	0.97

	Validation Report Land Surface Temperature (LST) Edition 2	Doc.: SAF/CM/MeteoSwiss/VAL/MET/LST/2.0 Issue: 2.1 Date: 07.09.2023
---	---	---

5.2 Temporal Stability

Decadal stability requirements for the CM SAF LST are 1 K/decade, 0.3 K/decade and 0.1 K/decade for threshold, target and optimal, respectively.

Decadal stability reveals the change of LST accuracy (bias) over 10 years. To assess the decadal stability we used the following data sets:

- Homogenized EUSTACE T2m for the period 1983-2019
- ESA CCI MODIS LST for the period 2003-2018

All the analysis presented in this study is performed on anomalies to remove the seasonal cycle in the data, which dominates the temperature signal. The seasonal correction is performed by adjusting the monthly means with the corresponding mean monthly value over the entire reference period.

The stability of the LST datasets is evaluated by comparing monthly mean anomalies for the CM SAF LST and the reference data. Monthly rather than daily data are used for this analysis to ensure a clear signal over time that is not dwarfed by the 'noise' that is present in a daily time series (Good et al. 2022).

The trend of this time series of differences is calculated as above using Theil- Sen estimates (Theil 1950) and the significance of a trend is estimated with the Mann- Kendall test (Kendall 1938, Mann 1945, 5% significance level). Significant trend in this LST-T2m difference is assumed to indicate that the dataset is inhomogeneous.

5.2.1 Stability against EUSTACE 2m air temperature

Good et al. (2022) has clearly outlined that the homogenized EUSTACE T2m measurements can be used to assess the stability of Land Surface Temperature CDRs:

- 1) There is a strong correlation between satellite-based LST and T2m in anomaly space. Good et al. (2022) found excellent correlations between daily LST and T2m anomalies ~0.8 to ~0.9 for various ESA CCI LST data. The results are very consistent across different datasets and are very similar to those reported by Good et al. (2017).
- 2) There is no evidence of a clear-sky bias for trends calculated from LST data. Satellite sensors measure LST only under clear sky conditions compared to "all weather" T2m station measurements (Good et al. 2022). Trends calculated in anomaly space from "all sky" T2m observations are almost identical to satellite-based LST trends for the ESA CCI satellite-data sets that are stable in time (Good et al. 2022).
- 3) The difference between the T2m and LST contains information on the land-atmosphere coupling, which in turn may indeed change over time. With e.g. less regions to be fully snow covered the atmospheric coupling between T2m and LST changes due to different surface properties. Jiang et al. (2022) claim to see a trend in T2m/LST difference of 0.009 K/decade from 1981 to 2020. This change is very small and therefore does not affect this stability analysis.

We strictly followed the methodology proposed by Good et al. (2022) to assess the stability of ESA CCI LST climate data records. To ensure a good temporal coverage we only included

those stations, which have continuous measurements for the entire 1983 to 2018 period. This reduces the station number from more than 2000 to 466.

Monthly mean anomalies are calculated for each time series (Figure 5-9 above). These monthly means are averaged over all available stations and seasonal corrected to obtain monthly anomalies (Figure 5-9 middle). The slope (trend) and the p-value between LST and T2m monthly mean anomalies are calculated together with the bias and bc-RMS of the LST-minus-T2m difference time series (Figure 5-9 below). A non-zero bias indicates an offset between LST and T2m anomalies. Ideally the trend of the difference should be zero or statistically not significant, indicating that the trends in LST and T2m anomalies are identical. A summary of the results is presented in Figure 5-9 and Table 5-6.

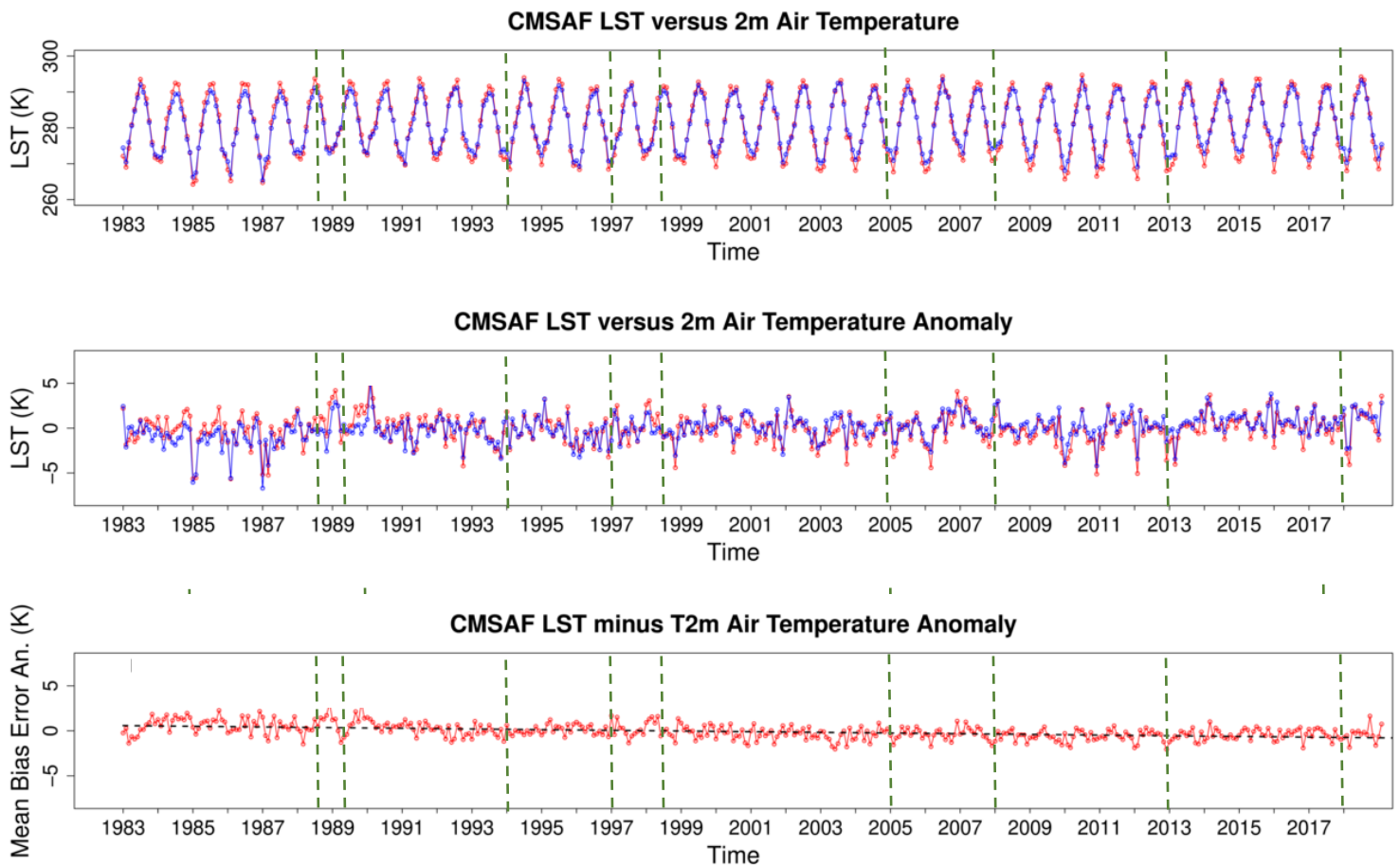


Figure 5-9: Monthly daily mean time series of the monthly CM SAF LST (red) as compared to EUSTACE T2m air temperature measurements (blue) at 466 stations over Europe. The air temperature measurements, which originate from the European Climate Assessment and Dataset (ECA&D), were homogenized in the EUSTACE project. The black dashed line represents the Theil-Sen linear trend. a) Monthly means, b) Temperature anomalies (seasonal corrected) and c) CM SAF minus T2m temperature anomalies. The green dashed lines show major satellite changes with the transition MVIRI/SEVIRI in 2005.


Table 5-6: Slope (trend), bias and bc-RMS for the CM SAF LST minus T2m monthly anomalies averaged spatially for different regions. The trend ($t_{CM\ SAF-T2m\ anom}$) is shown for the period 1983-1998 (left), 1999-2019 (middle) and 1983-2019 (right). In addition, we show trends calculated for T2m (t_{T2m}) and CM SAF LST (t_{CMSAF}) anomalies separately for each period. Statically significant trends ($p < 0.05$) are marked with a *. The last five rows relate to the average over all stations. Compliance with the requirements is presented in colours (threshold, target, optimal).

Region	1983-1998	1999-2019	1983-2020
Southern Europe			
bias $_{CM\ SAF-T2m\ anom}$ (K)			-0.03
bc-RMS $_{CM\ SAF-T2m\ anom}$ (K)			0.86
$t_{CM\ SAF-T2m\ anom}$ (K/dec)	-0.28	0.07	-0.26
t_{T2m} (K/dec)	0.55 K*	0.33 *	0.31 *
t_{CMSAF} (K/dec)	0.16 K*	0.44 *	0.03 *
Eastern Europe			
bias $_{CM\ SAF-T2m\ anom}$ (K)			-0.07
bc-RMS $_{CM\ SAF-T2m\ anom}$ (K)			1.26
$t_{CM\ SAF-T2m\ anom}$ (K/dec)	-0.39	0.13	-0.33
t_{T2m} (K/dec)	0.22	0.48 *	0.49 *
t_{CMSAF} (K/dec)	-0.27	0.59 *	0.16
Western Europe			
bias $_{CM\ SAF-T2m\ anom}$ (K)			-0.06
bc-RMS $_{CM\ SAF-T2m\ anom}$ (K)			0.99
$t_{CM\ SAF-T2m\ anom}$ (K/dec)	-0.31	0.13	-0.30
t_{T2m} (K/dec)	0.53 *	0.36 *	0.40 *
t_{CMSAF} (K/dec)	0.16	0.48 *	0.08 *
Central Europe			
bias $_{CM\ SAF-T2m\ anom}$ (K)			-0.07
bc-RMS $_{CM\ SAF-T2m\ anom}$ (K)			1.16
$t_{CM\ SAF-T2m\ anom}$ (K/dec)	-0.40 K	0.07	-0.30
t_{T2m} (K/dec)	0.47 K	0.36	0.43 *
t_{CMSAF} (K/dec)	-0.05 K*	0.42	0.11
Europe (all stations)			
bias $_{CM\ SAF-T2m\ anom}$ (K)			-0.06
bc-RMS $_{CM\ SAF-T2m\ anom}$ (K)			0.85
$t_{CM\ SAF-T2m\ anom}$ (K/dec)	-0.51	-0.05	-0.34
t_{T2m} (K/dec)	0.36	0.37 *	0.44 *
t_{CMSAF} (K/dec)	-0.10 *	0.34 *	-0.06 *

There is an excellent agreement between the monthly CM SAF LSTs and the T2m (Figure 5-9 above). Minimum CM SAF LSTs are slightly colder, maximum LSTs are slightly warmer than T2m for the entire time series, which reflects the different thermal behaviour of T2m and LST and is in line with the findings of Good et al. (2017).

There is also a very good agreement ($R=0.91$) between CM SAF LST and T2m anomalies from 1991 onward (Figure 5-9 middle). This strong agreement clearly outlines that CM SAF LSTs can be used to characterize temperature anomalies over Europe for the new WMO climatological norm period starting in 1991. In the mid-80s there are a couple of years, where the CM SAF maximum temperature anomalies are overestimated by 1 K to 2 K.

For the period 1983 to 1998, we observe a -0.3 K/decade to -0.5 K/decade trend in bias. Although this trend is statistically not significant, this is an indication that there are some instability in the CM SAF LST CDR for this period. For 1983 to 1998 the trends calculated from

	<p style="text-align: center;">Validation Report Land Surface Temperature (LST) Edition 2</p>	<p>Doc.: SAF/CM/MeteoSwiss/VAL/MET/LST/2.0 Issue: 2.1 Date: 07.09.2023</p>
---	--	--

the EUSTACE T2m and CM SAF LSTs do not match (Table 5-6), which also points to an instability in the CM SAF LST CDR. Please note that the observed decadal trends in bias are small and well within the 1 K/decade stability threshold requirement. This negative bias reflects the slight CM SAF LST overestimation in the mid-80s (Figure 5-9), which is likely caused by some instabilities for MFG3 to MFG5.

For the period 1999 to 2019 the CM SAF LST CDR is very stable. We observe a statistically not significant -0.1 K/decade trend in the CM SAF minus T2m bias. This stability is within the optimal requirement. For 1999 to 2019 the decadal T2m trends are almost identical to CM SAF LST trends for all regions in Europe (Table 5-6). Trends calculated for all regions agree within +/- 0.1 K/decade. Averaged over all stations, a trend of 0.37 K/decade is obtained for CM SAF LSTs, which compares well with the equivalent T2m trend of 0.34 K/decade. This suggests that the CM SAF CDR is stable and can be used to calculate temperature trends over Europe from 1999 onward.

5.2.2 Stability against ESA CCI MODIS AQUA LST

Monthly diurnal cycle CM SAF LSTs were carefully resampled in space and time to match the monthly the ESA CCI MODIS Aqua LSTs. The stability analysis is performed on night-time data (01:30 am).

The stability analysis confirms the findings from the T2m analysis over Europe. Mean monthly CM SAF LST anomalies are very closely aligned ($R=0.94$) with MODIS LST for the period 2003 to 2018 in Europe (Figure 5-10). The decadal trend in bias for the CM SAF minus MODIS LST over Europe (lat 39° to 54°, lon 9° to 20°) is statistically not significant and with -0.1 K/decade within the optimal requirement (Figure 5-11, Table 5-7). For the Tropics (lat 7° to 5°, lon 13° to 30°), we also observe a -0.1 K/decade trend in bias, which is not significant. In Southern Africa (lat -30° to -18°, lon 17° to 30°), the 0.3 K bias trend is significant. In Northern Africa (lat 5° to 30°, lon 0° to 15°) we observe a significant 0.8 K/decade bias trend. We conclude here, that the CM SAF CDR is very stable for Europe, the Tropics and Southern Africa. For Northern Africa, there are some instabilities in either the CM SAF or MODIS ESA CCI CDR, which need to be further investigated.

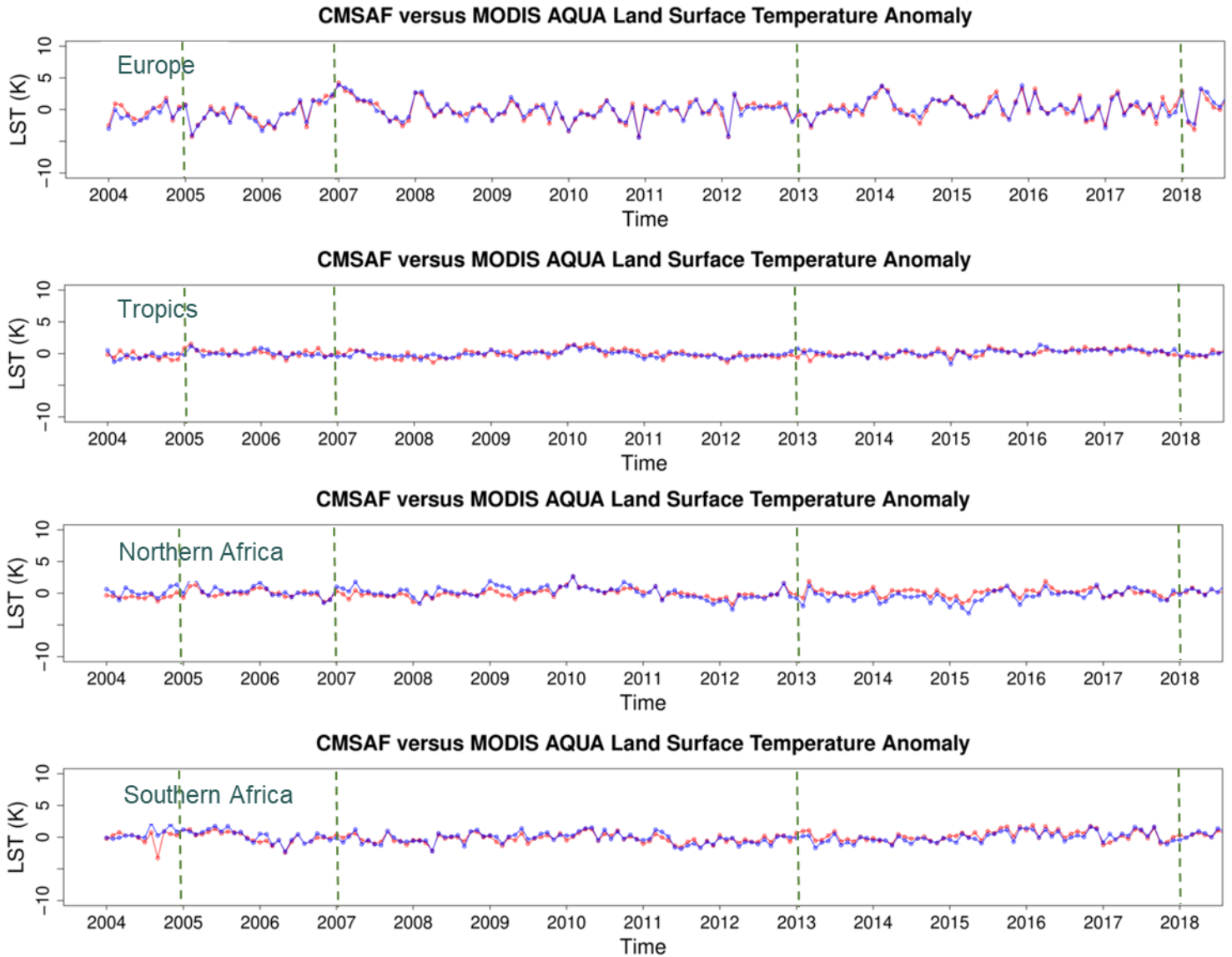


Figure 5-10: Monthly daily mean anomaly time series of the CM SAF LST (red) as compared ESA CCI MODIS Aqua (blue) in Europe (upper), the Tropics (second), Northern Africa (third) and Southern Africa (fourth). The analysis is performed at night-time (01:30 am local time). The dashed green lines show major satellite transitions with the change MVIRI/SEVIRI in 2005.

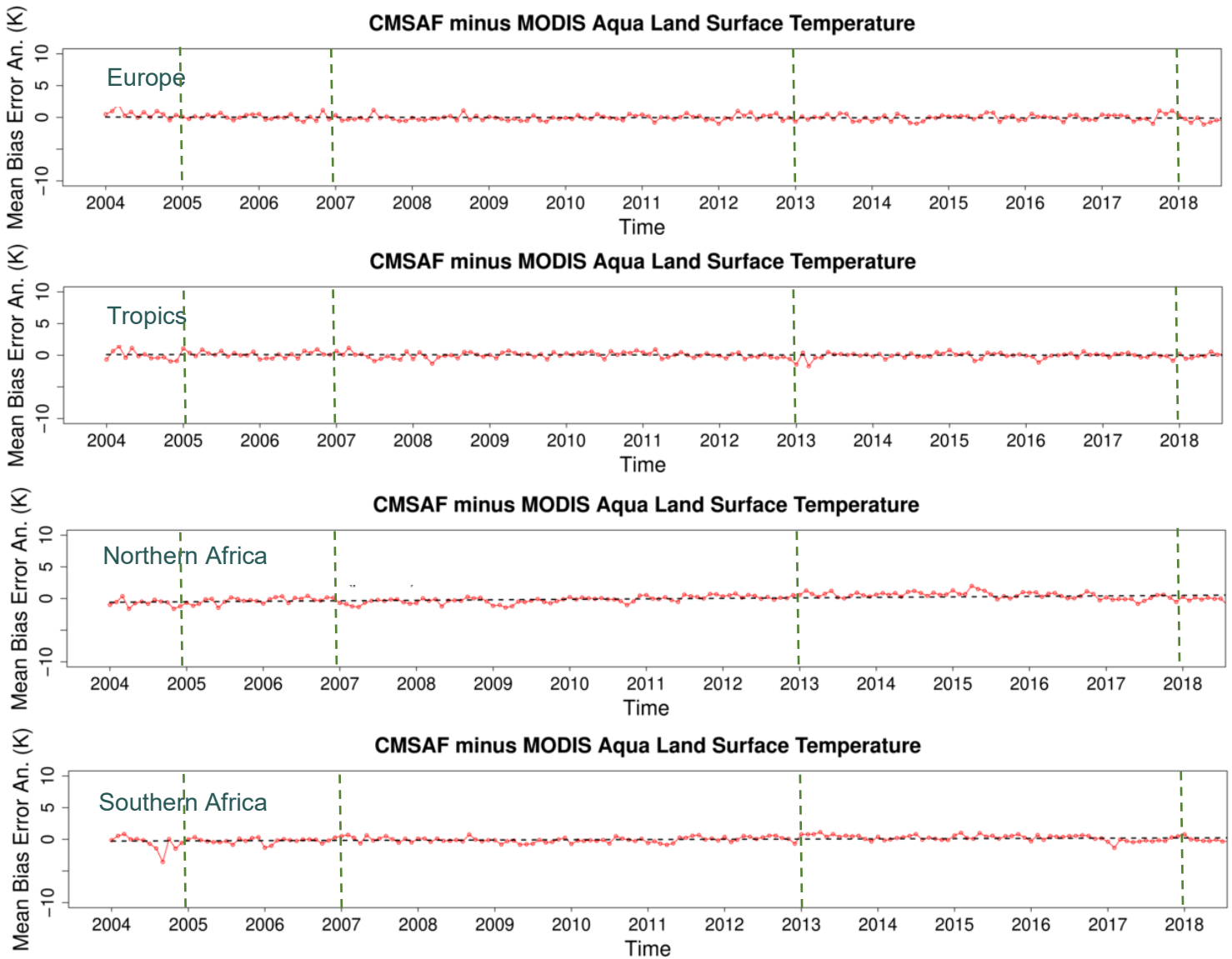


Figure 5-11: Bias of the monthly daily mean anomaly time series of the CM SAF minus ESA CCI MODIS Aqua LST in Europe, the Tropics, Northern Africa and Southern Africa. The dashed green lines show major satellite transitions with the change MVIRI/SEVIRI in 2005.

Table 5-7: Slope (trend), bias and bc-RMS for the CM SAF LST minus ESA CCI MODIS monthly anomalies. The trend is calculated for the period 2004 to 2018. Statistically significant trends ($p < 0.05$) are marked with a *. Compliance with the requirements is presented in colours (threshold, target, optimal)

CMSAF – MODIS LST anomaly 2003-1018	Europe	Tropics	Northern Africa	Southern Africa
$t_{\text{CMSAF-T2m anom}}$ (K/dec)	-0.11	-0.08	0.78*	0.30*
$\text{Bias}_{\text{CMSAF-T2m anom}}$ (K)	0.03	0.02	0.04	0.01
$\text{Bc-RMS}_{\text{CMSAF-T2m anom}}$ (K)	0.50	0.49	0.65	0.55

5.2.3 Summary Stability

The stability of the CM SAF LST data, as evaluated against EUSTACE T2m and ESA CCI MODIS Aqua LSTs is well within the threshold requirements as outlined in Table 5-8. In Europe, we reach the optimal 0.1 K stability for the period 1999 to 2019. For the period 1999-2019 trends in CM SAF LST of 0.34 K/decade are obtained, which compare well with the equivalent station-based T2m trends of 0.37 K/decade. The comparison with ESA CCIs MODIS Aqua outlines that the CM SAF LST CDR is also within the optimal requirement for 2003-2018 for the Tropics, while the 0.3 K target requirement is reached in Southern Africa. In North Africa, we observe a significant 0.8 K bias.

Table 5-8: Overall CM SAF LST stability compared to EUSTACE T2m and ESA CCI MODIS Aqua. The stability is given in terms of decadal trend in bias of the CM SAF minus reference data anomaly. Compliance with the requirements is presented in colours (threshold, target, optimal)

Trend of bias	Europe	Tropics	North Africa	South Africa
EUSTACE T2m 1999-2019	-0.1 K/decade	X	X	X
EUSTACE T2m 1983-1998	-0.5 K/decade	X	X	X
EUSTACE T2m 1983-2019	-0.3 K/decade	X	X	X
ESA CCI MODIS Aqua 2003-2019	-0.1 K/decade	-0.1 K/decade	0.8* K/decade	0.3* K/decade

5.3 Comparison against other LST CDRs

We compared the monthly diurnal cycle CM SAF LST CDR against:

- ESA CCIs combined LEO & GEO monthly LST climatology 2009-2020
- ESA CCIs MODIS Aqua monthly LST climatology 2003-2018
- LSA SAFs SEVIRI monthly LST climatology 2004-2019

Comparisons of the CM SAF LST CDR with other satellite LST CDRs were performed for the entire Meteosat disk for different climatological periods depending on the availability of the reference data. The ESA CCI combined LEO & GEO LSTs and the LSA SAF LSTs match the CM SAF LSTs in acquisition time and grid. The ESA CCI LEO & GEO data are available as monthly files for three hourly time steps, the LSA SAF LST climatology is provided as hourly diurnal cycle composites. Here we selected the 0 am and 0 pm time steps for comparisons. For the ESA CCI MODIS comparison the monthly diurnal CM SAF were carefully resampled in time to match the MODIS acquisition time. Note that the monthly CM SAF data are available for every hour of the date. Hence, the maximum deviation with the MODIS LST acquisition time (1:30 am local time) is +/- 30 minutes. As LSTs can vary strongly within 30 min during the day due to solar heating, the MODIS comparison was only performed during night-time.

Table 5-9: Statistics CM SAF – ESA CCI LEO & GEO LST (fulldisk, 2009 to 2020).

	DAY 12:00 UTC (K) VZA < 60° (fulldisk)		NIGHT 00:00 UTC (K) VZA < 60° (fulldisk)	
	bias	bc-RMS	Bias	bc-RMS
January	0.1 (-0.3)	2.6 (2.8)	-0.8 (-0.2)	2.2 (2.4)
March	1.7 (1.2)	3.3 (3.5)	-0.9 (-0.2)	2.2 (2.5)
June	0.0 (0.4)	2.4 (2.8)	0.5 (0.6)	1.4 (1.6)
September	-0.5 (0.2)	2.6 (2.7)	0.4 (0.7)	1.6 (1.8)
All	0.3 (0.4)	2.7 (2.9)	-0.2/0.2	1.9/2.1

Table 5-10: Statistics CM SAF – ESA MODIS AQUA LST (fulldisk, 2003 to 2018).

	NIGHT 01:30 local time (K) VZA < 60° (fulldisk)	
	bias	bc-RMS
January	0.3 (0.3)	1.2 (1.8)
March	0.4 (0.7)	1.4 (1.8)
June	0.8 (1.5)	1.8 (1.6)
September	0.7 (0.8)	1.6 (1.6)
ALL	0.5 (0.9)	1.5 (1.7)

Table 5-11: Statistics CM SAF – LSA SAF LST (fulldisk, 2004 to 2019).

	DAY 12:00 UTC (K) VZA < 60 (fulldisk)		NIGHT 00:00 UTC (K) VZA < 60° (fulldisk)	
	bias	bc-RMS	Bias	bc-RMS
January	0.0 (-0.3)	1.6 (2.9)	-0.1 (1.1)	1.1 (2.9)
March	0.1 (-0.6)	1.9 (2.7)	-0.1 (0.5)	1.0 (2.1)
June	0.4 (-0.3)	2.1 (2.6)	0.9 (1.4)	1.2 (1.6)
September	-0.7 (-0.9)	2.0 (2.2)	0.4 (1.1)	1.2 (1.6)
All	-0.1 (-0.5)	1.9 (2.6)	0.3 (1.0)	1.2 (2.0)

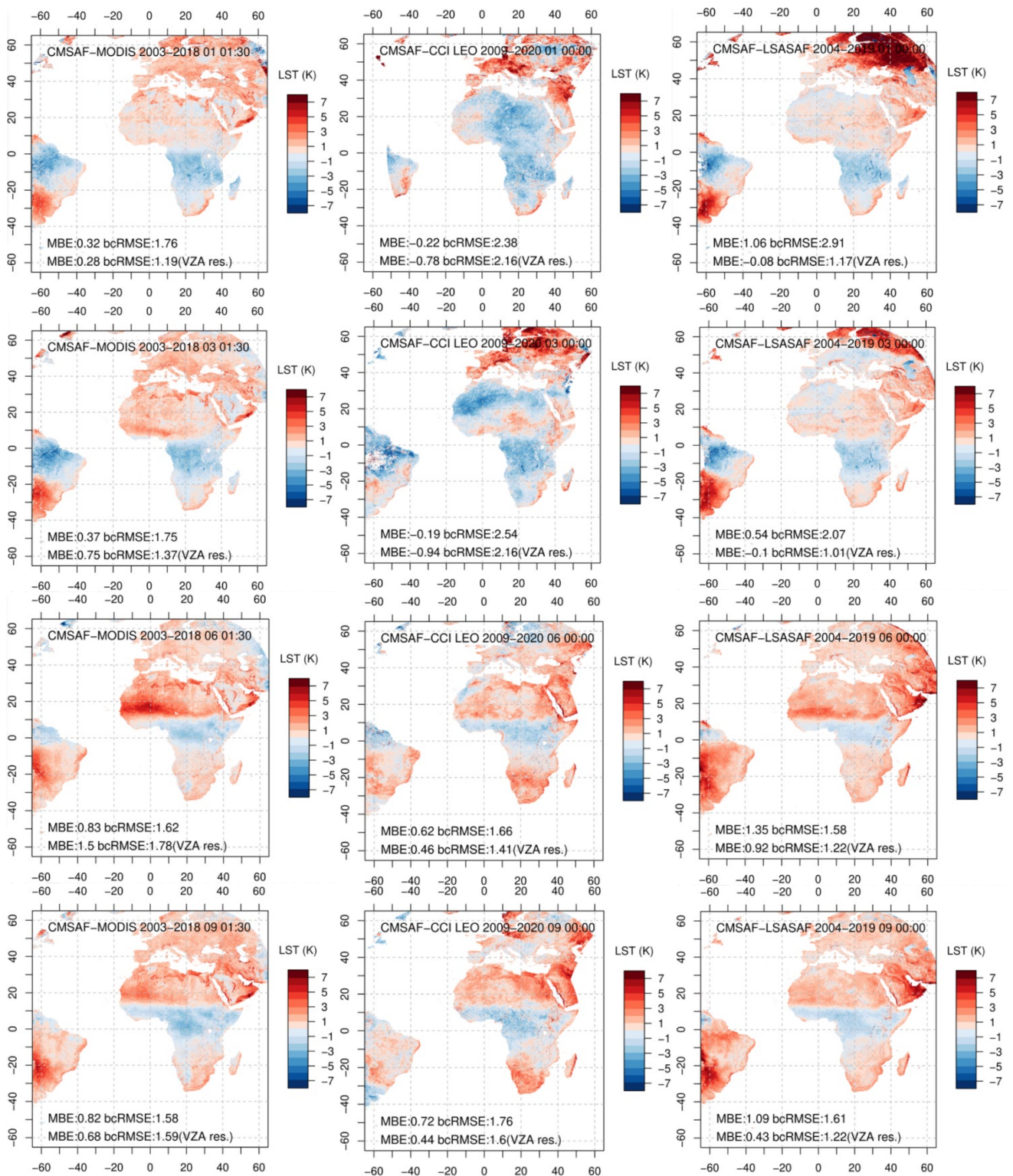



Figure 5-12: Satellite-based comparison for January, March, June and September. Left: CM SAF minus ESA CCI MODIS LST (1:30 am local time, 2003-2018), middle: CM SAF minus ESA CCI LEO & GEO LST (0 am, 2009-2020) and right: CM SAF LST minus LSA SAF LST (0 am, 2004-2019).

	Validation Report Land Surface Temperature (LST) Edition 2	Doc.: SAF/CM/MeteoSwiss/VAL/MET/LST/2.0 Issue: 2.1 Date: 07.09.2023
---	---	---

The differences between the different CDRs are illustrated in Figure 5-12. Detailed statistics of the satellite comparison are provided in Table 5-9 (ESA CCI LEO & GEO CDR), Table 5-10 (ESA CCI MODIS Aqua CDR) and Table 5-11 (LSA SAF CDR).

The overall biases between the CM SAF and the reference CDRs is below 0.5 K for most satellite-based comparison for VZAs < 60°. Seasonal biases are generally below 1 K showing the excellent overall agreement of the CM SAF CDR with the reference CDRs. The precision is with a bc-RMS of about 1.5 K best for the CM SAF – MODIS night-time comparison for VZAs < 60°. A low 1.2 K bc-RMS is also observed for the CM SAF – LSA SAF LSTs for viewing angles < 60°. For the daytime comparison the bc-RMS is in the order of 2 K to 3 K. This likely reflects the high variability of LSTs during daytime.

There are distinct regional and seasonal pattern in the observed bias (Figure 5-12). For January and March, we observe a distinct positive bias in Northern Europe for the CM SAF minus ESA CCI LEO and the LSA SAF CDRs. Please note that those large positive biases (> 3 K) are not reproduced in the CM SAF minus MODIS comparison. Here we observe an overall 0.7 K positive bias for all seasons. We therefore conclude that the CM SAF CDR is accurate at high VZAs in Northern Europe.

Over the Sahara, there is a distinct 1 K to 4 K positive bias, which is reproduced in all CDR comparison for June and September (Figure 5-12). The bias is strongest for the CM SAF minus MODIS comparison for June. As the ESA CCI MODIS LSTs are generated using the same surface emissivity as the CM SAF, we assume here that this bias reflects a deficit of the CM SAF single channel model to correctly represent the surface overheating in desert regions. Please note that for January there is no overall positive bias in the CM SAF minus LSA SAF comparison.

For the Tropics and Subtropics, we observe a 1 K to 3 K negative bias in all CM SAF minus reference data comparison (Figure 5-12). The location of this bias varies in time and space. A larger cloud contamination of the CM SAF LSTs could explain this negative bias. Moreover, we have also observed a negative bias for very high TCWV at the station Dahra during the moist season (Figure 5-5). Single channel LST models have large uncertainties for high TCWV (Duguay-Tetzlaff et al. 2015). This negative bias likely reflect those uncertainties.

Overall, we observe a very good agreement between the CM SAF and ESA CCI MODIS LST CDR outside of desert regions for VZAs < 60° (bias 0.5 K, bc-RMS 1.5 K). There is also an excellent agreement between the LSA SAF and the CM SAF CDR for VZAs < 60° during the night (bias 0.3 K, bc-RMS 1.2 K). The positive bias in desert regions is less pronounced for the CM SAF minus LSA SAF data. For large VZAs we observe distinct differences between the LSA SAF and CM SAF CDRs which need to be further investigated.

6 Conclusions

This report validates the CM SAF LST CDR. The study uses independent observation sources and other satellite-based LST datasets as references. The comparisons were performed against four reference data sets: in-situ LST from KITs validation stations, EUSTACE T2m measurements, ESA CCI MODIS Aqua LST CDR, ESA CCI combined LEO & GEO CDR and the LSA SAF SEVIRI LST CDR. With these, the best possible effort to assess accuracy, precision and temporal stability of the derived LST products has been made.

The validation results can be summarized as:

- **Precision:**

The CM SAF LST products fulfil the 2.0 K threshold (hourly) and 1.0 K target (monthly) precision when compared to more than 50,000 in-situ LST measurements from KITs validation sites including a large variety of atmospheric conditions (Table 6-1). Comparisons against in situ data show a precision for the CM SAF LST very close or within the uncertainty of the ground estimates.

Table 6-1: Summary of CM SAF LST precision (bias-corrected RMS error) compared to target requirements (threshold, target, optimal) evaluated at KITs validation stations.

Bc-RMS	hourly	monthly
CM SAF LST	1.9	1.0

- **Accuracy:**

The CM SAF LST products fulfil the 1.0 K target (hourly) and 0.5 K optimal (monthly) accuracy when compared with more than 50,000 in-situ LST measurements from KITs validation sites (Table 6-2).

Table 6-2: Summary of CM SAF LST accuracy (mean bias error) compared to target requirements (threshold, target, optimal) evaluated at KITs validation stations.


Bias	hourly	monthly
CM SAF LST	0.6	0.4

- **Decadal stability:**

The CM SAF LST products fulfil the 1 K/decade target decadal stability requirement when compared against EUSTACE T2m and ESA CCI MODIS LSTs (Table 6-3). In Europe the CM SAF LSTs are stable from 1999 onward (0.1 K/decade), while the trend in bias reaches a maximum of 0.8 K/decade over Northern Africa.

Table 6-3: Summary of CM SAF LST stability (decadal trend in bias) compared to target requirements (threshold, target, optimal) when evaluated over the various sites and regions.

Trend in bias	EUSTACE T2m 1983-1998	EUSTACE T2m 1998-2019	EUSTACE T2m 1983-2019	ESA CCI MODIS LST 2003-2018	
	Europe	Europe	Europe	Europe	North Africa
CM SAF LST	-0.5 K/dec	-0.1 K/dec	-0.3 K/dec	-0.1 K/dec	0.8 K/dec

	Validation Report Land Surface Temperature (LST) Edition 2	Doc.: SAF/CM/MeteoSwiss/VAL/MET/LST/2.0 Issue: 2.1 Date: 07.09.2023
---	---	---

Whereas the validation results presented for the CM SAF LST show that these reach their respective target accuracies for most of the encountered situations, the report also revealed the following limitations:

- Low precision and accuracy for very moist atmospheres (TCWV > 45 mm): bc-RMS of about 3 K and bias of about - 2 K for very moist atmospheres at the tropical station Dahra.
- LST overestimation during summer in desert regions. A 2-3 K seasonal bias at high temperatures (> 310 K) found for the desert station RMZ. This bias is also present in satellite inter-comparisons in the Sahara desert.

This study has outlined that monthly temperature anomalies calculated from CM SAF data agree very well with station-based air temperature measurements and other satellite data. We therefore conclude that the CM SAF LST CDR is very well suited to assess temperature anomalies for the new WMO 1991-2020 norm period. The study also suggests that CM SAF LST data can be used to assess warming trends from 1999 onward. For Europe (1999 to 2019) significant trends in LST of 0.37 K/decade are obtained, which match the T2m trends of 0.34 K/decade. Trend analysis before 1999 or in desert regions is not recommended. Here the CM SAF LST data do not have the required homogeneity.

7 References

Dee, D. P., and 35 co-authors. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. R. Meteorol. Soc.* 2011, 137, 553-597.

Duguay-Tetzlaff, A., V. A. Bento, F. M. Göttsche, R. Stöckli, J. P. A. Martins, I. Trigo, F. Olesen, J. S. Bojanowski, C. da Camara and H. Kunz, *Meteosat land surface temperature climate data record: Achievable accuracy and potential uncertainties*, *Remote Sens.*, 2015

Dodd, E., Ghent, D., Jimenez, C., and S. Ermida, S., *LST_cci Algorithm Theoretical Basis Document v1.2*. ESA, 2019.

Dodd, E., Ermida, S., Jimenez, C. and Ghent, D., *LST_cci Product User Guide v1.2*. ESA. 2021.

Eckardt, F., Soderberg, K., Coop, L., Muller, A., Vickery, K., Grandin, R., Jack, C., Kapalanga, T., Henschel, J. (. The nature of moisture at Gobabeb, in the central Namib Desert. *J. Arid Environ.*, 2013, 93, 7–19.

Ermida, S.L., Trigo, I.F., DaCamara, C.C., Göttsche, F.M., Olesen, F.S., Hulley, G. Validation of remotely sensed surface temperature over an oak woodland landscape—The problem of viewing and illumination geometries. *Remote Sensing of Environment*, 2014, 148, 16-27.

Ermida, S.L., Trigo, I.F., DaCamara, C.C. and J.-L. Roujean. Assessing the potential of parametric models to correct directional effects on local to global remotely sensed LST. *Remote Sens. Environ*, 2018a, 209.

Ermida, S.L., Trigo, I.F., DaCamara, C.C. and A.C. Pires, 2018b. A Methodology to Simulate LST Directional Effects Based on Parametric Models and Landscape Properties. *Remote Sens.* 10, 2018b. 1114.

Freitas, S., Trigo, I., Bioucas-Dias, J., Göttsche, F. Quantifying the uncertainty of land surface temperature retrievals from SEVIRI/METEOSAT. *IEEE Transactions on Geoscience and Remote Sensing* 2010, 48, 523-534.

Good, E.J., Ghent, D.J., Bulgin, C.E., & Remedios, J.J. A spatiotemporal analysis of the relationship between near-surface air temperature and satellite land surface temperatures using 17 years of data from the ATSR series. *Journal of Geophysical Research: Atmospheres*, 2017. 122(17), 9185 – 9210.

Good, E.J., Aldred, F.M., Ghent, D.J., Veal, K.L., Jimenez. An Analysis of the Stability and Trends in the LST_cci Land Surface Temperature Datasets Over Europe. *EARTH AND SPACE SCIENCE*, 2022, 9(9), e2022EA002317.

Göttsche, F., Hulley, G. Validation of six satellite-retrieved land surface emissivity products over two land cover types in a hyper-arid region. *Remote Sens. Environ.* 2012, 124, 149– 158.

Göttsche, F., F. Olesen, I. Trigo, Bork-Unkelbach, A., Martin M. Long term validation of land surface temperature retrieved from MSG/SEVIRI with continuous in-situ measurements in Africa. *Remote Sensing* 2016, 8, 410.

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

Göttsche, F., Olesen, F., Bork-Unkelbach, A. Validation of operational Land Surface Temperature products with three years of continuous in-situ measurements. In Proceedings of the EUMETSAT Meteorological Satellite Conference, Oslo Norway, 2011, Abstract number P59_S2_06.

Göttsche, F.-M., Olesen, F.S., Trigo, I.F., Bork-Unkelbach, A., Martin, M.A., Long Term Validation of Land Surface Temperature Retrieved from MSG/SEVIRI with Continuous in-Situ Measurements in Africa. *Remote Sensing*, 2016, 8(5), 410-437.

Govaerts, Y., Clerici, M., and N. Clerbaux, Operational calibration of the Meteosat radiometer VIS band. *IEEE Transactions On Geoscience and Remote Sensing*, 42(9):2004.

Guillevic, P.C., Bork-Unkelbach, A., Göttsche, F.M., Hulley, G., Gastellu-Etchegorry, J.-P., Olesen, F.S., Privette, J.L. Directional viewing effects on satellite land surface temperature products over sparse vegetation canopies—A multisensor analysis. *IEEE Geoscience and Remote Sensing Letters* 2013, 10 (6), 1464-1468.

Jiang, K., Pan, Z., Pan, F., Wang, J., Han, G., Song, Y., Zhang, Z., Huang, N., Ma, S., Chen, X., Zhang, Z., Men, J. The global spatiotemporal heterogeneity of land surface-air temperature difference and its influencing factors. *Science of The Total Environment*, 2022. 838, Part 2.

John, V.O., Tabata, T., Rüthrich, F., Roebeling, R., Hewison, T., Stöckli, R., Schulz, J. On the Methods for Recalibrating Geostationary Longwave Channels Using Polar Orbiting Infrared Sounders. *Remote Sens.* 2019, 11, 1171.

Kabsch, E., Olesen, F., Prata, F. Initial results of the land surface temperature (LST) validation with the Evora, Portugal ground-truth station measurements. *Int. J. Remote Sens.* 2008, 29, 5329–5345.

Kendall, M. A new measure of rank correlation. *Biometrika* 1938, 30, 81–93.

Kilpatrick, K.A., Podestá, G., Walsh, S., Williams, E., Halliwell, V., Szczodrak, M., Brown, O.B., Minnett, P.J., Evans, R. A decade of sea surface temperature from MODIS. *Remote Sens.*, 2015, 165, 27–41.

Klok, E.J. and Klein Tank, A.M.G. Updated and extended European dataset of daily climate observations. *International Journal of Climatology*, 2008, 29, 1182–1191.

Kondratyev, K. *Radiation in the Atmosphere*. Academic Press: New York, NY, USA, 1969.
Mann, H. Nonparametric tests against trend. *Econometrica* 1945, 13, 245–259.

Rayner, N. A., Auchmann, R., Bessembinder, J., Brönnimann, S., Brugnara, Y., Capponi, F., Carrea, L., Dodd, E. M. A., Ghent, D., Good, E., Høyer, J. L., Kennedy, J. J., Kent, E. C., Killick, R. E., van der Linden, P., Lindgren, F., Madsen, K. S., Merchant, C. J., Mitchelson, J. R., Morice, C. P., Nielsen-Englyst, P., Ortiz, P. F., Remedios, J. J., van der Schrier, G., Squintu, A. A., Stephens, A., Thorne, P. W., Tonboe, R. T., Trent, T., Veal, K. L., Waterfall, A. M., Winfield, K., Winn, J., & Woolway, R. I. The EUSTACE Project: Delivering Global, Daily Information on Surface Air Temperature, *Bulletin of the American Meteorological Society*, 2020, 101(11), E1924-E1947.

Stante, F., Ermida, vS. 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 Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2023, 16, 4064-4079.

Seemann, S.W., Borbas E. E., R. O. Knuteson, G. R. Stephenson, Huang H. L. Development of a global Infrared land surface emissivity database for application to clear sky sounding retrievals from multi-spectral satellite radiance measurements. J. of Appl. Meteor. and Climatol. 2008, 47.

Squintu, A.A., van der Schrier, G., & van den Besselaar, E.J. EUSTACE / ECA&D: European land station daily air temperature measurements, homogenised. Retrieved 22.11, 2023, 2019a, from European Climate Assessment & Dataset: <https://catalogue.ceda.ac.uk/uuid/81784e3642bd465aa69c7fd40ffe1b1b>

Squintu, A.A., van der Schrier, G., Brugnara, Y., & Klein Tank, A. Homogenization of daily temperature series in the European Climate Assessment & Dataset. International Journal of Climatology, 2019b, 39(3), 1243-1261.

Theocharous, E., Usadi, E., Fox, N.P. CEOS Comparison of IR brightness temperature measurements in Support of satellite validation. Part. I: Laboratory and ocean surface temperature comparison of radiation thermometers. National Physical Laboratory: Teddington, UK, 2010.

Trigo, I., Peres, L., DaCarnara, C., Freitas, S. Thermal land surface emissivity retrieved from SEVIRI/METEOSAT. IEEE Transactions on Geoscience and Remote Sensing 2008, 46, 307-315.

Trigo, I., Freitas, S.C., Bioucas-Dias, J., Barroso, C., Monteiro, I.T., Viterbo, P., J.P. Martins, A Pires and S. Ermida. Algorithm theoretical basis document for Land Surface Temperature Directional Effects (SAF/LAND/IPMA/ATBD_MLST/1.0). Land Surface Analysis Satellite Application Facility (LSA- SAF) 2018.

Theil, H. A rank-invariant method of linear and polynomial regression analysis. Parts 1–3. Ned Akad Wetensch 1950, 53, 386–392, 521–525, 1397–1412.

Veal, K., Ermida, S., Ghent, D., Perry, M., and I, Trigo. ESA Land Surface Temperature Climate Change Initiative (LST_cci): Monthly multisensor Infra-Red (IR) Low Earth Orbit (LEO) and Geostationary Earth Orbit (GEO) land surface temperature (LST) level 3 supercollated (L3S) global product (2009-2020), version 1.00. 11 July 2022. NERC EDS Centre for Environmental Data Analysis.

Wan, Z.; Dozier, J. A Generalized Split-Window algorithm for retrieving land-surface temperature from space. IEEE Transactions on Geoscience and Remote Sensing 1996, 34, 892-905.

Wan, Z., Zhang, Y., Zhang, Y.Q., Li, Z.-L. Quality assessment and validation of the global land surface temperature. International Journal of Remote Sensing 2004, 25, 261-274.

8 Glossary

ATBD	Algorithm Theoretical Baseline Document
AVHRR	Advanced Very High Resolution Radiometer
BC-RMS	Bias-Corrected RMS
CSIRO	Commonwealth Scientific and Industrial Research Organisation
CDO	Climate Data Operators
CM SAF	Satellite Application Facility on Climate Monitoring
DRR	Delivery Readiness Review
DWD	Deutscher Wetterdienst (German MetService)
ECMWF	European Centre for Medium Range Forecast
ECV	Essential Climate Variable
ERA5	ECMWF Re-Analysis dataset
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FCDR	Fundamental Climate Data Record
GCOS	Global Climate Observing System
GOES	Geostationary Operational Environmental Satellite
GSW	Generalised Split-Window
HIRS	High Resolution Infrared Radiation Sounder
IFS	Integrated Forecast System
INRA	French National Institute for Agricultural Research
IPMA	Portuguese Institute for the Ocean and Atmosphere (IPMA-Instituto Português do Mar e da Atmosfera)
KIT	Karlsruhe Institute of Technology
LST	Land Surface Temperature
LTS	Statistical Land Surface Temperature
LTP	Physical Land Surface Temperature
LSA SAF	Land Surface Analysis Satellite Applications Facility
MBE	Mean bias error
MetOp	Meteorological Operational satellite
MVIRI	Meteosat Visible and InfraRed Imager
MFG	Meteosat First Generation
MSG	Meteosat Second Generation
MODIS	Moderate Resolution Imaging Spectroradiometer

NOAA	National Oceanic and Atmospheric Administration
NWP	Numerical Weather Prediction
PMW	Physical Mono-Window Model
PRD	Product Requirement Document
PUM	Product User Manual
RTTOV	Radiative Transfer for TOVS
SEVIRI	Spinning Enhanced Visible and InfraRed Imager
SMW	Statistical Mono-Window Model
SZA	Solar Zenith Angle
TCDR	Thematic Climate Data Record
TCWV	Total atmospheric column water vapour
VZA	Viewing Zenith Angle