fiducial reference measurements for satellite ocean colour

D-170: Technical Report TR-6 "Results from the First FRM4SOC Field Ocean Colour Radiometer Verification Round Robin Campaign"

Title	Technical Report TR-6 "Results from the First FRM4SOC Field Ocean Colour Radiometer Verification Round Robin Campaign"
Document reference	FRM4SOC-TR6
Project	ESA – FRM4SOC
Contract	ESRIN/Contract No. 4000117454/16/1-SBo
Deliverable	D-170 (Technical Report TR-6)
ATTN	Tânia G. D. Casal ESA/ESTEC Technical Officer Keplerlaan 1 2201 AZ Noordwijk The Netherlands
Version	**1.0**
Date issued	01.10.2018

	Prepared by	**Signed By**	**Approved by**
Name:	Joel Kuusk	Riho Vendt	Tânia G. D. Casal
Organisation:	Tartu Observatory, University of Tartu	Tartu Observatory, University of Tartu	ESA/ESTEC
Position:	WP leader	Project manager	Technical Officer
Date:	01.10.2018	Digitally signed by Riho Vendt Date: 2018.10.01 19:56:30 +03'00'	
Signature:			

UNIVERSITY OF TARTU
Tartu Observatory

Fiducial Reference Measurements for Satellite Ocean Colour (FRM4SOC) – Laboratory Calibration Exercise 2 (LCE-2): Verification of Fiducial Reference Measurement Ocean Colour Radiometers (FRM OCR)

D-170: Results from the First FRM4SOC Field Ocean Colour Radiometer Verification Round Robin Campaign (TR-6)

TECHNICAL REPORT

Joel Kuusk, Viktor Vabson, Ilmar Ansko, Riho Vendt, Krista Alikas

Data and processing by Krista Alikas, Ilmar Ansko, Ave Ansper, Mariano Bresciani, Henning Burmester, Maycira Costa, Davide D'Alimonte, Girogio Dall'Olmo, Bahaiddin Damiri, Tilman Dinter, Claudia Giardino, Kersti Kangro, Joel Kuusk, Martin Ligi, Birgot Paavel, Kevin Ruddick, Gavin Tilstone, Viktor Vabson, Ronnie Van Dommelen, Sonja Wiegmann

Photos by Viljo Allik, Joel Kuusk, Martin Ligi

Document Control Table

Title	Technical Report TR-6 "Results from the First FRM4SOC Field Ocean Colour Radiometer Verification Round Robin Campaign"
Document reference	FRM4SOC-TR6
Project	ESA – FRM4SOC
Contract	ESRIN/Contract No. 4000117454/16/1-SBo
Deliverable	D-170 Technical Report TR-6
Version	**1.0**
Date Issued	01.10.2018

Document Change Record

Index		Issue	Revision	Date	Brief description	Issued by
1.		0	1	01.08.2018	DRAFT	Viktor Vabson
2.		0	2	07.09.2018	Draft sent to participants	Joel Kuusk
3.		1	0	01.10.2018	Updated according to revision from Andrew Barnard and Ronnie Van Dommelen (Sea Bird Scientific)	Joel Kuusk

Distribution List

Company/Organisation	Name	Format	No. of Copies
Tartu Observatory, UT	Riho Vendt	Electronic file – pdf and original (WORD) file	1 pdf 1 WORD file
Tartu Observatory, UT	Joel Kuusk	Electronic file – pdf	1
RBINS	Kevin Ruddick	Electronic file – pdf	1
NPL	Agnieszka Bialek	Electronic file – pdf	1
PML	Gavin Tilstone	Electronic file – pdf	1
ACRI-ST	Christophe Lerebourg	Electronic file – pdf	1
ESA ESTEC	Craig Donlon	Electronic file – pdf and original (WORD) file	1 pdf 1 WORD file
ESA ESTEC	Tânia Casal	Electronic file – pdf and original (WORD) file	1 pdf 1 WORD file
ESA ESRIN	Nathalie Schwarz	Electronic file – pdf and originaal (WORD) file	1 pdf 1 WORD file

Contents

Executive Summary

This document, the D-170: Technical Report TR-6 "Results from the First FRM4SOC Field Ocean Colour Radiometer Verification Round Robin Campaign" is written following the Contract No. 4000117454/16/I-SBo between the European Space Agency (ESA) and University of Tartu as stated in the Statement of Work, for the ESA Invitation to Tender (ITT) ESA/AO/1-8500/15/I-SBo Fiducial Reference Measurements for Satellite Ocean Colour (FRM4SOC). The document addresses the requirements to:

- report in full the activities of the LCE-2 and the results obtained;
- conclude with a set of activities to improve future LCE and any actions required to bring OCR used for satellite validation up to FRM standards;
- address any other aspect required to ensure that TR-6 is complete.

The Laboratory Comparison Esperiment LCE2 to verify the SI traceability of Ocean Colour Radiometers (OCR) used for satellite OCR validation took place in Estonia from 08. to 13. May 2017. The exercise was organised by Tartu Observatory (TO) of the University of Tartu that also acted as the pilot laboratory for the comparison. The activities were supervised and supported by NPL.

The LCE-2 consisted of three sub-tasks:

1) SI-traceable radiometric calibration of all participating radiometers just before the start of comparison measurements (done by TO using a set of calibration standards provided by NPL) 02. – 07.05.2017;
2) indoor intercomparison of measuring stable radiance and irradiance sources in controlled environment was performed on 09. – 10.05.2017;
3) outdoor intercomparison over terrestrial water surface was performed on 11. – 12.05.2017.

In total, 11 organisations from 8 different countries with 28 radiance and 16 irradiance sensors participated in the comparison. Measurement results and information about measurement parameters were reported back to the pilot laboratory by participants in the form of spreadsheet files. Measured spectra in raw counts were also requested for unified data handling carried out by the pilot. All uncertainties had to be computed and reported according to the "Guide to the Expression of Uncertainty in Measurement".

Agreement between irradiance and radiance sensors was mostly affected by sensor calibration. For example, factory calibrations made at different times can cause differences more than ±10 %. Former calibrations at different labs and several years ago can cause differences around ±3 %. Different calculation schemes (corrections for non-linearity, stray light or for OLCI band values) can cause differences about ±1…2 % each factor. The best agreement between participants 0.5…0.8 % was achieved when a unified procedure for measurements and data handling was applied. Additionally, major part of sensors involved in LCE-2 were recalibrated at TO a year later for FICE-AAOT field intercomparison allowing to evaluate the stability of sensors. Major part of sensors (over 80 %) changed during a year less than ±1 %.

For outdoor measurements the variability between radiance sensors was about two times larger than during indoor exercise, and this can be partly explained by particular differences in data handling carried out by different participants, when large effects of outside influence factors e.g. temperature, stray light, and nonlinearity effects are not corrected.

Variability between irradiance sensors was about five times larger than during indoor exercise. It is highly likely that unexpectedly large differences revealed between sensors were caused by rather high inter-instrument variability of cosine response and different

construction of input optics of different types of sensors. Variability cannot be fully explained by influence factors like temperature, nonlinearity, and stray light.

A list of recommendations is proposed to be followed in organising future outdoor intercomparison campaigns as lessons learned from the presented exercise.

Acronyms and Abbreviations

Acronym	Abbreviation
AAOT	Aqua Alta Oceanographic Tower
ADC	Analog-to-digital converter
AERONET	Aerosol Robotic Network
CDOM	Colored dissolved organic matter
EO	Earth Observation
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FICE	Field Inter-Comparison Experiment
FOV	Field of View
FWHM	Full Width at Half Maximum
FRM4SOC	Fiducial Reference Measurements for Satellite Ocean Colour
IOP	Inherent Optical Properties
ISO	International Organization for Standardization
LCE	Laboratory Comparison Experiment
LSF	Line Spread Function
NPL	National Physical Laboratory
OC	Ocean Colour
OCR	Ocean Colour Radiometer
OLCI	Ocean and Land Colour Instrument
PAR	Photosynthetically active radiation
QTH	Quartz Tungsten Halogen
SI	Systeme International d'Unites
TO	Tartu Observatory
TR	Technical Report
TSM	Total Suspended Matter
UTC	Coordinated Universal Time

List of symbols

Symbol	Definition
E	Irradiance
E_d	Downwelling irradiance
λ	Wavelength
L	Radiance
L_d	Downwelling radiance
L_u	Upwelling radiance
L_w	Water-leaving radiance
R_{RS}	Remote sensing reflectance

1. Introduction

Laboratory Comparison Experiment LCE-2 was organized in the frame of the FRM4SOC project with the aim to link the OC field measurements to the SI-traceable calibration and verify whether different instruments measuring the same object will provide consistent results within the expected uncertainty limits. As an outcome, methodologies used by participants for the measurements and data handling were also critically reviewed. In preparation of this report [1], [2] were used.

All the radiometers participating in LCE-2 were radiometrically calibrated in TO just before the experiment using SI-traceable radiometric calibration standards provided by NPL. The LCE-2 experiment was divided into two parts – indoor and outdoor experiment. The indoor experiment was used to determine best possible agreement between commonly used OC radiometers. The factors that could possibly affect the measurements were minimised and the measurements were carried out in a controlled laboratory environment closely matching the conditions during radiometric calibration. The outdoor experiment was a step towards actual in situ OC measurements. Downwelling solar irradiance and sky and water radiance were measured and intercompared, but no derived OC parameters such as water-leaving radiance or remote sensing reflectance were calculated. There were a few notable differences compared to the indoor experiment:

1. ambient temperature was roughly 15 ˚C lower than during the radiometric calibration of the radiometers;
2. spectral composition of the target signal (sky, water) was different compared to the radiometric calibration standard (incandescent source);
3. the angular distribution of downwelling irradiance was significantly different than during radiometric calibration (normal illumination);
4. due to variable nature of natural illumination it was not possible to measure the targets using different integration times.

Due to non-ideal performance of radiometers (nonlinear response, temperature dependence, spectral stray light, deviation from cosine response, etc.) all these differences between conditions during laboratory calibration and field measurements contribute to the measurement uncertainty. The known measurement errors should be corrected and the unknown or residual errors have to be adequately assessed in the uncertainty budget.

2. Organisation and protocol

LCE-2 took place on 08. – 13.05.2017. TO was the main organizer for this comparison, and, supported by NPL, TO acted also as the pilot laboratory.

LCE-2 consisted of three sub-tasks:

1. SI-traceable radiometric calibration of participating radiometers just before the LCE-2 was carried out on 02. – 07.05.2017;
2. indoor intercomparison of measuring stable radiance and irradiance sources in controlled environment was performed on 09. – 10.05.2017;
3. outdoor intercomparison over terrestrial water surface was performed on 11. – 12.05.2017.

The measurement results and information about measurement parameters were reported back to the pilot laboratory by participants (Table 1) in the form of spreadsheet files. Measured spectra in raw counts were also claimed for unified data handling carried out by the pilot. All uncertainties had to be computed and reported according to the "Guide to the Expression of Uncertainty in Measurement" [3].

Instruments used for the intercomparison are listed in Table 2, and some important technical parameters of the participating radiometers are given in Table 3.

3. Participants and instruments

Table 1. Participants of the intercomparision.

Participant	Acronym	Country	Contact person
Tartu Observatory (pilot)	TO	Estonia	Joel Kuusk
Alfred-Wegener-Institut	AWI	Germany	Sonja Wiegmann, Tilman Dinter
Royal Belgian Institute of Natural Sciences	RBINS	Belgia	Kevin Ruddick
National Research Council of Italy	CNR	Italy	Claudia Giardino, Mariano Bresciani
University of Algarve	CIMA	Portugal	Davide D'Alimonte
University of Victoria	UVIC	Canada	Maycira Costa
Satlantic; Sea Bird Scientific	Satlantic	Canada	Ronnie Van Dommelen
Plymouth Marine Laboratory	PML	UK	Gavin Tilstone
Helmholtz-Zentrum Geesthacht	HZG	Germany	Henning Burmester
University of Tartu	UT	Estonia	Birgot Paavel
Cimel Electronique S.A.S	Cimel	France	Bahaiddin Damiri

Figure 1. Participants and organizers of the intercomparison LCE-2. Additional participants not shown in the photograph: Mariano Bresciani, Claudia Giardino, Ronnie Van Dommelen.

Table 2. Instruments used for the intercomparision.

Participant	Sensor type
TO	RAMSES: 2 radiance, 1 irradiance, WISP3 (2 radiance, 1 irradiance)
AWI	RAMSES: 2 radiance, 2 irradiance
RBINS	RAMSES: 7 radiance, 4 irradiance
CNR	SR-3500 (1 radiance, 1 irradiance), WISP3 (2 radiance, 1 irradiance)
CIMA	RAMSES: 2 radiance, 1 irradiance
UVIC	HyperOCR, 2 radiance 1 irradiance
Satlantic	HyperOCR, 2 radiance 1 irradiance
PML	HyperOCR, 2 radiance 1 irradiance
HZG	RAMSES: 2 radiance, 1 irradiance
UT	RAMSES: 1 radiance, 1 irradiance
Cimel	SeaPRISM (1 radiance)
In total	**28 radiance and 16 irradiance sensors**

Table 3. Technical parameters of the participating radiometers

Parameter	RAMSES	HyperOCR	WISP-3	SR-3500	SeaPRISM
Field of View	7°/cos	6°[1]/23°/cos	3°/cos	5°/cos	1.2°
Manual integration time	yes	Yes	no	yes	no
Adaptive integration time	yes	yes	yes	yes	yes
Min. integration time, ms	4	4	0.1	7.5	NA
Max. integration time, ms	4096	4096	NA	1000	NA
Min. sampling interval, s	5	5	10	2	NA
Internal shutter	no	yes	no	yes	yes
no of channels	256	256	2048	1024	12
Wavelength range, nm	320.. 1050	320..1050	200..880	350..2500	400..1020
wavelength step, nm	3.3	3.3	0.4	1.2/3.8/2.4	NA
spectral resolution, nm	10	10	3	3/8/6	10

Figure 2. Instruments participating in the LCE-2 intercomparison.

[1] According to the manufacturer, the HyperOCR radiance sensors 444 and 445 have 6° FOV.

4. Radiometric calibration of participating radiometers

In the frame of LCE-2, absolute radiometric calibration of the participating radiometers (Table 2, Table 3) was performed in the optical radiometry laboratory of Tartu Observatory, Estonia. Calibration measurements were performed at the room temperature of 21.5 °C ± 1.5 °C in EN ISO 14644 Class 8 cleanroom environment. The calibration method and data processing are described in details in [2].

4.1. Calibration of irradiance sensors

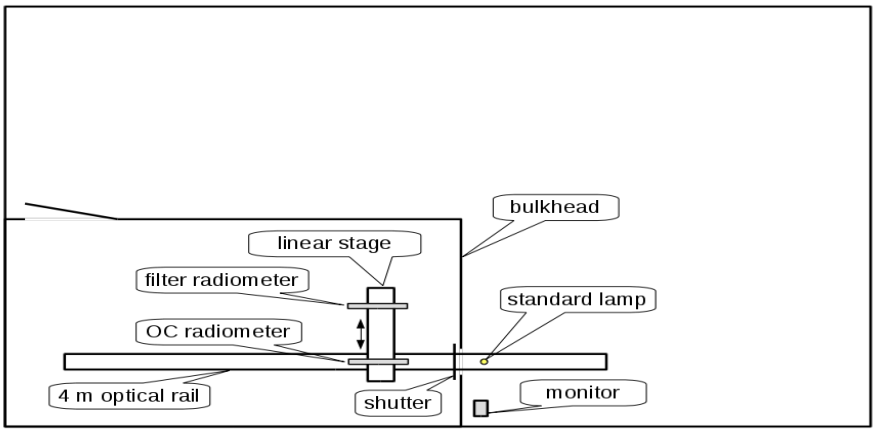


Figure 3. Irradiance calibration setup.

FEL type 1000 W quartz tungsten halogen spectral irradiance standard lamp was used for radiometric calibration of the radiometers. The lamp was powered by a stabilized radiometric power supply Newport/Oriel 69935 ensuring proper polarity as marked on the lamp. The lamp was operated in constant current mode. A custom designed circuit was used for monitoring the lamp current through a 10 mΩ shunt resistor P310 and providing feedback to the power supply. Lamp current was stabilized to better than ±1 mA. The same feedback unit was used for logging the lamp current and voltage. Voltage was measured with a 4-wire sensing method from the connector of the lamp socket. The power supply was turned on and slowly ramped-up to the working current of the lamp. Calibration measurements were started after at least a 20 min warm-up time. During calibration the voltage across the lamp terminals was also measured, and compared to the voltage measured during the last calibration of the lamp. A significant change in the lamp's operating voltage would have indicated that it was no longer usable as a reliable working standard of spectral irradiance. On completion of the calibration, the lamp current was slowly ramped down to avoid thermally shocking the filament.

The lamp and OC radiometer being calibrated were mounted on an optical rail that passed through a bulkhead which separated the lamp and radiometer during calibration (Figure 3). A computer-controlled electronic shutter with a Ø60 mm aperture was attached to the bulkhead. The shutter was used for dark signal measurements during calibration. Two additional baffles with Ø60 mm apertures were placed between the bulkhead and the radiometer at 50 mm and 100 mm distances from the bulkhead.

The OC radiometer being calibrated was mounted next to a filter radiometer on a computer-controlled linear translation stage which allowed perpendicular movement with respect to the optical rail. The positions of both radiometers were carefully adjusted before calibration

and the translation stage positions saved in the controlling software. This allowed fast and accurate swapping of the radiometers when the lamp was turned on. Many radiometers were of the same make and model (TriOS RAMSES group and Satlantic HyperOCR group) and all the instruments within group had identical outside diameter. This allowed using a V-block for mounting the radiometers during calibration. Distance from the lamp was measured individually for each sensor before the lamp was turned on and a clamp at appropriate position was attached to the sensor. During calibration the radiometers of the same type were swapped without turning off the lamp. Placing the clamp against the end of the V-block ensured proper distance between the lamp and the radiometer during calibration.

The distance between the lamp and the radiometer was measured with a custom designed measurement probe. One end of the probe was placed against the socket of the lamp and the other end of the probe had two lasers with beams intersecting at 120° angle (Figure 8). The point of intersection defined the other endpoint of the probe. Such a design allowed contactless distance measurement and there was no need for touching the diffuser surface of the radiometer. The measurement accuracy of the distance probe was better than 0.2 mm.

The filter radiometer was used for monitoring possible long term drifts of the standard lamp. The filter radiometer was based on a 3-element trap detector with Hamamatsu S1337-11 windowless Si photodiodes and temperature-controlled bandpass filters with peak transmittances at nominal wavelengths 340 nm, 350 nm, 360 nm, 380 nm, 400 nm, 450 nm, 500 nm, 550 nm, 600 nm, 710 nm, 800 nm, 840 nm, 880 nm, 940 nm, and 980 nm. The photocurrent of the filter radiometer was amplified and digitized with a Bentham 487 current amplifier with integrating ADC. Newport 350B temperature controller was used for stabilizing the temperature of the bandpass filters. The filters were changed manually and it took about two minutes for the temperature of the filter to stabilize. As the OC radiometer and filter radiometer could not been used simultaneously, an additional monitor detector was used for recording short-term changes in the lamp intensity during calibration.

At least two different integration times were used for each sensor (except in the case of the SeaPRISM and WISP3 instruments for which the provided standard measurement programs were used). After a warm up time, at least 30 (10 in the case of WISP-3, internal averaging) spectral measurements were collected measuring the radiation from the lamp. Next, the shutter in front of the lamp was closed and the same number of spectral measurements were collected, in order to estimate dark signal and ambient stray light. All measurements were repeated at least twice, including readjustment of the lamp and the sensor.

NPL provided two Gigahertz-Optik BN9101-2 FEL-type irradiance calibration standard lamps with S/N 399 and 401 for the LCE-2 exercise. The lamps were calibrated by NPL and had not been used since the last calibration. Differences of responsivity in the range from 340 nm to 980 nm determined by these lamps for a precision filter radiometer with 3-element trap detector were less than ±0.5 %. The drift of the irradiance values (at 500 nm) measured during the calibration campaign was ~0.1 % which is close to the detection limit of the filter radiometer. In certificates issued for LCE-2 radiometers, arithmetic mean of the responsivity measured by the two lamps was used.

4.2. Calibration of radiance sensors

Radiance sensor calibration setup (Figure 4) was based on the lamp/plaque method and utilized the same components as the irradiance sensor calibration setup. A Sphere Optics sg3151 (200×200) mm calibrated white reflectance standard was mounted on the linear translation stage next to the filter radiometer. Normal incidence for the illumination and 45°

from normal for viewing were used. The panel was calibrated in the same illumination and viewing conditions by NPL during LCE-1. A mirror in a special holder and an alignment laser were used for aligning the plaque and radiance sensor. As in the case of irradiance sensors, at least 30 spectra were acquired using two different integration times (3 readings for SeaPRISM and 10 spectra for WISP-3, automatic integration time) and the background spectra. All measurements were repeated at least twice, including readjustment of lamp, plaque, and sensor.

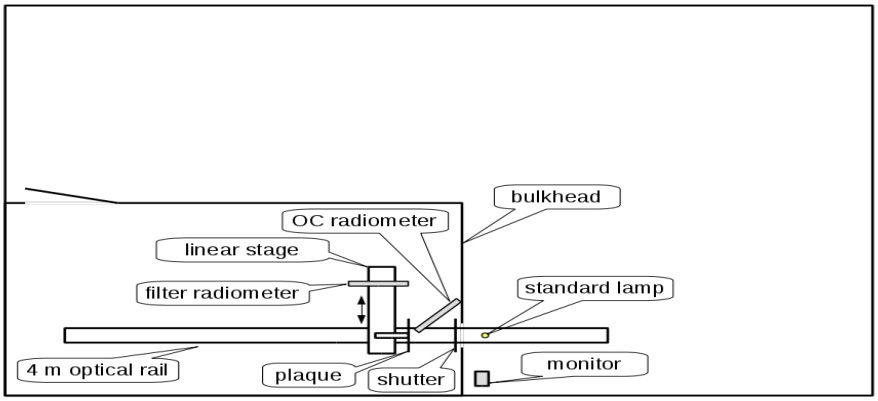


Figure 4. Radiance calibration setup.

4.3. Stability of radiometric calibration

Analysis of the LCE-2 calibration results, comparing them with former calibrations, including the factory calibrations, and also with calibrations carried out by TO one year later (before the FRM4SOC FICE-AAOT intercomparison) reveals interesting information about stability of the sensors. Some uncertainty contributions characteristic to calibration can also be estimated.

In Figure 5, variability of calibration coefficients of radiance and irradiance sensors due to adjustment of the lamps, plaques, and sensors, and due to short-term instability of the lamps and sensors are depicted. Before LCE-2 all the radiometers were calibrated using the same pair of lamps (Chapters 4.1 and 4.2). For each sensor two sets of calibration coefficients were obtained and the difference between the lamps was calculated as the ratio of these coefficients. The curves in Figure 5 are calculated as standard deviations from the ratios of calibration coefficients determined with two different standard lamps. Using standard deviation allows to exclude the systematic difference between lamps (traceability to SI) and illustrate the other components related to individual setup and measurement. Data in Figure 5 include calibration of more than 25 sensors for LCE-2 intercomparison and for FICE-AAOT intercomparison one year later when different pair of lamps was used, but the procedure was similar. Remarkable is the rapid increase of variability in the UV region of the spectrum.

Figure 6 shows average long-term variability of calibration coefficients of TriOS RAMSES and Satlantic HyperOCR radiance and irradiance sensors. All the radiometers had previous radiometric calibration certificates of various origin and age. The curves in Figure 6 are calculated similarly to Figure 5 as standard deviations of the calibration coefficient ratios. It has to be noted, however, that systematic differences between lamps is not excluded when the radiometric calibration coefficients obtained at TO are compared to last known calibrations, because the previous calibrations are not made using the same calibration standards for all the radiometers. Many of the RAMSES and HyperOCR radiometers that participated in LCE-

2 also took part in the FICE-AAOT field intercomparison experiment one year later. Those sensors were radiometrically calibrated at TO in June 2018 before the beginning of the field campaign. This gave a good opportunity to estimate the long-term stability of the sensors while minimising other possible factors influencing the calibration result. The sensors were calibrated in the same laboratory by the same operator in similar environmental conditions using the same calibration setup and methodology. Only the FEL type irradiance calibration standards were not the same lamps no. 399 and 401 used during LCE-2. Nevertheless, all the sensor were calibrated once again using common pair of lamps and the systematic differences between lamps used during LCE-2 and FICE-AAOT calibrations are excluded for the L_one_year and E_one_year curves in Figure 6. Comparing the two calibrations done in the same lab one year apart showed that most of the sensors (over 80 %) have changed less than ±1 % over the year. This allows to conclude that the inherent long-term stability of the sensors is much better than 5 % to 10 % that was revealed from the calibration history and most of the differences are caused by other factors such as different calibration standards, environmental conditions, calibration setups and methodologies, etc. However, stepwise changes in the responsivity of some TriOS RAMSES irradiance sensors may cause even larger deviations which cannot be explained by other factors than the instability of the sensor itself. No stepwise changes were observed for the RAMSES radiance sensors, however, even after exclusion of obvious outliers from the irradiance sensors, the stability of RAMSES radiance sensors is still better compared to the irradiance sensors. Some effects causing variability of the results (also during radiometric calibration) are discussed in Chapter 7.1.

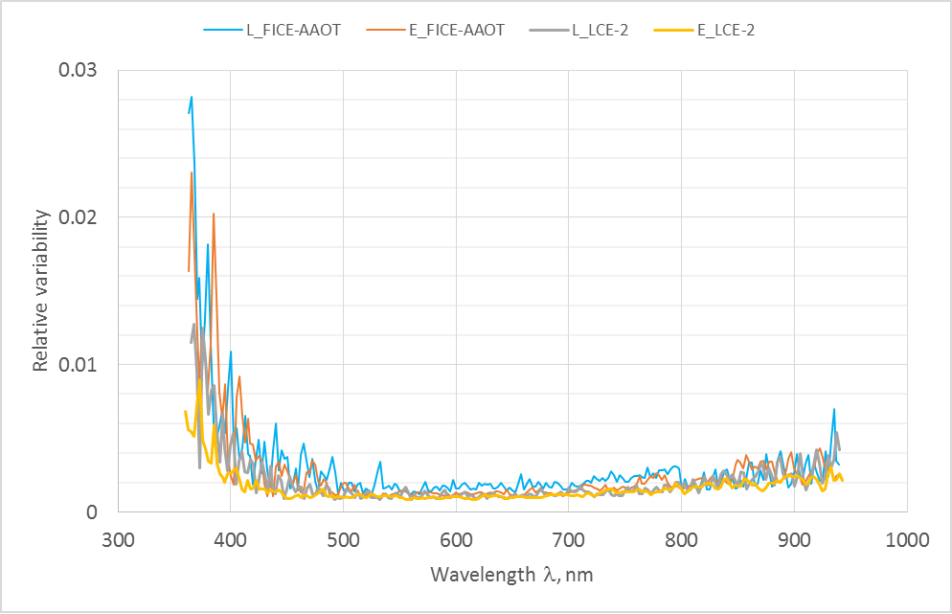


Figure 5. Relative variability of calibration coefficients of radiance (L) and irradiance (E) sensors with two different lamps used for calibration before LCE-2 and before FICE-AAOT.

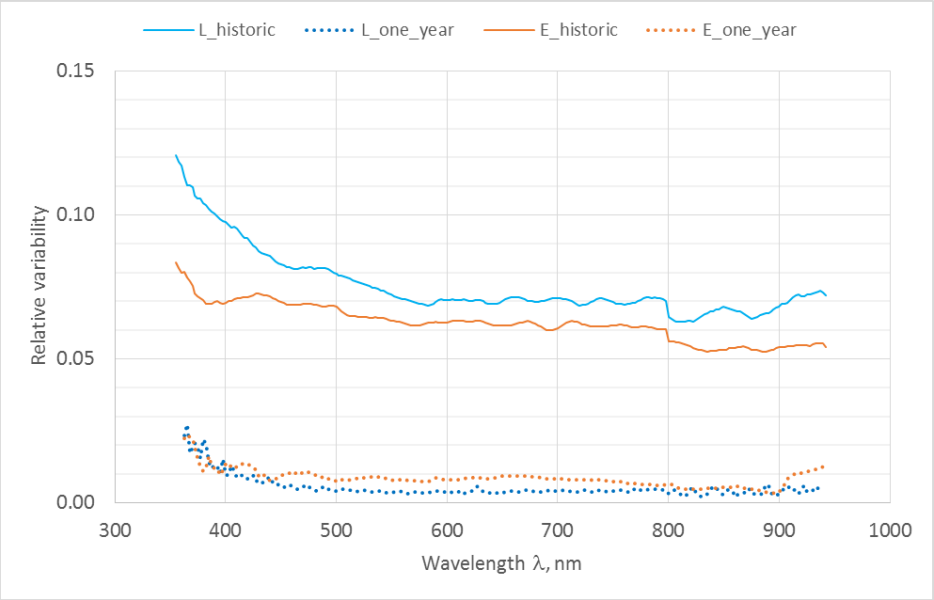


Figure 6. Relative variability of calibration coefficients of radiance (L) and irradiance (E) sensors: historic - difference of previous known calibrations and results of LCE-2 calibration; one_year - changes during one year after LCE-2 calibrations, some extra-large changes excluded.

5. LCE-2 indoor exercise

5.1. Venue and measurement setup

The indoor intercomparison (Figure 7) took place at Tartu Observatory, Estonia. Stable radiance and irradiance sources were used for verifying the performance of OC radiometers.

Figure 7. Indoor intercomparison exercise.

The irradiance setup can be seen in Figure 8. An FEL lamp was used as a stable irradiance source for indoor intercomparison. The power supply, current feedback unit, monitor detector, and distance measurement probe were the same as used during the radiometric calibration, but the FEL lamp and measurement distance were different. In order to change and align the radiometers without switching off the lamp, an additional alignment jig was placed between the shutter and the radiometer. When the shutter was closed, it was possible to change and realign the radiometer with respect to the jig. The alignment jig support was fixed to the optical rail during the whole intercomparison experiment and used as a reference

plane for distance measurement. During the intercomparison the FEL source was switched off only once in the evening of May 9, the first day of the indoor exercise.

Each participant measured the irradiance source using two different integration times (with corresponding shutter measurements) and one series with the instrument rotated by 90° around the optical axis. The latter was used to estimate the uncertainty related to the polarization sensitivity of the irradiance sensors. Each series was expected to contain at least 30 readings. As an exception, for the WISP3 instruments two series (including re-alignment) of 10 readings were recorded and one series with the shutter closed.

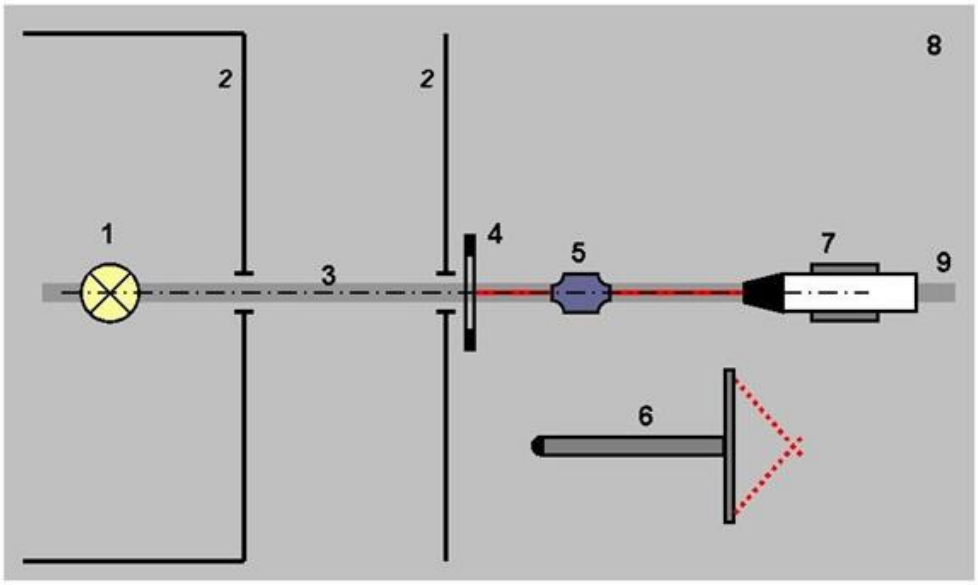


Figure 8. Indoor irradiance comparison. 1 - FEL lamp; 2 - baffles; 3 - main optical axis; 4 - alignment jig; 5 - alignment laser; 6 - distance tool; 7 - radiometer on the support; 8 - optical table; 9 - optical rail.

The radiance setup for indoor intercomparison is depicted in Figure 9. A Bentham ULS-300 integrating sphere with internal illumination was used as a stable radiance source. ULS-300 is a 300 mm integrating sphere with Ø100 mm target port. According to the manufacturer, the uniformity of radiance over the output aperture is ±0.05 % independent of the intensity setting. The sphere has a single 150 W quartz tungsten halogen light source (Osram Sylvania HLX 64640) and an 8-branch fibre bundle for transporting the light into the sphere. The sphere has a variable mechanical slit between the light source and the fibre bundle which allows changing the intensity of light inside the sphere while maintaining the spectral composition of light which corresponds to correlated colour temperature (3100±20) K. The lamp was powered by a Bentham 605 stabilized power supply at 6.3 A constant direct current. A Gigahertz-Optik VL-3701-1 broadband illuminance sensor attached directly to the sphere was used as a monitor detector. The monitor detector current was recorded by an Agilent 3458A multimeter, lamp voltage was measured by a Fluke 45 multimeter. Each participant measured the sphere source at two radiance levels and two distances from the sphere. The monitor detector current reading was used for setting the same sphere radiance levels for all the participants. 1 µA monitor current was used for low radiance measurements corresponding roughly to the typical water radiance during field measurements whereas 10 µA monitor current was used to simulate typical sky radiance. Obviously, the spectral composition of the incandescent sphere source did not match the field spectra, but was rather similar to the emission of the FEL-type calibration standard. In addition to sphere radiance, dark measurements were recorded by placing a baffle between the sphere and the

radiometer. The sphere radiance was measured at two distances, typically 17 cm and 22 cm from the sphere port. Although the radiance measurement should not depend on measurement distance as long as the sphere port overfills the FOV of the radiometer, the results measured at two distances were used to estimate the uncertainty component caused by back-reflection from the radiometer into the sphere.

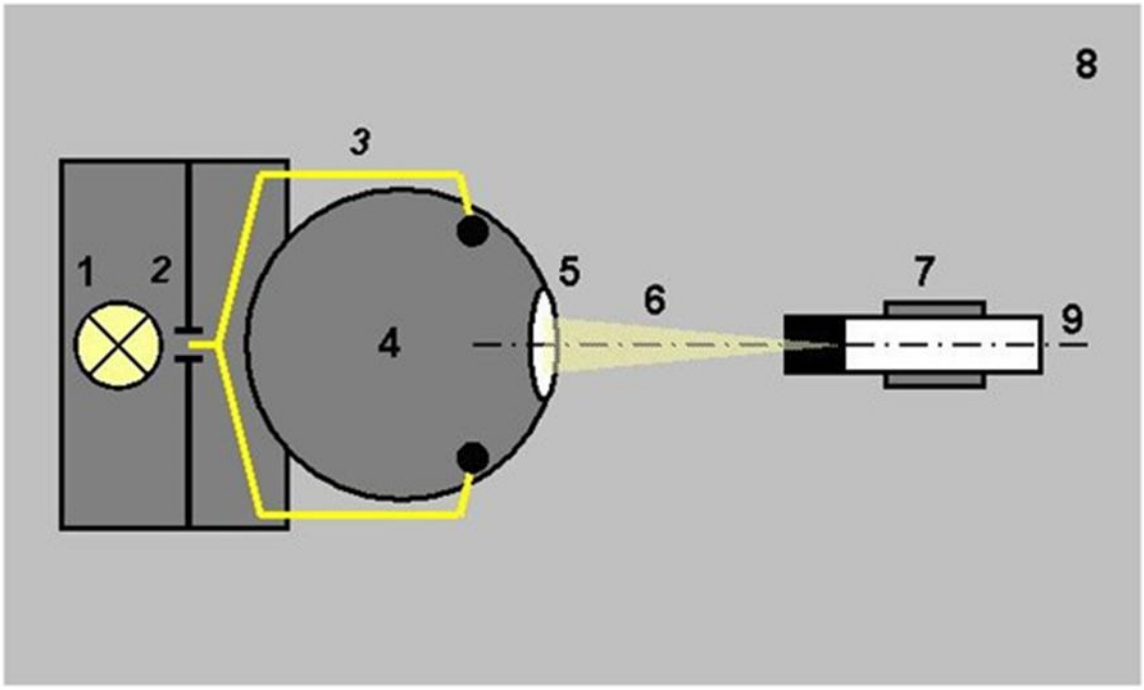


Figure 9. Indoor radiance comparison. 1 - quartz tungsten halogen lamp; 2 - variable slit; 3 - optical fibre; 4 - integrating sphere; 5 - output port; 6 - FOV of the radiometer; 7 - radiometer on the support; 8 - optical table; 9 - main optical axis.

5.2. Results

The measurement results, including measurement uncertainty and information about measurement parameters, were reported back to the pilot laboratory by most of the participants (for 33 out of 44 sensors involved) in the form of spreadsheet files. For the rest, the pilot carried out the data analysis based on the raw instrument data. In the case of discrepancies, pilot repeated the calculations on raw user data applying unified data handling described in the next chapter. Nevertheless, due to differences in hardware and software of the participating radiometers, fully unified data handling was not possible.

5.3. Data processing

The participants were encouraged to perform the data processing for their radiometers and report back the radiance/irradiance values with uncertainty estimates. However, in few cases, TO accomplished/repeated the calculations for other participants too. Data processing for the RAMSES, HyperOCR and WISP-3 instruments at TO was fully automatized by using the purpose-designed computer programs. The source code of the programs is freely available for the project partners.

Data processing was performed on the per-instrument basis and included the following steps:

- Separation the raw datafiles based on the scene (e.g. low/high radiance, distance), integration time, shutter measurements;
- pairing the raw data with corresponding shutter measurement;
- dark signal subtraction;
- linearity correction whenever applicable;
- division by radiometric responsivity;
- recalculation for the MERIS/OLCI spectral bands;
- averaging;
- evaluation of the uncertainty.

Device-specific issues are explained in the following sections. Uncertainty analysis is described in Chapter 7.

5.3.1 TriOS RAMSES

The RAMSES series of radiometers include both the radiance (ARC) and irradiance (ACC) sensors. The raw spectra are stored in the ASCII and/or Microsoft ACCESS database files. The ASCII files were used for the calculations. Data processing for these radiometers is fully unified based on the measured data (2-byte integer numbers) and calibration files provided by factory and TO. The detailed procedure to derive the calibrated results is described in [2]. RAMSES instruments do not include internal mechanical shutter, black-painted pixels on the photodiode array are used instead to compensate for the dark signal and electronical drifts. The background spectrum (with the external shutter closed) was subtracted as well. For subtraction, only the spectra with matching integration times were used. Before division by the responsivity coefficients, linearity correction was applied to the data as described in Chapter 7.1.

5.3.2 Satlantic HyperOCR/OCR3000

As with the RAMSES instruments, the HyperOCR and its older version OCR3000 (UVIC) product family contains both the radiance and irradiance sensor, sharing similar data chain. The raw spectra are stored in binary files and were converted to ASCII by participants using the proprietary manufacturer's software. Data processing for the HyperOCR was based on the calibration file provided by TO and is similar to the one of TriOS RAMSES. The HyperOCR radiometers are equipped with an internal mechanical shutter, which is deployed automatically after every 5th target spectrum. The shutter measurements were detected in the datafiles and the closest shutter measurement was subtracted from each raw spectrum before the next steps.

5.3.3 WISP-3

The WISP-3 instrument contains a three-channel Ocean Optics JAZ module spectrometer and computer. Two of the spectrometer's input channels are connected to the radiance inputs while the third one is attached to the irradiance adaptor. Acquisition of the spectra is started by the user by pressing a button, the internal computer is taking care of the measurement sequence, determining the integration times, and storing the data. All three channels are acquired simultaneously and the data stored into a single ASCII file. The spectrometers are

taking advantage of the painted pixels as in the case of the RAMSES radiometers. The internal dark signal is subtracted automatically and resulting data is stored in the form of floating point numbers. The only operation needed was the division by the responsivity coefficients determined by TO using the same manual measurement sequence. The linearity correction described in Chapter 7.1 was not used. Finally, the radiometrically calibrated ambient background (with the external shutter closed) was subtracted.

5.3.4 Spectral Evolution SR-3500

The SR-3500 spectrometer is equipped with an optical fibre input and interchangeable radiance and irradiance foreoptics. Thus, the data processing for the radiance and irradiance measurements are identical. The spectral output is stored in the ASCII files and can contain both the raw and radiometrically calibrated results (using the coefficients stored internally). The dark signal is subtracted internally using an integrated mechanical shutter. Each target measurement is automatically followed by a dedicated dark measurement. During the radiometric calibration at TO, calibration factors to the existing coefficients were derived. The calibrated data in the files was multiplied by these factors, and finally, the linearity correction scheme (Chapter 7.1) was used.

5.3.5 CIMEL CE318 (SeaPRISM)

The CIMEL CE318 binary output was converted by the owner and returned to the pilot in the form of ASCII files. Based on these data, TO derived the radiometric calibration coefficients. Neither linearity correction scheme nor re-calculation for the OLCI spectral bands was used.

5.3.6 OLCI spectral bands

As the final data processing step, the radiance and irradiance values were re-calculated for the OLCI spectral bands for each radiometer except for the multispectral CE318, in which case the initial band values were used. Based on the given CWL-s of the spectroradiometer and the OLCI channel definition [4], the weight factors were found for each pixel:

$$C(n) = O(\lambda_n), \quad K(n) = \frac{C(n)}{\sum_n C(n)},$$

where n is the pixel number with CWL of λ_n, $O(\lambda_n)$ is the responsivity of the corresponding OLCI channel interpolated to λ_n, and $K(n)$ - the normalized weight coefficient for n'th pixel. Finally the radiance/irradiance value for the corresponding OLCI channel was calculated as

$$I = \sum_n I(n) \cdot K(n),$$

where $I(n)$ denotes the measured radiance/irradiance at the n'th pixel.

5.4. Consensus and reference values used for the analysis

Consensus values were calculated as median [5] of all presented comparison values. Reference values were applicable only for the indoor irradiance measurements (Figure 8), when the measurand used for this exercise was during comparison measured also with the precision filter radiometer serving as a reference. The LCE-1 round robin transfer radiometer was planned to be used as the radiance reference instrument but unfortunately it did not arrive before LCE-2.

In spite of different sensor types, as the radiation sources used for indoor comparison were spectrally very similar to calibration sources, agreement between sensors was satisfactory for radiance and for irradiance sensors, see Figure 10 - Figure 13. No outliers were present.

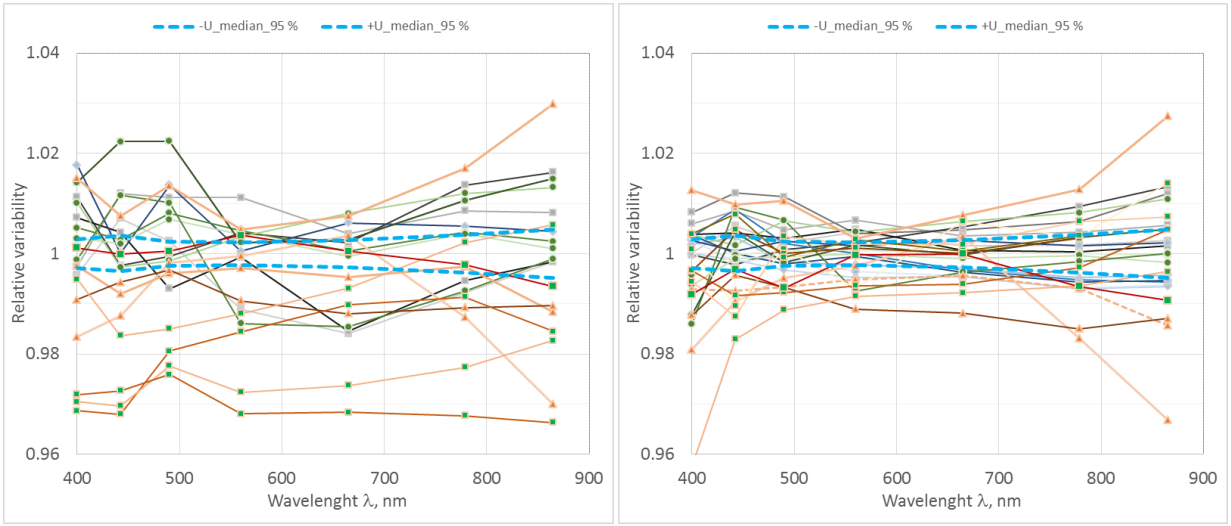


Figure 10. Low intensity radiance; agreement just after receiving data from participants (left), and after reviewing data by pilot, corrections submitted by participants and/or unified data handling by pilot (right). Blue dotted lines - expanded uncertainty of the median consensus value on the right graph.

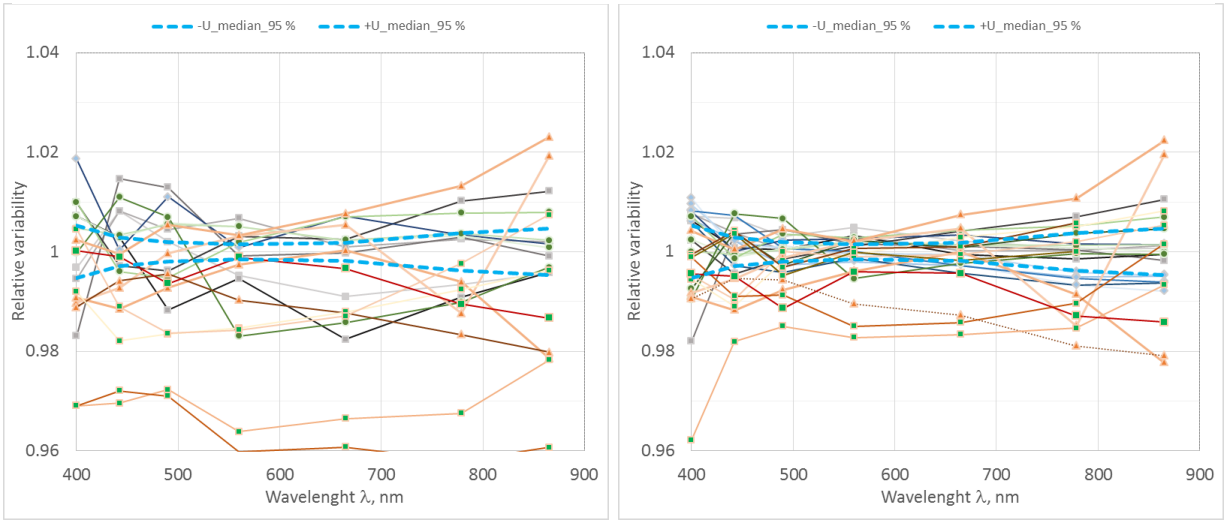


Figure 11. High intensity radiance; agreement just after receiving data from participants (left), and after reviewing data by pilot, corrections submitted by participants and/or unified data handling by pilot (right). Blue dotted lines - expanded uncertainty of the median consensus value on the right graph.

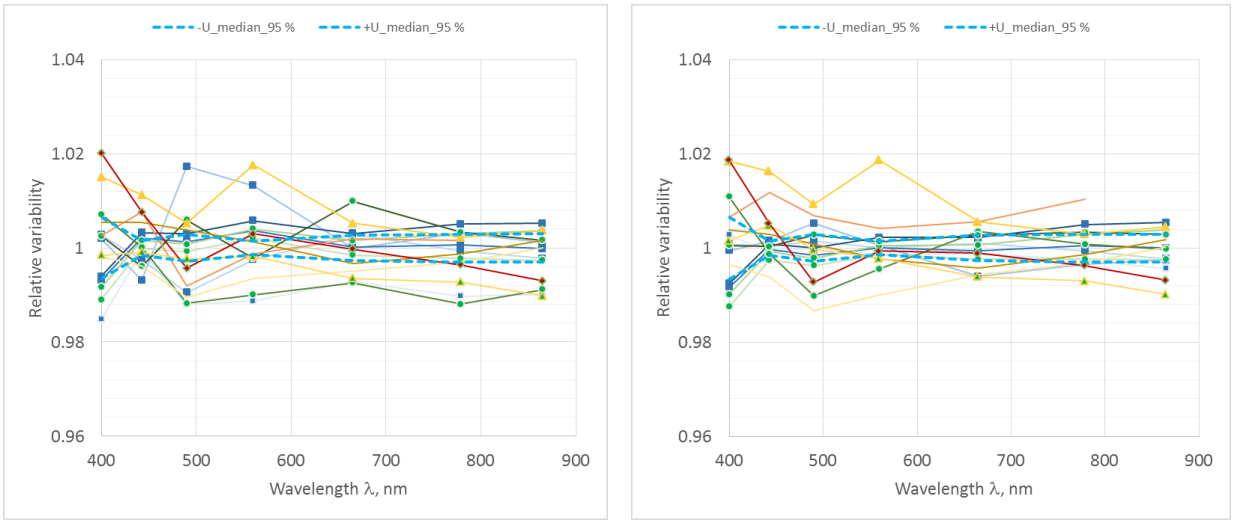


Figure 12. Irradiance sensors; agreement just after receiving data from participants (left), and after reviewing data by pilot, corrections submitted by participants and/or unified data handling by pilot (right). Blue dotted lines - expanded uncertainty of the median consensus value on the right graph.

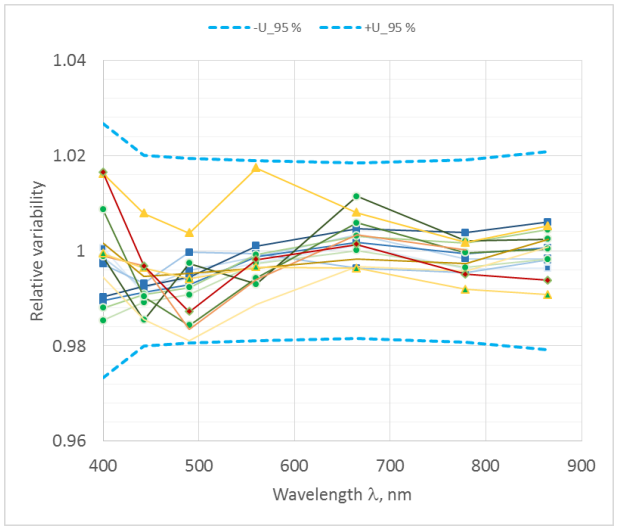


Figure 13. Irradiance sensors; agreement with reference values of the filter radiometer. *Blue dotted lines - expanded uncertainty covering 95 % of all data points. Uncertainty of radiometric calibration is included.*

Larger variability of the results initially reported by participants was caused by applying out-of-date calibration coefficients, by differently applying or not applying the linearity correction (Chapter 7.1), by differences in calculating the OLCI band values. For unified data handling (Capter 5.3) carried out by the pilot, the calibration results obtained during LCE-2 were used, non-linearity correction was applied, OLCI band values were calculated by using individual weights as determined from the wavelength scale of each radiometer (Chapter 5.3.6). After unified data handling, agreement between comparison results was significantly improved for the radiance sensors, see Figure 10 - Figure 11. There was almost no improvement in the case of the irradiance sensors, see Figure 12.

6 LCE-2 outdoor exercise

Initially [2], the outdoor intercomparison was planned in two phases:

- primary intercomparison – intercompared quantities are downwelling irradiance E_d, downwelling sky radiance L_d, and total upwelling water radiance L_u directly measured by the radiometers;
- secondary intercomparison – intercompared quantities are remote sensing reflectance R_{RS} and water-leaving radiance L_w derived from simultaneously measured E_d, L_d, and L_u.

At the first day of the outdoor measurements, 7 casts of simultaneous E_d, L_d, and L_u measurements at typical above-water 3-radiometer configuration were recorded but due to variable illumination conditions caused by cumulus clouds, none of the casts were considered suitable for the intercomparison analysis. At the second day of the outdoor experiment, priority was given to the primary intercomparison measurements and all the radiance sensors were simultaneously measuring either L_u or L_d.

6.1 Venue and measurement setup

The outdoor exercise took place at Lake Kääriku, Estonia, 58° 0' 5" N, 26° 23' 55" E on 11-12.05.2017. Lake Kääriku is a small eutrophic lake, it has 19.8 ha surface area. Maximum depth is 5.9 m, with an average of 2.6 m. The water colour is greenish-yellow, measured transparency (Secchi disk depth) was 2.6 m. The average chlorophyll content Chl = 7.3 mg m^{-3}, total suspended matter content TSM = 3.9 g m^{-3}, absorption of the colored dissolved organic matter aCDOM(442 nm) = 1.7 m^{-1}, diffuse attenuation coefficient of downwelling irradiance Kd(PAR) = 1.3 m^{-1}. The bottom is muddy.

Lake Kääriku has a 50 m long pier and a diving platform on the southern coast. The diving platform has two levels. During LCE-2 the upper level was used for the instruments, computers and instrument operators were located on the lower level and the pier below the tower (Figure 14). The instruments were located roughly 7.5 m above the water surface. Depth of water around the diving platform is 2.6 m to 3.6 m. The closest trees are about 65 m south of the platform, the treetops are less than 20° above the horizon when viewed from the upper level of the platform.

Purpose-built frames were used for mounting and aligning the participating radiometers (Figure 15 - Figure 16). The irradiance sensors were mounted in a fixed frame that ensured the levelling of the cosine collectors. The front surfaces of all the cosine collectors were set at the same height so that the illumination conditions were equal and the instruments were not shadowing each other.

Figure 14. Pier and diving platform at the southern coast of Lake Kääriku.

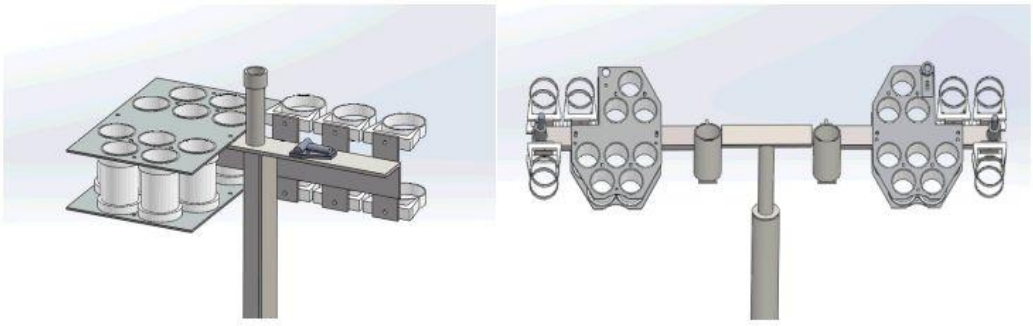


Figure 15. 3D CAD drawings of the frames for mounting irradiance (left) and radiance (right) sensors during the outdoor experiment.

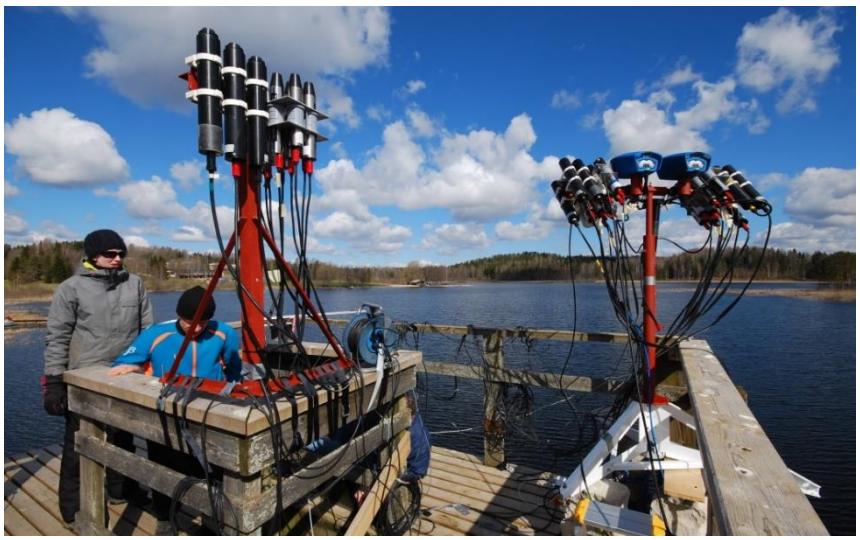


Figure 16. All the radiance and irradiance radiometers were mounted in common frames during the LCE-2 outdoor experiment. Left frame – irradiance sensors; right frame – radiance sensors.

The radiance sensors were mounted on the frame in two groups which could be moved independently in the zenith direction or the relative zenith angle between the two groups could be fixed and both groups tilted together. The relative azimuth angle between the two groups of sensors was fixed to 0° and in azimuth direction all the radiance sensors could be moved only simultaneously. The design of the radiance frame allowed mounting the L_u radiometers to one group and L_d radiometers to another group for measuring L_w and R_{RS} in a typical 3-radiometer above-water configuration. On the first day of outdoor experiment, a few casts were recorded in such configuration, but the weather conditions were not stable enough to perform useful intercomparison analysis and on second day both radiance sensor groups were aligned collinear and pointed simultaneously to either sky or water. The collinearity of groups was set by visual observation from the side of the frame and was better than 1°. Due to flexibility of the plastic clamps used to fix the HyperOCR radiometers, slight misalignment of HyperOCR and RAMSES sensors within the groups was noticed during the experiment (visually much larger than misalignment between the groups). The misalignment between HyperOCR and RAMSES sensors was measured to be 1.3° using Figure 17, the HyperOCR sensors were poining lower than the RAMSES instruments. Image taken from the other side of the frame revealed that the HyperOCR sensors in the other group were pointing about 1.1° higher than the RAMSES instruments. The left and right radiance frames were visually aligned by the topmost RAMSES instruments, thus, the misalignment between the HyperOCR instruments on the frames could have been up to 2.5°. Although this is ten times smaller compared to the nominal FOV of a standard HyperOCR instrument, it can have significant impact when measuring nonuniform targets.

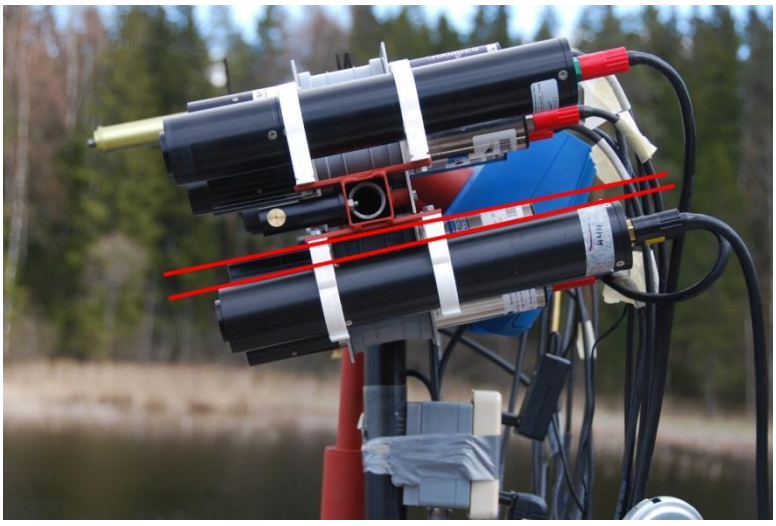


Figure 17. The angle between red lines marking the alignment of HyperOCR and RAMSES sensors was measured to be 1.3° from this image.

6.2 Environmental conditions and selection of casts

The environmental conditions during the outdoor experiment were not ideal, mainly due to the cumulus clouds. The aerosol content was low, average daily AOD_{500} was 0.077 on May 11 and 0.071 on May 12 (measured at Tõravere AERONET station, 30 km north of Lake Kääriku [6]). Air temperature was rather low and remained between 5 °C and 9 °C, water temperature was around 11 °C. Wind speed was mainly between 0.5 m s^{-1} and 4 m s^{-1} with occasional gusts of up to 7 m s^{-1}.

The outdoor measurements were performed in 5-minute casts, with the exception of 25-minute irradiance cast no. 14. The beginning and end times of casts were announced and during the casts all the participants recorded the radiance and irradiance data at their usual fieldwork data rate. 30 casts were recorded in total, but only 7 of them were included in the intercomparison. The selection of casts was based on the time series of 550 nm spectral band. The pilot received the 550 nm time series data for 16 radiance and 10 irradiance sensors. Only the casts with most stable signal and least missing data were selected for further analysis. All the selected casts were measured on May 12 - the second day of the outdoor experiment. The all-sky camera images captured in the middle of the selected casts can be seen in Figure 18.

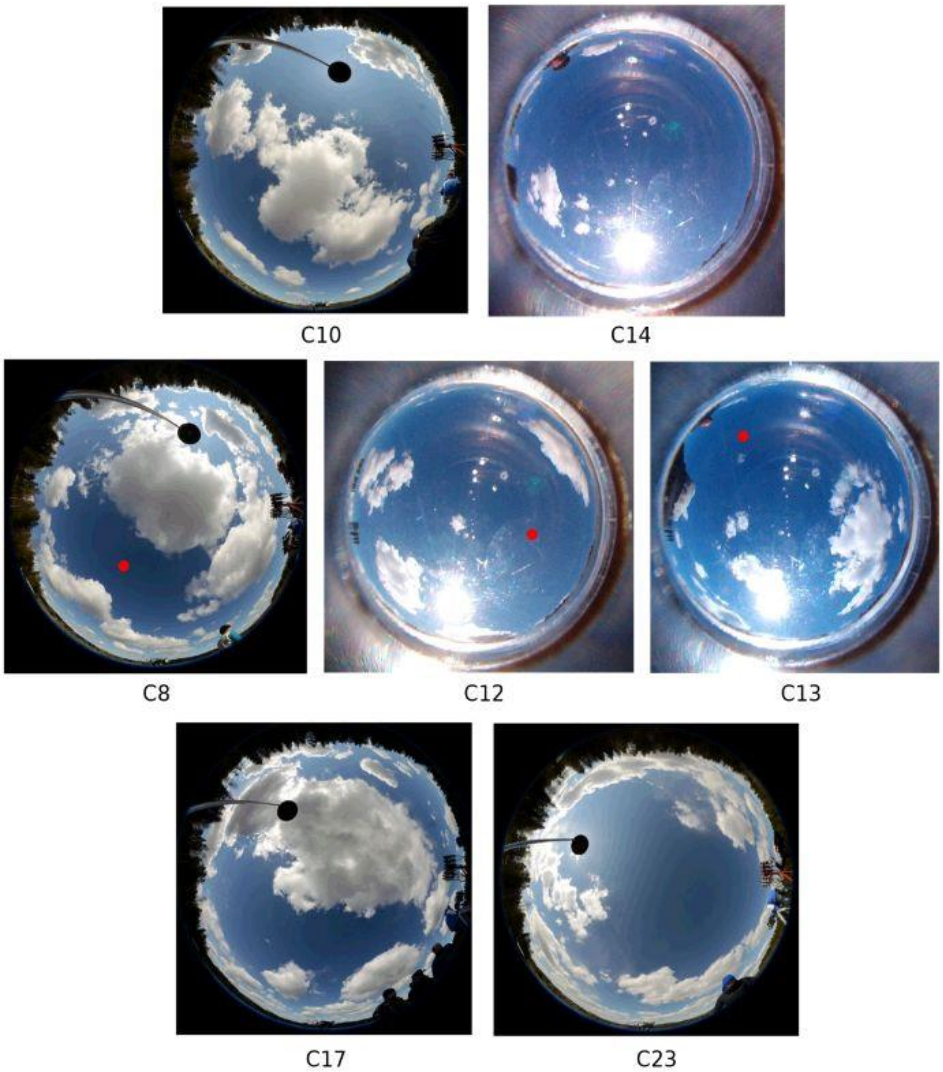


Figure 18. All-sky camera images captured in the middle of the casts used in the intercomparison analysis. Irradiance - C10, C12, C13, C14; blue sky radiance - C8, C12, C13; water radiance - C17, C23. Red dots in C8, C12, C13 indicate approximate view direction of the radiance sensors.

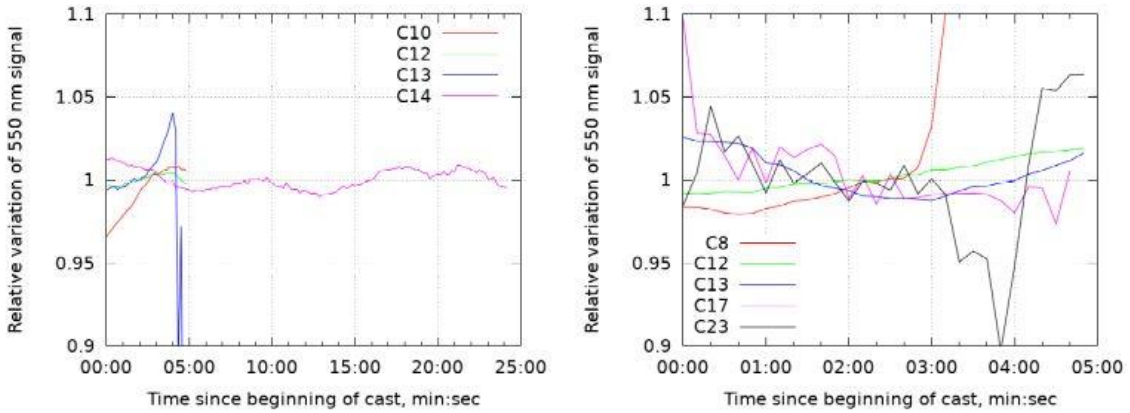


Figure 19. Relative variation of 550 nm signal of a RAMSES sensor during irradiance (left) and radiance (right; C8, C12, C13 blue sky; C17 water in cloud shadow; C23 sunlit water) casts selected for intercomparison analysis.

The casts used in the intercomparison are listed in Table 4. Four casts (C10, C12, C13, and C14) were chosen for irradiance intercomparison, all recorded with direct sunlight. Five casts were chosen for radiance intercomparison: three casts (C8, C12, and C13) recorded with blue sky as a target, one (C17) measurement of water surface in cloud shadow, and one (C23) measurement of sunlit water. The 550 nm time series of one irradiance (RAMSES SAM_8329) and one radiance (RAMSES SAM_81B0) sensor for all the radiance and irradiance casts used for intercomparison are plotted in Figure 19. The initial cast start and stop times were adjusted based on Figure 19 to exclude the intervals with high temporal variability. Photographs of the radiance targets can be seen in Figure 20. Approximate FOV footprints for WISP-3 (3°), RAMSES (7°), and HyperOCR (23°) are shown in Figure 20 as well. The images were taken with a handheld Nikon D40X DSLR camera equipped with a Nikkor 18-200 mm zoom lens. According to the EXIF metainfo of the images the lens was completely zoomed out to 18 mm for C8, C12, C13, and C23. Considering the parameters of the lens and the camera, the horizontal FOV of these images is 67°. The lens was zoomed to 32 mm for C17 which corresponds to 41° horizontal FOV of the image. As the camera was not fixed to the frame along with the radiometers, exact co-alignment of the camera and the radiometers is not certain and the actual FOV footprints of the radiometers may have been slightly different than in Figure 20.

Table 4. Casts used in the intercomparison analysis.

Cast	Target	Time (UTC)	SZA	SAA	Relative VAA from Sun	VZA	Wind speed
C8	L_d (blue sky)	07:46:00–07:49:25	48°	131°	162°	43°	NA
C10	E_d	08:07:00-08:12:00	46°	137°	NA	NA	NA
C12	E_d, L_d (blue sky)	08:50:00-08:55:00	43°	151°	90°	43°	2.5 m s^{-1}
C13	E_d, L_d (blue sky)	09:00:00-09:03:05	42°	154°	134°	58°	NA
C14	E_d	09:22:30-09:47:30	41°	162°	NA	NA	2 m s^{-1}
C17	L_u (shadow)	10:30:00-10:35:00	40°	187°	107°	139°	2 m s^{-1}
C23	L_u (sunlit)	11:56:00-12:01:00	44°	217°	143°	130°	1 m s^{-1}

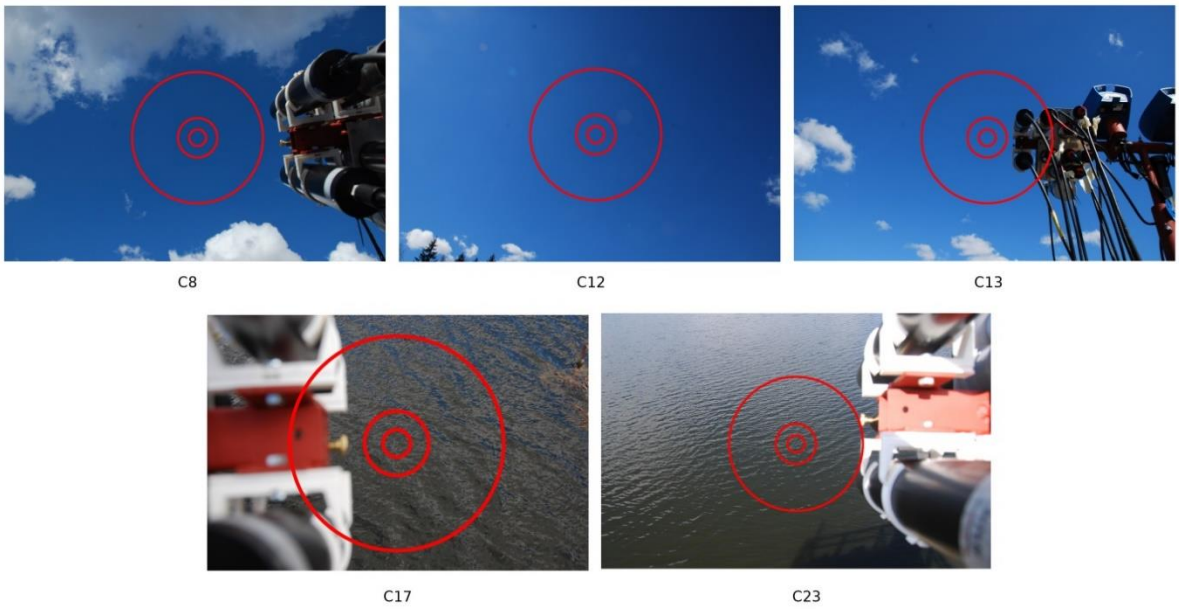


Figure 20. Photographs of radiance targets used in the intercomparison analysis. The circles denote approximate FOV of WISP-3 (smallest), RAMSES, and HyperOCR (largest).

6.3 Results

The participants were notified of the selected casts and expected to report back the calibrated spectra corresponding to each cast, including timestamps, with no averaging. Evaluation of the uncertainty for the outdoor results was not obligatory. As in total, data for 40 out of 44 radiometers were reported back to the pilot. For the rest, pilot carried out the data handling based on the provided raw files.

6.4 Data processing

The data processing details were described in Chapter 5.3. The outdoor data processing chain contained the following steps:

- Separation of the raw datafiles based on the casts' start and stop timestamps;
- dark signal subtraction;
- division by radiometric responsivity;
- recalculation for the OLCI spectral bands.

6.5 Consensus and reference values used for the analysis

Group median was used as the consensus value, with few outliers removed. Compared to the indoor measurements, outdoor variability between radiance sensors on average was about two times larger, and for irradiance sensors more than five times larger. Two irradiance and one radiance sensor were not accounted for in the variability estimate, because they had extremely large deviations from the group median.

The rather different behaviour of RAMSES and HyperOCR sensor groups should be noted. For the irradiance measurements, the deviation of HyperOCR sensors from consensus value of the group was very small, and for the increase of mean variability the group of RAMSES sensors was responsible, see Figure 21. At the same time, for the RAMSES group, variability

of the radiance sensors during indoor and outdoor exercises were almost at the same level, and the increase of the outdoor variability was caused largely by the HyperOCR sensors, see Figure 22.

All the irradiance casts in Figure 21 were measured with direct sunshine and no big difference between casts can be observed. The group of HyperOCR sensors is shown with dotted lines, These sensors are more consistent (less relative variability) with the consensus mean than the RAMSES sensors (higher variability across sensors).

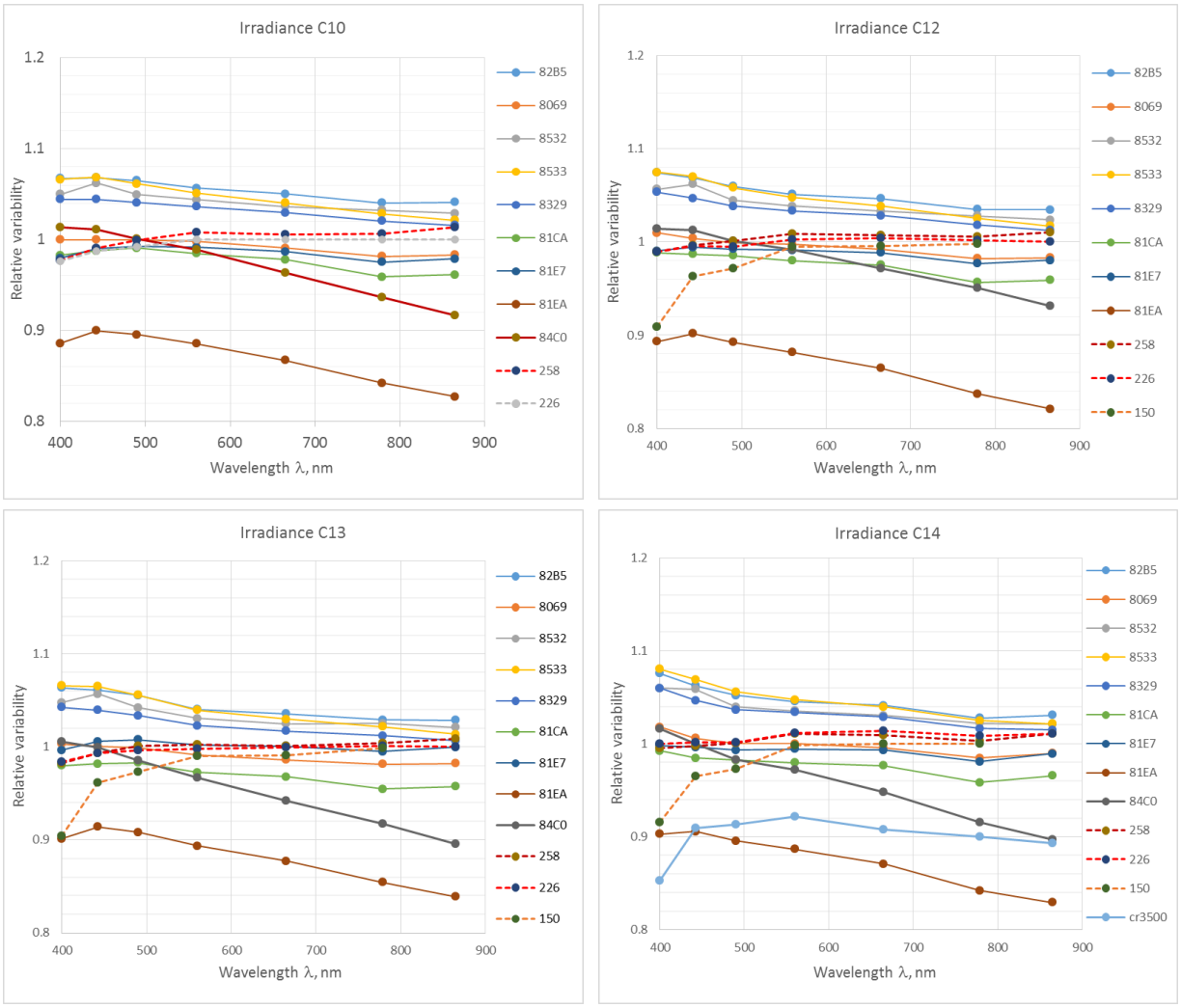


Figure 21. Irradiance sensors; agreement with the consensus value. Dotted lines - HyperOCR sensors.

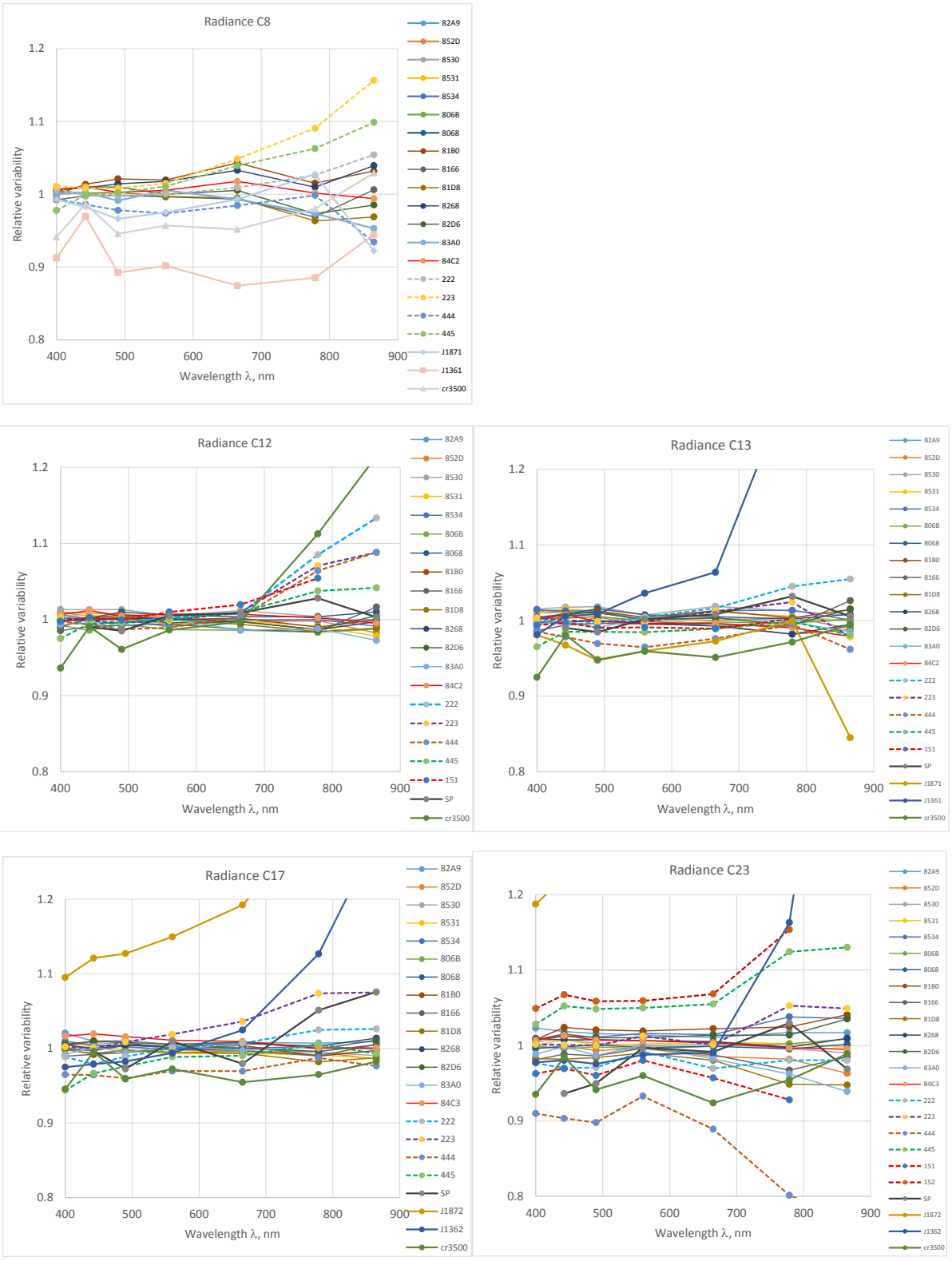


Figure 22. Radiance sensors; *agreement* with the consensus value in the outdoor experiment. *C8, C12, C13 - blue sky; C17 – water in cloud shadow; C23 – sunlit water. Dotted lines - HyperOCR sensors.*

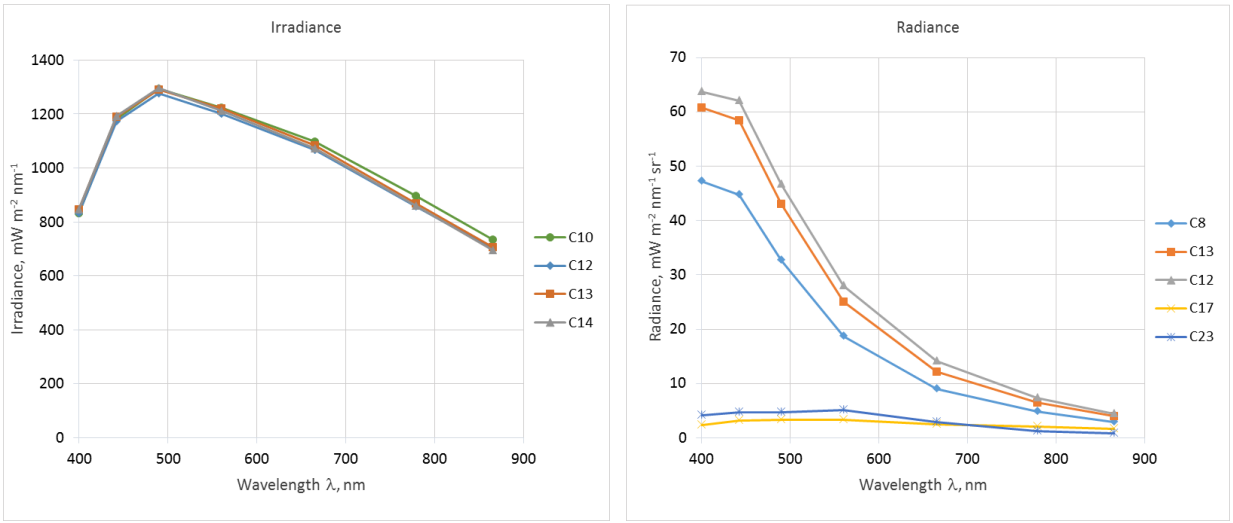


Figure 23. Irradiance and radiance consensus values in the outdoor experiment.

The difference between the casts of radiance sensors measuring the sky and water is evident in Figure 23. Radiation from water with blue sky gave the smallest signal, and consequently, the largest relative variability.

The variability between HyperOCR radiance sensors was unexpectedly large when measuring sunlit water (C23 in Figure 22). This is probably caused by misalignment of the sensors (Figure 17) and nonuniformity of the target (C23 in Figure 20). This assumption is supported by the fact that radiometers 151, 222, and 444 which are below the consensus mean in Figure 22 were mounted on the left frame and radiometers 152, 223, and 445 which all remain above the consensus mean in Figure 22 were mounted on the right frame.

7 Uncertainty analysis

The uncertainty analysis has been carried out according to the ISO Guide to the Expression of Uncertainty in Measurement [3], and to the EA guide EA 4/02. Evaluation is based on the measurement model, which describes the output quantity y as a function f of input quantities x_i: $y = f(x_1, x_2, x_3 \ldots)$. For every input quantity standard uncertainty is evaluated separately. There are two types of standard uncertainties: Type A is of statistical origin; Type B is determined by other means. Both types of uncertainties are indicated as standard deviation, denoted correspondingly by s and u. Measuring radiation sources using array spectrometers, the uncertainty contributions arising from averaging of a large number of repeatedly measured spectra is considered as Type A. Contributions from calibration certificates (lamp, current shunt, multimeter, diffuse reflectance panel etc.), but also from instability and spatial non-uniformity of the radiation sources are considered of Type B. For input quantities relative standard uncertainties are estimated. The relative combined standard uncertainty of output quantity is calculated by combining relative standard uncertainty of each input estimate by using formula (12) of [3]. Uncertainty of final result is given as relative expanded uncertainty with a coverage factor $k = 2$.

7.1 Effects causing variability of the results

Factors, causing the variation in the calibration and field results were listed in [2]. During the calibration and indoor measurements, uncertainty of the source is the dominant factor in the uncertainty budget, assuming the ambient temperature is known and stable within ±1 °C. Based on the experience from LCE-2 and the following FICE activities, differences between the calibration sources can reach ±2 % in the wavelength range of (350...900) nm. Nevertheless, in some cases sharp changes in the responsivity of radiometers were detected, clearly exceeding the expected contributions from the calibration source, alignment, contamination, temperature effects etc. (Figure 24).

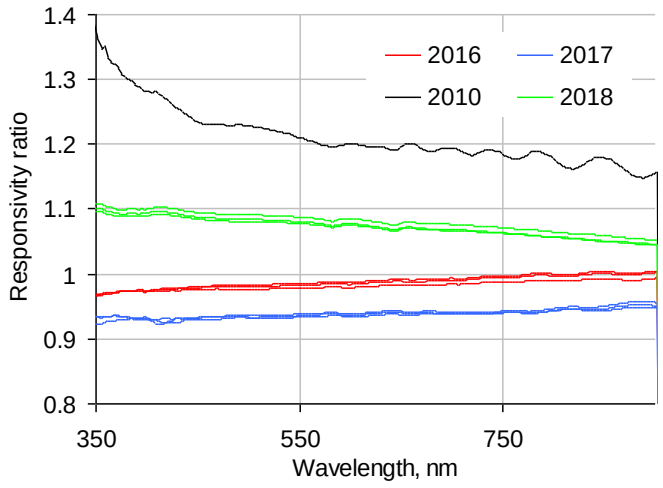


Figure 24. Changes in the spectral irradiance responsivity of the TriOS RAMSES SAM_8329.

FOV of the participating radiometers was not characterized during the FRM4SOC project. Uncertainty contributions for the calibration and indoor results were added to the budgets, based on the measurement geometry and class-specific parameters of the instruments [2]. For the outdoor exercise, application of the FOV contribution is complicated because of the high environmental variability during the measurements. [7] can be used to estimate the

influence of non-ideal cosine response of the irradiance sensors in the case of clear sky conditions.

Variation of the calibration coefficients as a function of temperature was not determined. Temperature effects were evaluated based on [8], see Figure 25.

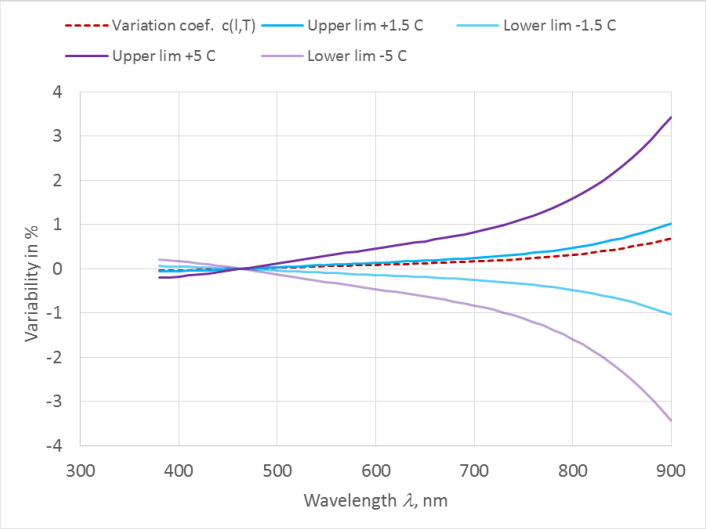


Figure 25. Relative variability of calibration coefficients due to temperature deviations from the reference temperature 21.5 °C.

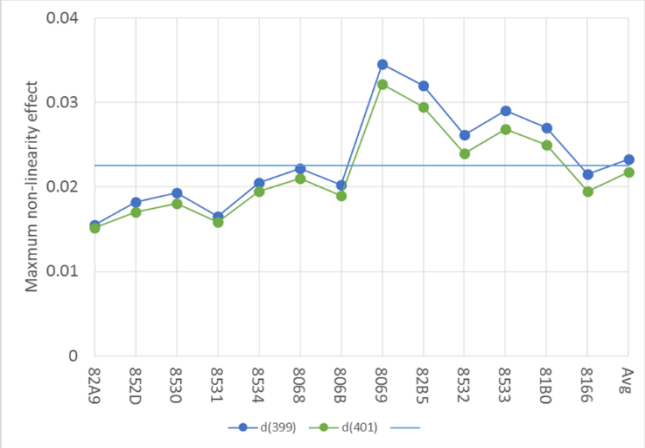


Figure 26. Maximum nonlinearity effect determined from calibration spectra using NPL lamps 399 and 401.

Maximum value of nonlinearity effect due to integration times determined from calibration spectra of TriOS RAMSES radiometers was in the range of (1.5...4) %, see Figure 26. Variability between the instruments due to this effect, if not corrected, will mostly be in the range of ±1 %. For certain types of radiometers, the effect can be corrected to 0.1 % by using the special formula, provided that for the same source at least two spectra with different integration times are available. Linearity corrected raw spectrum $S_{1,2}(\lambda)$ is calculated as

$$S_{1,2}(\lambda) = \left[1 - \left(\frac{S_2(\lambda)}{S_1(\lambda)} - 1\right)\left(\frac{1}{t_2/t_1 - 1}\right)\right] S_1(\lambda).$$

Here $S_1(\lambda)$ and $S_2(\lambda)$ are the initial spectra measured with integration times t_1 and t_2. Minimal ratio is usually t_2/t_1 =2, but may be also 4, 8, 16, etc. For large ratios $t_2/t_1 > 8$ the spectrum

$S_1(\lambda)$ is close to corrected spectrum $S_{1,2}(\lambda)$ and application of nonlinearity correction is not needed. The formula has been thoroughly tested for TriOS RAMSES and Satlantic HyperOCR radiometers, and in the range of 400 nm to 800 nm performs quite effectively. Unfortunately, it cannot be applied for outdoor measurements, as variability of natural radiation invalidates measurements using different integration times.

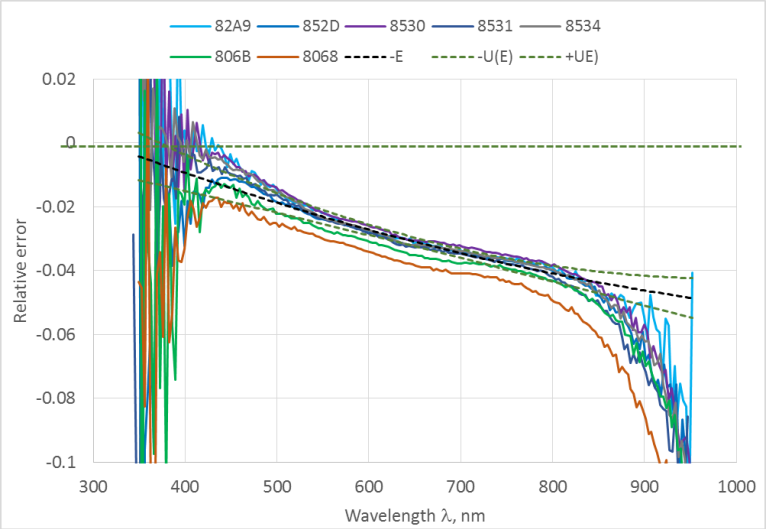


Figure 27. Non-linearity errors of different radiance sensors scaled to full-range value. Dashed lines are fitted model with uncertainty.

From the analysis of the calibration data it became evident that non-linearity errors scaled to full-range value of different radiance sensors behave in similar way. This behaviour serves as a basis for nonlinearity correction applicable to single spectrum that can also be used for the outdoor measurements, see Figure 27.

Relative nonlinearity correction $\delta x_{\max}(\lambda)$:

$$\delta x_{\max}(\lambda) = -5{,}1 \cdot 10^{-8}\lambda^2 + 0{,}00014 \cdot \lambda - 0{,}0355.$$

Relative nonlinearity correction $\delta x(\lambda)$ for the signal $x(\lambda)$:

$$\delta x(\lambda) = \frac{x}{x_{\max}} \delta x_{\max}(\lambda).$$

Corrected signal $x_{\mathrm{cor}}(\lambda)$ can be expressed as

$$x_{\mathrm{cor}}(\lambda) = x(\lambda)\left[1 + \frac{x}{x_{\max}} \delta x_{\max}(\lambda)\right].$$

The formula has been thoroughly tested on the TriOS RAMSES calibration data, and is effective in the range of 400 nm to 800 nm, correcting nonlinearity mostly better than to 0.2 %. The model can be fitted to all the studied RAMSES instruments by adjusting only the constant term.

The spectral stray light matrix was known from previous characterisation for the RAMSES sensors of TO and HyperOCR sensors of PML. Figure 28 presents the impact of stray light correction for these sensors, evaluated both for indoor and outdoor measurements. The indoor radiance and irradiance sources were spectrally similar to the calibration sources, therefore, the stray light correction has relatively small impact. The stray light effect in outdoor

measurements is an order of magnitude stronger due to significantly different spectral shape of the target and calibration signals. General impact of the stray light correction is similar for RAMSES and HyperOCR radiometers, but the variability between sensors, but also between measurements of different targets increases significantly in the NIR spectral region. This is probably related to the uncertainty associated with the stray light correction procedure and is not characteristic to the actual impact of spectral stray light. The spectral stray light matrices of HyperOCR sensors used in the analysis have higher noise level compared to the matrices of the RAMSES instruments. WISP-3, SR-3500, and SeaPRISM have different optical design, thus, their spectral stray light properties can have different nature compared to the data presented in Figure 28.

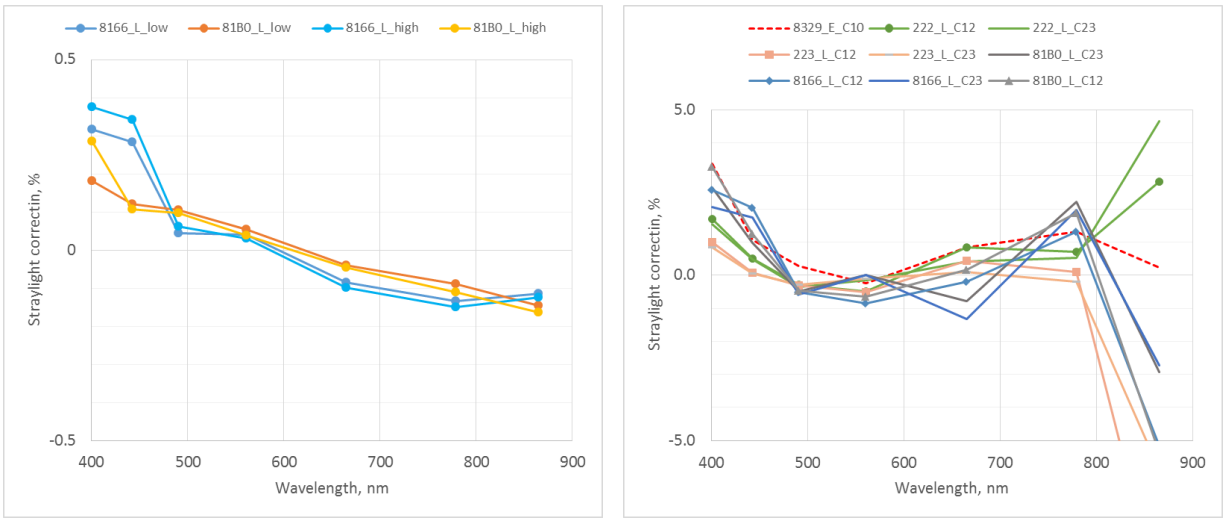


Figure 28. Stray light effects estimated for some RAMSES and HyperOCR sensors. Left – indoor radiance measuements with two RAMSES sensors at high and low sphere radiance; right – one irradiance sensor (RAMSES 8329, dashed line) and four radiance sensors (two RAMSES, two HyperOCR) in outdoor experiment, solid lines with symbols – blue sky (C12), solid lines without symbols – sunlit water (C23).

Figure 29 shows the change of the OLCI band values of a spectrum as a function of the wavelength scale error of a radiometer, determined for a single RAMSES radiance sensor for the casts C8, C12, C17, and C23. Precision of the wavelength scale of the MMS-1 module stated by manufacturer is 0.3 nm. For ±0.3 nm shift of the scale, the changes of the OLCI band values for the different spectra remain less than ±0.5 % except for the 400 nm spectral band where the radiance changes rapidly with wavelength and the effect of shifting the wavelength scale is stronger.

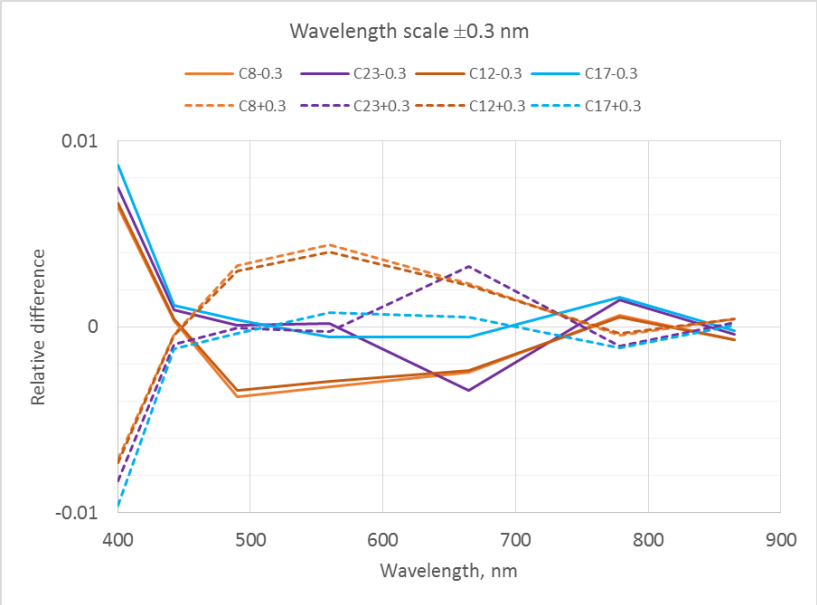


Figure 29. Relative variability due to wavelength error of ±0.3 nm of a radiance sensor.

For calibration of the radiometers during LCE-2, different types of purpose-built and proprietary software were used for the instrument control and acquisition. The calibration data were presented in the form suitable for the specific software used for the sensor. Due to differences in the sensor design and software, procedure of later data handling cannot be strictly the same for all radiometer types. Dark signal is estimated using optically masked pixels in the RAMSES and WISP-3 instruments whereas internal mechanical shutter is used in HyperOCR, SR-3500, and SeaPRISM. All the participating instruments were hyperspectral radiometers except the multispectral SeaPRISM which already had a filter set matched to the OLCI bands and no weighted binning of spectral bands was necessary. The proprietary data processing workflow inside WISP-3 includes probably dark subtraction, nonlinearity correction, spectral and temporal averaging, but exact algorithm is not known to the pilot. The SR-3500 was operated in calibrated radiance and irradiance modes, therefore, dark subtraction, temporal averaging, and conversion from raw DN to radiance and irradiance units using factory calibration coefficients were done by the proprietary software. The radiometric calibration coefficients of SR-3500 are stored inside the instrument and cannot be updated by the user. Thus, correction factors were provided for SR-3500 to apply the radiometric calibration done at TO before the start of LCE-2. The SeaPRISM data were downloaded in proprietary binary format K8 which pilot was not able to decode. The data was decoded by the participant and returned to the pilot as dark-corrected digital counts for two different electronic gain modes. Data recorded in high gain mode were used in the intercomparison analysis.

7.2 Uncertainty budgets for comparison tasks

Uncertainty analysis is made for two indoor and three outdoor measurement tasks with the aim to see which contributions will explain the variability revealed between sensors during the intercomparison. In laboratory, uncertainty estimates for irradiance sensors measuring an FEL source at approximately 1 m distance are given in Table 5, and for radiance sensors measuring integrating sphere, in Table 6. The outdoor solar irradiance uncertainty estimates are presented in Table 7, Table 8 corresponds to the blue sky radiance and Table 9 to the radiance of sunlit water. All the uncertainty estimations in Table 5 - Table 9 are based on

experimental variability data of TriOS RAMSES sensors and information from [9]–[15]. For the other radiometer models that took part in the intercomparison very little publicly available information can be found regarding various instrument characteristics that influence the measurement results. In addition, the RAMSES was the only sensor model that was represented in sufficiently large number for statistical analysis.

The uncertainty is calculated from the contributions originating from the spectral responsivity of the radiometer, including data from the calibration certificate, from interpolation of the spectral responsivity values to the designated wavelengths and/or spectral bands, from instability of the array spectroradiometer, from contribution to the spectral irradiance and/or radiance due to setting and measurement of lamp current of the lamps, from measurement of the distance between the lamp and input aperture of the radiometer, from the spatial uniformity of the irradiance at 1 m distance, and from reproducibility of the alignment. For the radiometer, uncertainty contributions arising from the non-linearity, temperature effects, spectral stray light, and from dark measurements, from repeatability and reproducibility of averaged signal are included.

The radiometric calibration uncertainty in the following tables is excluded from the combined uncertainty because all participating radiometers were calibrated using common standards shortly before the intercomparison. While the calibration uncertainty is relevant for traceability to SI units, the uncertainty budgets in Table 5 - Table 9 describe variability between the sensors and systematic effects which can influence all the instruments in similar way are not accounted for.

Table 5. Typical relative uncertainty budget (in percent) for the variability between irradiance sensors during indoor comparison, contribution of radiometer calibration certificate omitted.

		400 nm	442.5 nm	490 nm	560 nm	665 nm	778.8 nm	865 nm
Radiometer	Certificate	0.88	0.68	0.65	0.62	0.59	0.62	0.56
	Interpolation	0.5	0.2	0.3	0.2	0.2	0.1	0.1
	Instability	0.05	0.03	0.04	0.03	0.04	0.03	0.02
	Alignment	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	Nonlinearity	0.2	0.15	0.15	0.15	0.15	0.15	0.2
	Stray light	0.2	0.2	0.2	0.2	0.2	0.2	0.2
	Temperature	0.02	0.01	0.01	0.03	0.09	0.2	0.38
	Signal, uA	0.11	0.04	0.02	0.02	0.01	0.02	0.04
Source	Instability	0.14	0.14	0.12	0.11	0.1	0.09	0.08
	Uniformity	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	Stray light	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	Combined (k=1)	0.63	0.39	0.45	0.38	0.39	0.39	0.52
	Expanded (k=2)	1.3	0.8	0.9	0.8	0.8	0.8	1.0

Table 6. Typical relative uncertainty budget (in percent) for the variability between radiance sensors during indoor comparison, contribution of radiometer calibration certificate omitted.

		400 nm	442.5 nm	490 nm	560 nm	665 nm	778.8 nm	865 nm
Radiometer	Certificate	1.2	0.78	0.76	0.73	0.71	0.73	1.35
	Interpolation	0.5	0.2	0.3	0.2	0.2	0.1	0.1
	Instability	0.04	0.03	0.02	0.01	0.01	0.02	0.01
	Back-reflection	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	Alignment	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	Nonlinearity	0.2	0.15	0.15	0.15	0.15	0.15	0.2
	Stray light	0.2	0.2	0.2	0.2	0.2	0.2	0.2
	Temperature	0.02	0.01	0.01	0.03	0.09	0.2	0.38
	Signal, uA	0.12	0.07	0.04	0.02	0.03	0.03	0.06
Source	Instability	0.14	0.14	0.12	0.11	0.1	0.09	0.08
	Uniformity	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	Stray light	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	Combined (k=1)	0.64	0.41	0.46	0.39	0.40	0.40	0.53
	Expanded (k=2)	1.3	0.8	0.9	0.8	0.8	0.8	1.1

Table 7. Typical relative uncertainty budget (in percent) for the variability between irradiance sensors during outdoor comparison, contribution of radiometer calibration certificate omitted.

		400 nm	442.5 nm	490 nm	560 nm	665 nm	778.8 nm	865 nm
Radiometer	Certificate	0.88	0.68	0.65	0.62	0.59	0.62	0.56
	Interpolation	0.5	0.2	0.3	0.2	0.2	0.1	0.1
	Instability	0.05	0.03	0.04	0.03	0.04	0.03	0.02
	Polarisation	0.1	0.1	0.1	0.1	0.1	0.1	0.2
	Alignment	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	Nonlinearity	0.4	0.3	0.3	0.3	0.3	0.3	0.2
	Stray light	0.9	0.7	0.3	0.3	0.7	0.9	1.0
	Temperature	0.4	0.2	0.2	0.2	0.2	0.4	0.8
	Cosine error	2.5	2	1.5	1	1	1	0.5
	Signal, uA	0.01	0.01	0.01	0.01	0.01	0.02	0.02
	Combined (k=1)	2.77	2.16	1.61	1.13	1.30	1.45	1.41
	Expanded (k=2)	5.5	4.3	3.2	2.3	2.6	2.9	2.8

Table 8. Typical relative uncertainty budget (in percent) for the variability of radiance sensors during outdoor comparison measuring the blue sky, contribution of radiometer calibration certificate omitted.

		400 nm	442.5 nm	490 nm	560 nm	665 nm	778.8 nm	865 nm
Radiometer	Certificate	1.2	0.78	0.76	0.73	0.71	0.73	1.35
	Interpolation	0.5	0.1	0.3	0.3	0.2	0.1	0.1
	Instability	0.04	0.03	0.02	0.01	0.01	0.02	0.01
	Polarisation	0.1	0.1	0.2	0.2	0.4	0.4	0.4
	Nonlinearity	0.1	0.1	0.1	0.1	0.15	0.15	0.2
	Stray light	0.9	0.7	0.3	0.3	0.7	0.9	1.0
	Temperature	0.1	0.1	0.1	0.1	0.1	0.2	0.4
	Signal, uA	0.01	0.01	0.01	0.01	0.02	0.11	0.20
	Alignment and target uniformity	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	Combined (k=1)	1.05	0.74	0.50	0.50	0.86	1.03	1.19
	Expanded (k=2)	2.1	1.5	1.0	1.0	1.7	2.1	2.4

Table 9. Typical relative uncertainty budget (in percent) for the variability of radiance sensors during outdoor comparison measuring the water radiance with blue sky, contribution of radiometer calibration certificate omitted.

		400 nm	442.5 nm	490 nm	560 nm	665 nm	778.8 nm	865 nm
Radiometer	Certificate	1.2	0.78	0.76	0.73	0.71	0.73	1.35
	Interpolation	0.6	0.2	0.1	0.1	0.3	0.2	0.1
	Instability	0.04	0.03	0.02	0.01	0.01	0.02	0.01
	Polarisation	0.1	0.1	0.1	0.1	0.2	0.2	0.2
	Nonlinearity	0.2	0.2	0.3	0.3	0.3	0.3	0.3
	Stray light	0.9	0.7	0.3	0.3	0.7	0.9	1.0
	Temperature	0.2	0.2	0.2	0.2	0.2	0.4	0.9
	Signal, uA	0.04	0.07	0.11	0.11	0.21	0.55	0.72
	Alignment and target uniformity	0.3	0.3	0.3	0.3	0.4	0.8	1.1
	Combined (k=1)	1.16	0.85	0.59	0.58	0.98	1.44	1.92
	Expanded (k=2)	2.3	1.7	1.2	1.2	2.0	2.9	3.8

For RAMSES group, variability of radiance sensors during indoor (Figure 10 - Figure 11) and outdoor exercises (Figure 22, except C8 and C23) was very close. Therefore, variability due to influence factors like temperature, nonlinearity, and stray light, and respective estimates used in uncertainty budgets, can be considered practically the same. For example, during outdoor measurements, temperature was rather stable varying from 5 °C to 9 °C, a range fairly comparable with variation of temperature during indoor exercise (21 °C to 24 °C).

As the construction of radiance and irradiance sensors (except the input optics) is similar, the outdoor estimates are likely suitable also for variability due to temperature, nonlinearity, and stray light between irradiance sensors. Major differences in combined uncertainty estimates

for outdoor measurements are caused by different FOV of the sensors (including deviation from cosine response for irradiance instruments) and due to instability and nonuniformity of the scene.

The uncertainty components in Table 5 - Table 9 are explained below.

7.2.1 Radiometer calibration certificate

Calibration certificate provides calibration points following the individual wavelength scale of the radiometer. During relatively short time needed for intercomparison, this uncertainty component normally is not contributing to the variability between freshly calibrated radiometric sensors. At the same time, this component is highly relevant for traceability of results to the SI.

7.2.2 Interpolation of radiometer data

Due to differences between spectral response functions, direct comparison between sensors is impossible. Therefore, measured values were transferred for comparison to a common scale basis (OLCI bands), see Chapter 5.3.6. The uncertainty contribution associated with interpolation of spectra is estimated from calculations using different interpolation algorithms. The weights used for binning hyperspectral data to OLCI bands depend on the wavelength scale and exact pixel positions of the hyperspectral sensor. The effect of wavelength scale uncertainty on calculating OLCI bands is estimated from data presented in Figure 29 (Chapter 7.1). The interpolation components in Table 5 - Table 9 include interpolation as well as wavelength scale uncertainty contributions.

7.2.3 Radiometer instability

The radiometric responsivity of the instruments can change over time, which can be observed from time series of repeated radiometric calibrations. For LCE-2 the instruments were calibrated just before the intercomparison exercise and only uncertainty component related to short-term instability relevant for the time needed for intercomparison measurements has to be considered. The values are derived from the data collected in calibration sessions of LCE-2 and FICE-AAOT, see Chapter 4.3 and Figure 6. The variability over two weeks was interpolated from the yearly variability data. In addition to instability of the sensors the data in Figure 6 includes other uncertainty components related to the calibration setup (e.g. alignment, short-term lamp instability, etc). Overall, the radiometer instability component in LCE-2 intercomparison is small compared to some other contributions and the data presented in Chapter 4.3 is currently the best available estimate for this uncertainty component.

7.2.4 Back-reflection

Radiation reflecting from the radiance sensor back into the integrating sphere, was estimated using different distances between the sphere and the radiometer.

7.2.5 Polarisation

The indoor irradiance measurements were repeated with the radiometer rotated 90° around the main axis to get some information about the combined effect of alignment and polarisation (Chapter 5.1). According to [16] the FEL emission is about 3 % polarised and [12] reports the polarisation sensitivity of RAMSES irradiance sensors increasing from (0.05…0.3) % at 400 nm to (0.3…0.6) % at 750 nm which is less than polarisation sensitivity

of RAMSES radiance sensors due to the depolarising nature of the cosine diffuser. This allows to conclude that the observed differences at repeated measurements are mostly caused by other uncertainty components (alignment, short-term instability of source emission, etc.) and the polarisation component can be omitted from the indoor irradiance uncertainty budget. Likewise, the polarisation is not included in the indoor radiance uncertainty budget as the integrating sphere is a strong depolariser due to multiple internal reflections and the source radiance is assumed to be unpolarised.

For the outdoor measurements, the uncertainty contribution caused by polarisation sensitivity is estimated using worst-case data published in [12]. It is very difficult to estimate the polarisation effect for the outdoor irradiance measurement as the degree of linear polarisation (DoLP) depends on various factors such as wavelength, SZA, AOD, amount and location of clouds to name a few. In addition, the DoLP varies a lot over the hemisphere being largest at 90° from the Sun due to Rayleigh scattering and decreasing to zero for the direct solar flux. However, according to [12] the polarisation sensitivity of RAMSES irradiance sensors is rather small, hence, the contribution of polarisation effect in the uncertainty budget is also small regardless of the DoLP value of downwelling irradiance. For estimating the uncertainty component associated with polarisation in Table 7, $DoLP_{Ed}$ = 0.25 has been assumed.

7.2.6 Alignment

Distance between the source and the reference plane of the irradiance radiometer's cosine diffuser, measured along the optical axis (x-axis) is important alignment parameter. Alignment errors of the lamp source across optical axes less than ±1 mm in the y or z directions, rotation of the lamp around x and z axes less than ±0.1°, and around y axes (parallel with the filament of the lamp) less than 2° will cause uncertainty in the irradiance less than 0.1 % [17]. Positioning errors of the input optics of the radiometer lead to an additional irradiance uncertainty of about 0.1 %. This accuracy can be achieved only by very careful alignment by means of special adjustment laser.

7.2.7 Nonlinearity

For some hyperspectral radiometers, spectra measured at different integration times and converted to the same scale using respective gain factors show relative differences up to 4 %. According to recommendations, the non-linearity effects of good sensors should be correctable to less than 0.1 % (see Chapter 7.1 for details).

7.2.8 Spectral stray light

Spectral stray light of sensors is commonly not very relevant for measurements when the calibration and target source emissions have similar spectral composition. In the case of outdoor measurements, the effect may be rather significant, but variability may be reduced due to the correlated class-specific behaviour between the sensors. Value is estimated from Figure 28 and [18].

7.2.9 Temperature

For array spectroradiometers with silicon detectors, the present estimate for standard uncertainty due to temperature variability (±1.5 °C) in the spectral region from 400 nm to 700 nm is around 0.1 % and will increase up to 0.6 % for longer wavelengths (950 nm) [8].

7.2.10 Cosine error

The irradiance sensors are calibrated using normal illumination, but during outdoor solar irradiance measurements the radiation arriving from hemisphere has to be measured with the angular dependence of responsivity corresponding to cosine of incidence angle. Typical class-specific values of uncertainty related to deviation of cosine response are derived from [7].

7.2.11 Type A uncertainty of signal

Present estimate for standard uncertainty due to random effects is valid for indoor measurements, where typically white noise can be expected. Analysis of individual measurements in a measurement series has indicated that the measurements are not completely independent and the autocorrelation of time series has been taken into account. The correlation is likely caused by small changes in the lamp emission (see also next chapter). If there is autocorrelation in the time series, the effective number of independent measurements n_e has to be considered instead of actual number of samples n_t in the series [19]:

$$n_e \approx n_t \frac{1 - r_1}{1 + r_1},$$

where r_1 is the lag-1 autocorrelation of the time series.

The type A uncertainty is estimated differently for outdoor measurements due to relatively large variability of the target signal and synchronous measurement with all the instruments. While there is strong autocorrelation in individual time series due to the instable nature of natural illumination, the variability of source signal is the same for all the instruments. The autocorrelation in the ratio of two sensors is significantly smaller than the autocorrelation in the individual time series of both sensors. The type A uncertainty of signal in Table 7 - Table 9 was estimated from the ratio of two RAMSES radiometers. There was almost no correlation between individual ratios during one cast and the effective number of measurements was close to actual number of data points in the time series.

7.2.12 Source instability

The indoor measurements were not made simultaneously, thus, the short-term instability of the source, relevant for the time needed for intercomparison measurements, including power cycling the source between the two days of indoor experiment, has to be considered. This uncertainty component was estimated using the uncertainty in setting the lamp current and its effect on lamp emission.

The instability of the target signal during outdoor measurements was significantly larger, however, all the instruments measured simultaneously and the impact of source variability affected all the radiometers in similar manner without causing differences between the sensors. This was verified by separately analysing some shorter and more stable sections of the selected casts, no reduction of variability between the sensors was observed.

7.2.13 Source stray light

The uncertainty component associated with the stray light in the laboratory during the indoor experiment has been estimated in previous experiments made in the laboratory of Tartu Observatory.

7.2.14 Uniformity and/or alignment

Due to nonuniformity of FEL irradiance at the measurement plane, the differences in the cosine diffuser dimensions and positioning of the radiometers at the measurement plane contribute to the indoor irradiance measurement uncertainty. For indoor radiance measurement this uncertainty component is related to the nonuniformity of the sphere radiance. In the outdoor radiance experiment the target nonuniformity has much larger contribution to the uncertainty budget due to possible misalignment and differences in the FOV of the radiance sensors (Figure 20).

8 Conclusions

The LCE-2 exercise consisted of three sub-tasks: SI-traceable radiometric calibration of participating radiometers just before the intercomparison; indoor intercomparison of stable sources in controlled environment; outdoor intercomparison over terrestrial water surface. Additionally, major part of sensors involved in LCE-2 were recalibrated at TO a year later (for FICE-AAOT) revealing valuable information about their long-term stability. Most of the sensors (more than 80 %) changed less than ±1 % in one year.

Agreement between irradiance and radiance sensors is mostly affected by sensor calibration. For example, factory calibrations made at different times can cause differences exceeding ±10 %. Former calibrations in different labs and several years ago can cause differences around ±3 %. Different calculation schemes (corrections for non-linearity, stray light or for OLCI band values) can cause differences about ±1...2 % each factor. The best agreement of 0.5...0.8 % between participants has been achieved when a unified procedure for measurements and data handling are applied.

For outdoor measurements the variability between radiance sensors was about two times larger than during indoor exercise, this can be explained by larger effects of outside influence factors like temperature, stray light and nonlinearity which all have not been corrected.

Variability between irradiance sensors was about five times larger than during indoor exercise. Most likely the revealed unexpectedly large differences between sensors were caused by the rather high inter-instrument variability of cosine response described in [7].

Different behaviour of RAMSES and HyperOCR sensor groups was clearly revealed during LCE-2 exercise. For RAMSES group, variability of radiance sensors during indoor and outdoor exercises was very similar, and larger variability for outdoor measurements was mostly caused by HyperOCR and WISP-3 sensors. For irradiance measurements, the deviation of HyperOCR sensors from consensus value of the group was very small, and increase in variability was mostly caused by the group of RAMSES sensors.

Rather large variability between sensors during outdoor exercise cannot be explained by poor stability of sensors, as stability check in lab conditions a year later has showed much smaller changes than during outdoor measurements some days after calibration. Variability cannot be fully explained by influence factors like temperature, nonlinearity, and stray light either. Most likely, the different behaviour of RAMSES and HyperOCR sensors is largely due to different construction of input optics of these sensors.

In order to help in interpretation of the results, the following suggestions are proposed keeping in mind the future outdoor intercomparison campaigns:

- The instruments' temperatures should be logged whenever possible;
- during the responsivity calibration, using different ambient temperatures is recommended;
- acquisition of the data for all instruments should start synchronously within ±1 s and sampling interval should be the same, which makes it possible to compare the individual spectra instead of temporal averages;
- characterization of the angular response of the radiometers is important, especially in the case of variable sky conditions;

- irradiance measurements under the clear sky conditions covering large span of solar zenith angles are necessary to assess the uncertainties caused by the irradiance entrance optics;
- intercomparisons should be done in varying water optical property conditions;
- calibration history for each participating radiometer is vital in order to detect possible instrument misbehaviour and remove outliers;
- using a well-characterized reference instrument is highly recommended;
- using an aligned photo- or video camera to continuously record the measurement scene during outdoor experiments;
- the data processing algorithms should be well defined and agreed between the participants.

9 References

[1] D. Antoine *et al.*, "IMOS Radiometry Task Team," Final report, Jun. 2017.
[2] J. Kuusk, I. Ansko, V. Vabson, M. Ligi, and R. Vendt, "Protocols and Procedures to Verify the Performance of Fiducial Reference Measurement (FRM) Field Ocean Colour Radiometers (OCR) used for Satellite Validation," Tartu Observatory, Tõravere, Technical Report TR-5, Feb. 2017.
[3] "Guide to the Expression of Uncertainty in Measurement." International Organization for Standardization (ISO), Geneva-1995.
[4] "Spectral Response Function Data." [Online]. Available: https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-3-olci/olci-instrument/spectral-response-function-data. [Accessed: 03-Sep-2018].
[5] J. W. Müller, "Possible Advantages of a Robust Evaluation of Comparisons," *J. Res. Natl. Inst. Stand. Technol.*, vol. 105, no. 4, pp. 551–555, 2000.
[6] "Aerosol Robotic Network (AERONET) Homepage." [Online]. Available: https://aeronet.gsfc.nasa.gov/. [Accessed: 09-Aug-2018].
[7] S. Mekaoui and G. Zibordi, "Cosine error for a class of hyperspectral irradiance sensors," *Metrologia*, vol. 50, no. 3, p. 187, 2013.
[8] G. Zibordi, M. Talone, and L. Jankowski, "Response to Temperature of a Class of In Situ Hyperspectral Radiometers," *J. Atmospheric Ocean. Technol.*, vol. 34, no. 8, pp. 1795–1805, May 2017.
[9] G. Zibordi *et al.*, "In situ determination of the remote sensing reflectance: an inter-comparison," *Ocean Sci.*, vol. 8, pp. 567–586, 2012.
[10] G. Zibordi and K. J. Voss, "Chapter 3.1 - In situ Optical Radiometry in the Visible and Near Infrared," in *Optical Radiometry for Ocean Climate Measurements*, vol. 47, C. J. D. and A. C. P. Giuseppe Zibordi, Ed. Academic Press, 2014, pp. 247–304.
[11] M. Gergely and G. Zibordi, "Assessment of AERONET-OC L WN uncertainties," *Metrologia*, vol. 51, no. 1, p. 40, 2014.
[12] M. Talone and G. Zibordi, "Polarimetric characteristics of a class of hyperspectral radiometers," *Appl. Opt.*, vol. 55, no. 35, pp. 10092–10104, Dec. 2016.
[13] "IOCCG, International Network for Sensor Inter-comparison and Uncertainty Assessment for Ocean Color Radiometry (INSITU-OCR) White Paper, available from http://www.ioccg.org/groups/INSITU-OCR_White-Paper.pdf, 2012." [Online]. Available: http://www.ioccg.org/groups/INSITU-OCR_White-Paper.pdf. [Accessed: 07-Feb-2017].
[14] K. G. Ruddick, V. De Cauwer, Y.-J. Park, and G. Moore, "Seaborne measurements of near infrared water-leaving reflectance: The similarity spectrum for turbid waters," *Limnol. Oceanogr.*, vol. 51, no. 2, pp. 1167–1179, Mar. 2006.
[15] K. Alikas *et al.*, "Validation of Sentinel-3A/OLCI data over Estonian inland waters," presented at the AMT4 Sentinel FRM workshop, Plymouth, United Kingdom, 2017.
[16] K. J. Voss and L. B. da Costa, "Polarization properties of FEL lamps as applied to radiometric calibration," *Appl. Opt.*, vol. 55, no. 31, pp. 8829–8832, Nov. 2016.
[17] G. Bernhard and G. Seckmeyer, "Uncertainty of measurements of spectral solar UV irradiance," *J. Geophys. Res. Atmospheres*, vol. 104, no. D12, pp. 14321–14345, Jun. 1999.
[18] M. Talone, G. Zibordi, I. Ansko, A. C. Banks, and J. Kuusk, "Stray light effects in above-water remote-sensing reflectance from hyperspectral radiometers," *Appl. Opt.*, vol. 55, no. 15, pp. 3966–3977, May 2016.
[19] B. D. Santer *et al.*, "Statistical significance of trends and trend differences in layer-average atmospheric temperature time series," *J. Geophys. Res. Atmospheres*, vol. 105, no. D6, pp. 7337–7356, Mar. 2000.

Appendix A User feedback

Feedback for the FRM4SOC project Laboratory Comparison Exercise LCE-2

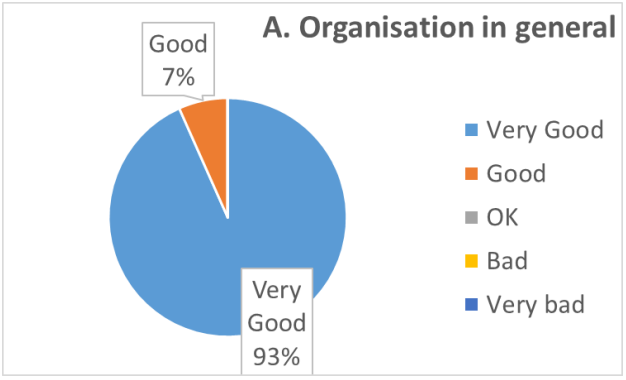
8 – 13 May 2017 in Estonia

1. **How was this event useful for you?**

- Improve my knowledge (Science, Technology). Meet excellent people.
- To create a good Network. To evaluate the accuracy of the instruments. To improve the measurement protocol.
- Interesting to find out how people use those radiometers and how they take uncertainties into account.
- Talk to customers/users. See sensors used in field. Learn best methods of measurements. Learn about our own sensors.
- It was very interesting; learn the calibration procedures in the lab environment. Meet the group from Tartu and the other countries. Great opportunity. The usefulness will be defined depending on the outcomes.
- It was interesting to observe how laboratory comparison exercise was conducted in your field of research.

2. **Please give us feedback about the organisation of the event.**

A. Organisation in general

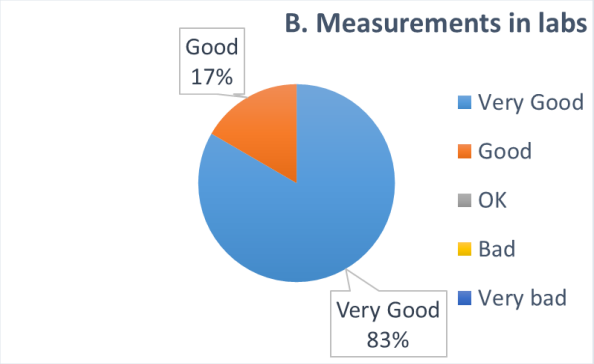


Your comments:

- Everything clear and smooth.
- You have a very good equipment and experience
- It is very difficult to organise such a complex activity with so many participants but it was very well done.

- Fantastic. An easy event to attend. Friendly, helpful staff always willing to help.
- Very interesting, especially discussions around lab cal.
- Excellent team, well organised, very welcoming.
- All very organized; the first day could be more practical with more details about the activities (methods) to achieve the goal.

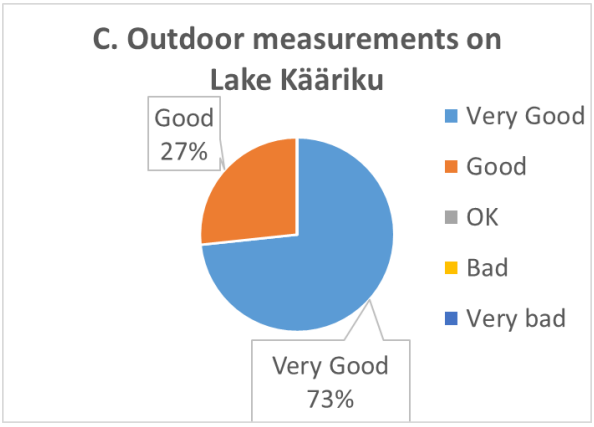
B. Measurements in labs



Your comments:

- Excellent experts assistance
- Carefully thought out procedures
- Sensors settings were changed without notice. It took a long time to debug. Quality of measurements is high, though, I think.
- Really interesting. I missed some information about the methods and expected results; a simulator of results could be useful.

C. Outdoor measurements on Lake Kääriku



Your comments:

- Excellent facilities and support
- Took a long time to set up each morning. Unlucky with weather
- It would have been better to have more time for better chance of clear sky. But, Tartu staff worked very hard to make this run smooth – and Fun
- Need 5 days instead of 2, but 2 better than 0
- Really interesting. I missed a connector between the methods and what are the expected results. This would help on the analyses
- I really enjoyed all the organisation in general

3. Suggestion for improvement in the future (organisational, technical, etc):

- None
- Choose much different typology of instruments
- Perhaps more days in the field to allow contingency days for the weather
- More time. First time experiments (lab or field) always take a long time
- Clarify post-experiment data processing and deadlines before end of week
- Organisational – it was very good. I would recommend pair the Tartu grad students with the researchers (one student per group) so the students have more integration and learn in a more effective way
- More detailed list of the field work plan

4. How to benefit from this kind of comparison exercise in the future:

- The exercise might be used as prime example for similar activities
- Keep doing them!
- It is extremely important; I would recommend that discussion about the next steps, for instance, data analyses, publication, happens during the organization of the event
- It is important to organize and participate these comparison exercise regular basis