Report from the Field Inter-Comparison Experiment (FICE) for Land Surface Temperature

Frank-M. Göttsche[1], Folke Olesen[1], Laurent Poutier[2], Stéphane Langlois[2], Werenfrid Wimmer[3], Vicente Garcia Santos[4], César Coll[4], Raquel Niclos[4], Manuel Arbelo[5] and Jean-Pierre Monchau[6]

[1]Karlsruhe Institute of Technology, Germany,
[2]Office National d'Etudes et Recherches Aérospatiales, France,
[3]National Oceanographic Centre, United Kingdom,
[4]University of Valencia, Spain, [5]Universidad de la Laguna, Spain,
[6]THEMACS Ingénierie, France

Digitally signed by Device Identity
9c6f8ccd4f5eaa531db70b0929b3d2a06f1e6010
Reason: I am approving this document
Location: Noordwijk, The Netherlands
Date: 2018.06.05 15:13:58 +02'00'

Abstract

A land surface temperature (LST) field inter-comparison experiment (FICE) was conducted in June 2017 on the highly homogeneous Namib gravel plains near Gobabeb Research and Training Centre (GRTC), Namibia. Five different Fiducial Reference Measurement (FRM) Thermal Infrared (TIR) radiometers, a Fourier Transform TIR spectrometer and a novel 'emissiometer' participated in the experiment. The FICE was the first of its kind to be performed over the Namib gravel plains. Weather conditions were typical for a hyper-arid desert environment during southern hemisphere winter, with air temperatures between 10°C and 35°C, surface temperatures between 10°C and 45°C, with overcast skies on the 17-18 of June and predominantly clear-sky situations during the remaining days of the experiment. In-situ land surface emissivities were obtained from dedicated measurements with a FTIR spectrometer and an 'emissiometer', which is a novel instrument combining an oscillating TIR radiance source with digital signal processing to determine the band-effective emissivity of a radiometer. Additionally, emissivity spectra of soil samples were retrieved in the laboratory.

The first part of the LST FICE consisted of a four day inter-comparison between parallel TIR measurements from a wind mast. Five radiometers were mounted at heights between 11m and 15m to increase their footprints on the gravel surface to more than 2m^2. With respect to the chosen reference, a radiometer continuously calibrated by two blackbodies (ISAR: 'infrared sea surface temperature autonomous radiometer'), the in-situ LST RMSE were less than 1.0 K for the other four instruments, less than 0.7 K for three instruments, and about 0.3 K for two radiometers. The results show that in-situ LST can be retrieved with RMSEs of about 0.5 K, if the deployed radiometers are well-aligned, have narrow spectral bands and view angles, observe a surface area of more than 2 m^2 and accurate channel-specific in-situ emissivities are available.

The second part of the FICE investigated the impact of spatial variability on retrieved in-situ LST: five radiometers were mounted to horizontal booms fixed to the roofs of two vehicles and then driven three times for about 20km across the Namib gravel plains north-east of GRTC. The set-up did not allow an alignment of the radiometers and the movement of the vehicles caused vibrations. In order to obtain spatially representative in-situ LST, in-situ LST obtained over about 200 m driving distance were averaged, which yielded RMSEs of 0.6 K or less with respect to the mean of the in-situ LSTs provided by the five teams, thereby giving comparable results to the inter-comparison experiment performed at the wind mast.

© KIT 2017

Karlsruhe Institute of Technology

Hermann-von-Helmholtz-Platz 1

76344 Eggenstein-Leopoldshafen, Germany

Extracts from this report may be reproduced provided the source is acknowledged and the extract is not taken out of context.

Table of Contents

4

1 MOTIVATION

Satellite remote sensing of surface parameters is an essential part of the global observation system and provides inputs for weather forecast, climate studies and many other applications. One of the important variables is surface temperature. Satellites have been monitoring global surface temperature for several decades and have established sufficient consistency and accuracy between in-flight sensors to claim that it is of "climate quality". However, it is essential that such quantities are fully anchored to SI units and that there is a direct correlation with "true" surface/in-situ based quantities, which must be derived from completely independent measurements, i.e. without the involvement of any satellite observations. The most accurate surface based land surface temperatures (LST) are derived from Fiducial Reference Measurements (FRM) with field radiometers. These are in principle calibrated traceably to SI units, generally through a reference radiance blackbody. Such instrumentation is of varying design and operated by different teams in various parts of the globe. However, so far neither FRM field radiometers nor their field deployment have been compared and there are no established standards to ensure SI-traceability. Therefore, the overarching objective of the Field Inter-comparison Experiments (FICE) performed within the FRM4STS project is "to coordinate and demonstrate field inter-comparison activities for TIR FRM".

Field inter-comparisons cannot be controlled to the same extent as inter-comparisons in the laboratory: therefore, selecting naturally homogenous sites is of key importance [8]. This report describes the LST FICE performed in June 2017 in Namibia as a contribution to the FRM4STS project (www.FRM4STS.org). KIT operates a permanent LST validation station at Gobabeb, Namibia and characterised the site with a number of field campaigns: these showed that the station LST, which are determined over a surface of about $12m^2$, are representative of the highly homogeneous Namib gravel plains over a broad range of spatial scales up to several ten kilometres [10]. The experiments described here build on this knowledge and, for the first time, inter-compare different FRM TIR radiometers under desert conditions and assess the retrieved in-situ LSTs and their associated uncertainties.

2 EXPERIMENTS

The field program was carried out between Friday, the 16th of June 2017 and Sunday, the 25th of June 2017 at Gobabeb, Namibia. Six research groups participated in the LST FICE: Karlsruhe Institute of Technology (KIT, Lead), Office National d'Etudes et Recherches Aérospatiales (ONERA), National Center for Oceanography Southampton (NOCS), Thermal Remote Sensing Group of the University of Valencia (TRSG-UV), Grupo de Observacion de la Tierra y la Atmosfera Universidad de la Laguna (GOTA-UL), and 'Thermique, Environnement, Matériaux, Contrôle de Structures Ingénierie' (THEMACS). For logistical reasons the experiment was performed in two parts, separated by three days during which automated

measurements were performed. The LST FICE measurement program consisted of three components:

- **Temporal inter-comparison (four days).** Five field radiometers were mounted to the mast of the permanent LST validation station between 11m and 15m height; all instruments were aligned with a laser and observed surface areas > 2 m^2.
- **Spatial inter-comparison and variability (two days).** Five field radiometers and a TIR camera were driven along a track between GRTC and 'Mirabib' to assess LST spatial variability across the gravel plains.
- **Determination of in-situ emissivity.** Emissivities of 49 samples at the two masts, GRTC and along the track (spectrometer and novel 'emissiometer') were obtained.

2.1 Program

Due to the very tight accommodation situation at GRTC, the LST FICE had to be shifted several times and finally took place between Friday the 16th and Monday the 26th of June 2017. Most participants arrived in Windhoek early in the morning on Thursday the 15th to pick up shipped equipment and buy required materials and supplies. The transfer to Gobabeb on the 16th of June was followed by two intense days of setting up equipment and instruments at 'GBB wind' mast. Due to a lack of accommodation the LST FICE participants had to leave GRTC on Sunday the 18th and returned to the site on Thursday the 22nd of June. During the four days of absence from the site continuous automatic measurements were successfully performed by all participating instruments. Table 1 gives an overview of the various activities and measurements performed during the LST FICE 2017.

Table 1: *LST FICE 2017 activities and measurements*

Day of experiment	Activities and measurements
Fri. 16.06	Teams arrive at various times of the day at Gobabeb Research and Training Centre (GRTC). Unpacking of shipped boxes and equipment from storage; checking and preparing of instruments; inspection of main mast & site. After a briefing about the work programme and safety instructions the day closes with a welcome dinner.
Sat. 17.06	Transportation of equipment, tools and safety equipment to GBB wind mast (about 2 km northeast of GRTC). Starting to set up instruments and mount radiometers to the mast (at 11m to 15m height). Parallel calibration of radiometers against Mikron 345x4 blackbody.
Sun. 18.06	Continuing to set-up the instruments and installing battery power supplies for several days of automated operation. Measurements started at 09:00 UTC. In parallel the

	emissivity of gravel around GBB wind mast is determined with the emissiometer and the TIR camera is calibrated. On Sunday afternoon all teams leave GRTC.
Mon. 19.06 – Wed 21.06	Automated measurements.
Thu. 22.06	Teams return to GRTC at 7:00h UTC. Briefing and inspection of instruments, downloading of recorded measurements. Measurements are stopped at 09:00h UTC. Work is performed in three groups: group 1 removes the instruments from the mast, group 2 prepares two vehicles for mobile measurements (mounting horizontal booms and instruments), group 3 performs emissivity measurements.
Fri. 23.06	08:00h - 14:00h UTC: first mobile measurement along the track Gobabeb – Mirabib, twice in each direction. Completing the unmounting of equipment from GBB wind mast. Emissivity measurements at masts GBB Wind and GBB Plains.
Sat. 24.06	06:30h - 11:00h UTC: second mobile measurement along the track Gobabeb – Mirabib, once in each direction. Unmounting of instruments and equipment from vehicles. Emissivity measurements at both masts.
Sun. 25.06	Disassembly of equipment and instruments. Equipment is returned into storage and instruments are prepared to be shipped home. After a debrief most participants leave for Windhoek to ship their instruments and fly home.
Mon. 26.06	Official end of LST FICE 2017; remaining participants leave GRTC.

2.2 Experiment Site

Gobabeb Research and Training Centre (GRTC; www.gobabebtrc.org) in Namibia is the only permanently staffed desert research station worldwide. GRTC is located on a sharp transition between the vast Namib sand sea with its up to 300 m high dunes and adjacent 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 ([17]; [22]), the site is spatially and temporally highly stable and, therefore, ideal for long-term validation studies of satellite products ([14]; [13]). The long-term average annual temperature at Gobabeb is 21.1°C [18] whereas the average annual precipitation is less than 100 mm [5] and highly variable ([20]; [26]).

Consequently, the relatively frequent fog events are of special importance for the water balance of the Namib [5].

Continuous in-situ measurements are performed at KIT's two permanent LST validation stations 'GBB Wind' (23.551° S, 15.051° E, 450m asl) and 'GBB Plains' (23.519° S, 15.083° E, 450m asl). GBB Wind uses a 30m high wind profiling tower about 2 km north-east of GRTC, while GBB Plains uses a 25m high telescopic mast about 7km north-east of GRTC. Both stations are at the edge of several thousand km^2 of gravel plains, which are covered by a highly homogeneous mixture of gravel, sand and sparse desiccated grass (see Figure 1). 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 characterized. Using a mobile radiometer system, several field experiments were performed during which a radiometer was driven along tracks of up to 40 km length across the gravel plains. The results showed a high level of homogeneity and a stable relationship between GBB Wind LST and the LST along the tracks with biases between -0.1°C and 0.8°C [11].

Figure 1: 360 degree panorama at KIT station 'GBB Plains'; the mast is located left of the car.

Clear sky conditions are preferable for field measurements since down-welling radiance is then easier to determine and varies relatively slowly and smoothly. Furthermore, LST retrieval from passive TIR satellite sensors also requires clear-sky situations. Since Gobabeb is located in the Namib Desert it offers frequent clear sky conditions almost all year around, which is ideal for LST validation.

2.3 Measurement protocol

The LST FICE closely followed '3.6.1 Measurement protocol for In situ LST' provided in [8]. In order to minimise differences due to LST anisotropy, all measurements were performed at near-nadir view angle (about 30°). Since LST is not directly measured but derived from measurements of surface brightness temperature (BT), sky BT and emissivity, the participants provided their corresponding estimates along with time (UTC) and geolocation (decimal degrees lat/lon). Three different approaches for obtaining hemispherical sky radiance were used: BT at representative zenith angle of 53°, zenith observation of BT, and BT of diffuse gold plate [6]. The participants obtained their own instrument-specific Land Surface Emissivities (LSE), e.g. using the TES method, the emissivity box method [25] [9], or by convolving their instruments spectral response function with in-situ emissivity spectra

provided by ONERA. In order to ensure representative surfaces that are homogeneous on the spatial scale (i.e. the FOVs) of the participating radiometers, these were mounted to 11-15 m height for the measurements performed at GBB Wind. For the mobile measurements from the 4x4 vehicles the FOVs were considerably smaller (about 30 cm in diameter), but temporal averaging and resampling to a common resolution of 1 minute reduced the impact of small scale spatial variations considerably.

3 PARTICIPATING INSTRUMENTS

Five research groups with different instruments for measuring thermal infrared radiance (TIR) participated in the inter-comparison experiments. Additionally the permanent validation station provided up- and downwelling shortwave and longwave broadband radiances as well as basic meteorological parameters.

The instruments deployed by the participants were an ISAR (NOCS), three Heitronics KT19.85 II (ONERA), two Heitronics KT15.85 IIP (KIT), two CIMEL CE 312-1 (TRSG-UV), and one CIMEL CE 312-2 (GOTA-UL). Furthermore, in-situ emissivity spectra were determined with an FTIR field spectrometer (ONERA) while channel-specific 'broadband' emissivities were obtained with a novel emissiometer (THEMACS) [4] [21]. Details about the participating instruments and their characteristics are provided in Table 2 and in the following sections.

Radiometer	**Institution**	**LST sampling rate station / mobile**	**Spectral range (µm)**	**Measured parameters**
ISAR	NOCS	2-3 min	9.8-11.5	Surface BT / sky BT
KT19.85 II	ONERA	1 min / 1 sec	9.6-11.5	Surface BT / sky BT
KT15.85 IIP	KIT	1 min / 1 sec	9.6-11.5	Surface BT / sky BT
CIMEL CE 312-1	TRSG-UV	20 min / 2-3 min	8-14 (4 bands)	Surface BT / sky BT
CIMEL CE 312-2	GOTA-UL	43 min / 2-3 min	7-13 (6 bands)	Surface BT
Other instruments	**Institution**	**sampling rate**	**Spectral range (µm)**	**Obtained parameter**
BOMEM MR304SC FTIR	ONERA	1 min	3 to 13	Emissivity spectra
EM-3 emissiometer	THEMACS	1 min	1-50 (broadband) 8-14 (band III)	Channel-specific LSE
TIR Camera	THEMACS	1 sec	7.5-13.0 (band III)	Surface BT

Table 2 Instruments participating in the LST FICE 2017

3.1 ISAR radiometer (NOCS)

The infrared Sea Surface Temperature (SST) autonomous radiometer (ISAR) is a self-calibrating instrument measuring infrared emission from the surface and atmosphere. It employs two reference blackbody cavities to maintain the radiance calibration of a Heitronics KT15.85D radiometer with an accuracy of ±0.1K and measures IR emission in the spectral waveband 9.8-11.5 µm [29][31].

3.1.1 Operational methodology

One measurement cycle views the land surface for ~70 seconds, the sky temperature for about 20 seconds and the two black bodies for about 60 seconds each. Afterwards for each cycle one referenced surface brightness temperature and land surface temperature value is calculated using post processing software. The land incidence angle is 35° from nadir while sky incidence angle is 53° from zenith. The post processing software also provides uncertainty values for each referenced measurement. These values have been derived for LST in analogy to SST [30].

3.2 Heitronics KT15.85 II P radiometer (KIT)

The KT15.85 IIP is a single channel radiometer based on a pyroelectric infrared detector and had a L6 lens with a full-view angle of 8.3°, which is well-suited for directional measurements. This type of sensor links radiance measurements via beam-chopping to internal reference temperature measurements and thermal drift can practically be eliminated. The KT15.85 IIP covers the spectral range from 9.6 µm to 11.5 µm, has an uncertainty of about ±0.3K over the temperature range relevant to land surfaces [29] and offers excellent long-term stability.

3.2.1 Breakdown of LST uncertainty

The uncertainties in Table 3 are based on estimates and laboratory measurements obtained with KITs Landcal Blackbody Source P80P (May 2017) and during the “Laboratory Intercomparison Exercise” held in NPL (June 2016) with the reference radiance blackbody calibrated traceable to SI.

At the first day of the LST FICE the KT15.85 IIP were compared against ONERAs Mikron M345x4 blackbody: for blackbody temperatures (uncertainty ±0.2°C) set to 20°C and 50°C the radiometers had RMSE of 0.3°C and 0.2°C, respectively. The KT15.85 IIP emissivity of 0.916 ± 0.007 determined for pure gravel (no grass component at the time the LST FICE) in a previous field experiment [9] was used for all LST retrievals performed KIT during the FICE. The used data processing methods are detailed in [11] and [10].

Uncertainty Contribution	Type A Uncertainty in Value / %	Type B Uncertainty in Value / (appropriate units)	Uncertainty in Brightness temperature K
Repeatability of measurement	0.57		0.143
Reproducibility of measurement	0.57		0.143
Primary calibration		0.250 K	0.250
Target emissivity		0.5 %	0.333
Linearity of radiometer		0.070 K	0.070
Drift since calibration		0.179 K	0.179
Resolution of radiometer		0.035 K	0.035
Ambient temperature fluctuations		0.035 K	0.035
Atmospheric absorption/emission		0.035 K	0.035
Down-welling sky radiance		0.011 K	0.011
RMS total	0.81 %		0.505 K

Table 3 Typical uncertainty contributions for LST determination with the Heitronics KT15.85 IIP radiometer.

3.2.2 Operational methodology

For the temporal inter-comparison experiment at GBB Wind (section 5.1) the KT15.85 IIP radiometer was mounted at 15 m height and aligned to the target area with a laser, yielding a view angle of about 32° and a footprint on the gravel surface north of the mast of at least 2 m^2. The response time of the radiometer was set to 10 sec and its temperature range to -25°C to +100°C. Sky brightness temperatures (BT) were obtained from another KT15.85 IIP (temperature range -100°C to +100°C; at representative zenith angle of 53°) installed at station GBB Wind. Additionally, sky BT measured by ONERA with a KT19.85 II at 55° zenith angle and a KT19.85 II observing a diffuse gold plate were available. For the temporal inter-comparison experiment BTs were recorded every 10 sec by a Campbell Scientific CR1000 data logger and then averaged over 1 minute.

For the measurements performed along the track across the gravel plains (section 5.2) the KT15.85 IIPs response time and the sampling rate were set to 1 sec; for the inter-comparisons the obtained in-situ LSTs were averaged and resampled to a lower temporal resolution. The radiometer was mounted to the end of a horizontal boom on a 4x4 vehicle and aligned to a viewing angle of 30°. Simultaneously sky brightness temperatures were obtained at nadir (zenith angle of 0°) by a separate KT15.85 IIP with a temperature range of -100°C to +100°C.

3.3 Heitronics KT19.85 II radiometer (ONERA)

ONERA provided three KT19.85 IIP radiometers for the LST FICE. This type of radiometer has a 95 mm target diameter at a 2 m range and a spectral range of 9.6 to 11.5 µm.

3.3.1 Breakdown of uncertainty for the brightness temperature measurements

The breakdown of uncertainty was assessed during the laboratory comparison of radiation thermometers held at NPL in June 2016 uncertainty.

The repeatability of measurements was evaluated for each of the three radiometers using the standard deviation of continuous measurements in front of the ONERA Mikron M345x4 blackbody regulated at various temperature values ranging between 12°C and 55°C. This repeatability is better than 0.05K (including the blackbody temporal instability).

The reproducibility error is obtained by considering 5 sets of measurements in front of the Mikron blackbody with the same set temperature. The standard deviation between theses 5 experiments is less than 0.05K.

The primary calibration 2-sigma uncertainty of measurements made by these radiometers is given by the manufacturer as ±0.5 °C + 0.7 % of the difference between target and housing temperature. The final uncertainty breakdown is summarized in Table 4.

Uncertainty Contribution	Uncertainty in Brightness temperature K
Repeatability of measurement	0.05K
Reproducibility of measurement	0.05K
Primary calibration	$0.25°C+0.35\% \times \Delta T_{target-inst}$
rms	0.33K for $\Delta T_{target-inst} = 20K$

Table 4 Typical uncertainty contributions for brigthness temperature measurement with the Heitronics KT19.85 IIP radiometer.

Indeed, the FRM4STS Laboratory inter comparison of radiometers held at NPL in June 2016 showed a difference with the NPL standard blackbody less then 200mK for these 3 radiometers in the range [0°C, 45°C] [28].

3.3.2 Operational methodology

During the intercomparison at GBB Wind one KT19.85 II was mounted next to KITs KT15.85 IIP radiometer at 15m height and aligned with a target laser. The other two KT19.85 II radiometers were mounted at about 2m height to measure downwelling sky radiance at the representative zenith angle of 55° and via a diffuse gold plate equipped with a temperature probe to compensate for its self-emission. The general methodology is illustrated in Figure 2.

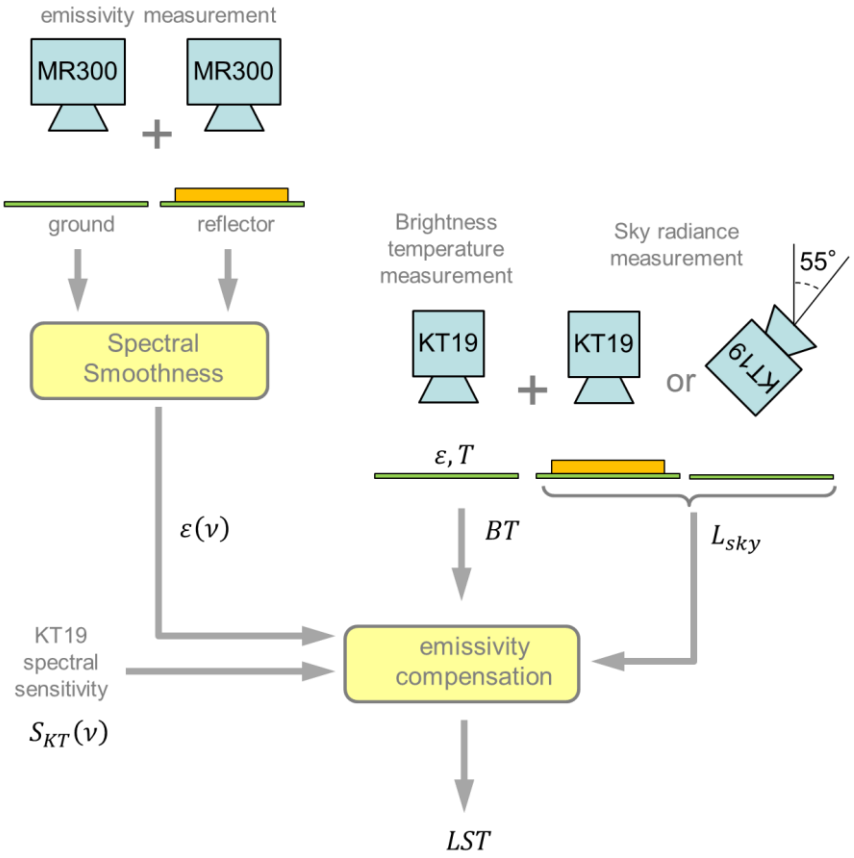

Figure 2 General methodology employed by ONERA for the retrieval of in-situ LST.

The LST retrieval method takes into account the spectral emissivity, by using two Look-Up tables for the two following relations:

$$T \leftrightarrow \frac{\int \varepsilon_{target}(\nu).B(\nu,T)S_{KT}(\nu)d\nu}{\int S_{KT}(\nu)d\nu}$$

$$T \leftrightarrow \frac{\int \left(1-\varepsilon_{target}(\nu)\right).B(\nu,T)S_{KT}(\nu)d\nu}{\int S_{KT}(\nu)d\nu}$$

Where $B(\nu, T)$ is the Planck law for wavenumber ν, $S_{KT}(\nu)$ is the radiometer broadband normalized spectral sensitivity and $\varepsilon_{target}(\nu)$ is the spectral emissivity assessed with the Bomem spectroradiometer.

The first relation represents the emissive contribution of the measured radiance, T being the surface temperature; the second relation represents the reflective contribution of the measured radiance and T is the sky brightness temperature retrieved by the sky radiometers.

3.3.3 Error budget assessment for the retrieved surface temperature

The error budget takes into account the following sources of error:

- emissivities: determined at the mast and along the track with the BOMEM MR304sc spectrometer; uncertainty ≤ ±0.02 (see §4.2)
- time: average per minute for DRIVE experiments: [hh:mm:00 - hh:mm:59] => hh:mm
- KT19 measurements: uncertainty = 0.25K + 0.35% (T_{target} - $T_{ambiant}$), using the weather station values for $T_{ambiant}$, see :

  www.sasscalweathernet.org/weatherstat_hourly_we.php?loggerid_crit=8893

- KT19 spectral response from Heitronics. Uncertainty: spectral shift of +-50nm.

The downwelling sky brightness temperature uncertainty is obtained using the following sources of error (in addition to the KT19 spectral response and measurement uncertainties described above):

- For sky brightness temperature using the indirect method :
  - infragold reflectance: from laboratory measurements; uncertainty = ±0.03
  - infragold surface temperature: temperature probe; uncertainty = ±2K (at MAST only)
- For sky brightness temperature using the direct method – 55° viewing angle):
  - Angular uncertainty of ±2° : 3.7% in sky radiance

3.4 CIMEL CE 312-2 radiometer (GOTA-UL)

The CIMEL CE 312-2 radiometer consists of an optical head and a control unit. The detector is a thermopile that measures radiance coming from a target through one of five narrow spectral channels with effective wavelengths at 8.3, 8.5, 9.1, 10.5, 11.1 µm and a wide band

between 7 and 13 µm. The radiometer has a field of view of 10° and a measurement integration time of 1 s. The instrument has a built-in radiance reference made of a retractable gold-coated mirror that enables comparison between the target radiance and the reference radiation from inside the detector cavity. The temperature of the detector is measured with a PRT, thus allowing compensation for the cavity radiation. A detailed description of the radiometer and its performance derived from laboratory and in situ measurements are given in [3] [19] [27] and under www.cimel.fr/?instrument=radiometer-ir-climat-benchmark&lang=en .

3.4.1 Breakdown of uncertainty

The uncertainties shown in Table 5 are based on estimates and laboratory measurements with our Landcal Blackbody Source P80P (May 2017) and on results from the "Laboratory Intercomparison Exercise" held in NPL (June 2016), which were obtained for a reference radiance blackbody calibrated traceable to SI.

The Temperature Emissivity Separation (TES) algorithm was applied to the two sets of radiometric measurements (wind tower and tracks). An average value of the retrieved spectral emissivities was calculated as well as its standard deviation for each site. On the other hand, ONERA emissivity spectra were convolved with our CIMEL CE312-2 spectral response functions, in order to obtain an average value of the spectral emissivity for each site (wind tower and tracks).

Uncertainty Contribution	**Type A**	**Type B**	
	Uncertainty in %	**Uncertainty in (K)**	**Uncertainty in Brightness temperature (K)**
Repeatability of measurement	0.03		0.09
Reproducibility of measurement	0.01		0.03
Primary calibration		0.33	0.33
Linearity of radiometer		0.1	0.1
Drift since calibration			
Ambient temperature fluctuations		0.1	0.1
Size-of-Source Effect			
Atmospheric absorption/emission			
RMS total	0.09 K/0.03 %	0.36 K	0.37 K

Table 5 Breakdown of uncertainties for the CIMEL CE312-2 radiometer (GOTA-UL).

Table 6 shows the emissivities obtained with both approaches (TES & ONERA); these were used to retrieve LST from channel 2 and 3 measurements of the CIMEL radiometer. The in-situ LST inter-compared in the section 5 were retrieved with channel 2 emissivities obtained from ONERA emissivity spectra.

CIMEL channel	TES (mast)	ONERA (mast)	TES (track)	ONERA (track)
2 (10.8-11.5µm)	0.957 ±0.001	0.952 ±0.017	0.956 ±0.001	0.946 ±0.016
3 (10.15-10.9µm)	0.937 ±0.003	0.925 ±0.017	0.942 ±0.006	0.923 ±0.016

Table 6 CIMEL CE312-2 emissivities determined by GOTA-UL with the TES method and from ONERA emissivity spectra.

3.4.2 Operational methodology

Four measurements of the surface radiance per spectral band were made consecutively over each site at an observation angle close to 35°. Each channel measurement lasted approximately 19-20 s (entire data takes near 2 min (20 s × 6 bands)).

Continuous measurements (sampling rate set to 43 minutes) were performed from 11 meters height from the wind tower ('GBB Wind') between 09:29 UTC on the 18/06/2017 and 08:48 UTC on the 22/06/2017, yielding 134 measurements for each channel. Simultaneous sky brightness temperatures were not obtained, but approximated from measurements performed with a CIMEL radiometer by the Universidad de Valencia, which observed the sky at a zenith angle of about 53°.

Measurements along different tracks were made on the 23rd and the 24th of June 2017. The methodology used was the same as for the wind tower, but a different sampling rate was set for these measurements: 2.30 minutes. A similar time series of sky radiometric measurements from Universidad de Valencia's CIMEL were used to estimate the downwelling atmospheric radiance.

3.5 CIMEL CE 312-1 radiometer (TRSG-UV)

The CIMEL CE 312-1 is a four-band radiometer (b1: 8-13 µm, b2: 11.5-12.5 µm, b3: 10.5-11.5 µm and b4: 8.2-9.2 µm) with field of view of 10°. The instrument has a built-in radiance reference made of a concealable gold-coated mirror which enables comparison between the target radiance and the reference radiation from inside the detector cavity. The temperature of the detector is measured with a calibrated PRT, thus allowing compensation for the cavity radiation. Laboratory calibration exercises using a Landcal P80P blackbody (itself calibrated against a reference radiometer during the "Laboratory Intercomparison Exercise" held in NPL in June 2016) provide re-calibration equations to account for the decrease of detector's sensitivity with time. After re-calibration, CE 312-1 shows good accuracy and stability. Two units of CE 312-1 were used in the present comparison. For the measurements from the mast on June 18-22, the optical head of both radiometers was kept inside a protective shield

with a 4-mm-thick ZnSe window, which was selected because of its good transmission in the thermal infrared window. The effect of the window on the CE 312-1 measurements was compensated through calibration equations derived for each band and radiometer from laboratory measurements against the P80P blackbody.

3.5.1 Breakdown of uncertainty

Together with the LST data measured by the TRSG-UV team and the determined in-situ emissivity values (3.5.2), the associated LST uncertainties were estimated for each case taking into account three error sources:

a) The uncertainty in the re-calibration of the CE-312-1 radiometers resulting from the calibration experiments carried out in the laboratory against the P80P blackbody and in the field at Gobabeb with the ONERA blackbody on June 17. Results show an uncertainty <0.2 K in the relevant temperature range for the two units and all bands.
b) The uncertainties in the emissivity values used, according to the field emissivity measurements performed during the experiment (see 3.5.2). This was the largest source of error, yielding uncertainties around 1.0 K for all surfaces and bands.
c) The uncertainties in the sky downwelling radiance measured in the field. We assumed an uncertainty of 10 K in the equivalent sky temperature, from which the resulting uncertainty in LST ranged from 0.04 K for the narrow band with high emissivity to 0.4 K for the wide band with low emissivity.

We also evaluated the impact on LST of the atmospheric absorption and emission in the path from the radiometer position on the mast (path length 17 m, zenith angle of 30°) to the ground using MODTRAN 5.0 and NCEP profiles over Gobabeb for the experiment days at 00:00 and 12:00 UTC. The simulations showed that the impact was negligible both for the broad (b1) and a narrow band (b2) (<0.04 K).

3.5.2 Operational methodology

In order to retrieve LSTs from thermal infrared in-situ measurements of surface-leaving radiances, an emissivity estimate is required for each surface type, together with simultaneous measurements of downwelling sky radiances. The TRSG-UV team used two CE-312-1 multiband radiometers for the surface and sky radiance measurements. The required emissivity data were obtained from in-situ measurements performed by different methods for the various targets considered in the experiment. The emissivity values selected for each case are described below.

For the LST measurements performed from the mast on June 18-22, we used the CE-312-1 wide band b1 (8-13 µm) and the narrow band b2 (11.5-12.5 µm). The appropriate emissivity values were obtained from the spectral measurements performed by ONERA for the target labelled as MAST_TARGET_GLOB, which were integrated with the filter functions of the CE-312 bands, yielding 0.893±0.017 for b1 and 0.974±0.016 for b2.

For the LST measurements performed along transects on June 23 and 24, we used two approaches: (i) Simultaneous derivation of LST and band emissivities from the temperature-emissivity separation (TES) method (Gillespie et al., 1998) using the four bands of the CE-312-1 radiometers. The LST was obtained for the 11.5-12.5 µm band (maximum emissivity among all bands), for which the emissivity was estimated as 0.967±0.013 for both days. (ii) Using the spectral measurements performed by ONERA for the target labelled as ROAD, which were integrated with the filter function of the 11.5-12.5 µm band, yielding 0.974±0.016.

3.6 BOMEM MR304SC FTIR spectroradiometer (ONERA)

The spectral emissivity is derived from two radiance spectra measured sequentially: one looking down at the sample, Rs, the other one looking at a diffuse reflector, $Rref$, in order to estimate the downwelling irradiance. The use of a reflector in place of the target has the advantage of taking into account the contribution of the instrumentation in the downwelling irradiance estimation [15] [16]. The radiance spectra are acquired by a BOMEM MR304SC FTIR spectroradiometer equipped with a 75mrad FOV telescope and a 45° flat mirror. With this setup, the ground target surface is viewed at nadir and the diameter of the analyzed area is approximately 10cm. The reflector is a Labsphere Infragold 10'x10' standard equipped with a temperature probe. Figure 3 illustrates the instrumentation during the experiment.

Figure 3. BOMEM MR304SC FTIR spectroradiometer at the Gobabeb site.

3.6.1 Operational methodology

The two radiance spectra are acquired sequentially, typically within a 1 minute interval. The temporal variations of the atmospheric conditions are assumed negligible in this time interval.

Radiometric calibration uses two acquisitions of a single MIKRON M345 4'x4'blackbody set at two different temperatures, done alternately before and after the actual measurements. For the calibration measurements the flat mirror is tilted and the blackbody active surface is vertical. The blackbody emissivity is assumed to be spectrally constant and set to its nominal value of 0.983 [23][28][24]. The reflective contribution is supposed to come from an environment at a brightness temperature equal to ambient temperature, which is obtained from Gobabeb meteorological station.

3.7 Emissiometer EM3 (THEMACS Ingénierie)

The device used is the emissivity measuring device EM3 from THEMACS Ingénierie. This device is the result of a patent filed in 2012 by CERTES (University Paris Est Créteil). The patent [4] describes the apparatus and [21] its operation. Figure 4 shows the sensor head of the apparatus (A) and its operation during the LST FICE at Gobabeb (B).

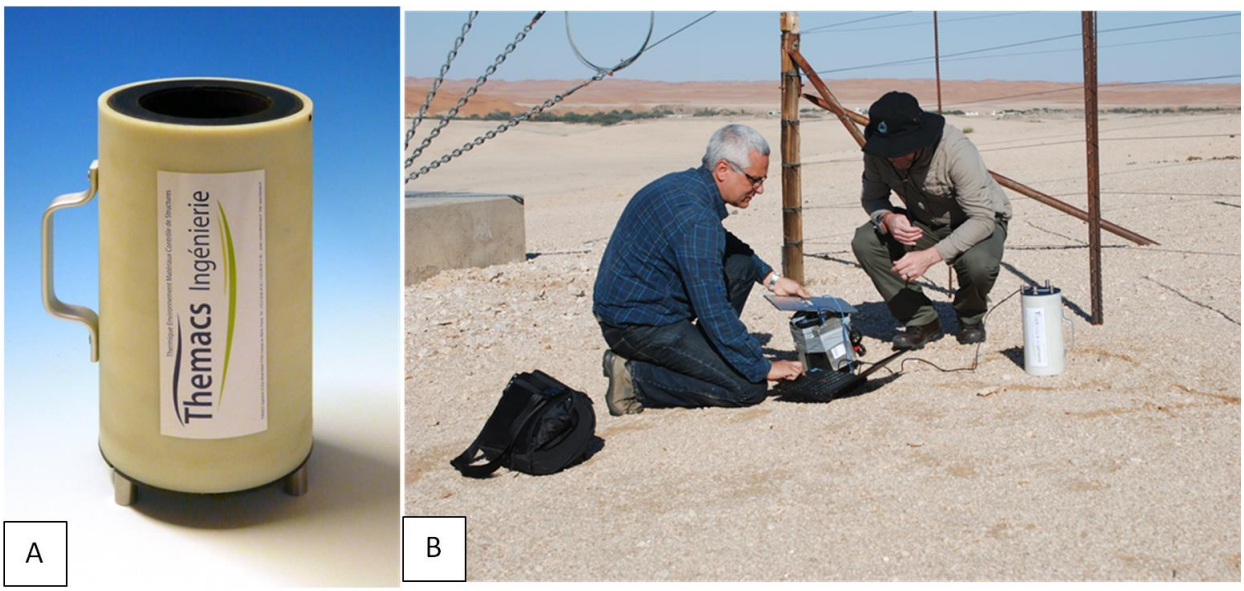

Figure 4. A: the sensor head of the EM3 emissiometer. B: field measurement on Gobabeb site

4 EMISSIVITY MEASUREMENTS

During the LST FICE 2017 at Gobabeb two of the participants, ONERA and THEMACS, performed emissivity measurements at various locations of the gravel plains, e.g. around

both of the permanent validation stations and along the tracks driven for the mobile LST measurements during the second part of the experiment. Whereas ONERA used an FTIR field spectrometer for emissivity determination, THEMACS measured channel-specific 'broadband' emissivities using novel emissiometer [4] [21]. In addition to the in-situ measurements, THEMACS obtained emissivity spectra of several samples in the laboratory using a FTIR spectrometer.

4.1 KT15.85 IIP radiometer with emissivity box (KIT)

In November 2011 in-situ measurements with the 'one-lid emissivity box' method [25] were performed to determine the emissivities of relevant surface types at Gobabeb. Based on these measurements and assuming a dry grass fraction of 25% [2], the land surface emissivity (LSE) for the gravel plains was estimated as 0.944 ± 0.015 for SEVIRI channel 9 [9]. This value was also shown to be in good agreement with LSE derived with the temperature emissivity separation (TES) algorithm [7] from ASTER and MODIS data [12]. The corresponding LSE for the KT15.85IIP radiometer in 2011 / 2012 was estimated as 0.940 ± 0.015. However, in June 2017 the dry grass fraction over the gravel plains was zero, i.e. the surface was bare (Figure 5 and Figure 6). Therefore, the KT15.85IIP emissivity value of 0.916 ± 0.007 for pure, undisturbed gravel [9] was taken as representative for the conditions encountered during the LST FICE.

Figure 5. The main mast 'GBB Wind' and the measurement area on the 17th of June 2017.

Figure 6. The secondary mast 'GBB Plains' on the Namib gravel plains (June 2017).

4.2 BOMEM MR304SC FTIR spectroradiometer (ONERA)

The emissivity is obtained with an iterative algorithm based on the spectral smoothness in accordance to Borel's method [1]. The downwelling spectral irradiance I is estimated by correcting the upwelling radiance $Rref$ from the residual self-emission of the reflecting plate:

$$I = \frac{\pi}{\rho_{ref}}\left[R_{ref} - [1 -]B(T_{ref})\right]$$

where ρ_{ref} is the infragold spectral reflectance and T_{ref} is its surface temperature. The reflectance was measured in the laboratory with a Bruker Equinox 55 spectrometer equipped with an integrating sphere while the surface temperature was monitored during the experiment with a PRT probe in contact with the back of the metal plate.

For a given surface temperature, the spectral emissivity of the sample, ε_s, is solution of:

$$\varepsilon_S = \frac{R_S - I/\pi}{B(T_S) - I/\pi}$$

The emissivity of a solid is spectrally smooth. If the chosen temperature is far from the real surface temperature, the above equation will generate high frequency water vapor

absorption features in the emissivity, coming from the downwelling irradiance. Therefore an iterative algorithm selects the smoothest spectrum and its associated temperature as the solution.

4.2.1 Error budget assessment

The error of the method has been assessed through inter comparison with laboratory measurements over a set of manmade and natural targets [4][6]. Statistically, the emissivity deviation to laboratory remains less than 0.02 in the LWIR band. An error budget was evaluated by considering the error sources given in Table 1.

Blackbody temperature accuracy	±0.2 K (independent for the 2 reference measurements)
Blackbody emissivity accuracy	±0.015
Ambient temperature uncertainty	±3 K
Reflector reflectance uncertainty	±0.03
Reflector surface temperature uncertainty	±2 K

Table 7. Sources of uncertainty considered for the error budget.

Thus, for each sample, 3^6 =729 emissivities were calculated, considering the 6 sources of error (the blackbody uncertainty is applied independently to each temperature) and for each parameter, the nominal value and the nominal plus or minus the uncertainty. By considering these three values with a constant weight for each parameter, the associated probability density is assumed to be a uniform distribution with a standard deviation reduced by a factor $\sqrt{(2/3)} \cong 0.82$.

The mean value and the standard deviation (amplified by a factor of 1/0.82 =1.22) of the dataset respectively represent the best estimate and the rms uncertainty of the retrieved emissivity spectra.

The instrument noise and temporal instability in outdoor conditions is not included in this error budget. Therefore a standard deviation threshold is systematically applied. This threshold is assumed to be half of the deviation with the lab measurement (0.01) and account for the global deviation to the lab measurement mentioned previously, implicitly balancing the deviation in equal parts for the in-lab and outdoor measurements.

4.2.2 Results

The locations and samples for which in-situ emissivities were determined are described in Table 8.

Date	Area	Number of measurements	Comments
17/06/2017	Mast 'GBB Wind'	14	7 spots around the fence, shared with Themacs; 2 measurements per spot
22/06/2017	Mast2 'GBB Plains' (north mast)	5	Along a 20 m line
22/06/2017	Mast2 'GBB Plains' (north mast)	4	measurements on disturbed soil (the gravel is covered by sand/dust)
23/06/2017	GRTC	16	3 sets of samples. 30m between sets, each set covering a 5 m line
24/06/2017	ROAD	10	Along a 30m line at the starting point of the road experiment

Table 8. Description of locations where emissivity measurements were performed.

Individual and mean emissivity spectra retrieved for the 5 different measurement locations around Gobabeb are shown in Figure 7 and Figure 8.

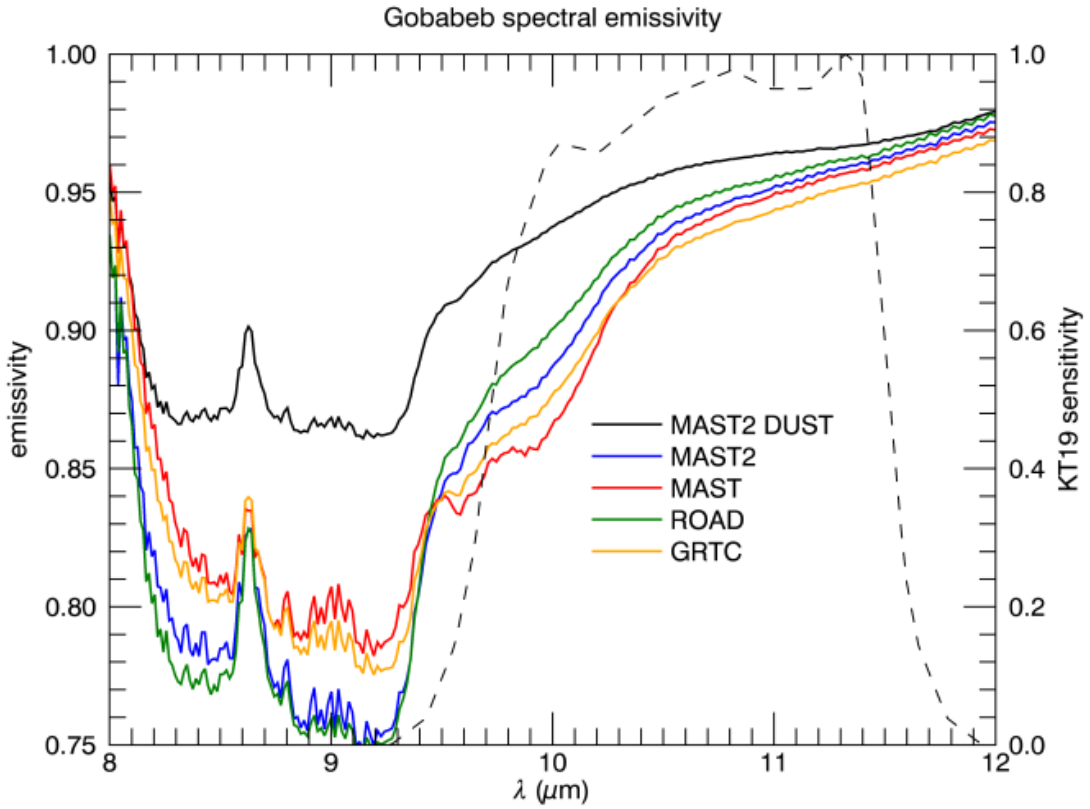

Figure 7. Mean spectral emissivities for each of the 5 Gobabeb datasets described in Table 8 . The broken lines indicate the spectral response function of the KT19.85II.

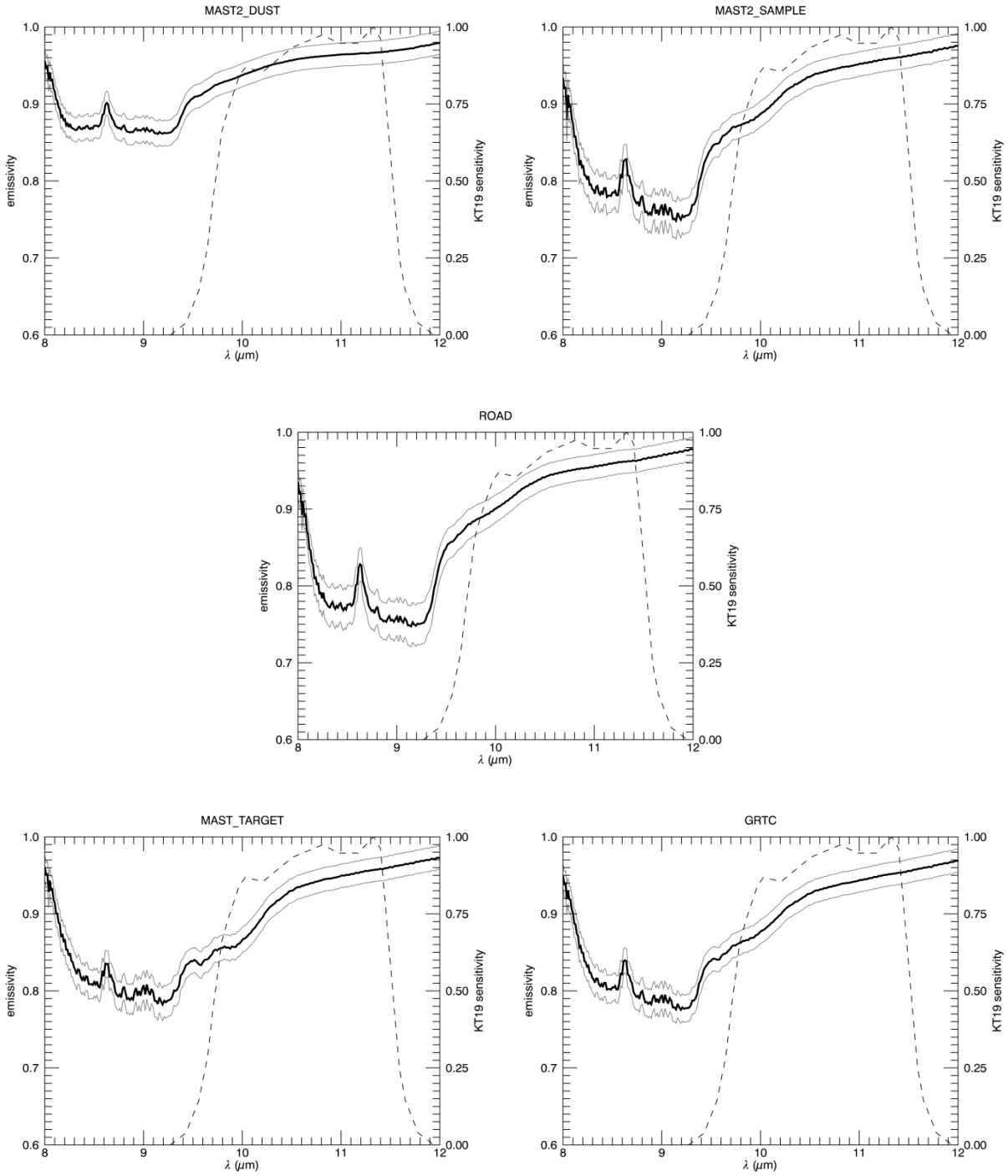

Figure 8. Spectral emissivities for the 5 datasets described in Table 8. The broken lines indicate the spectral response function of the KT19.85II. 'MAST' corresponds to GBB Wind shown in Figure 5.

The uncertainty in retrieved emissivity is composed of both, the variability of the dataset and measurement uncertainty. Table 9 shows the equivalent emissivity computed for the KT19.85II spectral response function provided by the manufacturer. The error budget should be enlarged due to spectral response uncertainty.

Target	Mean	Standard deviation
MAST2_DUST	0.953	0.015
MAST2_SAMPLE	0.927	0.016
MAST_TARGET_GLOB	0.918	0.016
ROAD	0.933	0.016
GRTC	0.918	0.015

Table 9. Equivalent emissivities for the KT19.85II radiometer at various locations. (MAST2 = GBB Plains, MAST = GBB Wind, GRTC = Gobabeb Research and Training Centre)

The mast emissivity measurement has been used in the LST retrieval for both experiments. This solution was preferred because of the FICE objective; the other teams used published or measured data obtained in the vicinity of the "GBB wind" mast. Therefore, in order not to introduce a bias due to the origin of the soil emissivity, the average of the mast measurements was selected for the LST processing. Of course, for the purpose of a comparison with satellite data, for the drive experiments the track emissivity should be preferred.

4.3 EM3 emissometer (THEMACS)

Broadband emissivity measurements were performed around GBB Wind (Figure 5). The numbers of the measurement points correspond to the points marked around the fence surrounding the mast. The same points were used by Laurent Poutier and Stéphane Langlois from ONERA for their spectral measurements. The standard deviation on the EM3 emissivity is between 0.005 and 0.009. This brings the expanded uncertainty to a level of 0.015 to 0.027.

In view of the measurements made on this type of soil, an average emissivity of 0.915 may be considered as the reference value. The raw EM3 broadband (2-17µm) measurements are given in Table 10. The averages of the emissivity measurements made at each point at different times are given in Table 11.

N° measurement	Location	Emissivity
1	Point1	0.923
2	Point2	0.904
3	Point2	0.923
4	Point2	0.919
5	Point3	0.901
6	Point3	0.915
7	Point4	0.927
8	Point4	0.914
9	Point5	0.913
10	Point5	0.925
11	Point6	0.902
12	Point6	0.918
13	Point6	0.926
14	Point7	0.918
15	Point7	0.907
	Mean	**0.917**
	standard dev.	**0.009**

Table 10. Raw EM3 broadband (1-50µm) emissivity measurements.

Location	Emissivity
Point1	0.923
Point2	0.915
Point3	0.908
Point4	0.921
Point5	0.919
Point6	0.915
Point7	0.912
Mean	**0.916**
standard dev.	**0.005**

Table 11. Broadband emissivity averages at the 7 locations.

4.4 Band III emissivities from EM-3 emissometer (THEMACS)

Additional emissivity measurements were carried out using a different thermopile with a narrower 'band III' (8-14µm, see Figure 9). The corresponding filter function provided by the thermopile manufacturer Dexter Research is shown in Figure 9. The spectral characteristics of the band III thermopile are more similar to those of the Heitronics radiometers KT19.85II and KT15.85IIP (both 9.6-11.5µm), which were deployed by ONERA and KIT, respectively. However, this band III still differs considerably from the KT19.85II band (Figure 10).

Therefore, the emissivities determined for the different sensors cannot be expected to be identical. The raw band III emissivity measurements are provided in Table 12 and the corresponding emissivity averages for each location are given in Table 13.

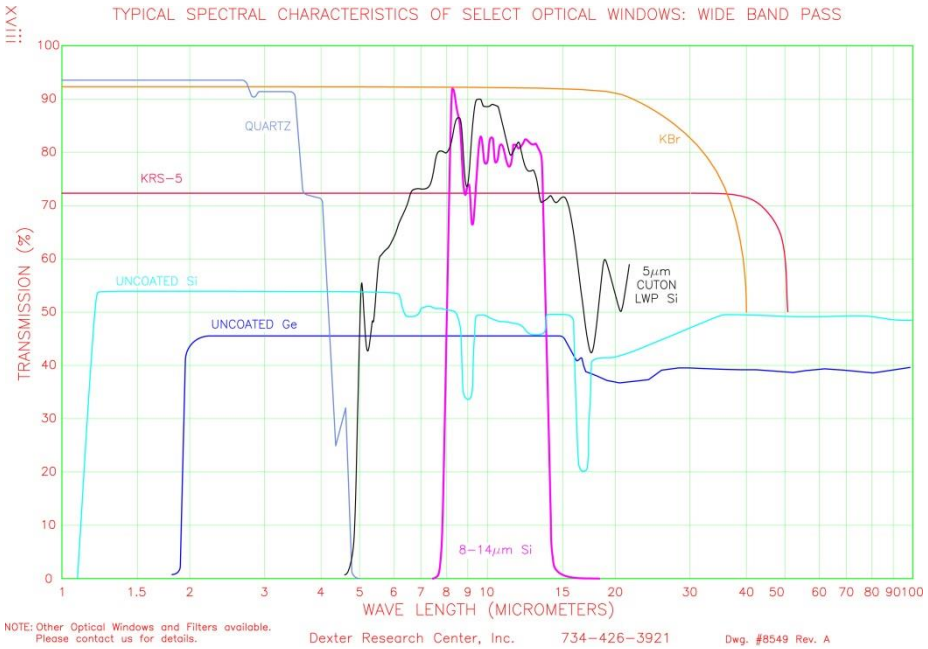

Figure 9. The spectral 'band III' (purple, annotated '8-14µm Si') of the thermopile from Dexter Research Center Incorporation. The Heitronics KT19.85II spectral response function (9.6-11.5µm) is shown in Figure 7 and Figure 8.

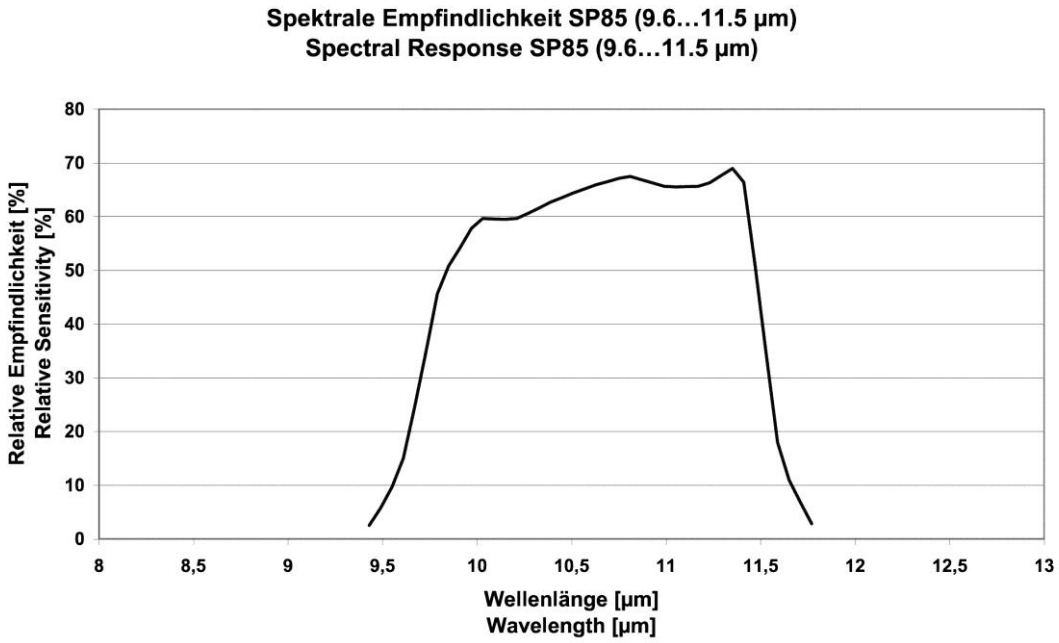

Figure 10. Sensitivity of KT19-85.II from Heitronics® documentation.

Location	Emissivity
point1	0.928
point1	0.907
point1	0.910
point1	0.911
point2	0.900
point2	0.899
point3	0.890
point3	0.892
point3	0.902
point3	0.887
point4	0.917
point5	0.877
point6	0.902
point6	0.902
point6	0.902
mean	**0.902**
standard dev.	**0.012**

Table 12. Raw band III (8-14µm) emissivities measured with the EM3 and their mean & standard deviation.

Location	Emissivity
point1	0.914
point2	0.900
point3	0.893
point4	0.917
point5	0.877
point6	0.902
mean	**0.901**
standard dev.	**0.015**

Table 13. Average band III (8-14µm) emissivities measured with the EM3 by location and their mean & standard deviation.

4.5 Laboratory measurements with Frontier Perkin-Elmer spectrometer (THEMACS)

Additional measurements were performed on samples in the laboratory, i.e. not under in-situ conditions in the field. The employed spectrometer was a Frontier® Perkin-Elmer® (Figure 11A) with an integrating sphere Pike®. Measurements were made separately on sand and gravel (supposedly quartz) pieces (Figure 11 B-D).

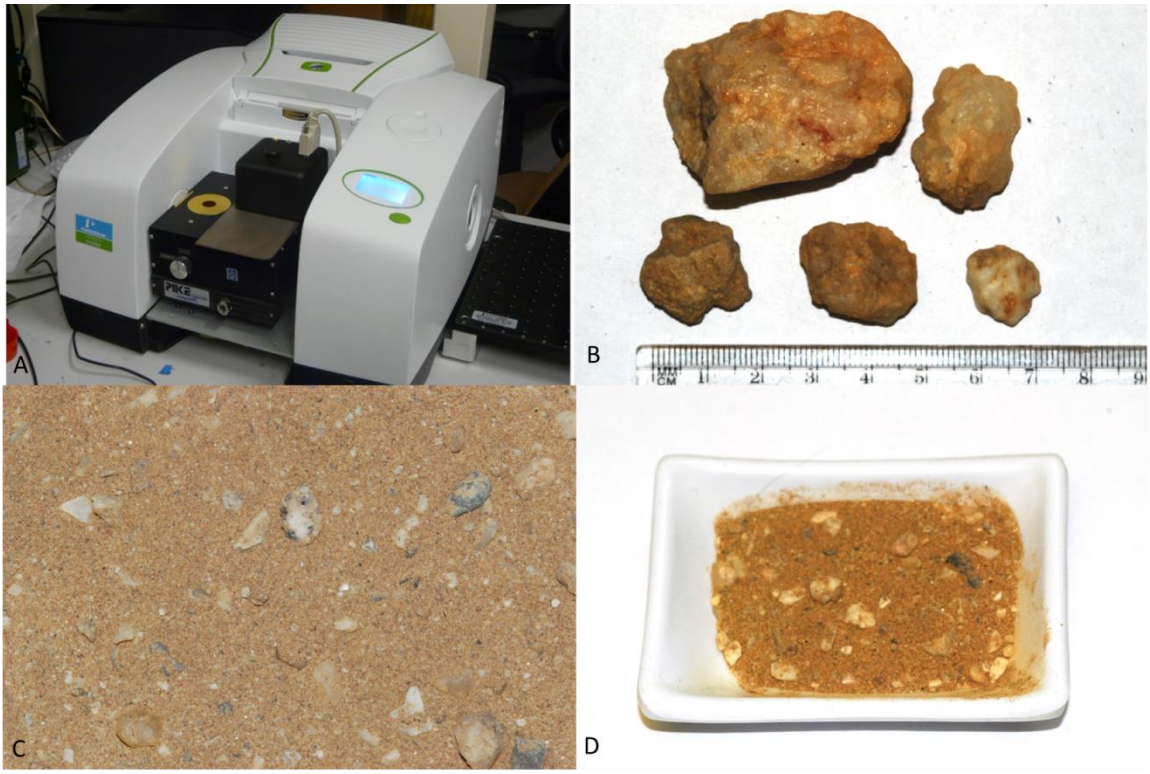

Figure 11. A: Frontier® spectrometer, B: Gobabeb quartz gravel, C and D: Gobabeb sand.

The obtained emissivity spectra for the samples of gravel (4) and sand (3) are shown in Figure 12 and Figure 13, respectively. It can be seen that the gravel found at Gobabeb has a high reflection factor (around 0.5) and is highly variable in the 8-14 µm band, whereas the sand has a considerably lower reflection factor (between 0.04 and 0.15) over the same spectral range. Consequently, the ratio between quartz gravel and sand strongly influences a samples reflection factor and, therefore, its emissivity, which can be approximated as the average of the emissivity of sand and gravel weighted by the respective fractional surface areas within the sensors FOV. From the emissivity spectra obtained for the samples band-effective emissivities for the following four spectral bands were derived:

- 2-17µm: valid range for the integrating sphere; EM3 broad band sensitivity
- 8-14µm: Sensitivity of EM3 when used with band III thermopile
- 8-12µm: Sensitivity of the A35 thermal camera
- 9.6-11.5µm: Sensitivity of the KT-19.85II

The corresponding band-effective emissivities for sand and gravel at a temperature of 303K are presented in Table 14. Calculations for band 2-17µm were performed because this spectral range is similar to the broadband measurements of the EM3 emissiometer. The large reflectance values observed for the sand at lower wavelengths, e.g. at 2µm more than 0.4, explain the differences between the broadband in-situ emissivity obtained with the EM3 over the undisturbed surface at Gobabeb (ε_{ground}=0.916; Table 11) and the broadband emissivities obtained from the laboratory spectra for sand (ε_{sand}=0.937) and gravel (ε_{gravel}=0.873). Finally, apparent emissivities representative for the KT-19.85II and KT15.85IIP radiometers (9.6-11.5µm; see Figure 10) were calculated.

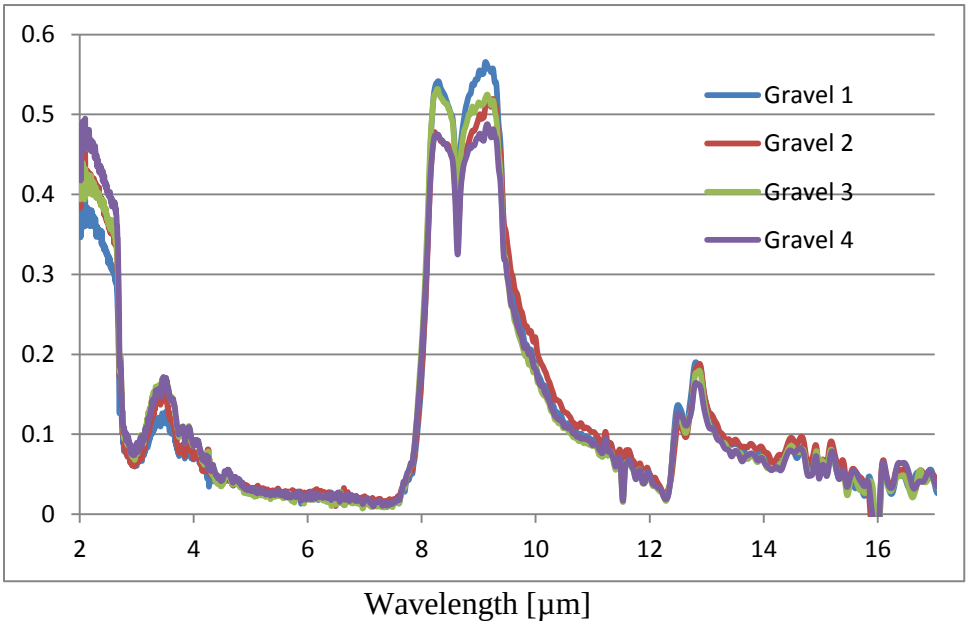

Figure 12. Spectral reflectance of four gravel samples obtained in the laboratory.

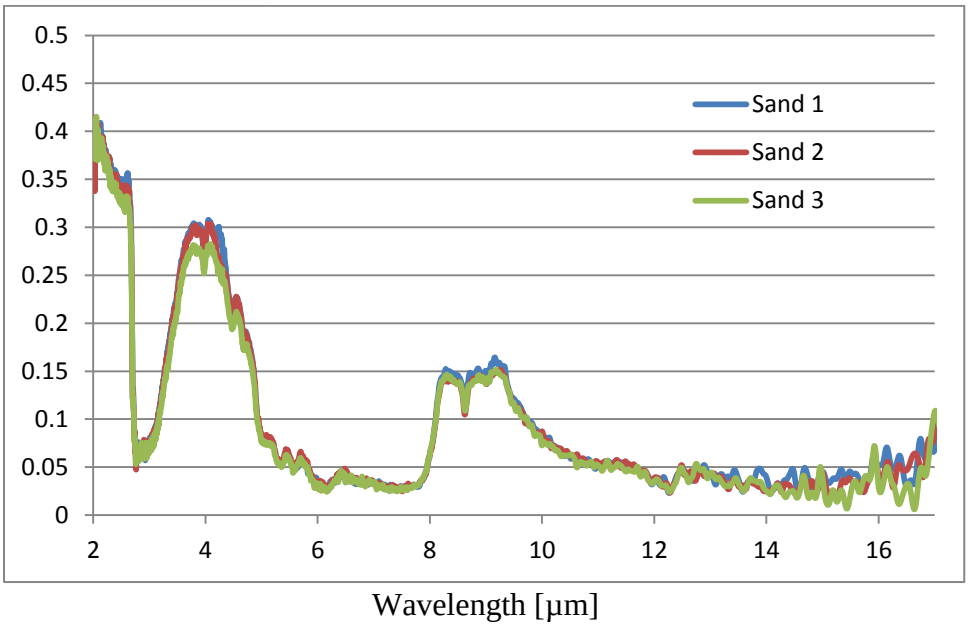

Figure 13. Spectral reflectance of three sand samples obtained in the laboratory.

	Emissivity for spectral band			
Material	2-17µm	8-14µm	8-12µm	9.6-11.5µm
Sand	0.937	0.925	0.909	0.937
Gravel	0.873	0.806	0.758	0.894

Table 14. Emissivities obtained from the spectra in Figure 12 and Figure 13 for different bands.

The different substrates of sand and gravel and the large differences in their spectral behaviour result in a strong dependence of in-situ emissivity on the ratio of these two components. The most meaningful comparison is for the 8-14µm (band III) spectral range: the EM3 band III measurements on the gravel plains yield an average in-situ emissivity of 0.901, while the band-effective emissivity derived from the spectral laboratory measurements is 0.925 for sand (Table 14), showing that the by far lower emissivity of 0.806 of gravel needs to be accounted for. We have supposed a flat response of the detector in EM3.

The differences between the various radiometers stem from the different spectral responses of their respective detectors. Therefore, the emissivity measurement device needs to be fitted with the same (type of) radiometer used for the brightness temperature measurements. For example, for the 9.6-11.5µm spectral range (e.g. a KT19.85 II radiometer) a band-effective emissivity of 0.937 is derived from the emissivity spectra of sand (Table 14). The corresponding in-situ emissivity is lower if the surface is covered by more gravel and increases with surface roughness. The emissivity values in Table 14 are in good agreement with previous measurements performed for the Namib gravel plains [9].

Further investigations are required to explain some of the remaining differences observed in the EM3 measurements.

5 LST INTER-COMPARISONS

The 2017 LST FICE in Namibia covered two parts: a temporal inter-comparison over four days, during which five field radiometers were mounted to the mast of the permanent LST validation station 'GBB Wind', and a spatial inter-comparison, during which the five field radiometers were driven along a track north-east of GRTC. The locations of GBB Wind, GBB Plains as well as the nominal locations of ISAR measurements along the track are shown in Figure 14.

Figure 14 Locations of GBB Wind and GBB Plains and nominal locations of ISAR measurements for the two 'drive experiments' across the Namib gravel plains (red circles).

The main purpose of the FICE was to inter-compare in-situ LST determined by the different measurement teams under field conditions. The temporal inter-comparison compared daytime and night-time in-situ LST retrieved from radiometers aligned to view the same natural target. The spatial inter-comparison investigated if representative in-situ LST can be obtained from the five non-aligned and vibrating radiometers.

Figure 15 shows some basic meteorological parameters measured at GBB Plains and GBB Wind during the LST FICE 2017. The broadband down-welling shortwave radiance (lower panel, black) shows that only the 18th and 19th of June 2017 had a substantial amount of clouds: for these two days the otherwise sinusoidal curves of solar irradiance are disturbed. The observed dips around 10h UTC are caused by a shadow falling onto the radiance sensor at this time of the day. On the 23rd and 24th of June around 10h UTC wind speed (lower panel, grey) increased to more than 10 m/s. Together with relative humidity values (upper panel, blue) of 10% - 70% this indicates that during the seven days there were no fog events.

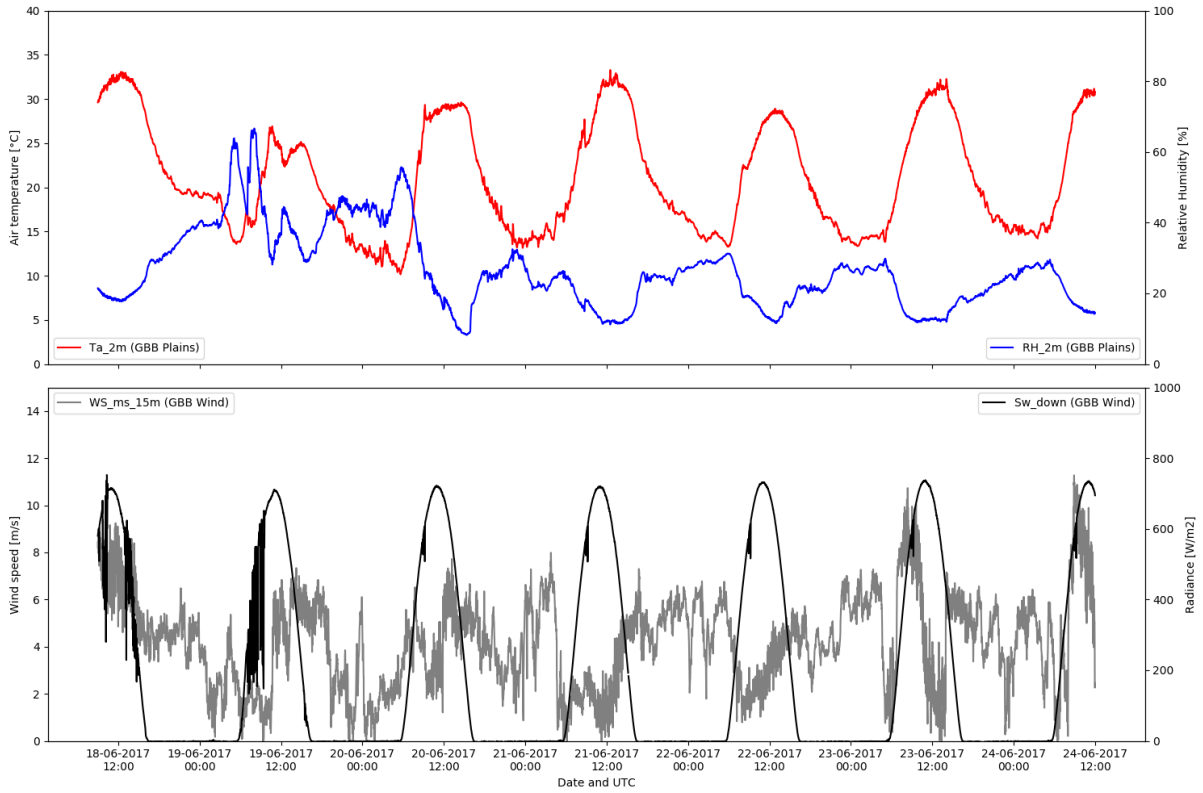

Figure 15 Top: Air temperature (red, left axis) and relative humidity (blue, right axis) at GBB Plains (2m height) during the LST FICE 2017 at Gobabeb, Namibia. Bottom: Wind speed at 15m height (grey, left axis) and broadband down-welling shortwave radiance (black, right axis) measured at GBB Wind.

5.1 Temporal inter-comparison

The first part of the inter-comparison took place at mast GBB Wind: the radiometers were mounted between 11m and 15m height and aligned to observe the undisturbed gravel surface north of the mast (see Figure 5, Figure 16 and the cover page of this report). The viewing direction was determined by the ISAR, which weighs about 20 kg and was mounted to the mast without offering the possibility for an azimuth alignment. The ISAR was programmed by NOCS to observe the surface at a scan (viewing) angle of 35°, which also determined the surface spot to be observed by the other radiometers. All radiometers were aligned to the ISAR spot by calculating the spots distance from the foot of the mast (8.4 m), marking it with an aluminum reflector (Figure 16) and then targeting it with the help of a laser designed for KT15.85 IIP and KT19.85 II radiometers.

Figure 16 Left: radiometers at GBB Wind mast; the ISAR (white cylinder) is at 12m height. Right: the FOVs of the radiometers were aligned to the spot marked by the aluminium reflector (centre-right; removed for measurements).

The in-situ LST provided by the participants were first re-sampled to a temporal resolution of 1 minute using linear interpolation, smoothed with a moving average filter (3 minute window), and then re-sampled to a final temporal resolution of 3 minutes.

Figure 17 shows the in-situ LST determined by the five teams for the temporal inter-comparison at GBB Wind from 9:00 UTC on the 18th of June 2017 to 9:00 UTC on the 22nd of June 2017. Until about noon of the 19th of June there was considerable cloud cover (see Figure 16, left); afterwards clear-sky situations dominated, resulting in smooth diurnal LST cycles (Figure 15). There is good agreement between the in-situ LST retrieved by NOCS, KIT, ONERA, and GOTA. In contrast, in-situ LST 'UV-ES broad' (i.e. from CIMEL 312-1 broadband channel 8-13 µm) deviate systematically from NOCS LST (the reference) and are on average 3°C higher; only around noon the level of agreement is similar to that of GOTA, who used the same type of radiometer (label 'GOTA_B2_ONERA' indicates in-situ LST retrieved from GOTA CIMEL band 2 using ONERA emissivity spectra). However, because in the spatial inter-comparisons (section 5.2) there is good agreement between 'UV-ES broad', GBB Wind and the other radiometers, the observed differences are considered to be real and not caused by a malfunctioning instrument. In order to analyse these differences, the corresponding in-situ LST from GBB Winds radiometer, which was not aligned with the ISAR and observes a different surface area, are also shown in Figure 17 (broken purple line): apart from the time around noon, there is good agreement between UV-ES broad and GBB Wind. However, the TRSG-UV team confirmed that they aligned the CIMEL 312-1 to the same surface spot as the other four radiometers. A possible explanation for the observed 'UV-ES broad' deviations could be non-negligible atmospheric path length (e.g. 17 m for the CIMEL), but this was ruled out by the TRSG-UV team on the basis of radiative transfer modelling (see 3.5.1). The in-situ

LST retrieved from the CIMELs narrow band channel (b2: 11.5-12.5 µm) are in considerably better agreement with those retrieved by the other teams. However, even though 'UV-ES narrow' LST is much closer to the reference, from the 20th of June onwards it deviates periodically from NOCS LST with amplitude of about 2°C (positive in the morning, negative in the afternoon).

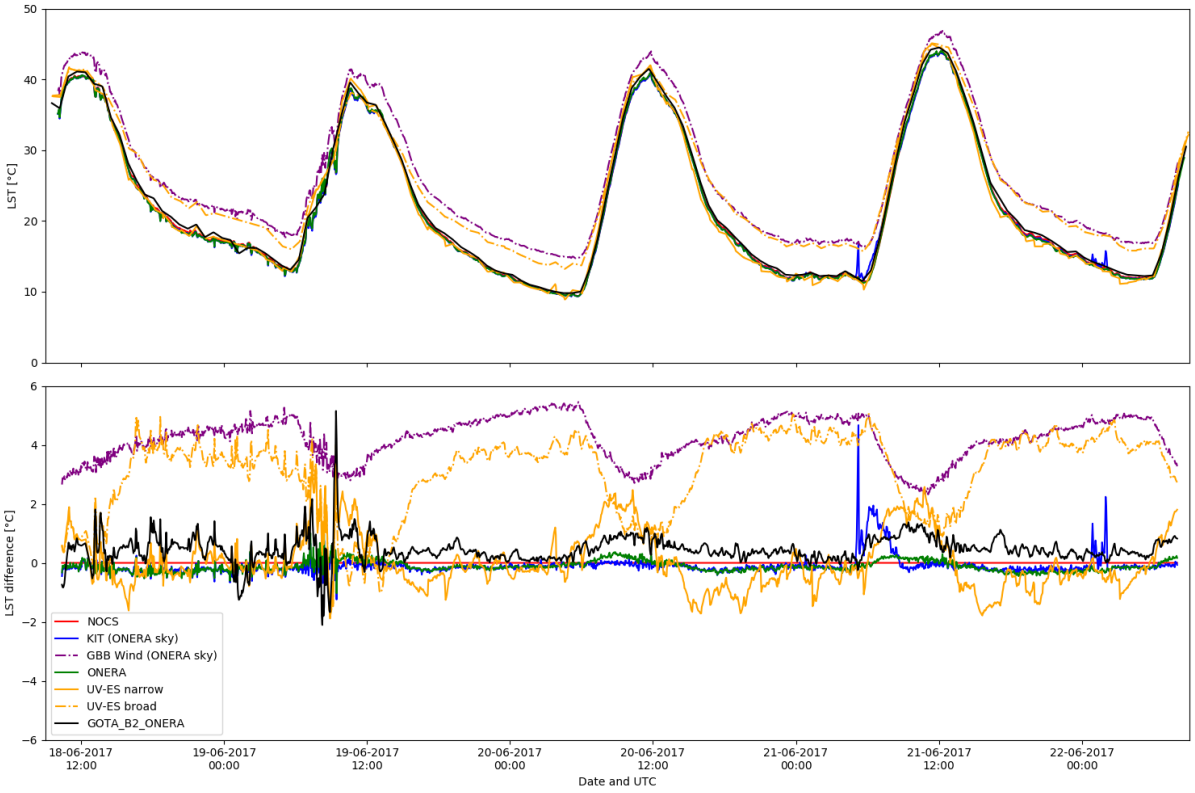

Figure 17 Top: in-situ LST retrieved by the five teams from the mast. The GBB Wind radiometer (KT15.85 IIP) observes a surface area about 7 m east from the spot observed by the other radiometers. Bottom: differences between the various in-situ LST and the ISAR in-situ LST (NOCS), which serves as reference.

On the 21st of June from about 06:00 UTC to 10:00 UTC and on the 22nd between 01:00 UTC to 02:00 UTC the LST retrieved by KIT deviates from NOCS LST (e.g. positive spikes): this is thought to be due to a piece of loose adhesive tape from an improvised sun shield, which may have partially obstructed the KT15.85 IIP radiometers view. The events also coincide with changes in wind speed (Figure 15).

Table 15 summarises the results obtained for the temporal inter-comparison at GBB wind mast. The statistics are with reference to the in-situ LST retrieved from the ISAR measurements, since this radiometer is continuously stabilised by two internal blackbodies [30] and, therefore, regarded as the most accurate. Interpolation and resampling to a 3 minute interval resulted in 1873 data points for the inter-comparison. Table 15 shows that the LST retrieved from narrow band measurements performed by the FICE participants generally agree well with ISAR LST: average deviation was 0.08 ± 0.47°C. Furthermore, when outliers of KITs KT15.85 IIP (due to loose adhesive tape) are ignored, the LST retrieved by KIT

and ONERA have similar bias and standard deviation of about -0.1°C and 0.2°C, respectively. The results for GBB Plains also demonstrate good agreement with ISAR LST, suggesting that the two observed surfaces, which are 5 km apart from each other, had very similar LST. In contrast, the in-situ LST retrieved from 'UV-ES broad' and 'GBB Wind' (the latter observes a surface area about 7 m east of the ISAR) differ systematically from ISAR LST, i.e. LST retrieved for 'GBB Wind' and 'UV-ES broad' had a mean difference of 4.21°C of 3.08°C, respectively. In contrast, the in-situ LST retrieved for CIMEL narrow band b2 ('UV-ES narrow') agrees considerably better with the reference (mean difference 0.04°C), but has a higher standard deviation (0.93°C): Figure 17 shows that this is due to an oscillation of UV-ES narrow LST on the 20th and 21st of June (and probably on the 22nd), which suggests that the effect of the protective window (see section 3.5) could not be completely compensated. Additionally, the good agreement of GOTA LST with NOCS LST (mean difference 0.46°C, standard deviation 0.45°C) suggests that the deviations of UV-ES LST are not related to radiometer type but to temperature variations of the protective window around noon.

Team / Instrument	Mean difference [°C]	Stdev of difference [°C]
KIT / KT15.85 IIP	-0.09	0.32
GBB Wind / KT15.85 IIP	4.21	0.72
ONERA / KT19.85 II	-0.11	0.18
UV-ES broad / CIMEL 312-1	3.08	1.36
UV-ES narrow / CIMEL 312-1	0.04	0.93
GOTA / CIMEL 312-2	0.46	0.45
GBB Plains / KT15.85 IIP	0.36	0.98

Table 15 Mean and standard deviation of the differences between the various in-situ LST and NOCS in-situ LST (ISAR = reference). The two in-situ LST from the permanent GBB stations (light blue) observe different surface areas. 'UV-ES broad' LST was derived from CIMEL broadband channel (8-13 µm). 'UV-ES narrow' LST was derived from the narrow band channel (11.5-12.5 µm).

5.2 Spatial inter-comparisons

The second part of the inter-comparison consisted of three return trips performed on the 23rd and 24th of June 2017 at different times of the day with two 4x4 vehicles. For these mobile measurements the radiometers of the five teams and a TIR camera (THEMACS) were mounted to horizontal booms (Figure 18). During the experiment the vehicles were kept at a (fairly) constant speed of about 12 km/h and generally followed each other within about 30 seconds (about 100 m distance between the vehicles). In order to exclude measurements recorded while the vehicles stopped or moved unsteadily, only in-situ LST obtained at speeds larger than 9.3 km/h are used in the analyses. The radiometers observed the ground from 1.8 m under a view angle of 35°, e.g. resulting in footprint diameter of about 30 cm for the KT15.85 IIP (8.5° full view angle). Figure 18 shows the two cars equipped with the radiometers on the gravel plains north-east of GRTC (tracks shown in Figure 14).

Figure 18 Left: vehicle with ISAR (NOCS), KT19.85 II (ONERA) and CIMEL 312-2 (GOTA). Right: vehicle with KT15.85 IIP (KIT), CIMEL 312-1 (UV-ES) and TIR camera (THEMACS). On the gravel plains between GRTC and the Inselberg Mirabib, Namibia.

Figure 19 shows close-ups of the ISAR (NOCS), KT19.85 II (ONERA) and CIMEL 312-2 (GOTA) on the one vehicle and of the KT15.85 IIP (KIT) and CIMEL 312-1 (TRSG-UV) on the other vehicle. Unlike in the temporal inter-comparison, the TRSG-UV CIMEL radiometer was deployed without a protective window (see section 3.5): therefore, the instruments calibration was accurate for all bands and retrieved LSTs agreed within their respective uncertainty limits with the mean of the five in-situ LSTs. Using the TES method, TRSG-UV additionally obtained an emissivity of 0.967±0.013 for CIMELs narrow band b2 over the two transects, which is close to the corresponding value of 0.974±0.016 obtained from ONERAs spectral measurements; consequently, the LSTs obtained with these two emissivity estimates are close to each other (mean difference of 0.49±0.04 K) and within their combined uncertainty limits.

Figure 19 Close-ups of the radiometers mounted to the 4x4 vehicles. Left: ISAR (NOCS), KT19.85 II (ONERA) and CIMEL 312-2 (GOTA). Right: KT15.85 IIP (KIT) and CIMEL 312-1 (TRSG-UV).

On the 23rd of June the vehicles drove 10 km in direction Mirabib and reversed; then they followed the same track for 8 km and reversed (i.e. a total of 36 km were covered). For these journeys the in-situ LST and their differences with respect to average in-situ LST are shown in 39

Figure 20 while Table 16 provides the corresponding statistics. From Figure 20 it can be seen that most of the in-situ LST are within ±1°C of their average; with a few exceptions, this also applies to the in-situ LST from GBB Wind.

Figure 20 Top: in-situ LST retrieved on the 23rd of June 2017 while driving across the gravel plains. The blue points refer to the right axis and give an estimate of vehicle speed: only in-situ LST for speeds > 2 deg/day (about 9.3 km/h) are taken into account. Bottom: differences between in-situ LST and mean in-situ LST determined. GBB Wind in-situ LST is shown for reference only (i.e. not included in mean LST).

Table 16 summarises the results obtained on the 23rd of June 2017 for the spatial inter-comparison across the gravel plains. The statistics are with reference to average in-situ LST, since the radiometers were not aligned and observed different parts of the land surface while being driven across the gravel plains. Interpolation and resampling of the five in-situ LST time series to a 1 minute interval yielded 156 valid data points for the inter-comparison (i.e. all five in-situ LST were present and the car sufficiently fast); at an estimated speed of 12 km/h averaging over 1 minute is equivalent to averaging over 200 m driving distance. Table 16 shows that the in-situ LST retrieved from the five radiometers agree well with each other with an average absolute deviation of 0.36°C from their mean and an average standard deviation of ± 0.41°C.

Team / Instrument	Mean difference [°C]	Stdev of difference [°C]
KIT / KT15.85 IIP	-0.21	0.35
ONERA / KT19.85 II	-0.05	0.31
UV-ES / CIMEL 312-1	-0.83	0.26
NOCS / ISAR	0.46	0.76
GOTA / CIMEL 312-2	0.63	0.39
GBB Wind / KT15.85 IIP	0.37	1.12

Table 16 Mean and standard deviation of the differences between team and mean in-situ LST (23rd of June 2017; Figure 20). The corresponding results for the permanent validation station GBB Wind are shown for reference only (light blue). Interpolation and resampling of the data to a 3 minute interval resulted in 156 data points for the inter-comparison.

On the 24th of June the vehicles drove 12 km in direction Mirabib and reversed (i.e. a total of 24 km were covered). For these journeys the in-situ LST and their differences with respect to average in-situ LST are shown in Figure 21 and Table 17 provides the corresponding statistics. From Figure 21 it can be seen that practically all in-situ LSTs are within ±1°C of their average: with a few exceptions, this also applies to the in-situ LST from GBB Wind. The slightly larger deviations at the beginning of the plot in Figure 21 (top panel) are thought to be caused by a more heterogeneous surface near the beginning of the track (at about 07:20 UTC). Over all, on June the 24th the in-situ LST obtained by the five participating teams are more similar to each other and show less spatial variability than those obtained on the 23rd of June, which is mainly due to spatially more homogeneous LST early in the morning (sunrise at 05:45 UTC).

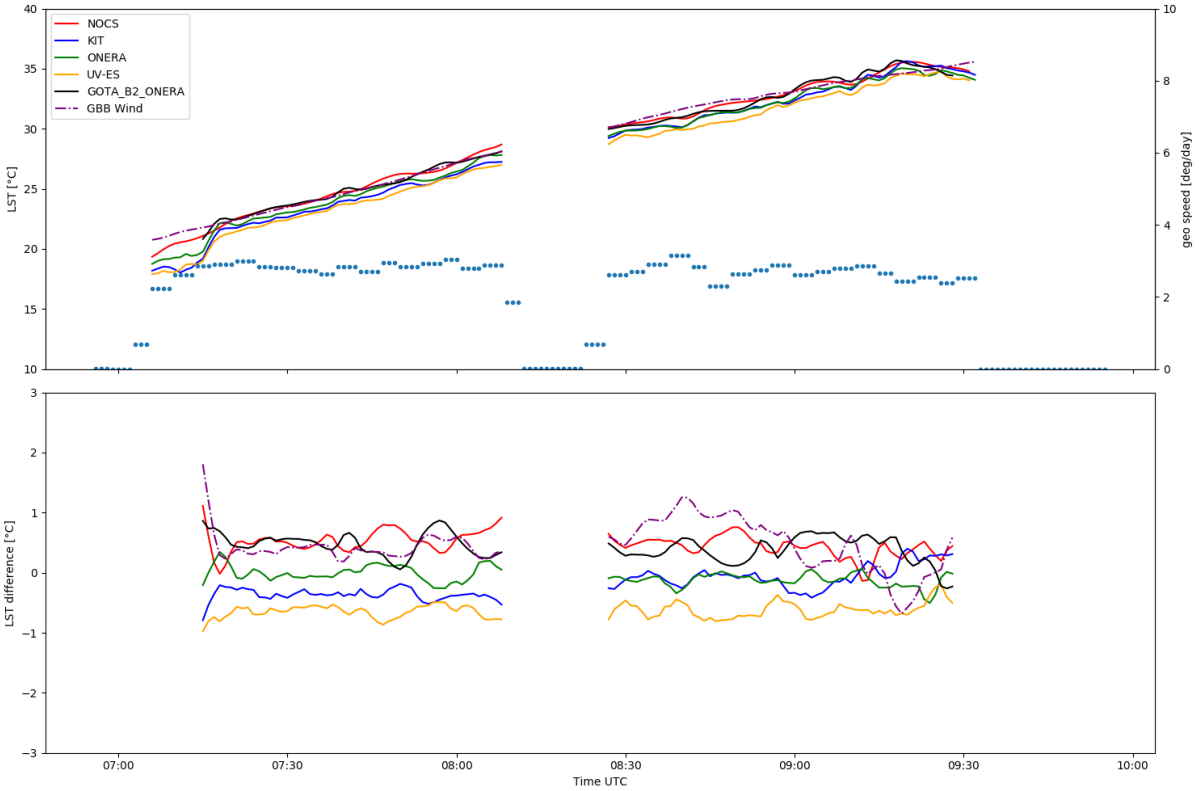

Figure 21 Top: in-situ LST retrieved on the 24th of June 2017 while driving across the gravel plains. The blue points refer to the right axis and give an estimate of vehicle speed: only in-situ LST for speeds > 2 deg/day (about 9.3 km/h) are taken into account. Bottom: differences between in-situ LST and mean in-situ LST determined. GBB Wind in-situ LST is shown for reference only (i.e. not included in mean LST).

Table 17 summarises the results obtained the 24th of June 2017 for the spatial inter-comparison across the gravel plains. The statistics are with reference to average in-situ LST, since the radiometers were not aligned and observed different parts of the land surface while being driven across the gravel plains. Interpolation and resampling of the five in-situ LST time series to a 1 minute interval yielded between 116 valid data points for the inter-comparison (i.e. all five in-situ LST were present and the car sufficiently fast). Table 17 shows that the in-situ LST retrieved from the five radiometers agree well with each other with an average absolute deviation of 0.44 °C from their mean and an average standard deviation of ± 0.18°C.

Team / Instrument	Mean difference [°C]	Stdev of difference [°C]
KIT / KT15.85 IIP	-0.20	0.22
ONERA / KT19.85 II	-0.07	0.14
UV-ES / CIMEL 312-1	-0.64	0.12
NOCS / ISAR	0.48	0.19
GOTA / CIMEL 312-2	0.43	0.22
GBB Wind / KT15.85 IIP	0.45	0.39

Table 17 Mean and standard deviation of the differences between team and mean in-situ LST (24th of June 2017; Figure 21). The corresponding results for the permanent validation station GBB Wind are shown for reference (light blue). Interpolation and resampling of the data to a 3 minute interval resulted in 116 data points for the inter-comparison.

6 SUMMARY AND CONCLUSIONS

The primary objective of the LST field inter-comparison experiment (FICE) was to perform inter-comparisons of TIR field radiometers and to verify their capability to provide fiducial reference measurements for validating land surface temperature (LST) products, e.g. as retrieved from Sentinel 3 and the Meteosat series. Five teams deployed different types of field radiometers on the gravel plains near Gobabeb Training and Research Centre (GRTC), Namibia: an Infrared Sea Surface Temperature Autonomous Radiometer (ISAR), a Heitronics KT 19.85IIP, a Heitronics KT 15.85IIP, a CIMEL 312-2 and two CIMEL 312-1. It was the first deployment of an ISAR over a desert site. Due to their individual technical specifications the various instruments measured at different sampling rates. In order to allow quantitative inter-comparisons, the in-situ LST provided by the participants were interpolated and resampled to a common sampling interval. All radiometers operated continuously during a four day temporal inter-comparison from a mast and during two spatial inter-comparisons, for which the instruments were mounted to vehicles and driven 60 km across the highly homogeneous Namib gravel plains. Additionally, in-situ land surface emissivities were obtained with a FTIR spectrometer (BOMEM MR304SC) and an 'emissiometer' (EM-3), a novel instrument that utilizes an active TIR radiance source. Furthermore, emissivity spectra of seven soil samples were retrieved in the laboratory.

For the temporal LST inter-comparison at the wind mast the five radiometers were mounted to heights between 11m and 15m, thereby ensuring that the observed surface areas exceeded $2m^2$, and aligned by laser to point to a common target. For the four day LST inter-comparison the average deviation from the chosen reference (ISAR) was 0.08°C ± 0.47°C. After removing outliers identified in one of the time series, two of the provided in-situ LSTs had standard deviations of about ± 0.2°C. Based on these results and assuming the measurement protocol [8] is followed, it is concluded that the investigated gravel surface is sufficiently homogeneous to inter-compare field radiometers. A larger bias of about 3 K was observed for in-situ LSTs retrieved by TRSG-UV from CIMELs wide band (b1: 8-13 µm) with an additional protective window, despite of applying a correction. The corresponding corrected LSTs retrieved from CIMELs narrow band (b2: 11.5-12.5 µm) were much closer to the reference, but still exhibited a oscillation that appears to be related to temperature fluctuations of the protective window. Therefore, it is recommended to avoid the use of protective windows.

For the spatial inter-comparisons the instruments could not be aligned with each other and observed different surface areas. Therefore, the mean of the five LST series was used as reference. In order to obtain spatially representative in-situ LST for the inter-comparison, in-situ LST were spatially averaged over about 200 m driving distance. The first set of measurements yielded an average absolute deviation and standard deviation of 0.36°C and ± 0.41°C, respectively. The second set of measurements was performed shortly after sunrise and yielded an average absolute deviation and standard deviation of 0.44°C and ± 0.18°C,

respectively. The smaller standard deviation is attributed to the spatially more homogenous LST field in the early morning.

The LST FICE again highlighted the importance of in-situ emissivities, since the corresponding values obtained for samples in the laboratory differed considerably from these. Furthermore, approximating in-situ emissivity as a weighted sum of endmembers (e.g. sand and gravel) does not reproduce effects from natural surface structure and texture. The field experiments provided an excellent opportunity to compare the performance of various instruments and measurement approaches and helped the participants to identify and understand differences between their results.

7 REFERENCES

[1] C.C. Borel. Surface emissivity and temperature retrieval for a hyperspectral sensor. In *1998 IEEE International Geoscience and Remote Sensing Symposium*, Los Alamos and New Mexico 87545 and USA, 1998. Los Alamos National Laboratory and NIS-2 and MS C323.

[2] Annika Bork-Unkelbach. *Extrapolation von in-situ Landoberflächentemperaturen auf Satellitenpixel*. Phd thesis, Karlsruher Institut für Technologie, 2012.

[3] Gérard Brogniez, Christophe Pietras, Michel Legrand, Philippe Dubuisson, and Martial Haeffelin. A high-accuracy multiwavelength radiometer for in situ measurements in the thermal infrared. part ii: Behavior in field experiments. *Journal of Atmospheric and Oceanic Technology*, 20(7):1023–1033, 2003.

[4] Y. Candau, L. Ibos, and J.P.Monchau. Dispositif de mesure de l'émissivité ou de la réflectivité d'une surface, patent 1251476, 2013.

[5] F.D. Eckardt, K. Soderberg, L.J. Coop, A.A. Muller, K.J. Vickery, R.D. Grandin, C. Jack, T.S. Kapalanga, and J. Henschel. The nature of moisture at Gobabeb, in the central Namib Desert. *Journal of Arid Environments*, 93:7–19, jun 2013.

[6] Vicente Garcia-Santos, Enric Valor, Vicente Caselles, Maria Mira, Joan Miquel Galve, and Cesar Coll. Evaluation of Different Methods to Retrieve the Hemispherical Downwelling Irradiance in the Thermal Infrared Region for Field Measurements. *IEEE Transactions on Geoscience and Remote Sensing*, 51(4):2155–2165, Apr 2013.

[7] A. Gillespie, S. Rokugawa, T. Matsunaga, J. S. Cothern, S. Hook, and A. B. Kahle. A temperature and emissivity separation algorithm for advanced spaceborne thermal emission and reflection radiometer (aster) images. *IEEE Transactions on Geoscience and Remote Sensing*, 36(4):1113–1126, 1998.

[8] Frank Göttsche, Folke S. Olesen, Jacob L. Høyer, Werenfred Wimmer, and Tim Nightingale. Fiducial Reference Measurements for Validation of Surface Temperature from Satellites (FRM4STS) - A Framework to Verify the Field Performance of TIR FRM. Technical Report 3, National Physical Laboratories (NPL), Teddington, UK, 2017.

[9] Frank-M. Göttsche and Glynn C. Hulley. Validation of six satellite-retrieved land surface emissivity products over two land cover types in a hyper-arid region. *Remote Sensing of Environment*, 124:149–158, sep 2012.

[10] Frank-M. Göttsche, Folke-S. Olesen, Isabel Trigo, Annika Bork-Unkelbach, and Maria Martin. Long term validation of land surface temperature retrieved from MSG/SEVIRI with

continuous in-situ measurements in Africa. *Remote Sensing*, 8(5):1–27, may 2016. Article number 410.

[11] Frank-Michael Göttsche, Folke-Sören Olesen, and Annika Bork-Unkelbach. Validation of land surface temperature derived from MSG/SEVIRI with in situ measurements at Gobabeb, Namibia. *International Journal of Remote Sensing*, 34(9-10):3069–3083, May 2013.

[12] Glynn C. Hulley and Simon J. Hook. Generating Consistent Land Surface Temperature and Emissivity Products Between ASTER and MODIS Data for Earth Science Research. *IEEE Transactions on Geoscience and Remote Sensing*, 49(4):1304–1315, Apr 2011.

[13] Glynn C. Hulley, Simon J. Hook, and Alice M. Baldridge. Validation of the North American ASTER Land Surface Emissivity Database (NAALSED) version 2.0 using pseudo-invariant sand dune sites. *Remote Sensing of Environment*, 113(10):2224–2233, Oct 2009.

[14] Glynn C. Hulley, Simon J. Hook, Evan Manning, Sung-Yung Lee, and Eric Fetzer. Validation of the Atmospheric Infrared Sounder (AIRS) version 5 land surface emissivity product over the Namib and Kalahari deserts. *Journal of Geophysical Research*, 114(D19104):1–11, 2009.

[15] Keyvan Kanani. *Utilisation de la très haute résolution spectrale pour la mesure en environnement extérieur de l'émissivité de surface dans la bande infrarouge 3-13 Mum: Méthodes et validation expérimentale*. PhD thesis, Université Louis Pasteur (Strasbourg), 2005.

[16] Keyvan Kanani, Laurent Poutier, Francoise Nerry, and Marc-Philippe Stoll. Directional effects consideration to improve out-doors emissivity retrieval in the 3-13 µm domain. *Optics Express*, 15(19):12464–12482, September 2007.

[17] W. Köppen. *Das geographische System der Klimate*. Handbuch der Klimatologie. Gebr. Bornträger, 1936.

[18] J. Lancaster, N. Lancaster, and M.K. Seely. Climate of the central namib desert. *Madoqua*, 14(1):5–61, 1984.

[19] Michel Legrand, Christophe Pietras, Gérard Brogniez, Martial Haeffelin, Nader Khalil Abuhassan, and Michaël Sicard. A high-accuracy multiwavelength radiometer for in situ measurements in the thermal infrared. part i: Characterization of the instrument. *Journal of Atmospheric and Oceanic Technology*, 17(9):1203–1214, 2000.

[20] S. J. Mason and M. R. Jury. Climate variability and change over southern africa: a reflection on underlying processes. *Progress in Physical Geography*, 21:25–50, 1997.

[21] Jean-Pierre Monchau, Mario Marchetti, Laurent Ibos, Jean Dumoulin, Vincent Feuillet, and Yves Candau. Infrared emissivity measurements of building and civil engineering

materials: A new device for measuring emissivity. *International Journal of Thermophysics*, 35(9-10):1817–1831, jun 2014.

[22] M. C. Peel, B. L. Finlayson, and T. A. McMahon. Updated world map of the köppen-geiger climate classification. *Hydrology and Earth System Sciences*, 11:1633–1644, 2007.

[23] L. Poutier. Laboratory and land surface measurements at npl for frm4sts project. Technical Report RT 1/26303 DOTA, ONERA, 2016.

[24] Laurent Poutier. Complementary investigations of the mikron m345 x4 onera blackbody. Technical Report 2/26303, ONERA, 2017.

[25] E. Rubio, V. Caselles, and C. Badenas. Emissivity measurements of several soils and vegetation types in the 8-14 micrometer wave band: Analysis of two field methods. *Remote Sensing of Environment*, 59:490–521, 1997.

[26] M. K. Seely. Standing crop as an index of precipitation in the central namib grassland. *Madoqua*, 9(3):5–13, 1978.

[27] M. Sicard, P.R. Spyak, M. Legrand, N.K. Abuhassan, C. Pietras, and J.-P. Buis. Thermal-infrared field radiometer for vicarious cross-calibration: characterization and comparisons with other field instruments. *Optical Engineering*, 38(2):345–356, 1999.

[28] E. Theocharous, I. Barker Snook, and N.P. Fox. 2016 comparison of ir brightness temperature measurements in support of satellite validation. part 1: Blackbody laboratory comparison. NPL report ENV12, National Physical Laboratory (NPL), 2016.

[29] E. Theocharous, E. Usadi, and N.P. Fox. 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 OP 3. Technical Report ISSN: 1754-2944, National Physical Laboratory, Teddington, UK, 2010.

[30] Werenfrid Wimmer and Ian S. Robinson. The ISAR Instrument Uncertainty Model. *Journal of Atmospheric and Oceanic Technology*, 33(11):2415–2433, 2016.

[31] Werenfrid Wimmer, Ian S. Robinson, and Craig J. Donlon. Long-term validation of AATSR SST data products using shipborne radiometry in the Bay of Biscay and English Channel. *Remote Sensing of Environment*, 116:17–31, Jan 2012.