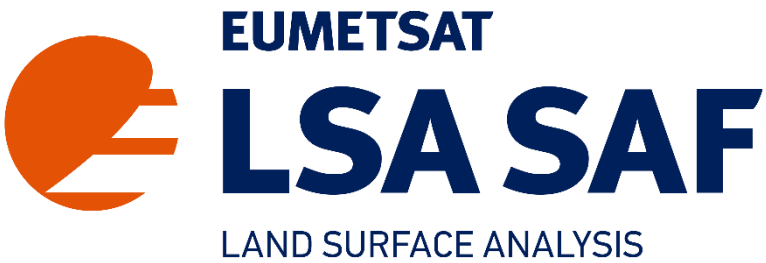
VALIDATION REPORT

LST (LSA-001), ELST (LSA-002)



Reference Number:	SAF/LAND/IM/VR_LST/v1.5
Issue/Revision Index:	Issue II/2016
Last Change:	19/12/2016

DOCUMENT SIGNATURE TABLE

	Name	Date	Signature
Prepared by :	I. F. Trigo, I. Monteiro, S. Coelho, F. Olesen, F. Göttsche, R. Torres		
Approved by :	LSA SAF Project Manager (IM)		

DOCUMENTATION CHANGE RECORD

Issue / Revision	Date	Description:
Version I/2009	17/03/2008	Version to be presented to OR-3
V1.4, I/2016	11/04/2016	LSA-001: updated validation results; LSA-002 validation of revised product to be presented at ELST ORR (longer time period; comparison with in situ estimates; note on the quality of the AVHRR cloud mask).
V1.5, II/2016	19/12/2016	Changes following ORR_EPS, with the section on ELST (AVHRR LST) totally re-written: - longer period for ELST validation. - the comparison between LST_AVHRR and LST_SEVIRI is performed taking into account the viewing geometry, orography and land-cover - LST_AVHRR is also assessed through comparison with in situ measurements. Tables with algorithm versioning description (Table 2 and Table 3) added.

Executive Summary

Land Surface Temperature retrieved from SEVIRI/Meteosat, LST_SEVIRI (LSA-001 product), is generated on an operational basis since February 2005 for the European region and since July 2005 for the whole Meteosat disk. The regular generation of LST from AVHRR/MetOp, LST_AVHRR, began in September 2007, however, the production of daily composites (ELST, LSA-002 product) is available from 2015 onwards. The main algorithm for LST estimation from both sensors is based on a Generalized Split Window (GSW) that uses the difference between two adjacent window channels to correct the atmospheric absorption. This document presents the validation results obtained for the LSA-SAF LST products. LST_SEVIRI is compared with that retrieved from MODIS and with in-situ LST obtained from the LST validation station at Evora (Southern Portugal) and in Africa (Senegal and Namibia). LST_AVHRR is compared with LST_SEVIRI and for available match-ups with in situ measurements.

The comparison of SEVIRI and MODIS LST retrievals with in situ observations is consistent with the analysis performed for the three selected areas. The differences between ground and satellite-derived values show high variability for daytime for both sensors, with LST_SEVIRI overestimating in situ values. The differences between satellite and in situ LST's are lower for night-time observations. In this case, both sensors tend to underestimate local measurements, with colder values obtained with MODIS.

Performing separate radiance measurements over the relevant endmembers at Evora, LST_SEVIRI has been validated with in-situ data from April 2009 to October 2012. Furthermore, the dependence of LST observations on viewing and illumination geometries has been investigated (Ermida et al., 2012), showing that the modelling of dynamic cover fractions to estimate in situ LST within a satellite sensor's FOVs reduces the bias in daytime LST considerably and allows more meaningful comparisons between LST obtained under different viewing geometries.

Up to five years of in-situ LST from KIT's long term validation stations in Africa have been used to validate LST_SEVIRI. Typically thousands of monthly match-ups between satellite LST and in-situ LST were available at each validation site and yielded highly linear relationships between the two quantities. Ignoring rainy seasons at Dahra (sub-tropical), the highest mean rmse for the African sites for daytime and night-time data combined was 1.6°C and the corresponding highest mean absolute bias was 0.1°C. The large number of match-ups allows temporally resolved validations of LST_SEVIRI which highlight seasonal differences in the retrieval algorithm's performance.

The assessment of AVHRR/Metop LST (LSA-002) is presented for data generated by the LSA SAF system for the period between January 2015 and November 2016. The validation is based on a comparison between LSR_AVHRR and LST_SEVIRI for 6 areas (~10º longitude x 10º latitude) within the MSG disk covering a wide variety of surface and atmospheric conditions. Overall, average differences between night-time LST range between -1.4ºC and 0.2ºC, while their standard deviation lies around 1.0ºC for most cases. As expected, daytime estimates show a larger discrepancies and also larger variability in space and time. The use of static emissivity fields in this version of LST_AVHRR leads to a seasonal variability in the differences to SEVIRI LST, particularly in regions where changes in vegetation, and therefore in emissivity, are more pronounced. Directional effects on LST are also clearly seen in the AVHRR – SEVIRI comparison, similar to those observed when comparing SEVIRI and MODIS LST products.

The ELST product has been compared to in situ estimates taken at KIT stations, specifically designed for the validation of LST satellite products. Differences are again larger for daytime than for night-time. The latter are less influenced by directional effects and therefore a more reliable

measure of the satellite product accuracy. On average LST_AVHRR underestimated night-time in situ estimates, with mean differences ranging between -0.8ºC and -1.7ºC. The root mean square differences for night-time LST lies within the target accuracy of 2ºC (2.0ºC obtained for Évora and Heimat, and 1.8ºC for Gobabeb).

TABLE OF CONTENTS

List of Tables

List of Figures

1 Introduction

LSA-SAF obtains Land Surface Temperature (LST) from directional surface emitted TIR radiances that are derived from cloud free SEVIRI/Meteosat (LSA-001) or AVHRR/MetOp (LSA-002) measurements. 'Surface emitted radiances' means that the measured top-of-atmosphere (TOA) radiances have been corrected for atmospheric attenuation along the path and reflected down-welling radiance has been removed. The 'surface' is formed by all elements that emit IR radiance (Norman and Becker, 1995). Thus, LST is the radiative skin temperature of the land surface, as measured in the direction of the remote sensor. Following Becker and Li (1995), directional radiometric temperature provides the best approximation of thermodynamic temperature that can be obtained from a radiance measurement. The main algorithm for LST estimation, from both SEVIRI/Meteosat and from AVHRR/MetOp, is based on the 'Generalised Split-Window' (GSW) formulation first developed for MODIS and AVHRR by Wan and Dozier (1996). Thus LSA-SAF LST is estimated as a linear function of clear-sky top of the atmosphere (TOA) brightness temperatures measured by the split-window channels available on SEVIRI (10.8 and 12.0 μm) and on AVHRR (channels 4 and 5), assuming surface emissivity is known for both bands (Trigo et al., 2008a). The estimation of the GSW parameters relies on linear regressions of synthetic brightness temperatures, obtained from radiative transfer simulations (using MODTRAN) over a wide range of surface and atmospheric conditions. The GSW algorithm used by the LSA SAF generates LST from SEVIRI/Meteosat measurements on an operational basis. The LST are freely available at http://landsaf.ipma.pt, along with estimated uncertainties based on an error propagation analysis (Freitas et al., 2010).

This document presents validation results obtained for LSA-SAF LST retrieved from SEVIRI/Meteosat (LSA-001, MLST) and AVHRR/Metop (LSA-002, ELST), hereafter referred as LST_SEVIRI and LST_AVHRR, respectively. The methodology followed for the validation of LST_SEVIRI is described in Trigo et al. (2008b), Freitas et al. (2010), and Göttsche et al. (2013). Given the wide range of validation data gathered for LST_SEVIRI, and the rather short time-series of daily composites of AVHRR/Metop LST product, the validation of LST_AVHRR is heavily based on its inter-comparison with LST_SEVIRI.

The requirements of LSA SAF products, both SEVIRI (LSA-001) and AVHRR-based (LSA-002) are summarized in Table 1.

Table 1 Product Requirements for LSA SAF LST products, LSA-001 and LSA-002, in terms of area coverage, resolution and accuracy (Product Requirements Document version 2.9, SAF/LAND/PRD/2.9).

LST Product	Coverage	Resolution		Accuracy		
		Temporal	Spatial	Threshold	Target	Optimal
MLST (LSA-001): LST_SEVIRI	MSG disk	15 min	MSG pixel resolution	4K	2 K	1K
ELST (LSA-002): LST_AVHRR	GLOBE	Daytime/ Night-time	1-km	4K	2 K	1K

In the case of LST_AVHRR (LSA-002), at any given time, only the primary satellite is processed.

This validation report is applicable to LST products generated by the LSA SAF operational chain. The detailed indication of the algorithm version used at any given time is provided in Table 2 for LST_SEVIRI (LSA-001) and in Table 3 for LST_AVHRR (LSA-002).

Table 2 Algorithm versions of MLST (LSA-001) product implemented in the LSA SAF operational chain, the respective beginning of data processing and description.

SEVIRI/MSG LST (LSA-001)		
Product Version	**Beginning of Processing**	**Description**
1.0	20060105	Pre-operational version of SEVIRI LST GSW algorithm.
5.0	20061013	First Operational version of SEVIRI LST GSW algorithm.
4.2	20071004	Product wrongly reverted to a previous version.
6.2	20080214	Updated to remove the emissivity calculations from this algorithm. The emissivity maps are now read as input files.
6.4	20080407	Correct global attributes. Now the attributes are dynamically read from input files.
7.0	20080707	Include a new dataset with the error bars associated to the LST estimations.
7.1	20090217	Processing inland waters.
7.3	20130215	Updated to handle MSG3 satellite.
7.4	20130219	Correct a bug on reading GSW coefficients from input files.
7.5	20130307	Correct a bug on reading global attributes.
7.7	20130731	Correct a bug on the handling of brightness temperatures missing values.
7.13	20151111	Adapted to a new MSG system: 1) Product generated in a single region: MSG-Disk; 2) Distribution through EUMETCast in the 4 geographical regions: NAfr, SAfr, SAme, Euro.
7.14.0	20160113 20160707	-Minor change to allow the use of MSG4 data as input. -Corrected a bug on the interpolation of forecasted atmospheric input data in the coastal regions.

Table 3 Algorithm versions of ELST (LSA-002) product implemented in the LSA SAF operational chain, the respective beginning of data processing and description.

AVHRR/Metop LST (LSA-002)		
Product Version	**Beginning of Processing**	**Description**
0.0.8	20150105	Pre-operational version of AVHRR LST GSW algorithm.

2 LST validation stations

In-situ LST obtained from ground measurements are the most conclusive and independent datasets for validating remotely sensed LST (Schneider et al., 2012). However, such validation exercises have their own limitations: (i) the low number of high-quality field data sets and the lack of global representativeness and (ii) up-scaling of the in-situ point measurements to satellite pixel size; given the high thermal heterogeneity of natural land surfaces, these issues introduce uncertainties that are not easily quantified (Göttsche et al., 2013; Ermida et al., 2014; Jimenez-Munoz et al., 2014).

In order to be able to validate satellite-derived LST products over a wide range of surface and climatic conditions, Karlsruhe Institute of Technology (KIT) set up four permanent LST validation stations in areas characterised by naturally homogeneous land use and land cover in different climate zones. The stations are part of LSA-SAF's validation effort and were chosen and designed to validate LST derived from MSG/SEVIRI, but are equally well suited to validate LST products from other sensors. Figure 1 shows the stations' locations within the field of view (FOV) of the METEOSAT satellites: Evora (Portugal, since 2005; cork-oak trees and grass), Dahra (Senegal, since 2008; tiger bush), Gobabeb (Namibia, since 2007; gravel plain), and RMZ farm / farm Heimat (Namibia, since 2009; Kalahari bush). The station Evora is in temperate Mediterranean climate (CSh), Dahra in semi-arid climate (BSh), and Gobabeb & Kalahari are in warm desert climate (BWh) climate zones, respectively (Köppen, 1936).

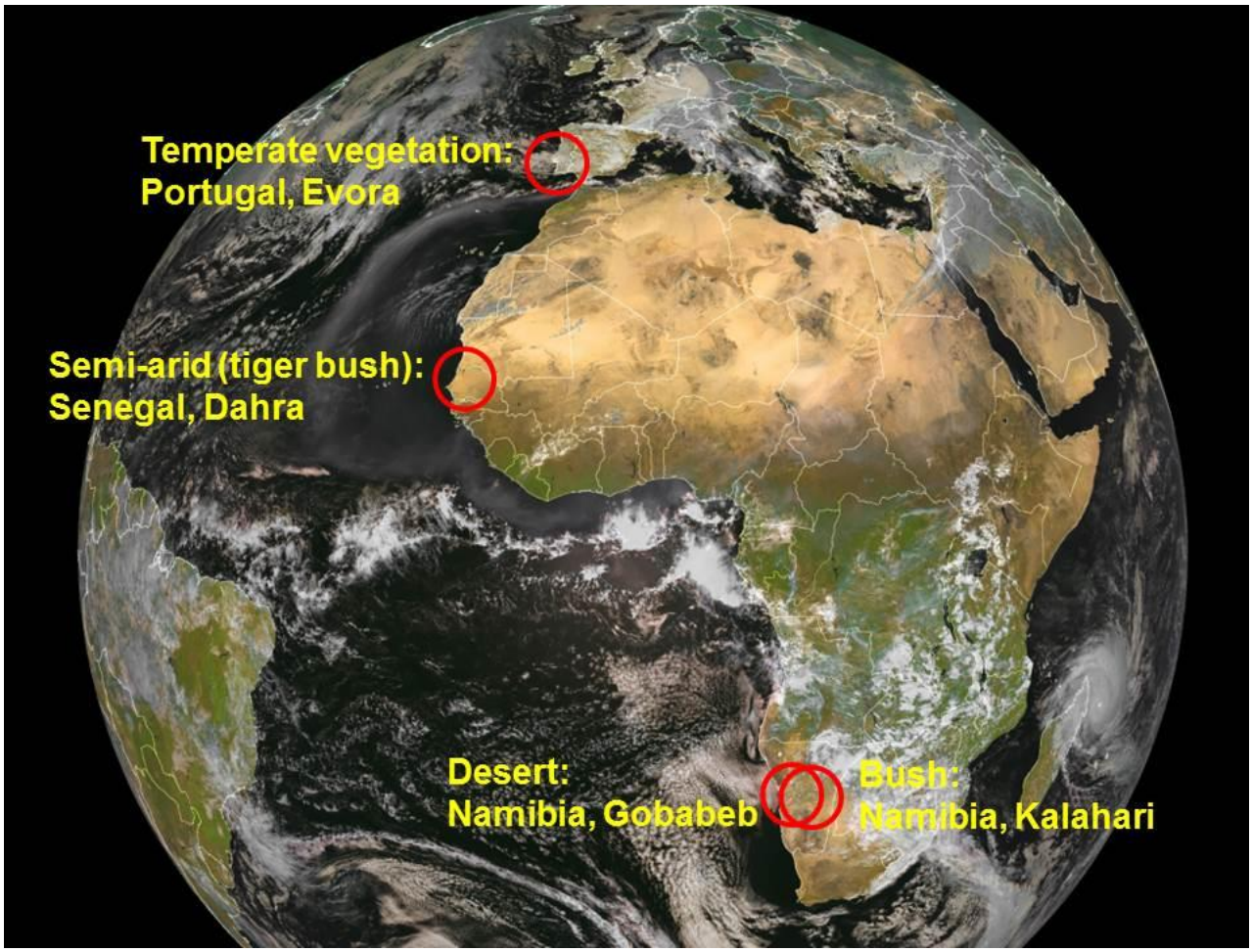


Figure 1 Locations of KIT's validation stations on MSG/SEVIRI earth disk.

Accurate estimations of land surface emissivity (LSE) and the continuous measurement of down-welling radiance are essential for the validation of satellite LST&E products, but also to limit the uncertainty of ground-based LST observations. Especially sites with larger fractions of bare ground are prone to be misrepresented in satellite-retrieved LSEs. In-situ measurements performed at Gobabeb revealed that LSE estimations over arid regions can be wrong by more than 3% (Göttsche and Hulley, 2012): this typically causes LST errors between 1°C and 2°C (Schädlich et al., 2001). In order to minimize such errors, in-situ LSEs of the dominant surface cover types at Gobabeb and Dahra were obtained using the so-called 'emissivity box method' (Rubio et al., 1997) and from emissivity spectra of soil samples (Göttsche and Hulley, 2012). However, given the high variability of LST in space and in time, the in-situ validation of LST estimated from remote sensing observations remains a challenging problem. Particularly at daytime over complex and structured land surfaces, e.g. the cork-oak tree forest at Evora station, large thermal gradients between different land surface covers and the dependence of their fractions on viewing and illumination geometry may have to be accounted for (Ermida et al., 2014; Guillevic et al., 2013).

2.1 In-situ measurements and LST determination

The main instrument for the in-situ determination of land surface temperature at KIT's validation stations is the precision radiometer 'KT15.85 IIP' produced by Heitronics GmbH, Wiesbaden, Germany. The radiometers measure thermal infra-red radiance between 9.6 µm and 11.5 µm, have a temperature resolution of 0.03 °C and an accuracy of ±0.3°C over the relevant temperature range (Theocharous et al., 2010). The KT15.85 IIP has a drift of less than 0.01% per month: the high stability is achieved by linking the radiance measurements via beam-chopping (a differential method) to internal reference temperature measurements and was confirmed by a long-term parallel run with the self-calibrating radiometer 'RotRad' from CSIRO, which was 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).

The KT15.85 IIP radiometers are mounted between 15 m and 28 m height from where they have a field of view (FOV) of about 4 m2 to 10 m2. Since the KT15.85 IIP's spectral response function lies within the atmospheric window and the distance between the radiometers and the surface is relatively small, the attenuation of the surface-leaving TIR radiation is negligible. However, the measurements of the KT15.85 IIPs observing the surface contain its emitted radiance (i.e. the target signal) as well as reflected down-welling IR radiance from the atmosphere, which needs to be corrected for: depending on target emissivity and on down-welling longwave radiance (e.g. a cold clear sky vs. a warm humid atmosphere), the reflected component can cause differences of several degrees Kelvin (Becker, 1987; Schädlich et al., 2001). Therefore, at each station an additional KT-15.85 IIP measures down-welling longwave IR radiance from the atmosphere at 53° zenith angle: measurements under that specific zenith angle are directly related to down-welling hemispherical radiance (Kondratyev, 1969) so that no ancillary data for deriving ground truth LST are needed.

In order to obtain in-situ LSTs which are representative for the land surface within the satellite pixel, we use the extrapolation methodology described by Bork-Unkelbach (2012) to scale the point-like station measurements to satellite pixel resolution. This so-called End-Member-Cover method is based on a linear spectral mixing approach and assumes that the total IR radiance emitted by the land surface within a satellite pixel can be reasonably well approximated by a linear mixture of the IR radiance emitted by the relevant surface cover types within that area. The relevant surface cover types, also called spectral end-members, can be trees, grassland or different kinds of rock or soil and are determined from an independent component analysis of high-resolution satellite data in the visible

and near infrared. The cover fractions of the relevant end-members are determined by land cover classification and then used as weights for mixing the measured radiances.

2.1.1 LST derivation from in-situ measurements

Planck's law relates the radiance emitted by a black body (emissivity ε = 1) to its temperature. However, most objects relevant to remote sensing applications are non-black bodies with 0 < ε(λ) < 1. Spectral emissivity ε(λ) is defined as the ratio between the spectral radiance Rk emitted by surface component k at wavelength λ, and the spectral radiance emitted by a black body at the same wavelength and temperature. Spectral radiance emitted by a non-black body can be obtained by multiplying Planck's function B(Tk, λ) with ε (Dash et al., 2002):

$$R_k(\mathrm{T_k}, \lambda) = \varepsilon(\lambda) \cdot \mathrm{B}(\mathrm{T_k}, \lambda) \qquad 1$$

where R_k is in W m^{-3} sr^{-1}, T_k is the measured component temperature in Kelvin, and λ is the wavelength in meters. For a sensor located near the surface and measuring within an atmospheric TIR window, the influence of the atmosphere can be neglected. With known emissivity, the simplified radiative transfer equation (Dash et al., 2002) can be used to account for reflected down-welling TIR radiance from the atmosphere and for the non-black body behaviour of the surface. Therefore, the blackbody equivalent spectral radiance B_k emitted by end-member k at temperature T_k is given by:

$$\mathrm{B_k}(\mathrm{T_k}, \lambda) = \frac{R_k(\mathrm{T_k}, \lambda) - \big(1 - \varepsilon(\lambda)\big) \cdot R_{sky}(\lambda)}{\varepsilon(\lambda)} \qquad 2$$

where R_k is the end-member's measured spectral radiance and R_{sky} is sky radiance, which in practice is measured by a dedicated KT15.85 IIP radiometer aligned at the zenith angle of 53°. Once B_k is known, inverting Planck's law gives the temperature T_k of surface component *k*. The spectral response function of the KT15.85 IIP radiometers is approximately symmetric and the Planck function as well as the spectral emissivity of natural surfaces varies slowly over the radiometers spectral range. Therefore, land surface temperature is retrieved by evaluating Planck's function at the radiometer's centre wavelength of 10.55 µm (Göttsche and Hulley, 2012).

2.2 Evora, Portugal

Evora LST validation station (38.540° N, 8.003° W, 230m a.s.l.) is located about 12 km south-west of the town of Evora in the Alentejo region, Portugal. Among several other potential European sites within Meteosat's field of view Evora was chosen for setting up an LST ground-truth site taking into account that (Dash et al., 2004): (i) large Meteosat zenith angles correspond to suboptimal conditions for LST retrievals (only SZA up to 60o are admitted in the LSA SAF LST algorithm) and should be avoided; (ii) the area around the station must be homogeneous in terms of land cover, ensuring equal temperature dispersion; (iii) mountainous regions should be avoided since heterogeneous orography causes additional geometrical distortions in the satellite images; (iv) observations should be carried out continuously, preferably over years, and thus areas with a relatively stable land cover should be preferred; (v) it is also important that the area experiences long clear-sky periods and low aerosol loads. The Evora site fulfils all of the above-mentioned criteria (Dash et al., 2004).

The dominant vegetation types at Evora station are isolated groups of evergreen oak trees (quercus ilex, quercus rotundifolia, quercus suber) and grassland, which is mainly used for grazing cattle. The cork oak trees are protected by law and, therefore, have a stable age distribution and a high

average age. About three kilometres south and east of the station there are agricultural areas whereas to the West there are plantations of fast growing eucalyptus trees. Therefore, up-scaling of station LST to MSG/SEVIRI spatial scale is only recommended for the pixel centred on the station and the one directly adjacent to the north. The climate at Evora station is warm temperate (Peel et al., 2007) with hot, dry summers, annual temperature averages between 15°C and 16°C and an average annual precipitation of 669 mm (Pereira et al., 2007). From about May to September the grass is usually desiccated so that the oak trees are the dominant green vegetation. In contrast, November to March are wet months with a rapidly developing vegetation cover: the combined effect is a strong annual amplitude of green vegetation.

2.2.1 Rotating radiometer 'RotRad'

In 2005 a suite of instruments was added to an existing FLUXNET tower of 28 m height, including a rotating radiometer 'RotRad', which has been specifically designed by the 'Commonwealth Scientific and Industrial Research Organization (CSIRO), Australia, for LST-validation (Kabsch et al., 2008). The radiometer head is able to rotate about an axis perpendicular to the viewing direction (Figure 2a), allowing the scene to be viewed at varying zenith angles and to point to two black bodies as well as to the sky. The instrument calibrates itself automatically at every circle of measurements, making use of two blackbodies; one of them is heated (42ºC), while the other is close to the environment temperature; both temperatures are measured. The sensor takes measurements within the 8-12 μm spectral range, with an expected accuracy of at least 0.2 K. The RotRad measures the brightness temperature for 3 positions on the ground (Figure 2b) with an instantaneous field of view (IFOV) of the order of 6 m, and with 2-minute periodicity. The 3 scenes on the ground correspond to (i) tree crown; (ii) grass (always sunlit during summer); and (iii) a mixture of shadow grass and tree crown. SEVIRI (and MODIS) pixels are essentially composed of these three end-members. For comparison with the satellite-derived LST, we consider the in situ surface brightness temperature, TRR_sfc, to be a weighted average of the brightness temperatures of these 3 scenes, taken within each circle of measurements. The estimation of the weights – 0.37 for 'tree spot' and 0.315 for each 'grass spot' – used in the average is based on the percent of tree crowns observed in an IKONOS image (1m-resolution), for an area surrounding the station equivalent to that of SEVIRI/Meteosat pixel. These tree crown/grass fractions are in agreement with an independent analysis performed using Landsat Thematic Mapper (TM) data from 1995, which suggests a tree cover of the order of 40% for the same region (Carreiras et al., 2006).

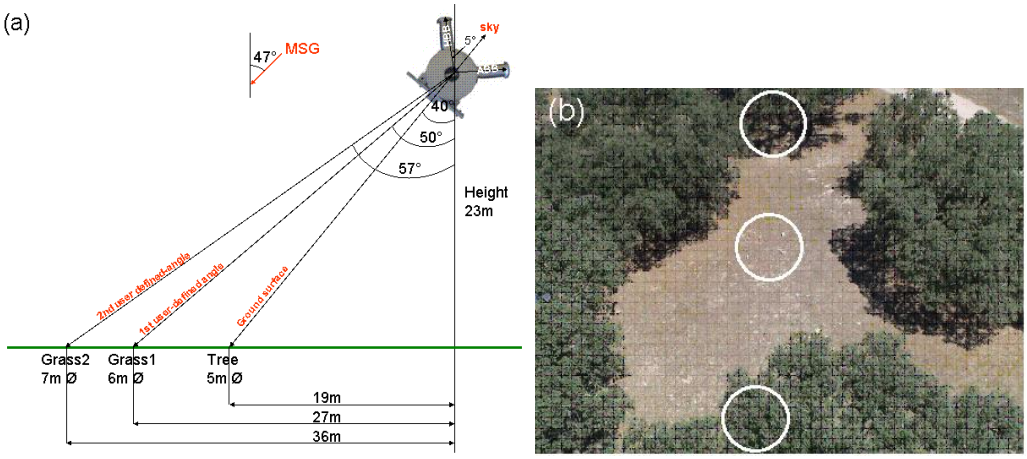


Figure 2 (a): Schematic diagram of the measurements taken by the rotating radiometer installed in the tower at Evora ground station; (b) view of the three spots on the ground corresponding to tree crown, sunlit grass and a mixture of sunlit/shadow grass.

The radiance measured by the radiometer (in the 8-12 μm band) is given by:

$$L_{RR} = L_{RR}\left(T_{RR\_sfc}\right) = \varepsilon_{RR\_sfc}\, L_{RR}\left(T_{sfc}\right) + \left(1-\varepsilon_{RR\_sfc}\right) L^{\downarrow}_{RR\_atm} \quad (1)$$

where T_{RR_sfc} and ε_{RR_sfc} are the effective brightness temperature and emissivity of a surface consisting of an ensemble of the scenes described above (Figure 2b), in the radiometer band and; T_{sfc} is the respective surface temperature; and $L^{\downarrow}_{RR\_atm}$ is the downward atmospheric radiance in the RotRad band. The latter is estimated from a fourth measurement of the radiometer during each circle, taken with the sensor facing the sky at a 40º zenith angle (close to the average atmospheric thermal path).

The in situ land surface observations at Evora ground station are obtained by resolving equation (1). The values used for surface emissivity, εRR_sfc, are approximated from the LSA SAF estimations for the region (for the SEVIRI channel centred at 10.8 μm), taking into consideration the vegetation types and respective fraction (Peres and DaCamara, 2005; Trigo et al., 2008). The impact of emissivity uncertainties on LST observation errors is discussed in section 4, where the comparison between clear sky LST retrievals (from SEVIRI and MODIS) and Evora in situ observations is analysed. The comparison is carried out for five 7-day periods between September 2005 and May 2006, when both data types (satellite and ground-based) are available.

2.2.2 LST_SEVIRI and LST_MODIS versus 'RotRad' in situ LST

The variability of LST and emissivity within the pixel is one of the major obstacles to the validation of LST satellite retrievals with ground-based instruments (Wan et al., 2002). To partially overcome this problem, the in situ data at Evora are collected from three spots on the ground, corresponding to the most relevant end members at the pixel subscale (Figure 2b). We then use a single emissivity that represents the 'soil/grass and canopy' combined scene to correct the radiometer measurements, taken as the average of sensed temperatures of 'tree crown', 'sunlit grass/soil' and 'shadow grass/crown' (Figure 2). Emissivity values, corresponding to Land-SAF estimations for the SEVIRI channel centred at 10.8 μm over an area surrounding Evora station, range from 0.9628 for the driest period in September, to 0.9684 for the greenest phase in May. The emissivity computations take into account the type and fraction of vegetation cover within each pixel, following the vegetation cover method described in (Peres and DaCamara, 2005). Emissivity error bars are estimated considering the uncertainty in the fraction of vegetation (maximum absolute errors of 0.1), emissivity variability among the different types of vegetation/bare soil within the pixels surrounding the station and the inherent uncertainty of the vegetation cover method for emissivity (discussed in Trigo et al., 2008a). The resulting uncertainties in the emissivity values are of the order of 1.1% to 1.3%.

A sensitivity analysis of the final LSTInSitu values to each uncertainty source is performed for each measurement, allowing us to characterise the observation errors associated to (i) emissivity, δLST_{ε}, (ii) the radiometer noise, $\delta RotRad$, (iii) the variability of the radiometer measurements within the 10-minute intervals, which were then averaged to get each single observation, $\delta LST_{InSituVarT}$, and (iv) the spatial variability of the RotRad measurements, $\delta LST_{InSituVarSp}$, assuming

an error in the fraction of tree crowns up to 0.1. The total uncertainty of each LST in situ observation is then given by:

$$\delta(LST_{InSitu}) = \left[(\delta LST_{\varepsilon})^2 + (\delta LST_{InSituVarT})^2 + (\delta LST_{InSituVarSp})^2 + (\delta RotRad)^2 \right]^{1/2} \quad (2)$$

Table 4 and Table 5 show the average values of LSTInSitu and LST retrievals obtained from SEVIRI and MODIS for the studied periods, corresponding to a total sample of 8 (16) cases around the MODIS daytime (night-time) passage over Evora; the time elapsed between in situ observations and satellite retrievals is within ±7 minutes for SEVIRI and ±2 minutes for MODIS. The availability of data for the comparison of satellite versus 'in situ' data is subject to the existence of the four ground measurements – 'tree crown', 'sunlit grass/soil', 'shadow grass/crown' and sky brightness temperature – and (clear sky) LST retrievals from both sensors collocated with the station. Despite the existence of systematic differences, the satellite retrievals follow quite well the in situ measurements (Figure 3). Night-time estimations tend to be colder than ground observations, while daytime SEVIRI LST tends to be warmer.

As suggested by the analysis of the comparison between MODIS and SEVIRI LSTs (Trigo et al., 2008b), the sun-satellite viewing geometry does influence the retrievals. Although not shown, when MODIS and ground values are compared taking into account the MODIS zenith angle, we obtain averaged differences 'satellite minus in situ' of –2.8ºC (–0.2ºC) for positive (negative) angles, i.e. for scenes viewed from West (East). Accordingly, SEVIRI LST generally presents a warm bias. The variability of satellite – in situ discrepancies within the whole studied period is also higher for daytime values. The root mean square differences (RMSD) between MODIS and in situ LST are within the 0.0 to 3.7ºC range. The RMSD for SEVIRI estimations are higher, with values varying from 1.6 ºC in November, to 4.9ºC in September (Table 4). It is worth mentioning that 1 (out of 2) SEVIRI LST value within the latter period overestimate the ground observations by about 7ºC (Figure 2a); in this particular case, the TOA brightness temperature of SEVIRI channel centred at 10.8 μm also exceeds the ground observations by nearly 3ºC. We cannot fully understand the largest discrepancies between SEVIRI and in situ observations obtained for September. However, it should be kept in mind that this is the driest period under study, when the temperature (and emissivity) contrasts between the canopy and the ground are more pronounced, the morning heating rate is highest, and thus, the uncertainty of the observations is largest.

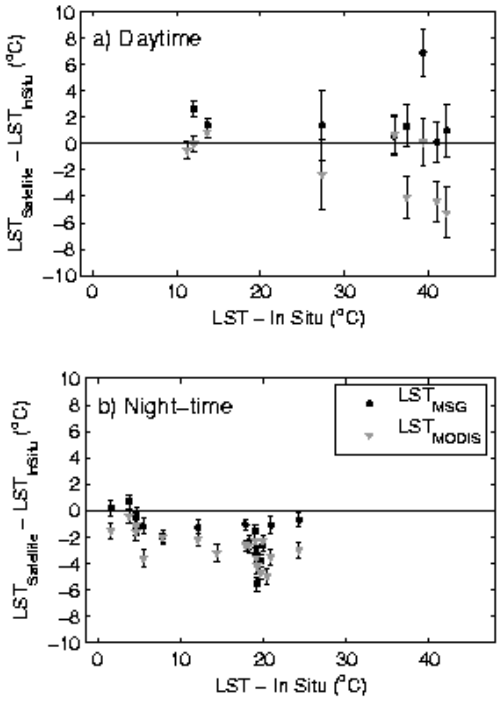


Figure 3 Differences between LST satellite retrievals and ground measurements (ºC) as a function of in situ observations taken in Evora, Southern Portugal. Black dots and grey triangles correspond to SEVIRI and MODIS LST, respectively. The error bars represent the estimated uncertainty of each in situ observation.

Table 4 Daytime statistics per study period, including: mean LST observation (°C) and respective mean uncertainty (ΔLST; °C); SEVIRI and MODIS average LST (°C) and root mean square difference against the observations (RMSD; °C). The 1st column shows the 7-day periods under study and the respective number of cases available.

Period	OBS		SEVIRI		MODIS	
(no. obs)	LST	ΔLST	LST	RMSD	LST	RMSD
14-20 Sep 05 (2)	40.8	1.85	44.7	4.9	38.2	3.7
11-17 Nov 05 (1)	13.7	0.44	15.1	1.4	14.5	0.9
23-31 Jan 06 (1)	12.1	0.62	14.7	2.7	12.1	0.0
23-29 May 06 (4)	35.4	1.82	36.3	1.0	32.9	3.2

Night-time LST retrievals from SEVIRI and MODIS are consistently below in situ observations throughout the whole studied period (Table 5 and Figure 3). Such cold bias ranges from 0.4 to 2.5ºC and 1.6 to.3.3ºC for SEVIRI and MODIS, respectively. In contrast with Evora ground observations obtained for daytime MODIS passages, night-time values have relatively low uncertainties associated (of the order of 0.5ºC). These are essentially associated with emissivity uncertainties, particularly in November and January when the temporal and spatial variability of in situ observations are lowest.

Table 5: As in Table 4, but for night-time.

Period	OBS		SEVIRI		MODIS	
(no. obs)	LST	ΔLST	LST	RMSD	LST	RMSD
14-20 Sep 05 (7)	18.2	0.52	15.7	2.9	15.2	3.1
11-17 Nov 05 (1)	7.8	0.50	5.9	1.9	5.8	2.0
23-31 Jan 06 (5)	4.0	0.59	3.6	0.9	2.4	2.0
23-29 May 06 (3)	21.0	0.51	19.0	2.4	17.7	3.5

2.2.3 In-situ validation at Evora from 2009 onwards

In-situ measurements at Evora started in 2005, but the set-up at the current location and orientation of radiometers is in operation since March 2009. It is worth mentioning that these results were obtained after the changes in level 1.5 SEVIRI radiance definition in May 2005 took place, which were shown (Barroso et al., 2008) to have a rather low impact on derived LST (generally less than 0.5 K). However, there is a tendency for high LST (> 30ºC) to become cooler and colder LST values to become warmer, leading to lower LST amplitudes during the warm season. At the end of 2007 the FLUXNET tower at Evora was dismantled and the instruments for LST validation had to be moved to a KIT owned mast at a nearby location. Figure 4 shows the set-up at Evora before October 2011, when the two masts were replaced by a single larger mast at the same location.

In order to determine the relevant end-members of the surface in the validation area, Bork-Unkelbach (2012) performed an independent component analysis (ICA) of a high-resolution multispectral IKONOS dataset covering the area of the MSG/SEVIRI pixel located over the validation station. Bork-Unkelbach identified trees and grassland as the two relevant end members. Streets, buildings and lakes were also identified as end-members, but can be neglected due to their small fractional coverage. The fractional coverage of the validation area with cover types trees and grasslands was then determined using an object-based image analysis approach, yielding fractions of 32% for trees and 68% for grasslands for the MSG/SEVIRI pixel under investigation. The results were verified in a quality assessment and agree well with a previous study (Carreiras et al., 2006) about the tree crown cover (TCC) fraction in the validation area. Bare ground is usually not observed at Evora validation station: even in August the surface is covered by dry and relatively high grass, which increases effective emissivity via the cavity effect (French et al., 2000; Olioso et al., 2007). The end-members observed by the Heitronics KT15.85 IIP radiometers are grass and tree crown. The emissivities of the two observed end-members are set to the values retrieved operationally by LSA-SAF: over the course of the year LSA-SAF emissivities for MSG/SEVIRI channel 9 vary between about 0.96 (August; desiccated grass) and 0.98 (April; green grass). These values are in good agreement with literature values (French et al., 2000; Olioso et al., 2007; Salisbury and D'Aria, 1992). Furthermore, a 9 minute delay between actual satellite acquisition time for Evora (SEVIRI scans the Earth from South to North) and nominal product time has been accounted for and in-situ LST and LSA SAF LST were matched to better than 1 minute.

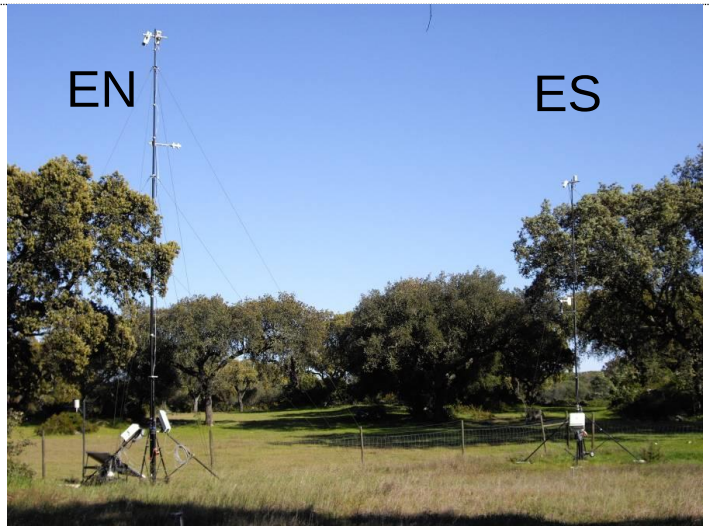


Figure 4: Evora LST validation station, Portugal. Radiometers on the two masts (Evora North and Evora South) measure brightness temperatures of tree crown, grass (2x), and sky.

2.2.4 Validation results for fixed end-member fractions

Figure 5, Figure 6 and Figure 7 show plots of LSA SAF LST derived from MSG/SEVIRI against Evora station LST for May, August, and November 2011, respectively. The three months represent Spring, Summer and Autumn at Evora and usually have a high number of clear sky situations. For May 2011 (Figure 5) there are 1154 match-ups between satellite and in-situ LST and bias and rmse are 0.16°C (satellite slightly warmer) and 1.32°C, respectively, i.e. LSA SAF meets its LST target accuracy of 2°C. However, at low LST (night-time) LSA SAF LST is systematically higher than in-situ LST.

For August 2011 (Figure 6) there are 966 match-ups between satellite and in-situ LST and bias and rmse are 1.0°C and 2.0°C, respectively, which means that LSA SAF (just) meets its target

accuracy of 2°C. LSA SAF LST are systematically higher than in-situ LST at high LST (above 30°C), which is the main cause of the overall larger bias and rmse.

For November 2011 (Figure 7) there are 891 match-ups between satellite and in-situ LST and bias and rmse are 0.8°C (satellite slightly warmer) and 1.3°C, respectively, and LSA SAF meets its target accuracy of 2°C.

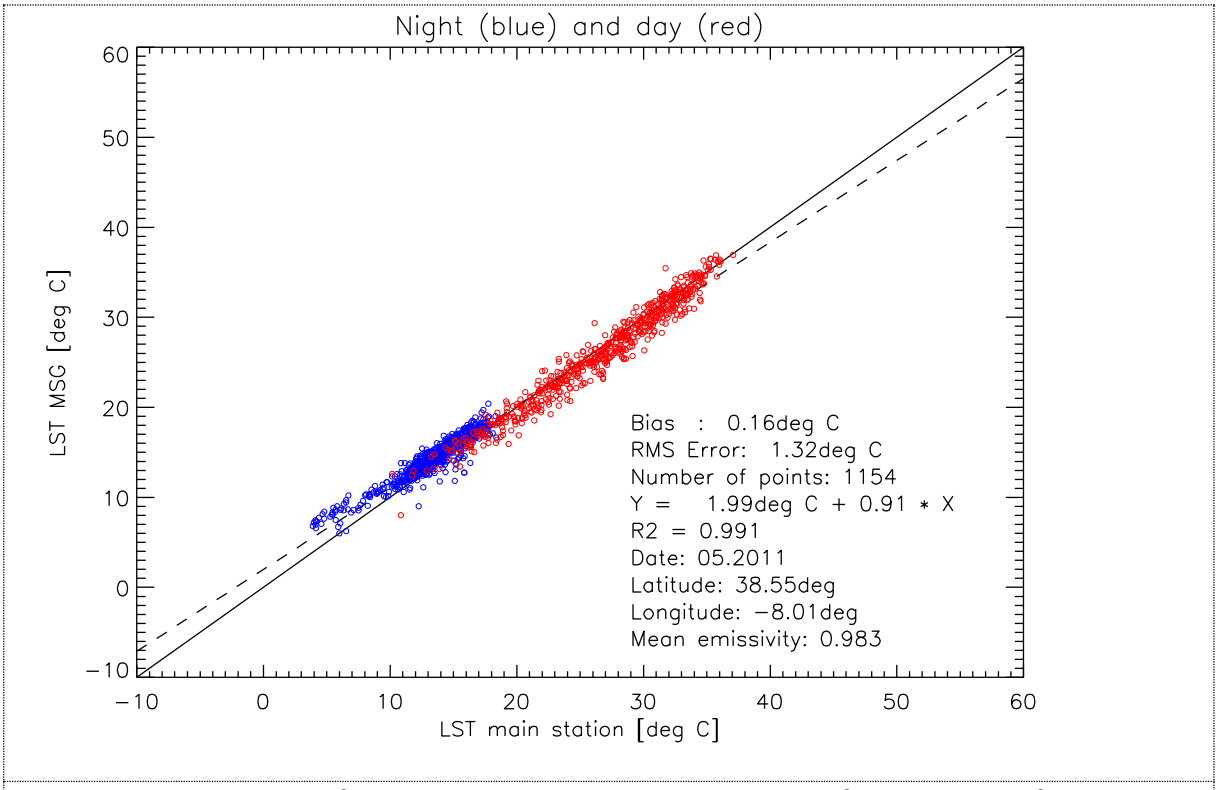


Figure 5 LSA SAF LST from MSG against Evora in-situ LST for May 2011 for night-time (blue circles) and daytime (red circles).

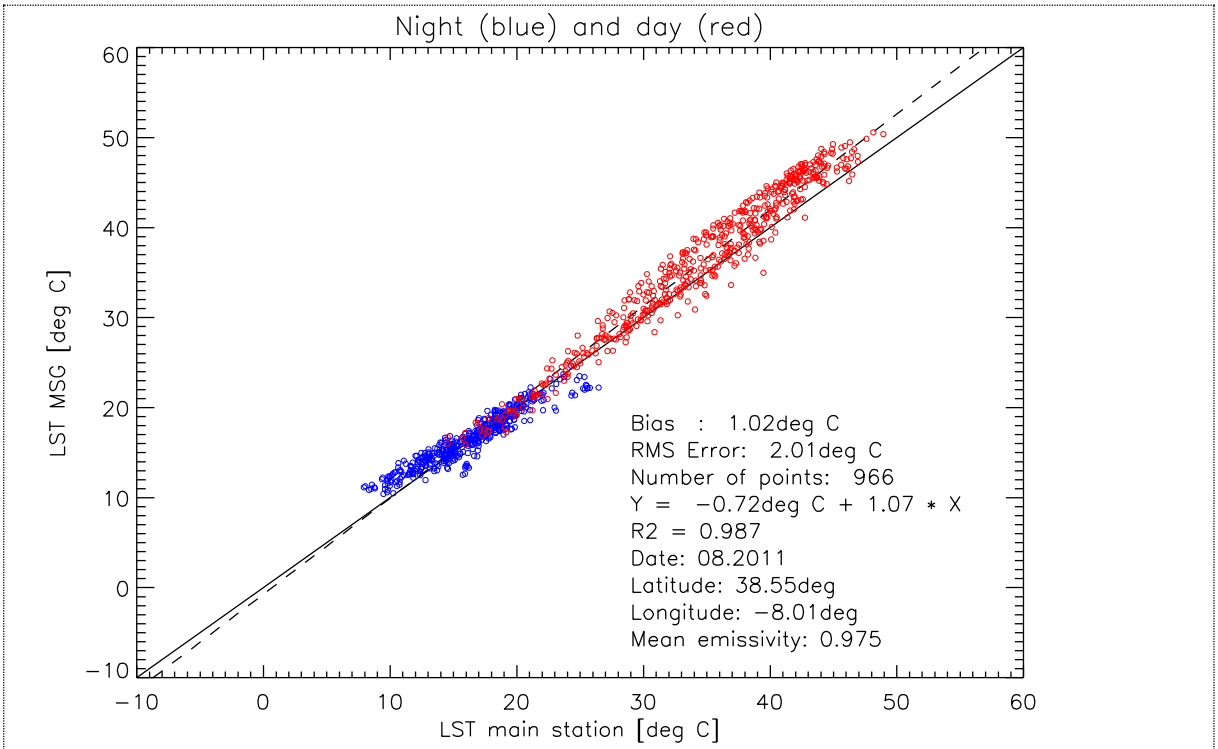


Figure 6 LSA SAF LST against Evora in-situ LST for August 2011 for night-time (blue circles) and daytime (red circles).

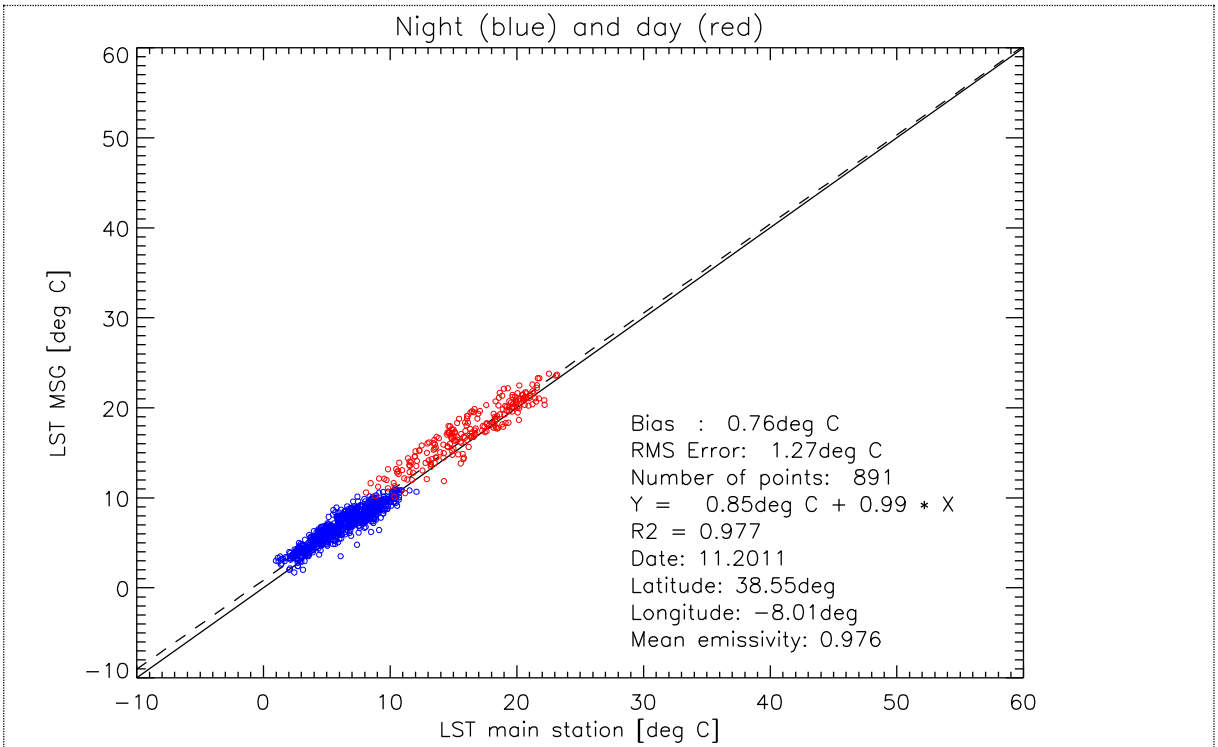


Figure 7 LSA SAF LST against Evora in-situ LST for November 2011 for night-time (blue circles) and daytime (red circles).

In order to validate LSA SAF's LST algorithm with several years of in-situ data (April 2009 – October 2012) while still being able to assess possible variations (e.g. seasonal) in its performance, Figure 8 displays monthly biases & rmse together with the corresponding numbers of valid data points, i.e. monthly LST match-ups, for all available data (night and day). For the shown period of time the mean bias is 0.6°C, mean stdev 1.5°C, and mean rmse 1.9°C. For 2009 and 2010 the number of match-ups appears to be seasonal (most match-ups in summer), while this behavior is considerably less pronounced in 2011 and 2012. No obvious seasonality of bias and rmse is observed. Between May and September 2012 there is sudden change to a negative bias of about -2K while rmse increases. This is thought to be caused by cattle breaking through the fence around Evora station, grazing within the FOV of the KT15.85 IIP radiometer, and leaving behind short, dry grass with more visible background soil: this caused an increase of in-situ LST, making it unrepresentative of the wider area around Evora station.

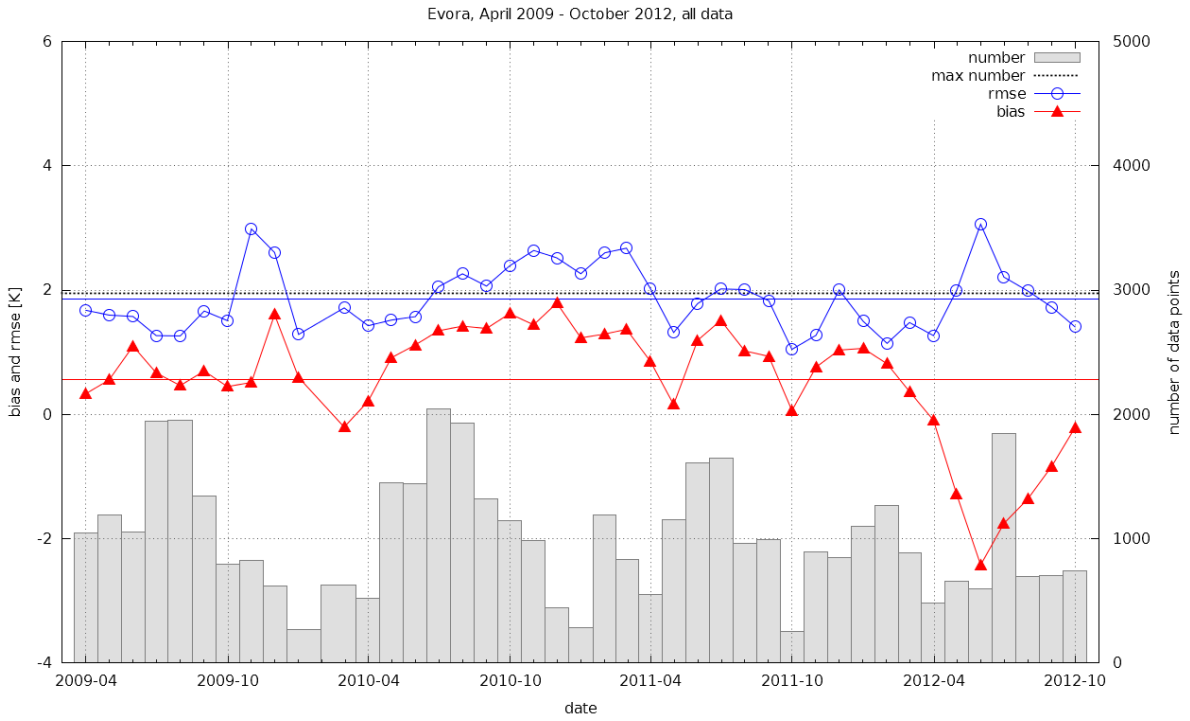


Figure 8 Monthly statistics at Evora station, Portugal, for daytime and night-time LSA SAF LST. Mean bias (red triangles) and rmse (blue circles) refer to the left y-axis, the number of match-ups (grey bars) to the right y-axis.

2.2.5 Modelling of projected end-member fractions

Remotely sensed LST is a directional variable, unless some sort of compositing of observations from different viewing angles is performed (Ermida et al., 2014). Especially over regions with sparse canopies there are often strong temperature differences between sunlit background, shaded background and tree crowns. Therefore, LST retrievals obtained for the same scene, using the same sensor, but at different viewing angles would likely produce different temperature values. Ermida et al. (2014) used a geometrical model to estimate the projected areas of the end-members at Evora station, using parallel-ray geometry to describe the illumination of a three-dimensional vegetation element and the shadow it casts. The model is an effective means for the

correction of LST differences between sensors associated with their viewing geometries as well for matching in-situ LST to the viewing geometry of a satellite.

At Evora the temperature of shaded background is not directly measured, but it is generally very close to air or tree crown temperature (Guillevic et al., 2013). Therefore, Ermida et al. (2014) estimate shaded background temperature from air temperature and correct its time lag w.r.t. ground temperature. Since there is no direct incoming solar radiation on shaded surfaces, we expect their temperature to be close to radiative equilibrium with air (Figure 9).

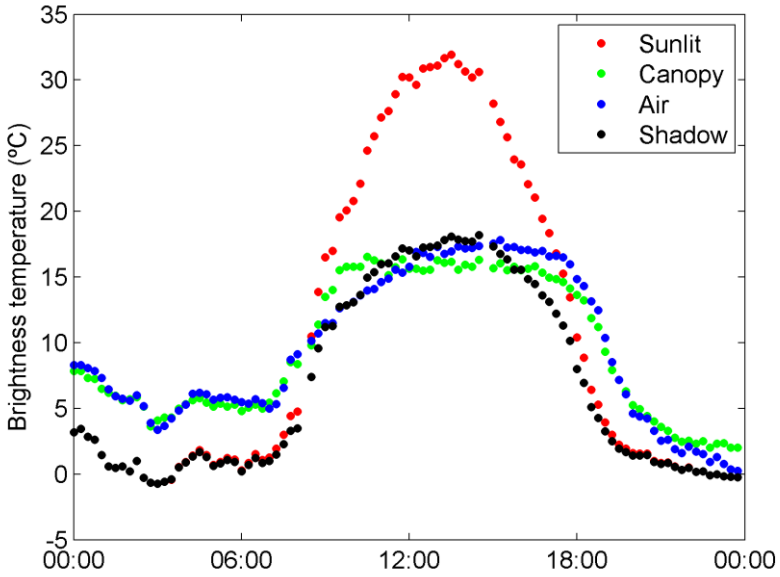


Figure 9: Diurnal cycle of Air, Canopy, and Sunlit ground temperatures (°C) and estimated shaded surface temperature at Évora, 20th of March 2011. From Ermida et al. (2014).

The appropriate ground-based LST to validate values retrieved from any given space-borne sensor (e.g. SEVIRI or MODIS) is obtained by compositing the measured in situ component temperatures (e.g. sunlit background, shaded background and tree canopy), giving different weights according to the different fractions seen by the sensor. This is achieved by means of a geometric model that takes into account the viewing and illumination conditions, the Geometrical-Optical (GO) part of the Geometrical-Optical Radiative Transfer (GORT) model (Ni et al., 1999). The sunlit and shaded parts of canopy should in principle also be treated separately (Jones and Vaughan, 2010), but differences between the two parts are negligible when compared to the differences between sunlit background, shaded background and canopy. Previous works by Pinheiro et al. (2006) and Guillevic et al. (2013) have shown that the above mentioned three components suffice to capture the scene angular variability. Therefore, the pixel's radiance measured by a sensor can be estimated as a linear combination of the radiances emitted by each of the scene components weighted by their respective projected scene fractions (Pinheiro et al., 2006):

$$L_{avg} = F_{sunlit} * L_{sunlit} + F_{shadow} * L_{shadow} + F_{canopy} * L_{canopy} L_{\text{avg}} = f_{\text{sunlit}} * L_{\text{sunlit}} + f_{\text{shadow}} * L_{\text{shadow}} + f_{\text{canopy}} * L_{\text{canopy}} \quad (4)$$

where $L_{avg}\, L_{\text{avg}}$ is the pixel's radiance within a sensor FOV, $L_{sunlit} L_{\text{sunlit}}$, $L_{shadow} L_{\text{shadow}}$ and $L_{canopy}\, L_{\text{canopy}}$ are sunlit background, shaded background and canopy radiances, respectively, and $F_{sunlit}\, L_{\text{sunlit}}$, $F_{shadow} L_{\text{shadow}}$ and F_{canopy} are the corresponding component fractions, as seen by each space-

borne sensor. Sunlit and tree canopy radiances are obtained from *in situ* measurements of brightness temperature using Planck's Law. Here we use a representative wavelength of 10.55 µm 10.55 $\mu m$for channel-effective emissivities for the KT-15.85 IIP radiometer band (Göttsche et al., 2013). The radiance from the shaded background L_{shadow} is obtained as described in Ermida et al. (2014).

2.2.6 Validation results for projected end-member fractions

Composited values of surface temperature obtained with the geometric model of Ermida et al. (2012) were then used to assess MSG and MODIS LST products. The comparison between satellite and ground observations is performed for pixels closest to the Évora site and using the respective sensor viewing geometry to set up the appropriate composite in situ temperature. For both sensors, SEVIRI and MODIS, the composite temperature is calculated using the effective emissivity obtained with the vegetation cover fraction method (Freitas et al., 2010), yielding values between 0.9691 for the driest period in September, and 0.9773 for the greenest phase in April. It is assumed that this range reflects well the seasonal variability between dry and green understory that characterize the region. However, it is acknowledged that emissivity uncertainties may be an important source of error for the in situ composite temperatures. For reference, we also show the comparison between satellite LST and ground composites following a procedure where neither the daily and seasonal variations in the illumination geometry, nor the actual sensor viewing angles are taken into account, i.e. with fixed land cover fractions as in section 2.2.4 and in Trigo et al. (2008a). Figure 10 presents scatterplots of satellite LST versus in situ temperature values obtained using the geometric model (lower panels d, c) and using the above mentioned weighted temperature average where the effects of viewing and illumination geometry are not taken into account (upper panels a, b).

MODSW LST (i.e. MOD11A1 and MYD11A1 LST products) and SEVIRI-derived LST are considerably closer to in situ composites obtained with the model (Figure 10c and Figure 10d), which demonstrates the need to consider the directional character of LST products. This is further confirmed by the corresponding statistics shown in Table 6: taking all LST satellite products together, the daytime absolute bias (i.e. average of satellite LST minus in situ LST) and root mean square error (RMSE) decrease by 1.5 to 2.5°C when the viewing and illumination geometries of the scene are considered. The standard deviation of the difference between MODSW LST and in situ daytime temperature decreases by about 1.2 °C when the model is applied, whereas the corresponding decrease for SEVIRI and MODTES daytime LST temperatures is smaller. This can be explained by the fact that SEVIRI observes a scene from a fixed perspective and the limitation of the MODTES algorithm to view angles ≤40° (Hulley et al., 2011). As expected, the impact of the geometric correction on the night-time statistics is very small, while the large improvements at daytime considerably impact the overall statistics obtained for day and night-time data ("TOTAL" line in Table 6).

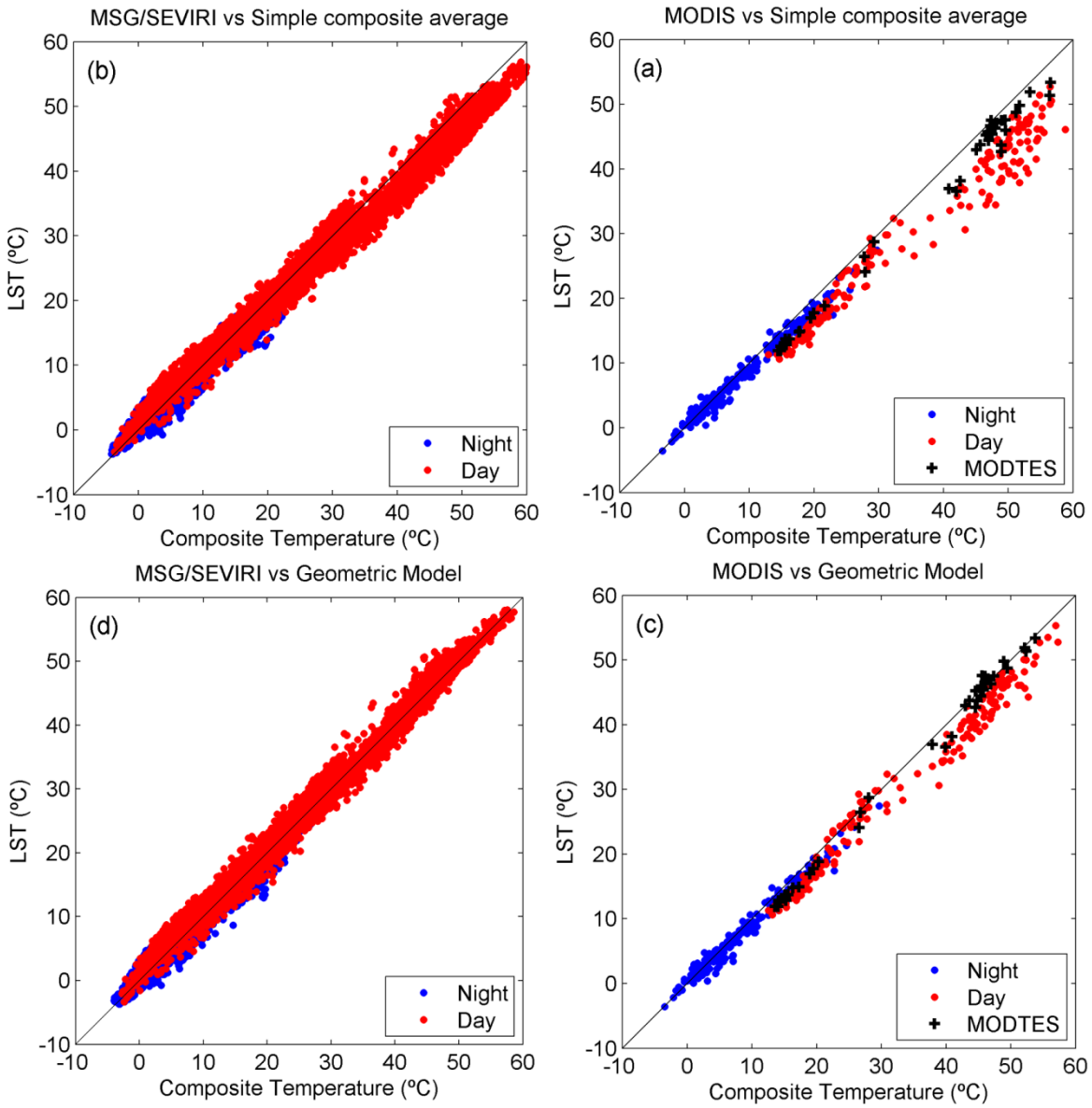


Figure 10 –Scatterplots of LST (°C) derived from MODIS (a, c) and MSG/SEVIRI (b, d) against in-situ composite LST. In (c, d) In-situ composite LST was obtained using the geometric model whereas in (a, b) fixed fractions of surface elements were used. From Ermida et al. (2014).

Overall MSG shows a better agreement with in situ observations than MODSW (i.e., MOD11A1 and MYD11A1), presenting a lower RMSE, error STD and bias for both daytime and night-time values (Table 6). MODSW LST tends to be cooler than composite temperature, keeping a bias of about -2.7°C (-0.7°C) for daytime (night-time) passages. In contrast, when the model is considered the biases of daytime SEVIRI and MODTES LST values (about +0.5 and -0.8°C, respectively) are close to the uncertainty of in situ temperatures; RMSE are of the order of 1.5°C in both (daytime) cases. These results are not in agreement with the recent work by Guillevic et al. (2013), where MODSW gridded LST (Collection 5) data are compared with in situ measurements taken in Évora; in that study, the application of a geometric model to upscale Évora measurements to MODIS observations leads to a negligible bias of satellite retrievals with respect to the in situ estimations. In the geometric model used by Guillevic et al. (2013), the area surrounding the station is populated with trees (similar percent to that used here) with crowns simulated as spheres of radius

6m; Ermida et al. (2014) suggest that 6m tree crowns may be oversized and that the traditional pruning of the trees leads to a shape resembling an ellipsoid rather than a sphere. The reasonable agreement between SEVIRI and MODTES supports the proposed approach where LST is estimated based on in situ temperature composites using a geometric model relying on a small number of parameters. Moreover, the statistics shown in Table 6 indicate that MODTES and MSG (daily) LST estimates are fairly close, and discrepancies in the RMSE, standard deviation or bias are within the uncertainty of the in situ estimations (ranging between 0.5 and 1.5°C). However, this is not the case when we compare the bias or RMSE of MODSW with those of MODTES or MSG.

Table 6 – Root Mean Square Error (RMSE), error Standard Deviation (STD) and bias for LST versus *in situ* composite temperature (°C) using the model (**bold**) and using the composite with fixed fractions of surface elements (*italics*). The values in parentheses correspond to the validation of MSG only using data for which MODIS observations are also available. (Ermida et al., 2014)

	RMSE		STD		BIAS	
	"simple composite"	modeled	"simple composite"	modeled	"simple composite"	modeled
	MODSW					
Daytime	*5.89*	**3.24**	*3.05*	**1.85**	*-5.04*	**-2.66**
Night-time	*1.34*	**1.35**	*1.19*	**1.17**	*-0.63*	**-0.68**
Total	*4.02*	**2.37**	*3.11*	**1.80**	*-2.56*	**-1.54**
	MODTES					
Daytime	*2.87*	**1.48**	*1.38*	**1.25**	*-2.53*	**-0.81**
	MSG					
Daytime	*2.48 (2.49)*	**1.50 (1.51)**	*2.17 (2.16)*	**1.42 (1.33)**	*-1.19 (-1.25)*	**0.50 (0.72)**
Night-time	*1.27 (1.21)*	**1.19 (1.21)**	*1.27 (1.22)*	**1.19 (1.21)**	*-0.05 (-0.04)*	**0.06 (-0.08)**
Total	*1.90 (1.88)*	**1.34 (1.35)**	*1.82 (1.80)*	**1.31 (1.32)**	*-0.55 (-0.57)*	**0.26 (0.27)**

2.2.7 Satellite inter-comparison

Here we use the Ermida et al. (2014) model to compare temporally matched MSG LST against MODSW LST (i.e., MOD11A1 and MYD11A1) over Évora station. The dependence of daytime LST differences between MSG and MODSW on MODIS viewing geometry has been analyzed by Trigo et al. (2008a) and Guillevic et al. (2013). In order to further understand this effect, Figure 11 displays the discrepancies between the two satellite products as a function of MODIS zenith and azimuth angles, per season. As pointed out by Trigo et al. (2008a) and by Guillevic et al. (2013), SEVIRI/MSG LST values are generally warmer than the corresponding MODIS estimates, with larger discrepancies obtained for larger MODIS view zenith angles. The results also suggest a clear seasonal variability with the highest differences being obtained during summer and, to a lesser extent, during (late) spring.

MODIS AQUA observations are performed approximately between 14 and 15 UTC and correspond to the set of values nearly aligned with MSG orthogonal plane (Figure 11). TERRA observations in turn are performed around 11-12 UTC. In both cases, high MODIS – SEVIRI/MSG discrepancies occur for extreme MODIS zenith angles (about 60°), when the uncertainties associated to the atmospheric correction are higher (see also Trigo et al., 2008a; Guillevic et al., 2013). MODSW LST coolest values with respect to MSG tend to be obtained for MODIS azimuth angles favoring the

viewing of shaded surfaces, i.e., around 285° in the case of TERRA and 75° in the case of AQUA. In contrast, observations with an azimuth angle of about 250° in the case of AQUA and 100° in the case of TERRA will be affected by an opposite effect which results in smaller LST differences. Only daytime LST differences exhibit a strong dependence on MODIS view zenith angle, because temperature contrasts at night-time contrasts are usually much smaller. The LST differences are within the range of those that were obtained using the model.

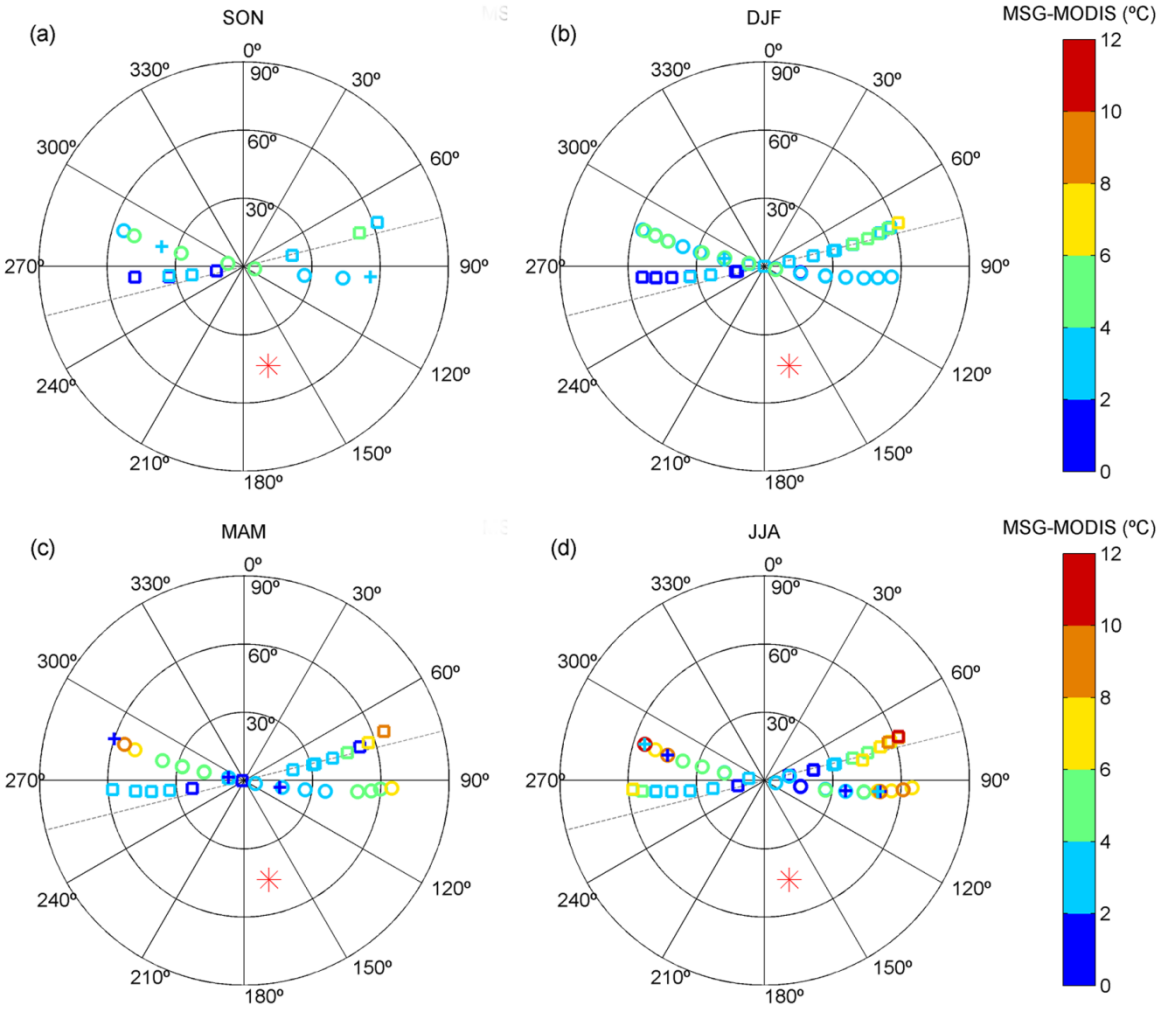


Figure 11 – Differences of daytime LST (MSG minus MODIS) in °C (colorbar) as a function of MODIS viewing geometry, for (a) autumn, (b) winter, (c) spring and (d) summer. The zenith angle is represented by the distance to the center and the azimuth angle is represented by the (clockwise) angle with respect to the vertical diameter of each panel. The circles refer to MODSW/TERRA (MOD11 product), the squares refer to MODSW/AQUA (MYD11 product) and the crosses to MODTES. The red star indicates the MSG viewing geometry at the Évora site. The grey dashed line represents the MSG orthogonal plane. (Ermida et al., 2014)

Using the Évora *in situ* measurements, the developed model also allows calculating the expected deviations from MSG associated to a change in view zenith and azimuth angles. Figure 12 shows the impact of correcting MODIS LST using the estimated deviations related to viewing geometry. This correction results in a significant reduction of the differences between the two LST products (Table 7). Because the LST differences depend on viewing geometry, which is variable in the case of MODIS observations, this correction leads to a reduction in the error standard deviation.

However, there is still a considerable bias, which indicates a systematic source of error other than differences in viewing geometry. As shown in Figure 12, the dispersion in at higher temperatures is significantly reduced, in agreement with the reduction in error standard deviation (Table 7).

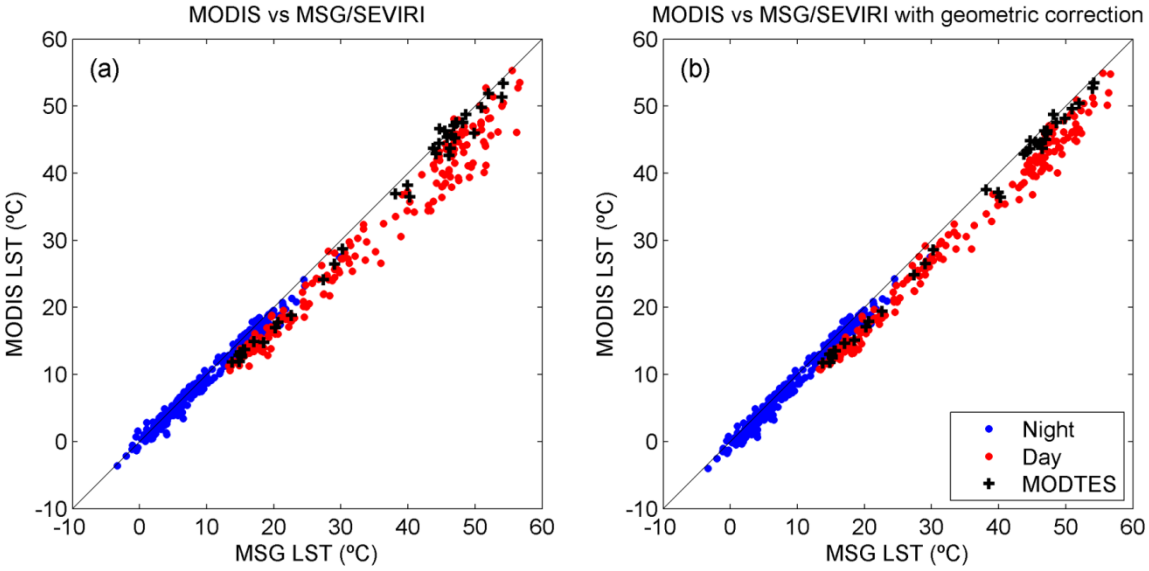


Figure 12 – MODIS LST versus MSG/SEVIRI LST before (a) and after (b) using the model to remove differences related to the viewing geometry. Blue dots indicate night-time MODSW measurements whereas red dots respect to daytime MODSW observations. The black crosses represent daytime MODTES LST (Ermida et al., 2014).

Table 7 –Root Mean Square Error (RMSE), error Standard Deviation (STD) and bias for the LST difference (MSG minus MODIS) in °C, before correcting for angular effects (*italics*) and after correcting with the geometric model (**bold**). From Ermida et al. (2014).

	RMSE		STD		BIAS	
	Without correction	With correction	Without correction	With correction	Without correction	With correction
	MODSW					
Daytime	*4.46*	**3.83**	*2.34*	**1.53**	*3.80*	**3.52**
Night-time	*1.04*	**1.14**	*0.87*	**0.87**	*0.56*	**0.75**
Total	*3.05*	**2.68**	*2.32*	**1.83**	*1.98*	**1.96**
	MODTES					
Daytime	*2.15*	**2.03**	*1.45*	**0.97**	*1.60*	**1.79**

Using the model to estimate the in-situ LSTs corresponding to the FOVs of SEVIRI and MODIS sensors reduces the bias of SEVIRI and MODSW daytime LST values by 1 to 2.5 °C compared to not taking viewing geometry and shadowing effects into account. When the model is used to account for the different viewing geometries of MODIS and MSG there is a significant reduction in LST differences between the two sensors.

However, the comparison between geometry adjusted SEVIRI/MSG LST and MODSW gridded products (MOD11A1 and MYD11A1 Collection 5) shows a systematically higher temperature of the former of about 3.5°C for daytime (0.8°C for night-time) observations. These results may be partly explained by the differences between MODSW and LSA-SAF emissivities,

which over the study period varied between 0.005 and 0.01, the values of MODSW emissivity being always higher.

2.3 Gobabeb, Namibia

Gobabeb LST validation station (23.551° S, 15.051° E, 450m a.s.l.) lies about 2 km north-east of Gobabeb Training & Research Centre (www.gobabebtrc.org) in the Namib Desert, Namibia. The validation station is located on large gravel plains (several thousand km2), which are covered by a highly homogeneous mixture of gravel, sand and sparse desiccated grass (Figure 13). There is a sharp transition between the vast Namib sand sea with its up to 300 m high dunes and the gravel plains: this natural boundary is maintained by irregular flows of the ephemeral Kuiseb River (a few days every other year), which wash the advancing sand into the South Atlantic Ocean. Due to the hyper-arid desert climate (Köppen, 1936; Peel et al, 2007), the site is spatially and temporally highly stable and, therefore, ideal for long-term validation studies of satellite products (Hulley et al., 2009). The long-term average annual temperature at Gobabeb is 21.1°C (Lancaster et al., 1984) whereas the average annual precipitation is less than 100 mm (Eckardt et al., 2013) and highly variable (Peel et al., 2007). Consequently, the relatively frequent fog events are of special importance for the water balance of the Namib (Eckardt et al., 2013). Continuous in-situ measurements from Gobabeb are available since the beginning of 2008 and the homogeneity of the site was investigated with several field campaigns.

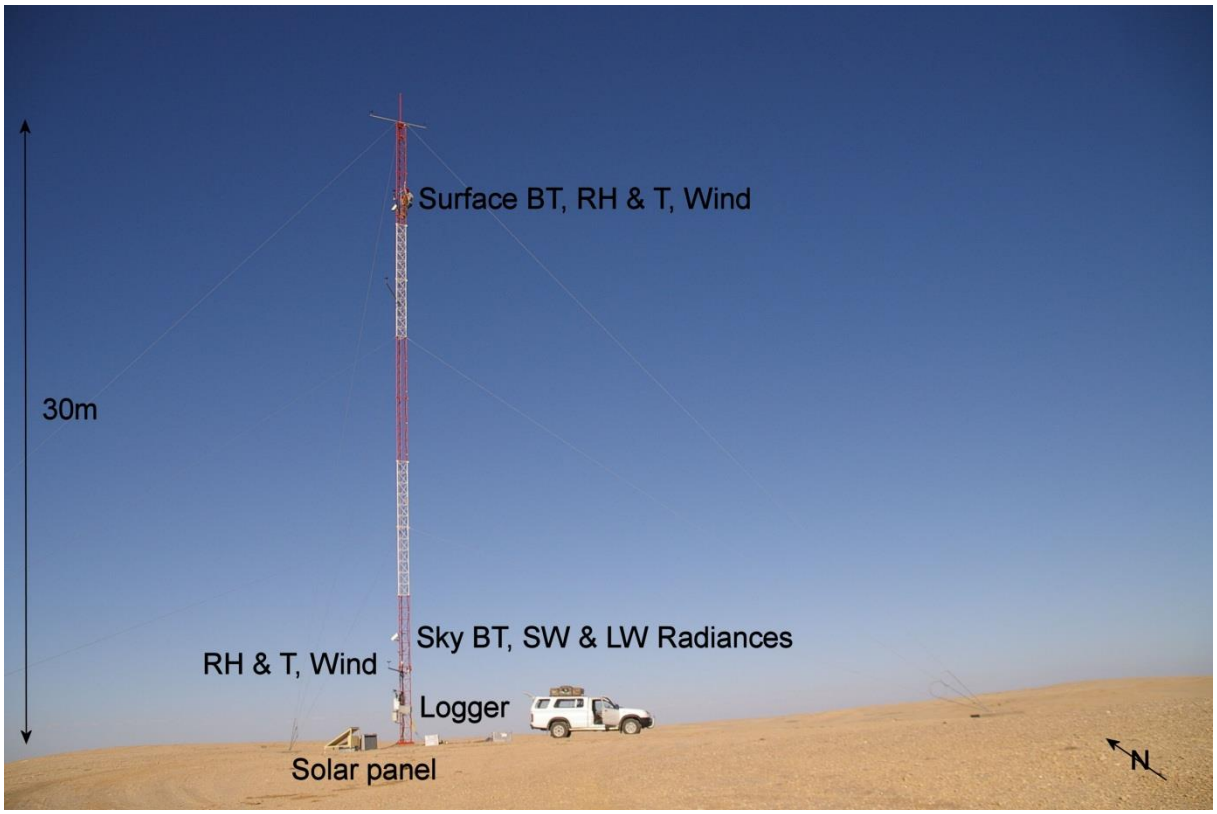


Figure 13: Gobabeb LST validation station ‘GBB Wind’ at Gobabeb, Namibia. The station is located on highly homogeneous gravel plains.

The pointing of the radiometers to the assumed surface end-members has not been changed since the setup of the station in December 2007. A view direction of the ground measurements close

to North was chosen to observe an undisturbed surface: this implies that the radiometers are not aligned with MSG's line of sight to the validation site. The radiometers are mounted at 25m height and observe neighbouring surface areas of about 13m² each under an angle of 30° (Figure 14).

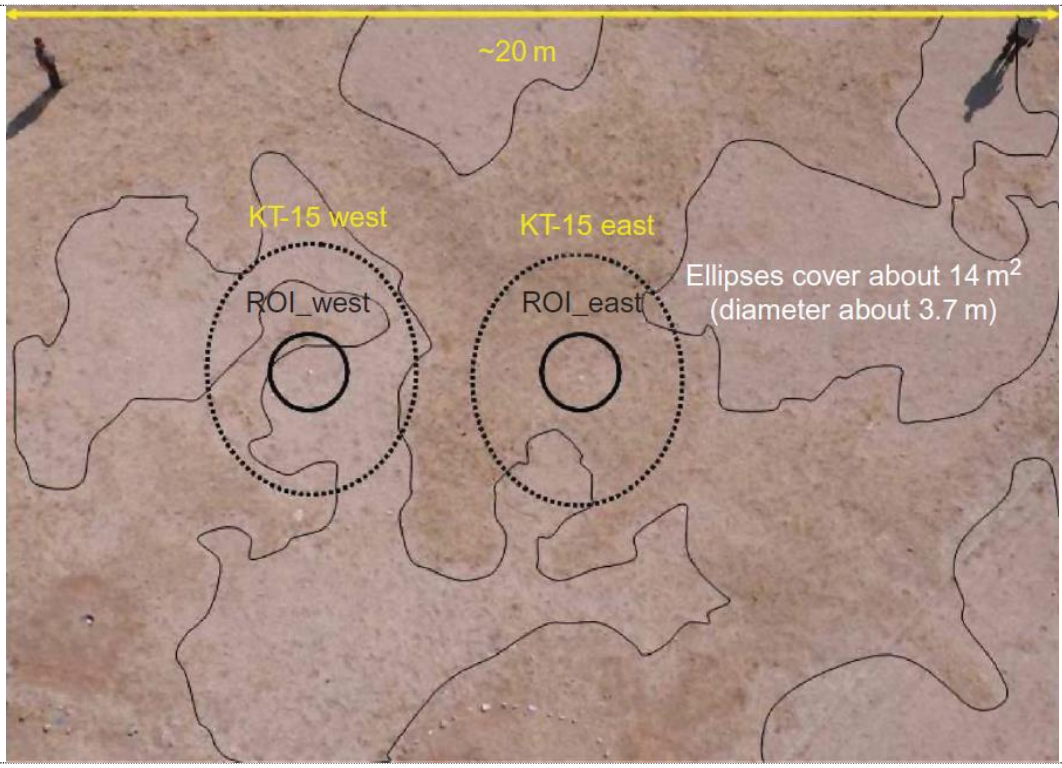


Figure 14 FOV of the KT-15 radiometers at Gobabeb station and smaller ROI on 'grass' and 'gravel' used with a thermal infrared camera. The hand-drawn lines roughly outline the (sparse) grass-pure gravel boundary (Göttsche et al., 2013).

2.3.1 Initial validation results for Gobabeb

Freitas et al. (2010) compared Gobabeb in-situ LST against LSA SAF LST for the nearest SEVIRI/MSG pixel for the period between May 2008 and March 2009. They obtained in situ LST as the average of measurements taken by the downward looking radiometers, previously corrected for surface emissivity and down-welling radiance. Since in-situ emissivity was not available, it was set to 0.959 as representative for the gravel plains (further details see Trigo et al., 2008a). Figure 15 presents the resulting scatterplots of in-situ LST versus satellite LST retrievals for 6 months, representative of the full year (Freitas et al., 2010). Overall there is very good agreement between the two datasets, with root mean square differences (RMS) between 1 and 2ºC. The high cloud cover during the rainy season reduces significantly the number of retrieved LST values, explaining the significantly fewer points in the July 2008 plot (Figure 15). The larger discrepancies between the two datasets generally correspond to an underestimation of in-situ LST by SEVIRI/MSG LST (points clearly above the regression line in the panels, e.g. for Sep 08 and Jan 09). These are likely to be associated to cloud contamination or cloud shadow within the MSG pixel, and, thus, are also linked to scaling discrepancies between point and pixel measurements.

The area surrounding Gobabeb station is dominated by bareground gravel plains, which are associated with relatively large emissivity uncertainties. These constitute the major source of error

for LST satellite estimations, with the highest impact observed during the dryer months of the year (Freitas et al. 2010; Göttsche and Hulley, 2012).

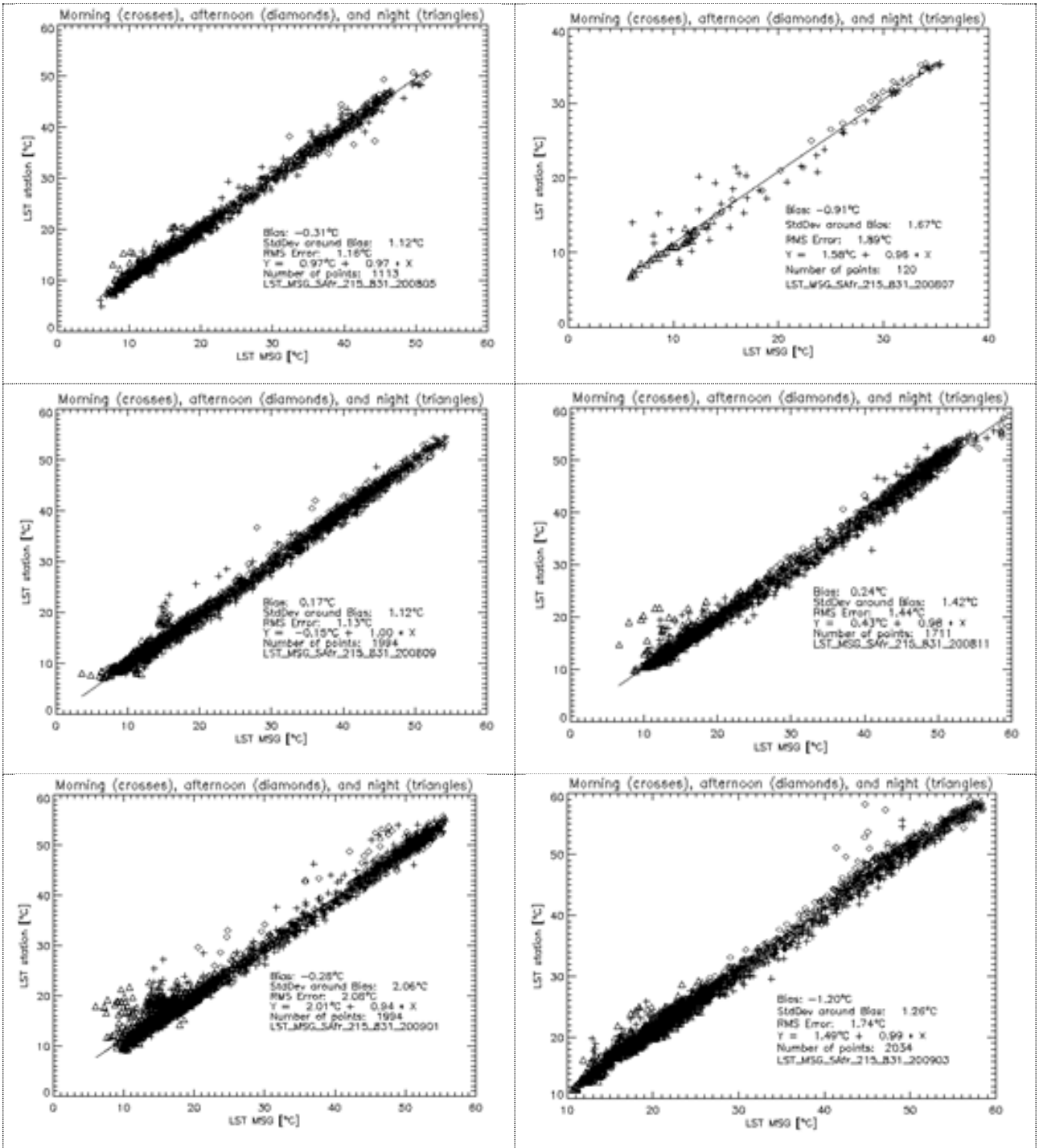


Figure 15 In-situ LST (ºC) from Gobabeb (y-axis) versus SEVIRI/MSG LST retrievals (x-axis). From top left to bottom right: May 2008, July 2008, September 2008, November 2008, January 2009 and March 2009. Each subplot shows mean difference between satellite retrievals and in situ observations (bias), standard deviation, root mean square (RMS) differences and the number of available matchups (Freitas et al., 2010).

2.3.2 Land Surface Emissivity determination at Gobabeb

In order to assess the accuracy of the operational LSA-SAF land surface emissivity (LSE) for SEVIRI channel 10.8 over the gravel plains, which is quasi-static at 0.949, Göttsche and Hulley (2012) determined in-situ emissivity of relevant surface types at Gobabeb with the 'one-lid emissivity box' method. Assuming a dry grass fraction of 25% (Bork-Unkelbach, 2012) and accounting for additional uncertainty in the fraction of dry grass, the LSE of the gravel plains for SEVIRI channel 10.8 is estimated as 0.944±0.015, which is in very good agreement with the corresponding ASTER-TES and MODTES emissivities (Göttsche and Hulley, 2012). Combining in-situ measurements performed in 2011 and 2012 at Gobabeb and assuming a dry grass fraction of 25%, the emissivity for the KT15.85IIP is estimated as 0.940 ± 0.015: this is the value used to derive in-situ LST at Gobabeb (e.g. Figure 18).

2.3.3 Gobabeb site characterisation

The gravel plains are highly homogeneous in space and time, which makes them ideal for validating a broad range of satellite-derived products (Göttsche and Hulley, 2012). Nevertheless, for reliable product validation the effect of the small scale variation of surface materials (e.g. dry grass, rock outcrops) and topography needs to be fully characterised. Motivated by initial analyses of LST from the Gobabeb site (Göttsche and Olesen 2009; Freitas et al. 2010) as well as by the observed temperature differences between 'KT-15 east' and 'KT-15 west' (Figure 14), a field survey was performed in March 2010. The aim of the survey was to characterize the gravel plains more closely and to determine which KT-15 radiometer (or weighted average) is most representative for the gravel plains (Göttsche et al., 2013). Additional in situ measurements with a telescopic mast were performed at various locations of the gravel plains, and the retrieved LSTs were compared with LSTs from the main mast (Figure 16) as well as with LSTs retrieved from MSG/SEVIRI by LSA SAF.

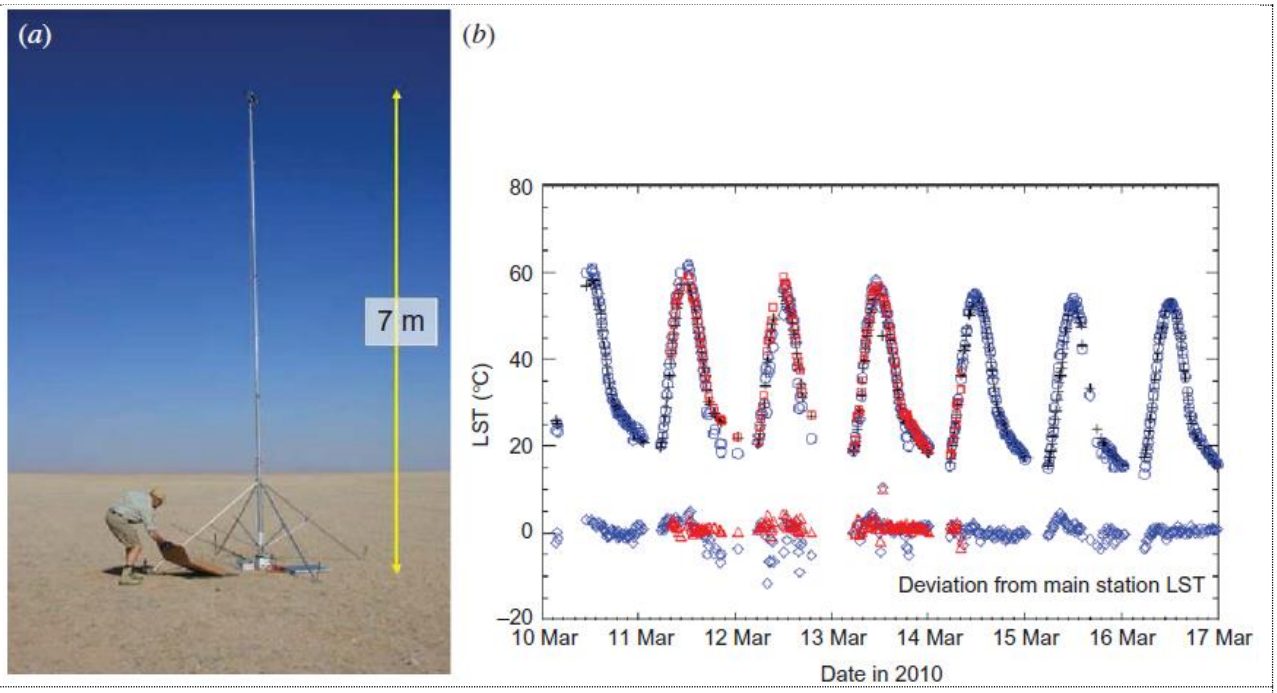


Figure 16 (a) Set-up of the mobile mast system at location on the highly homogeneous gravel plains. (b) Available LST for the time of the field survey resampled to match the acquisition times of MSG/SEVIRI. (Gobabeb main station, black; mobile station, red; MSG blue.) The lower red and blue points show the differences mobile LST – station LST andMSG LST – station LST, respectively. From Göttsche et al. (2013).

Compared to LST obtained from the 'KT-15 east' radiometer the mobile station LST (Figure 16) had a positive bias of about 1.4 K and a standard deviation of 1.2 K, which means that the gravel plain at the mobile mast is systematically warmer than the surface observed by 'KT-15 east'; for 'KT-15 west' the deviations were larger. However, when comparing LSA SAF LST derived from the SEVIRI pixel at the mobile mast location against Gobabeb main station LST, LSA SAF LST had a positive bias of about 0.4 K and a standard deviation of 2.0 K. These comparisons highlight the difficulty of finding locations which are representative for spatially coarse satellite pixels, e.g. with a resolution of 25 km2 as for MSG/SEVIRI. In order to obtain a more comprehensive picture of the LST distribution over the gravel plains, in situ LST were measured along a 40 km track (Figure 17). The results in Figure 18 show that MSG/SEVIRI LSTs retrieved by LSA SAF are very close to Gobabeb main station LSTs and well within LSA SAF's target accuracy of ±2 K (Göttsche et al., 2013). For the results in Figure 18 the determined bias is 0.5°C (other campaigns: -0.1°C to 0.8°C) and standard deviation 1.2°C, showing a high level of homogeneity and a stable relationship between Gobabeb station LST and LST obtained along the tracks. The bias between LSA SAF LST and in situ LST along the track is less than 0.5 °C, indicating that the LST obtained from the in-situ measurements are representative for the LST derived from MSG/SEVIRI pixel, i.e. over large areas.

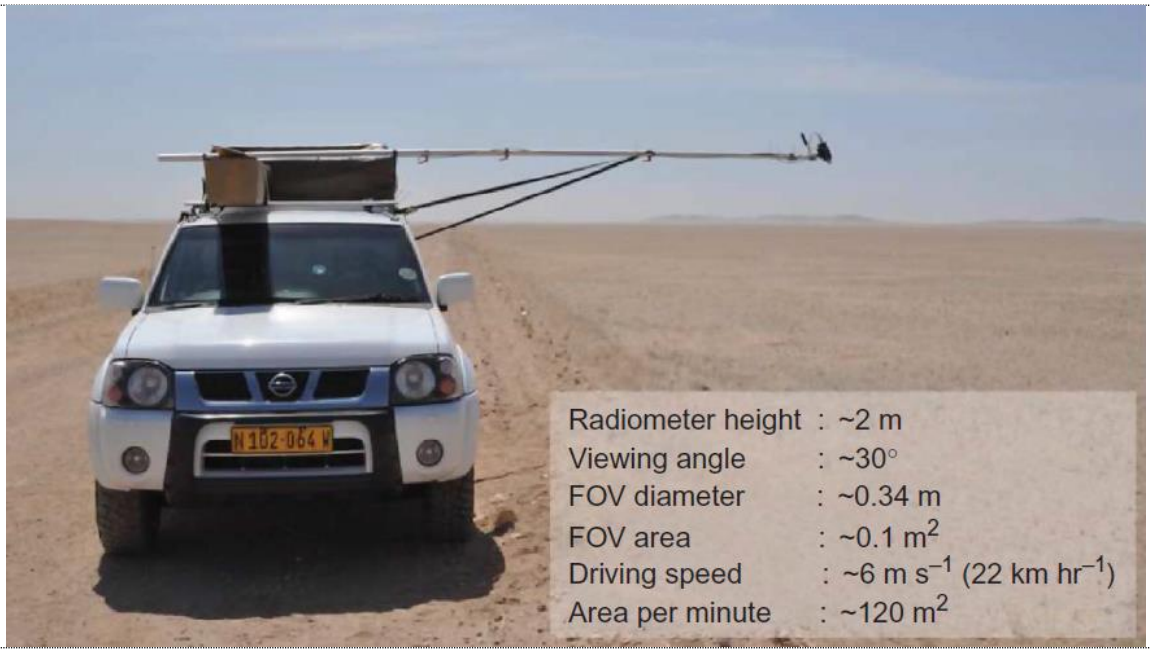


Figure 17 The 4WD with the mobile mast system on the gravel plain near Gobabeb, Namibia (Göttsche et al., 2013).

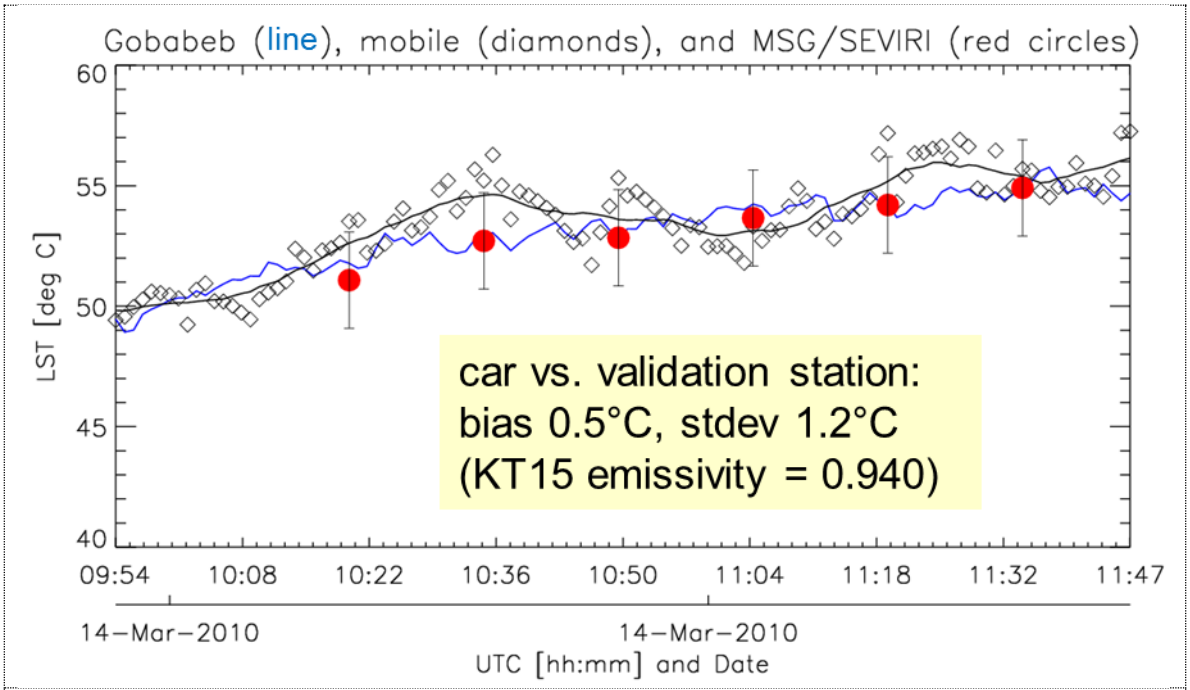


Figure 18 LST measured from along a distance of about 40 km on 14 March 2010. Gobabeb main station LST (line), LST obtained along the driven track (diamonds; black curve: 15 min moving average), and the seven co-located LSA SAF LSTs (filled circles with error bars of 2 K; missing data at 10:00 UTC).

2.3.4 Long-term validation results for Gobabeb

For Gobabeb a 5 minute delay between actual satellite acquisition time and nominal product time is accounted for and in-situ LST and LSA SAF LST are matched to better than 1 minute. Figure 19, Figure 20 and Figure 21 show plots of LSA SAF LST derived from MSG/SEVIRI against Gobabeb station LST for May, August, and November 2011, respectively. The three months represent Autumn, Winter and Spring at Gobabeb (southern hemisphere) and usually have a high number of clear sky situations. The in-situ LST at Gobabeb were derived from KT15.85 IIP measurements using a static emissivity of 0.940 (Göttsche and Hulley, 2012). For May 2011 (Figure 19) there are 2464 match-ups between satellite and in-situ LST for which bias and rmse are determined as -0.84°C (satellite colder) and 1.45°C, respectively. From Figure 19 it can be seen that the negative bias is mainly due to lower satellite LST estimates at night-time.

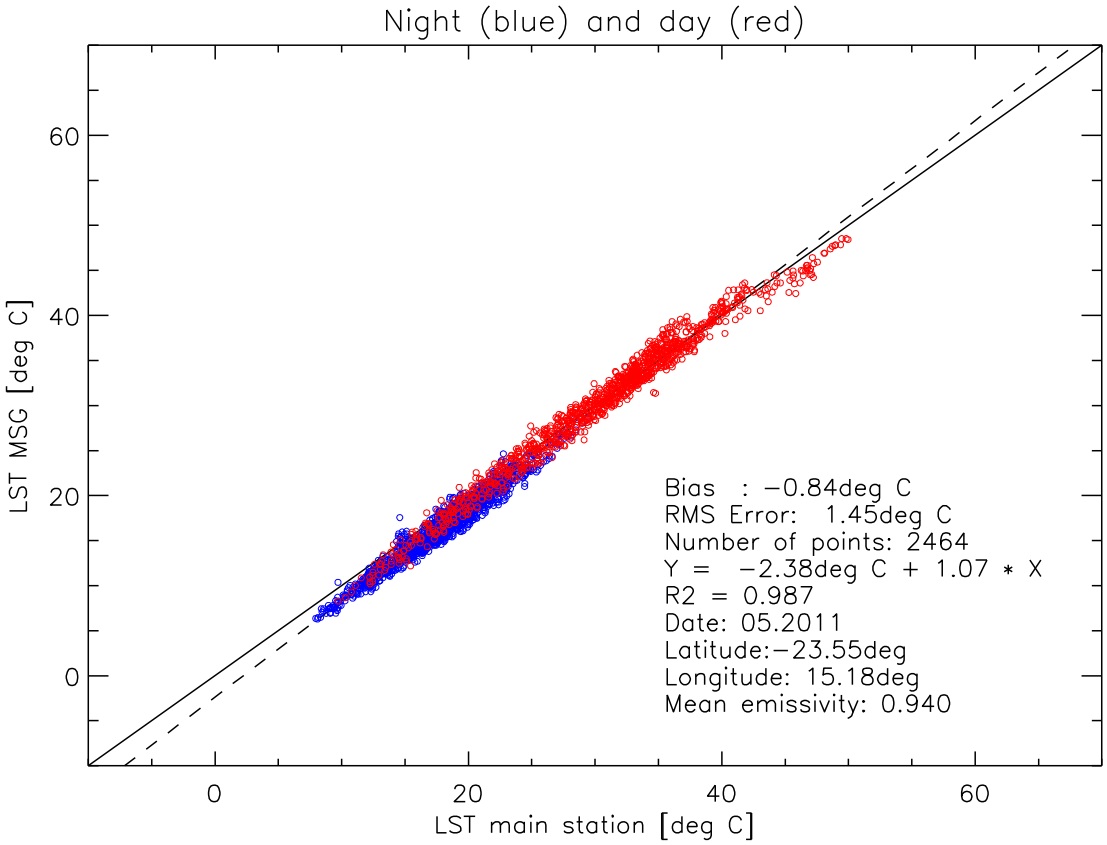


Figure 19 LSA SAF LST against Gobabeb in-situ LST with for May 2011 (blue circles: night-time, red circles: daytime).

For August 2011 (Figure 20) there are 2506 match-ups (i.e. 84% of all potential satellite observations are cloud-free) between LSA SAF LST and in-situ LST; the determined bias and rmse are 0.13°C (satellite colder) and 1.46°C, respectively. However, Figure 20 also shows that LSA SAF LST estimates tend to be slightly too low at night time and slightly too high at daytime.

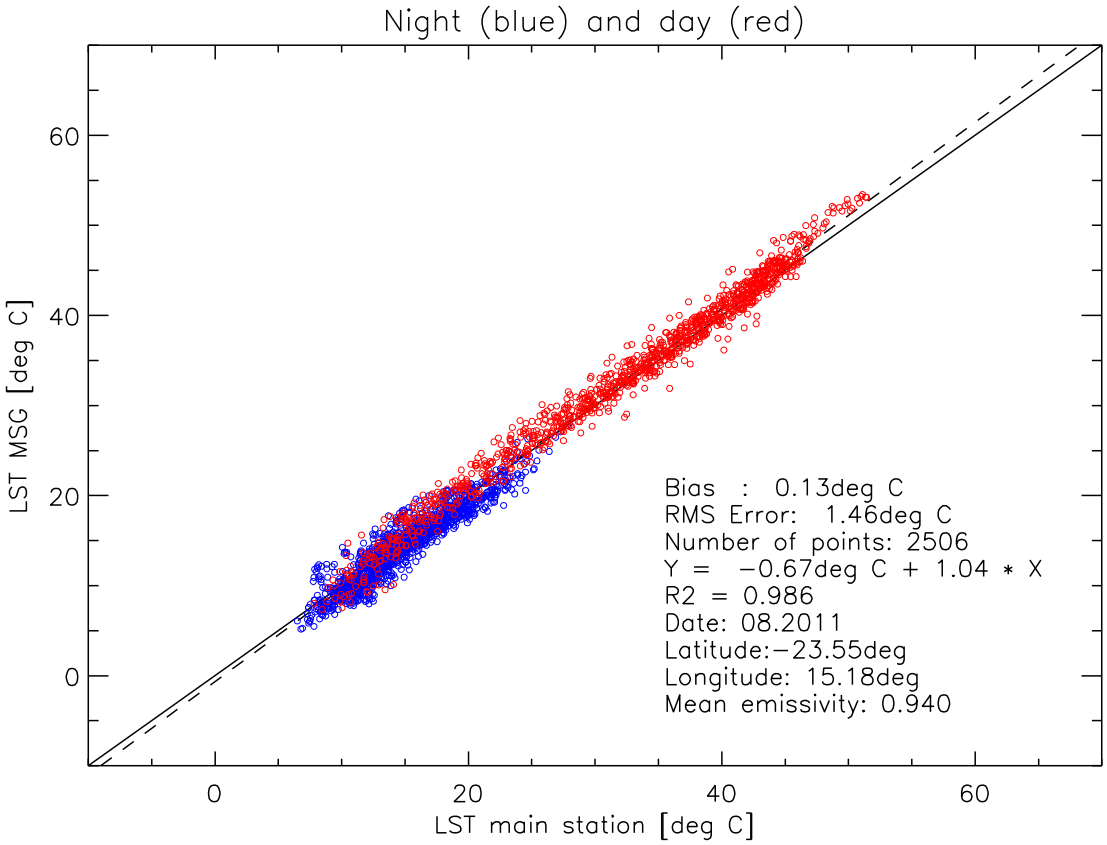


Figure 20 LSA SAF LST against Gobabeb in-situ LST with for August 2011 (blue circles: night-time, red circles: daytime).

For November 2011 (Figure 21) there are 2158 match-ups between satellite and in-situ LST and bias and rmse are 0.27°C (satellite colder) and 1.33°C, respectively. As for August 2011, LSA SAF LST estimates are slightly lower than in-situ LST at night time and slightly higher at daytime.

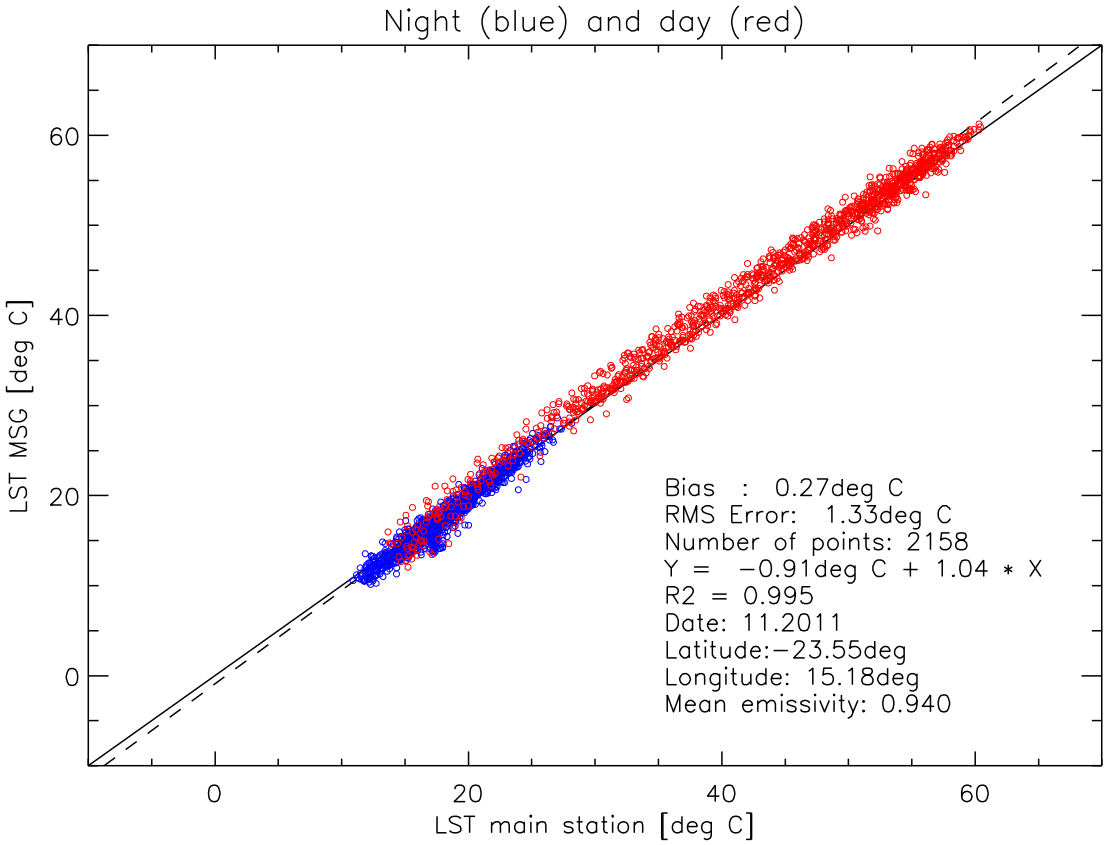


Figure 21 LSA SAF LST against Gobabeb in-situ LST with for November 2011 (blue circles: night-time, red circles: daytime).

Figure 19, Figure 20 and Figure 21 clearly demonstrate that the LSA SAF LST are in excellent agreement with Gobabeb station LST. This means that the LSA SAF GSW algorithm performs well in desert regions, i.e. at high LST (>50°C) and with considerable surface overheating w.r.t. air temperature (>20°C). Furthermore, the results show that the chosen station location and instrumentation provide in-situ LST that are representative for spatially coarse satellite LST products, which was also demonstrated with additional measurements across the gravel plains (Göttsche et al., 2013). The high level of agreement gives confidence in the station concept and in measurements performed at the other, more complex validation sites.

Figure 22, Figure 23 and Figure 24 display monthly biases & rmse together with the corresponding numbers of valid data points for Gobabeb between June 2008 and July 2014. Figure 22 shows the results for night-time and daytime data. The bias appears to be seasonal with minimum values of about -1.0°C around May and positive values of about +0.5°C around November; the overall mean bias and rmse are 0.1°C and 1.6°C, respectively. The data gap around December 2012 is due to technical problems. The higher bias observed around October 2010 (about +1.0°C) and the rmse increase between December 2010 and March 2011 (to about 3.5°C) are due to an exceptionally wet rainy season (Namibia's 'small' rainy season is in Oct/Nov and its 'big' rainy season between Jan and Apr). The increased rmse & bias in Jan/Feb 2014 is also caused by cloud-contamination; however, there was insufficient precipitation for triggering a substantial growth of grass. The gravel plains usually receive little or no rain, i.e. the long term annual average at Gobabeb is 25mm per year (Eckardt et al, 2013). According to Eckardt et al., cumulated rainfall in 2011 was about 175 mm and 80 mm in 2008, which resulted in an unusual growth of grass on the gravel plains. The negative bias after these two rainy seasons is more pronounced and lasted longer than in the other years (see Figure

22). This could be explained by the different effect that grass, which desiccates rapidly at Gobabeb, has on SEVIRI channel 9 and the KT15.85 IIP. According to (Olioso et al., 2007) and (Hulley et al., 2014) the emissivity of dry grass decreases considerably between 8µm and 11µm. For dry grass the wider spectral response function of the KT15.85 IIP, therefore, leads to an increase in effective emissivity. In contrast, a higher dry grass fraction affects the spectrally narrower SEVIRI channel 9 less and may even reduce its corresponding emissivity, since the SEVIRI channel 9 emissivity of the background gravel & sand may be slightly higher. This is supported by Figure 23, which shows the corresponding results for night-time data: no consistent seasonality of the bias is observed, but there are pronounced periods of negative bias during & after the big rainy seasons in 2009 and 2011. The bias during these rainy periods reaches negative values of around -1.5°C and -2.5°C, respectively; since Figure 23 only shows night-time data, illumination effects can be ruled out. LSA SAF's dynamic LSE product, which is based on the vegetation cover method, is practically constant over the shown data period.

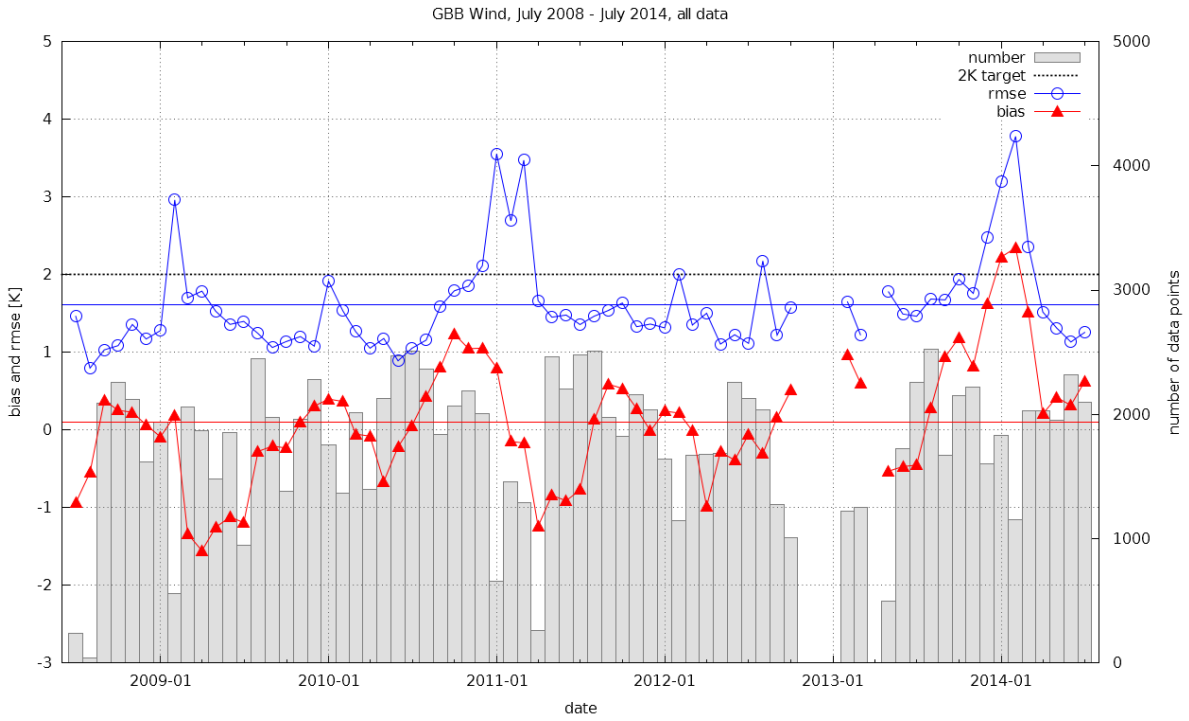


Figure 22 Daytime and night-time monthly statistics at Gobabeb station, Namibia, for LSA SAF LST. Mean bias (red triangles) and rmse (blue circles) refer to the left y-axis, the number of match-ups (grey bars) to the right y-axis.

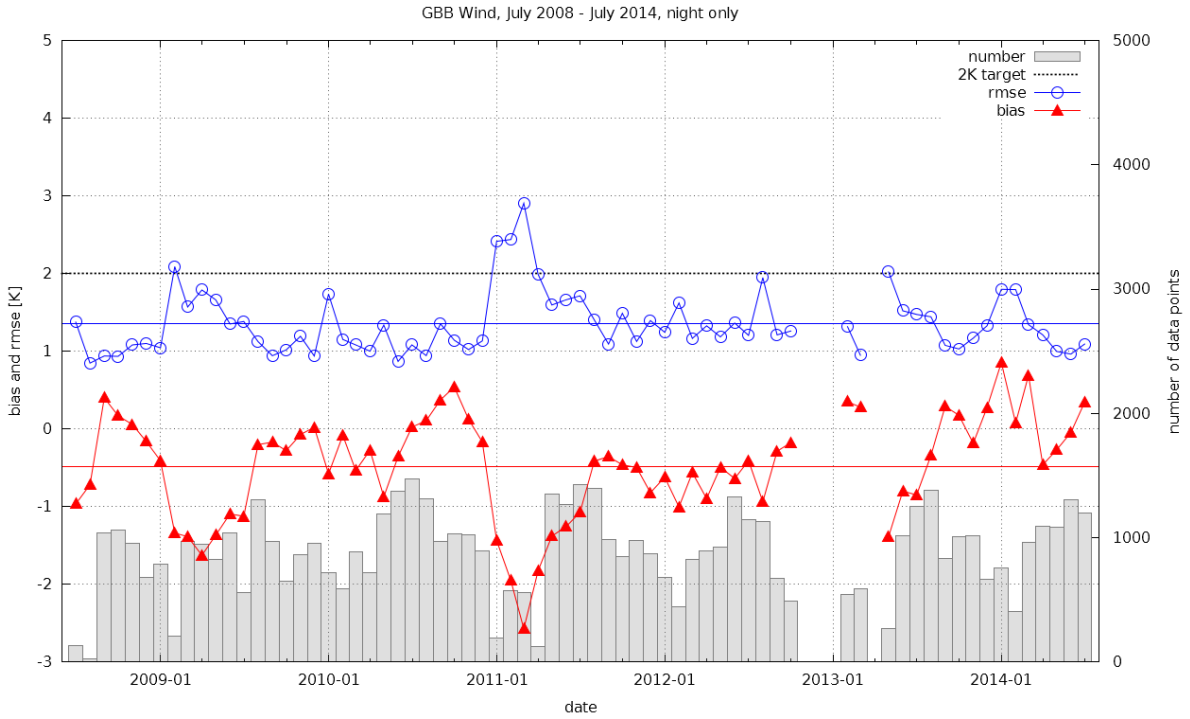


Figure 23 Night-time monthly statistics at Gobabeb station, Namibia, for LSA SAF LST. Mean bias (red triangles) and rmse (blue circles) refer to the left y-axis, the number of match-ups (grey bars) to the right y-axis.

Figure 24 shows monthly biases & rmse for the daytime data: the bias shows seasonal behavior, which is superimposed with the signal from the exceptional rainy seasons 2009 & 2011 and, as for the night-time data, periods of negative bias are more pronounced around these rainy seasons. During and just after the rainy seasons the bias is generally negative, whereas it tends to be slightly positive otherwise (ignoring the extreme rainy season 2010/2011). This periodic behavior of the bias for the day-time data could be caused by a seasonal variation in atmospheric correction of LSA SAF's LST retrieval algorithm, i.e. a systematic overestimation of water vapor content during the rainy season. This would also help to explain the strong increase in bias from Sep to Nov 2010, since an overestimation of water atmospheric vapor content would result in too high satellite LST.

From Figure 22, Figure 23 and Figure 24 it is concluded that for Gobabeb the LSA SAF LST algorithm meets its target accuracy of 2°C when considering all data together as well as when considering daytime and night-time data separately.

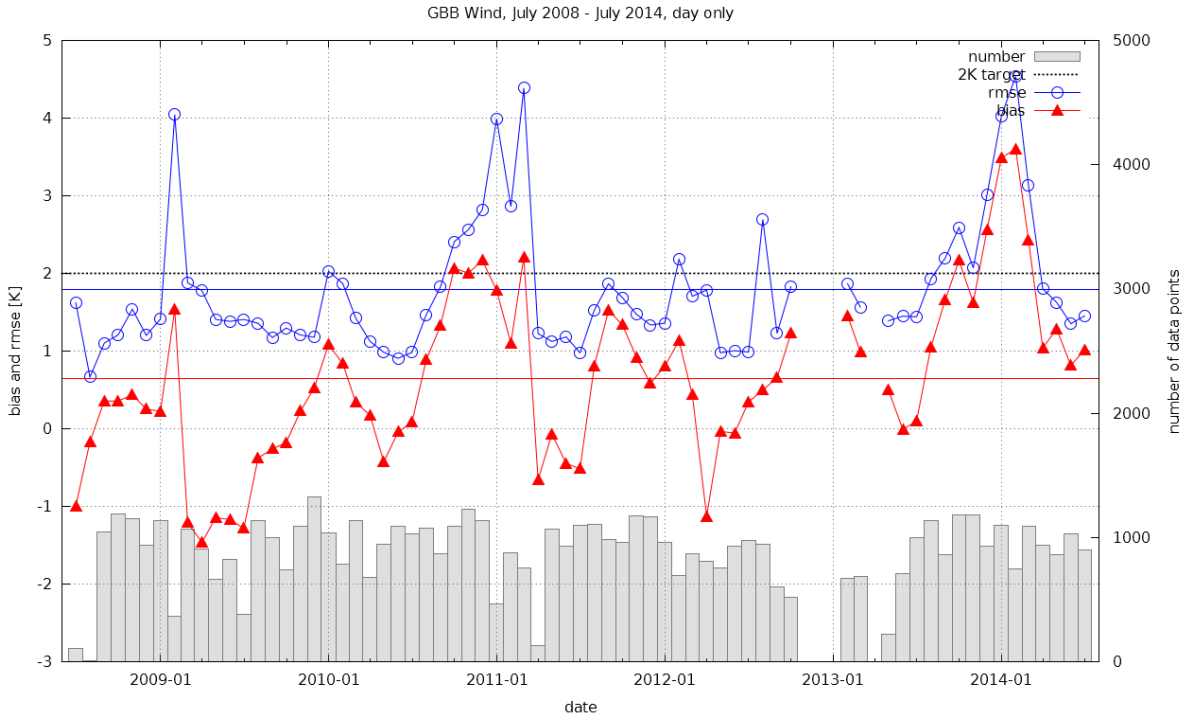


Figure 24 Daytime monthly statistics at Gobabeb station, Namibia, for LSA SAF LST. Mean bias (red triangles) and rmse (blue circles) refer to the left y-axis, the number of match-ups (grey bars) to the right y-axis.

2.4 Farm Heimat, Namibia

Farm Heimat (22.933° S, 17.992° E, 1380 m a.s.l.) lies about 100 km south-east from Windhuk on a plateau in the Kalahari semi-desert. The Kalahari is characterised by hot and arid climate, which exhibits a natural seasonality: there is a small rainy season with very little rain (September to November) and a big rainy season (January and March) with possible flooding. Outside the big rainy season the Kalahari bush is dry and the grass desiccates quickly. Farm Heimat produces livestock (cattle & sheep) and is also used for hunting game (mainly springbok and Oryx). The farm itself has a size of about 50 km2, but the land cover ('Kalahari bush') and land use in a wide area (thousands of km2) around farm Heimat are identical. Cattle are carefully managed and moved systematically between fenced off 'camps' to avoid overgrazing. In-situ measurements at farm Heimat started in February 2011. The station is located in a typical Kalahari land scape and a wide area around the mast is mainly covered by patchy, desiccated grass dotted with bushes and isolated camel thorn trees. Due to the station's high elevation winter temperatures (June - August) frequently drop well below freezing point. Furthermore, the remaining water vapour column between the surface and a satellite is quite small, i.e. about half of the atmosphere's water vapour is contained in the lowest 2 km.

Figure 25 shows a view over the Kalahari bush from the mast at farm Heimat. The station is equipped with radiometers measuring the brightness temperatures of the crown of a small tree, grass (2x), and the sky. Standard meteorology, a rain gauge (tipping bucket), and a radiation balance are also available.

Figure 25: View from mast of 'Farm Heimat' LST validation station, Kalahari, Namibia (February 2011;during wet season). In front: radiation balance sensor 'Hukseflux NR01'.

2.4.1 Estimation of land surface cover and representative in-situ LSTs

For the area of the MSG/SEVIRI pixel located over farm Heimat, the relevant surface cover types and their fractional coverage were determined from publicly available Google Earth (earth.google.com) imagery. Using object based image analysis (Bork-Unkelbach, 2012), it was found that the validation area is covered by 15% trees, 22% bush and 63% grass & sand; the latter form a joint class, since the cover fraction of grass strongly depends on season and annual rainfall. Therefore, the brightness temperature of a representative area of grass & sand is measured with a single in-situ radiometer. Furthermore, trees at farm Heimat are generally small and dry (Figure 25) and are thermodynamically similar to bushes; both are usually close to air temperature. Therefore, the tree and bush cover fractions are treated as a single cover fraction of 37%.

2.4.2 Land Surface Emissivity at Farm Heimat

The end-members observed by the Heitronics KT15.85 IIP radiometers are grass/bare ground and bush/tree crown. More bare ground may be observed during the dry season, whereas during the wet season the surface can be completely covered by relatively high grass (up to 1 m), which increases effective emissivity via the volume effect. Currently the KT15.85 IIP's emissivities of the two observed end-members are set to the value retrieved operationally by LSA-SAF for MSG/SEVIRI channel 9: over the course of the year this value varies between about 0.973 – 0984, which is in good agreement with literature values for vegetation.

2.4.3 Long-term validation results for farm Heimat

For Farm Heimat a 5 minute delay between actual satellite acquisition time and nominal product time is accounted for and in-situ LST and LSA SAF LST are matched to better than 1 minute. Figure 26, Figure 27 and Figure 28 show scatter plots of LSA SAF LST derived from MSG/SEVIRI against Farm Heimat in-situ LST for May, August, and November 2011, respectively. The three months represent Autumn, Winter and Spring at Farm Heimat and usually have a high number of clear sky situations. In-situ LST were derived using LSA SAF's emissivity product, which at Farm Heimat is about 0.975. For May 2011 (Figure 26) there are 1977 match-ups between satellite and in-situ LST, for which the determined bias and rmse are -0.24°C and 1.39°C, respectively. From Figure 26 it can be seen that bias is positive at night time (LSA SAF LST warmer) and negative at daytime (LSA SAF LST colder). The bias and rmse for August 2011 are 0.56°C and 1.52°C, respectively, and there is very good agreement between LSA SAF daytime LST and in-situ LST (Figure 27). For November 2011 (Figure 28) bias and rmse are 0.33°C and 1.41°C, respectively. For all three months (Figure 26, Figure 27 and Figure 28) there is a pronounced positive night-time bias, which indicates that LSA SAF LST are systematically warmer than in-situ LST; at daytime LSA-SAF LST are (slightly) colder than in-situ LST.

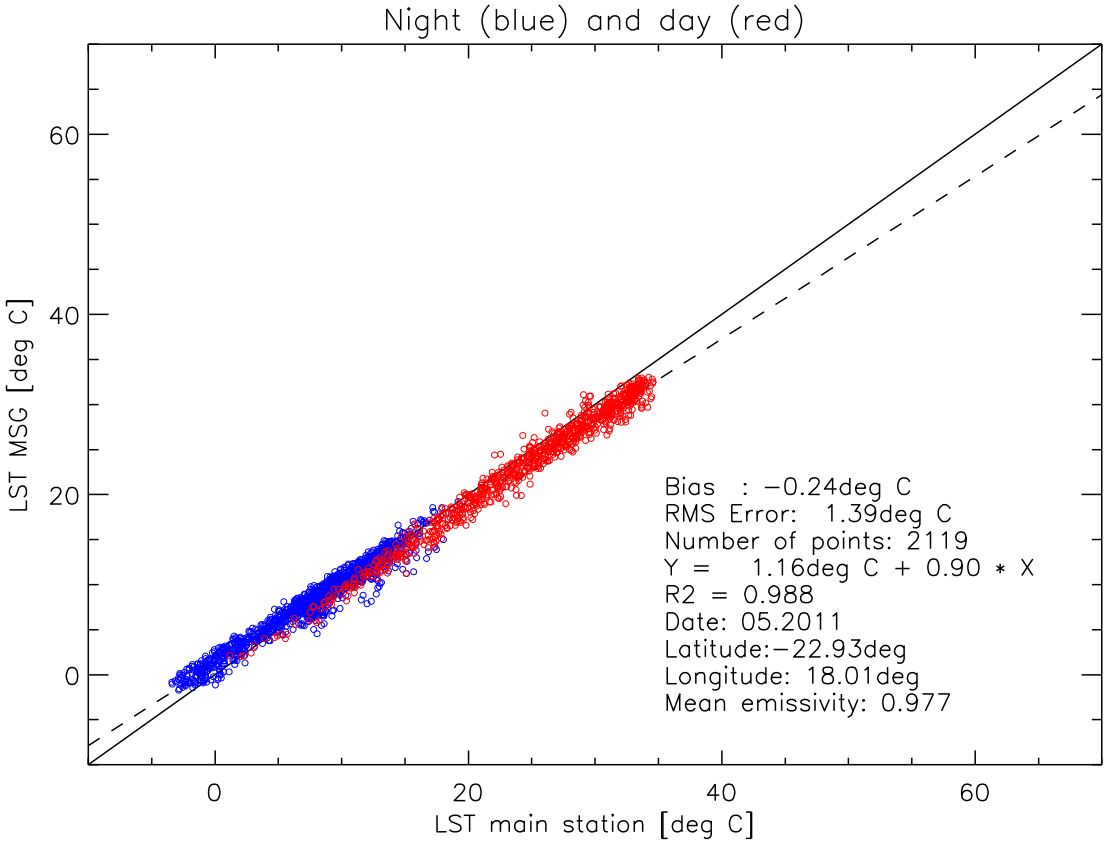


Figure 26 LSA SAF LST against farm Heimat in-situ LST for May 2011 (blue circles: night-time, red circles: daytime).

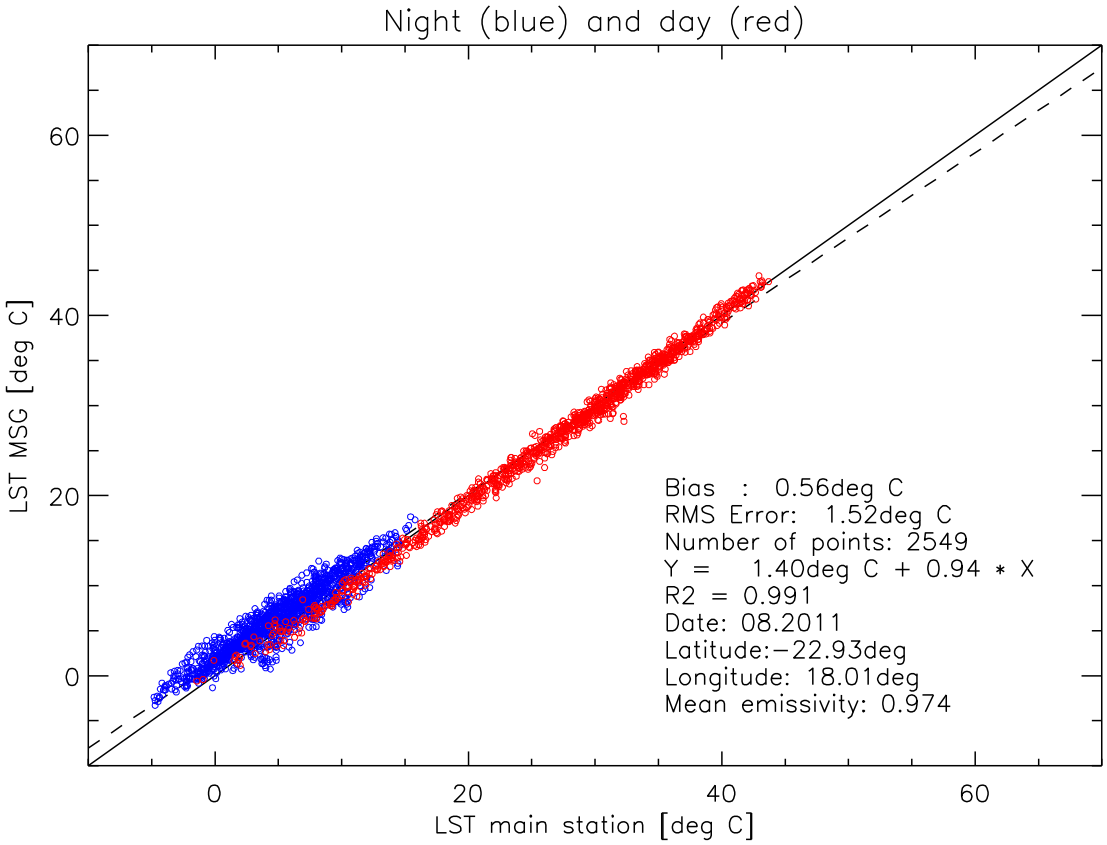


Figure 27 LSA SAF LST against farm Heimat in-situ LST for August 2011 (blue circles: night-time, red circles: daytime).

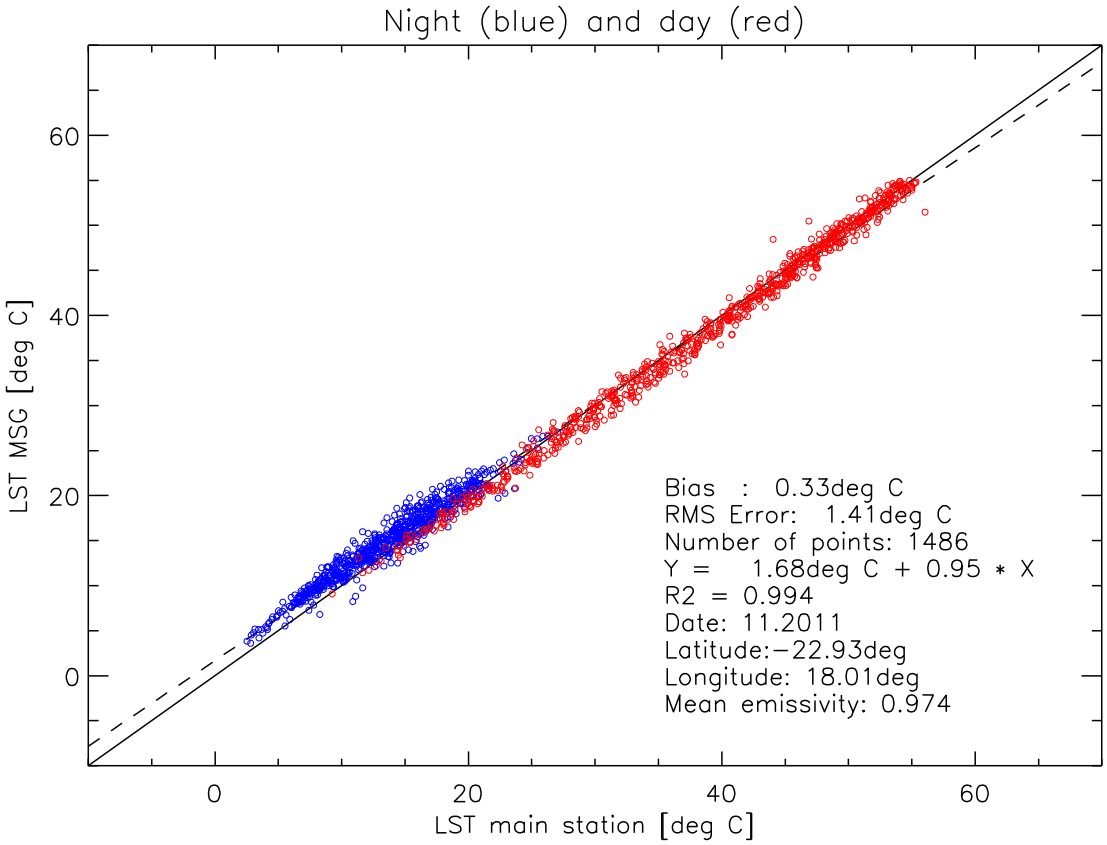


Figure 28 LSA SAF LST against farm Heimat in-situ LST for November 2011 (blue circles: night-time, red circles: daytime).

Figure 29, Figure 30 and Figure 31 display monthly biases & rmse together with the corresponding numbers of valid data points for farm Heimat between March 2011 and June 2014. Figure 29 shows the results for night-time and daytime data: with the exception of March and April 2011, which were affected by the exceptional rainy season 2010/2011 in Namibia, the monthly bias varies relatively little (amplitude about 0.5°C) and has a mean of about 0.1°C. The monthly rmse in Figure 29 also shows little variation and has a mean of 1.2 °C. However, the small rmse for the combined daytime and night-time data partially results from compensation between different night-time (Figure 30) and daytime (Figure 31) biases: whereas mean night-time bias is 0.7°C, mean daytime bias is -0.5°C. The corresponding means of rmse are 1.3°C and 1.4°C for night-time and daytime, respectively. Finally, there appears to be a seasonal increase of night-time bias and rmse (Figure 30) around southern hemisphere Winter (June – August).

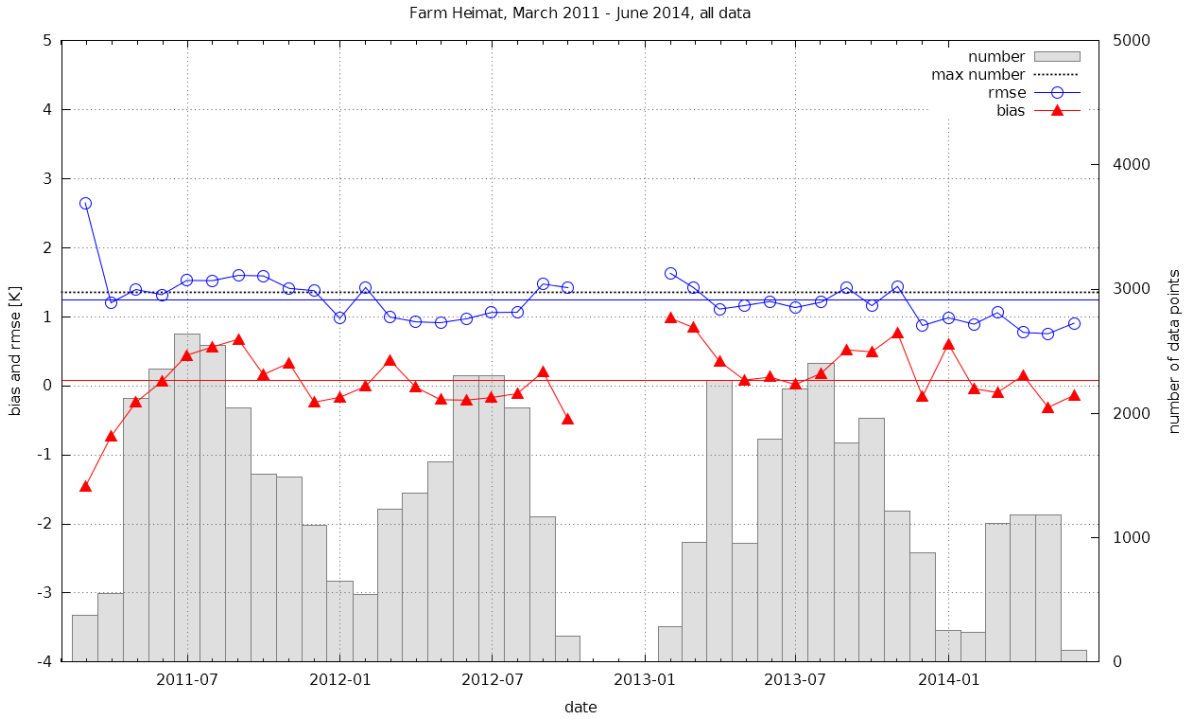


Figure 29 Daytime and night-time monthly statistics at farm Heimat, Namibia, for LSA SAF LST. Mean bias (red triangles) and rmse (blue circles) refer to the left y-axis, the number of match-ups (grey bars) to the right y-axis.

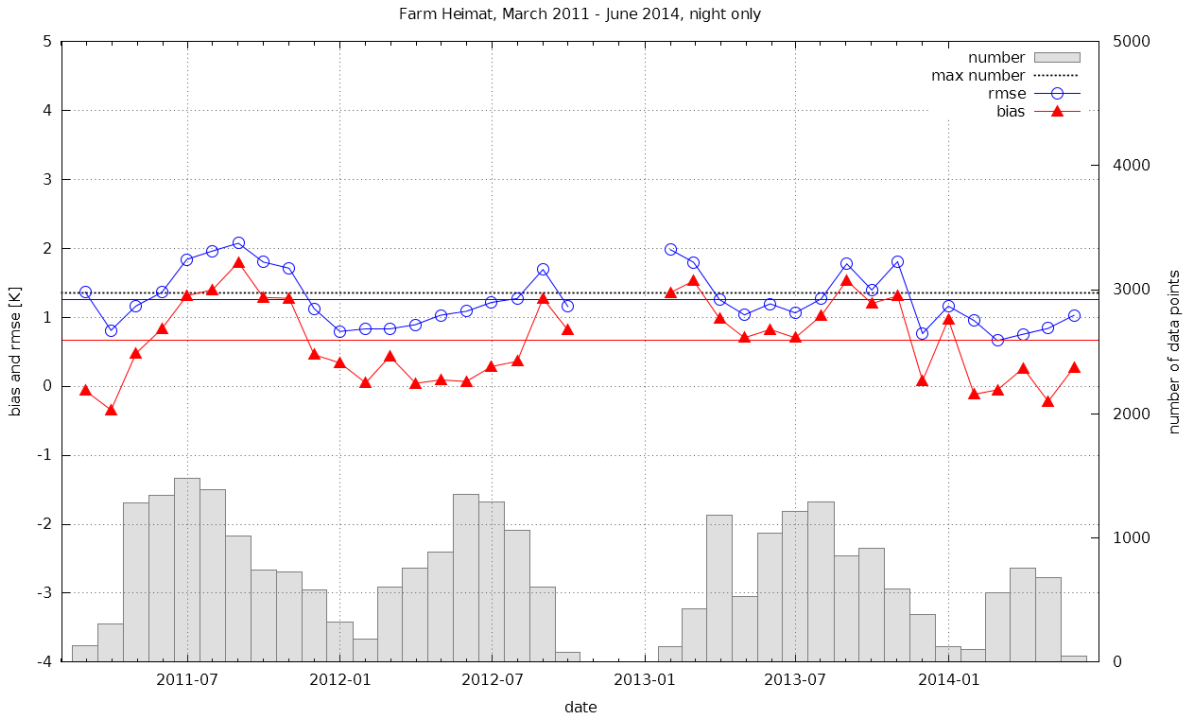


Figure 30 Night-time monthly statistics at farm Heimat, Namibia, for LSA SAF LST. Mean bias (red triangles) and rmse (blue circles) refer to the left y-axis, the number of match-ups (grey bars) to the right y-axis.

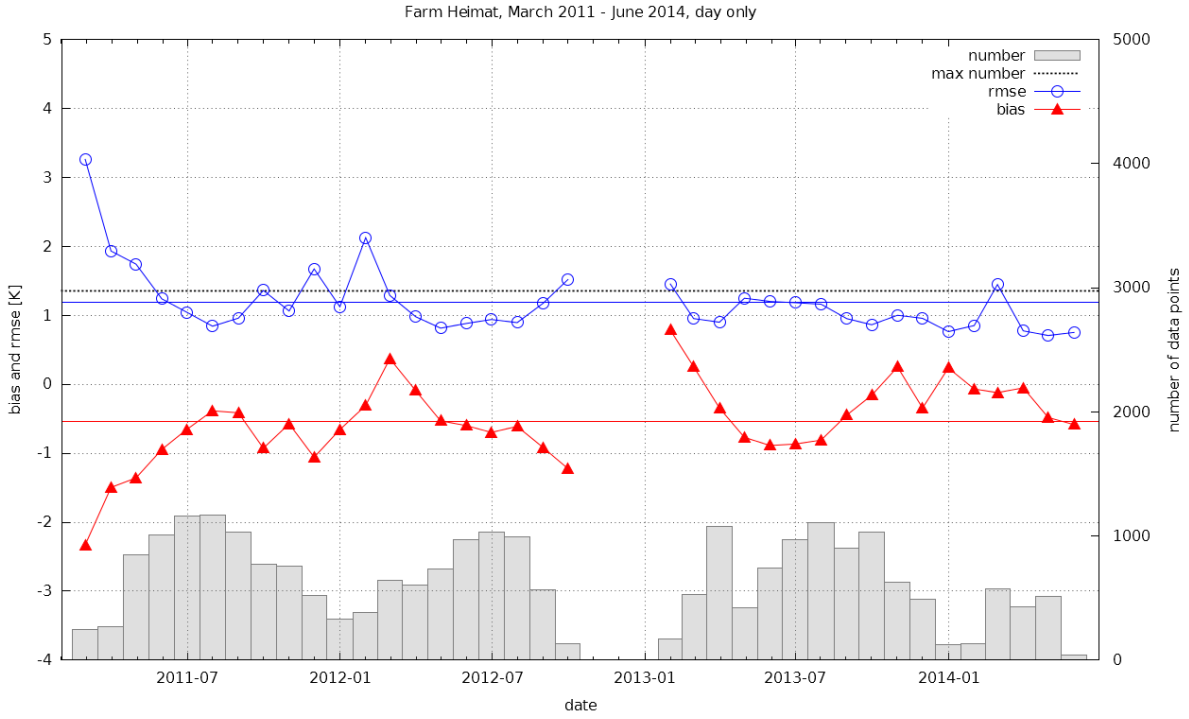


Figure 31 Daytime monthly statistics at farm Heimat, Namibia, for LSA SAF LST. Mean bias (red triangles) and rmse (blue circles) refer to the left y-axis, the number of match-ups (grey bars) to the right y-axis.

2.5 Dahra, Senegal

Dahra LST validation station (15.402° N, 15.443° W, 45 m a.s.l.) is located about 7 km north-east of the town of Dahra, Senegal. The field site is hosted by the Centre de Recherches Zootechniques de Dahra, Institut Senegalais de Recherches Agricoles (ISRA), and also includes two towers operated by the University of Copenhagen for validating satellite products (Stisen et al., 2008; Fensholt and Sandholt, 2005; Tagesson et al., 2015). The towers are equipped with instruments for validating satellite products in the visible, near-infrared, and the thermal domain.

The area around the station (Figure 32) is practically unpopulated and the dominant surface cover types are seasonal grass and sparse trees, which are mostly 'Acacia raddiana', 'Acacia Senegal', 'Balanites aegyptiaca' (Rasmussen et al., 2011). The soil is sandy and reddish in colour and was classified as an Arenosol (Batjes, 2011). The entire site is grazed by cattle and sheep while migrating camels from the northern Sahel feed on the leaves of the trees. The trees are scattered in the landscape, either as isolated trees or as small clumps. In some cases the distribution of the bushes and trees follows ancient dunes, which causes stripes of high vegetation - hence the name 'tiger bush'. According to the classification by Köppen and Geiger (Köppen, 1936; Peel et al., 2007), Dahra is characterised by the hot-arid, steppe-prairie climate typical of the Sahel region. During the dry season from October to March the climate is especially hot and the grass desiccates rapidly (see Figure 32), whereas the trees stay usually green throughout the year.

Figure 32: Dahra LST validation station, Senegal. Radiometers observe tree crown, grass and sky BT. Standard meteorology is available from a nearby station of the University of Copenhagen.

In contrast, the wet season (July to October) is strongly influenced by the monsoon, which is characterised by very humid atmospheres and strong and persistent cloud-cover, and the grass grows about 1m high (Rasmussen et al., 2011). The long-term average annual precipitation is about 370 mm and has a high inter-annual variability. Due to the distinct dry and wet seasons the validation site exhibits a strong annual vegetation cycle. Usually only the trees stay green all year while in the rainy season the grass grows dense and the entire site is covered by vegetation.

The current set-up of instruments at Dahra is in operation since July 2009, but due to technical problems and theft there are considerable data gaps. The sensor configuration for validating LST at the Dahra site consists of four 'KT-15.85 IIP' IR-radiometers (self-calibrating, chopped radiometers, Heitronics GmbH). The targets observed by the three surface facing KT-15 are two patches of grass / soil (direction south and direction west) and a canopy of a Acacia raddiana tree from south west. The fourth radiometer measures downwelling longwave radiance from the sky at 53° zenith angle. Dahra's low elevation of about 45 m a.s.l. results in long atmospheric paths and the atmospheric water vapour load varies strongly between the rainy season and the dry season: especially during the warm (about 40° C) and humid rainy season (up to 90 % relative humidity) the atmospheric correction of satellite TIR data is extremely challenging. Occasional outbreaks of Sahara dust complicate cloud detection further.

2.5.1 Estimation of land surface cover and representative in-situ LSTs

In the validation area at Dahra, the relevant surface cover types were determined from an Independent Component Analysis of high-resolution multispectral Quickbird data to consist of trees and grasslands. The fractional coverage of these end-members was subsequently determined as

described in (Bork-Unkelbach, 2012): the land surface of the MSG/SEVIRI pixel located over the KIT station Dahra is covered by 97% grasslands and 3% trees. The results were verified in a quality assessment and the TCC agrees well with a previous tree survey in the validation area, where the authors found a TCC between 3 and 4.5% by manually classifying 5 sample regions (Rasmussen et al., 2011a). Due to the small TCC value, the effect of varying TCC between 3% and 6% on retrieved LST is negligible: this was also verified for low sun and view angle combinations (Rasmussen et al., 2011).

2.5.2 Land Surface Emissivity at Dahra

Even though the trees at Dahra usually stay green throughout the year, they display some temporal variation in greenness. This variation is mainly caused by changes in crown density, but also by some species which are out of phase with the rainy season (Rasmussen et al., 2011a). While the emissivity of the tree crowns can be set to a literature value of 0.98, the emissivity of the grass & soil background varies significantly over the year (Xu et al., 2014). This is mainly due to seasonal changes in soil moisture and the strongly varying grass cover, which both greatly affects land surface emissivity. The operational LSA-SAF emissivity product for MSG/SEVIRI uses the vegetation cover fraction method to capture this variation. For an estimated tree crown cover (TCC) at Dahra of 4%, an emissivity of 0.98 for the tree crown and an assumed emissivity of 0.95 for the soil/grass background the error associated with integrating over entire pixel is about 0.001 (Rasmussen et al., 2011). LSA-SAF emissivities for MSG/SEVIRI channel 9 vary annually from about 0.969 (July) to 0.984 (September), which is largely related to seasonal changes of the grass. However, in-situ emissivities performed by KIT in January 2013 with the ‘one-lid emissivity box method’ (Rubio et al., 1997) indicate that LSA-SAF emissivities for the dry season at Dahra are too high (Xu et al., 2014): for unaltered, dominantly low lying, dry grass & sand mixtures an emissivity of 0.941 ±0.005 was determined for a KT15.85II P radiometer (Jimenez-Munoz et al., 2014), which has a similar, but slightly broader response function as SEVIRI channel 9 (Göttsche and Hulley, 2012).

2.5.3 Long-term validation results for Dahra

For Dahra a 7 minute delay between actual satellite acquisition time and nominal product time has been accounted for and in-situ LST and LSA SAF LST were matched to better than 1 minute. Generally, there are considerably more clear sky situations during the dry season then during the rainy season. Therefore, Figure 33, Figure 34 and Figure 35 show plots of LSA SAF LST derived from MSG/SEVIRI against Gobabeb station LST for November 2010 (begin of dry season), March 2011 (near middle of dry season), and May 2011 (near end of dry season), respectively. In-situ LST were derived using LSA SAF’s emissivity product, which at Dahra varies from about 0.984 at the end of the rainy season to 0.969 near the end of the dry season (e.g. May, see Figure 35). However, this limited emissivity range does not sufficiently reflect the strong seasonal vegetation cycle at Dahra (i.e. from near full vegetation cover to near bare soil). Furthermore, the vegetation cover method only responds to green vegetation, which may result in a wrong emissivity for desiccated grass. In-situ measurements of emissivity during the dry season (Jimenez-Munoz et al., 2014) and comparisons with other emissivity products (Xu et al., 2014) have shown that a realistic annual range for SEVIRI channel 9 LSE is from about 0.95 (end of dry season) to 0.985 (middle/end of rainy season).

For November 2010 (Figure 33) there are 1591 match-ups between satellite and in-situ LST and bias and rmse are determined as -0.11°C and 1.37°C, respectively. This suggests that in November – after the rainy season – the LSA SAF LSE of 0.975 approximates the actual emissivities for SEVIRI channel 9 and the KT15.85 IIP well, which is plausible since in November the site usually still has a cover of desiccated grass, for which (Göttsche and Hulley, 2012) determined a KT15.85 IIP emissivity of 0.962±0.013. Some outliers are still observed, indicating small, undetected clouds

(shadow) over the station (satellite LST higher) or sub-pixel clouds (in-situ LST higher). For March 2011 (Figure 34) there are 458 match-ups and bias and rmse are 0.05°C and 1.00°C, respectively. However, LSA SAF emissivity for March is 0.970, which is considered too high for the dry season. In May 2011 (Figure 35) emissivity has nearly the same value (0.969), there are 1876 match-ups and bias and rmse are -1.24°C and 2.12°C, respectively. The negative bias means that LSA SAF LST are systematically too low, which may be caused by warm and moist air already approaching with the monsoon from the South.

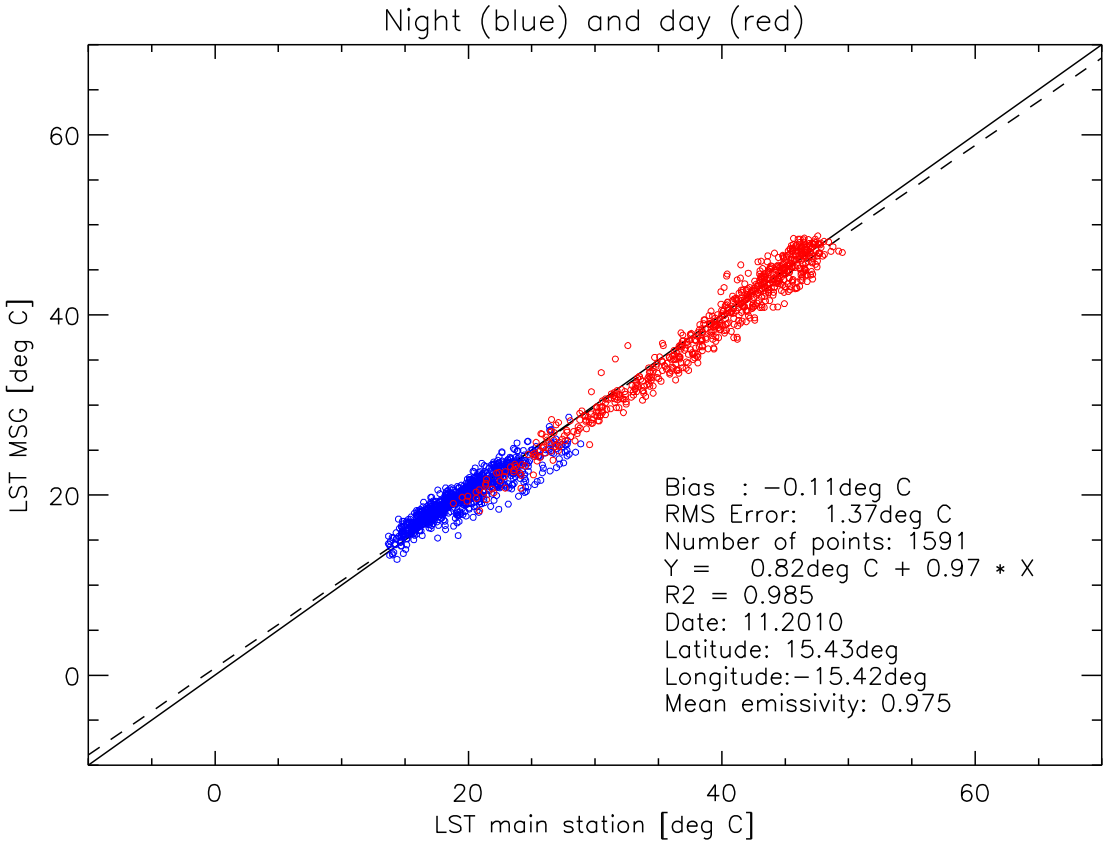


Figure 33 LSA SAF LST against Dahra in-situ LST with for November 2010 (blue circles: night-time, red circles: daytime).

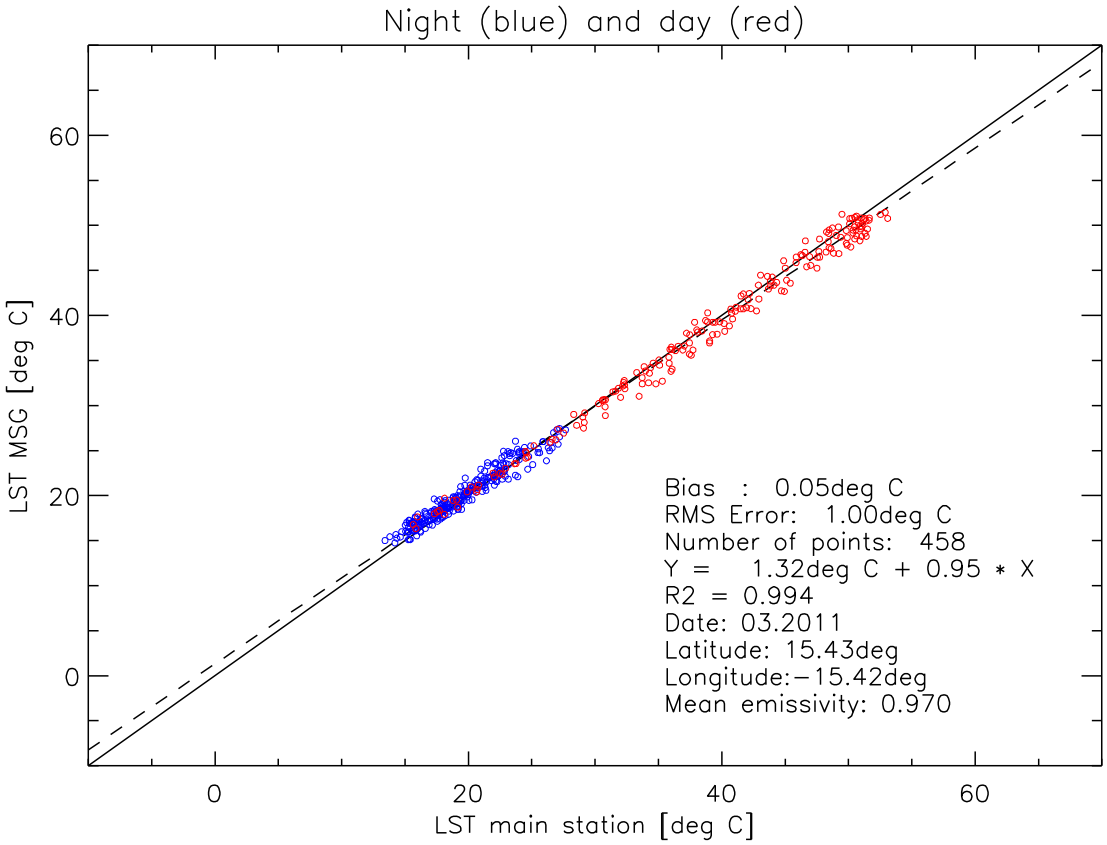


Figure 34 LSA SAF LST against Dahra in-situ LST for March 2011 (blue circles: night-time, red circles: daytime).

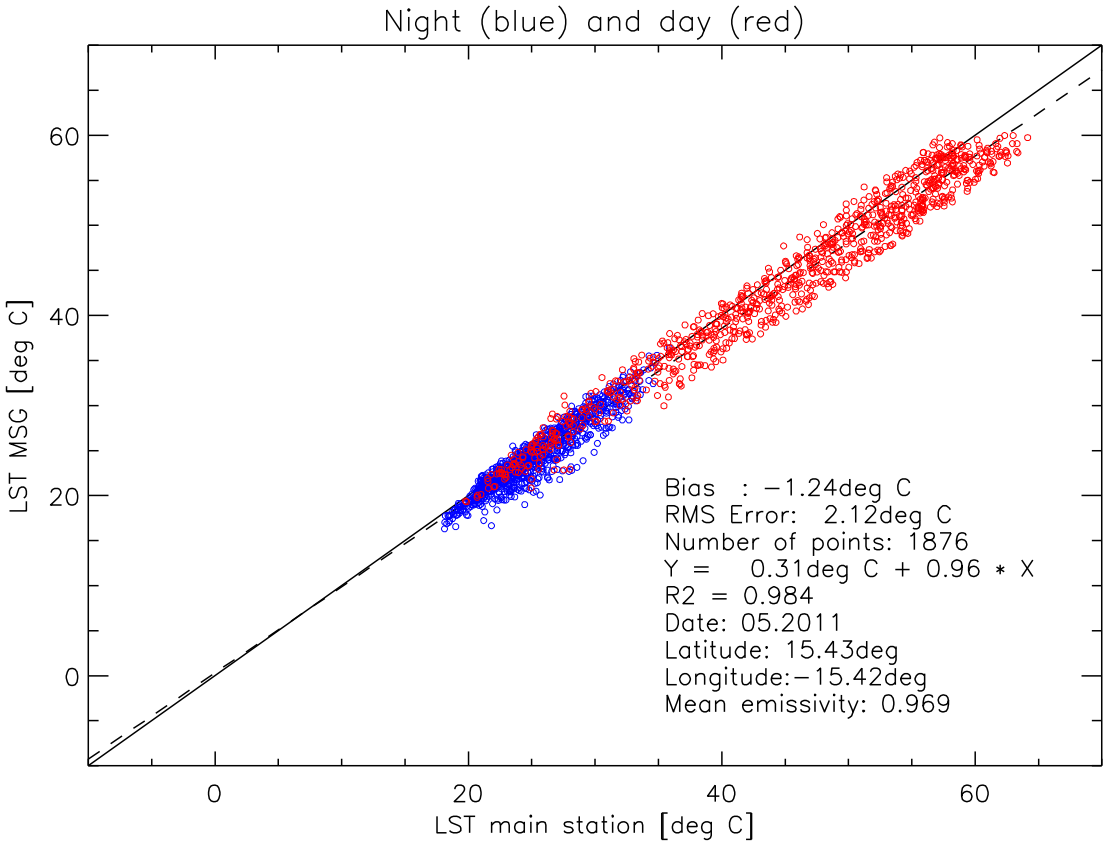


Figure 35 LSA SAF LST against Dahra in-situ LST for May 2011 (blue circles: night-time, red circles: daytime).

Figure 36, Figure 37 and Figure 38 display monthly biases & rmse for Dahra between July 2009 and July 2014: each data point in the plots represents the results for an entire month (e.g. a plot as in Figure 35) and the number of valid match-ups between LSA SAF LST and in-situ LST is shown as a grey bar. The large data gaps are due to stolen solar panels and technical problems with the station. In January 2014 the mast was damaged by cattle, which changed viewing geometry and observed areas: however, this appears to have had no obvious effect on the measurements from February 2014 onwards. Figure 36 shows the results for night-time and daytime data: monthly bias and rmse clearly vary seasonally, with strong negative biases during the rainy seasons, e.g. -6°C in August 2009, and rmse reaching up to 7°C; mean bias and rmse are -2.0°C and 3.2°C, respectively. However, limiting the data to the dry season, which is defined here as November to April, yields considerably a smaller mean bias and rmse of -0.04°C and 1.43°C, respectively. Analyzing only the night-time data (Figure 37), we obtain a mean bias and rmse of -1.5°C (dry season: 0.4°C) and 2.5°C (dry season: 1.1°C), respectively. For the day time data (Figure 38), we obtain a mean bias and rmse of -2.6°C (dry season: -0.6°C) and 4.1°C (dry season: 1.7°C), respectively. The highest daytime bias (-9.7°C) and rmse (10.7°C) are reached in August 2009 (rainy season). From these results we conclude that for Dahra the LSA SAF LST product meets its target accuracy of 2°C only during the dry season.

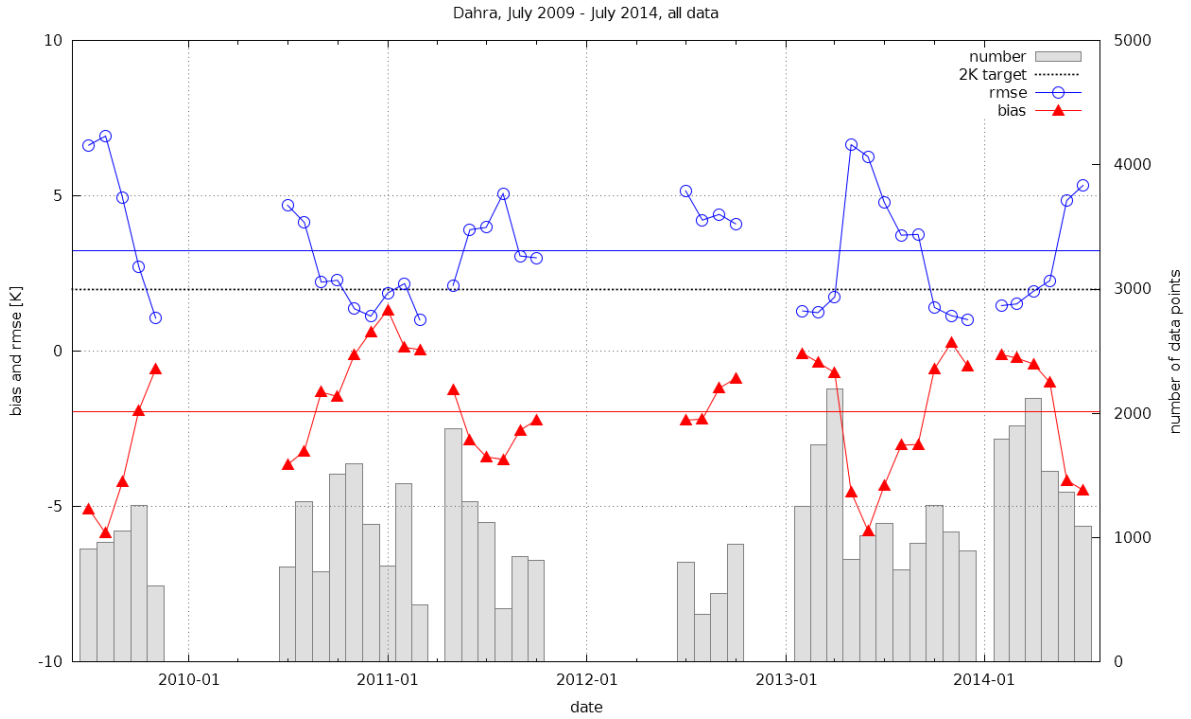


Figure 36 Daytime and night-time monthly statistics at Dahra station, Senegal, for LSA SAF LST. Mean bias (red triangles) and rmse (blue circles) refer to the left y-axis, the number of match-ups (grey bars) to the right y-axis.

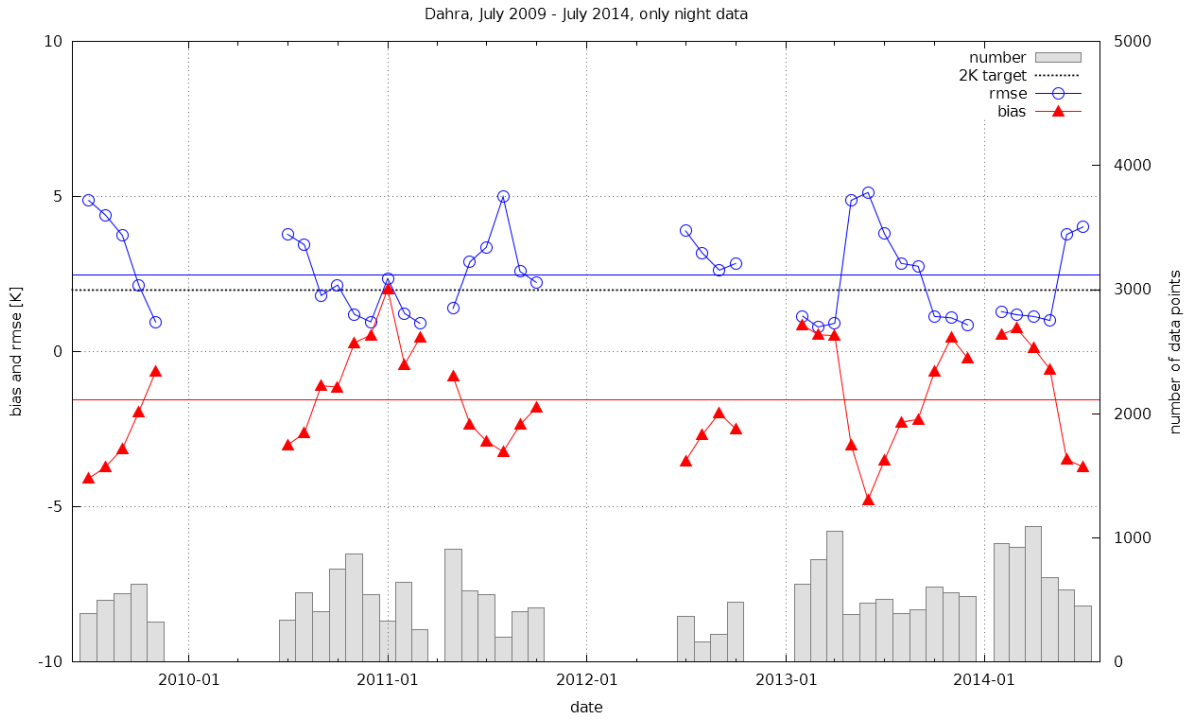


Figure 37 Night-time monthly statistics at Dahra station, Senegal, for LSA SAF LST. Mean bias (red triangles) and rmse (blue circles) refer to the left y-axis, the number of match-ups (grey bars) to the right y-axis.

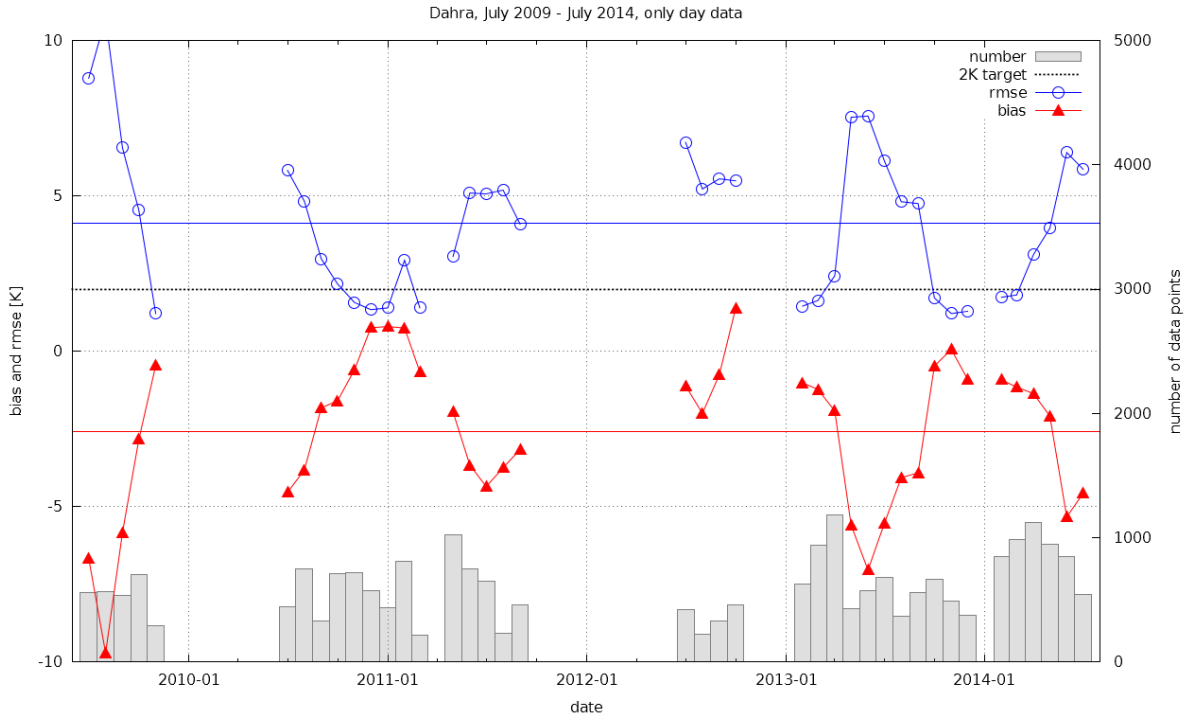


Figure 38 Daytime monthly statistics at Dahra station, Senegal, for LSA SAF LST. Mean bias (red triangles) and rmse (blue circles) refer to the left y-axis, the number of match-ups (grey bars) to the right y-axis.

Table 8 summarises the validation results for LSA SAF LST for validation stations Dahra, Gobabeb and farm Heimat. It can be seen that the LSA SAF LST product achieves its target accuracy of 2°C rmse for Gobabeb and Heimat, regardless whether LST are validated as night and day time data separately or together. For Dahra the results show that LSA SAF achieves its target accuracy during the dry season, but not during the wet season. However, when limiting the analysis for Dahra to dry seasons (Table 8, values in brackets) the results are comparable to those found for Gobabeb and Heimat.

Table 8 Mean multi-annual biases and root mean square errors for Dahra, Gobabeb, and farm Heimat. The values in brackets for Dahra give the results for dry seasons only (November-April).

Station	**Bias [°C]**			**RMSE [°C]**		
	All data	Night	Day	All data	Night	Day
Dahra	-2.0 (0.0)	-1.5 (0.4)	-2.6 (-0.6)	3.2 (1.4)	2.5 (1.1)	4.1 (1.7)
Gobabeb	0.1	-0.5	0.6	1.6	1.4	1.8
Heimat	0.1	0.7	-0.5	1.2	1.3	1.2

3 AVHRR/Metop LST (ESLT, LSA-002)

The LSA SAF is processing LST from Metop/AVHRR: product ELST, LSA-002. The ELST/LSA-002 product (hereafter LST_AVHRR) consists of daily composites of LST values retrieved from individual AVHRR/Metop orbits (or more specifically from Product Distribution Units, PDUs). The PDU-based values are then organized by night-time and daytime values (according to the respective solar zenith angle) and projected over a sinusoidal grid. All daytime (night-time) retrievals overlapping the same grid-point in the sinusoidal grid are then averaged; the same procedure is applied to observation time and viewing angle, as described in the Product User Manual (PUM_LST). Only the primary Metop satellite is processed; in the period considered in this Validation Report (January 2015 – November 2016), ELST was processed from Metop-B.

The validation of ELST presented in this Validation Report covers the period between Jan 2015 and Nov 2016 and includes both, inter-comparison with other satellite LST products (SEVIRI/MSG) and with data gathered at the in situ stations maintained by KIT (Figure 1).

The identification of cloudy pixels is an essential process of the LST generation chain. To optimize resources within the SAF Network, the LSA SAF relies on cloud mask algorithms developed by the NWC SAF team, and their respective validation results also published by the NWC SAF (www.nwcsaf.org/scidocs/Documentation/NWC-CDOP2-PPS-SMHI-SCI-VR-Cloud_v1_0.pdf). The Probability of Detection (POD) of the cloud mask when compared with synops (as reported by NWC SAF) is generally of the order of 94%. However, the POD varies with illumination conditions, and decreases to ~90% for twilight conditions. These values suggest that there are still a number of cloudy pixels which may not be masked and users should be made aware of this. The most frequent cases of missed cloudy pixels actually occur in the vicinity of identified clouds; the LST validation exercise, for both SEVIRI/MSG and AVHRR/Metop products, takes into account that cloud contamination may occur in those cases.

3.1 Intercomparison of LST_AVHRR and LST_SEVIRI

Given the maturity of the LST_SEVIRI product and the wide range of validation exercises performed for that product (please see previous sections in this report), the consistency of LST_AVHRR will be assessed taking into account a thorough comparison with LST_SEVIRI (LSA-001) product. The high frequency of LST_SEVIRI product (15 minutes) increases the chances of getting good matchups between the two satellites. In contrast, collocation in space and time between polar-orbiter observations is much more difficult, and not included in this report.

The LSA-001 and LSA-002 products were collocated in space and time according to the following criteria:

- ELST (LST) is re-projected onto the geostationary grid (3-km at sub-satellite point), by averaging 3 x 3 LST_AVHRR values centred within each SEVIRI pixel. The resulting LST_AVHRR field is smoother than that obtained with a nearest neighbour approach; using a 9 km^2 area insures that the AVHRR LST are within all SEVIRI pixels (which become larger towards the edge of the disk).
- The difference in time between LST_SEVIRI and LST_AVHRR is 7.5 minutes or lower.

The comparison between the two LSA SAF products is performed for a set of 10º longitude x 10º latitude areas indicated in Figure 39, which were chosen to: (i) include a wide range of biomes and different atmospheric conditions within the Meteosat disk; (ii) include in situ stations used to validate LST_SEVIRI.

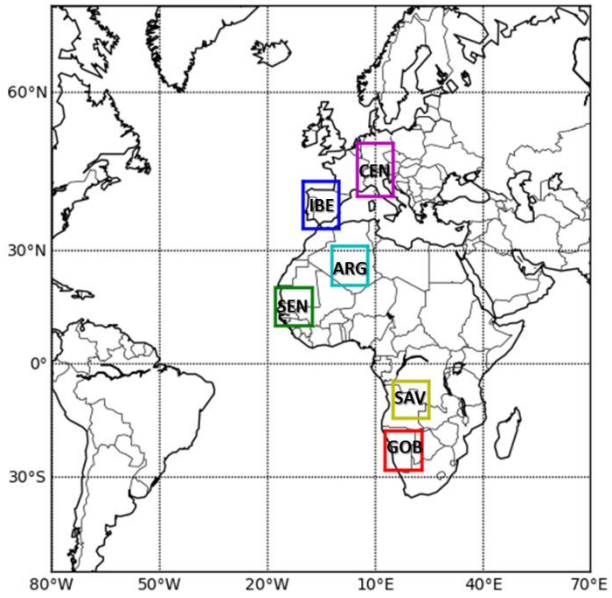


Figure 39 Areas (10º longitude x10º latitude) considered for comparison between LST_AVHRR and LST_SEVIRI, centred in the following locations: Iberian Peninsula (IBE) 5ºW, 40ºN, which include Évora station; Central Europe (CEN) 10ºE, 45ºN; Algeria (ARG) 3ºE, 26ºN; Senegal (SEN) 12.5ºE, 15.5+N, which includes Dahra station; Souther Africa /Savannah (SAV) 20ºE, 10ºS; and Namibia (GOB), including Gobabeb and Farm Heimat stations 18ºE, 23.5ºS.

As a general illustration of the spatial variability of AVHRR and SEVIRI LST fields, we show in Figure 40 to Figure 43 the corresponding monthly values collocated in space and time, together with the respective difference and root mean square difference. The fields are presented for night and daytime observations respectively, and for two contrasting months – January and July 2016. The results suggest that discrepancies among SEVIRI and AVHRR retrievals are within the expected range (mostly within 2ºC or 2.5 ºC). However, and for any given point, these are also shown to vary with time of the day and period of the year, while within each region considered, AVHRR – SEVIRI differences vary with local topography and land cover.

The seasonal fluctuations in the LST_AVHRR – LST_SEVIRI differences are put into evidence in Figure 44 and Figure 45. These results cover nearly two years of data, which show very similar seasonal variability, indicative of the stability of the retrievals. As indicated above, night-time LST estimates present better agreement and lower seasonal variability than daytime ones. This is not surprising, since daytime LST retrievals are subject to more pronounced directional effects associated with differences in view-illumination geometries seen by each sensor and to contrasts in the temperature of surface elements within each pixel, both evolving with time of the year. As such, the seasonal variability of LST_AVHRR – LST _SEVIRI differences is smother in areas where the surface tends to be homogeneous, both in terms of land cover and orography (see, e.g., the case of Algeria in North Africa, Figure 45).

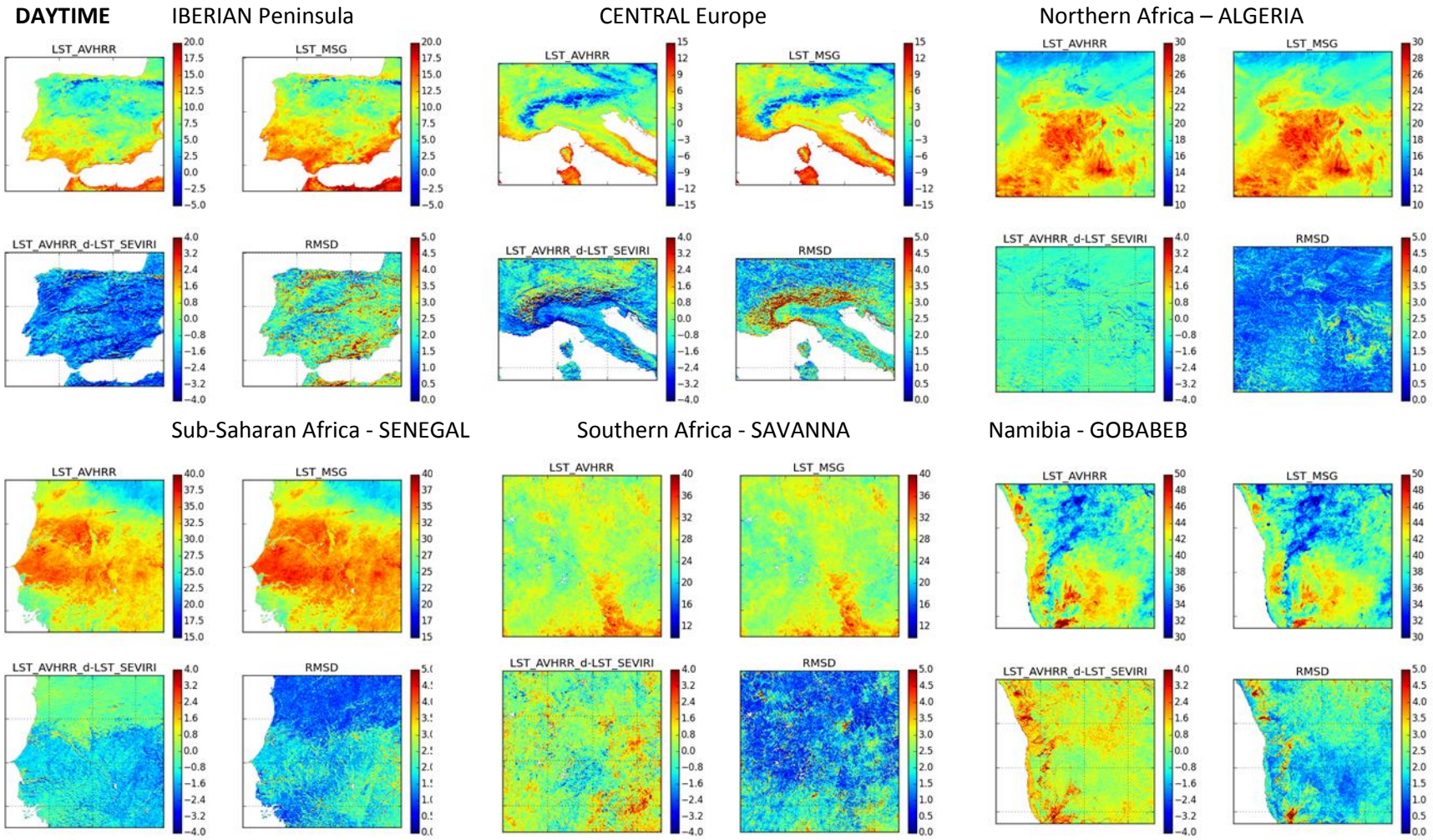


Figure 40. Statistics gathered for January 2016 for each of the areas indicated in Figure 39, corresponding to (clockwise from top left): averaged LST_AVHRR daytime retrievals; averaged LST_SEVIRI retrievals matching Metop/AVHRR estimates (in space and time); mean difference (LST_AVHRR – LST_SEVIRI); and root mean square of the differences. All values are in ºC.

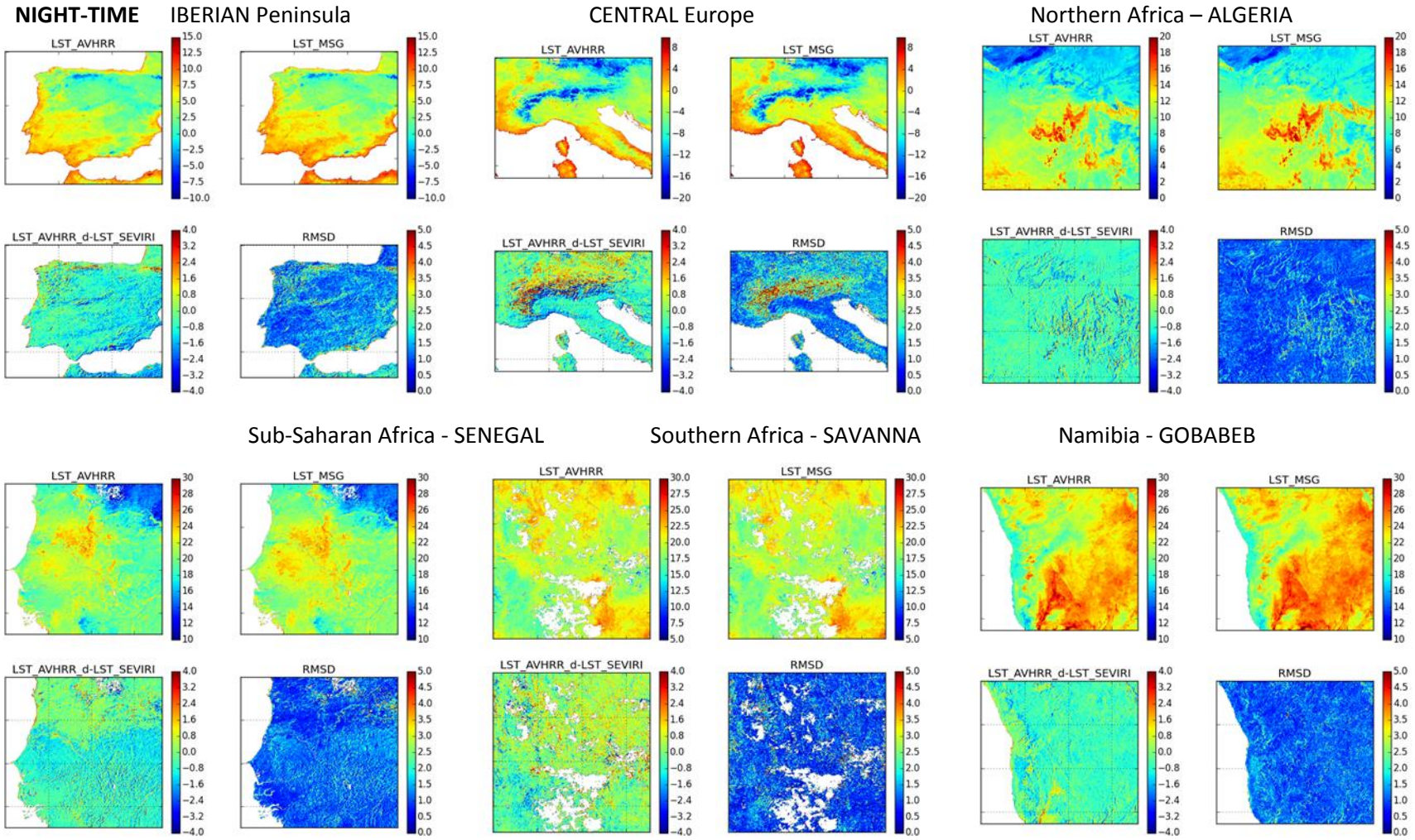
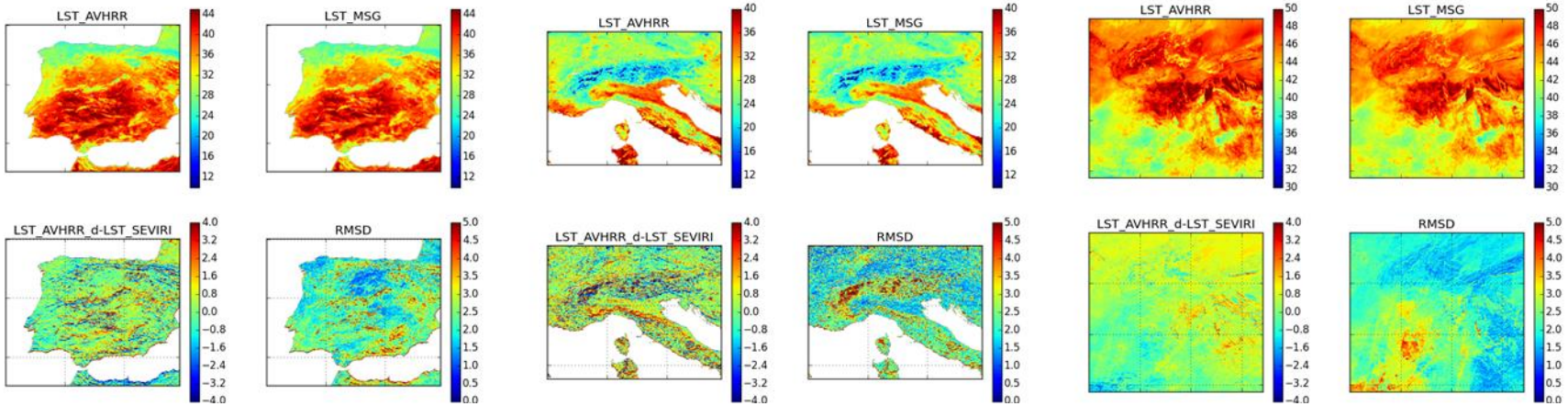


Figure 41 As in Figure 40, but for Metop/AVHRR night-time retrievals in January 2016.

DAYTIME IBERIAN Peninsula CENTRAL Europe Northern Africa – ALGERIA



Sub-Saharan Africa - SENEGAL Southern Africa - SAVANNA Namibia – GOBABEB

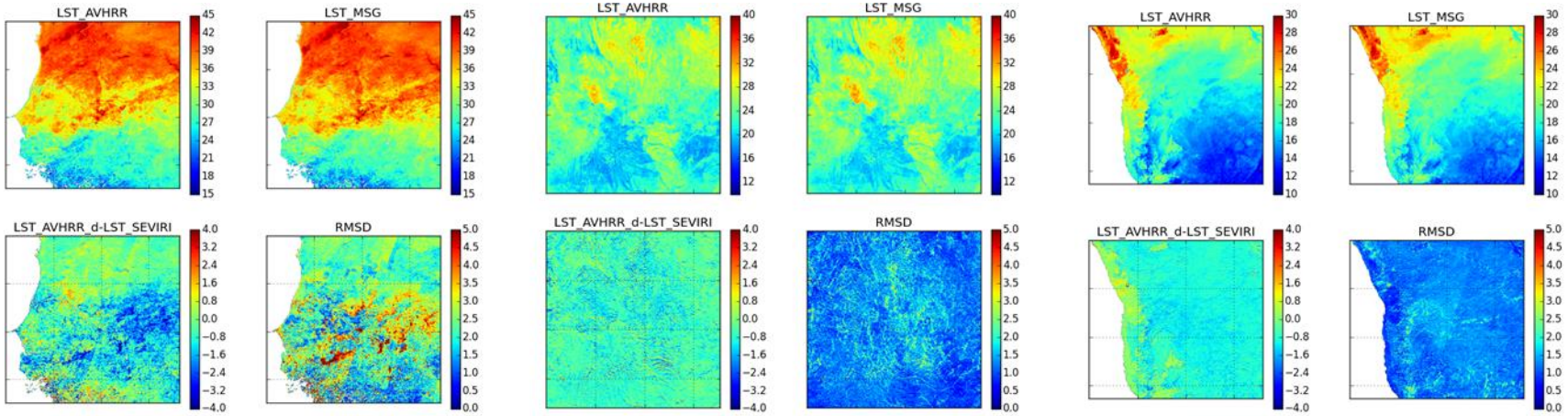


Figure 42 As in Figure 40, but for Metop/AVHRR daytime retrievals in July 2016.

NIGHT-TIME IBERIAN Peninsula CENTRAL Europe Northern Africa – ALGERIA

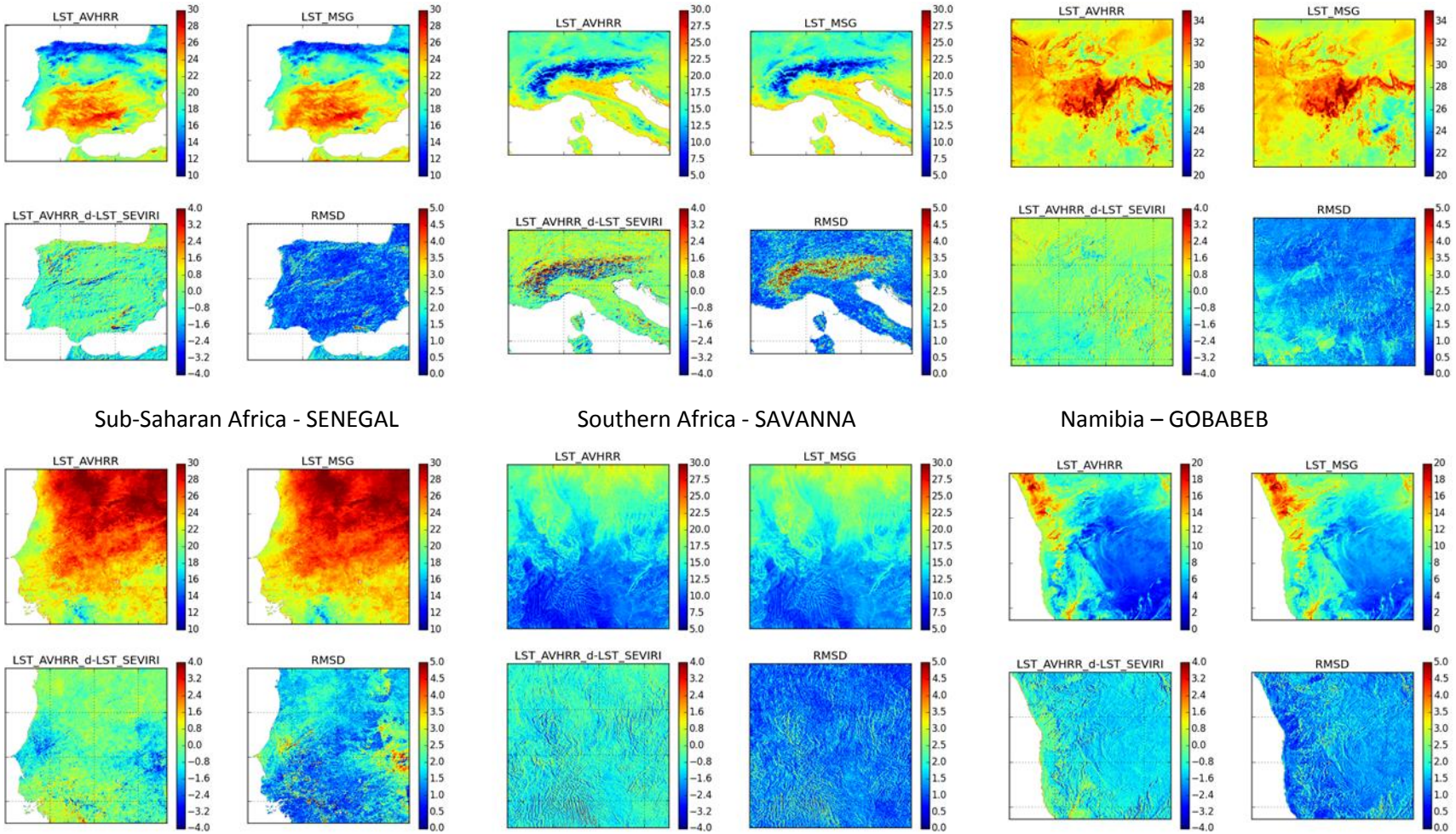


Figure 43 As in Figure 40, but for Metop/AVHRR night-time retrievals in July 2016.

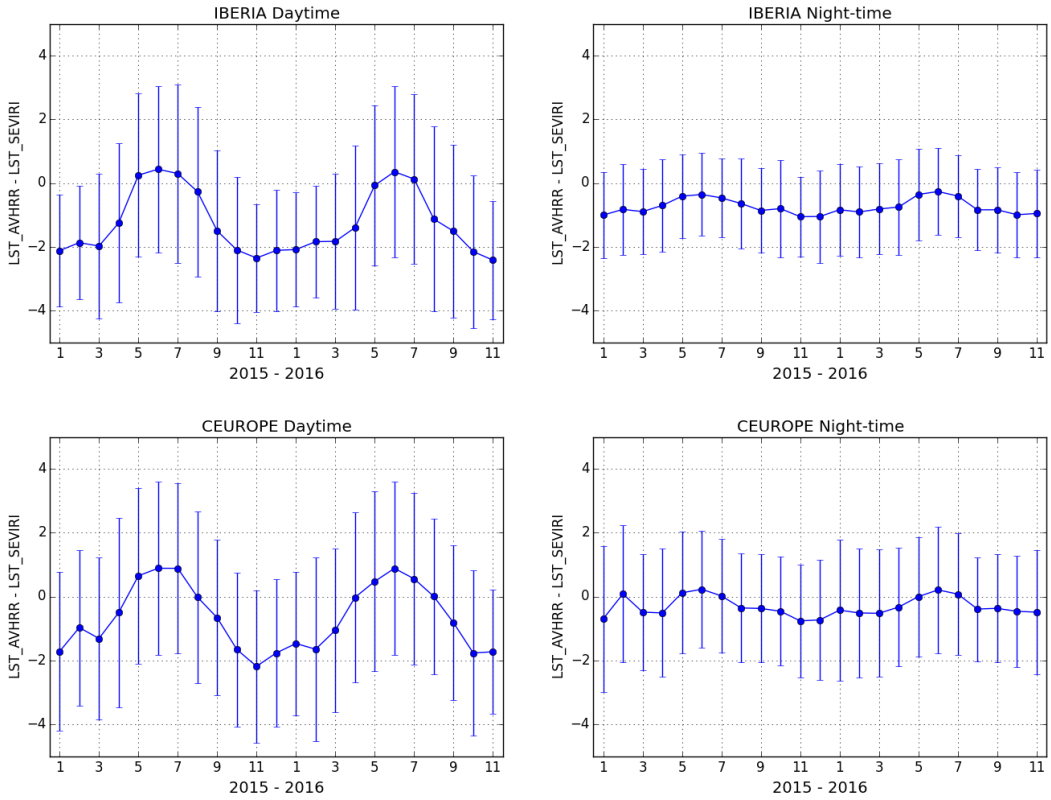


Figure 44 Monthly average (dots) and standard deviation (error-bars) of the differences between LSA_AVHRR and LST_SEVIRI (ºC) for the period between January 2015 and November 2016, for the Euroepan areas indicated at the top of each panel. Left and right panels show the statistics for daytime and night-time, respectively.

The average of night-time statistics are the most reliable to assess the agreement between the two products since they are far less dependent on viewing geometry than daytime observations. The monthly average of night-time LST differences, considering all the regions shown in Figure 39, ranges between -1.4ºC (Senegal, April 2015) and +0.2ºC (Algeria , August 2015 and Southern Africa/Savanna, February-March, 2016), while the standard deviation of night-time differences lies around 1.0 ºC for most cases, reaching values over 2ºC (maximum of 2.4 ºC) for some particular cases, most notably: regions and periods of the year where the spatial variability is particular high, even at night-time and when the probability of pixels being contaminated by clouds is highest, as is the case of Central Europe in winter and West Africa / Senegal region during the wet season (particular in June – July). The fact that LST_AVHRR analysed here still uses static emissivity fields explains part of the seasonal variability in the differences to SEVIRI LST, particularly in regions where changes in vegetation, and therefore in emissivity are more pronounced.

In order to have further insight into the impact of viewing geometry on LST retrievals, we show average and standard deviation of AVHRR and SEVIRI LST differences for a number of selected areas and months, with statistics estimated for different classes of AVHRR/Metop view zenith angle. In this exercise we consider the SEVIRI viewing geometry to be roughly constant within each 10ºx10º area. The statistics are estimated for night-time and daytime observations, separately.

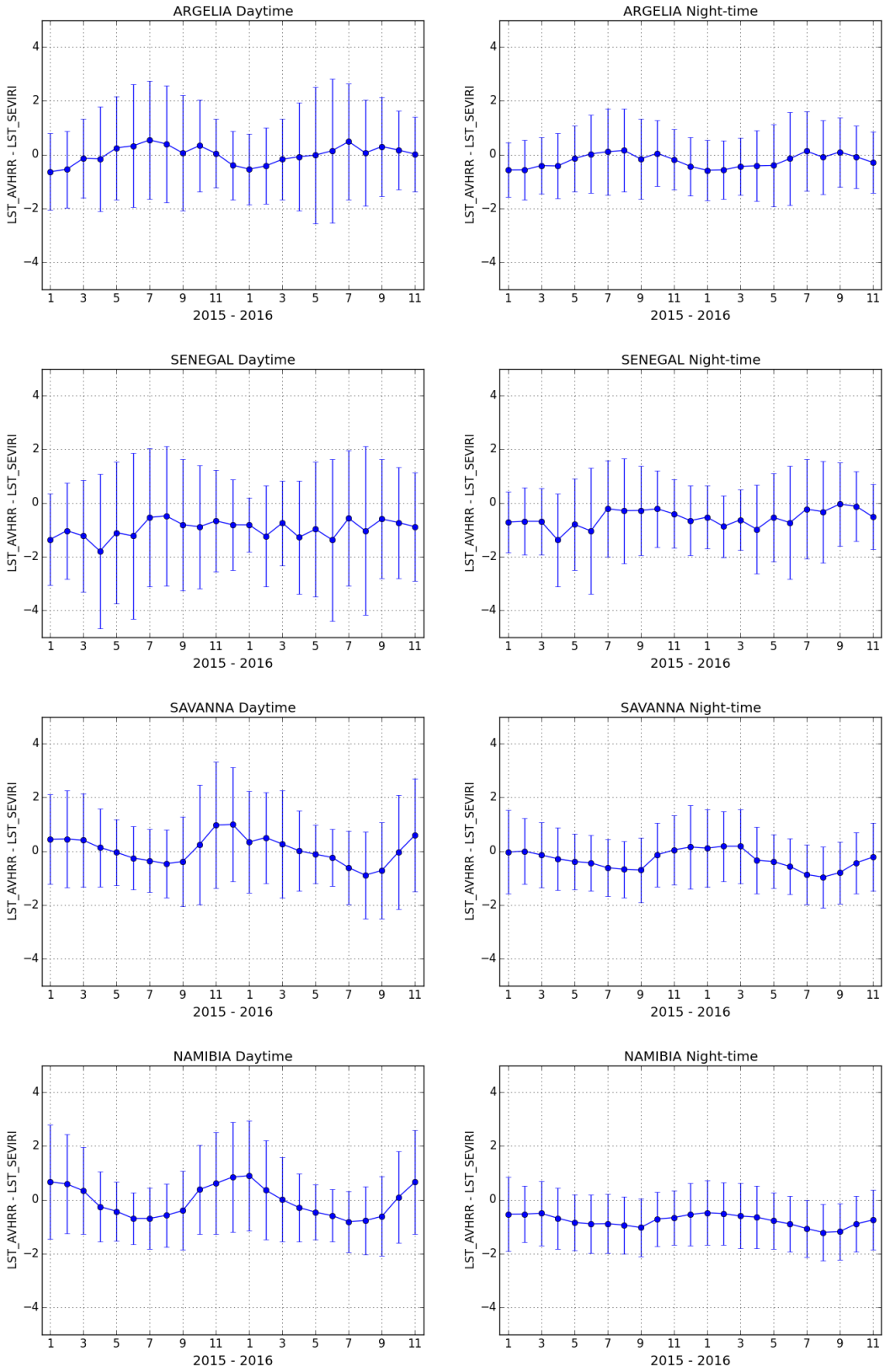


Figure 45 As in Figure 44, but for the areas located in Africa (see top of each panel).

January 2016 April 2016

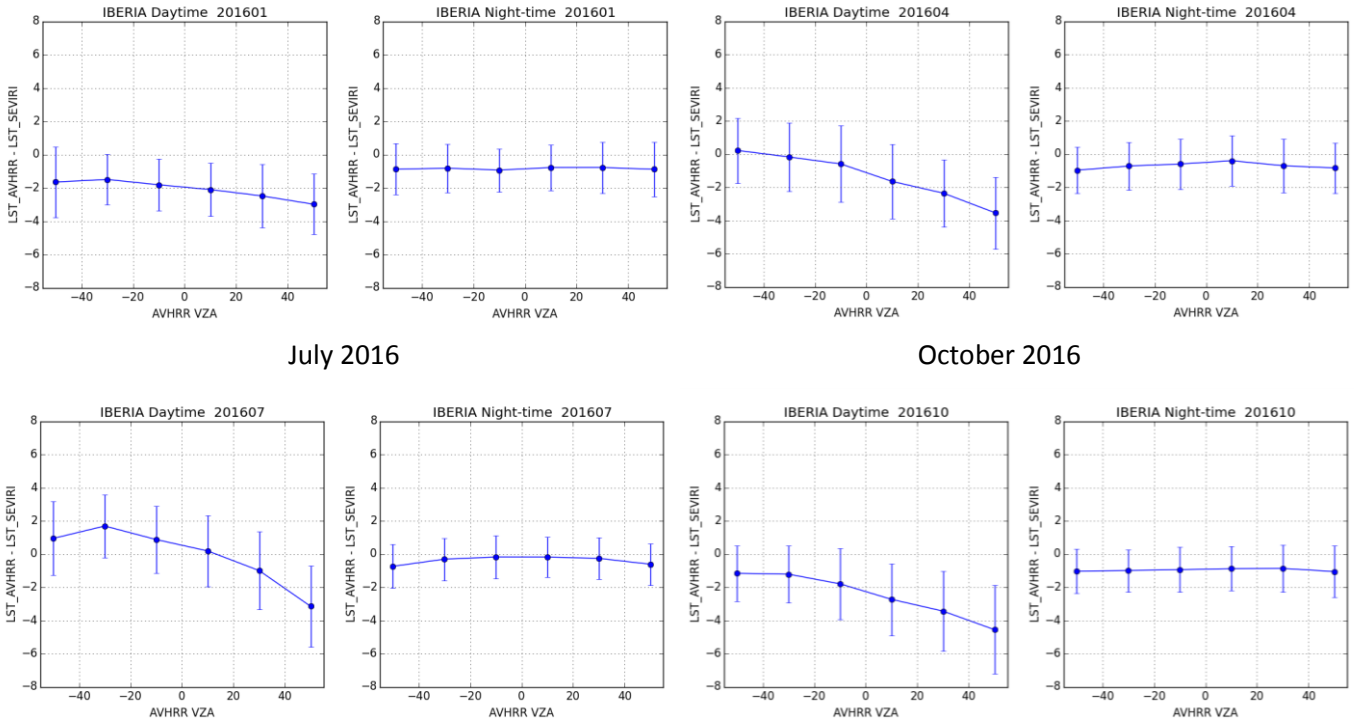


Figure 46 Average (dots) and standard deviation (error-bars) of differences of [LST_AVHRR – LST_SEVIRI] (ºC) for the Iberian region, grouped by AVHRR view zenith angle (vza), taking into account the following classes: vza below -40º, vza between -40º and -20º, vza between -20º and nadir, vza between nadir and 20º, vza between 20º and 40º, and vza above 40º. The data are displayed for January, April, July and October 2016, for daytime and night-time, respectively (please see panel title).

Figure 46 presents the LST statistics per AVHRR view zenith angle (vza) for the Iberian region. In order to get a reasonable overview of the changes over the year, the results are shown for January, April, July and October 2016. Positive AVHRR view zenith angles indicate the scene is viewed from west, while negative values indicate the scene is view from east. The LST differences are nearly insensitive to AVHRR vza in case of night-time observations, while daytime cases present the signature of the viewing and illumination geometries. Keeping in mind that Metop overpass occurs during mid-morning (close to 10 a.m. local time), [LST_AVHRR – LST_SEVIRI] are higher when AVHRR/Metop retrievals are taken from East (vza < 0) and therefore facing a higher fraction of illuminated surfaces; the opposite occurs for AVHRR/Metop estimates taken from the West view (vza > 0). The higher contrast in surface temperatures occur in summer months, and therefore the impact of directional effects is highest in July. When compared to winter, when the land surface over Iberia tends to be general more homogeneous and therefore the standard deviation of LST differences is smaller in daytime January, when compared to those in daytime July. In summer, dry soils allow for the sunlit ground to get much warmer than the surrounding canopies – see section on LST_SEVIRI validation at Évora station (2.2). The results shown in Figure 46 are consistent with those obtained by Trigo et al (2008b) when assessing SEVIRI and MODIS LST products.

Although only 2016 is shown in Figure 46, the results for 2015 are very similar, as suggested by the stability of the inter-annual statistics shown in Figure 44 and Figure 45. It should also be noted that, April and October present intermediate results to those obtained for January and July. This is also the case for all the areas considered in the report, and therefore, only the statistics for the latter two months will be shown in the figures below.

January 2016 Central Europe July 2016

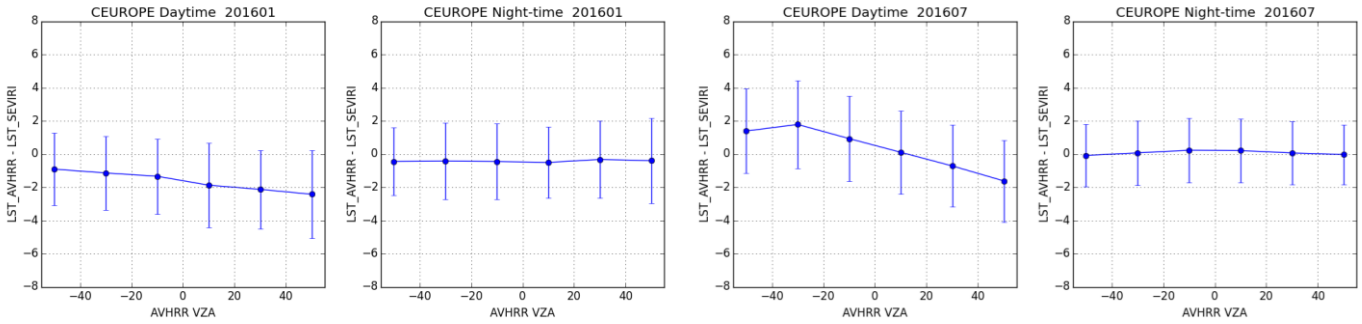


Figure 47 As in Figure 46, but the Central European region. Only January and July 2016 statistics are shown (2 left and 2 right panels, respectively), again separated by daytime and night-time observations (see top of each panel).

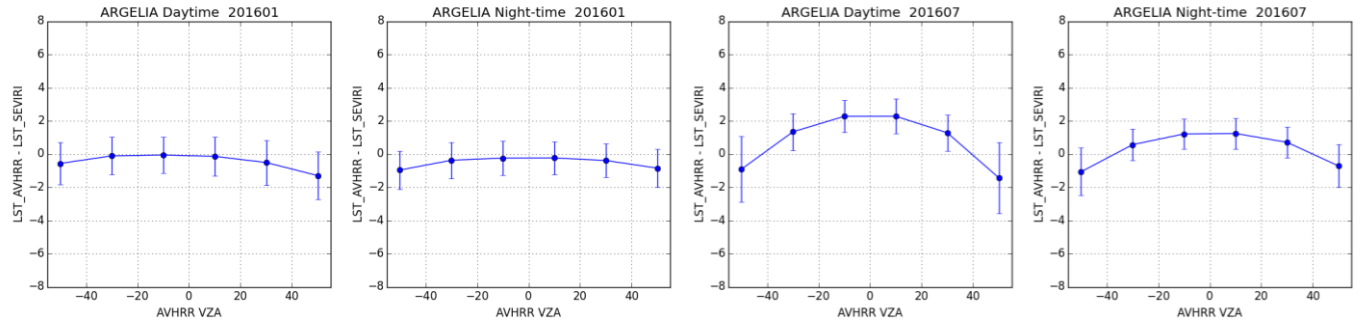
The results shown for most areas reveal the same pattern as that described for the Iberia Peninsula. It is worth looking at the results for Algeria (Northern Africa), which show a nearly symmetric behavior in LST differences, particularly pronounced in daytime July. A similar, but somehow smoother pattern is seen in night-time LST differences in July and to a lesser extent in January statistics. In this case, the directional effects are likely to be associated to changes in emissivity with angle, which are currently being ignored in LSA SAF LST retrievals (the only exception is for snow/ice emissivity, where we consider tabulated values per gross classes of view zenith angle). In fact, over barren surfaces, such as those observed in Northern Africa, emissivity may decrease for high viewing angles (typically for vza above 40º, see e.g., García-Santos et al., 2012). Since this effect is not taken into account, LST_AVHRR retrieval tend to decrease towards the edge of the scan. As a result we get a very different angular dependency from the asymmetric pattern seen for Iberia and Central Europe.

The barren / desert area, which is still in the northern part of the West Africa/ Senegal region, explains why it also shows a similar behavior of LST_AVHRR – LST_SEVIRI differences to that described for Algeria.

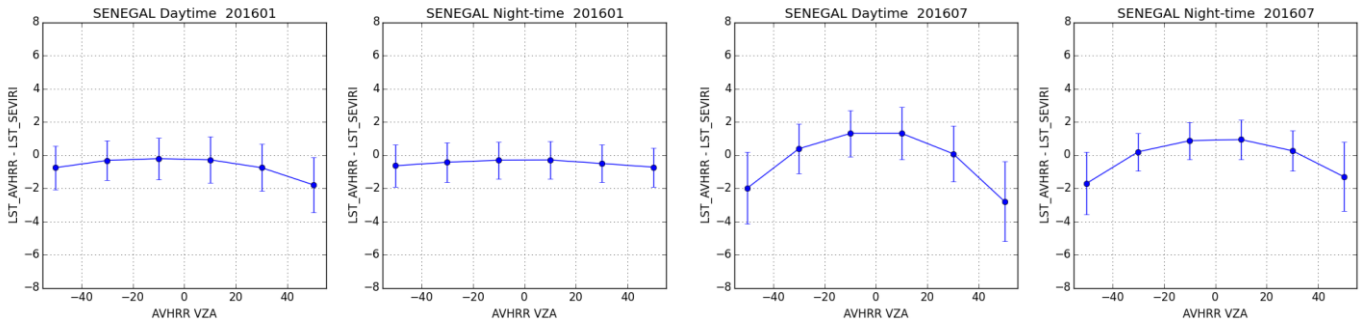
The directional characteristics seen in all infra-red based LST products, result from a blend of different effects. The most relevant in most cases arises from a combination of surface heterogeneity and view-illumination geometries. However, under certain conditions, namely bare surfaces in deserts (particularly rocky surfaces) and snow/ice surfaces, emissivity may change significantly for high view zenith angles. The emissivity variation is however not straight-forward, and therefore difficult to take into account in operational retrievals.

Land cover and surface orography play an important role in determining surface temperature heterogeneity, and therefore on the impact of directional effects related to view-illumination geometries.

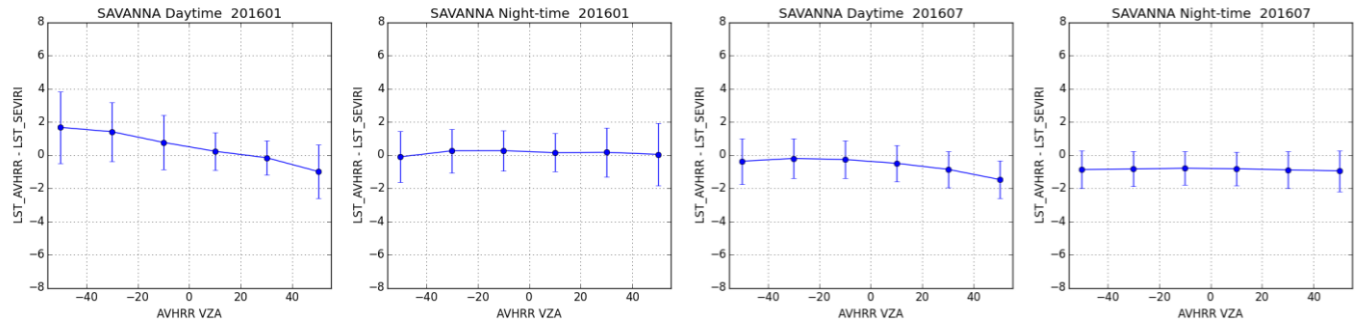
January 2016 Algeria July 2016



West Africa / Senegal



Southern Africa / Savanna



Southern Africa / Namibia

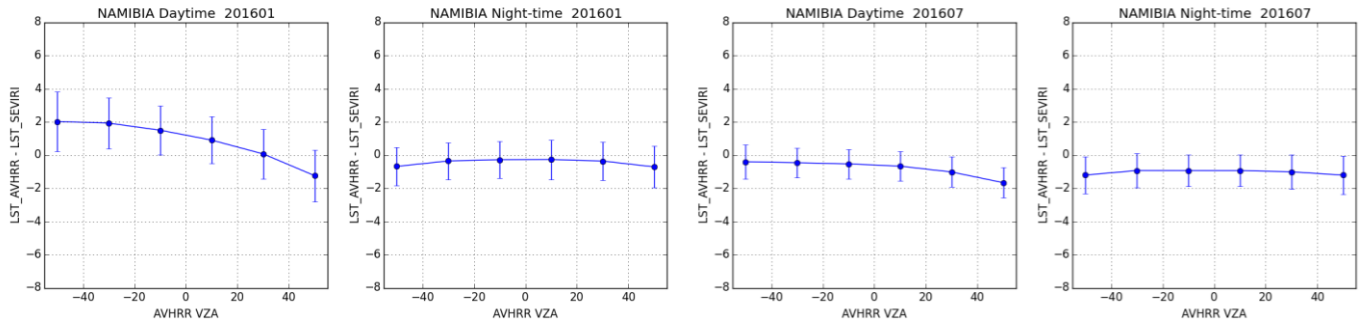


Figure 48 As in Figure 47, but for the African regions indicated at the top of the panels.

January 2016 Iberia July 2016

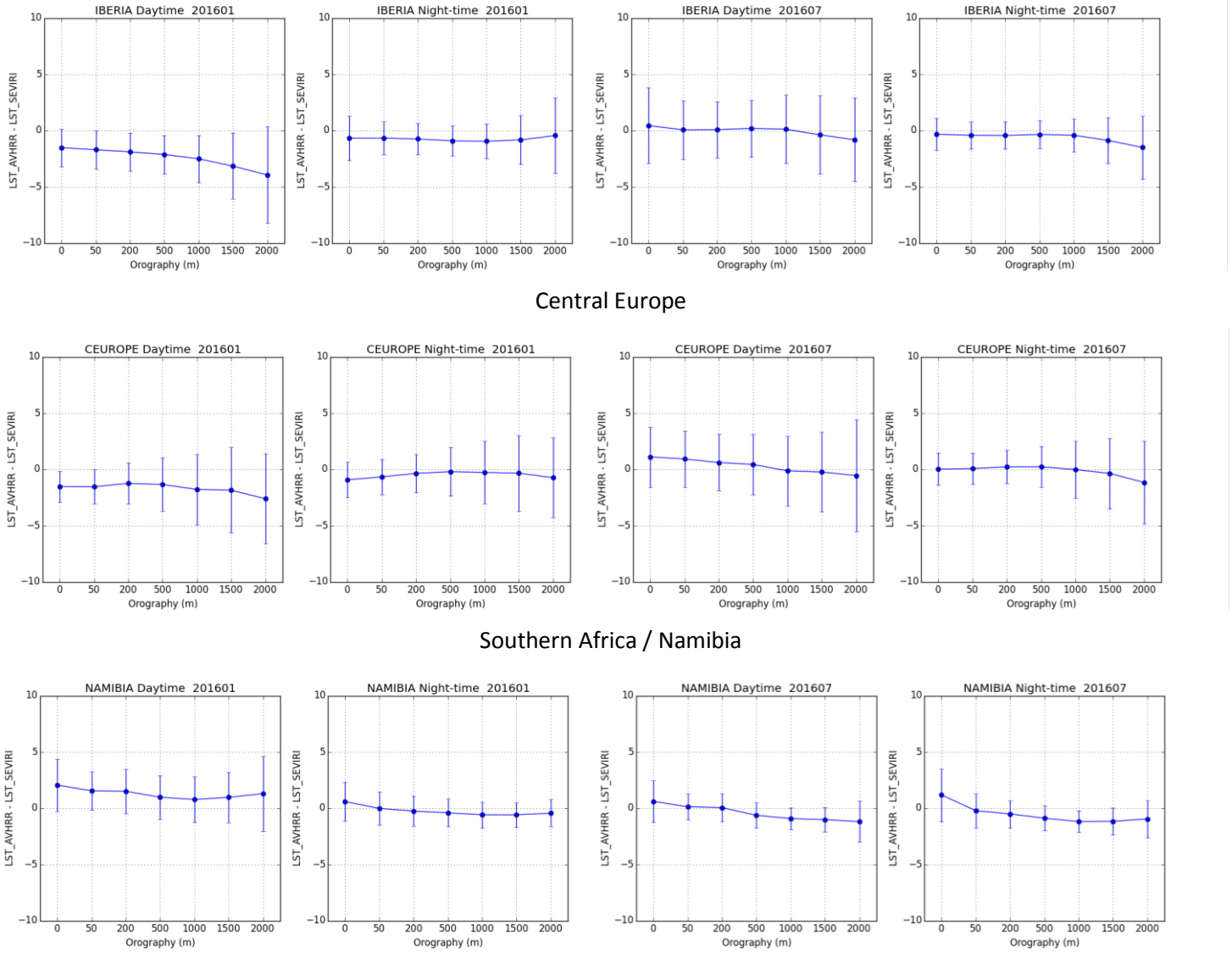


Figure 49 Average (dots) and standard deviation of the difference between LST_AVHRR and LST_SEVIRI (ºC) for the areas and months indicated in the top of each panel, for classes of surface elevation. The x-axis indicates the lower limit (m) of each elevation class.

In general, the average and standard deviation of differences between SEVIRI and AVHRR LST products increase with pixel elevation, regardless of season and for both daytime and night-time observations (Figure 49). Higher elevation areas generally corresponds to irregular and potentially highly heterogeneous landscapes. Daytime observations will be affected by terrain illumination and shading, while both daytime and night-time observations may correspond LST measured at different height – either because of the sensor footpring, or due to uncertainties in the geolocation. While the sign of the [LST_AVHRR – LST_SEVIRI] difference depends on local effects, including the variation of LST with height, the area heterogeneity contributes to an increase in their standard deviation with height, seen in Iberia and Central Europe regions.

The Namibian region corresponds to a slightly different case: the area is dominated by a plateau (above 1000 – 1200 m) and the steepest terrain lies either near the coast (0 – 100 m), or near the highest peaks (above 1800 m).

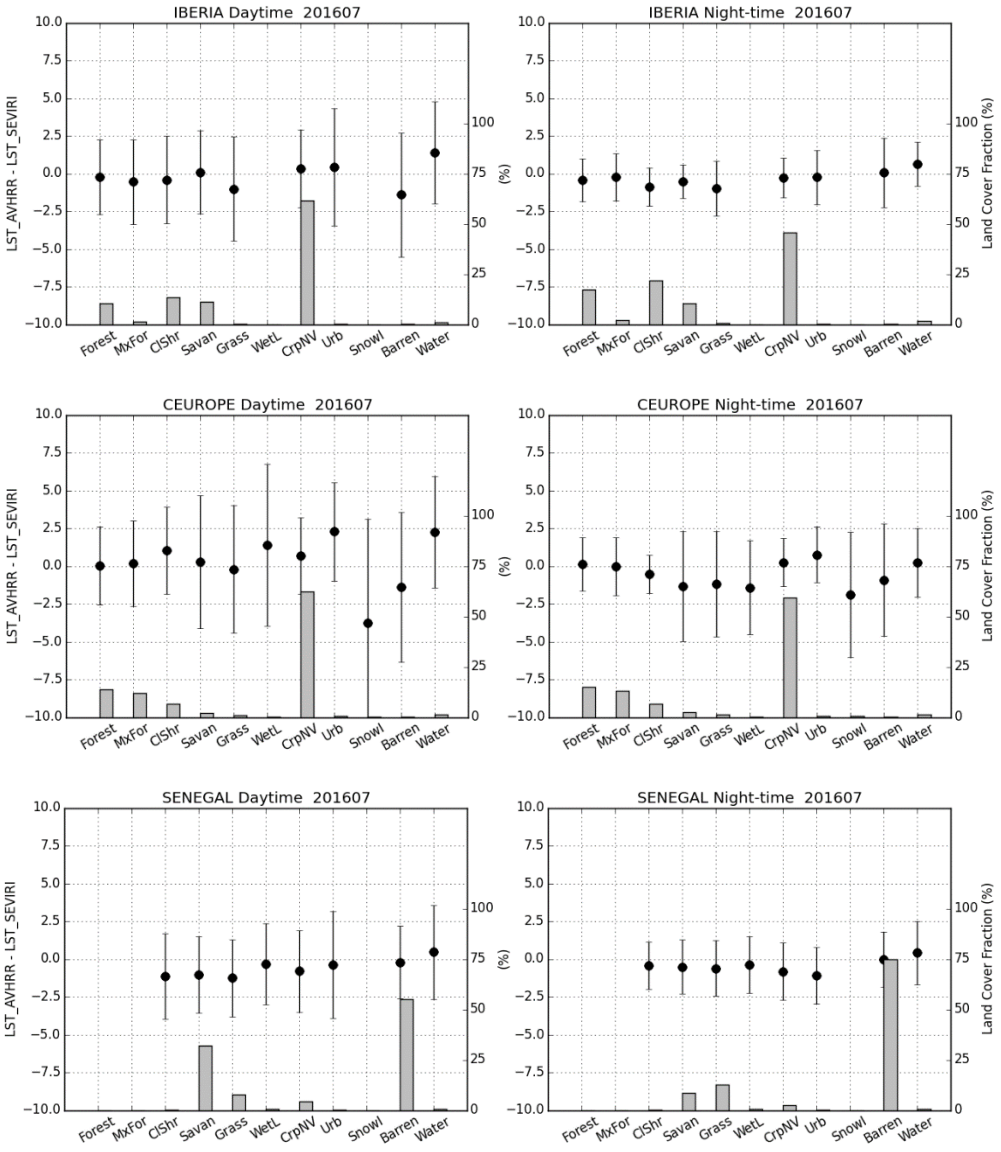


Figure 50 Average (dots) and standard deviation of the difference between LST_AVHRR and LST_SEVIRI (ºC) for Iberia (top), Central Europe (middle) and West Africa/Senegal (bottom), for July 2016. The statistics are estimated per land cover class (x-axis). The bars indicate the percentage of the land cover class per each 10ºx10º area (right axis) with (clear sky) LST estimations.

For completeness, Figure 50 shows the statistics of AVHRR – SEVIRI LST comparison per land cover, within some of the regions identified in Figure 39. Only areas with a relatively high variety of land cover types were selected; the results are shown for July 2016, which coincides with the period of the year when the variability of LST differences is higher. Forests and closed shrublands present relatively low variability. The standard deviation tends to be higher for less frequent land covers type, and therefore with lower statistical significance. It is worth noting the Snow/Ice case in Central Europe, which also coincides with high elevated areas referred before, and the wetlands and inland water case, corresponding often to mixed pixels near rivers or lakes. Both of these tend to present high standard deviation in [LST_AVHRR – LST_SEVIRI].

3.2 Intercomparison of LST_AVHRR and In Situ Observations

This section presents the comparison of LST_AVHRR and in situ measurements (see section 2.1). We consider observations gathered at stations maintained by KIT within the context of the LSA SAF, for the specific purpose of LST validation, overlapping with the available LSA-002 product – please see Table 9. The stations location is shown in Figure 1.

Table 9 In situ observations at KIT Stations overlapping with currently processed ELST (LSA-002) product. The average difference between LST_AVHRR and in situ observations (bias) and the respective standard deviation is also for daytime and night-time, and for each station.

Station	Period	Bias (ºC)		StDev (ºC)	
		Day	Night	Day	Night
Évora (South Portugal)	Sep 2015 to Oct 2016	-3.2	-0.8	2.8	1.8
Gobabeb (Namibia)	Jan 2015 to Sep 2016	2.8	-1.3	2.0	1.3
Heimat (Kalahari, Namib)	Feb 2015 to Sep 2016	-1.5	-1.7	1.7	1.1

The comparison of LST_AVHRR with the station observations is show in Table 9, and in Figure 51 and Figure 52. The results are consistent with those described before, revealing overall a better agreement and lower variability for night-time than for daytime observations.

The Evora station results are particularly affected by directional effects in daytime (see discussion in section 2.2). The large negative bias obtained for daytime estimates over that station seems to be affected by a few points in winter (see LST_AVHRR < 10ºC) which are cooler than in situ estimates – either a case of cloud contamination or spatial variability within the AVHRR pixel, not represented in the station point measurements. This effect is not observed in LST_SEVIRI estimates, and needs to be further assessed.

Overall LST_SEVIRI collocated with LST_AVHRR and station observations presents a better agreement with in situ measurements – both in terms of average differences and their standard deviation.

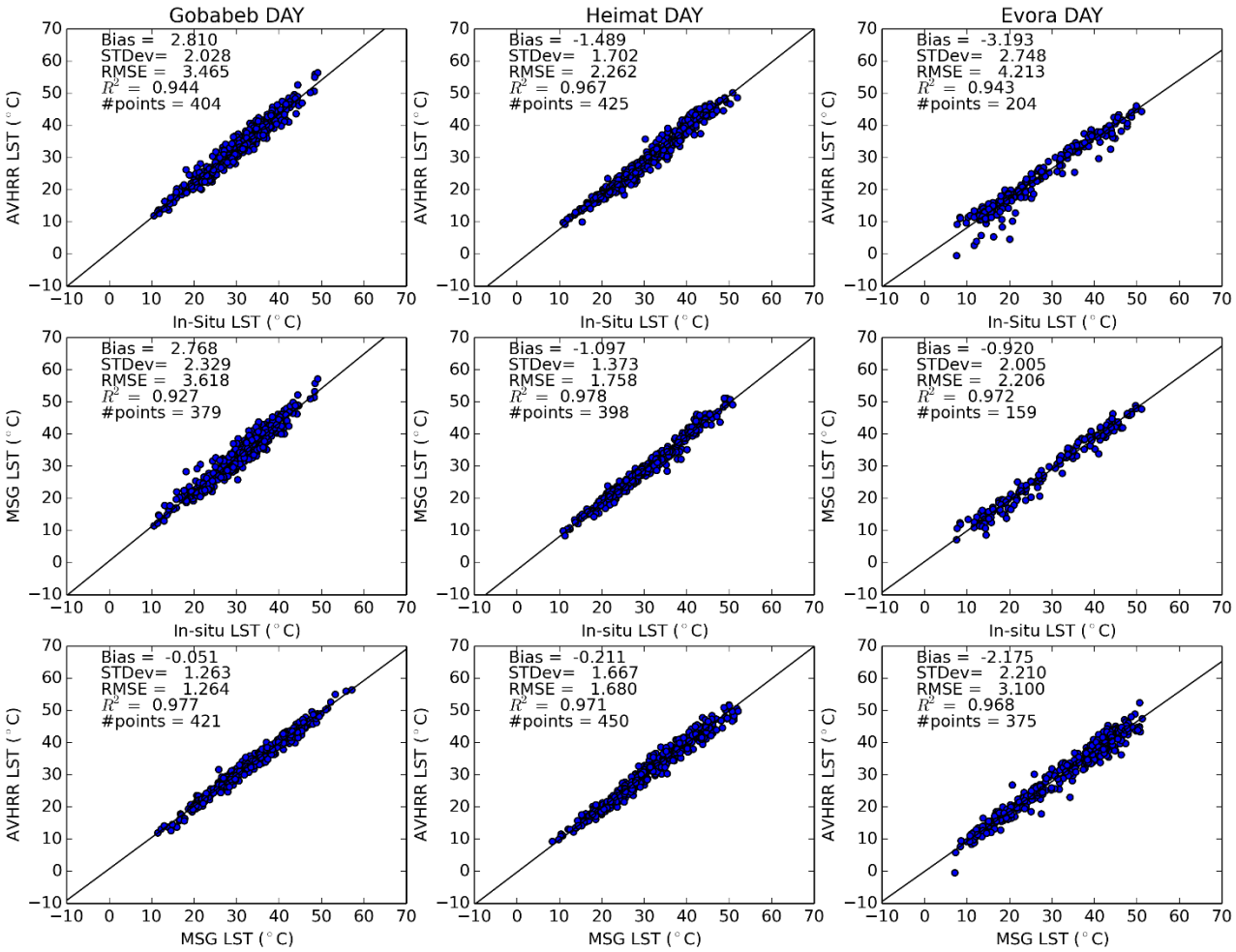


Figure 51 Scatterplots of daytime (from top to bottom): LST_AVHRR (ºC) versus in situ observations; LST_SEVIRI (collocated with LST_AVHRR values) versus in situ observations; and LST_AVHRR versus LST_SEVIRI estimates at the in situ stations. The results for Gobabeb/Namibia, Heimant/Namibia and Evora are shown in the first, second and third column, respectively.

Figure 53 presents a zoom of the satellite LST versus in situ comparisons for July 2016, which as seen before, corresponds to one of the periods where the variability of LST and of LST statistics is high. In Évora site, this also coincides with the period where angular effects on LST products are most pronounced. The AVHRR overpass in the morning overlaps with the steepest heating rate, which combined with variable AVHRR view zenith angles (i.e. with variable fraction of sunlit surfaces in the morning) results in the fluctuation of the green dots in daytime Évora with respect to the in situ ones (blue). SEVIRI has a fixed view angle, so although directional effects are relevant for LST_SEVIRI, the day-to-day variability is low.

Gobabeb is located in the most homogeneous site and therefore the effect referred above is not detectable there. Heimat/Namibia site is within a less homogeneous landscape, and day-to-day fluctuations in the AVHRR – in situ match are again visible.

The night-time LST_AVHRR retrievals show negative biases with respect to observations, which may be attributable to an overestimation of local emissivities, something that needs to be further investigated.

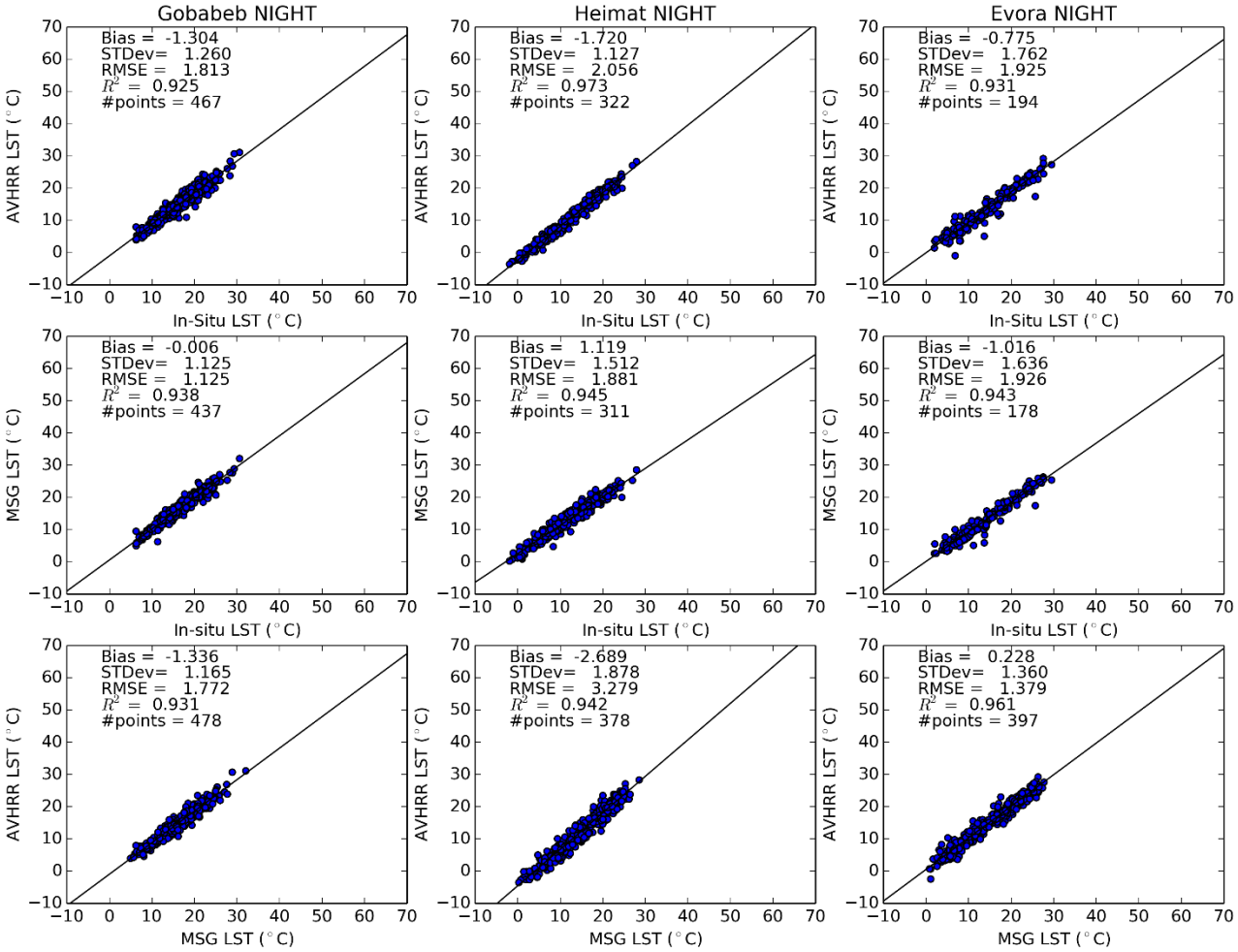


Figure 52 As in Figure 51, but for night-time observations.

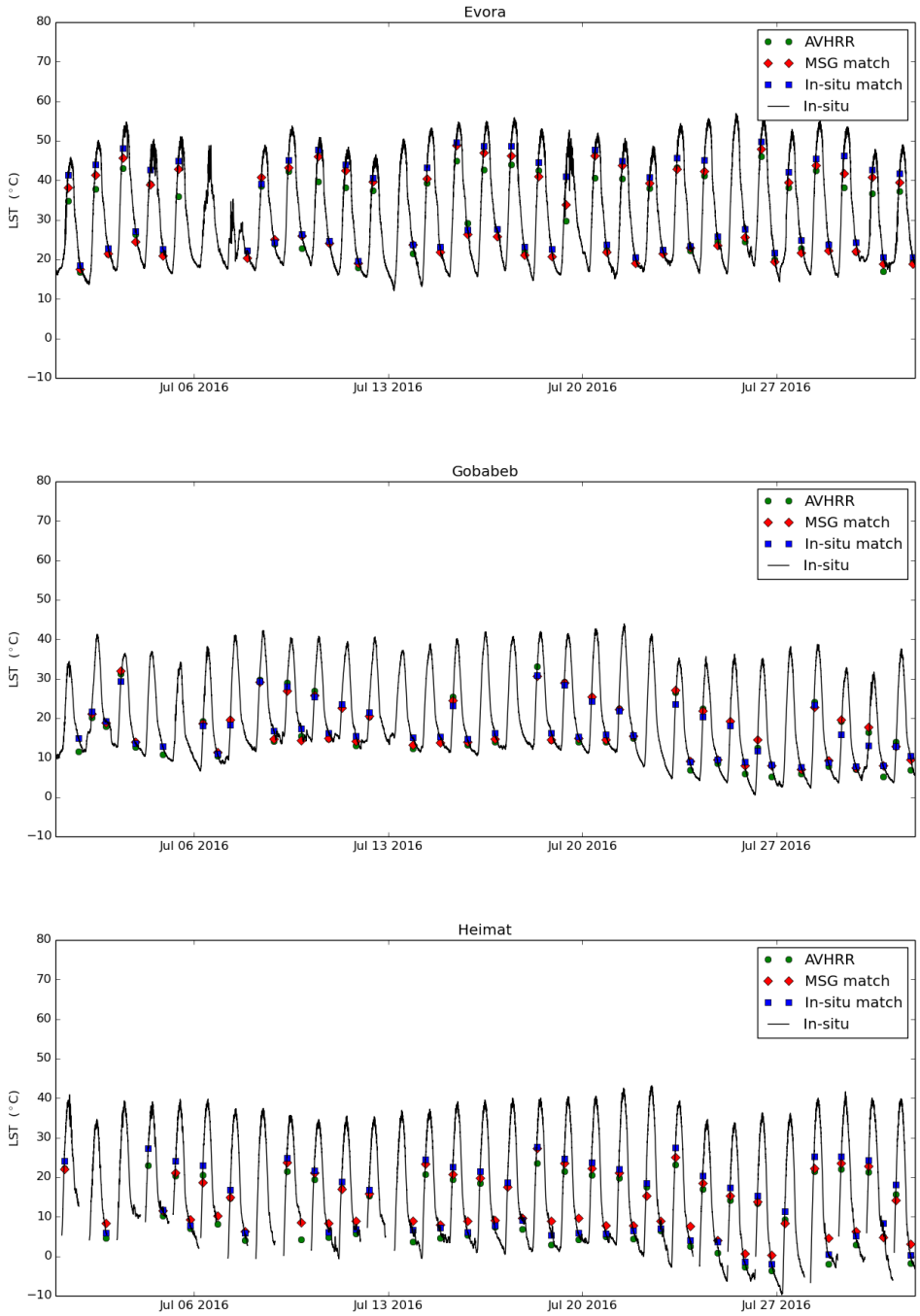


Figure 53 Time-series of in situ measurements and AVHRR (green dots) and SEVIRI (red dots; only retrievals collocated with LST_AVHRR) LST (ºC) estimates for July 2016. In situ measurements corresponding to AVHRR/Metop overpasses are highlighted as blue dots.

As a first exercise to assess the quality of ELST product over high-latitude regions, Figure 54 shows the comparison of LSA SAF LST_AVHRR product with in situ estimates obtained for the Baseline Surface Radiation Network (BSRN) station Ny-Ålesund located at 78.9ºN; 11.9ºE. The station was chosen taking into account the overlapping period between LSA SAF ELST product and the observations; here we analyze the January 2015 – March 2016 period. In situ LST is estimated

from measurements of up-welling and down-welling long-wave, considering a constant emissivity of 0.99. The comparison indicates LST_AVHRR underestimates local estimates, with an overall bias of –1.3 ºC. The standard deviation of the LST differences is close to 4ºC. The reasons for these discrepancies between in situ and satellite observations need to be further assessed. In particular, the evaluation of the station representativeness for pixel estimates with the full characterization of its surrounding is still to be performed. The evaluation of LST_AVHRR with further in situ data will be performed for the next release of the product.

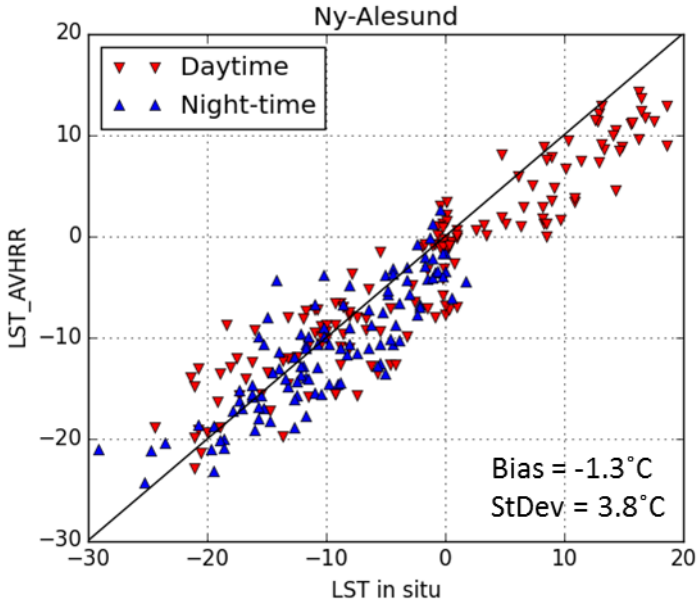


Figure 54 Scatterplot of in situ estimates of LST versus AVHRR LST over Ny-Ålesund station. Daytime (night-time) observations are marked in red (blue). The average LST differences and standard deviation are also indicated. The scatterplot covers the period between January 2015 and March 2016 (determined by LST_AVHRR and station data availability).

4 Concluding Remarks

4.1 SEVIRI/MSG LST (LSA-001)

Comparisons of satellite LST retrievals from SEVIRI and MODIS were compared with in situ observations made at Evora ground station (Portugal). Initially in-situ measurements were obtained using the rotating radiometer 'RotRad' (section 2.2.1). The comparison was consistent with inter-comparisons of LST_SEVIRI and LST_MODIS performed for the three areas in Trigo et al. (2008b). The differences between ground and satellite-derived values showed high variability for daytime for both sensors, with LST_SEVIRI overestimating in situ LST. This was mainly a result of the strong contrasts between tree crown and grass temperatures, particularly during the dryer months, as well as the variable fraction of shaded background in the sensors' FOV, which was not accounted for. The differences between satellite and Evora in-situ LST's were lower for night-time observations and both

satellite LSTs tended to underestimate in-situ LST, with colder values obtained for MODIS. These results agree with other studies that compare MODIS LST and ground observation over land, which also suggest an overall underestimation (Noyes et al., 2006; Bosilovich, 2006). Such cold bias may be associated to an overestimation of MODIS surface emissivity based on land cover classification, a problem that has been identified particularly for semi-arid regions (Wan et al., 2002, 2004; Xu et al., 2014). LST_SEVIRIs underestimate night-time observations by 0.5ºC and by ~3ºC for measurements taken within January and September periods, respectively. Furthermore, the comparison between SEVIRI and LST in situ measurements presented for 'RotRad' suggest an overestimation of the amplitude of LST daily cycle, with night-time (daytime) values colder (warmer) than in-situ LST.

It is well known that sparse canopies, e.g. the cork-oak tree forest at Evora (Portugal), often exhibit strong temperature differences between sunlit background, shaded background and tree crowns. From the 2008 onwards in situ radiances at Evora are measured with a set of separate Heitronics KT-15.85 IIP precision radiometers over the relevant endmembers. LSA SAF's LST_SEVIRI product has been validated using the in-situ data from April 2009 to October 2012: even though fixed fractions for the endmembers were used, the mean bias for this period was 0.6°C and mean rmse 1.9°C. However, there were extended periods of time for which mean rmse exceeded 2°C and between May and September 2012 bias abruptly changed to -2°C. This is thought to be caused by cattle breaking through the fence around Evora station and grazing within the FOV of a radiometer, which made the obtained in situ LST unrepresentative of the wider area around Evora station. In order to account for the dependence of LST observations on viewing and illumination geometries, Ermida et al. (2014) matched in situ LST to specific satellite sensors by modelling the end-member fractions within their respective FOV: compared to in situ LST obtained with static cover fractions, the use of modelled (dynamic) cover fractions to estimate in situ LST within the FOVs of SEVIRI and MODIS reduced the bias of SEVIRI and MODSW daytime LST values by 1°C to 2.5°C. Accounting for differences in viewing geometries also reduced the differences between of MODIS and MSG LST retrievals. However, SEVIRI/MSG LST showed a systematically higher temperature of about 3.5°C for daytime (0.8°C for night-time) observations compared to MODIS LST (MOD11A1 and MYD11A1 Collection 5). When LST was derived with the TES algorithm for MODIS (Hulley et al., 2011), the respective LST/emissivity products showed a better agreement with in situ modelled LST and SEVIRI/MSG LST.

Freitas et al. (2012) compared LST_SEVIRI with in-situ LST from Gobabeb station (Namibia), over a 1-year period. The site is located within an arid region, characterised by vast gravel plains. Satellite and in situ LST estimations show good agreement, with biases and root mean square differences of less than 1°C and 2ºC, respectively, i.e. the LST_SEVIRI were well within the estimated error bar (Freitas et al., 2010). Furthermore, up to five years of in-situ LST from KITs long term validation stations in Africa have been used to validate the operational LST product retrieved by the Land Surface Satellite Application Facility (LSA SAF) from MSG/SEVIRI data. The validation stations represent different surface cover types and climates and are located in flat, homogeneous terrains at the scale of several MSG-SEVIRI pixels. Typically thousands of monthly match-ups between satellite LST and in-situ LST were available at each validation site and yielded highly linear relationships between the two quantities. Furthermore, the large number of match-ups allowed seasonally resolved validations of LSA SAF LST; among others, this highlighted seasonal differences in the retrieval algorithm's performance, lower performance during rainy seasons as a consequence of increased cloud contamination. After strong rainy seasons there was an increase in night-time bias at Gobabeb (Namibia): the resulting increase in grass fraction is thought to have increased effective emissivity of the in-situ radiometer, which is currently treated as constant. However, these changes only occurred after unusually strong rain seasons and lasted for about 6 months.

The uncertainty of the currently available satellite-retrieved emissivities (or a constant emissivity in the case of Gobabeb) already allows validating LST with an absolute accuracy of 0.8K. Ignoring rainy seasons at Dahra (sub-tropical), the mean night-time rmse for LSA SAF's operational LST product at the three sites were at most 1.4°C while mean absolute biases were up to 0.7°C. The highest mean daytime rmse was 1.8°C and the corresponding highest mean absolute biases 0.6°C. The highest mean rmse for daytime and night-time data combined was 1.6°C and the corresponding highest mean absolute bias was 0.1°C.

4.2 AVHRR/Metop LST (LSA-002)

The assessment of AVHRR/Metop LST (LSA-002) has been performed for data covering the January 2015 – November 2016 period. LSA-002 corresponds to a daily composite of LST values retrieved from single AVHRR/Metop observations, which are later projected onto a sinusoidal grid and aggregated to form a daytime and night-time dataset. The validation exercise presented in this report is based on comparisons with: (i) the LSA SEVIRI/MSG LST product (LSA-001), a mature product which has been thoroughly assessed and characterised during the last 10-years; (ii) in situ measurements taken at KIT stations maintained within the LSA SAF.

A comparison between LST_AVHRR and LST_SEVIRI was performed for 6 areas (~10º longitude x 10º latitude) within the MSG disk, which were selected to represent a wide variety of surface and atmospheric conditions. Overall, average differences between night-time LST range between -1.4ºC and 0.2ºC, while their standard deviation lies around 1.0ºC for most cases. Daytime estimates show a larger discrepancies and also larger variability in space and time. The use of static emissivity fields in this version of LST_AVHRR leads to a seasonal variability in the differences to SEVIRI LST, particularly in regions where changes in vegetation, and therefore in emissivity, are more pronounced. This effect appears to be stronger for daytime values, possibly due to non-linear effects in the LST retrieval.

The higher spatial heterogeneity in daytime LST fields leads to higher and more complex differences between LST_AVHRR and LST_SEVIRI. It is shown that these depend on the view-illumination geometries, and also on the local orography and land cover type. Directional effects on LST are clearly seen in the AVHRR – SEVIRI comparison analysed here: [LST_AVHRR – LST_SEVIRI] are higher when AVHRR/Metop retrievals are taken from East, i.e., when AVHRR field-of-view presents a higher fraction of illuminated surfaces; the opposite occurs for AVHRR/Metop estimates taken from the West view. This effect is highest during the warm season, when temperature contrasts within surface elements are largest.

The ELST product has been compared to in situ estimates taken at KIT stations. These were specifically designed for the validation of LST satellite products, which means that not only the instruments are carefully calibrated and maintained, but the sites are carefully characterized to aid the matching between point and satellite measurements. Differences are again larger for daytime than for night-time. The latter are less influenced by directional effects and therefore a more reliable measure of the satellite product accuracy. On average LST_AVHRR underestimated night-time in situ estimates, with mean differences ranging between -0.8ºC and -1.7ºC. The root mean square differences for night-time LST lies within the target accuracy of 2ºC (2.0ºC obtained for Évora and Heimat, and 1.8ºC for Gobabeb).

5 References

Baldocchi, D. and 28 co-authors (2001), FLUXNET: A New Tool to Study the Temporal and Spatial Variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities, *Bull. Amer. Meteor. Soc.*, *82*, 2415-2432.

Barroso, C, I. F. Trigo, S. C. Freitas, P. Viterbo (2008), Impact of changes to SEVIIRI level 1.5 effective radiances on the Land-SAF Land Surface Temperature. Proceedings of the 2008 EUMETSAT Meteorological Satellite Conference, Darmstadt, 8 - 12 September 2008 (http://www.eumetsat.int).

Batjes, N.H. (2001). Options for increasing carbon sequestration in West African soils: an exploratory study with special focus on Senegal. *Land Degradation & Development*, Vol. 12(2), pp. 131–142.

Becker, F. (1987). The impact of spectral emissivity on the measurement of land surface temperature from a satellite. *International Journal of Remote Sensing*, Vol. 8(10), pp. 1509–1522.

Bork-Unkelbach, A. (2012). *Extrapolation von in-situ Landoberflächentemperaturen auf Satellitenpixel*. Phd thesis (in German), Karlsruher Institut für Technologie.

Carreiras, J. M. B., Pereira, J. M. C. and Pereira, J. S. (2006), Estimation of tree canopy cover in evergreen oak woodlands using remote sensing. *Forest Ecology and Management*, *223*, 45 – 53.

Dash, P., Göttsche, F.-M., Olesen, F.S., and Fischer, H. (2002). Land surface temperature and emissivity estimation from passive sensor data: theory and practice - current trends. *International Journal of Remote Sensing*, Vol. 23(13), pp. 2536–2594.

Dash, P., F.-S. Olesen and A. J. Prata (2004), Optimal land surface temperature validation site in Europe for MSG, Proceedings of the 2004 EUMETSAT Meteorological Satellite Conference, Prague, 31 May - 04 June 2004 (http://www.eumetsat.int).

Eckardt, F.D., Soderberg, K.L., Coop, J., Muller, A.A., Vickery, K.J., Grandin, R.D., Jack, C., Kapalanga, T.S., and Henschel, J. (2013). The nature of moisture at Gobabeb, in the central Namib Desert. *Journal of Arid Environments*, Vol. 93.

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

Fensholt, R., and Sandholt, I. (2005). Evaluation of MODIS and NOAA AVHRR vegetation indices with in situ measurements in a semi-arid environment. *International Journal of Remote Sensing*, Vol. 26(12), pp. 2561–2594.

Freitas, S. C., I. F. Trigo, J. M. Bioucas-Dias, and F. Göttsche, 2010, Quantifying the Uncertainty of Land Surface Temperature Retrievals from SEVIRI/Meteosat. *IEEE Trans. Geosci. Remote Sens.*, Vol. 48, Num. 1, pp. 523-534.

French, A.N., Schmugge, T. J., and Kustas, W.P. (2000). Discrimination of Senescent Vegetation Using Thermal Emissivity Contrast. *Remote Sensing of Environment*, Vol. 74, pp. 249–254.

García-Santos, V., Valor, E., Caselles, V., Ángeles Burgos, M., & Coll, C. (2012). On the angular variation of thermal infrared emissivity of inorganic soils. Journal of Geophysical Research, 117, D19116. doi:10.1029/2012JD017931.

Göttsche, F.-M., and Olesen, F.-S. (2009). Modelling the Effect of Optical Thickness on Diurnal Cycles of Land Surface Temperature, *Remote Sensing of Environment*, Vol. 113, pp. 2306–16.

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

Göttsche, F.-M., Olesen, F.-S., and Bork-Unkelbach, A. (2013). Validation of land surface temperature derived from MSG/SEVIRI with in situ measurements at Gobabeb, Namibia. *Inter. J. of Remote Sens.*, Vol. 34, Issue 9-10, pp. 3069-3083.

Guillevic, P.C., Bork-Unkelbach, A., Göttsche, F.M., Hulley, G., Gastellu-Etchegorry, J.-P., Privette, J.L., and Olesen, F.S. (2013). Directional viewing effects on Land Surface Temperature products over sparse vegetation canopies - A multi-sensors analysis from field to polar to geostationary satellites measurements. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 10(6), pp. 1464–1468.

Hulley, G.C., Hook, S.J., Manning, E., Lee, S.-Y., and Fetzer, E. (2009). Validation of the Atmospheric Infrared Sounder (AIRS) version 5 land surface emissivity product over the Namib and Kalahari deserts. *Journal of Geophysical Research*, Vol. 114(D19).

Hulley, G.C., Hook, S.J., and Baldridge, A.M. (2011). Generating consistent land surface temperature and emissivity products between ASTER and MODIS data for Earth science research. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 49(4), pp. 1304-1315.

Hulley, G., Veraverbeke, S., and Hook, S. (2014). Thermal-based techniques for land cover change detection using a new dynamic MODIS multispectral emissivity product (MOD21). *Remote Sensing of Environment*, Vol. 140, pp. 755–765.

Lancaster, J., Lancaster, N., and Seely, M.K. (1984). Climate of the central Namib Desert. *Madoqua*, Vol. 14(1), pp. 5–61.

Jimenez-Munoz, J.C., Sobrino, J.A., Mattar, C., Hulley, G., and Gottsche, F.-M. (2014). Temperature and emissivity separation from MSG/SEVIRI data. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 52(9), pp. 5937–5951.

Kabsch, E., Olesen, F.-S., and Prata, F. (2008). Initial results of land surface temperature (LST) validation with the Evora, Portugal ground-truth station measurements. *International Journal of Remote Sensing*, Vol. 29(17), pp. 5329–5345.
Kondratyev, K.Y. (1969). *Radiation in the atmosphere*. Academic Press, New York, USA.

Köppen, W. (1936). Das geographische System der Klimate. *Handbuch der Klimatologie*. Gebr. Bornträger.

Norman, J., and Becker, F. (1995). Terminology in thermal infrared remote sensing of natural surface. Agriculture and Forest Meteorology, Vol. 77, pp. 153–176.

Ni, W., Li, X., Woodcock, C. E., Caetano, M. R., and Strahler, A. H. (1999). An Analytical Hybrid GORT Model for Bidirectional Reflectance over Discontinuous Plant Canopies. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37(2), pp. 987-999.

Olioso, A., Sòria, G., Sobrino, J., and Duchemin, B. (2007). Evidence of low land surface thermal infrared emissivity in the presence of dry vegetation. *Geoscience and Remote Sensing Letters*, IEEE, Vol. 4(1), pp. 112–116.

Peel, M.C., Finlayson, B.L., and McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*, 11:1633–1644, 2007.

Pereira, J.S., Mateus, J.A., Aires, L.M., Pita, G., Pio, C., David, J.S., Andrade, V., Banza, J., David, T.S., Paco, T.A., and Rodrigues, A. (2005). Net ecosystem carbon exchange in three contrasting Mediterranean ecosystems - the effect of drought. *Biogeosciences*, 4:791–802, 2007.

Peres L. F., and C. C. DaCamara (2005), Emissivity maps to retrieve land-surface temperature from MSG/SEVIRI, *IEEE Trans. Geosci. Remote Sens.*, *43*, doi: 10.1109/TGRS.2005.851172.

Pinheiro, A. C. T., Privette, J. P., and Guillevic, P. C. (2006). Modeling the Observed Angular Anisotropy of Land Surface Temperature in a Savanna. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 44(4), pp. 1036-10467.

Rasmussen, M.O., Göttsche, F.-M., Olesen, F.-S., and Sandholt, I. (2011). Directional Effects on Land Surface Temperature Estimation From Meteosat Second Generation for Savanna Landscapes. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 49(11), pp. 4458–4468.

Rasmussen, M.O., Göttsche, F.-M., Diop, D., Mbow, C., Olesen, F.-S., Fensholt, R., and Sandholt, I. (2011a). Tree survey and allometric models for tiger bush in northern Senegal and comparison with tree parameters derived from high resolution satellite data. *International Journal of Applied Earth Observation and Geoinformation*, Vol. 13(4), pp. 517–527.

Rubio, E., Caselles, V., and Badenas, C. (1997). Emissivity measurements of several soils and vegetation types in the 8–14 µm wave band: analysis of two field methods. *Remote Sensing of Environment*, Vol. 59, pp. 490–521.

Salisbury, W., and D'Aria, D. M. (1992). Emissivity of terrestrial materials in the 8-14 micrometer atmospheric window. *Remote Sensing of Environment*, Vol. 42, pp. 83–106.

Schädlich, S., Göttsche, F.-M., and Olesen, F.-S. (2001). Influence of Land Surface Parameters and Atmosphere on METEOSAT Brightness Temperatures and Generation of Land Surface Temperature Maps by Temporally and Spatially Interpolating Atmospheric Correction. *Remote Sensing of Environment*, Vol. 75(1), pp. 39–46.

Schneider, P., Ghent, D., Corlett, G., Prata, F., and Remedios, J. (2012). AATSR Validation: LST Validation Protocol. Report, ESA Contract Number: 19054/05/NL/F (UL-NILU-ESA-LST-LVP Issue1 Rev0), pp. 1-39.

Stisen, S., Sandholt, I., Noergaard, A., Fensholt, R., and Jensen, K.H. (2008). Combining the triangle method with thermal inertia to estimate regional evapotranspiration – applied to MSG-SEVIRI data in the Senegal River basin. *Remote Sensing of Environment*, Vol. 112(3), pp. 1242–1255.

Theocharous, E., Usadi, E., and Fox, N. P. (2010). CEOS comparison of IR brightness temperature measurements in support of satellite validation. Part I: Laboratory and ocean surface temperature comparison of radiation thermometers. *NPL REPORT OP3*. Technical Report ISSN: 1754-2944, National Physical Laboratory, Teddington, UK.

Tagesson, T., Fensholt, R., Guiro, I., Rasmussen, M.O., Huber, S., Mbow, C., Garcia, M., Horion, S., Sandholt, I., Holm-Rasmussen, B., Göttsche, F.M., Ridler, M.E., Olén, N., Lundegard, O.J., Ehammer, A., Madsen, M., Olesen, F.S., and Ardö, J. (2015). Ecosystem properties of semiarid savanna grassland in West Africa and its relationship with environmental variability. *Global Change Biololgy*, Vol. 21(1), pp. 250-264.

Trigo, I. F., L. F. Peres, C. C. DaCamara, and S. C. Freitas (2008a), Thermal Land Surface Emissivity retrieved from SEVIRI/Meteosat, *IEEE Trans. Geosci. Remote Sens.*, *46*, doi: 10.1109/TGRS.2007.905197.

Trigo, I. F., I. T. Monteiro, F. Olesen, and E. Kabsch (2008b), An assessment of remotely sensed Land Surface Temperature. *J. Geophys Res.*, **113**, **D17108**, doi:10.1029/2008JD010035.

Wan, Z., (1999), *MODIS Land-Surface Temperature*, Algorithm Theoretical Basis Document (LST ATBD), Version 3.3, Contract Number: NAS5-31370, 75 pp.

Wan, Z., and J. Dozier (1996), A generalized split-window algorithm for retrieving land surface temperature from space, *IEEE Trans. Geosci. Remote Sens.*, *34*, 892–905.

Wan, Z., Y. Zhang, Q. Zhang, and Z.-L. Li (2002), Validation of the land-surface temperature products retrieved from Terra Moderate Resolution Imaging Spectroradiometer data. *Remote Sens. Environ.*, *83*, 163-180.

Xu, H., Yu, Y., Tarpley, D., Gottsche, F.-M., and Olesen, F-S (2014). Evaluation of GOES-R Land Surface Temperature Algorithm Using SEVIRI Satellite Retrievals With In Situ Measurements. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 52(7), pp. 3812–3822.