AMERICAN
METEOROLOGICAL
SOCIETY

Journal of Atmospheric and Oceanic Technology

The DOI for this manuscript is doi: 10.1175/JTECH-D-18-0032.1

The final published version of this manuscript will replace the preliminary version at the above DOI once it is available.

The 2016 CEOS infrared radiometer comparison: Part 2: Laboratory comparison of radiation thermometers.

by

E. Theocharous[1], N. P. Fox[1], I. Barker-Snook[1], R. Niclòs[2], V. Garcia Santos[2], P. J. Minnett[3], F. M. Göttsche[4], L. Poutier[5], N. Morgan[6], T. Nightingale[7], W. Wimmer[8], J. Høyer[9], K. Zhang[10], M. Yang[10], L. Guan[10], M. Arbelo[11] , C. J. Donlon[12]

[1] NPL, Hampton Road, Teddington, TW11 0LW, UK.

[2] Faculty of Physics, University of Valencia, Dr. Moliner, 50, 46100 Burjassot, Spain.

[3] Rosenstiel School, University of Miami, 4600 Rickenbacker Causeway, Miami, FL 33149, USA.

[4] IMK-ASF, Karlsruhe Institute of Technology (KIT), Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany.

[5] ONERA, 2, avenue Edouard Belin, 31055 Toulouse Cedex4, France.

[6] CSIRO Oceans and Atmosphere, Castray Esplanade, Hobart 7000, Australia.

[7] STFC Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, UK.

[8] Southampton University, European Way, Southampton, SO14 3ZH, UK.

[9] Danish Meteorological Institute (DMI), Centre for Ocean and Ice, Lyngbyvej 100, 2100 København Denmark.

[10] Ocean University of China, 238 Songling Road, Qingdao, 266100, China.

[11] Grupo de Observacion de la Tierra y la Atmosfera (GOTA), ULL, Spain.

[12] European Space Agency, Keplerlaan 1, 2201 AZ, Noordwijk, The Netherlands

Corresponding author address: E. Theocharous, Environmental Climate and Optical Group, NPL, Hampton Road, Teddington TW11 0LW, UK.

E-mail: e.theo@npl.co.uk

Abstract

To ensure confidence, measurements carried out by imaging radiometers mounted on satellites require robust validation using 'fiducial quality' measurements of the same 'in-situ' parameter. For surface temperature measurements this is optimally carried out by radiometers measuring radiation emitted in the infrared region of the spectrum, co-located to that of a satellite overpass. For ocean surface temperatures the radiometers are usually on-board ships to sample large areas but for Land and Ice they are typically deployed at defined geographical sites. It is of course critical that the validation measurements and associated instrumentation are internationally consistent and traceable to international standards. The Committee on Earth Observation Satellites (CEOS) facilitates this process and over the last two decades has organised a series of comparisons, initially to develop and share best practise, but now to assess metrological uncertainties and degree of consistency of all the participants. The fourth CEOS comparison of validation instrumentation: blackbodies and infrared radiometers, was held at the National Physical Laboratory (NPL) during June and July 2016 sponsored by the European Space Agency (ESA). The 2016 campaign was completed over a period of three weeks and included not only laboratory based measurements but also representative measurements carried out in field conditions, over land and water. This paper is one of a series and reports the results obtained when radiometers participating in this comparison were used to measure the radiance temperature of the NPL ammonia heat-pipe blackbody during the 2016 comparison activities i.e. an assessment of radiometer performance compared to international standards. This

comparison showed that the differences between the participating radiometer readings and the corresponding temperature of the reference blackbody were within the uncertainty of the measurements but there were a few exception, particularly for a reference blackbody temperature of -30 ^{o}C. Reasons which give rise to the discrepancies observed at the low blackbody temperatures were identified.

1 INTRODUCTION

The measurement of the Earth's surface temperature and, more fundamentally, its temporal and spatial variation, is a critical operational product for meteorology and an essential parameter for climate monitoring (Yoder et al., 2014). Satellites have been monitoring global surface temperature for some time. However, it is essential for long-term records that such measurements are fully anchored to international physical standards as represented by the Systeme International (SI) units. Field-deployed infrared radiometers[1] currently provide the most accurate measurements of the Sea Surface Temperature and are currently used for calibration and validation of Earth observation radiometers (Minnett and Corlett, 2012). These radiometers are in principle calibrated traceably to SI units, generally through a blackbody radiator. However, they are of varying design and are operated by different teams in different parts of the globe, and the quality of the blackbody radiator can be variable. It is essential for the integrity of their use, that any differences in their measurements are understood, so that any potential biases are removed and are not transferred to satellite sensors (Minnett et al., 2002). One way of ensuring this is for the radiometers to be calibrated against a common high quality

1 This report describes the comparison of instruments which are referred to by participants as "radiometers". However, radiometers generally measure and report radiometric parameters in radiometric units (W, Wm^{-2}, etc.). The instruments dealt with here measure temperature (in units of degrees C or K) so they are thermometers or "radiation thermometers". However, in view of the common usage of the terminology for this application, this report will continue to use the term "radiometer".

SI traceable blackbody and be tested alongside each other under field conditions. As part of this process, it is also essential that each radiometer and its procedure for use is well-documented and a detailed uncertainty budget related to the traceability of its measurements to SI units is created. To recognise this rigour and distinguish such measurements from other 'in-situ' measurements the term 'fiducial reference measurements' (FRM) has been established (https://earth.esa.int/web/sppa/activities/frm) and is being used for similar measurements of other Earth Observation parameters e.g. Ocean colour, Sea height etc.

Previous CEOS comparisons of terrestrial based infrared (IR) radiometric instrumentation used to support calibration and validation of satellite borne sensors, with emphasis on sea/water surface temperature, were completed in Miami (Florida, USA) in 2001 (Barton et al., 2004) (Rice et al., 2004) and at the National Physical Laboratory (NPL) and Miami (Florida, USA) in 2009 (Theocharous and Fox, 2010) (Theocharous et al., 2010). However, seven years had passed, and many of the satellite sensors originally underpinned were at best nearing the end of their life. Under the auspices of the Committee of Earth Observation Satellites (CEOS), the European Space Agency (ESA) established a new comparison of terrestrial based Infra Red (IR) radiometric instrumentation, in this case with their use expanded to support calibration and validation of satellite borne sensors for sea/water/land/ice surface temperature, this was completed at NPL during June and July 2016. The expansion of applications reflected the capabilities of new sensors such as the Sea and Land Surface Temperature Radiometer (SLSTR) on the Copernicus Sentinel 3 satellite and the increasing importance of Land and Ice temperature measurements, particularly for climate monitoring. The objectives of the 2016 comparison were to establish the "degree of equivalence" between terrestrially based IR Calibration/Validation (Cal/Val) measurements made in support of satellite observations of the Earth's surface temperature and to ensure their traceability to SI units through the participation of National Metrology Institutes (NMIs). The comparison was organised through an ESA project called Fiducial Reference Measurements for Surface Temperatures derived by Satellite (FRM4STS) which also carried out a critical review of community measurement practises, details can be found at http://www.FRM4STS.org .

During the 2016 comparison, NPL acted as the pilot laboratory and provided traceability to SI units during laboratory comparisons. Stage 1 consisted of Lab comparisons, and took place at NPL during the week starting on 20th June 2016. This Stage involved laboratory measurements of participants' blackbodies calibrated using the NPL Absolute Measurement of a Blackbody Emitted Radiance (AMBER) reference transfer radiometer (Theocharous et al., 1998) and the Physikalisch-Technische Bundesanstalt (PTB) infrared radiometer. In another exercise run concurrently, participants' radiometers were calibrated using the NPL ammonia heat-pipe reference blackbody (Chu and Machin, 1999). Stage 2 took place at Wraysbury reservoir (Spelthorne, TW19 5NX, UK) during the week starting on 27th June 2016 and involved field measurements of the temperature of the surface of the water. Stage 2 included the testing of the same radiometers alongside each other, completing direct daytime and night-time measurements of the surface temperature of the water. Stage 3 took place in the grounds of NPL during the week starting on 4th July 2016 and involved field measurements of the temperature of the surface of a number of solid targets. Stage 3 included the testing of the same radiometers alongside each other, completing direct daytime and night-time measurements of the surface temperature of short grass, clover, soil, sand, gravel and tarmac/asphalt.

This paper provides the results of the comparison of the participants' radiometers while they were viewing the NPL ammonia heat-pipe reference blackbody. All measurements reported by the participants, along with their associated uncertainties, were analysed by the pilot laboratory and are presented in this report.

The findings described in this paper are important because they confirm performance of the radiometers which participated in the comparison. This is a critical requirement because these radiometers are used to validate the surface temperature measurements provided by imaging radiometers mounted on satellites.

Section 2 of this paper summarises the organisation of the radiometer comparison, while Section 3 provides the measurement procedure which was employed during this comparison. Section 4 describes the characteristics of the radiometers which took part in the comparison while Section 5 compares and discusses the findings of the comparison.

2 ORGANISATION OF THE COMPARISON

Recognising the increasing reliance of satellite operators and their customers/users on the quality of the data that comes from the satellite sensors it is essential that measurements used for their validation can be relied upon over a wide range of operational environments. Investments in projects which support the long term delivery of data for decades to come, such as the European Union (EU) Copernicus program, have encouraged the community to subject such measurements to the scrutiny and practises common to other sectors of commerce i.e. comparison and/or audit by independent experts. The international metrology community has a responsibility to support such initiatives and therefore undertake regular comparisons between themselves of key quantities and report the results in open literature to ensure global consistency and transparency to the SI (https://kcdb.bipm.org/). To support this process, they have established procedures and guidance on how to optimally carry out such comparisons and analyse the results. The Earth Observation (EO) community is taking advantage of this knowledge and adopting the guidance to meet its needs. The Quality Assurance Framework for Earth Observation (QA4EO) [http://qa4eo.org/] developed by CEOS is the embodiment of this, and the comparison described below was organised following these metrology-based guidelines and practises.

This meant that before the comparison took place, a formal protocol describing the nature of the comparison, timelines, measurements to be undertaken, reporting format and, in particular, guidance on the content and presentation of an uncertainty budget was developed and agreed by all participants. Such protocols can then be subsequently used, with minor modifications, for similar comparisons in the future and will ensure a degree of consistency in how to interpret results.

During the 2016 comparison, NPL acted as the pilot laboratory and, with the aid of PTB, provided formal traceability to SI units during the laboratory comparisons at NPL. NPL was supported with specialist application advice from University of Southampton, Rutherford and Appleton Lab (RAL) and Karlsruhe Institute of Technology (KIT) during the development of the necessary protocols.

This report provides the results, together with uncertainties as provided by the participants, of the radiometer measurements of the NPL ammonia heat-pipe blackbody operating at seven fixed temperatures as performed in one of NPL's temperature-controlled laboratories during the week beginning 20th June 2016. The laboratory comparison of the participants' blackbodies, as measured by the NPL AMBER radiometer and the PTB infrared radiometer, as well as the Water Surface Temperature (WST) comparison at Wraysbury reservoir and the Land Surface Temperature (LST) comparison that took place in the NPL grounds are being presented elsewhere.

During the 2016 comparison, all participants were encouraged to develop uncertainty budgets for all measurements they reported. In order to achieve optimum comparability, tables containing the principal influence parameters for the measurements were provided to all participants, highlighting the importance of including in their uncertainty budgets uncertainty contributions due to the primary calibration of the radiometer, the linearity of response of the radiometer, drift since the last calibration, effects due to ambient temperature fluctuations, atmospheric absorption/emission, as well as the repeatability and reproducibility of their measurements. All measurements reported by the participants, along with their associated uncertainties, were analysed by the pilot laboratory, blind to all participants, and are presented in this report.

3 MEASUREMENT PROCEDURE FOR THE RADIOMETER LAB COMPARISON

The NPL ammonia heat-pipe reference blackbody (Chu and Machin, 1999) was used in the comparison of the participating radiometers. A schematic of this blackbody is shown in Figure 1. This blackbody uses a heat-pipe to control the blackbody cavity temperature which results in negligible temperature gradients along the length of the cavity. The length of the ammonia heat-pipe blackbody cavity is 300 mm, and it has a 75 mm internal diameter with a 120° cone angle at the end wall. The blackbody cavity is coated with a high-emissivity Nextel black paint. The emissivity of the blackbody cavity has been calculated using the series integral method (Berry, 1981). The effective emissivity of the cavity was estimated to be 0.9993, assuming an emissivity of 0.96 for the Nextel black coating (Betts, *et al.,* 1985).

The temperature of the blackbody cavity was obtained from an ITS-90 calibrated Platinum Resistance Thermometer (PRT) which was inserted into a well of 150 mm depth in the rear of the cavity. The front of the blackbody contained a circular support which allowed aperture plates with different diameters to be positioned in front of the blackbody cavity. The blackbody had a 75 mm diameter aperture mounted on the blackbody casing. There was a total distance of approximately 75 mm from the front of this aperture to the actual blackbody cavity. This, in turn, meant that if radiometers with a large field of view were measuring the reference blackbody, then there was a possibility that they could be seeing parts which were outside of the blackbody cavity, even when they were placed right up against the front of the blackbody casing. While participants were free to position and align their blackbodies at any position in front of the reference blackbody, most of the participants placed their radiometers right up against the reference blackbody, in order to ensure that the blackbody cavity overfilled the entire Field of View of their radiometers.

The temperature of the blackbody cavity was controlled by a cylindrical heat exchanger which fitted closely around the blackbody cavity. Heat transfer fluid was circulated through a continuous 6 mm wide helical groove which was machined in the surface of the internal cylinder. Full information on the ammonia heat-pipe blackbody can be found elsewhere (Chu and Machin, 1999).

At sub-ambient temperatures i.e. at temperatures below the Dew point, the blackbody cavity was purged with dry nitrogen, in order to prevent water from condensing on the internal surfaces of the cavity which could damage the internal black coating and change the effective emissivity. The dry nitrogen gas was fed into the blackbody cavity from the rear. Its temperature was iso-thermalised within the feed tube which was embedded within the wall of the heat pipe. The gas was introduced into the front of the blackbody cavity via a gas distribution ring consisting of 12 holes of 1.5 mm diameter. In order to reduce the effect of convection currents from the surroundings, the aperture of the blackbody cavity was open whilst measurements were being made but was blocked at all other times with an insulation plug.

For each comparison point, the reference blackbody was set at a nominal temperature known only to NPL and enough time was allowed for its cavity temperature to stabilise to the new setting. Once the operating temperature had been selected, the system required just 30 minutes to reach temperatures greater than 0 °C, but as much as 3 hours to reach temperatures on the region of −30 °C. Once the set-point had been reached, the blackbody required another 0.5 to 1 hour to stabilize at the new temperature.

Once the temperature of the reference blackbody was stabilised at a particular temperature, each participant was allowed a maximum period of 30 minutes to position their radiometer, align it to the aperture of the blackbody and take measurements at that particular temperature setting. The order with which radiometers completed the measurements at the beginning of the comparison depended on the readiness of the radiometers of the different participants to do measurements at that particular time. Towards the end of the comparison, participants were allocated 30 minute periods, according to timetables which were circulated to all participants. Participants with more than one radiometer were asked to arrange for the 30 minute measurement period to be shared between all their measuring radiometers. Figure 2 shows the Rosenstiel School of Marine and Atmospheric Sciences (RSMAS) M-AERI radiometer viewing the ammonia heat-pipe blackbody during the comparison.

The temperature of the reference blackbody was continuously logged referenced to Universal Time Coordinated (UTC) and the participants were asked to use the same time reference. This allowed the direct comparison of the measurements of each participant with the corresponding measurements of the reference blackbody.

Participants were asked to provide their measurements in pre-defined spreadsheets. The top of each spreadsheet indicated the date on which the measurements shown in the spreadsheet were performed. Each spreadsheet consisted of a minimum of three columns. The first column indicated the time of the measurement, in a UTC format. The second column gave the brightness temperature of the reference blackbody, as measured by the participant, at the time indicated in the first column. The third column provided the combined standard uncertainty of the measurement of the brightness temperature estimated by the participant corresponding to the measurement indicated in the second column.

Participants were encouraged to develop and provide full uncertainty budgets for their measurements. In order to help participants to do this, tables were provided listing the parameters which were likely to contribute to the uncertainty of the measurement. Some participants provided completed tables, providing extensive information on each uncertainty contribution, while other participants provided considerably less information on their uncertainty budgets, and this is recognised by the community as an area where more work is needed. Full information on the uncertainty budgets provided by participants can be found elsewhere (Barker Snook, et al., 2017).

The measurements were carried out in a lab whose temperature was controlled to ±1 °C around 20 °C and the humidity was controlled to ±5% around 45 % during these measurements.

4 PARTICIPANTS' RADIOMETERS AND MEASUREMENTS

A total of 19 radiometers operating on 24 different measurement channels took part in the 2016 radiometer lab comparison. This section gives brief descriptions of the participating radiometers. A summary of the most important parameters of the participating radiometers is given in Table 1.

The University of Valencia CIMEL Electronique CE312-2 radiometers

Two radiometers were provided by the Dept. of Earth Physics and Thermodynamics of the University of Valencia, Spain. Both radiometers were of the CIMEL Electronique CE312-2 type and operated in six spectral bands, 8.0 µm to 13.3 µm, 10.9 µm to 11.7 µm, 10.2 µm to 11.0 µm, 9.0 µm to 9.3 µm, 8.5 µm to 8.9 µm, and 8.3µm to 8.6 µm. These radiometers were able to provide measurements of the brightness temperature of the reference blackbody for each of the six bands on which they were able to operate. Both radiometers employed germanium windows and used narrow band filters with zinc sulphide substrates to select the different wavelength bands. Both instruments had a 10 degree full angle Field of View and included a built-in radiance reference made of a concealable gold-coated mirror which enabled comparison between the target radiance and the reference radiation from inside the detector cavity. The temperature of the detector was measured with a calibrated PRT, thus allowing compensation for the cavity radiation. The relevant outputs of the radiometer were the detector temperature and the difference in digital counts between the signals from the target and the detector cavity. The quoted uncertainty of measurements made by the first radiometer (Unit 1) was 370 mK, while the corresponding value for the second radiometer (Unit 2) was 360 mK (Barker-Snook et al., 2017). Further information on these radiometers can be found in Sicard et al., (1999) and in Legrand et al., (2000).

The KIT Heitronics KT15.85 IIP radiometer

The radiometer provided by the Institute of Meteorology and Climate Research - Atmospheric Trace Gases and Remote Sensing (IMK-ASF), Karlsruhe Institute of Technology (KIT), Germany was a Heitronics KT15.85 IIP radiometer with L6 lens (https://www.heitronics.com/fileadmin/content/Prospekte/KT15IIP_e_V510.pdf). This was a single channel radiometer based on a pyroelectric infrared detector. This type of sensor links radiance measurements via beam-chopping to internal reference temperature measurements and thermal drift can practically be eliminated. The field of view of this radiometer was 8.3° (full angle). The KT15.85 IIP responded in the 9.6 µm to 11.5 µm spectral range, had a quoted uncertainty of approximately 0.3 K (Barker Snook, et al., 2017, page 29) over the temperature range relevant to land surfaces and claimed good long-term stability.

The ONERA radiometers

Four radiometers were provided by the Office National d'Etudes et de Recherches Aérospatiales (ONERA), France. The first three radiometers were Heitronics KT19.85 II (https://www.heitronics.com/fileadmin/content/Prospekte/KT15IIP_e_V510.pdf) which had a 95 mm target diameter when viewing a target at a distance of 2 m. These radiometers operated in the 9.6 µm to 11.5 µm spectral band and offered a 60 mK temperature resolution. Their quoted measurement uncertainty was ±0.5 °C + 0.7 % of the difference between target and housing temperature. The fourth ONERA radiometer was a BOMEM MR304SC Spectroradiometer (https://library.e.abb.com/public/654dfb800019d7168525712d00693379/4314%20MR304SC%20Spec.pdf) covering the 3 µm to 13 µm wavelength range with two detectors, one InSb and one MCT detector, with a 4 cm^{-1} resolution. This radiometer had a 20° (full angle) FoV. The measured radiance spectrum was converted into brightness temperature and averaged over the 9.6 µm to 11.5 µm wavelength range of the Heitronics radiometers. The temperature uncertainty was quoted for each measurement and ranged from 0.2 K to 0.4 K, depending on the blackbody set temperature (Barker-Snook et al., 2017).

The CSIRO Infrared Sea surface temperature Autonomous Radiometer (ISAR)

The radiometer provided by the Marine National Facility, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia was an ISAR 5D radiometer. This radiometer had a Field of View of 7 degrees (full angle) and responded to wavelengths in the 9.6 µm to 11.5 µm spectral band. This radiometer offered a 10 mK temperature resolution. The measurement uncertainty was quoted for each blackbody temperature measured and ranged from 85 mK at -30 ºC to 57 mK at 45 ºC. Full information on this type of radiometer is given by Donlon et al., (2008) and Wimmer and Robinson, (2016).

The STFC RAL SISTeR radiometer

The radiometer provided by the Science and Technology Facilities Council Rutherford Appleton Laboratory, UK, was the Scanning Infrared Sea Surface Temperature Radiometer (SISTeR). SISTeR was a chopped, self-calibrating filter radiometer manufactured by RAL Space. It had a single-element Deuterated Lanthanum Alanine-doped TriGlycine Sulphate (DLATGS) pyroelectric detector, a filter wheel containing up to six band-defining filters and two internal reference blackbodies, one operating at ambient temperature and the other heated to approximately 17 K above ambient. During operation, the radiometer selected (with a scan mirror) successive views of each of the blackbodies and the external scene in a repeated sequence. For Sea Surface Temperature (SST) measurements, the external measurements included views of the sea surface and the sky at the complementary angle. The instrument field of view was approximately 13° (full angle). During this comparison, a filter centred at 10.8 µm was used. The measurement uncertainty was quoted for each blackbody temperature measured and ranged from 128 mK at -30 ºC to 19 mK at 45 ºC. Further information on the SISTeR radiometer can be found in http://www.stfc.ac.uk/research/environment/sister/

The Southampton University ISAR radiometer

The radiometer provided by the National Oceanography Centre of Southampton University, UK, was an ISAR-5D with a Field of View of 7 degrees (full angle). The radiometer responded to wavelengths in the 9.6 µm to 11.5 µm spectral band. This radiometer offered a 10 mK temperature resolution. The measurement uncertainty was quoted for each blackbody temperature measured and ranged from 120 mK at -30 °C to 60 mK at 30 °C. Full information on the ISAR radiometer can be found in the papers by Donlon et al., (2008) and Wimmer and Robinson, (2016).

The DMI radiometers

Two radiometers were provided by the Danish Meteorological Institute (DMI), Denmark. The first radiometer was an ISAR-5D. The measurement uncertainty of this radiometer was quoted for each blackbody temperature measured and ranged from 83 mK at -15 °C to 59 mK at 45 °C. Full information on this radiometer can be found in the paper by Donlon et al., (2008) and in the sections dealing with the CSIRO and Southampton University radiometers. The second radiometer was a Campbell Scientific IR120. This was a broadband radiometer measuring over the 8 µm to 14 µm wavelength range. This radiometer offered a 10 mK temperature resolution and a quoted measurement uncertainty of 200 mK. For further information on the Campbell Scientific IR120 radiometer see:

https://s.campbellsci.com/documents/eu/manuals/ir100_ir120.pdf

The OUC, Qingdao radiometers

Two radiometers were provided by the Ocean University of China (OUC), Qingdao, China. The first radiometer was an ISAR 5C radiometer. This radiometer had a Field of View of 7 degrees (full angle) and responded to wavelengths in the 9.6 µm to 11.5 µm spectral band. This radiometer offered a 10 mK temperature resolution and a quoted measurement uncertainty of 100 mK for all blackbody temperatures measured. Full information on this radiometer can be found in Donlon et al., (2008) and in Wimmer and Robinson, (2016).

The second radiometer provided by the OUC was an Ocean University of China First Infrared Radiometer (OUCFIRST) developed for measurements of the sea surface temperature. The OUCFIRST radiometer was similar to the ISAR radiometer and was based on the Heitronics KT15.85 IIP detector which responds in the 9.6 µm to 11.5µm wavelength range. The OUCFIRST radiometer also included two internal reference blackbody sources. This radiometer was calibrated before and after each measurement campaign using an external blackbody. The quoted measurement uncertainty of this radiometer was 100 mK for all blackbody temperatures measured.

The GOTA CIMEL Electronique CE312-2 radiometer

The radiometer provided by Grupo de Observacion de la Tierra y la Atmosfera (GOTA) Universidad de La Laguna, Spain was a CIMEL Electronique CE312-2 radiometer. This radiometer incorporated a thermopile detector and was able to operate over six wavelength bands spread over the 8 µm to 13 µm wavelength range. The measurement uncertainty of this radiometer was quoted for each blackbody temperature measured and ranged from 400 mK to 500 mK for all measurements completed by this radiometer. Further information on this radiometer can be found in Sicard et al., (1999) and Legrand et al., (2000) as well as in the section dealing with the University of Valencia radiometers.

The RSMAS M-AERI radiometer

The radiometer provided by the Rosenstiel School of Marine and Atmospheric Science (RSMAS), University of Miami, USA was a Marine-Atmospheric Emitted Radiance Interferometer (M-AERI) Mk-3. This radiometer, like its predecessors, was based on a Fourier-Transform Infrared Spectro-radiometer, which uses a Michelson-Morley interferometer design, with the path differences generated by an oscillating yoke with a corner-cube reflector on each arm. Wavelength calibration is accomplished using a He-Ne laser. Radiometric calibration is achieved by using two blackbodies whose cavities are maintained at known temperatures at each of which the field of view of the interferometer is directed sequentially before and after scene measurements. It had a 25 mm diameter entrance aperture and a spectral resolution of 0.5 cm^{-1}. Its temperature resolution was quoted as 5 mK. The M-AERI Mk-3 had a field of view of 2.58 degrees (full angle) and responded over the 3300 cm^{-1} to 525 cm^{-1} range (3 μm to 19 μm wavelength range). The brightness temperature of the ammonia heat-pipe blackbody was provided at two wavenumbers, 1000 cm^{-1} (10 μm) and 1302 cm^{-1} (7.68 μm) with quoted combined uncertainties of 18 mK and 40 mK respectively. Full information on this radiometer can be found in Minnett et al., (2001).

5 RESULTS AND DISCUSSION

Figure 3 plots, as an example, the measurements provided by the STFC RAL SISTeR radiometer (orange circles) when viewing the NPL blackbody maintained at about 10 °C and the corresponding measurements of the cavity temperature made by the NPL (blue dashes). Similar plots corresponding to all participating radiometers and for all ammonia heat-pipe blackbody temperatures for which measurements were made can be found elsewhere (Barker Snook, et al., 2017). Also plotted in the same figure are the combined uncertainty values of the measurements made by SISTeR (orange error bars) and those of the NPL blackbody measurements (blue error bars). From the measurements shown in Figure 3, the difference between the average of the measurements made by the SISTeR radiometer over this time period and the average of the corresponding NPL measurements of the blackbody temperature was estimated to be 60 mK.

Figures 4 to 10 show the plots of the mean of the differences between the radiometer readings and the corresponding NPL measurements of the temperature of the ammonia heat-pipe reference blackbody, for all the blackbody temperatures at which the radiometers were compared. The uncertainty bars shown in these Figures represent the combined standard uncertainty ($k = 1$) of the measurements provided by the participants and includes the uncertainty contribution due to the ammonia heat-pipe blackbody.

It is clear that the uncertainty of measurements reported by radiometers which included internal blackbodies for continuous calibration of their responsivity is significantly lower that the corresponding uncertainty of radiometers which did not include internal references. This was to be expected because the responsivity of infrared detectors is known to drift due to a number of reasons (see for example, Theocharous and Theocharous, 2006). The use of internal references such as the blackbodies included within the radiometers allowed the effects of these drifts to be arrested, thus reducing the combined uncertainty of their measurements. This is also reflected in the difference between the measurements made by these radiometers and the temperature of the ammonia heat pipe blackbody. Measurements made by radiometers with internal blackbodies generally provided better agreement compared to measurements reported by radiometers which did not have internal references.

Examination of Figures 4 to 10 indicate that some cases exists in which the measurements reported by the ammonia heat-pipe blackbody were well outside the uncertainty bars of measurements reported by participating radiometers, even radiometers which included internal reference blackbodies. A major part of this discrepancy can be explained on the basis that the uncertainty bars shown in Figures 4 to 10 represent the one-sigma (k=1) uncertainty values. If the uncertainty bars were extended to represent the three-sigma (k=3) case, then the uncertainty bars of all measurements reported by radiometers which included internal reference blackbodies would have included the corresponding measurements reported by the ammonia heat-pipe blackbody.

Figures 4 to 10 show that the differences between the participants' radiometer readings and the corresponding temperature of the NPL reference blackbody became progressively larger, particularly as the reference blackbody temperature decreased to -15 °C and -30 °C. This observation is not altogether surprising because measurements were made in a lab, with the measuring radiometers operating at ambient temperatures. This means that the internal blackbodies within the participating radiometers (which provided the reference against which the radiometers were basing their measurements) were also operating at near ambient temperatures; hence for low temperatures of the ammonia heat-pipe blackbody, the difference between the temperature of the test blackbody and the internal reference blackbodies increased, probably leading to the observed discrepancies. The discrepancies are likely to arise due to the large extrapolation ranges (up to 50 °C) and may be enhanced by other effects. If, for example, the out-of-band response of a radiometer was measured incorrectly or had a small undetected spectral leak, then discrepancies are likely to arise. It is estimated that the output of a radiometer responding in the 10 µm to 11 µm region, which is calibrated at 30 ºC and extrapolated to -30 ºC, will be 0.26 % different from the output obtained if the radiometer had an out-of-band response in the 5 µm to 6 µm region which was just 1 % of the response in the 10 µm to 11 µm band.

It is important to point out that if the radiometers were used to measure low temperature targets, such as the surface temperature of ice in the arctic, then the radiometers (as well as the internal blackbodies) will also be at low temperatures so the extrapolation will not be over a significant temperature range. This means that the discrepancies between the radiometer measurement of the ice and the true surface temperature of ice are likely to be much smaller. For future comparisons where such low temperatures are important, consideration should be given to how the ambient temperature of the radiometers can be reduced to be more representative of the operational environment.

Moreover, as the temperature of the reference blackbody decreases, the signal detected by the photodetectors within the radiometers also decreases, resulting in poorer signal-to-noise ratios. The poorer signal-to-noise ratios would result in measurements with poorer Type A uncertainty and thus more unreliable measurements due to the resulting higher combined uncertainty.

It is important to note that the NPL AMBER radiometer was used in the past to measure the temperature of the same ammonia heat-pipe reference blackbody used in this comparison and the agreement between the NPL AMBER measurements and the blackbody measurements was good, indicating its reliability. In fact the difference between the NPL AMBER measurements and the reference blackbody measurements are included in the Figures for blackbody temperatures of -30 ºC, 0 ºC, 10 ºC, 20 ºC and 30 ºC. The agreement between the AMBER and the reference blackbody measurements indicates that the discrepancies observed in the measurements of some radiometers (which can be as large as 2 K for blackbody temperatures around -30 °C) do not arise due to issues with the blackbody but are likely to be associated with the participants' measurements. Furthermore, NPL AMBER radiometer was used to measure the temperature of the ammonia heat-pipe blackbody of PTB, the German national standards lab, and that comparison also showed good agreement between the measurements provided by NPL AMBER and those provided by the PTB reference blackbody, as shown in Figure 4 in the paper by Gutschwager (Gutschwager et al., 2013) which deals with that comparison.

The NPL reference blackbody had an aperture of 75 mm in diameter which could be decreased by adding apertures with diameters smaller than 75 mm on the blackbody casing (see Figure 1). The distance between the front of the blackbody cavity and the aperture formed/mounted on the blackbody casing was also 75 mm, meaning that the Field of View (FoV) of a radiometer placed against the casing would be overfilled by the blackbody cavity, provided its half angle was less than 26.5° (53° full angle). Although the 75 mm diameter of the blackbody and its position were defined and open for review in the protocol before the measurements took place, this could be a source of error for radiometers with a large angle field of view (e.g. the ONERA MR354SC, IR120, SISTeR and CE312-2 radiometers), as well as radiometers which could not be positioned close to the blackbody casing aperture (e.g. M-AERI Mk-3). For these radiometers, the measurements taken would likely capture the edges of the blackbody cavity, as well as radiation emitted by blackbody cavity, thus introducing biases to the measurements. To avoid this problem, some participants made their measurements with their radiometers as close to the blackbody front aperture as possible**.** Although this was considered to be a satisfactory compromise, care should be taken because the emissivity of the NPL reference blackbody is not unity. This meant that when a test radiometer was brought very close to the blackbody aperture it partly "saw itself" reflected by the blackbody because the blackbody is no longer exposed to ambient temperature but the temperature of the radiometer. This is a particular problem with radiometers which operate at cryogenic temperatures.

For the temperatures below 0 °C, ice began to form near the aperture of the reference blackbody cavity. While the ice only formed near the entrance to the cavity (the cavity was continuously purged with dry nitrogen gas), the presence of the ice may have affected the effective emissivity of the areas of the blackbody cavity on which ice was deposited and thus alter the effective emissivity of the reference blackbody for radiometers with very large FoVs. This may also have impacted some of the results associated with the measurement of the temperature of blackbody cavity. However, the same measurements were made using the NPL AMBER radiometer and no discrepancies were observed for blackbody temperatures as low as -45 ºC, indicating that no ice was formed inside the reference blackbody cavity.

For the majority of radiometers being compared, their intended use was for sea surface temperature measurements. For this reason, the majority of the participants used blackbodies to calibrate their radiometers which could not operate below 0 ºC, while some participants used blackbodies which could not operate below ambient temperature. This meant that the temperature range over which the majority of radiometers were calibrated was for temperatures above 0 °C and in some cases temperatures above ambient. This means that some measurements taken during this laboratory comparison were outside the range of calibrated temperatures for these instruments so measurements made at the lower temperatures relied on the extrapolation of the calibrations at higher temperatures. Any consideration of irregularities with the values for measurements and their associated uncertainties made below 0 °C should take this into account.

During the 2016 radiometer comparison, a 30 minute period was allocated to each participant to allow for the alignment of the radiometer to the reference blackbody aperture and the making of the measurements at a particular blackbody temperature. Some participants reported that 30 minutes was not enough. However, because of the number of radiometers participating in the 2016 comparison and the number of temperatures which had to be completed over the week-long comparison, the 30 minute period could not be extended. It is recommended that in future comparisons, participants should be asked to state how long they would ideally require in order to align and complete a measurement (at a particular blackbody temperature). If the total duration of the comparison could not be extended, or the number of participating radiometers could not be reduced, then the number of reference blackbody temperatures at which measurements are done should be reduced to allow participants the extra time periods they require to complete their measurements.

6 CONCLUSIONS

The performance of a number of radiometers was compared in the lab by measuring the brightness temperature of the NPL ammonia heat-pipe blackbody at a number of temperatures in the -30 ^{o}C to +45 ^{o}C temperature range. The results show that measurements of the reference blackbody for cavity temperatures above 0 ^{o}C reported by radiometers which include internal blackbodies exhibit superior measurement uncertainties and provide better agreement with measurements reported by the ammonia heat-pipe blackbody compared to radiometers which rely on infrequent re-calibration using external blackbodies. Furthermore, although the Figures indicate that some cases exists in which the measurements reported by the ammonia heat-pipe blackbody were well outside the uncertainty bars of measurements reported by participating radiometers (even radiometers which included internal reference blackbodies), this can be explained on the basis that the uncertainty bars shown in the Figures represent the one-sigma (k=1) uncertainty values. If the uncertainty bars were extended to represent the three-sigma (k=3) case, then the uncertainty bars of all measurements reported by radiometers which included internal reference blackbodies would have included the corresponding measurements reported by the ammonia heat-pipe blackbody.

Participants were encouraged to provide detail uncertainty budgets for all measurements they provided. Although uncertainty estimates were provided by all participants for all measurements they reported as part of the 2016 comparison, the level of detail which was included in the uncertainty budgets varied significantly from one participant to the next, with some participants providing only a value for the estimate of the uncertainty of their measurements. It is recommended that participation in future comparisons should be made conditional on participants providing full uncertainty budgets for all measurements they provide as part of the comparison activity.

The 2016 comparison showed that the differences between the readings of the participating radiometer and the corresponding temperature of the reference blackbody increased, particularly for measurements corresponding to reference blackbody temperatures below 0 °C. Reasons for the discrepancies observed at low blackbody temperatures were put forward, including the extrapolation from the calibration of the radiometers using blackbodies operating at ambient temperatures, combined with the absence of any information on the relative spectral responsivity of the radiometers. These discrepancies are not expected to arise, if the radiometers were calibrated with a reference blackbody operating at these low temperatures. Furthermore, any discrepancies which were measured at low blackbody temperatures may be considered irrelevant because the majority of the radiometers taking part in this comparison will be used to measure sea surface temperature i.e. temperatures above 0 °C.

7 ACKNOWLEDGEMENTS

The authors wish to thank Miss Helen McEvoy and Mr Jamie McMillan (both of NPL) for their assistance in operating the ammonia heat-pipe blackbody and with the testing surrounding the radiometer laboratory comparison. The authors also wish to thank ESA for funding this work under ESA Contract No: 4000113848_15I-LG. The University of Valencia participation in the 2016 FRM4STS comparison was funded by the Spanish Ministerio de Economía y Competitividad and the European Regional Development Fund (FEDER) through the project CGL2015-64268-R (MINECO/FEDER, UE) and by the Ministerio de Economía y Competitividad through the project CGL2013-46862-C2-1-P (MINECO). The Ocean University of China participation in the 2016 FRM4STS comparison was funded by the National Natural Science Foundation of China under Grant 41376105 and the Global Change Research Program of China under Grant 2015CB953901. The Universidad de La Laguna participation was supported by Spanish Ministerio de Economía y Competitividad under project CGL2013-48202-C2-1-R.

8 REFERENCES

Barker Snook, I., E. Theocharous and N. P. Fox, 2017: 2016 comparison of IR brightness temperature measurements in support of satellite validation. Part 2: Laboratory comparison of radiation thermometers. *NPL Report ENV 14*, 118 pp. see link http://www.frm4sts.org/wp-content/uploads/sites/3/2017/12/FRM4STS_D100_TR-2_Part2_Radiometer_23Jun17-signed.pdf

Barton, I. J., P. J. Minnett, K. A. Maillet, C. J. Donlon, S. J. Hook, A. T. Jessup and T. J. Nightingale, 2004: The Miami 2001 infrared radiometer calibration and intercomparison: Part II Shipboard results. *J. Atmos. Oceanic Technol.*, **21**, 268-283.

Berry, K., 1981: Emissivity of a cylindrical black-body cavity with a re-entrant cone end face. *J. Phys. E: Sci. Instrum.* **4,** 629–632.

Betts D. B., F. J. J. Clarke, L. J. Cox and J. A. Larkin, 1985: Infrared reflection properties of five types of black coating for radiometric detectors. *J. Phys. E: Sci. Instrum.,* **18,** 689–696.

Chu, B and G. Machin, 1999: A low-temperature blackbody reference source to -40 ^{o}C. *Meas. Science Technol.*, **10**, 1–6.

Donlon, C., I. S. Robinson, W. Wimmer, G. Fisher, M. Reynolds, R. Edwards and T. J. Nightingale, 2008: An infrared sea surface temperature autonomous radiometer (ISAR) for deployment aboard volunteer observing ships (VOS). *J. Atmos. Oceanic Technol.*, **25**, 93-113.

Gutschwager, B., E. Theocharous, C. Monte, A. Adibekyan, M. Reiniger, N. P. Fox and J. Hollandt, 2013: Comparison of the radiation temperature scales of the PTB and the NPL in the temperature range from -57 °C to 50 °C. *Meas. Science Technol.*, **24**, Article No 095002.

Legrand, M., C. Pietras, G. Brogniez, M. Haeffelin, N. K. Abuhassan. And M. Sicard, 2000: A high-accuracy multiwavelength radiometer for in situ measurements in the thermal infrared. Part I: characterization of the instrument, J. *Atmos. Ocean Technol.,* **17**, 1203-1214.

Minnett, P. J., R. O. Knuteson, F. A. Best, B. J. Osborne, J. A. Hanafin and O. B Brown, 2001: The Marine-Atmosphere Emitted Radiance Interferometer (M-AERI), a high accuracy, sea-going infrared spectroradiometer. *J. Atmos. Oceanic. Technol.*, **18**, 994-1013.

Minnett, P. J., C. Gentemann, T. J. Nightingale, I. J. Barton, B Ward, B and M. J. Murray, 2002: Toward improved validation of satellite sea surface skin temperature measurements for climate research. *J. Climate,* **15**, 353-369.

Minnett, P.J. and G. K. Corlett, 2012: A pathway to generating Climate Data Records of sea-surface temperature from satellite measurements. *Deep Sea Research Part II: Topical Studies in Oceanography*, **77–80**, 44-51

Rice, J. P., J. I. Butler, B. C. Johnson, P. J. Minnett, K. A. Maillet, T. J. Nightingale, S. J. Hook, A. Abtahi, C. J. Donlon and I. J. Barton, 2004: The Miami 2001 infrared radiometer calibration and intercomparison. Part I: Laboratory characterisation of blackbody targets. *J. Atmos. Oceanic Technol.*, **21**, 258-267.

Sicard, M., P. R. Spyak, G. Brogniez, M. Legrand, N. K. Abuhassan, C, Pietras and J. P. Buis, 1999: Thermal infrared field radiometer for vicarious cross-calibration: characterization and comparisons with other field instruments. *Opt. Engin.*, **38**, 345-356.

Theocharous, E., N. P. Fox, V. I. Sapritsky, S. N. Mekhontsev and S. P. Morozova, 1998: Absolute measurements of black-body emitted radiance *Metrologia*, **35**, 549-554.

Theocharous, E. and Theocharous, O. J., 2006: Practical limit of the accuracy of radiometric measurements using HgCdTe detectors *Applied Optics*, **45**, 7753-7759Theocharous, E., E. Usadi and N. P Fox, 2010: CEOS comparison of IR brightness temperature measurements in support of satellite validation. Part I: Laboratory and ocean surface temperature comparison of radiation thermometers", *NPL Report COM OP3*, 130 pp.

Theocharous, E. and N. P. Fox, 2010: CEOS comparison of IR brightness temperature measurements in support of satellite validation. Part II: Laboratory comparison of the brightness temperature of blackbodies. *NPL Report COM OP4,* 43 pp.

Wimmer, W., and I. Robinson, 2016: The ISAR instrument uncertainty model. *J. Atmos. Oceanic Technol*., **33,** 2415-2433.

Yoder, J. A., K. S. Casey and M. D. Dowell, 2014: Ocean Climate and satellite optical radiometry, *Optical radiometry for Ocean climate measurements*, edited by G. Zibori, G. J. Donlon and A.C. Parr, Published by Academic press, 3-12.

Table 1. Instruments involved in the 2016 CEOS infrared radiometers laboratory comparison.

Radiometer	Institute	Waveband (μm)	Detector	Field of View (°)	Reference
CE312-2 (2 units)	UV[a]	8.0 – 13.3	Thermopile	10	Sicard et al. (1999)
KT15.85 IIP	KIT[b]	9.6 – 11.5	Pyroelectric	8.3	https://www.heitronics.com/fileadmin/content/Prospekte/KT15IIP_e_V510.pdf
KT19.85 II (3 units)	ONERA[c]	9.6 – 11.5	Pyroelectric	1.36	https://www.heitronics.com/fileadmin/content/Prospekte/KT15IIP_e_V510.pdf
MR354SC	ONERA[c]	3.0 – 13.0	MCT Photoconductive	20	https://library.e.abb.com/public/654dfb800019d7168525712d00693379/4314%20MR304SC%20Spec.pdf
ISAR-5D	CSIRO[d]	9.6 – 11.5	Pyroelectric	7	Wimmer & Robinson, (2
SISTeR	STFC/RAL[e]	10.8	Pyroelectric	13	Barton et al., (2004)
ISAR-5D	NOC[f]	9.6 – 11.5	Pyroelectric	7	Wimmer and Robinson, (2016)
ISAR-5D	DMI[g]	9.6 – 11.5	Pyroelectric	7	Wimmer and Robinson, (2016)
IR120	DMI[g]	8.0 – 14.0	Thermopile	20	www.campbellsci.eu/ir120
ISAR-5C	OUC[h]	9.6 – 11.5	Pyroelectric	7	Wimmer and Robinson, (2016)
OUCFIRST	OUC[h]	9.6 – 11.5	Pyroelectric	7	
CE312-2	GOTA/ULL[i]	8.1 – 11.7	Thermopile	10	Sicard et al. (1999)
M-AERI	RSMAS[j]	3.0 – 18.0	Cooled InSb and HgCdTe	2.58	Minnett et al., (2001)

[a] University of Valencia (Spain)
[b] Karlsruhe Institute of Technology (Germany)
[c] Office National d'études et de Recherches Aérospatiales (France)
[d] Commonwealth Scientific and Industrial Research Organisation (Australia)
[e] Science and Technology Facilities Council Rutherford Appleton Laboratory (UK)
[f] National Oceanography Centre (UK)
[g] Danish Meteorological Institute (Denmark)
[h] Ocean University of China (China)
[i] Grupo de Observación de la Tierra y la Atmósfera/Universidad de La Laguna (Spain)
[j] Rosenstiel School of Marine and Atmospheric Sciences (USA)

Figure captions:

Figure 1: Schematic of the ammonia heat-pipe blackbody.

Figure 2: The RASMAS M-AERI radiometer viewing the ammonia heat-pipe blackbody during the 2016 radiometer comparison.

Figure 3: Measurements of the STFC RAL SISTeR radiometer viewing the NPL reference blackbody maintained at approximately 10 °C (in orange) and the corresponding measurements made by NPL of the blackbody temperature (in blue).

Figure 4: Plot of the mean of the differences of the radiometer readings from the temperature of the NPL reference blackbody, maintained at a nominal temperature of -30 °C.

Figure 5: Plot of the mean of the differences of the radiometer readings from the temperature of the NPL reference blackbody, maintained at a nominal temperature of -15 °C.

Figure 6: Plot of the mean of the differences of the radiometer readings from the temperature of the NPL reference blackbody, maintained at a nominal temperature of 0 °C.

Figure 7: Plot of the mean of the differences of the radiometer readings from the temperature of the NPL reference blackbody, maintained at a nominal temperature of 10 °C.

Figure 8: Plot of the mean of the differences of the radiometer readings from the temperature of the NPL reference blackbody, maintained at a nominal temperature of 20 °C.

Figure 9: Plot of the mean of the differences of the radiometer readings from the temperature of the NPL reference blackbody, maintained at a nominal temperature of 30 °C.

Figure 10: Plot of the mean of the differences of the radiometer readings from the temperature of the NPL reference blackbody, maintained at a nominal temperature of 45 °C.

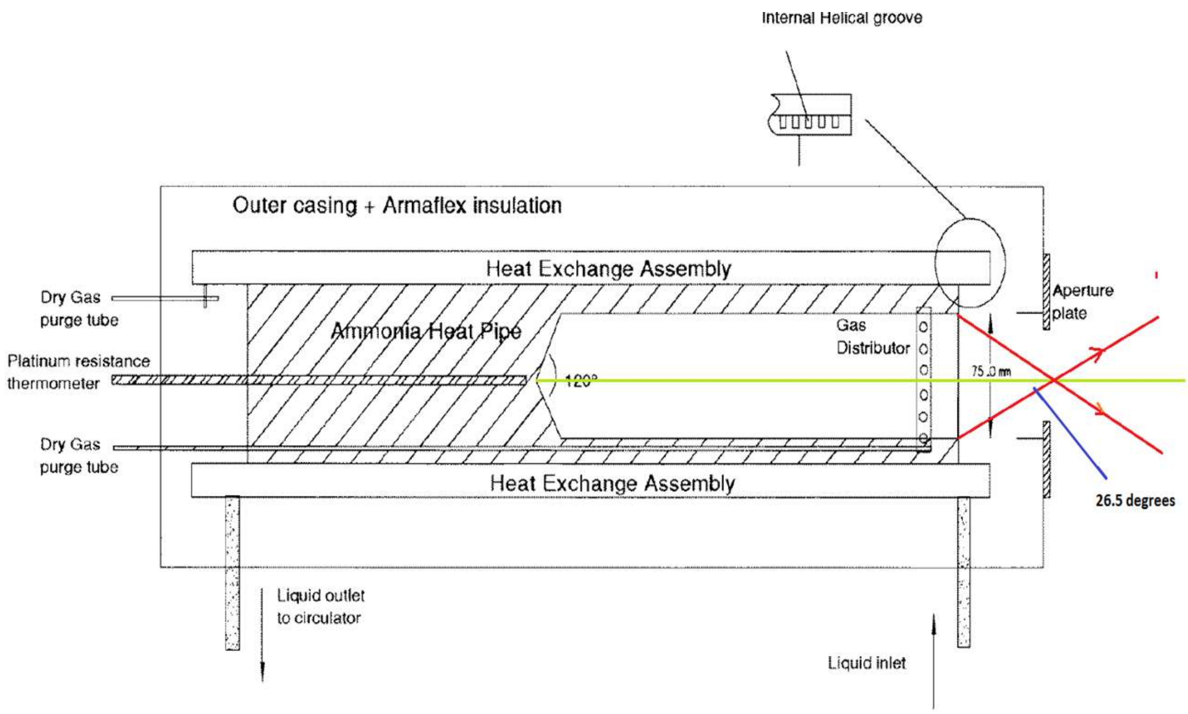

Figure 1: Schematic of the ammonia heat-pipe blackbody

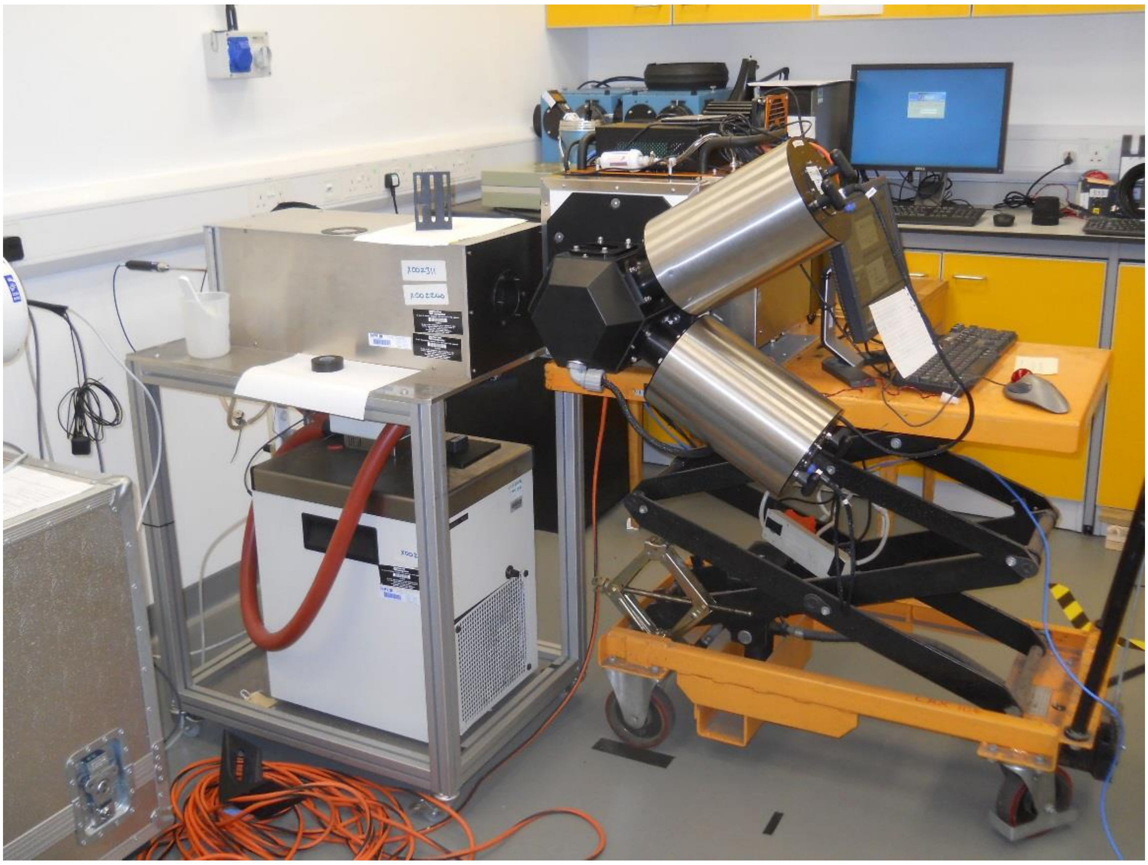

Figure 2: The RSMAS M-AERI radiometer viewing the ammonia heat-pipe blackbody during the 2016 radiometer comparison.

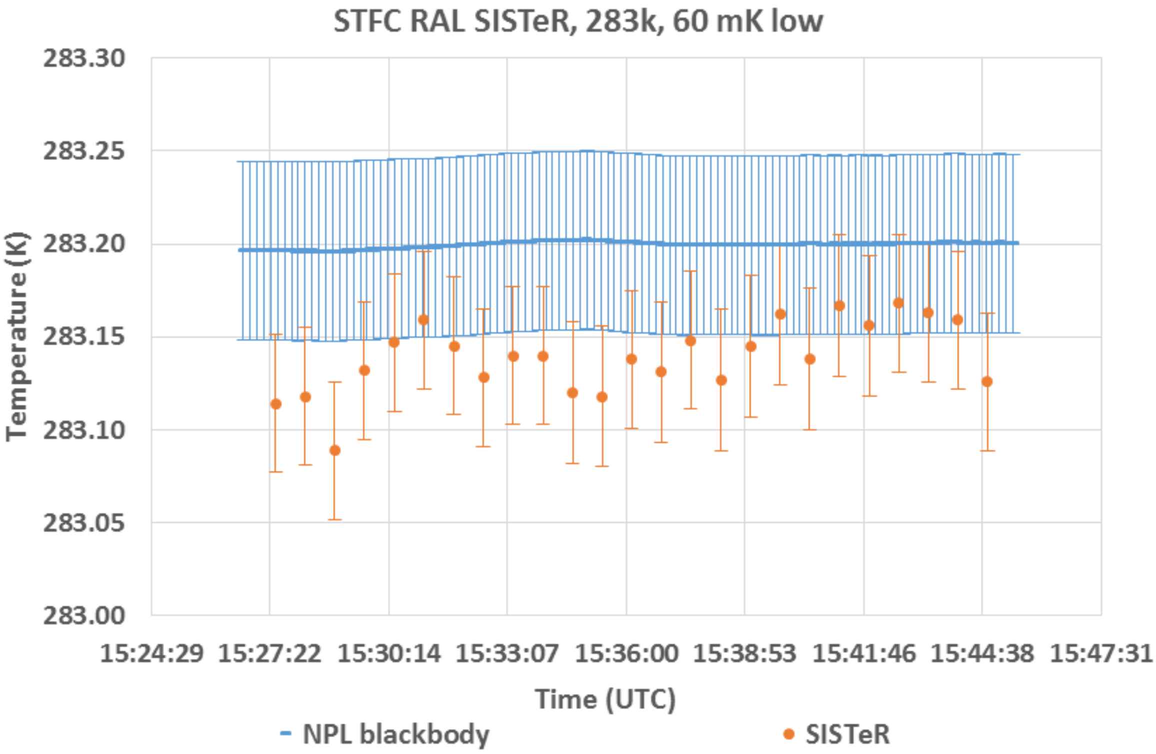

Figure 3: Measurements of the STFC RAL SISTeR radiometer viewing the NPL reference blackbody maintained at approximately 10 °C (in orange) and the corresponding measurements made by NPL of the blackbody temperature (in blue).

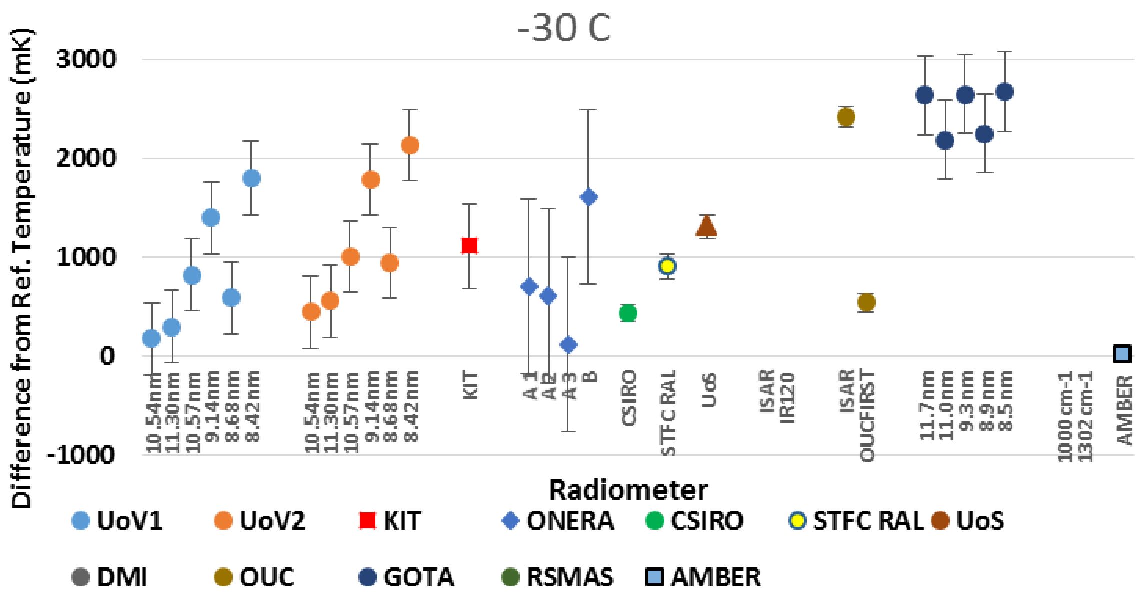

Figure 4: Plot of the mean of the differences of the radiometer readings from the temperature of the NPL reference blackbody, maintained at a nominal temperature of -30 °C.

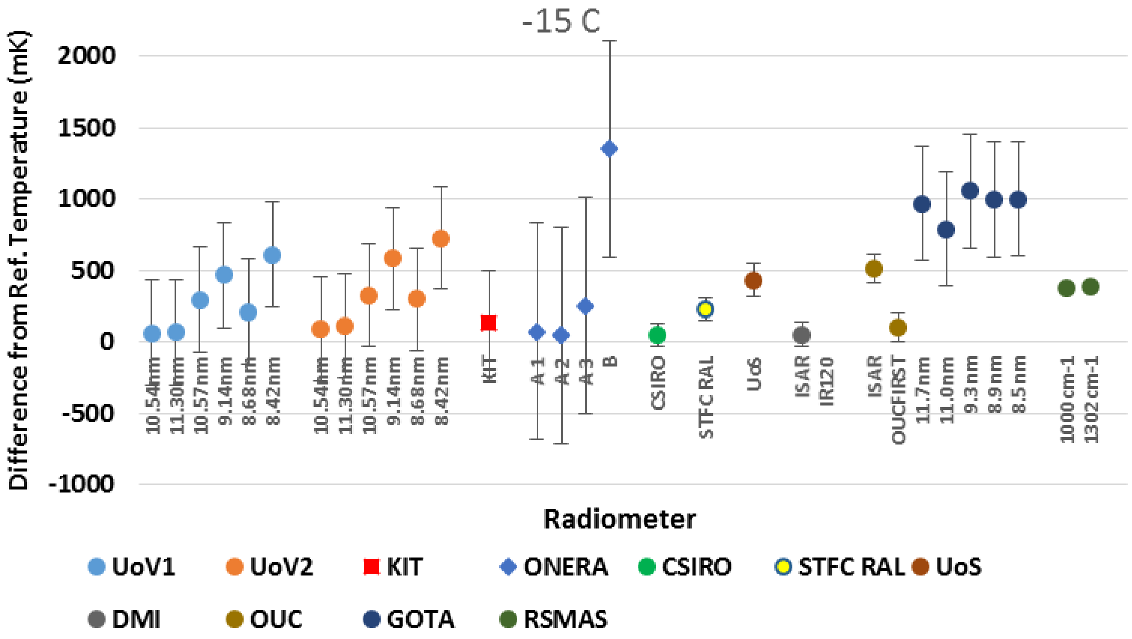

Figure 5: Plot of the mean of the differences of the radiometer readings from the temperature of the NPL reference blackbody, maintained at a nominal temperature of -15 °C.

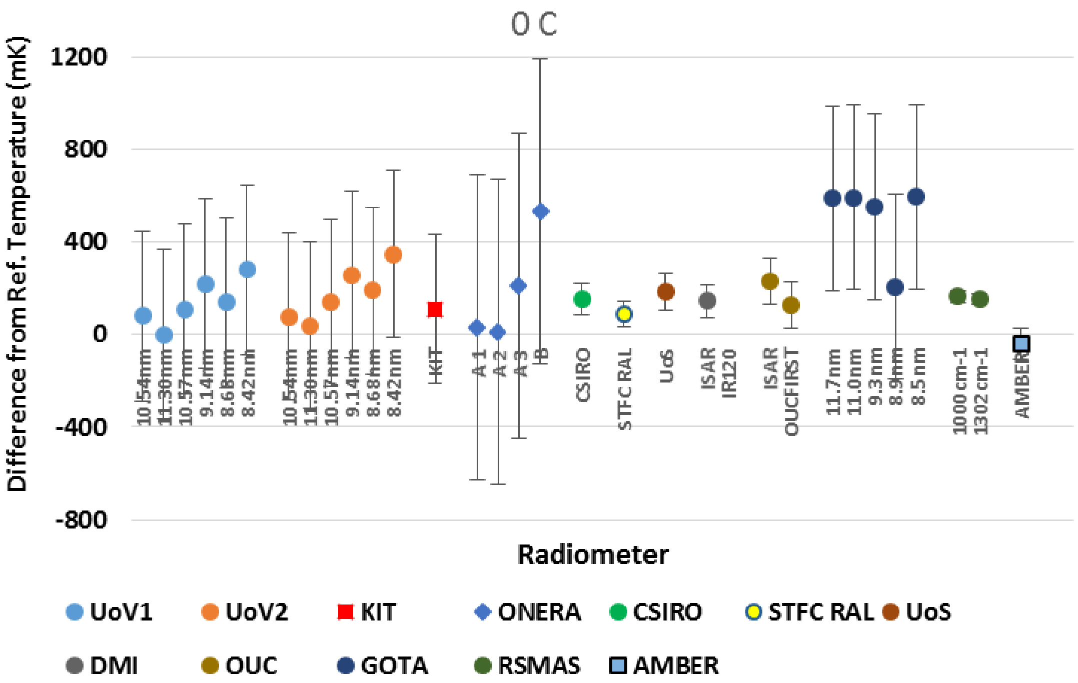

Figure 6: Plot of the mean of the differences of the radiometer readings from the temperature of the NPL reference blackbody, maintained at a nominal temperature of 0 °C.

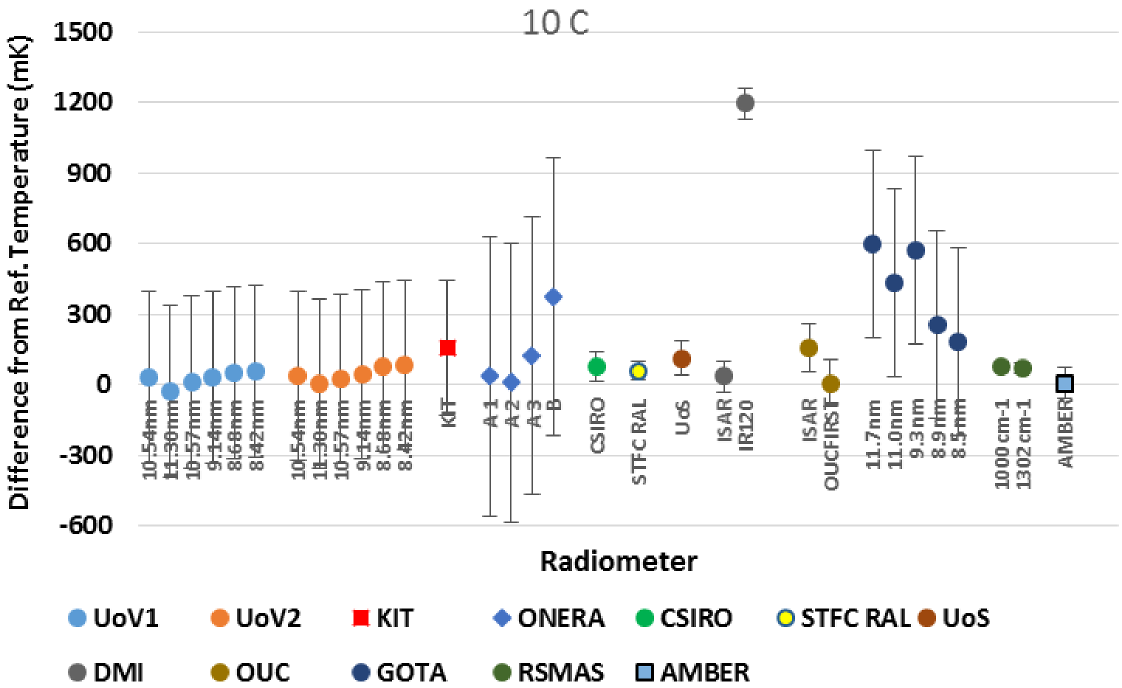

Figure 7: Plot of the mean of the differences of the radiometer readings from the temperature of the NPL reference blackbody, maintained at a nominal temperature of 10 °C.

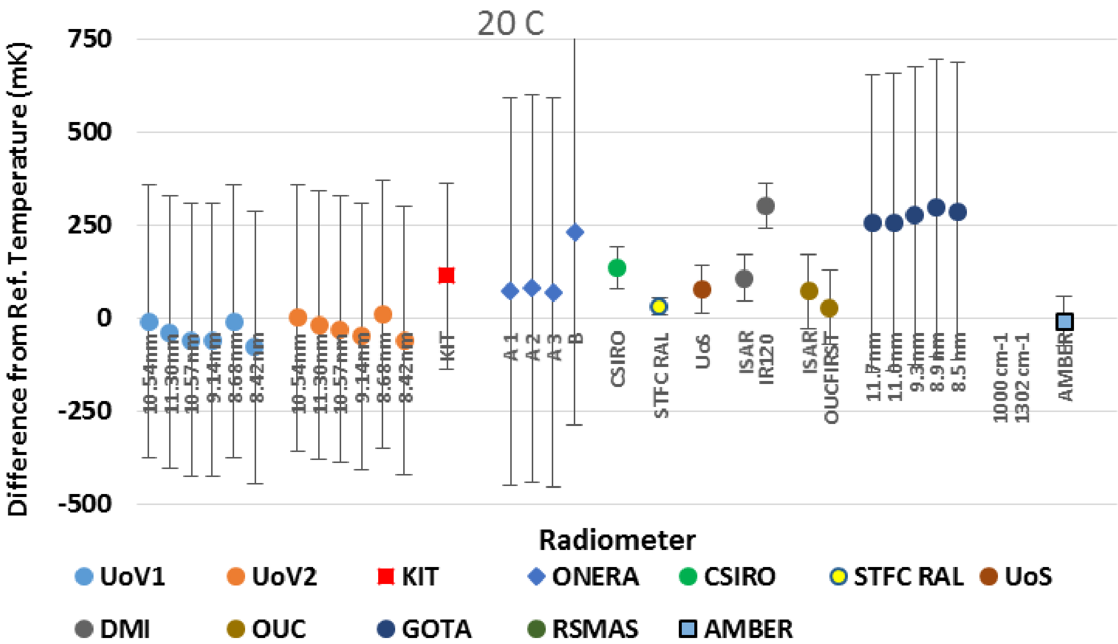

Figure 8: Plot of the mean of the differences of the radiometer readings from the temperature of the NPL reference blackbody, maintained at a nominal temperature of 20 °C.

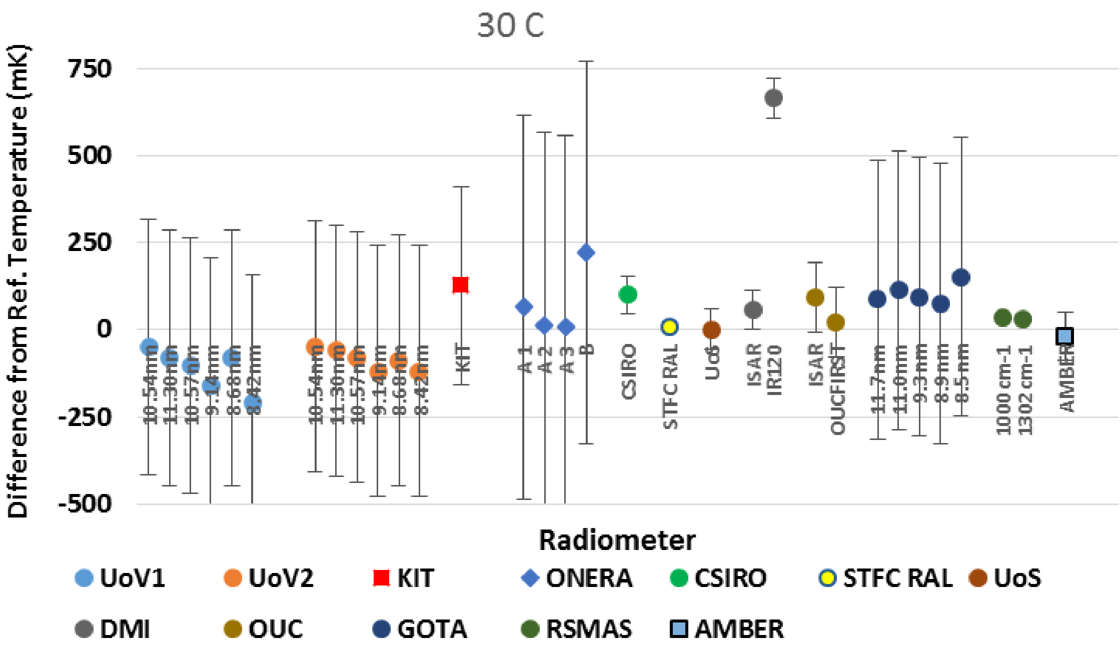

Figure 9: Plot of the mean of the differences of the radiometer readings from the temperature of the NPL reference blackbody, maintained at a nominal temperature of 30 °C.

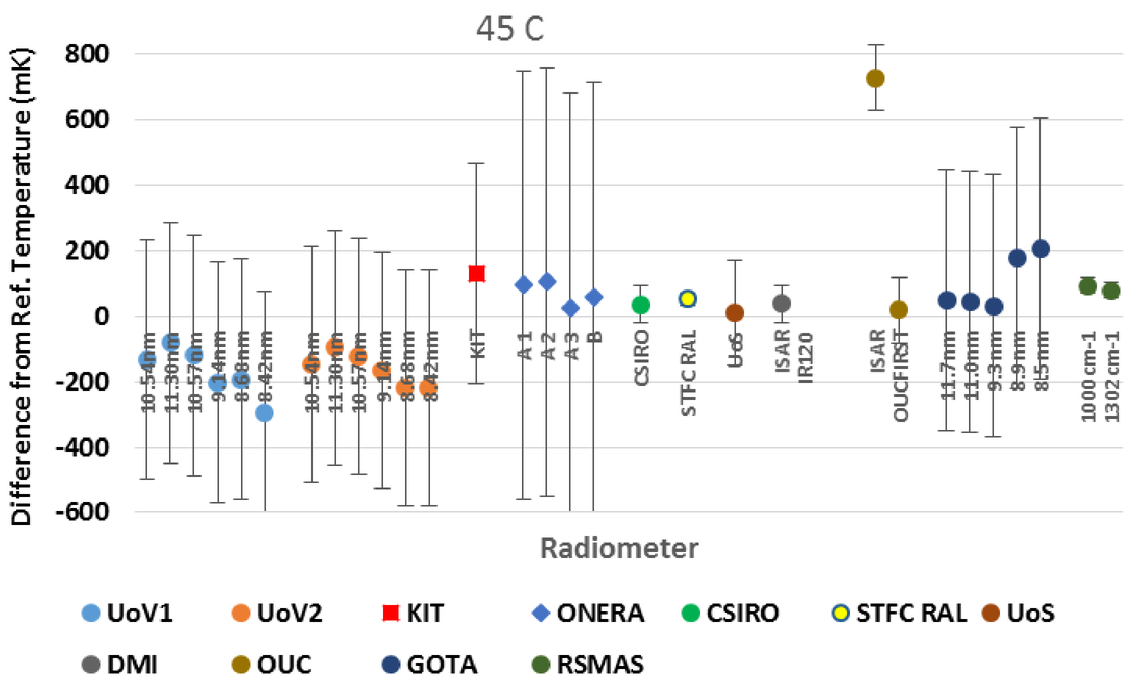

Figure 10: Plot of the mean of the differences of the radiometer readings from the temperature of the NPL reference blackbody, maintained at a nominal temperature of 45 °C.