D-290: FRM4SOC Final Report

Title	FRM4SOC Final Report
Document reference	FRM4SOC-FR
Project	ESA – FRM4SOC
Contract	ESRIN/Contract No. 4000117454/16/1-SBo
Deliverable	D-290
ATTN	Tânia G. D. Casal ESA/ESTEC Technical Officer Keplerlaan 1 2201 AZ Noordwijk The Netherlands
Version	1
Date issued	30.06.2020

	Signed By	**Approved by**
Name:	Riho Vendt	Tânia G. D. Casal
Organisation:	Tartu Observatory, University of Tartu	ESA/ESTEC
Position:	Project manager	Technical Officer
Date:		
Signature:		

Responsibility

Statements made in this report are the responsibility of the lead author and the FRM4SOC Project Team and do not necessarily represent the official views of the European Space Agency or the organisations of any of the scientists who contributed to this report. The FRM4SOC project is funded by the European Space Agency.

Authors

R. Vendt, V. Vabson, J. Kuusk, K. Alikas, I. Ansko, M. Ligi, A. Noorma, A. Bialek, A. C. Banks, N. Fox, C. Lerebourg, K. Ruddick, G. Tilstone, T. Casal, C. Donlon.

Acknowledgements

Input and comments from numerous scientists and experts are gratefully acknowledged. Especially G. Zibordi and D. Antoine, the FRM4SOC team, all co-authors of the FRM4SOC papers published in the special issue of the Multidisciplinary Digital Publishing Institute (MDPI) Journal *Remote Sensing*, all participants of the project events, and manufacturers of the ocean colour radiometers are kindly thanked for the valuable support and contribution to the project outcomes and the current document.

Citations

Papers of the Special Issue of the MDPI Journal *Remote Sensing* "Fiducial Reference Measurements for Satellite Ocean Colour" should be considered as the primary reference for citation of the study, where such reference to the related paper is provided in the beginning of relevant chapters.

1 Document Control Table

Title	FRM4SOC Final Report
Document reference	FRM4SOC-FR
Project	ESA – FRM4SOC
Contract	ESRIN/Contract No. 4000117454/16/1-SBo
Deliverable	D-290
Version	1
Date Issued	30.06.2020

2 Document Change Record

Index	Issue	Revision	Date	Brief description	Issued by
1.	0	1	14.04.2020	Initial draft.	Riho Vendt
2.	0.	2	24.04.2020	TO internal review	Riho Vendt
3.	0.	3.	06.05.2020	FRM4SOC partner review	Riho Vendt
4.	0	4.	10.06.2020	Final Draft	Riho Vendt
5.	1.	0.	30.06.2020	First version	Riho Vendt

3 Distribution List

Company/Organisation	Name	Format	No. of Copies
Tartu Observatory, UT	Riho Vendt	Electronic file – word	1 docx
NPL	Agnieszka Bialek	Electronic file – word	1 docx
NPL	Nigel Fox	Electronic file – word	1 docx
-	Andrew Banks	Electronic file – word	1 docx
ACRI-ST	Christophe Lerebourg	Electronic file – word	1 docx
RBINS	Kevin Ruddick	Electronic file – word	1 docx
PML	Gavin Tilstone	Electronic file – word	1 docx
ESA	Tânia Casal	Electronic file – word, pdf	1 docx, 1 pdf

4 Contents

5 Acronyms and Abbreviations

Acronym	Abbreviation
AAOT	Acqua Alta Oceanographic Tower
AC	Atmospheric Correction
ACRI-ST	Company in France, member of ACRI group
ADC	Analogue Digital Converter
AERONET	Aerosol Robotic Network
AERONET-OC	Aerosol Robotic Network Ocean Colour
AMT	Atlantic Meridional Transect
AOT	Aerosol Optical Thickness
ASD Fieldspec	Type of a spectroradiometer by Analytical Spectral Devices Inc.
Bio-ARGO	Bio-Optical Sensors on Argo Floats
BIPM	Bureau International des Poids et Mesures
BOUSSOLE	BOUée pour l'acquiSition d'une Série Optique à Long termE
BRDF	Bidirectional Reflectance Distribution Function
C2RCC-NN	Case 2 Regional Coast Colour Neural Network
C-OPS	Compact Optical Profiling System by Biospherical Instruments Inc.
CCD	Charge coupled device
CDR	Climate Data Records
CDOM	Coloured Dissolved Organic Matter
CEOS	Committee on Earth Observation Satellites
CEOS WGCV	Committee on Earth Observation Satellites Working Group on Calibration & Validation
CIMEL SeaPRISM	Above-water filter radiometer by Cimel Electronique
CMEMS	Copernicus Marine Environment Monitoring Service
CMODIS	Chinese Moderate Resolution Imaging Spectrometer
CNES	Centre National d' Etudes Spatiales
CNR	Consiglio Nazionale delle Ricerche
CNSA	China National Space Administration
COCTS	Chinese Ocean Colour and Temperature Scanner
COTS	Commercial Off the Shelf
CSV	Comma Separated Values
CZCS	Coastal Zone Color Scanner
CZI	Coastal Zone Imager
D-XXX	Reference to a FRM4SOC deliverable, where XXX denotes the particular number
ECMWF	European Centre for Medium-Range Weather Forecasts
ECV	Essential Climate Variables
EnMAP	Environmental Mapping and Analysis Program
ENVISAT	Environmental Satellite
EO	Earth Observation
EOEP	Earth Observation Envelope Programme
ESA	European Space Agency
ESRIN	ESA's centre for Earth observation
ESTEC	ESA's Space Research and Technology Centre
EU	European Union

Acronym	Abbreviation
EU/FP5	European Union Framework Programme 5
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FCDR	Fundamental Climate Data Record
FEL	An ANSI standard 1000 watt quartz tungsten halogen lamp
FICE	Field Inter-Comparison Experiment
FOV	Field of View
FR	Final Report
FRM	Fiducial Reference Measurements
FRM4SOC	Fiducial Reference Measurements for Satellite Ocean Colour
FWHM	Full Width at Half Maximum
GCOS	Global Climate Observing System
GEO	Group on Earth Observations
GEO-CAPE	Geostationary Coastal and Air Pollution Events
GEOSS	Global Earth Observation System of Systems
GIS	Geographic Information System
GLASS	Global Lake Sentinel Services
GLI	Global Imager
GUM	Guide to the Expression of Uncertainty in Measurement
GOCI	Geostationary Ocean Colour Imager
GSM	Global System for Mobile Communication
HICO	Hyperspectral Imager for the Coastal Ocean
HyperNav	Type of Hyperspectral radiometer by Sea-Bird Scientific
HYPERNETS	Hyperspectral radiometer integrated in automated networks
HyspIRI	The Hyperspectral Infrared Imager
IFOV	Instantaneous Field of View
IOP	Inherent Optical Properties
INSITU-OCR	International Network for Sensor Inter-comparison and Uncertainty Assessment for Ocean Colour Radiometry
ISO	International Standards Organisation
ISRO	Indian Space Research Organisation
ITT	Invitation to Tender
JAXA	Japan Aerospace Exploration Agency
JRC	Joint Research Centre
L2R	Level 2 Radiometric
LCE	Laboratory Calibration/Comparison Experiment/Exercise
KARI	Korea Aerospace Research Institute
KIOST	Korea Institute of Ocean Science and Technology
MAST	EU Marine Science and Technology Programme
MCM	Monte Carlo Method
MDPI	Multidisciplinary Digital Publishing Institute
MERIS	Medium Resolution Imaging Spectrometer (ESA ENVISAT mission)
MERMAID	MERIS MAtchup In-situ Database
MOBY	Marine Optical Buoy
MODIS	Moderate-resolution Imaging Spectroradiometer
MQC	Measurement Quality Control

Acronym	Abbreviation
MSI	Multispectral Imager
NA	Not Applicable / Not Available
NASA	National Aeronautics and Space Administration
NASDA	National Space Development Agency of Japan
NIR	Near Infrared
NEC	Nippon Electric Company
NIST	National Institute of Standards and Technology
NMI	National Metrology Institute
NOAA	National Oceanic and Atmospheric Administration
NPL	National Physical Laboratory
NRT	Near Real Time
OC	Ocean Colour
OCI	Ocean Colour Index
OCM, OCM-2	Ocean Colour Monitor/India
OCR	Ocean Colour Radiometer/Radiometry
OCR500	Type of multispectral radiometer by Satlantic Inc.
OCR-VC	Ocean Colour Radiometry-Virtual Constellation
OCTAC CMEMS	Ocean Colour Thematic Assembly Centre COPERNICUS Marine Environment Monitoring Service
OCTS	Ocean Colour and Temperature Scanner
OGC	Open Geospatial Consortium
OLCI	Ocean and Land Colour Instrument
OLI	Operational Land Imager
OSMI	Ocean Scanning Multi-spectral Imager
OSPREy	Optical Sensors for Planetary Radiant Energy by Biospherical Instruments Inc.
PACE	Plankton, Aerosol, Cloud ocean Ecosystem
PDF	Probability Distribution Function
POLDER	POLarization and Directionality of the Earth's Reflectances
POLYMER	POLYnomial-based algorithm applied to MERIS
PQC	Processing Quality Control
ProVal	Autonomous profiling float performing radiometric measurements
PML	Plymouth Marine Laboratory
QA4EO	Quality Assurance for Earth Observation
RBINS	Royal Belgian Institute of Natural Sciences
RefSpec	NPL Reference Spectroradiometer System
REVAMP	Regional validation of MERIS chlorophyll products in North Sea coastal waters
ROI	Return on Investment
S3MPC	Sentinel-3 Mission Performance Centre
S3VT	Sentinel-3 Validation Team
SAA	Solar Azimuth Angle
SBA	Skylight Blocked Approach
SDG	Sustainable Development Goal
SeaBASS	Bio-optical Archive and Storage System (of SeaWiFS)

Acronym	Abbreviation
SeaWiFS	Sea-Viewing Wide Field-of-View Sensor
SGLI	Second-generation Global Imager
SI	Systeme International d'Unites
SIMBADA	Type of an hand-held radiometer
SIMBIOS	Sensor Intercomparison for Marine Biological and Interdisciplinary Ocean Studies
SIO	Southern Indian Ocean (vicarious calibration site for NIR band)
SNR	Signal to Noise Ratio
SOA	State Oceanic Administration of China
SOR	Scientific Operational Roadmap
SOW	Statement of Work
SPG	South Pacific Gyre (vicarious calibration site for NIR band)
SPM	Suspended Particulate Matter
SRIPS	NPL Spectral Radiance and Irradiance Primary Standard
SRF	Spectral Response Function
SVC	System Vicarious Calibration
SVT	Sentinel Validation Team
SWIR	Short-wavelength infrared (1.4–3 μm)
SZA	Solar Zenith Angle
UN	United Nations
USGS	U.S. Geological Survey
UT	University of Tartu
VIM	International Vocabulary of Metrology
VIIRS	Visible Infrared Imaging Radiometer Suite
TO	Tartu Observatory
TOA	Top of Atmosphere
TR	Technical Report
TriOS	Manufacturer of optical metrology instruments
TriOS RAMSES	Hyperspectral Radiance and Irradiance Sensors for UV/VIS/NIR range by TriOS
UTC	Coordinated Universal Time
UV	Ultraviolet
VAA	View Azimuth Angle
VIS	Visible (spectral range)
VZA	View Zenith Angle
WCR	Water Colour Radiometry/Radiometers
WFS	Web Feature Service
WMS	Web Map Service
WebEx	Online teleconference and meeting platform
WISP	A portable water quality spectrometer (Water Insight)

6 Introduction

The Fiducial Reference Measurements for Satellite Ocean Colour (FRM4SOC) project was funded by the European Space Agency (ESA). It was structured to provide support for evaluating and improving the state of the art in satellite ocean colour validation through a series of comparisons under the auspices of the Committee on Earth Observation Satellites (CEOS) Working Group on Calibration & Validation (WGCV) and in support of the CEOS ocean colour virtual constellation (OCR-VC). FRM4SOC also strived to help fulfil the objectives of the International Ocean Colour Coordinating Group (IOCCG) in situ ocean colour radiometry white paper [1] and contribute to the relevant IOCCG working groups and task forces (e.g. the working group on uncertainties in ocean colour remote sensing and the ocean colour satellite sensor calibration task force).

The project made contribution to the European system for monitoring the Earth (Copernicus) through its core role of working to ensure that ground-based measurements of ocean colour parameters are traceable to SI standards [2]. This is in support of ensuring high quality and accurate Copernicus satellite mission data, in particular Sentinel-2 MSI and Sentinel-3 OLCI ocean colour products. The FRM4SOC project also contributes directly to the work of ESA and EUMETSAT to ensure that these instruments are validated in orbit.

The FRM4SOC project was carried out in 2016 – 2018 by the consortium consisting of four partners – the Royal Belgian Institute for Natural Sciences (RBINS), Belgium; National Physical Laboratory (NPL), UK; Plymouth Marine Laboratory (PML); ACRI-ST (France) and the lead partner University of Tartu[1] (UT), Estonia.

The current document D-290 "FRM4SOC Final Report" 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) [3].

The Final Report provides a thorough description of all work done and serves as a self-standing document, not requiring other project reports to be read meaningfully. It highlights all the activities conducted during the project (with reference to the deliverables of the contract) and results obtained. However, papers of the FRM4SOC Special Issue of the MDPI Journal *Remote Sensing* should preferably be considered as the primary reference for citation of the study. Such reference to the related paper is provided at the beginning of each relevant chapter.

[1] The lead partner Tartu Observatory (TO) was merged with the University of Tartu as a sub-institute since 01.01.2018.

7 Scientific background

Our modern society puts the limited natural resources on Earth under increasing pressure. We depend on these resources for our survival and development while our global population continues to grow, thus generating an ever-increasing demand for safe living space, fresh water, fertile land, and clean air. Society as a whole is facing numerous global threats, including climate change, energy crisis, potential food shortages, and a higher frequency and intensity of natural and manmade disasters. [4–7]

The United Nations (UN) General Assembly adopted the resolution "Transforming Our World: the 2030 Agenda for Sustainable Development" in September 2015. The 2030 Agenda also calls for new Earth Observation (EO) methods, data acquisition and exploitation of a wide range of data sources to support implementation. [4]

In particular, Article 76 states,

"*We will promote transparent and accountable scaling-up of appropriate public-private cooperation to exploit the contribution to be made by a wide range of data, including Earth Observation and geo-spatial information, while ensuring national ownership in supporting and tracking progress*". [4]

Vast amounts of global data are being collected from satellites as well as airborne, ground-based, and seaborne (*in-situ*) measurement systems all over the world as a key resource to support decision-making in addressing environmental challenges, and provide information for service providers, public authorities and other international organisations in improving the quality of life. Decision-making relies, and will continue to rely, on the ability of expert communities to utilize complex data from Earth observations and combine these with social and economic analyses. Sound, evidence-based decision-making will encourage sustainable behaviour by humankind in relation to Earth's resources, leading to economic benefits for all of society. [5,7]

It is recognised that collected EO data must be reliable and of high quality. For example if ground-based measurements are to be credibly used for satellite validation activities (particularly for assessment of climate data record stability, e.g. [8,9]) then they must be obtained contemporaneously, co-located with satellite measurements and be accurate and precise [3]. For that purpose, the Group on Earth Observation (GEO) has identified the need to develop and implement a data quality assurance strategy and works closely with the Committee on Earth Observation Satellites (CEOS). The mission of the CEOS Working Group on Calibration & Validation (WGCV) is to ensure long-term confidence in the accuracy and quality of Earth Observation data and products and to provide a forum for the exchange of information about calibration and validation, including the coordination of cooperative activities. The CEOS Quality Assurance Framework for Earth Observation (QA4EO) has set general principles for EO data quality assurance. [5,10–14]

As noted in 1995 at the 20th Conference Generale des Poids et Mesures [15], a recommendation was made that:

"*those responsible for studies of Earth resources, the environment, human wellbeing and related issues ensure that measurements made within their programs are in terms of well-characterized SI units so that they are reliable in the long term, are comparable world-wide and are linked to other areas of science and technology through the world's measurement system established and maintained under the Convention du Metre*".

This lays the foundation to relate satellite measurements to Systeme International d'Unites (SI) standards and gives the guiding principle for an EO data quality assurance strategy to

relate all data and derived products associated with them to SI reference standards. [2,3,10,16].
In addition, in order to have an objective indication for the quality level of data and compare available datasets meaningfully, all relevant measurement uncertainties must be evaluated [17]. The principles of metrological traceability and measurement uncertainty being a part of general, internationally recognised good practices for conducting measurements are developed and endorsed by the Bureau International des Poids et Mesures (BIPM) and National Metrology Institutes (NMI) [16–20].

As defined by BIPM and the International Vocabulary of Metrology (VIM) [19,21]:

Metrology *is the science of measurement, embracing both experimental and theoretical determinations at any level of uncertainty in any field of science and technology.*

Traceability *is a property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty.*

Calibration *is an operation that, under specified conditions, in a first step, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication.*

Unequivocally linking an "observation" to an invariant constant of nature (e.g. international system of units) with a robust estimate of uncertainty ensures the "measurement" can be: trusted, coherent and comparable with others, and have longevity "improving with age" [22].

We need to trust our collected measurement data:

- Fundamental Climate Data Records (FCDR),
- Data and time series,
- Data products

This can be achieved by implementation of the principles and methods of metrology including

establishment of the metrological traceability of the measurement data to the units of SI with related end-to-end uncertainty analysis.

8 Fiducial Reference Measurements

In order to build trust and ensure the quality of EO data, the concept of Fiducial Reference Measurements (FRM) has been established. [23–25]

FRM are a suite of independent, fully characterized, and traceable ground measurements that follow the guidelines outlined by the GEO/CEOS Quality Assurance framework for Earth Observation (QA4EO). These FRM provide the maximum Return On Investment (ROI) for a satellite mission by delivering, to users, the required confidence in data products, in the form of independent validation results and satellite measurement uncertainty estimation, over the entire end-to-end duration of a satellite mission. [23,24]

The defining mandatory characteristics for FRM are [23,24]:

- FRM have documented SI traceability (e.g. via calibration and/or round robin intercalibration of instruments) using metrology standards.
- FRM measurements are independent from the satellite geophysical retrieval process (noting the exception of L2 product vicarious adjustment that fundamentally depends on FRM ground based measurements).
- Uncertainty budgets for all FRM instruments and derived measurements are available and maintained, traceable where appropriate to SI, ideally directly through an NMI.
- FRM measurement protocols and community-wide management practices (measurement, processing, archive, documents, etc.) are defined, published openly and adhered to by FRM instrument deployments.
- FRM measurements are openly and freely available for independent scrutiny.

9 The FRM4SOC project

Copernicus is the European Union's Earth Observation Programme, looking at our planet and its environment for the ultimate benefit of all European citizens. It offers information services based on satellite Earth observation and in situ (non-space) data. The Programme is coordinated and managed by the European Commission (EC). It is implemented in partnership with the Member States, the European Space Agency (ESA), the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), the European Centre for Medium-Range Weather Forecasts (ECMWF), EU Agencies and their contractors. [6,7]

Within the context described above, ESA has initiated a series of projects targeting the validation of ESA altimetry, atmosphere, land, and ocean products [24]. The FRM4SOC project, with funding from ESA, has been structured to provide support for evaluating and improving the state of the art in OC validation through a series of comparisons under the auspices of CEOS WGCV and in support of the CEOS OCR virtual constellation [3,26].

The societal Benefits of Ocean Colour Radiometry (OCR) are well articulated (e.g. [27–29]) and include management of the marine ecosystem, the role of ocean ecosystems in climate change, aquaculture, fisheries, coastal zone water quality, and mapping and monitoring harmful algal blooms.

Addressing the need for reliable EO data – a series of recommendations on activities critical to ensure high accuracy and consistency for ocean colour mission products have been agreed under the guidance of the International Ocean Colour Coordinating Group (IOCCG), representatives of Space Agencies and Institutions supporting INSITU-OCR. [1]

The aim of the FRM4SOC project is: *"To establish and maintain SI traceability of Fiducial Reference Measurements for satellite ocean colour radiometry"* [3].

The Objectives of the FRM4SOC project were [3]:

Obj1. Design and document **measurement procedures and protocols for OCR FRM** used for satellite OCR validation activities.

Obj2. Document the design and **performance of OCR radiometers** commonly used for satellite OCR validation including a review of their known characterisation (e.g. immersion factor, cosine response, linearity, stray light, spectral, temperature sensitivity, dark currents etc.) and identify significant issues to address.

Obj3. Design, document protocols and procedures and implement a laboratory based (round robin) **comparison experiment to verify the performance of reference irradiance and radiance sources** (i.e. lamps, plaques, etc.) used to maintain the calibration of FRM OCR radiometers traceable to SI.

Obj4. Design, document protocols and procedures and implement a **laboratory based comparison experiment to verify the performance** (i.e. absolute radiometric calibration and characterisation) **of FRM Field Ocean Colour Radiometers (OCR)** used for Satellite Validation.

Obj5. Design, document protocols and procedures and implement **field intercomparisons of FRM OCR radiometers and build a database of OCR field radiometer performance** knowledge over a several years.

Obj6. **Conduct a full data analysis, derivation and specification of uncertainty budgets**, following agreed NMI protocols, for FRM OCR field measurements used for satellite OCR validation collected as part of FRM4SOC.

Obj7. Evaluate options for **long-term future European satellite OCR vicarious adjustment**.

10 Strategy for implementation of FRM

The main goal to ensure the high quality of EO data is achieved by implementing the following activity chain [30]:

1. Analysis and establishment of requirements.
2. Definition of measurement methods and protocols to meet the requirements.
3. Selection of instruments that meet the established requirements and protocols.
4. Establishment of the traceability chain to the units of SI by calibration (Figure 1.).
5. Evaluation of uncertainty sources (including characterisation of instruments) and compilation of end-to-end uncertainty budgets (including characterisation).
6. Validation of the methods and uncertainty budgets in comparison experiments.
7. Collection and database storage of measurement and comparison data.

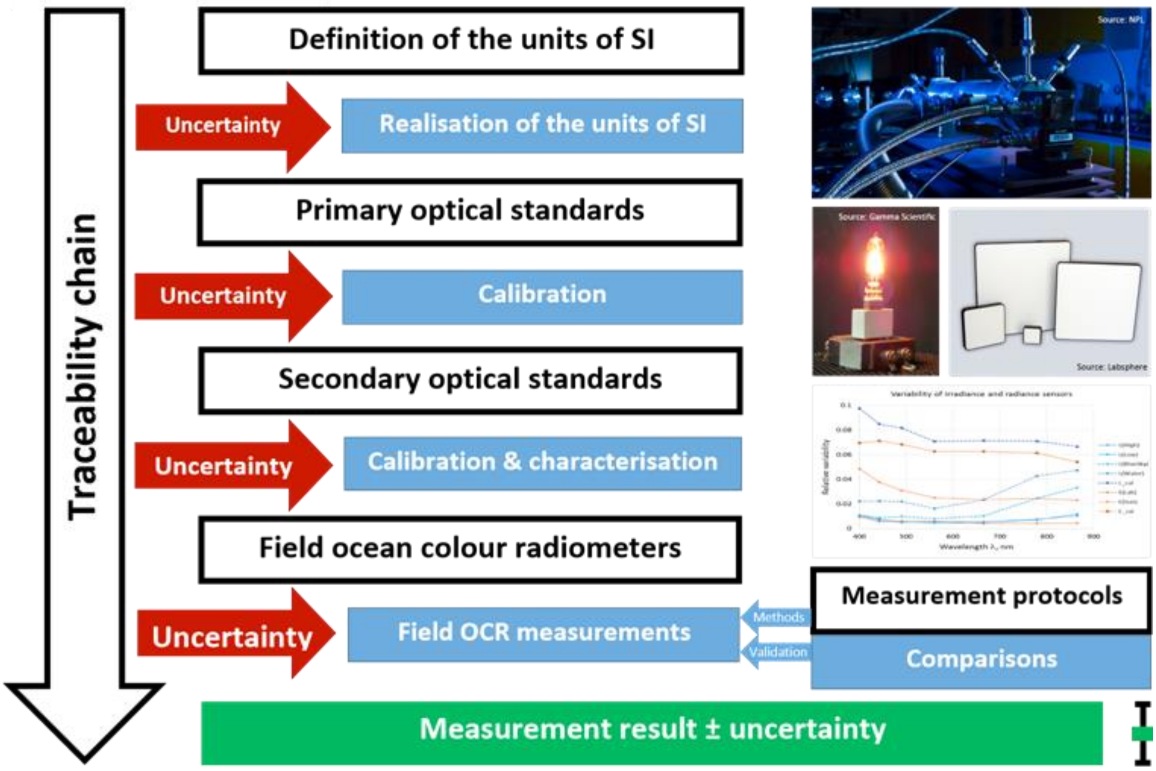


Figure 1. Establishment of the traceability chain with related uncertainty evaluation for collection of reliable in situ data.

11 Requirements and recommendations for infrastructure required for the long-term vicarious calibration of the Sentinel-3 OLCI and Sentinel-2 MSI A/B/C and D instruments

11.1 International workshop "Options for future infrastructure required for the long-term vicarious calibration of the Sentinel-3 OLCI and Sentinel-2 MSI A/B/C and D instruments" (WKP-1) [D-230]

An international workshop "*Options and approaches to the long-term vicarious calibration of Sentinel- OLCI & MSI A/B/C and D instruments*" was held on 21 – 23 February 2017 at ESA/ESRIN, Frascati, Italy [3,26,31]. The objective of the workshop are listed below:

1. Foster an open-forum, wide-ranging debate with the international ocean colour community Figure 2 and Figure 3;
2. Review of historical and contemporary approaches to vicarious adjustment;
3. Document lessons learned from international teams;
4. Review the strengths and weaknesses of alternative methods and approaches to OCR satellite vicarious adjustment;
5. Derive an optimum European location for OCR vicarious calibration infrastructure based on spatial and temporal distributions of chlorophyll, atmospheric aerosol loading and cloud cover (and other geophysical quantities if deemed appropriate);
6. Review and define justified traceability requirements for vicarious calibration measurements to be made in support of satellite OCR;
7. Review the costs to implement, operate and maintain a European satellite OCR vicarious adjustment infrastructure for Sentinel-2 and -3 missions;
8. Conclude with a consensus on the way forward to deliver the best scientific outcomes to support long-term Copernicus operations using European infrastructure Sentinel-2 and Sentinel-3 OCR vicarious calibration infrastructure.

Figure 2. The workshop gathered excellent world-class specialists from a large diversity of institutions and scientific fields.

The workshop gathered the world experts in ocean and satellite-borne optical radiometry to open a wide-ranging debate on the way forward to provide the best possible outcome of the Sentinel-2 and Sentinel-3 series. Ocean colour remote sensing relies on highly precise and accurate in situ measurements of the optical properties of the oceans (FRM), to optimize its reliability through indirect calibration, the so-called System Vicarious Calibration (SVC).

An important aspect of the meeting was to analyse the actual needs from the community (Sentinel-2 and Sentinel-3 Mission Performance centres, Copernicus – Marine Environment Monitoring Service, scientific users) and constraints for long term applications like the generation of Climate Data Records (CDR) of Essential Climate Variables (ECV) to specify the requirements for future System Vicarious Calibration (SVC).

On the oceanic aspect, great attention was paid to reviewing and learning from the experience of the existing reference sites for SVC: MOBY (the Marine Optical Buoy) deployed off the Hawaiian coast since 1996 and BOUSSOLE (Buoy for the acquisition of long-term optical times series) deployed in the Ligurian Sea since 2003. The two systems are currently being refreshed to better respond to the new challenges of operational ocean colour remote sensing. In addition, new emerging techniques based on autonomous profiling floats have also been reviewed as they offer potential for both vicarious adjustment and data validation [32,33].

The discussions held during the workshop have converged toward a consensus for future development of SVC infrastructure in Europe.

Figure 3. The discussions held during the workshop converged toward a consensus for future development of SVC infrastructure in Europe.

11.2 Proceedings of the workshop (PROC-1) [D-240]

The presentations and proceedings [29] of the workshop are available at the FRM4SOC project website https://frm4soc.org [23].

The conclusions from the workshop are summarised in the technical report TR-10. [34]

11.3 Report on requirements and recommendations for infrastructure required for the long-term vicarious adjustment of the Sentinel-3 OLCI and Sentinel-2 MSI A/B/C and D instruments (TR-10) [D-250], [34]

11.3.1 Introduction

Since the launch of SeaWiFS in 1997, spaceborne ocean colour sensors have provided continuous records of ocean optical properties and opened new research areas as well as industrial and technological innovation to support its development. Satellite ocean colour has provided the means for monitoring the spatial and temporal variability of remote areas of the world's oceans hardly accessible through conventional shipborne surveys on a daily basis. It has also enabled the collection of crucial data from other complex coastal regions where anthropogenic activities and interaction with terrestrial ecosystems have a major influence on the optical properties of water. A large international scientific community now relies on OCR data to perform activities covering from short or medium term coastal water quality monitoring to long-term analysis of climate change through Climate Data Records of Essential Climate Variables. Several Space Agencies have been actively supporting ocean colour remote sensing in the past: NASA (CZCS, SeaWiFS, HICO), ESA (MERIS), CNES (POLDER series), CNSA/SOA (CMODIS and COCTS), NASDA/NEC (GLI, OCI, OCTS), ISRO (OCM), KARI (OSMI). Ten OCR missions are currently in operation: SOA (COCTS CZI), JAXA (SGLI), KARI/KIOST (GOCI; the first geostationary ocean colour sensor), NASA (MODIS Aqua and MODIS Terra), ISRO (OCM-2), ESA/EUMETSAT (OLCI-A and OLCI-B), NOAA (VIIRS 1 and 2). Several other OCR missions are planned to launch in the near future. In this context, the European Commission Copernicus ocean colour program has the most ambitious objective in the long term with the Sentinel-3 series committed to maintain two identical sensors in orbit (OLCI) for the next decades. Two of this family are already in operation (OLCI-A and OLCI-B). OLCI-C and OLCI-D are currently being assembled by Thalès Alenia Space. With the OLCI series, scientists around the world are ensured that data will continuously flow in the future to support their activities. Also importantly, with European ocean colour going fully operational, public institutions and private businesses can invest in the development of environmental monitoring and services based on these data. CMEMS for instance already relies on OLCI data for its operational services.

In order to ensure the best possible data quality for scientific research, operational monitoring and commercial applications, it is crucial that OLCI data processing provides the best possible product quality. The sub-sections below present the general principle of Ocean Colour Radiometry (OCR) from spaceborne sensors and the justification for SVC.

11.3.2 Principle of OCR

Spaceborne sensors measure the radiance leaving the earth atmosphere, referred as the Top Of Atmosphere (TOA) radiance. Past (HICO), present (CHRIS, PRIMSA) or planned (EnMAP, PACE, GEO-CAPE/GLIMR, HyspIRI/SBG, CHIME) missions have provided or will provide hyperspectral data but the large majority of OCR sensors are the so-called multispectral sensors, measuring TOA radiance at discrete wavelengths (λ) ranging from visible (~400 nm) to near infrared (NIR; ~1000 nm) [35]. Most of the signal measured by an ocean colour sensor actually comes from the atmosphere that represents from 60% to 80% of the total signal in the visible spectral region depending on the wavelength. Atmospheric Correction (AC) needed to separate and retrieve the water signal is therefore the most important part of ocean colour data processing.

Historically, the NIR bands have been used to perform the atmospheric corrections based on the so-called black pixel assumption [36,37]. The black pixel assumption states that a marine pixel's signal is null in the NIR. The entire signal measured by a spaceborne sensor therefore comes from the atmosphere. The molecular (Rayleigh) and aerosol signal is estimated at these bands and then extrapolated toward the visible bands. While true in the open ocean, the black pixel assumption is not realistic for coastal and shallow water pixels due to the influence from the continent and most importantly river sediment discharge and sediment resuspension. Specific algorithms have to be applied to these pixels to account for the water signal in the NIR bands [38,39] or, for sensors such as MODIS and Sentinel-2/MSI, the black pixel assumption can be applied reliably to the Short Wave Infra-Red bands [40].

Alternatively atmospheric correction algorithms using all available wavebands can be implemented: C2RCC-NN [41], POLYMER [42]. This document will focus on the standard "NIR black pixel" AC and the derived methodology to perform SVC. SVC is indeed closely linked to the AC procedure itself. Vicarious gains computation defined in the following sections is therefore solely based on the standard AC.

Water-leaving radiance or reflectance in the visible spectrum are the core variables produced through ocean colour data from which are derived biogeochemical variables like the diffuse attenuation coefficient K_d, Inherent Optical Properties (IOP), Suspended Particulate Matter (SPM) concentration and the ECV chlorophyll concentration. If regional water quality monitoring and global change studies based on OCR are to be trusted, it is essential to ensure the quality of these products.

11.3.3 Optical pathways

Radiant power from the sun reaches the atmosphere where it is partially absorbed at specific wavelengths by aerosols and atmospheric gases like O_2, O_3, H_2O, and scattered by atmospheric molecules (Rayleigh scattering) and aerosols (Mie scattering) before it hits the sea surface. The sea surface, depending on the sea state and sun geometry, reflects part of this solar radiance back into the atmosphere while some of the energy penetrates the water. Within the water column, the solar energy can either be diffused or absorbed by water molecules and dissolved or particulate matter. The optically active components of the water are the water itself, dissolved organic matter output of phytoplankton decay of plants through river runoff, phytoplankton and suspended sediments. Air bubbles created by breaking waves can significantly influence light scattering in the water but OCR radiometry products generally exclude from the processing chain pixels where the wind speed is strong enough to generate significant scattering by bubbles. Part of the light penetrating the ocean is backscattered into the atmosphere where it undergoes the same physical phenomena prior to eventually reaching the space borne sensor. Additional geophysical phenomena, e.g., major volcanic eruptions enriching the troposphere with aerosols, are not mentioned here. They can also highly influence solar energy transfer through the atmosphere. Furthermore, in shallow and clear water, part of the solar energy can be reflected by the seabed.

Figure 4 summarises the complexity of the optical pathways through the atmosphere and consequently the complexity of ocean colour product retrieval.

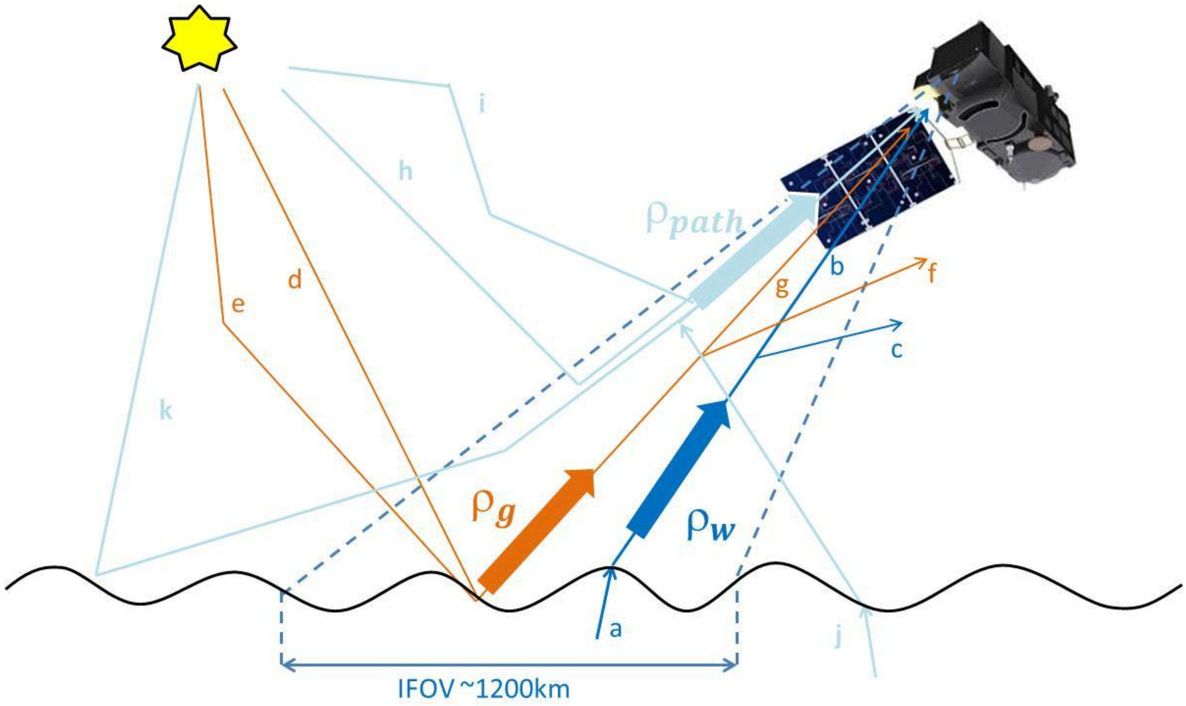


Figure 4. (Adapted from [41]) Optical pathways sun/water/sensor. (a) light rays originating from the water column within the sensor's Instantaneous Field of Fiew (IFOV) and scattered back toward the atmosphere. They constitute the water-leaving reflectance (ρ_w) once they have been refracted at the water/air interface. They may reach the sensor (b) if they are not absorbed or scattered out of the sensor's IFOV on the upward pathway (c). (e) and (d) constitute the glint. (d) are the sun's rays reflected by the water surface into the sensor (the sunglint) (e) are the sun's rays scattered by the atmosphere into the sensor's IFOV and reflected by the water surface into the sensor (skyglint). They constitute the glint reflectance (ρ_g). These rays partly reach the sensor (g) and partly are scattered out of the IFOV (f). (h) and (i) are rays reaching the sensor after single or multiple scattering in the atmosphere without interacting with the surface. (j) are rays emerging from water outside of the sensor's IFOV (either in the IFOV of another pixel or entirely outside the sensor's FOV). They reach the sensor after being scattered in the atmosphere within the sensor's IFOV. (k) are rays reflected from the water surface outside the IFOV that reach the sensor after being scattered within the IFOV. (h), (i), (j) and (k) constitute (ρ_{path}) atmospheric path reflectance.

For satellite data processing, a practical solution to account for all interaction of solar energy before it reaches the sensor is described and in Figure 5 [37], and can be summarized by the following operational equation considering a configuration that avoids sunglint and no contribution from whitecaps.

$$\rho_t(\lambda) = t_g(\lambda) \cdot \left(\rho_A(\lambda) + \rho_R(\lambda) + t_d(\lambda) \cdot \rho_w(\lambda)\right) = \\ = t_g(\lambda) \cdot \left(\rho_{path}(\lambda) + t_d(\lambda) \cdot \rho_w(\lambda)\right) = \frac{\pi . L_t(\lambda)}{\cos(\theta_s) . F_0}, \quad (1)$$

where $\rho_t(\lambda)$ is the TOA reflectance measured by the sensor, $t_g(\lambda)$ is the total gaseous transmittance of the atmosphere (upward and downward), $\rho_A(\lambda) = \rho_a(\lambda) + \rho_{ra}(\lambda)$ is the total aerosol reflectance signal including pure aerosol $\rho_a(\lambda)$ and multiple scattering between air molecules and aerosols $\rho_{ra}(\lambda)$, $\rho_R(\lambda)$ is the molecular (Rayleigh) scattering. $\rho_w(\lambda)$ is the water-leaving reflectance and $t_d(\lambda)$ is the total (direct and diffuse) upward transmittance of the atmosphere, L_t is the TOA radiance, F_0 is the TOA irradiance, and θ_S is the solar zenith angle.

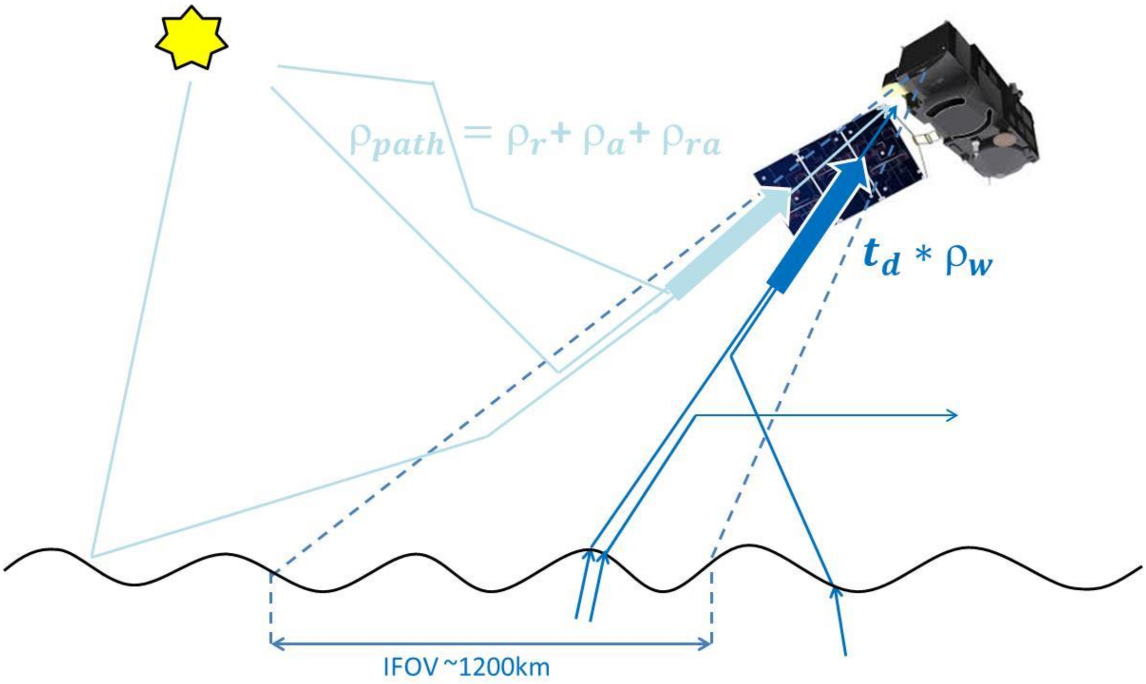


Figure 5. Practical solution to model the sun/water/sensor radiative transfer [43].

11.3.4 Space mission requirements

Space mission requirements have been described in [44] and reviewed by [45] and [46]. Essential requirements for the purpose of SVC are described below.

The primary aim of OCR spaceborne sensors is to retrieve the water-leaving signal from a Top Of Atmosphere (TOA) optical measurement over the visible (VIS) and NIR spectral domain. The authors of [45] have reviewed the historical requirements of ocean colour space missions. There are essentially three requirements available in the literature for ocean colour radiometry:

- 5 % uncertainty in satellite-derived ρ_w in the blue spectral region to allow for the determination of Chl-a concentration in oligotrophic waters with a standard uncertainty of 35 % quantified through the work of [47–49];
- 5 % spectrally independent uncertainty in satellite-derived ρ_w across the blue-red bands set as an objective (not a scientific requirement) of the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) mission [50]. This broad objective was later interpreted or set, as a scientific requirement for several missions;
- 5 % radiometric uncertainty in satellite derived ρ_w in the blue-green spectral bands in oceanic waters and 0.5 % radiometric stability over a decade for the creation of Climate CDR of ECV [20,51].

Actual uncertainty of remotely-sensed ρ_w depends mainly on the quality of both TOA acquisition (i.e. quality of the absolute and interband sensor calibration) and the atmospheric correction (i.e. ability to estimate and remove the atmospheric path contribution, see e.g. [37]). This can be made explicit by the decomposition of the signal in ideal conditions without sun specular reflection or white caps, and after correction of atmospheric gaseous absorption as defined in [42]

$$\rho_t(\lambda) = \rho_{path}(\lambda) + t_d(\lambda) \cdot \rho_w(\lambda). \tag{2}$$

Most of the current operational atmospheric correction algorithms consist in first, assessing the aerosol optical properties from the NIR bands; then, propagating the path reflectance $\rho_{\mathrm{path}}(\lambda)$ and total transmittance $t_{\mathrm{d}}(\lambda)$ at any wavelength λ in the visible spectral range; and finally, deducing the water-leaving signal by inversing Equation (1) (see [37] for MERIS and [36] for SeaWiFS). Hence, whatever the accuracy of the path reflectance retrieval, any uncertainty $u(\rho_{\mathrm{t}})$ on the total signal implies an uncertainty $u(\rho_{\mathrm{w}})$ on the water-leaving reflectance of:

$$\frac{u(\rho_w(\lambda))}{\rho_w(\lambda)} = \frac{u(\rho_t(\lambda))}{\rho_t(\lambda)} \Big/ \frac{t_d(\lambda)\rho_w(\lambda)}{\rho_t(\lambda)}. \tag{3}$$

The equation above demonstrates the statement from [52]: for typical open ocean waters, the water-leaving reflectance signal reaching the sensor is about 10 % of the TOA signal. Therefore, if an uncertainty of ±5% on ρ_w is to be achieved, a total uncertainty of ±0.5 % is expected on TOA signal. This cannot be achieved by prelaunch nor on-orbit calibration. This is where System Vicarious Calibration enters in the processing chain.

11.3.5 System vicarious calibration (SVC)

It is important in this document to clarify what is meant by SVC. On a routine basis "vicarious calibration" is often used to refer to methodologies using ground targets of known signals (hot and cold homogeneous deserts, ocean gyres, deep convective clouds) and radiative transfer models to simulate TOA radiance. These methodologies in an OCR perspective are more specifically used for calibration validation. They are not accurate enough to reach ocean colour uncertainty requirements on water-leaving reflectance.

The vicarious calibration implemented by the ocean colour community is referred to as a "System Vicarious Calibration" (SVC). This means that measurements in all visible spectral bands are adjusted by application of a multiplicative factor (a "SVC gain"). A single set of spectral gains is applied to all data for a given mission, which is aimed at correcting for possible errors in both the instrument response and in the atmospheric correction algorithm. As such, SVC is supposed to reduce uncertainty by dealing essentially with improving accuracy, not so much dealing with precision. SVC gains are applied to TOA reflectances before they enter the atmospheric correction process. The SVC process does not deal with possible temporal changes in the instrument response, which are dealt with by other means (e.g., using solar diffusers in the case of the S3/OLCI missions, or regular lunar views for sensors like MODIS). The "ground truth" is generally based on very high quality field radiometric measurements, the so-called FRM.

In general terms SVC aims to compute a target or theoretical $\rho_t^t(\lambda)$ through a direct model by using FRM data. While (1) represents the measured signal of a sensor in orbit, $\rho_t^t(\lambda)$ can be written as:

$$\rho_t^t(\lambda) = \rho_{path}(\lambda) + t_d(\lambda) \cdot \rho_w^{IS}(\lambda), \tag{4}$$

where $\rho_w^{IS}(\lambda)$ is the measured or modelled water signal. A time series of gains, $g(\lambda, i)$, where i is the number of a particular matchup, can then be calculated for a given pixel:

$$g(\lambda, i) = \frac{\rho_{\mathrm{t}}^{\mathrm{t}}(\lambda, i)}{\rho_t(\lambda, i)} \ . \tag{5}$$

Once a sufficient time series has been accumulated, a mean gain $\bar{g}(\lambda)$ can be derived and applied to the entire satellite time series (N is the total number of matchups):

$$\bar{g}(\lambda) = \frac{\sum_{i=1}^{N} g(\lambda, i)}{N}. \tag{6}$$

Ideally, FRM of both marine and atmospheric in situ variables (ρ_w and $AOT(\lambda)$) should be used. Technical and instrumental progress may provide this capacity in the future but for the time being, the marine FRM have been used solely for SVC and atmospheric values have been derived from the atmospheric correction model itself. Owing to the constraints of FRMs and SVC requirements, MOBY and BOUSSOLE radiometric time series are generally used for SVC. However, for sensors with design constraints like GOCI, a geostationary platform with neither BOUSSOLE nor MOBY in its field of view, high quality data acquired during field cruises have been used for SVC. Therefore, any FRM coming from cruises or permanent monitoring systems can support SVC provided that their level of uncertainty is low enough.

SVC is applied to the combined sensor and Level-2 processing chain, i.e. consequently after applying the radiometric calibration coefficients of the satellite sensor. It is therefore assumed that all possible efforts have been made in pre-launch and post launch instrument calibration and characterisation. This includes measuring the spectral response function, and out of band response when relevant, temperature sensitivity, dark signal, and on-board diffuser degradation. Temporal degradation has to be accounted for and regularly updated as this is the crucial element for applying a unique gain for the entire lifetime of the mission. Long-term stability of the SVC gain should also be monitored as any temporal trend would detect a temporal degradation correction failure.

Within the metrology community, there is not yet a consensus whether the procedure should be referred to as System Vicarious Calibration or System Vicarious Adjustment. This point is beyond the scope of this document and SVC will be used to stick to the historical denomination.

11.3.6 Challenges of operational OCR

Copernicus operational services rely on high quality OCR data. Among them a core service, OCTAC CMEMS is providing near real-time (NRT) and reprocessed level 3 and level 4 global multi-sensor products as well as regional single and multi-sensor products of European seas (Figure 6). OCTAC data are subsequently operationally used for data assimilation in bio-geochemical models (regional and global), and provision of marine environment indicator for marine policy and management of marine resources. The quality of CMEMS Ocean Colour products and all the downstream applications therefore strongly depends on the quality of upstream satellite data. Operational oceanography needs vicarious calibration for a stable calibration of the OC sensors and a prompt uncertainty assessment for operational NRT data. In an operational context, the vicarious calibration gains should be available as soon as possible and frequently updated to ensure the accuracy of the NRT operational data. All previously acquired data affected by SVC should be reprocessed to improve the gain accuracy and to ensure the accuracy required for climate observations.

While MERIS on-board ENVISAT was a scientific mission with no specific timeliness constraints, OLCI on-board Sentinel-3 has kicked-off the operational era with consequent strict constraints on product quality and tight timelines for product delivery. As an example, SVC was implemented in MERIS processing during its 3rd full mission reprocessing, which took place about eight years after its launch. In comparison, SVC had to be implemented in the OLCI processing chain before Level-2 product public release, which took place about a year after launch. This is clearly a challenging situation, primarily because all Level-1 products issues may not have been dealt with (on-board calibration, temporal degradation, temperature sensitivity, etc.). Such a short time range does not guarantee sufficient data to derive NIR gains computations, which solely rely on satellite data, then, it becomes extremely complicated for visible bands calibration needing FRMs, which are so far only provided by BOUSSOLE and MOBY.

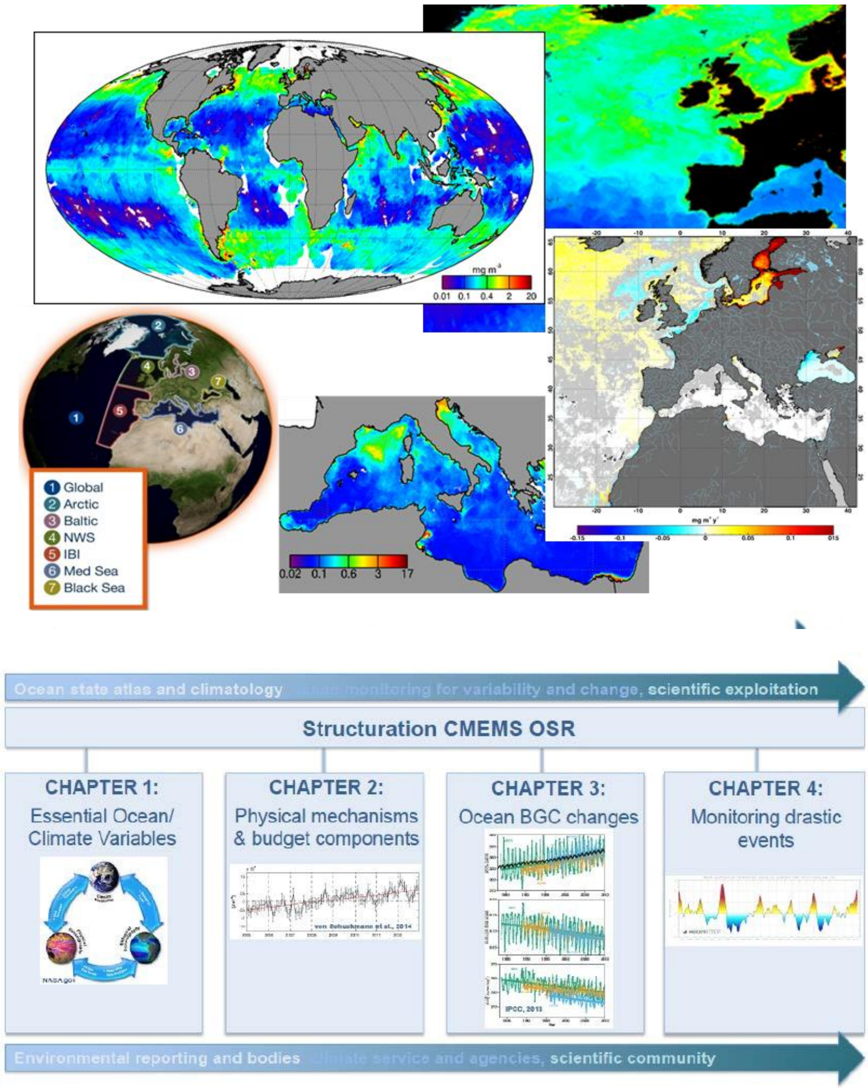


Figure 6. Use of OC products in CMEMS. [53]

11.3.7 Existing SVC infrastructures

The development of sea-going instrumentation dedicated to optical radiometry has been highly motivated by the development of applications using satellite data. Several manufacturers now provide instrumentation for this purpose. However, standard procedures for instrument characterisation implemented by the manufacturers generally do not satisfy the needs for SVC nor FRM and the full and costly characterisation generally has to be performed by the users themselves. When feasible, an increased interaction between manufacturers and users is key to achieving higher standards in instrument characterisation. A detailed review of existing instrumentation used in OCR is provided in [54] [D-70]. Nevertheless, the high level of requirements for SVC will surely impose post-factory characterisation.

The two reference infrastructures BOUSSOLE [55,56] and MOBY [57,58], (Figure 7) have adopted different concepts. For MOBY, dedicated instruments have been developed for the purpose of SVC, therefore, accounting from the beginning for the specific needs of SVC. MOBY was specifically designed for SVC and therefore programmed to collect data at the time of a satellite overpass. Detailed analysis of MOBY instrument calibration, uncertainty budgets and end to end processing are available in [59–68].

For BOUSSOLE, commercial off the shelf (COTS) instrumentation has been purchased and fully characterised before deployment of the infrastructure. BOUSSOLE is designed to support both SVC and scientific research. For this reason, it has been collecting radiometric data on a daily basis from dawn to dusk, therefore, generating a unique time series for now nearly two decades and contributing to multiple scientific publications. Detailed analysis of the BOUSSOLE infrastructure, instruments calibration and uncertainty budgets are available in [55,69–71].

Whatever the system (COTS instrumentation or specifically designed for SVC), the field instrumentation has to provide the lowest possible data uncertainty. This means that characterisation of the instrument including absolute radiometric calibration of spectral response function, wavelength scale calibration, testing diffuser cosine response, temperature dependence, stray light correction, nonlinearity, etc. has to be performed with state of the art procedures and with accurately derived uncertainties. Traceability to the International System of Units (SI) is obviously mandatory. This constitutes the first step towards achieving SVC grade data. Additional care in data reduction and quality control are also part of the system to ensure the lowest uncertainties. Several sets of instruments have to be available to ensure rotation of the system.

Hyperspectral radiometers should be used for SVC purpose to support SVC of multiple missions. While MOBY made the choice of custom designed hyperspectral radiometers from the beginning, BOUSSOLE, being initially designed for MERIS, opted for multispectral sensors. Hyperspectral sensors were added in 2009 and used alongside the multispectral sensors to ensure the continuity and consistency of the time series. The multi-spectral instruments were decommissioned in 2018. Hyperspectral instrumentation is obviously a logical choice for evolution of an SVC infrastructure. However, it brings additional challenges to the minimization of uncertainties.

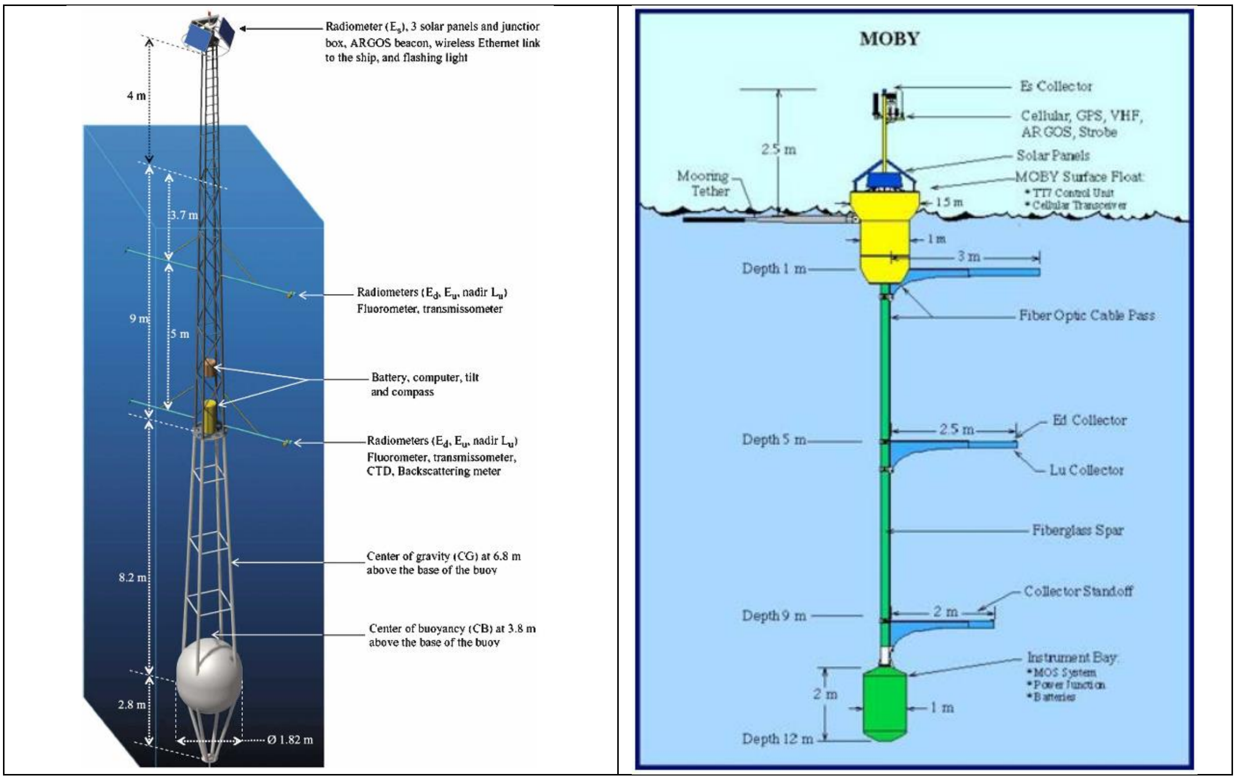


Figure 7. SVC infrastructure designs – BOUSSOLE (left) [55] and MOBY (right) [58].

11.3.8 Recommendations for SVC instrumentation

11.3.9 Spectral range

Current satellite ocean colour missions cover the (375…12 500) nm range while scheduled ocean colour missions will cover the (340…2500) nm range. Presently, OLCI covers the (400…1020) nm spectral region. For SVC purposes in situ radiometers should therefore cover the 340 nm to 700 nm spectral range to sustain the multi mission needs of current and planned OCR sensors. However, it should be noted that accurately deriving the water-leaving radiance (or reflectance) for wavelengths beyond about 600 nm is challenging due to the extremely small water-leaving signal at these bands in waters suitable for SVC (i.e., clear oceanic waters).

NIR and SWIR bands (>700 nm) FRM are not required for SVC: NIR band SVC can be performed over the clearest oligotrophic regions of the oceans where the NIR marine signal can be assumed negligible. Recent surveys [72] have suggested that the assumption of negligible signal in the NIR should nonetheless be demonstrated and the uncertainty associated with this assumption quantified. However, SVC sites used for NIR band calibration are generally in remote areas, such as South Pacific Gyre (SPG) and Southern Indian Ocean (SIO), and these requirements would probably be too costly to implement.

Recommendation:

- For SVC purposes, in situ radiometers should cover the 340 nm to 700 nm spectral range.

11.3.10 Spectral resolution

Ideally, radiometric instrumentation deployed on an SVC infrastructure should be hyperspectral in order to cover multi mission SVC needs. A detailed analysis of the impact of the spectral resolution and derived requirements for SVC needs have been published in [73]. That analysis was conducted on OLCI (a multispectral sensor with 10 nm bandwidth in the visible; 15 nm at 400 nm) and PACE (a planned hyperspectral mission targeting 5 nm bandwidth in the visible). Spectral resolution requirements for both validation and the more stringent SVC purpose were derived. While a target of less than 1 % difference between satellite and in situ remote sensing reflectance R_{rs} originating from spectral resolution only was assumed sufficient for data validation, less than 0.5 % was specified for hyperspectral radiometers supporting SVC. This 0.5 % maximum difference subsequently led to the requirement for spectral resolution better than 3 nm for an OLCI-like sensor. Additionally, in [73] it was argued that using water-leaving radiance L_w rather than R_{rs} also increases requirements to less than 1 nm spectral resolution.

Recommendations:

- A spectral resolution better than 3 nm is required for the OLCI satellite sensor.
- A spectral resolution better than 1 nm is devised for the PACE satellite sensor.
- Valid algorithms for the convolution of radiometric quantities with instruments spectral bands response functions are required [74].

11.3.11 Instrument characterisation

To fulfil SVC needs, radiometers must be fully characterised and regularly calibrated. This means that several sets of instruments must be within the SVC package to ensure continuity of the system when the instruments are in post deployment maintenance. Calibration includes spectral and radiometric calibration. Calibration should be with standard uncertainty lower than 2% traceable to SI scale, including uncertainty in the source, its transfer and error corrections. Instruments shall be with highly stable radiometric properties having maximum drift not exceeding 1 % per deployment and a target of 0.5 % [31,73]. Full characterisation includes spectral responsivity, spectral stray light (out-of-band response for filter radiometers), angular (cosine) response, immersion coefficients for in-water radiance and irradiance sensors, thermal stability, dark current, polarisation sensitivity, non-linearity response. A detailed list of SVC requirements for instrument characterisation has been provided in [72].

All measurements provided by an SVC infrastructure shall be traceable to SI standards:

- Radiometric instruments must be regularly calibrated spectrally and radiometrically.
- Radiometric instruments must be fully characterised.

11.3.12 Ancillary data

In addition to radiometry, it is advisable that the SVC field infrastructure is also equipped with instrumentation to determine IOPs of the water. The instrument set should include transmissometers and backscattering meters. An absorption meter could be considered but currently none of the existing technologies has proven to be adapted for long-term deployments. Chlorophyll fluorescence should also be monitored. Chlorophyll content is an important source of uncertainty in the SVC process, therefore, it is a fundamental parameter that should be carefully monitored. Meteorological and oceanographic data must be collected for quality control purposes. This includes wind speed and direction, atmospheric pressure, wave height, etc. Ideally, spectral Aerosol Optical Thickness (AOT) should be measured at

SVC sites but the current technology does not allow it to be performed autonomously (but feasible during servicing cruises to the SVC sites). Existing systems deployed on AERONET and AERONET-OC ground stations are not compatible with buoy deployment. Lidar systems might be envisaged in the future.

Recommendations:

- Inherent optical properties must be monitored at SVC sites.
- Chlorophyll content must be monitored at SVC sites.
- Meteorological and oceanographic properties must be monitored at SVC sites.

11.3.13 Recommendations for SVC infrastructure

While it is not excluded that FRM acquired during field campaigns can support SVC providing they have been performed in suitable conditions, the recommendations in this section will focus on permanent or semi-permanent system deployments.

11.3.13.1 Location

Historical requirements for SVC site location were defined by [52]. A suitable locations for an SVC sites should present the following environmental conditions:

- The SVC sites should be in a location with sufficient occurrence of cloud free days per year.
- The atmosphere should be clear with dominant marine aerosol types of AOT lower than 0.1 in the NIR and no absorbing aerosols to reduce uncertainties on atmospheric correction in the Level-2 processing chain. Consequently, SVC sites should not be located along the coast or in areas dominated by a continental atmosphere. This can be evaluated by assessing the aerosol angstrom exponent (maritime atmospheres having values of this exponent generally lower than about 0.5).
- The water body should be spatially homogenous so that a point measurement can reasonably be assumed representative of a satellite pixel or macropixel. Spatial homogeneity cannot be just claimed but should be characterised by field surveys.
- The SVC site should be located in oligotrophic to mesotrophic waters in order to minimize variability of the L_w signal in the blue-green region of the spectrum.
- The SVC site should additionally offer a sufficient number of days of low sea state or with low currents so as to limit the infrastructure tilt in case it is sensitive to currents, and low wind to minimise white caps formation and the probability of sunglint due to surface roughness.
- The SVC site should preferably be located in low latitudes to reduce the variability in solar zenith angle and therefore reduce uncertainties in atmospheric correction.

Local cloud climatology should be taken into account with the specific overpass time. Morning or afternoon overpasses can experience significant differences in cloudiness due to morning haze or afternoon evaporation in tropical regions (while the opposite is also frequently observed with clouds building up over the day).

In addition to the environmental conditions, practical considerations also have to be taken into account:

- SVC sites should be in the vicinity of a harbour to facilitate logistics. Regular “light weight” cruises are mandatory for instrument and infrastructure maintenance. Also, occasional “heavy weight” cruises will take place for platform turn-over.
- In addition to ship time, highly qualified scientists and technicians must be present near the SVC site to ensure the maintenance of the platform and instrumentation.
- Also importantly, the SVC site should be within GSM or other data link range to ensure NRT data transfer, although the use of satellite data transmission (e.g., Iridium) relieves this constraint.
- The site should be outside of commercial shipping routes or recreational fishing areas to avoid accidents or vandalism.
- The SVC infrastructure shall be autonomous, including power supply.

A detailed study with a selection of nine realistic locations relevant for SVC based on SeaWiFS time series analysis has been published in [73]. Based on the conclusions of this paper and the discussions held during the international workshop on SVC infrastructure (ESRIN, February 2017, [D-230] [26]), two particularly suitable locations were shortlisted in Europe: BOUSSOLE in the Ligurian sea and a location in the vicinity of Crete in the Eastern Mediterranean. A further, suitable site outside Europe was identified off the western coast of Australian.

Recommendations:

- Two locations were shortlisted in Europe: BOUSSOLE in the Ligurian Sea and a location in the vicinity of Crete in the Eastern Mediterranean.
- One location was shortlisted off the western coast of Australian.

11.3.13.2 Number of required SVC infrastructures

There are currently two types of long-term systems dedicated to radiometric measurements in the oceans: above and below water. Above-water systems like AERONET-OC have proven extremely valuable for providing operational data that can be retrieved in near real time for validation purposes. There is a community consensus though, that above-water systems should not be preferred for the SVC purpose. The discussion in this section will therefore focus on underwater systems and provide examples of the two reference ones: BOUSSOLE and MOBY.

During the commissioning phase of OLCI-A, SVC implementation could only rely on BOUSSOLE since MOBY was mostly unavailable due to infrastructure failure. Therefore, climatology data were needed to derive an initial OLCI-A SVC gain set, which is clearly not appropriate. During OLCI-B commissioning phase, BOUSSOLE suffered a system failure and was not available, while MOBY was only partly available. Gain derived from climatology data did not improve sufficiently products quality and SVC could not be implemented to OLCI-B prior to Level-2 public release. This situation does not stem from an inherent weakness of either MOBY or BOUSSOLE. It occurred because both infrastructures are aging and did not get in recent years the level of support needed to maintain them at the level of operation needed for regular delivery of SVC data.

Experiences from the early years of OLCI-A and OLCI-B operation have therefore proven that:

- operational SVC infrastructures are mandatory;

- SVC infrastructure shall also be redundant to ensure sufficient data provision in a short period of time;
- ideally, temporary autonomous radiometric systems like ProVal or HyperNav should be deployed in commissioning phase as well as throughout sensors lifetime to ensure sufficient in situ data provision for SVC.

With the Sentinel programme, Copernicus is at the forefront of Earth observation for the decades to come. To make the best use of European Union investments, it is mandatory that Copernicus mobilises resources to secure in the long-term SVC infrastructures as well as operational data validation capacity. The FRM4SOC project has documented general requirements for European SVC infrastructure [34]. In depth requirements for the future SVC infrastructures can be found in EUMETSAT requirements documents [72].

The current Copernicus operational system does not include a robust infrastructure for SVC but instead, relies on the MOBY infrastructure owned and operated by the United States NOAA in Hawaii, Pacific Ocean, and the quasi-operational research infrastructure of the BOUSSOLE buoy in the Mediterranean. This is a significant risk to the performance of Sentinel-3 OLCI Level-2 products in an operational context. It was discussed at the international FRM4SOC workshop on SVC (ESRIN, February 2017, [D-230], [D-240], [26,31]) that:

- Neither MOBY nor BOUSSOLE are directly supported by Copernicus. The risk of losing one or both and their associated expertise, and therefore losing the capacity to deliver robust EO products, must be taken into consideration.
- Assuming that the US MOBY infrastructure is secured in the long term, Copernicus should consider maintaining two operational SVC sites, resulting in a minimum of 3 sites globally. This will ensure system redundancy and robustness of ocean colour SVC as recommended by CEOS.
- Maintaining two sites in Europe will also: secure the existing expertise, knowledge and knowhow in Europe; develop new expertise; stimulate technical, scientific and industrial innovation.
- From a risk mitigation perspective, it is also essential that Copernicus controls (owns) its vicarious calibration capacity to ensure Sentinel-2 and Sentinel-3 product quality for the next two decades.
- For the development of these two proposed Copernicus operational SVC sites, it is clear that building upon existing systems and expertise (namely BOUSSOLE and MOBY) would be more cost effective.

Consequently, the final community recommendation for SVC development within the framework of Copernicus are:

- BOUSSOLE as the existing unique SVC site in Europe must be maintained in the long-term and upgraded to full operational status.
- Development and long-term operation of a second new European infrastructure in a suitable location to gain ideal SVC conditions and ensure operational redundancy is needed.
- For the second European SVC infrastructure, the results of studies to date [75,76] point to a site located in the Eastern Mediterranean Sea, near the island of Crete, as

the best candidate in European waters, although other options (for example in non-European waters) were not excluded at this stage.

- Sound metrological foundation with 'hands-on' involvement of NMIs at all stages of development and operation is a key element.
- FRM ensures SI traceability, full uncertainty characterisation and the best possible accuracy and precision for the SVC measurements and process. Note that the FRM element is limited to the in situ component of the SVC process.
- In situ radiometry should be hyperspectral, high spectral resolution, high quality, and of an SI-traceable FRM nature, with a full uncertainty budget and regular SI-traceable calibration.
- A MOBY-Net system [77], that includes the transportable modular optical system developed by NASA and the MOBY team, could be an option for the new site. It offers a technologically proven system within a realistic timeframe for Copernicus needs and it reinforces collaboration of world class experts and centres of excellence. In parallel, steps should be taken within the frame of Copernicus to develop a European solution in the mid-term.

However, the early stages of an ocean colour mission (typically a commissioning phase of 6 months to one year) generally lack sufficient data to achieve satisfactory gain computation. During the commissioning phase of each OCR sensor, additional temporary systems providing FRM shall be deployed to maximize the number of SVC grade matchups. This could be for instance autonomous systems like ProVal or HyperNav. Such systems have already been deployed and recovered after several weeks or months of deployment, therefore enabling post deployment calibration data quality control and re-deployment.

In-depth requirements for future SVC infrastructures are further described in the EUMETSAT requirements document [72].

11.3.14 Recommendation for maintenance and operations

Routine maintenance is a fundamental part of an operational infrastructure. MOBY has two identical systems that are deployed for a 3 to 4 month period and maintained alternately, therefore, with no overlap. BOUSSOLE performs a bi-annual rotation of the instruments and the upper buoy superstructure as well as monthly cruises where optical measurements are performed, water samples for biogeochemical analysis are taken (chlorophyll monitoring in the first place) and optical heads cleaned up. 4 to 6 month instrumentation rotation has proven to be sufficient in both cases. To ensure regular maintenance, SVC infrastructure shall be located in the vicinity of a harbour to facilitate regular maintenance as well as short notice intervention in case of a system failure, close enough from land to ensure real time or NRT data transmission through GSM or other data link but also far enough away to avoid perturbation by recreational and commercial marine activities.

Maintaining the expertise and therefore securing the human resources shall be a priority.

11.3.15 Recommendation for data reduction and distribution

The overall target combined standard uncertainty shall be of 3 % for L_w in the blue-green spectral regions and 4 % in the red. Data processing shall benefit from state-of-the-art data reduction and quality control schemes.

For traceability reasons, it is recommended that:

- Acquired data should be publicly available together with instrument calibration history.
- Measurement protocols and data processing source codes should be publicly available.

Finalized FRM shall be publicly available together with uncertainty budgets. Time delay for data delivery depends on the operational need and the mission needs. During a commissioning phase, a week between data acquisition and data distribution shall be a target, whereas monthly updates would be sufficient for routine operation.

11.3.16 Recommendation for gain computation

The general principle for vicarious gains computation has been described in Section 11.3.5. Assuming that all required efforts have been made in the generation of FRM of OCR time series (quality control, data reduction, spectral integration, etc.) the following section provides recommendations for the calculation of SVC gains to be applied to satellite time series.

- In situ data shall be converted to the satellite viewing geometry. In the case of OLCI, water-leaving reflectance is provided in the acquisition geometry. This means that FRMs shall be converted to OLCIs acquisition geometry. Uncertainty related to this conversion, which includes using a model of the BRDF, shall be quantified.
- In matchup generation, the time difference between in situ acquisition and satellite overpass shall be minimized. A maximum time difference of less than ±3 h was recommended in [78]. On a system like BOUSSOLE that acquires data every 15 min, the typical difference is less than 7.5 min. This reinforces the need for spatially homogeneous SVC sites to reduce uncertainty linked to time differences.
- Macropixels used to compute SVC gains shall be screened for cloud, glint, haze, white caps, high chlorophyll, etc. over a large area while the SVC gain itself is computed on a smaller macropixel to reduce adjacency effects. In [79] it was recommended to screen over a 15 × 15 macropixel area while the gains themselves should be computed over a 5 × 5 macropixel. Homogeneity of individual gains in a single macropixel shall be analysed.
- The minimum of SVC matchups is determined by the convergence of accumulated mean gains as described in [79]. In practise at least two years of data are needed to achieve stable gains. Experience from OLCI has proven that even using both BOUSSOLE and MOBY was not enough to derive stable gains within two years.

11.3.17 Data validation

OCR products validation shall not be neglected, particularly for operational missions. Having field infrastructures dedicated to SVC, operational services (S3MPC, CMEMS) are often left with little independent (not coming from SVC infrastructure) FRM data for post SVC validation. So far, AERONET-OC has proven to be only reliable source of FRM for routine validation activities, although the AERONET-OC stations are mostly located in coastal regions. Very little data are actually available for EO product validation in the open ocean.

New technologies are currently in development. They include, for instance, autonomous floats like ProVal and HyperNav and fixed systems like HYPERNET and PANTHYR [80]. These systems shall be supported to ensure the provision of FRMs for routine validation.

12 Measurement methods and protocols

12.1 Introduction

The FRM4SOC team was given a task to review the measurement requirements and protocols when operating FRM ocean colour radiometers for satellite validation. In response to the FRM4SOC Statement of Work (SOW) [3] the Technical Report TR-1

"Measurement Requirements and Protocols when Operating Fiducial Reference Measurement (FRM) Ocean Colour Radiometers (OCR) for Satellite Validation" [D-60] [81]

and two scientific papers [82,83] were published. The present chapter gives a summary on these documents. **For citation of the chapter 12**, the papers [82,83] should be considered as preferable references.

12.2 Theoretical background

Satellite remote sensing data are now used routinely for many applications, including monitoring of oceanic phytoplankton in the context of global climate change, detection of harmful algal blooms in coastal and inland waters, management of sediment transport in coastal waters, estuaries and ports, the optimization and monitoring of dredging operations, etc. [27]. To be able to trust and use the remote sensing data, this must be validated, usually by "matchup" comparison of simultaneous measurements by satellite and in situ. The terminology of "Fiducial Reference Measurements (FRM)" was introduced to establish the requirements on the in situ measurements that can be trusted for use in such validation. The defining mandatory characteristics for FRM are described in Section 8 [23–25].

In the following we focus on measurements of the standard Level-2 Radiometric (L2R) product from Sentinel-3, the "water-leaving radiance reflectance", ρ_w, or "directional reflectance", which is defined as:

$$\rho_w(\lambda,\theta,\varphi) = \pi \frac{L_w(\lambda,\theta,\varphi)}{E_d^{0+}(\lambda)}, \tag{7}$$

where $E_d^{0+}(\lambda)$ is the above-water downward irradiance, also called E_s or "surface/reference" irradiance in some studies, and $L_w(\lambda,\theta,\varphi)$ is the water-leaving radiance just above-water in the upward direction measured by the satellite sensor and defined by zenith angle θ and azimuth angle φ. In this terminology, further detailed by [84], the water-leaving radiance is the component of above-water directional upwelling radiance that has been transmitted across the air-water interface or, equivalently, is the above-water directional upwelling radiance, L_u^{0+}, after removal of the downward sky/sun radiance reflected at the air-water interface, L_r :

$$L_w = L_u^{0+} - L_r \, . \tag{8}$$

The latter term is called hereafter "skyglint", but may also include specularly reflected sunglint.

All radiometric quantities in this review are assumed to vary spectrally but for brevity the dependence on wavelength, λ, is generally not represented in the terminology unless cross-wavelength inelastic processes are important.

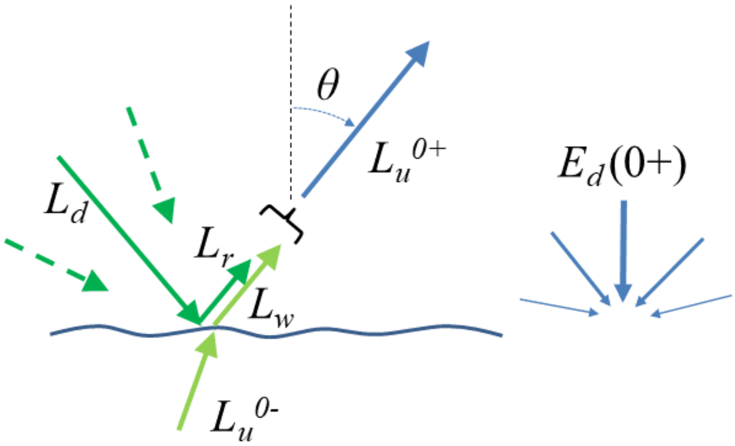


Figure 8. Illustration of definitions of water-leaving radiance, L_w; above and below water upwelling radiances, L_u^{0+} and L_u^{0-}; above-water downwelling (sky) radiance in the specular reflection direction, L_d; above-water upwelling radiance from reflection at the air-water interface ("skyglint") L_r; and downward irradiance, E_d^{o+}. See also [85].

Other missions or processing software may generate alternative L2R products such as normalised water-leaving radiance[2] (nL_w or L_{wN}) [86,87] or remote sensing reflectance (R_{rs}) which can easily be related to ρ_w and/or L_w by simple relationships:

$$R_{rs}(\lambda,\theta,\phi) = \frac{L_w(\lambda,\theta,\phi)}{E_d^{0+}(\lambda)} = \frac{\rho_w(\lambda,\theta,\phi)}{\pi}, \tag{9}$$

$$L_{wN}(\lambda,\theta,\phi) = \frac{L_w(\lambda,\theta,\phi)}{E_d^{0+}(\lambda)}\overline{F_0}(\lambda) = \frac{\rho_w(\lambda,\theta,\phi)}{\pi}\overline{F_0}(\lambda), \tag{10}$$

where $\overline{F_0}$ is the extraterrestrial spectral solar irradiance, which is assumed known to a specified uncertainty, from other studies, e.g. [88], and is possibly used in the generation of the satellite products. In equation (10), following the terminology and reasoning of [89] the viewing zenith angle, θ, and azimuth angle ϕ dependencies are retained. Corrections can then be made to estimate from L_{wN} the water-leaving radiance that would be measured for nadir viewing and in the case of a zenith sun. If such "bidirectional corrections" are made, the resulting parameter will be called "exact" normalised water-leaving radiance, L_{wN}^{ex}, as described in [89] and can be used for consistent time series.

All of these parameters require the measurement of in-water upwelling radiance and above-water downward irradiance. While there may be applications where measurement of upwelling radiance alone may be sufficient, or may be combined with satellite-derived downward irradiance to yield a reflectance product, a FRM L2R product clearly requires both upwelling radiance and downward irradiance to be based on in situ measurements.

[2] Notation for and definition of „normalised" water-leaving radiance may differ between references. In the current review, L_{wN}, is the (directional) <u>normalised</u> water-leaving radiance, as defined in equation (10), whereas the notation, L_{wn}, represents the <u>nadir</u>-viewing water-leaving radiance.

12.3 Previous protocol reviews

12.3.1 NASA Ocean Optics Protocols for Satellite Ocean Colour Sensor Validation

Most of the pre-2004 in situ measurements of water reflectance were made for the purpose of oceanic applications, and most aquatic optics investigators base their measurement protocol in some way on the NASA Ocean Optics Protocols [90] and the references contained within that multi-volume publication.

While there are no fully new methods for the measurement of E_d^{0+} since the NASA 2004 protocols collection, the current review aims to better reflect the current practices.

The main evolutions since 2004 include:

- more frequent use of unsupervised measurements for validation, e.g. AERONET-OC [91] and Bio-ARGO [92], instead of shipborne supervised measurements.
- greater need for validation measurements in coastal and inland waters rather than the prior focus on oceanic waters.
- preference for above-water measurement of E_d^{0+} rather than extrapolation from underwater profiles.
- reduction in the cost and size of radiometers, e.g. facilitating multi-sensor above-water radiometry and reducing self-shading problems for underwater radiometry, facilitating use of an irradiance sensor (instead of a radiance sensor and a reflectance plaque), and better availability of hyperspectral radiometers.

Whereas the methods for measurement of L_w from underwater radiometry using fixed depth measurements or vertical profiles were already well-established at the time of NASA-2004 collection, there has been considerable evolution of methods for above-water radiometry and development of the “skylight blocked approach (SBA)”.

A draft of new Protocols for Satellite Ocean Color Data Validation [93] has been released within the framework of the International Ocean Colour Coordinating Group (IOCCG), providing many updates on the previous NASA-2004 collection.

12.3.2 REVAMP/MERIS protocols

The EU/FP5-funded REVAMP Project (“Regional validation of MERIS chlorophyll products in North Sea coastal waters”) [94] compiled a set of protocols [95] for measurement of apparent and inherent optical properties and optically-relevant biogeochemical parameters (chlorophyll a concentration, total suspended matter).

The REVAMP protocols and the associated documentation of MERIS water products, validation strategies and sampling criteria [96] themselves draw heavily on the NASA Ocean Optics Protocols [90] and on protocols developed in the EU-funded Colors project (Coastal region long-term measurements for colour remote sensing development and validation [97]; funded by the EU Marine Science and Technology Programme MAST III Strategic Marine Research).

Whereas the NASA Ocean Optics Protocols are written generically as far as possible with only a few mentions of specific implementations, the REVAMP protocols are more focused on specific implementations with specific instruments, e.g. the so-called "TriOS method" and "SIMBADA method" are described in the section on above-water radiometry.

12.3.3 MERIS Optical Measurement Protocols and MERMAID database

The "MERIS MAtchup In-situ Database" (MERMAID) is supported by documentation describing the various datasets that have been archived [98]. This documentation describes many radiometric measurements, broken down by principal investigator. The information contains a description of the dataset, e.g. details of measurement locations and deployment methods, and, to different degrees, details of or references to the measurement protocol. Contributors were encouraged to supply information and data values for measurement uncertainty, although in many cases the latter are incomplete or denoted as "not yet available".

MERMAID is specifically designed to facilitate matchup validation for MERIS data and the in situ database is integrated with a tool to allow users to automatically extract MERIS satellite data at the date/time of the in situ measurements and generate "on-the-fly" scatterplots and validation statistics.

The in situ data is supplemented by a standard set of Measurement Quality Control (MQC) flags, denoting quality control checks made by the data provider, and Processing Quality Control (PQC) flags, denoting quality control performed by the database manager.

12.3.4 GLASS and MERIS Lake Water protocol documents

While the sea-going oceanographic community has traditionally been at the forefront of radiometric protocol development and community-wide harmonization, in particular via the NASA Ocean Optics Protocols [90], the inland water community also has significant expertise in aquatic radiometry. The advent of free and high quality data from the USGS/Landsat-8 sensor and the ESA/Sentinel-2 satellites has hugely enhanced the usage of satellite remote sensing for inland waters and generated a parallel need for high quality validation data and supporting protocols.

As an example, the GLASS project [99] collected and tested measurement protocols for measurement of the remote sensing reflectance, including above-water measurements with skyglint correction using a) a handheld 3-sensor system with integrated irradiance sensor [100], b) a single sensor system with reflectance panel measurement for estimation $E_d^{0+}(\lambda)$, c) a TriOS RAMSES 3-sensor system and also d) an underwater radiance measurement. The NASA Ocean Optics Protocols were generally used as guidelines, but the standard procedures were sometimes modified for practical reasons when using small boats.

The protocols used in the GLASS project differed also on the calculation of the R_{rs} from above-water measurements: 1) whether Fresnel reflectance coefficient ρ_F was considered as a constant or dependent on the wind speed, 2) selection of the outliers, 3) whether the fingerprint method [101] and whether the "NIR similarity spectrum" was applied [102].

As another example, the "MERIS Lake Water algorithms" project summarises some protocols used to make validation measurements in inland waters [103].

12.3.5 CEOS INSITU-OCR

The Committee on Earth Observation Satellites (CEOS) set up the "International Network for Sensor Inter-comparison and Uncertainty assessment for Ocean Color Radiometry (INSITU-OCR)" initiative to integrate and rationalize inter-agency efforts on satellite sensor inter-comparisons and uncertainty assessment for remote sensing products with particular emphasis on requirements addressing the generation of ocean colour Essential Climate Variables as proposed by the Global Climate Observing System (GCOS). This working group provides recommendations both on satellite measurements (calibration, development and assessment of satellite products) and on in situ measurements, with special consideration given to traceability, application and accessibility of the in situ measurements that are necessary for any ocean colour mission.

CEOS INSITU-OCR does not specify measurement protocols themselves but has provided in its white paper [1] a set of recommendations that have driven to a large extent the design of the FRM4SOC project SOW. These recommendations are reproduced verbatim in the following subsections in italic text, to denote that this text is not the original work of the FRM4SOC project and its collaborators.

12.4 Broad range of validation conditions

The scope of the current protocols review covers measurements made for validation of water surface radiance/reflectance data derived from calibrated satellite-borne optical sensors after atmospheric correction. This validation must be made over the full diversity of conditions where satellite optical products will be used.

One of the main drivers in development of validation measurement protocols has been the need for highly accurate measurements for the purposes of vicarious calibration with a strong heritage from open ocean measurements. [90]

The accurate measurement of water radiance/reflectance for open ocean waters remains vital for assessing the contribution of phytoplankton processes to the global carbon cycle [104] and for detecting changes in oceanic ecosystems, e.g. related to anthropogenic climate change. However, satellite-borne optical sensors are also used for many other applications in coastal and inland waters, including eutrophication assessment, harmful algal bloom detection, sediment transport, etc. [105].

The scope of the FRM4SOC protocols review on validation measurements is therefore quite different from previous NASA "Ocean" Optics protocol documents.

Whereas vicarious calibration measurements should be made in the best possible measurement conditions (horizontally and vertically homogeneous waters with low temporal variability, low and stable aerosol optical thickness, etc.), validation measurements must cover the full diversity of conditions where satellite optical products will be used, including coastal, estuarine and inland waters and suboptimal water and/or atmosphere conditions, where "suboptimal" means that conditions may not be optimal but the satellite data is still considered as usable and is not rejected by automated quality control procedures.

For example, validation is required for:

- aquatic conditions which include strong horizontal variability (onshore/offshore gradients, patchy waters, etc.), vertical variability (deep chlorophyll maxima, shallow river plumes, thermally stratified waters, etc.) and/or temporal variability (tidal waters, rapid algae blooms/declines);
- diverse aquatic constituents, including phytoplankton-dominated “case 1” waters, but also regions with high terrigenic Coloured Dissolved Organic Matter (CDOM), with non-algal particles, etc.;

- diverse and difficult atmospheric conditions including low/moderate/high and rapidly varying aerosol optical thickness, different aerosol type (marine/urban/dust, etc.) including absorbing aerosols, thin clouds, including cirrus, a wide range of sun zenith angles, conditions when the satellite measurement includes significant sunglint, etc.;
- water surfaces with moderate/high waves (if data is exploited in such conditions), and fetch-limited and/or developing surface wave fields, including estuarine and inland waters;
- locations and sun/viewing conditions with strong adjacency effects, where "adjacency" here refers to violation of typical atmospheric correction assumptions of a horizontally homogeneous water and atmosphere as may occur near land surfaces;
- locations where bottom reflectance may contribute to the water-leaving radiance;
- any other situations where the performance of atmospheric correction algorithms may be different.

Measurement protocols for radiometric validation therefore need to consider all such situations, and the "optimal" protocol may be highly situation- or location-specific. The FRM4SOC protocol review an attempt was made to cover a wide range of potential environmental conditions and a rather generic consideration of the four basic protocol families. For example, the MOBY [59] and BOUSSOLE [69] systems are obvious models for the underwater fixed-depth method and are both operating from platforms in deep, oligotrophic "case 1" waters with high performance and high cost infrastructure and instrumentation. However, the fixed-depth protocol can be applied in very different circumstances such as in very shallow inland waters (with much closer vertical spacing of radiometers) or from ground-fixed platforms (with negligible tilt). Similarly, the AERONET-OC [91] system is an obvious model for above-water radiometry and is characterised by fixed, coastal or offshore platforms with negligible tilt and no azimuthal rotation (of the platform itself). However, the above-water protocol can be applied in very different circumstances, e.g. from ships, or even small boats, with tilt and azimuthal rotation.

In view of the broad scope necessary for validation measurements, terminology specific to "ocean" colour or "marine" reflectance or the "sea" surface is therefore avoided wherever possible in favour of "aquatic", which can include oceanic, coastal and inland waters.

Unfortunately, because of the strong heritage from open ocean remote sensing the "ocean" colour terminology is often difficult to avoid and, for example, appears throughout the ESA SOW where OCR represents "Ocean Colour Radiometry" although the same SOW [3] does point out the importance of Sentinel-2 and coastal and inland waters. The importance and value of the IOCCG in structuring the aquatic optics community also reflects this strong "ocean" heritage.

Similarly the scope is not limited to the dedicated "ocean colour" medium-resolution multispectral polar-orbiting missions such as Sentinel-3/OLCI, MODIS/AQUA and VIIRS, but must consider all present and future satellite-borne optical sensors that are used for aquatic applications, including land-dedicated polar-orbiting missions such as Landsat-8 and Sentinel-2 (and many others), geostationary missions such as GOCI and GEO-CAPE/GLIMR, hyperspectral missions (PRISMA, PACE), etc.

12.5 Approach based on uncertainty estimates

The current protocol review does not try to identify the "best" protocol nor does it aim to prescribe mandatory requirements on specific aspects of a measurement protocol such as "acceptable tilt" or "minimum distance for ship shadow avoidance" or "correct azimuth and zenith angle for above-water radiometry". While such prescriptions have great value in encouraging convergence of methods and in challenging scientists to make good measurements, the diversity of aquatic and atmospheric conditions where validation is required, the diversity of instruments and platforms and the corresponding diversity of measurement protocols suggest that some flexibility may be needed. This flexibility is acceptable **provided that each measurement is traceable to SI and accompanied by an uncertainty budget that is a) based on a full analysis of the protocol and b) that is validated itself**, e.g. by measurement intercomparison exercises [106–109].

For a general treatment of uncertainties in measurements, including a recommended terminology (e.g. "expanded uncertainty") and generic methods for estimating each component uncertainty and combining uncertainties to achieve a total uncertainty the reader is referred to the Guide to the Expression of Uncertainty in Measurement (GUM) of the International Standards Organisation (ISO) [17].

12.6 FRM4SOC structured approach for addressing the methods for L_w and E_d separately

In the NASA Ocean Optics Protocols [90] methods were structured according to whether measurements were made underwater or above-water. Above-water radiometric methods were further grouped into 3 broad classes:

- **Method 1** "Calibrated radiance and irradiance measurements" – one calibrated irradiance radiometer (with a cosine collector head) measures directly E_d^{0+}, and one or two calibrated radiometers measure directly upwelling radiance, L_u^{0+}, and downwelling sky radiance, L_d^{+} (see Figure 8 for definitions). This straightforward method has been implemented by many scientists, e.g. [110,111] etc.
- **Method 2** "Uncalibrated radiance and reflectance plaque measurements" – in this variant on Method 1, the direct measurement of E_d^{0+} by an irradiance radiometer is replaced by a measurement of the radiance reflected from a calibrated reflectance plaque deployed horizontally. This method is typical of the earliest water reflectance measurements, e.g. [112], because of the obvious economy of using the same instrument for all 3 measurements, and is still typical of land surface reflectance measurements, e.g. [113], which are generally supervised.
- **Method 3** "Calibrated surface polarized radiance measurements with modelled irradiance and sky radiance" – in this method the upwelling radiance measurement, $L_{up}(0^+, \theta_v, \Delta\varphi)$, is made by a radiometer equipped with a polarizing filter set to pass only the vertically polarized component of viewed radiance. By viewing at a zenith angle close to the Brewster angle the skylight reflected at the air-water interface is significantly reduced. The measurement of E_d^{0+} is achieved by a direct sun measurement from a sunphotometer and use of a radiative transfer model to estimate E_d^{0+} from the aerosol optical thickness and potentially other auxiliary parameters (atmospheric pressure, cloud cover, etc.). This method was the basis of the specially-designed SIMBAD radiometer [114,115], and was subsequently upgraded to the SIMBADA radiometer.

A number of developments since the writing of this chapter of the NASA Ocean Optics Protocols [116], suggests that these classes of above-water radiometric methods need to be revised, particularly in the FRM context:

- In the FRM context there is really no justification for using uncalibrated instruments and the inclusion of this method with uncalibrated instruments in the NASA protocols contradicts the requirements of the same protocols series that E_d^{0+}, i.e. not just reflectance, is a required radiometric quantity [117] and that instruments should be adequately calibrated and characterised [118]. Method 2 should therefore be at least renamed to reflect that instruments should be radiometrically calibrated, even if the use of a reflectance plaque does effectively reduce uncertainties associated with potential inter-instrument calibration and/or sensitivity differences in Method 1.
- The original Method 3 introduces new ways of measuring both L_w and E_d^{0+} specifically tailored to the hand-held SIMBAD instrument. However, it is quite reasonable to adopt variant or hybrid methodologies with different instrumentation. e.g. direct measurements of E_d^{0+} could be made with an irradiance radiometer, alongside measurements of polarized upwelling radiance, $L_{up}(0^+, \theta_v, \Delta\varphi)$. Measurements could be made of both unpolarized, $L_u(0^+, \theta_v, \Delta\varphi)$, and polarized upwelling radiance, $L_{up}(0^+, \theta_v, \Delta\varphi)$ to better characterise the reflectance of skyglint at the air-water

interface. Sunphotometry derived E_d^{0+} could be combined with unpolarized measurements of $L_u(0^+, \theta_v, \Delta\varphi)$ and $L_d(0^+, -\theta_v, \Delta\varphi)$ as in the AERONET-OC methodology [91].

- To overcome the uncertainties associated with estimation of the skylight reflected at the air-water interface, [119,120], proposed a "Skylight Blocked Approach (SBA)" whereby the water-viewing radiometer is deployed in air, very close to the air-water interface, viewing at nadir, and is supplemented with a "skylight-blocking cone". This method requires careful self-shading corrections [121].
- Further variants on approaches for above-water radiometry render the former Method 1/2/3 structure inappropriate. e.g. the AERONET-OC protocol [91] combines the sunphotometry estimation of E_d^{0+} suggested in the NASA 2003 Method 3 [118], with an unpolarised measurement of L_u .

Moreover for underwater radiometry it is now generally accepted [46] that E_d^{0+} as used in the computation of R_{rs} should always be measured above water[3].

Because of this standardisation of using above-water measurements of E_d^{0+} in all cases, it is suggested here to structure the current document with one chapter for measurement of L_w, with sections for underwater and above-water methods, and one chapter for measurement of E_d^{0+}, relevant for all L_w methods.

This restructuring of the NASA Ocean Optics Protocols is illustrated in Figure 9.

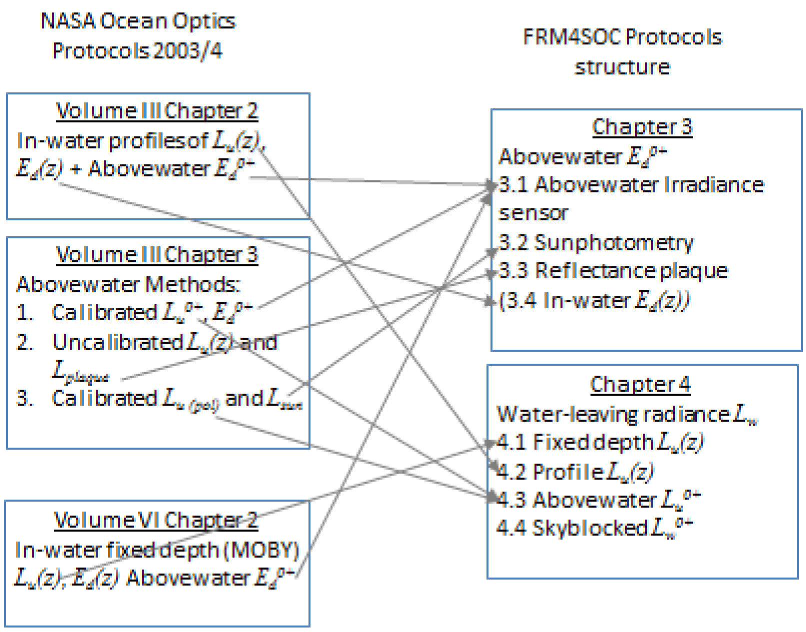


Figure 9. Illustration of restructuring of NASA Ocean Optics Protocols into FRM4SOC protocols review with separate chapters for $\boldsymbol{E_d^{0+}}$ and L_w.

[3] There are still good reasons to perform underwater measurements of $E_d\ (z)$, e.g. for determination of parameters such as the biologically important diffuse attenuation coefficient of downward irradiance, K_d . However, the above water measurement of E_d^{0+} is now always considered the reference for use in computation of R_{rs}.

12.7 Spectral range

Using the terminology of [35], the spectral ranges of primary interest here are the near UV and visible (380 nm to 760 nm) and near infrared (760 nm to 1400 nm) ranges.

The considerations for measurement of E_d^{0+} and L_w given here should be valid also for the near ultraviolet (300 nm to 400 nm) and short-wave infrared (1400 nm to 3000 nm), although the importance of the various uncertainty sources may be different because of the different intensity and angular distribution of downward irradiance and upwelling radiance and the instrumentation (radiance sensor detector and fore-optics) may have different properties in these ranges.

Although L_w is measurably non-zero in the range 1000 nm to 1100 nm in extremely turbid waters [122], L_w will be effectively negligible for the longer NIR from 1100 nm to 1400 nm and the SWIR wavelengths because of the very high pure water absorption in these spectral regions. The need for L_w measurements in the range 1100 nm to 3000 nm is very limited, because satellite R_{rs} data will typically be set to zero during atmospheric correction. However, there may be some interest in this range for quality control of above-water L_w measurements, with non-zero measurement indicating a data quality problem, e.g. skyglint or sunglint contamination or floating material, for the whole spectrum. Also, there may be some interest in the range 1100 nm to 3000 nm for applications such as measurement of floating aquatic vegetation, although this is not strictly speaking L_w and should be measured only using above-water radiometry and without a skyglint/sunglint correction for the percentage of surface covered by vegetation [123].

12.8 A Review of Protocols for Fiducial Reference Measurements of Downward Irradiance for the Validation of Satellite Remote Sensing Data over Water [81,82]

The current state-of-the-art protocols for the measurement of downward irradiance E_d^{0+} for the validation of satellite remote sensing data over water are grouped into three broad families of methods:

- direct above-water measurement of E_d^{0+} with an upward-pointing irradiance sensor ("Irradiance sensor method");
- estimation of E_d^{0+} using a downward-pointing radiance sensor and a reflective plaque ("Reflectance plaque method");
- estimation of E_d^{0+} from direct sunphotometry and a clear sky atmospheric model ("sunphotometry method").

A fourth measurement method that was previously used, estimating E_d^{0+} from the underwater vertical profiles of $E_d(z)$, is now considered inappropriate, and is no longer recommended. This method remains relevant for the measurement of $E_d(z)$ and related parameters such as diffuse attenuation coefficient, but not E_d^{0+}.

The methods are summarized in Table 1 [82], which lists the equipment needed, method variants, and any special issues. The measurement equation and the measurement parameters are described for each family of method in [82].

The elements that should be included for the estimation of total protocol-related measurement uncertainty (Figure 10) are also discussed in [82] with some key considerations, guidelines, and recommendations. Table 2 summarizes the components of the uncertainty estimation giving ideal conditions, recommendations for best practice, and approaches to estimating uncertainty [82]. However, uncertainties arising from radiometer imperfections, such as calibration, thermal sensitivity, spectral response (stray light/out of band effects), non-linearity, and angular (cosine) response must still be added to the overall uncertainty budget.

For the "irradiance sensor" and the "reflectance plaque" methods, the main challenge is to deploy the radiometer/plaque sufficiently high enough to avoid any shading. In this context, "shading" does not only refer to the obvious shadowing of direct solar beam, but also refers to the difference between the unobstructed hemisphere of direct and diffuse sun sky radiance and the reality of measuring in situations where the radiometer/plaque are not higher than all the other structures. For the "irradiance sensor" method, it is also a major challenge to have a sensor that is sufficiently well-designed and well-characterized as regards angular (cosine) response [124].

Table 1. Summary of the three measurement methods as regards equipment, method variants, and special issues. [82]

	Upward-Pointing Irradiance Sensor	Radiance Sensor and Reflective Plaque	Direct Sunphotometry
Equipment	Irradiance sensor (cosine response) Inclinometer	Radiance sensor Reflective plaque Inclinometer	• Sunphotometer (radiance) sensor • Pointing mechanism • Atmosphere radiative transfer model
Variants	• Surfacing of underwater drifting floats. • Shadowband for diffuse/direct.	White/grey plaques	Hand-held or robotic pointing
Other notes		***Note 1***: Uncalibrated radiometers? ***Note 2***: Plaque viewing nadir angle?	

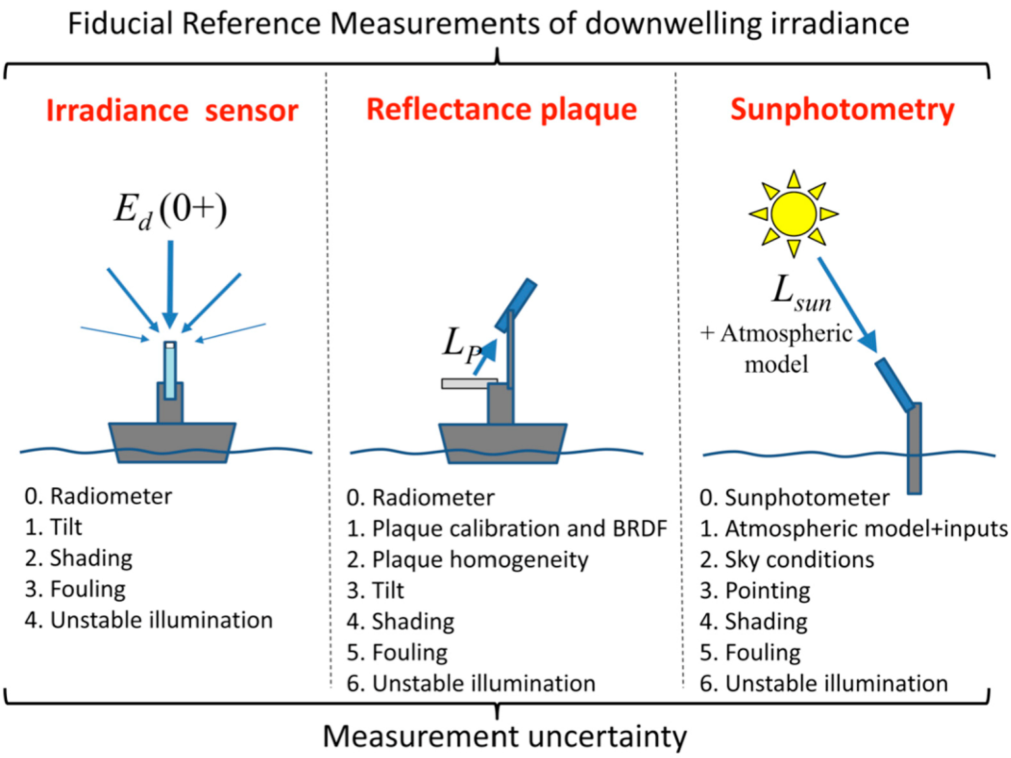


Figure 10. Summary of sources of uncertainty for the three generic families of method for measurement of downward irradiance. [82]

Table 2. Summary of the three measurement methods, including components that must be considered for the uncertainty estimation. Uncertainties arising from radiometer imperfections (calibration and characterisation) shall be still added to the overall uncertainty budget. BRDF: bidirectional reflectance distribution functions; I = Ideal conditions; R = Recommendations; U = Uncertainty estimation; Cal = calibration; FOV = field of view; AOT = aerosol optical thickness; r/t = radiative transfer; S.D. = standard deviation; N/A = Not Applicable. [82]

Method	Upward-Pointing Irradiance Sensor	Radiance Sensor and Reflective Plaque	Direct Sunphotometry
Plaque calibration and characterisation	N/A	I: BRDF-calibrated, homogeneous plaque fills FOV R: Tests to check FOV U: Plaque certificate including BRDF, experiments for homogeneity and height above plaque/FOV	N/A
Tilt/pointing	I: Deploy vertical R: Monitor with inclinometer U: Modelling/experiments	I: Deploy horizontal R: Monitor with inclinometer U: Modelling/experiments	I: Sensor FOV contains and centred on sun R: Small FOV, accurate pointing, check AOT U: Via estimation of AOT
Superstructure shading	I: Deploy above all structures R: Use mast and fish-eye photos U: Experiments (different heights/locations) and modelling	I: Deploy above all structures (except radiometer) R: Use mast and fish-eye photos U: Experiments (different heights/locations) and modelling	I: Clear radiometer–direct sun path R: Check with video surveillance and data QC U: N/A (if not rejected)
Fouling	I: Keep fore-optics clean R: Inspect/clean/protect, monitor with portable cal devices U: Pre-/post-cleaning cal of radiometer	I: Keep radiometer fore-optics and plaque clean R: Inspect/clean/protect, monitor radiometer with portable cal devices U: Pre-/post-cleaning cals for radiometer and plaque	I: Keep fore-optics clean R: Inspect/clean/protect U: Pre-/post-cleaning cals
Fast natural fluctuations	I: Reject if unstable illumination R: Compare replicates/time series U: S.D. of accepted measurements	I: Reject if unstable illumination R: Compare replicates/time series U: S.D. of accepted measurements	I: Reject if unstable illumination R: Compare replicates/time series U: S.D. of accepted measurements
Sky conditions and atmospheric r/t model	N/A	N/A	I: Perfectly cloud-free sky, horizontally homogeneous atmosphere and surface. Perfect r/t model and inputs R: Reject if clouds detected. Intercompare r/t models, check inputs U: Modelling/experiments

Note 1: Is It Necessary to Use a Calibrated Radiance Sensor?

The preparation of this review generated much discussion within the community regarding the question of whether an uncalibrated radiance sensor can be used to acquire measurements for satellite validation. This method was suggested in the NASA Ocean Optics protocols 2003 version "Method 2" [116] as being appropriate for the measurement of reflectance using an uncalibrated sensor. Indeed, R_{rs} can be calculated via Equation (9) from measurements of L_w and E_d^{0+} made by the same radiance sensor, even if this sensor is not calibrated, i.e., providing data for L_w and E_d^{0+} in (dark-corrected) digital counts rather than in SI-traceable units. While it is essential to characterize the sensor, e.g., for stray light, non-linearity, thermal effects, etc., it is not necessary to calibrate the sensor to perform radiometer-related corrections and uncertainty estimates. In fact, some radiometer-related uncertainties are best treated before calibration, e.g. non-linear effects may depend directly on the digital count data [125,126] (as compared to the maximum possible, saturated, digital counts), but not on the calibrated radiance. There is formally nothing in the FRM definition that would require a calibrated radiance sensor to be used for the measurement of R_{rs}.

However, the use of a calibrated radiance sensor does have two advantages:

- A calibrated radiance sensor will provide a calibrated E_d^{0+}, which can then be compared with clear sky models [127] for quality control purposes, and can be compared to satellite data to validate the computations of atmospheric transmittance (in addition to the more important R_{rs} products).
- The interpretation of in situ measurement intercomparison exercises [107], as required by the FRM process, necessitates a separation of uncertainties arising from L_w and E_d^{0+} measurements, e.g. comparing E_d^{0+} measurements from a vertically-mounted irradiance sensor (impacted by cosine angle uncertainties, etc.) with E_d^{0+} measurements deduced from a radiance sensor viewing a reflectance plaque (impacted by BRDF uncertainties, etc.). Moreover, it is noted [128] that the simple cancellation of unknown calibration factors used to calculate $R_{rs} = L_{\mathrm{w}}/E_d^{0+}$ in native spectral resolution no longer works precisely when spectrally convolving L_{w} and E_d^{0+} with a spectral response function, as needed for the validation of R_{rs} for individual spectral bands of satellite sensors.

Note 2: What Nadir Angle Should Be Used for Viewing a Reflectance Plaque?

The NASA 2003 protocols (Volume III, Section 3.3) [116] recommended that measurements of E_d^{0+} with a reflective plaque should be made with a vertical downward (nadir) pointing radiance sensor and a plaque with BRDF calibration for varying downwelling light distributions (typically characterized by sun zenith angle) and vertical upwelling reflected radiance. However, off-nadir viewing with the same nadir angle as water-viewing L_w measurements, typically 40°, has often been adopted for practical reasons, e.g. for easy switching between plaque and water-viewing modes for certain deployments. It is noted that [129] provides the scientific basis for a water-viewing nadir angle of 40° (and relative azimuth to sun of 135°) as a good geometry for sunglint avoidance, but does not give a scientific basis for a plaque-viewing nadir angle of 40° - the latter is merely suggested as practically convenient. On the other hand, an off-nadir plaque-viewing geometry may indeed be desirable for scientific reasons, since the radiometer shading of the plaque will be greater with nadir-viewing when the sun zenith angle is low [128]. For off-nadir plaque viewing, there seems to be no standardization of the viewing azimuth angle, although the same azimuth angle as used for L_{w} measurements (90° or 135° with respect to the sun) would be a typical choice for both practical and shadow-avoidance reasons. Optimal plaque-viewing geometry was investigated in [128] and recommends, for moderate sun zenith angles between 20°- 60°, a plaque-viewing nadir angle of 40° for a ~100 % reflective white plaque was recommended, to minimize operator/radiometer shading/reflection, but a nadir view for less reflective, grey plaques, where reflectivity may vary strongly with the viewing nadir angle. For both types of plaque, a viewing azimuth angle of 90° with respect to the sun was recommended. The FRM context does not prescribe a single viewing geometry (or any other specific aspect of a measurement protocol), but "simply" requires that, for whatever plaque-viewing geometry is adopted, the related uncertainties (radiometer and superstructure shading of plaque, plaque BRDF) be quantified.

Note 3: Irradiance Sensor or Reflectance Plaque?
The preparation of this review stimulated considerable discussion within the community on the pros/cons of the reflectance plaque method as compared to the irradiance sensor method in addition to the question of whether the reflectance plaque method radiance sensor needs to be calibrated (see Note 1.). When correctly applied, the reflectance plaque method can clearly meet the criteria expected of an FRM. However, in practice, this method has often been associated with less rigorous implementation. Specifically, recognizing that the reflectance plaque is performing the same function as the fore-optics of an irradiance sensor, which collects light from the upward hemisphere according to a zenith cosine weighting and directs that light to a photodetector, it is necessary that:

- there be no humans above the level of the reflectance plaque/irradiance sensor (and thereby affecting the sky radiance contributing to E_d^{0+} in a way that is highly variable and essentially not quantifiable in an uncertainty estimate);
- the reflectance plaque/irradiance sensor be mounted as high as possible on the ship/platform, typically higher than any superstructure elements with a significant solid angle as viewed from the plaque/sensor;
- the reflectance plaque/irradiance sensor be mounted on a fixed structure, not hand-held, and associated with an inclinometer allowing the estimation of uncertainties associated with non-horizontal/vertical measurements;
- the measurements made using the reflective plaque/irradiance sensor be supported by experiments and/or simulations to estimate the measurement uncertainties associated with any superstructure shading of the plaque/irradiance sensor.

12.9 A review of protocols for Fiducial Reference Measurements of water-leaving radiance for validation of satellite remote sensing data over water [81,83]

The current state of the art of protocols for the measurement of water-leaving radiance L_w for the validation of satellite remote sensing data over water are grouped into four broad families of method:

- underwater radiometry using fixed depth measurements ("Underwater fixed depths");
- underwater radiometry using vertical profiles ("Underwater profiling");
- above-water radiometry with sky radiance measurement and skyglint removal ("Above-water");
- on-water radiometry with skylight blocked ("Skylight blocked").

The methods are summarized in Table 3 [83], which lists the equipment needed, method variants, and any specific issues. The measurement equation and the measurement parameters are described for each family of method in [83].

The elements that should be included for the estimation of total protocol-related measurement uncertainty (Figure 11) are also discussed in [83] with some key considerations, guidelines, and recommendations. Table 4 summarizes the components of the uncertainty estimation giving ideal conditions, recommendations for best practice, and approaches to estimating uncertainty [83]. However, uncertainties arising from radiometer imperfections, such as calibration, thermal sensitivity, spectral response (stray light/out of band effects), non-linearity, and angular response must still be added to the overall uncertainty budget.

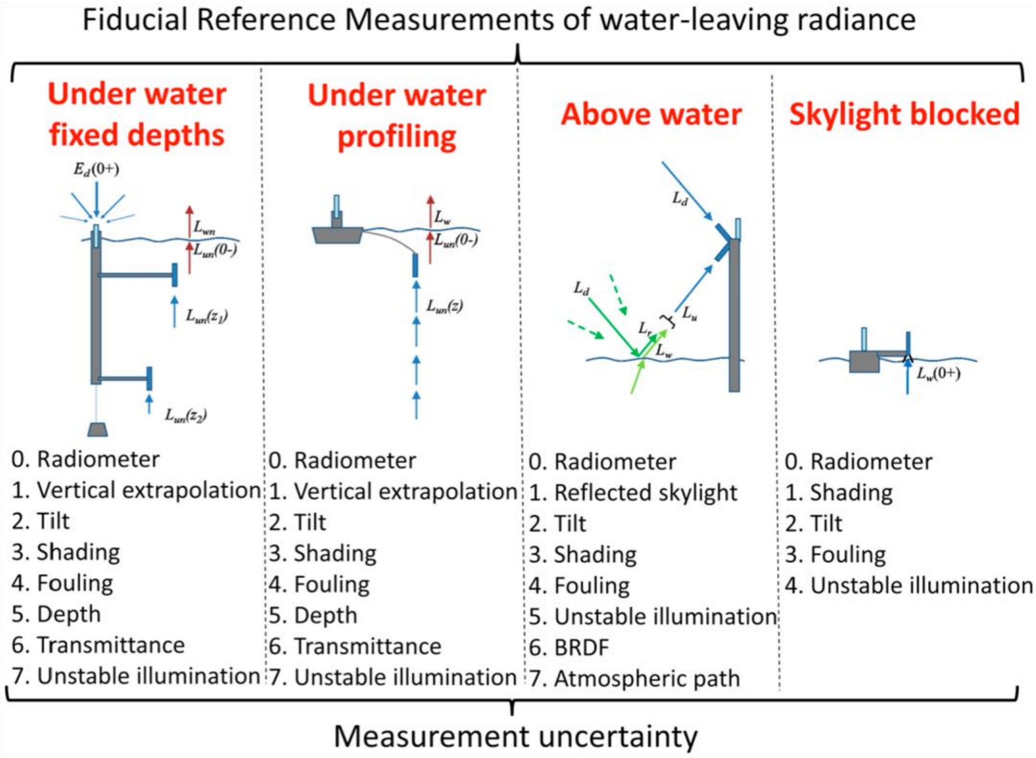


Figure 11. Summary of sources of uncertainty for the four generic families of method for measurement of water-leaving radiance. [83]

Table 3. Summary of the four measurement methods as regards: equipment; standard (S) and variant (V) methods; viewing geometry; protocol maturity/diversity; automation maturity; automation challenges; and challenging waters/wavelengths/conditions (see [83] for more details). The automation challenges refers to the protocol-specific challenges and excludes common challenges such as the logistics of maintenance visits, power supplies, hardware failures, radiometer calibration requirements, protection from damage, etc. [83]

	Underwater fixed depths	**Underwater profiling**	**Above-water**	**Skylight blocked**
Equipment (in addition to ship/platform/buoy)	2 radiance sensors Inclinometer Depth sensor	Radiance sensor and profiling platform Inclinometer Depth sensor	Radiance sensor and robotic/human pointing or 2 radiance sensors Inclinometer, Compass/protractor	Radiance sensor Sky-blocking cone/shield Inclinometer
Standard (S) and Variants (V)	S: tethered buoy, at least two fixed depths V: single very near-surface radiometer; single radiometer successively at different depths	S: free-fall away from ship V: platform/mooring-tethered vertical wire; horizontally drifting platforms	S: unpolarised radiometer V: vertical polarizer option	S: tethered buoy V: boats and other platforms
Viewing geometry	Nadir	Nadir	Off-nadir, usually $\theta_v = 40°$ and $\Delta\varphi = 90°$ or $135°$	Nadir (or off-nadir)
Protocol maturity/diversity	Mature	Mature	Mature basis but also diverse and evolving skyglint corrections	Mature
Automation maturity	Operational	Prototype	Operational	Feasible
Automation challenges	Fore-optics contamination	Fore-optics contamination Mechanical reliability of profiling (fixed location systems)	Fore-optics contamination	Fore-optics contamination
Challenging water types/wavelengths/ conditions	High K_{Lu}(high CDOM/NAP blue, red, near infrared) High waves Very shallow or stratified waters	High K_{Lu}(high CDOM/NAP blue, red, near infrared) High waves Very shallow or stratified waters	Low reflectance (high CDOM blue, low backscatter red/near infrared) High waves Scattered clouds in sky-viewing direction	High waves

Table 4. Summary of the four measurement methods as regards protocol-related uncertainty estimation. I = Ideal conditions; R = Recommendations; U = Uncertainty estimation. Cal = calibration. N/A = Not Applicable. Depth measurement and Fresnel transmittance should also be included in the uncertainty budget for the underwater fixed depth and profiling methods, but are not included in the table. Radiometer-related uncertainties must also be estimated for all methods but are beyond the scope of this review. [83]

	Underwater fixed depths	Underwater profiling	Above-water	Skylight Blocked
Non-exponential vertical variation	I: Known (e.g. exponential) variation R: Extra depths, profiles and modelling U: as R.	I: Known (e.g. exponential) variation R: Measure close to surface U: Goodness-of-fit tests, modelling	N/A	N/A
Tilt	I: Deploy vertical R: Monitor inclination and pressure U: Modelling, time series analysis	I: Deploy vertical R: Stable free-fall or wire-guided, Monitor inclination U: Modelling, time series analysis	I: Accurate pointing, stable platform R: Monitor inclination U: Modelling	I: Stable platform R: Monitor inclination U: Modelling, time series analysis
Self-shading from radiometer	I: Negligible size radiometer R: Small diameter radiometer U: Modelling	I: Negligible size radiometer R: Small diameter radiometer U: Modelling	N/A (in general)	I: Negligible size cone/shield R: Small diameter cone/shield U: Modelling
Self-shading from structure/platform	I: Negligible size superstructure R: Limit cross-section, horizontal arms, redundant radiometers U: Modelling, comparison of redundant radiometers	I: Negligible size superstructure R: Limit cross-section, deploy away from ship, redundant radiometers U: Modelling, comparison of redundant radiometers	I: Negligible size superstructure R: Target away from platform (masts) or ship (forward from prow), azimuth filtering to avoid shadow U: Modelling, experiments (different heights/positions/azimuths)	I: Negligible size platform R: Limit cross-section, horizontal arms, redundant radiometers U: Modelling, comparison of redundant radiometers
Fore-optics contamination	I: Keep fore-optics clean (in water) R: Inspect/clean/protect, monitor with portable cal devices U: Pre-/post-cleaning cal of radiometer	I: Keep fore-optics clean (in water) R: Inspect/clean/protect, monitor with portable cal devices U: Pre-/post-cleaning cal of radiometer	I: Keep fore-optics clean (in air) R: Inspect/clean/protect, monitor with portable cal devices U: Pre-/post-cleaning cal of radiometer	I: Keep fore-optics clean (in air, close to water) R: Inspect/clean/protect, monitor with portable cal devices U: Pre-/post-cleaning cal of radiometer
Temporal fluctuations	I: Clear sky, flat water R: Time series analysis U: Modelling, time series analysis	I: Clear sky, flat water R: Time series analysis, multi-casting U: Modelling, time series and multi-cast analysis	(here for sky, see below for waves) I: Clear, stable sky R: Replicates U: Standard deviation of replicates	I: Clear sky, flat water R: Time series analysis U: Modelling, time series analysis
Skylight reflection correction	N/A	N/A	I: Flat sea R: Very diverse, see text U: Very diverse, see text	N/A

12.10 Underwater or above-water measurement?

So which is the best approach to use? A newcomer to the field of water radiance measurements will typically be confronted with important decisions for:

- purchasing radiometers and associated equipment;
- purchasing, renting or arranging access to a deployment platform such as a fixed structure (offshore platform, jetty, pier, buoy, etc.), a ship (research vessel, small boat, passenger ferry "ship of opportunity", etc.), a drifting underwater platform, or even a low-altitude airborne vehicle (tethered balloon, drone, etc.);
- paying and training staff to make the measurements (if supervised) or to setup and maintain and monitor the measurement system (if unsupervised), including radiometer checks, calibration and characterisation and data processing, quality control, archiving and distribution.

The choice of protocol will affect both the quality and quantity of data and the setting and running costs of acquiring data. The choice of protocol will obviously be driven by the objectives of the measurement program and the environmental conditions (type of water: brightness, colour, depth, vertical homogeneity) as well as by any cost constraints and/or cost-sharing opportunities (such as the existence of platforms or other measurement programs).

The main fundamental differences in data quality that can be expected between the two underwater methods and the above-water (skyglint corrected) method, in their most generic implementations, can be related to the need for vertical extrapolation in the underwater methods and the need for skyglint correction in the above-water method:

- Uncertainties associated with vertical extrapolation in underwater methods will be highest for situations (water types, wavelengths) where the diffuse attenuation coefficient length scale, $1/K_{Lu}$, is small compared to the depth of the highest usable upwelling radiance measurement, z_1. Thus, the requirement for underwater measurements close to the surface becomes more and more demanding for waters/wavelengths with high K_{Lu}, including blue wavelengths in waters with high CDOM or high non-algal particle absorption and red and, a fortiori, NIR wavelengths in all waters. Self-shading also increases for high attenuation waters.
- Uncertainties of the skyglint correction in above-water methods will have a highest impact on the derivation of the water reflectance for low reflectance waters/wavelengths and for high sun zenith angle (as well as for cloudy and partially cloudy skies although these are supposed to be removed by quality control in the FRM context) and for blue wavelengths. Thus, the requirement for a highly accurate skyglint correction method becomes more and more demanding for blue wavelengths in waters with high CDOM absorption (and to a lesser extent high non-algae particle absorption) and for red and NIR wavelength in low particulate backscatter waters.

It is interesting to note that these two challenging conditions, high K_{Lu} and low reflectance, generally correlate in highly absorbing waters/wavelengths but anticorrelate in highly scattering waters.

Both the underwater methods and the above-water methods have uncertainties that increase with surface wave conditions because of wave focusing/defocusing effects and skyglint removal respectively.

The skylight blocked approach has quite different sensitivity to the water type and wavelength of measurement from the underwater and above-water approaches, because it requires neither vertical extrapolation nor skyglint removal. The most challenging conditions for this method will probably be practical deployment in high wave conditions and self-shading correction for low sun zenith and high K_{Lu} conditions.

12.11 Future perspectives

There has been considerable evolution and diversity of the L_w and E_d^{0+} measurement since the publication of the NASA Ocean Optics Protocols [90].

Future improvements to L_w and E_d^{0+} measurements are expected to come from the following developments:

- Improvements in the design and usage of calibration monitoring devices, which can be used in the field, are likely to improve the identification of fore-optics fouling and radiometer sensitivity changes.
- Model simulations of the 3D light field and dedicated experiments for the described protocols are likely to improve estimations of related uncertainties.
- Improvements in the stability and reduction in the cost of telescopic masts may allow developing of their use and therefore reduce superstructure shading effects.
- Reduction in the cost of pointing systems, thanks to the video camera surveillance industry, should facilitate multi-directional above-water radiometry [130] and improve the protection ("parking") of radiometers when not in use and thus reduce fouling for long-term deployments.
- Improvements in active gimbals might reduce the tilt effects for the irradiance sensor method and also for en-route above-water radiance measurements
- Greater use of full sky imaging cameras [131], whether calibrated (expensive) or not (typically inexpensive), potentially coupled with automated image analysis techniques, will allow better identification of suboptimal measurement conditions.
- Above-water imaging cameras may allow better characterisation of the air-water interface (wave field) and hence better removal of L_r in above-water radiometric measurements [132,133].

As regards to the future for the validation of water reflectance more generally:

- The tendency to move to highly automated systems with long-term, e.g., one year, essentially maintenance-free deployments is likely to significantly improve the quantity of data available for validation. Networks of such systems further increase the power and efficiency for validation purposes. Networks of automated systems are now already operational or in advanced prototype testing phases for systems based on the above-water, underwater profiling and underwater fixed depth methods and are conceptually feasible for the skylight blocked approach.

- The advent of operational satellite missions such as JPSS/VIIRS, Sentinel-3/OLCI, Sentinel-2/MSI, and Landsat-8/OLI with the need for a guaranteed long-term validated data stream will increase the need for FRM.

- The huge increase in optical satellite missions used for aquatic remote sensing will also increase the need for highly automated measurement systems and the economy of scale for such deployments – one in situ radiometer system can validate many, many satellite instruments.

As regards to the needs of the validation community, it is recommended to:

- Update this review, e.g., on a 10-year period, to take account of developments in the protocols, particularly in the estimation of uncertainties. Such an update is best preceded by community discussion at an international workshop.

- Organize regular, e.g., on a two-year period, intercomparison exercises to ensure that measurement protocols remain state of the art and scientists adopt them (as required by the FRM context).

Although not targeted by this review, it is possible that the considerations developed here may be useful for other applications where E_d^{0+} measurements are needed, including the validation of satellite-derived photosynthetically available radiation products [134], the validation of surface reflectance over land, and the monitoring of solar irradiance for the solar energy industry, for agriculture, for the building industry, for the estimation of the Earth's radiation budget, and absorbing atmospheric gases, etc.

In addition to the guidelines provided by the protocols themselves, there are some key recommendations from them for teams participating in satellite ocean colour validation activities that need to be considered when attempting to achieve FRM status for their measurements:

- Analyse carefully their present measurement protocol and construct an uncertainty budget including minimally the elements listed in the corresponding sections of the FRM4SOC protocols [81–83];
- Participate in intercomparison exercises to validate their uncertainty estimates against those of other methods/scientists;
- Consider the IOCCG/CEOS INSITU-OCR White Paper [1] and the FRM4SOC protocols [81–83] and provide comments for its improvement.

Further, it is recommended to ESA and other space agencies to:

- Facilitate discussion and adoption of best practice and uncertainty estimation by sponsoring intercomparison exercises with appropriate funding for post-measurement analysis of results;
- In the medium term encourage and stimulate the adoption of FRM requirements and in the long term, when sufficient progress and consensus is achieved, use only FRM for the routine validation of satellite ocean colour data.

Finally, it is recommended to the IOCCG:

- To adopt a terminology that reflects the generic nature of aquatic optical processes: "air-water interface" instead of "sea surface", "water colour/reflectance" instead of "ocean colour", "aquatic/water optics protocols" instead of "ocean optics protocols", etc.

13 Review of instruments used for Satellite Ocean Colour radiometer validation

13.1 Introduction

Confidence in data records is built by independent validation of measurement results and by establishing a traceability chain to the units of SI with accompanying uncertainty evaluation. Measurement instrumentation has a major role in this chain and therefore, it is of prime importance to estimate the uncertainties of measurements used for satellite Ocean Colour Radiometer (OCR) validation.

The FRM4SOC team, reflecting the FRM4SOC SOW [3] more generally, considers it of prime importance to estimate the uncertainties in measurements for satellite OCR data validation. This total uncertainty estimate includes components arising from: the type of instrument used, the instrument calibration, the protocol and data processing methods, and the spatio-temporal characteristics of the satellite-ground "matchup" measurements.

The present chapter focuses on the radiometers used for the in situ measurement as the summary of the FRM4SOC Technical Report

"A Review of Commonly used Fiducial Reference Measurement (FRM) Ocean Colour Radiometers (OCR) used for Satellite OCR Validation". [D-70] [54]

The aim of the current review [54] as defined by the FRM4SOC SOW [3] is to :

- Document the different designs and performance of Ocean Colour[4] Radiometers (OCR) commonly used for satellite validation including a review of their known characterisation (e.g. immersion factor, cosine response, linearity, stray light, spectral, temperature sensitivity, dark currents etc.) and identify significant issues to address.
- Highlight the technical strengths/weakness of each system.
- Building on available material, include a dedicated section on instrument characterisation and identify issues that must be addressed for each OCR system.
- Conclude with a justified set of actions to assure that each OCR used for satellite validation attains FRM status.
- Include any other aspect considered relevant to the FRM concept.

It is important to note that the review does not try in any way to identify a "best" instrument. This would be an impossible task, especially since there are no objective criteria on which to define "best" and, as will become apparent on reading this review, there is by no means sufficient information to perform any kind of fair comparison. [54]

[4] In compliance with the SOW and because of the strong heritage from the oceanographic optics community the terminology "Ocean Colour" is used throughout this document. However, it is noted that the exploitation of optical data from satellites is no longer restricted to oceanographic applications. The exploitation of optical data from coastal and inland waters implies a corresponding need for validation of such data. The terminology "Water Colour Radiometry/Radiometers (WCR)" would therefore be more appropriate. Clearly the same instruments are used for all water reflectance measurements and the contents of this report are equally applicable to measurements of oceanic, coastal and inland waters.

The review also does not document instrument costs and certainly does not attempt to evaluate "value for money", which is important for users but clearly outside the scope of our study. **Rather the objective of the present report is to document what is already known about the performance of the various instruments, according to traceable references, and identify what still needs to be characterised in order for validation users to construct a full uncertainty estimate for instrument-related factors** [54]**.**

13.2 Methodology

A list of instruments to be included in this study was compiled by checking the various team contributions to the Sentinel-3 Validation Plan and to the prior ESA/MERMAID protocols document, by searching relevant web sites, including those of NASA and NOAA, and by personal knowledge of scientists active in OCR validation work. An example of various types of radiometers in use is presented in Figure 12, [135].

Figure 12. Most of the commonly used OCR types were present in the FRM4SOC comparisons. An example from the FRM4SOC comparison experiment LCE-2.

Contact was then made by email with the manufacturers of currently available Commercial Off the Shelf (COTS) radiometers and followed-up by email/phone/WebEx. In the case of the manufacturers based in Europe (CIMEL, TriOS, Water Insight), one day site visit was made to discuss their instruments and to clarify information. The information collected on the basis of these contacts and of independent web-based search (peer-reviewed publications, technical reports, product data sheets, etc.), has been compiled into a standardised format and style. Where possible, documentation on tests performed by independent or semi-independent scientists published in peer-reviewed literature is preferred. However, it is clear that much information comes from the instrument manufacturer itself or from sources close to the manufacturer. It is left to the reader to assess the impartiality of any sources of information.

All manufacturers received first a draft of the information pertaining to <u>their own instrument</u> for comments and check for correctness and completeness. Manufacturers were informed that this report will be made public and were therefore warned that confidential information should not be communicated to the FRM4SOC team. It is, therefore, possible that more information exists for the characterisation of some instruments, but it has not been possible to include it here because of proprietary concerns. This approach of excluding confidential information that cannot be checked is entirely consistent with the FRM4SOC philosophy that uncertainty estimates should be based on traceable and open documentation.

Finally, all manufacturers were invited to comment on the complete report both by email and also by the physical/teleconference seminar. The seminar was held at ESTEC on 6 September 2017 (Figure 13) with physical participation from Water Insight and CIMEL as well as by tele-conference from IMO, Satlantic (Sea-Bird Scientific), and TriOS.

Figure 13. First round table seminar of the manufacturers of OCR at ESTEC on 6 September 2017. [54,136]

To our knowledge, it was also the first time when major manufacturers gathered to discuss uncertainty evaluation issues and common approaches for the characterisation of OCR. It is notable that manufacturers, despite the strong competition in the field, all considered it important to contribute together to the reliability of measurement data. The manufacturers stated that the gathering helped them to learn about the best measurement methods and see how uncertainties are taken into account. They also noted that it provided them with ideas of improving the measurement protocol and perhaps moving towards a new generation of instruments. The participants also benefited from the gathering by widening their network. It was pointed out that the exercise might be used as prime example for similar activities.

It is clearly considered important by the FRM4SOC team to gather information on all possible OCR instruments used for the validation of the Sentinels and of all other OC satellites. It is hoped that this FRM4SOC information-gathering activity will stimulate manufacturers and scientists to investigate in more detail the characterisation and uncertainty sources of their OCR instruments and hence improve the basis for FRM uncertainty estimates accounting for all uncertainty sources.

13.3 Definition of Radiometer Characteristics

13.3.1 Spectral response function and wavelength calibration

For hyperspectral spectrometer-based instruments the Spectral Response Function (SRF) is generally defined via the wavelength range and typical Full Width at Half Maximum (FWHM) of the spectral response function for each detector/pixel. The latter may, in reality, vary across the spectral range, however, the full dispersion relationship is not generally documented.

For multispectral filter-based instruments, the spectral response function is generally defined by the central wavelength and FWHM of the spectral response function for each (detector/filter) band.

The SRF of nearly all instruments is quite symmetrical and generally Gaussian (typical of spectrometers) or almost square (typical of filters). Apart from distinct stray light/out of band responses, which are considered separately in the following section, it is considered to be generally sufficient to know the central wavelengths and spectral width (FWHM).

Some details are given, where known, on wavelength calibration performed by the manufacturer of the instrument or by the manufacturer of components (filters, spectrometer) and in one case (WISP-3) a portable device for checking wavelength calibration is mentioned.

13.3.2 Spectral stray light/out of band response

Imperfections in instrument design and construction may lead to photons of one wavelength reaching the detector for a different wavelength. For the hyperspectral spectrometer systems this is generally termed as “(spectral) stray light” and can be characterised by illuminating the instrument in a laboratory with a tuneable monochromatic light source, scanning the necessary wavelength range, which may include wavelengths outside the nominal spectral range of the entire instrument (e.g. UV). For the filter-based systems such cross-wavelength effects are generally termed as “out of band response”, but can be characterised in the same way as for spectrometer-based systems.

Information on spectral stray light/out of band response is also available in some cases for certain critical instrument components (spectrometer, filter), although the full instrument is preferred, where available.

13.3.3 Radiometric calibration and immersion factor

Radiometric calibration consists of determining the conversion coefficients to transform the electrical signal recorded by an instrument into an absolute measurement of light, either radiance or (cosine) irradiance and is generally achieved by illuminating the radiometer in air with a light source of known intensity, traceable to an optical radiation primary standard, typically a cryogenic radiometer operated by an NMI such as NPL or NIST. Incandescent FEL type lamps (1000 W) are typically used for these “factory” calibrations of irradiance sensors, combined with calibrated diffuse reflectance plaques for the radiance sensors. Laboratory radiometric calibrations are discussed in [137]. For the present document, the scope is limited to giving a brief indication of current practice for the respective instrument manufacturers.

In general, such a radiometric calibration is performed by the instrument manufacturer on supply of a new instrument and, on request from the user, is typically repeated at annual intervals along with a general maintenance check of the instrument. Alternatively, users who are suitably equipped with their own calibration laboratory can perform the radiometric calibrations.

Radiometers which are operated underwater need also to be calibrated for the situation where the instrument fore-optics is in contact with water instead of air giving a typical decrease in responsivity of 40% for irradiance sensors and 70% for radiance sensors [46]. This effect is generally characterised by "immersion coefficients" to convert in-air calibration to in-water calibration. As explained in [46] and supporting references, the responsivity decrease for **irradiance** sensors is related to the reduced transmittance of the water-diffuser interface compared to an air-diffuser interface and can be measured in the laboratory with suitable equipment (water tank, stable light source). The resulting "immersion coefficients" needs to be measured for each sensor individually. As explained by [46] and supporting references, the responsivity decrease for **radiance** sensors operated underwater is primarily influenced by the decrease in solid angle field-of-view and additionally by an increase in the transmittance of the optical window when in water as compared to air. The corresponding immersion coefficient can be estimated theoretically from knowledge of the refractive index of the optical window [138,139], or, for higher accuracy, can be measured in the laboratory with suitable equipment, see [140] for an example. In contrast to irradiance sensors, the immersion coefficients for radiance sensors generally show less sensor-to-sensor differences for sensors from the same series [46]. Studies on immersion factors in further detail are reported in [54].

In addition to these typically annual calibrations, it is highly recommended that scientists check regularly, e.g. at the beginning/end of each measurement campaign, the radiometric stability of their instruments to reduce the uncertainty associated with responsivity changes between the time of absolute radiometric calibration and the time of measurement. In some cases (TriOS RAMSES, WaterInsight/WISP-3) the instrument manufacturers offer also portable light sources, fitting directly to the instrument, to facilitate rapid and frequent checks on the relative sensitivity of each sensor, e.g. Figure 14. The existence of such portable devices is reported in [54].

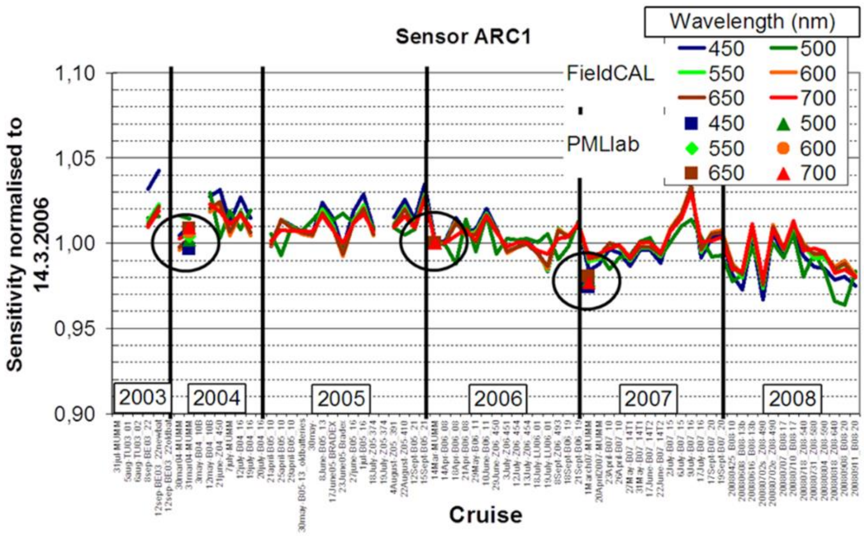


Figure 14. Example of a relative calibration time series for an RBINS-owned TriOS RAMSES instrument, combining absolute laboratory calibrations (symbols inside circles) with relative calibrations made at the beginning and end of each cruise using a portable calibration device. [54]

13.3.4 Radiometric Noise

The electrical signal generated within an instrument may contain various components which are not related to the incident light level.

The "dark current" which is generated even when all incident light is blocked, e.g. by an internal shutter or external lens cap, is often decomposed into a constant component, which is removed by the absolute calibration "offset" calibration, and a time-varying component, which can be removed by frequent "dark" measurements, e.g. automatically performed via an internal shutter or opaque filter (in a filter wheel design) or in some spectrometers, by permanently non-illuminated "black" pixels, or manually performed by placing a lens cap over the fore-optics. The dark current is generally sensitive to the internal temperature of the detector and associated electronic circuitry and can therefore have significant temporal variability for instruments which do not have internal temperature regulation. This temporal variability of dark current depends both on the ambient temperature and on the thermal inertia of the instrument and temperature-sensitive electronic components such as spectrometers – see [141] for a detailed description of dark signal temperature dependence for two spectrometer modules, one with and one without internal cooling.

In addition to these removable components of the electrical signal there will also be effectively random noise arising from optical and electrical processes, which are faster than the frequency of automatic/manual dark measurements. This random noise is often expressed by quoting a signal to noise ratio (SNR) for an instrument perhaps at a specified light level, although it is noted that the SNR is highly dependent on the incident light (and integration time of the instrument, if variable) and so should be determined at different light levels [46].

Finally, digitisation effects can arise from the discrete nature of recording raw data as digital numbers (integers) to be later converted into radiances using calibration coefficients. The level of digitisation, generally expressed in bits, will often be related by instrument design to the expected instrument noise level. These factors are reported in [54].

13.3.5 Radiometric linearity

Absolute radiometric calibration is generally performed at a single illumination intensity (in addition to the zero illumination dark condition) yielding a single "slope" calibration coefficient (in addition to the dark "offset" coefficient). However, the responsivity of instruments cannot be perfectly linear over all possible illumination intensities. Specifically the intensity that is used for calibration and the intensity during the measurement are different and non-linearity represents a source of additional measurement uncertainty, which needs to be estimated. CCD detectors used in hyperspectral instruments may have a slightly non-linear response, which is sufficiently constant in time as to be correctable, allowing reduction of the associated uncertainty. Other photodetectors or associated electronic circuitry may also have non-linear response, particularly if operating close to the maximum "saturation" light level.

Radiometric non-linearity and any procedure used to correct for this are reported in [54]. Such characterisation may be available only for certain critical system components, such as the photodetector, but should ideally be validated at the full instrument level, e.g. by laboratory tests at different, carefully controlled light intensities.

13.3.6 Thermal stability

Electrical components such as photodetectors and associated circuitry including Analogue to Digital Converters (ADC), may be significantly affected by temperature variations both via the dark current and via temperature-dependent responsivity/sensitivity, corresponding to the offset and slope coefficients derived from absolute calibration [54].

If the thermal variability of responsivity is sufficiently characterised, e.g. by laboratory tests made in a thermally controlled environment, a correction for this effect can significantly reduce the associated measurement uncertainties of an instrument. Uncertainties associated with thermal effects may also be reduced in instrument design via thermal regulation of the instrument itself, e.g. OSPREY [142], although this is not common in COTS instruments because of the associated construction and hence purchase costs as well as power requirements for autonomous deployments.

Thermal effects also depend on the range of ambient temperatures (generally larger in air than in water) although the relationship between the internal temperature of components such as photodetectors and the ambient temperature of the air/water environment in which an instrument is being used may be quite complex because of time lag effects associated with thermal inertia. In some instruments, e.g. CIMEL SeaPRISM and Satlantic (Sea-Bird Scientific) OCR500, there may be a measurement of internal instrument temperature, preferably close to the most thermally sensitive components. In other instruments it may be possible to estimate internal temperature or otherwise characterise thermal effects by analysis of dark current and/or noise [143].

13.3.7 Polarisation sensitivity

Most OCR validation studies are performed by comparing reflectances or radiances derived from the (scalar) intensity of light measured at water-level with the intensity of light deduced from satellite measurements, because these are the products required for derivation of nearly all ocean colour products. However, a full description of the light field should include the polarisation properties, as expressed via the Stokes vector. This becomes important in the OCR validation context when the light that reaches an instrument is significantly polarised, e.g. the sky viewed at 90° from sun or Fresnel reflection from the sea surface close to the Brewster angle (~53.3° incident angle for seawater viewed from air), and when the instrument is sensitive to this polarisation.

Certain instrument components, such as mirrors, gratings, slits or beam-splitters may generate significant polarisation sensitivity, while others, such as diffusers used in irradiance sensors, or fibre optics may reduce polarisation sensitivity.

Polarisation sensitivity can be measured in the laboratory by viewing a polarised light source, such as an FEL lamp viewed through a polarising filter with well-characterised properties, at various azimuthal rotation angles. See [144] for a description of experiments to determine polarisation and rotational uncertainties.

13.3.8 Angular response

For instruments measuring downward irradiance using a flat (cosine) collector head the angular response of the instrument can have a significant departure from the perfect cosine function and hence be a significant source of uncertainty. Characterisation of the angular response is therefore important and it may be possible to reduce measurement uncertainties by correcting to some extent [145] for an imperfect angular response using data from the characterisation and from the angular distribution of radiance at the time of a measurement, e.g. as expressed by a proxy such as average cosine of downward irradiance or direct/diffuse irradiance ratio or as measured by a hemispherical radiance camera. The angular response of an irradiance sensor is strongly affected by the material of the diffuser head and by the detailed internal geometry of fore-optics and any baffling elements. Characterisation of the angular response is generally performed in the laboratory by illuminating the instrument with a collimated light source, taking measurements for different angles of incidence.

For instruments measuring radiance the angular response generally has, by design, a narrow field of view (FOV), typically less than 10°, and very sharp angular cut-off with very little light reaching the detector for angles outside the stated FOV. FOV is typically defined by FWHM. Angular stray light from outside the FOV is not thought to be a significant source of uncertainty for radiance sensors viewing water or sky, although for completeness some tests should be performed and documented[5].

FOV itself is an important factor for water-viewing radiance instruments because this, combined with the distance of the instrument from the water target, determines the surface area of the water target and hence the extent to which spatial variability of surface waves will be resolved or averaged. Specifically, the accurate removal of sunglint from above-water radiometric measurements is facilitated by using an instrument with a small FOV (e.g. 2°) and rapid sampling (e.g. >1 Hz).

For sunphotometer radiance instruments used to estimate downward irradiance the angular response is much more important because of the very high angular variability of light coming from near the Sun's disk.

13.3.9 Instrument Operations

In addition to the main radiometer characteristics described above, there are also additional parameters that may have effect on measurements during instrument operations. These aspects include:

- **Instrument power supply**: Instrument performance may depend slightly on whether mains electricity or internal batteries are used and on the stability of such power supplies.
- **Instrument warm-up**: Some instruments may require some time to reach a stable operating temperature in the field [146]. The thermal characterisation once a stable temperature in equilibrium with the ambient temperature is reached has been considered previously in Section 13.3.6, but assumes that equilibrium has been reached.
- **Instrument cleaning**: Instrument characterisations are made in laboratory conditions with well-cleaned fore-optics. Field measurement protocols should describe instrument cleaning to achieve comparable performance. In the case of automated, unsupervised deployments, typically lasting days or months, post-deployment calibration should be made both before and after the cleaning of fore-optics to assess associated uncertainties. [81]

[5] Angular stray light from outside the nominal FOV should also be avoided during radiometric calibration of radiance sensors when viewing a diffuse reflectance plaque and the detailed angular response function may be relevant for determining calibration uncertainties if a diffuse reflectance plaque is non-uniformly illuminated (e.g. by a point source of light).

- **Instrument storage and transport**: Optical instruments are particularly prone to mechanical shock during transport because of sensitivity of measurements to the precise alignment and spacing of optical elements. Instruments should preferably be hand-carried whenever possible, although air and road freight is often necessary. Specially designed transport cases should be used to optimally protect from mechanical shock. Transport cases can be supplemented by "shock indicator" devices, which record when a box has been subjected to mechanical shock during transport (and which also act as a visible warning to transporters that shock should be avoided and a deterrent that shock will be noticed). Further complications may arise for instruments with internal batteries because of airline safety regulations. In addition to the use of shock indicators it is good practice to perform at least approximate calibration checks on instruments after transport, e.g. by use of a portable calibration device (if available) and/or by intercomparison of sky measurements with other instruments if post-transport absolute radiometric calibration in a laboratory is not feasible.

The user manual supplied with COTS instruments will typically cover the abovementioned aspects. Clearly, the specific recommendations of the manufacturer should be followed to achieve expected performance.

13.4 Summary of knowledge of all available systems and identification of gaps in knowledge

An overview of the systems/instruments documented in the report in [54] is given in Table 5.

Table 5. Key characteristics of the instruments and systems described in this review.

Manufacturer/ Instrument	Type	Deployment	Wavelength range	Spectral width (FWHM)	Radiance FOV in air (FWHM)
Biospherical C-OPS	Multispectral underwater system, L_u, E_d	Underwater, ship-tethered, slow free-fall	(305...1100) nm, (1100...1650) nm available with InGaAs detectors	10 nm	7°
CIMEL SeaPRISM	Multispectral system, sun/ sky radiance L_{sky}/ L_W	Above-water (fixed platform)	(412...1020) nm	(8...10) nm	1.3°
IMO DALEC	Hyperspectral system, E_d/L_{sky}/L_W	Above-water (ship)	(305...1050) nm, calibrated in (400...900) nm	10 nm	5°
Satlantic (Sea-Bird Scientific) HyperOCR	Hyperspectral instrument, L or E_d	Above-water, Underwater	(305...1100) nm, calibrated in (350...800) nm	10 nm	6°/23°
Satlantic (Sea-Bird Scientific) OCR500	Multispectral instrument, L or E_d	Above-water, Underwater	(380...865) nm, optional from 305 nm	10 nm (or 20 nm)	28°
TriOS RAMSES	Hyperspectral instrument, L or E_d	Above-water, Underwater	(320...950) nm	10 nm	7°
WaterInsight WISP-3	Hyperspectral system, E_d/L_{sky}/L_W	Above-water (handheld)	(380...800) nm	4.9 nm	3°

Even in cases where some information exists, e.g. stray light distribution function for the TriOS RAMSES and WISP-3 radiometers, this information may be limited to a single unit or a few units and will not necessarily be applicable to all instruments of this type being used for OCR validation. The question of transferability of stray light distribution function across units of the same class of radiometer is discussed by [147].

The detailed findings of the original FRM4SOC report [54] are not reproduced here because they may already be significantly out of date and hence misleading. The original report was written mainly in 2016 and was finalised in 2017. In 3-4 years there can be significant progress in instrument characterisation as well as new instruments or evolutions of the instruments already considered.

13.5 General conclusion

To our knowledge, the report [54] is the first attempt that has been made to compile information on all commonly used OCR to the level of detail that is required to construct a full uncertainty budget for instrument-specific aspects. This level of detail far surpasses the information that is generally made publicly available, e.g. on manufacturer websites. In many cases, sufficient information is just not available. In some cases, manufacturers have indeed performed radiometer characterisation tests but the information is not publicly available and/or is considered confidential, which is contrary to FRM requirements.

It is not our intention, and in fact would be neither feasible nor deontological, to recommend a "best" OCR or, a fortiori, a "best value for money" OCR. It is for the OCR users, as customers, to make such decisions. However, we hope that the present report will help understand what information is or currently is not available for preparation of a FRM uncertainty budget, so that these users will be able to make informed purchase decisions and request the relevant information on radiometer characterisation from their suppliers.

Similarly, this process should reward the efforts of the most conscientious instrument manufacturers, who perform careful characterisation tests and provide this information to their customers and to the scientific public and space agencies that use data from these instruments for satellite validation purposes.

This report is, therefore, a step in a process[6] towards FRM and will, hopefully, be followed up by activities of the instrument manufacturers, the FRM4SOC project team and individual validation scientists to better understand the performance and uncertainties of OCR.

13.6 The "missing" instruments

While the report [54] endeavours to include the OCR that are currently available as COTS units and are commonly used, e.g. within the Sentinel-3 Validation Team (S3VT) for satellite OCR validation, a number of "missing" instruments are mentioned here for completeness and for possible future updates of the current document:

- Many "legacy" OCR instruments have been discontinued by the respective manufacturers because improved instruments are now available, but are still in regular use by validation scientists. An important example is the Satlantic (Sea-Bird Scientific) OCR200 instrument family.

[6] The importance of the precursor NASA Ocean Optics Protocols Volume II Chapter 3 [137] is noted as the first important step in this process.

- The ASD Fieldspec instruments have not been included in the present report. Their use for satellite validation studies over water seems very limited although could warrant inclusion in an update of this report, e.g. if needed by the S3VT.
- The (non-commercial) SIMBADA handheld sunphotometer and radiometer [115] has been manufactured in quite a few units and has been used in the past for satellite validation studies. Its current usage is very limited although it could warrant inclusion in a future update of this report if usage becomes more widespread.
- The Biospherical/OSPREY above-water system uses some of the technology described already in the report on the Biospherical/C-OPS underwater system and its characterisation is described in detail in [148]. It is not clear if this instrument is in COTS production, although it could warrant inclusion in a future update of this report if usage becomes more widespread.
- The MOBY optical system [59] consists of two holographic reflective grating spectrometers which are integrated within a 14.5 m buoy/spar structure with multiple radiance and irradiance collectors at different depths connected to the spectrometer systems by fibre optics. The system does not consist of individual COTS ocean colour radiometers, which can be used independently, as in the OCR covered by the present report but is a fully integrated spectrometer-collectors-superstructure system. Information on characterisation of the MOBY system can be found in [59].
- The HyperNav system under development by Seabird Scientific consists of new hyperspectral radiometers integrated within an autonomous underwater profiling float. This system is under development at the time of writing but could be integrated in an update of the current review as information becomes available.
- Would you consider appropriate to include something here about the DALEC radiometer? (from in-situ marine optics).

13.7 General Recommendations

In order to ensure the reliability of measurement results i.e. traceability to the units of SI with the associated uncertainty evaluation it is recommended to instrument manufacturers:

- to characterise new types of instruments in well-equipped optics laboratories under stable reference conditions as well as under varied conditions similar to in-field measurements;
- to provide further public information on instrument performance and characterisation where necessary to fill gaps in present knowledge listed in [54].

Recommendations to instrument users:

- to perform regularly the radiometric calibration of instruments in well-equipped calibration laboratories, collect and carefully analyse the results;
- to request, as customers, detailed performance information from the instrument manufacturers;
- to verify specifications of instrument performance by performing independent tests. For scientists with access to a well-equipped optics laboratory these tests could be quite detailed, e.g. measurement of cosine response of irradiance sensors, measurement of thermal sensitivity, measurement of stray light/out of band response,

although it is fully recognised that such tests may be very time-consuming and will generally require external funding. For scientist without access to a well-equipped optics laboratory it is still possible to verify certain aspects of instrument performance, e.g. by intercomparison of measurements made by different instruments pointing at a uniform target such as a cloudless sky or by participation in multi-partner intercomparison activities (such as the LCE activities of the FRM4SOC project).

It is recommended to ESA and other space agencies or entities, including Copernicus Services, requiring FRM for satellite validation to fund and encourage:

- preparation of a guide document setting minimum requirements for most important properties of OCR instruments (like temporal stability, linearity, thermal stability, etc.);
- activities to test radiometers from all manufacturers according to standardised methodology;
- further development of OCR instruments, including a requirement that such developments provide FRM-compatible information on radiometer characterisation.

14 SI-traceable Laboratory Comparison Experiment (LCE-1) for verification of reference irradiance and radiance sources

14.1 Introduction

In order to reach Objective 3 of the FRM4SOC project – design, document measurement protocols and procedures, and implement a laboratory-based comparison experiment to verify the performance of reference irradiance and radiance sources (i.e. lamps, plaques, etc.) used to maintain the calibration of FRM OCR radiometers traceable to SI – the following documents were prepared:

- “Protocols and Procedures to Verify the Performance of Reference Irradiance Sources used by Fiducial Reference Measurement Ocean Colour Radiometers for Satellite Validation” [D-80], [149,150];
- “LCE-1 Implementation Plan (LCE-1-IP)” [D-90], [151];

Following the guidelines as provided by these two documents

- a laboratory comparison experiment (LCE-1) to verify the performance of reference radiance and irradiance sources (i.e. lamps, plaques etc.) used to maintain the calibration of FRM OCR radiometers traceable to SI [D-100] was organised.

The results of the of the LCE-1 are presented in the report

- “Results from the First FRM4SOC Reference Radiance and irradiance Source Verification Laboratory Calibration Experiment Campaign” [D-120], [152]

and also in the peer-reviewed paper [153] published in the FRM4SOC special issue of the MDPI journal Remote Sensing.

All data collected during the comparison experiment LCE-1 has been collected into the

- data package LCE-1 DATA [D-110].

14.2 Protocols and Procedures to Verify the Performance of Reference Irradiance and Radiance Sources used by Fiducial Reference Measurement (FRM) Ocean Colour Radiometers (OCR) for Satellite Validation. [D-80]

The document [D-80] [138], [139] addresses the requirements of the FRM4SOC SOW [3] to

- be written as a definitive handbook for those wishing to perform future LCE of this nature;
- critically review the exact methodology used by teams to practically verify the calibration of reference radiance/irradiance sources using external reference SI traceable calibration sources and/or other approaches;
- establish and document protocols and best practice to practically verify the performance of reference radiance sources using external reference SI traceable calibration sources and/or other approaches;
- establish and document protocols and best practice for verifying the performance of secondary standard transfer reference radiance sources used by researchers to validate the calibration of their FRM OCR.
- define how to establish, present and maintain uncertainty budgets for reference radiance calibration sources used by OCR.

The principles described in [D-80] are presented in detail in Section 14.3 and 14.4.

14.3 The laboratory comparison experiment (LCE-1) to verify the performance of reference radiance and irradiance sources (i.e. lamps, plaques, etc.) used to maintain the calibration of FRM OCR radiometers traceable to SI [D-100]

Following the requirements of the FRM4SOC SOW [7], the laboratory comparison experiment (LCE-1) [D-100] was organised.

In particular, as required by the FRM4SOC SOW [7], the comparison exercise LCE-1:

- was built on the experience and lessons learned from previous activities [107,154–157];
- followed the principles of QA4EO and in particular the guidelines [10,158–160];
- followed the procedures and protocols documented in [D-80] [138], [139];
- was coordinated by an NMI (NPL, UK) to provide SI traceability;
- provided required laboratory measurement facilities, including suitable SI traceable transfer standards (a reference calibration source and a reference transfer radiometer), that were linked to cross comparisons of relevant calibration reference sources;
- provided technical support to the participants (set-up of participants' instruments, perform measurements, evaluation of uncertainty).
- included a dedicated training session prior to the commencement of measurements to assure that all participants are fully aware of the relevant procedures and protocols;
- provided support to participants on all practical aspects (e.g. overview of the activity, dates, times, locations, customs and shipping aspects, hotels and travel details, visa requirements, etc.) as documented in "LCE-1 Implementation Plan (LCE-1-IP)" [D-140], [140].

LCE-1 addressed separately irradiance and radiance measurements comparisons. First NPL, the UK NMI, conducted a laboratory comparison of the irradiance sources involving measurements of all participating lamps at NPL.

The irradiance comparison took place in April 2017. Participants were encouraged to attend this comparison in person to hand carry the lamps to and from the comparison and to attend a training course in absolute radiometric calibration and uncertainty evaluation that was given at the same time. Remote participation in the irradiance comparison was allowed, however, the training course was given only to the seven participants present at NPL at the time.

Then, a round robin of each participant's radiance sources using ocean colour transfer radiometers was performed between May 2017 and October 2018. This involved two calibrated transfer radiometers sent back and forth in turn to each participant to perform radiance measurements. NPL served as a pilot and was responsible for inviting participants, circulating the transfer radiometers and for the analysis of data, following appropriate processing by individual participants. The experiment was conducted anonymously. NPL was the only organisation to have access to and to view all data from all participants.

The LCE-1 intended to verify the performance of irradiance and radiance sources used to calibrate ocean colour radiometers and took the form of international comparisons. Public announcements were made to publicise the activities and any laboratory that holds absolute calibration standards used for spectral irradiance and radiance calibration could participate in the comparison.

14.3.1 Participants

The list of participants in LCE-1 of FRM4SOC is shown in Table 6. Note that three of the institutes, DLR-IMF, JRC and NIVA participated in the radiance round robin only. Some of the initial contact persons have left their organisations by the time this report is published, therefore, two people are listed in the contact person column of Table 6.

Table 6. List of LCE-1 Participants

Contact Person	Institute	Contact Details
Agnieszka Bialek Andrew Banks	National Physical Laboratory (NPL), UK	agnieszka.bialek@npl.co.uk
Joel Kuusk	Tartu Observatory (TO), Estonia	joel.kuusk@to.ee
Giuseppe Zibordi	European Commission, Joint Research Centre (JRC)	giuseppe.zibordi@ec.europa.eu
Vincenzo Vellucci	Institut de la Mer de Villefranche (IMEV), France	enzo@obs-vlfr.fr
Ronnie Van Dommelen Andrew Barnard	Satlantic (Sea-Bird Scientific), Canada	abarnard@seabird.com
Ian Lau	Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia	ian.lau@csiro.au
Sabine Marty	Norsk Institutt for Vannforskning (NIVA), Norway	sabine.marty@niva.no
Christopher MacLellan (now at NPL)	Natural Environment Research Council's Field Spectroscopy Facility (NERC-FSF), UK	chris.maclellan@npl.co.uk
Michael Ondrusek	National Oceanic and Atmospheric Administration (NOAA), USA	michael.ondrusek@noaa.gov
Johannes Brachmann Peter Haschberger	Remote Sensing Technology Institute, Deutsches Zentrum für Luft und Raumfahrt (DLR-IMF), Germany	Peter.Haschberger@dlr.de

14.3.2 Irradiance sources comparison

All participants documented their traceability to SI for both irradiance and radiance measurements via appropriate calibration certificates.

In this comparison, 1000 W quartz tungsten halogen (QTH) lamps, so called FEL lamps (not an acronym) according to ANSI (American National Standard Institute) designation, are considered as irradiance sources and were used at the standard calibration distance of 500 mm measured from their reference plane.

The results of the comparison were expressed relative to the mean spectral irradiance values provided by the participants. Since the participants all use different (i.e. their own) lamps, the means to assess any differences between them were determined by NPL making measurements of all lamps. Furthermore, participant provided data was verified by NPL making measurements of all lamps. The mean ratio between the participants' measurements and those made at NPL was calculated and results for each lamp were then expressed relative to this mean ratio, so showing the degree to which the individual measurements agree with one another.

This approach was taken because:

1. The participants had various SI-traceability routes for their lamps, i.e. a number of different NMIs providing their calibration. If all results were shown relative to the NPL values, then this might give the impression that traceability to NPL is 'correct' or 'best' whereas traceability to any NMI should be regarded as equally acceptable, differing only by certificated uncertainty statements.
2. A few of the lamps were recently calibrated at NPL and using a simple ratio to the NPL scale would have shown them to be performing almost 'perfectly' and thus give a misleading and biased comparison.
3. Presenting the results in terms of the agreement between each lamp and the mean of all the lamps shows how well measurements of the different participants agree with each other, regardless of the traceability route. This was the key aim of the comparison and this form of presentation gives the clearest indication of that.
4. The ratio between the NPL scale and the mean of all the participants' lamps is also included, which gives the confidence that the linkage to SI is sound in all cases.

14.3.3 List of irradiance sources used

The FEL lamps from four different commercially available sources were used in this comparison: Gooch and Housego (now Optronic Laboratories) OL FEL 1000 W, Gigahertz BN-9101 FEL 1000 W, Gamma Scientific Model 5000 FEL 1000 W and LOT-Oriel 63350 FEL 1000 W. In addition, one lamp was custom built using a general purpose Osram Sylvania 1000 W FEL lamp by one participant. In total 14 lamps from the participants and two of the NPL standards were measured.

Figure 15 shows the three main types of lamp that are presently used as irradiance sources in calibration laboratories worldwide. Table 7 provides the details of individual lamps used. Note that each type of lamp can have a slightly different alignment procedure and reference plane for the 500 mm distance. For consistency, both the pilot and participants should follow what is specified on the calibration certificates for each lamp.

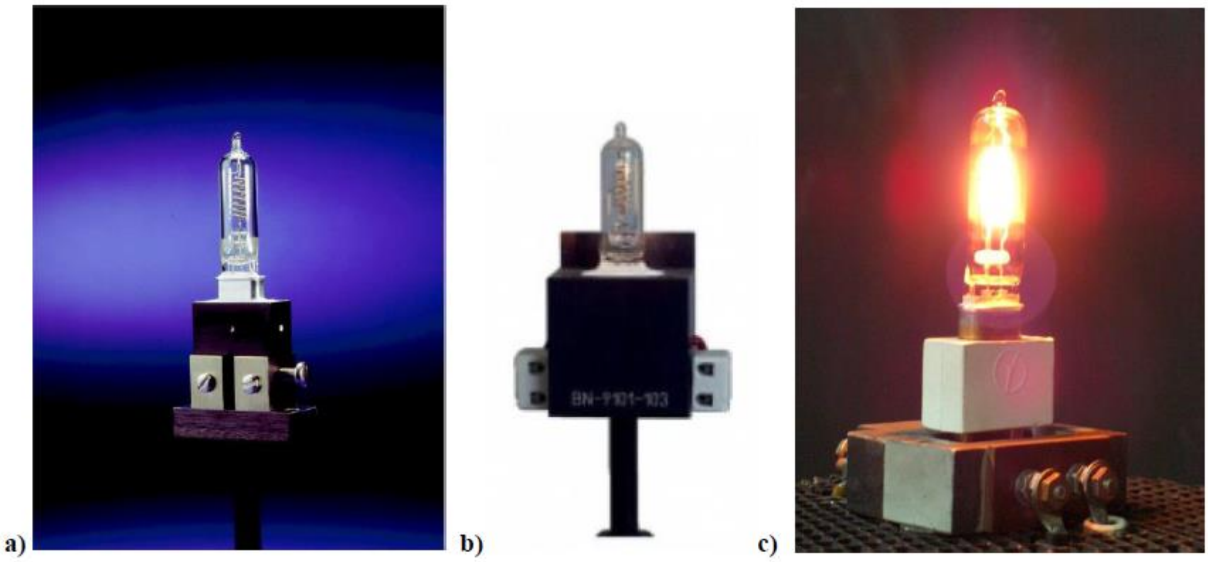


Figure 15. The three main types of FEL tungsten halogen lamps: a) Gooch and Housego OL FEL 1000 W, source Gooch and Housego; b) Gigahertz BN-9101 FEL 1000 W, source Gigahertz-Optik; c) Gamma Scientific Model 5000 FEL 1000 W, source Gamma Scientific.

Table 7. Detailed information about the lamps used in irradiance comparison LCE-1. In column 3, "RefSpec" and SRIPS are designated names of NPL measurement facilities, as indicated lamps were measured against the same references but sometimes using different facilities.

Manufacturer	Lamp serial number	Tested at NPL on	Last calibration date	Current
Gigahertz Optik	BN 9101-431	RefSpec	10.05.2010	8.1 A
LOT-Oriel 63350	FEL 7-1637	RefSpec	10.05.2010	8.2 A
FEL-TO-1 Osram 64743	FEL 17	RefSpec	01.04.2017	8.1 A
Optronics Labs FEL-M, G&H	F-921	RefSpec	04.12.2008	8.2 A
Optronics Labs FEL-M, G&H	F-1409	RefSpec	30.08.2016	8.2 A
Optronics Labs FEL-M, G&H	F-1425	RefSpec	02.02.2017	8.2 A
5000 FEL, Gamma Scientific	GS1122	RefSpec	2017	8.0 A
5000 FEL, Gamma Scientific	P1604	RefSpec	2017	8.0 A
5000 FEL, Gamma Scientific	GS1015	SRIPS	October 2013	8.0 A
5000 FEL, Gamma Scientific	GS1031	SRIPS	October 2013	8.0 A
Optronics Labs FEL-F, G&H	F-1291	SRIPS	2014	8.0 A
Optronics Labs FEL-F, G&H	F-1380	SRIPS	2016	8.0 A
Optronics Labs FEL-C, G&H	F-1173	SRIPS	2013	8.0 A
Optronics Labs FEL-C, G&H	F-1298	SRIPS	03.07.2014	8.0 A

14.3.4 NPL procedure for measurements of lamps

Depending on the lamp type, the appropriate alignment procedure was used which followed the lamp manufacturer instructions. The distance measurement for each type of lamp is defined by the manufacturers, thus for example for the G&H the removable alignment jig was placed vertically in the lamp mount, using a spirit level balanced on the top of the alignment jig to set the vertical alignment perpendicular to the optical axis. The lamp mount was then adjusted so that the measurement axis, defined as the axis passing through and perpendicular to the centre of the measurement aperture, passed through the centre of the alignment marks on the jig and was perpendicular to the jig. The jig was then carefully removed from the mount and the lamp inserted in its place. The calibration refers to the absolute spectral irradiance at a distance of 0.5 m, measured from the front face of the alignment jig.

On the other hand, for the Gigahertz type lamps, the lamp was positioned vertically, and adjusted so that the measurement axis, defined as the axis passing through and perpendicular to the centre of the measurement aperture, passed through the centre of the removable alignment jig and was perpendicular to that jig. The jig was placed on the rear of the lamp, with the scratched side towards an alignment laser behind the lamp, although for TO this was placed at the front. The back-reflected light was used to make the lamp perpendicular and the target used to make it central. The lamp was checked to be vertical using a spirit level balanced on the top of this alignment jig. The calibration refers to the absolute spectral irradiance at a distance of 0.5 m, measured from the plate at the front of the lamp.

For the Gamma Scientific lamp, the source was positioned with the exit port facing the measurement instrument. The measurement axis, defined as the axis passing through and perpendicular to the centre of the measurement aperture, passed through the centre of the source exit port and was perpendicular to the reference plane. The reference plane was defined as the front plate of the lamp alignment jig. The calibration refers to the absolute spectral irradiance at a distance of 0.5 m, measured along the measurement axis from the reference plane defined as above. This type of lamp can be purchased with a specially designed lamp enclosure. One participant requested the measurements to be made with the lamp enclosure as this followed their operational procedure.

The lamp was operated from an actively stabilised DC power supply at a specified current which varied between 8.000 A and 8.200 A depending on the lamp type (see Table 7 for individual lamps current setting). The polarity of the electrical current was as marked on the lamp and it was not changed. The lamp was gradually ramped up to operating current and run for 30 minutes before measurements commenced. The voltage was monitored during measurement and is given for checking purposes only.

Absolute spectral irradiance values were determined by reference to the NPL_{2010} spectral irradiance scale. The measurements were made using the NPL Spectral Radiance and Irradiance Primary Scales (SRIPS) or NPL's secondary RefSpec facility by direct comparison to NPL standard lamps that are radiometrically calibrated to a high temperature, high emissivity blackbody source operated at a temperature of approximately 3050 K.

Spectral irradiance measurements were made from 350 nm to 900 nm with an instrument spectral bandwidth of approximately 2.8 nm FWHM. Ambient temperature during measurement was (22 ± 2) °C.

14.3.5 Radiance sources round robin

The measurand was the calibration factor determined for the transfer radiometer using each participant's own spectral radiance reference (a lamp-panel combination). This exercise was performed sequentially by each participant at their own institute with two transfer radiometers circulated to them in turn. Each participant recorded the readings of the transfer radiometers when viewing a diffuse reflectance panel at 45° to the normal, with the panel illuminated at normal incidence using their irradiance lamp (the organisations who took part in the irradiance comparison used the same lamps). Figure 16 presents the radiance measurement setup. The recommended distance between the lamp and the panel was 0.5 m as this is the default distance for the irradiance calibration.

The recommended minimum distance between the radiometer and the panel was defined as approximately 250 mm. However, it was recognised that some participants did not perform their measurements at this distance and often did not have the capability to set it in their laboratories, thus, other distances were allowed as well.

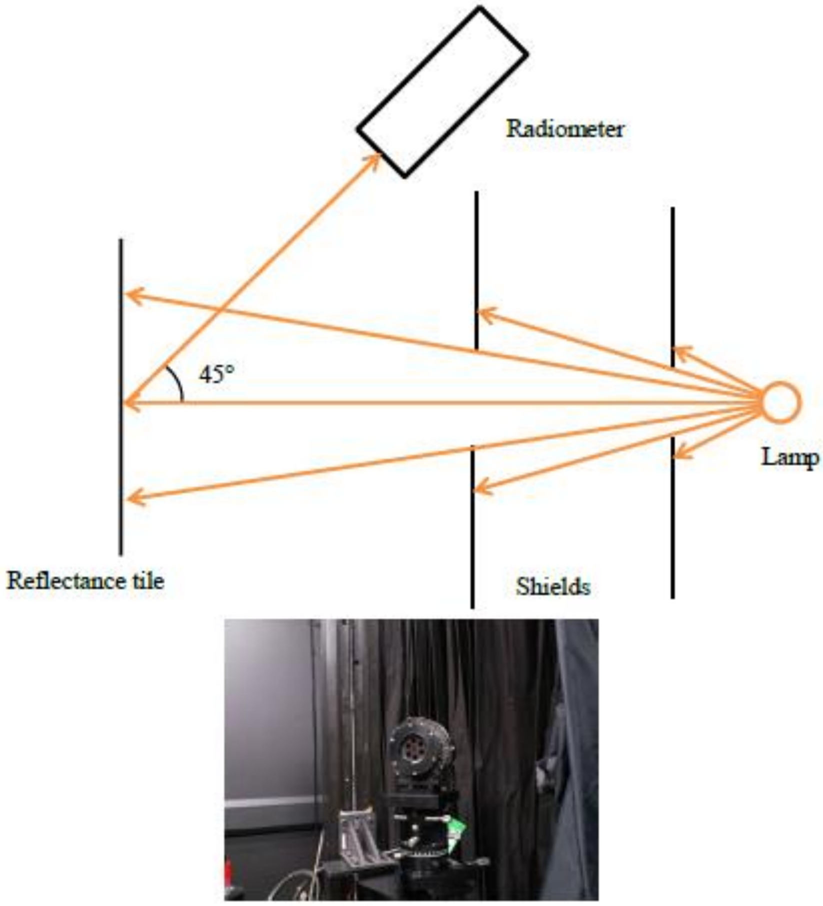


Figure 16. Radiance mode diagram of setup (top), multispectral radiometer (bottom).

Due to the different source calibration distance or other instrumental constraints, measurements were acquired at lamp-tile distances 500 mm, 750 mm, 1000 mm and 1300 mm by different participants, some participants obtaining measurements at more than one distance.

At least three independent measurements were made using each transfer radiometer by each participant. The measurements that were supplied to the pilot were:

1. three transfer radiometer readings for each radiometer and each spectral band obtained using the 0°:45° illumination and view geometry;
2. a table of spectral radiance as a function of wavelength from 400 nm – 700 nm as determined by the participant from the lamp irradiance combined with the panel reflectance for the 0°:45° geometry used.

For the participants who did not have their reference reflectance panels calibrated at 0°:45° reflectance factor geometry, the most common calibration for the 8º: hemispherical reflectance was allowed. A correction factor of 1.024 was applied to the diffuse reflectance calibration values to correct it for the proper measurement geometry. The value of that correction factor was established based on NPL internal data combined with published data by NIST [161].

Each participant was asked to evaluate the uncertainties associated with their radiance source operating in their own laboratory for these measurements. This included all the additional uncertainty components related to the alignment of the lamp, panel and radiometer, distance measurements, and other relevant laboratory specific factors such as power supply stability and accuracy.

The results of the comparisons are expressed as the percentage difference to the mean calibration coefficients obtained by taking an average of all participants results.

14.3.6 Transfer Radiometers

Two Satlantic (Sea-Bird Scientific) ocean colour radiometers (OCR-200) were used as transfer radiometers. As NPL does not own an OCR instrument, one of the participants, the Joint Research Centre of the European Commission (JRC), kindly agreed to provide two of their stable OC filter radiometers as the transfer radiometers. These are seven channel multispectral instruments. The general technical characteristics of this type of radiometer are shown in Figure 17, although the two particular instruments used for FRM4SOC have been customized by Satlantic (Sea-Bird Scientific) for JRC in terms of their spatial characteristics to provide a narrower (~3°) field of view (FOV) in air. Initial characterisation measurements to confirm this FOV have been carried out by NPL in air, and found to be 2.5° ± 0.3° at FWHM, with a close to Gaussian profile (see Figure 18).

Upwelling Radiance Sensor Characteristics

Sensor Model : OCR-200

Spatial Characteristics:	
- Field of view:	10° (0.100 steradians) in water
	14° (0.200 steradians) in air
- Entrance aperture:	9.5 mm diameter
- Detectors:	custom 13 mm^2 silicon photodiodes
Spectral Characteristics:	
- Bandwidth range:	400-700 nm
- Number of channels:	7
- Spectral bandwidth:	10 nm
- Filter Type:	custom low fluorescence interference
- Discrete wavelengths (centers):	412, 443, 490, 510, 560, 665, 683 nm
Optical Characteristics:	
- Out of band rejection:	10^{-6}
- Out of field rejection:	$5x10^{-4}$
- Typical saturation:	$5\mu Wcm^{-2}nm^{-1}sr^{-1}$ (customizable)
- Typical NER:	$1x10^{-4}\mu Wcm^{-2}nm^{-1}sr^{-1}$
Temporal Characteristics:	
- System time constant:	0.050 seconds
- -3dB frequency:	3 Hz
Physical Characteristics:	
- size:	8.9cm diameter, 12cm long
- weight:	1.2kg (in air)
- depth rating:	400m
- interface	Subconn micro 12-pin for remote/stand alone application
	or inline interface with DATA-100
- analog output range:	0-5V (customizable)

Figure 17. General specifications of the Satlantic (Sea-Bird Scientific) OCR-200 radiometers. The two particular instruments used for the LCE-1 had the customized ~3° FOV in air.

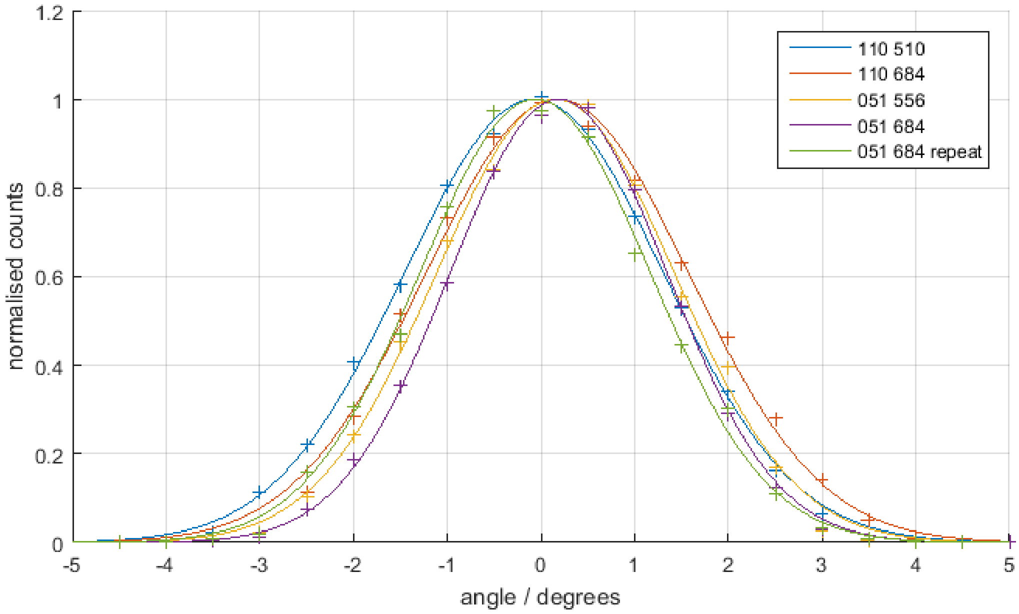


Figure 18. Measurement results to confirm the FOV of the transfer radiometers being used in the FRM4SOC LCE-1 radiance round robin.

14.3.7 Participants' laboratory setups

Each participant had its own equipped laboratory and used the available equipment to perform the round robin exercise. We present in a series of figures (Figure 19 to Figure 21) photographs from several laboratories to show differences in the radiance calibration setting. Figure 19 shows a custom-built radiance facility that is used for routine calibration and has a set of reference radiometers permanently mounted to monitor the radiance during the measurements.

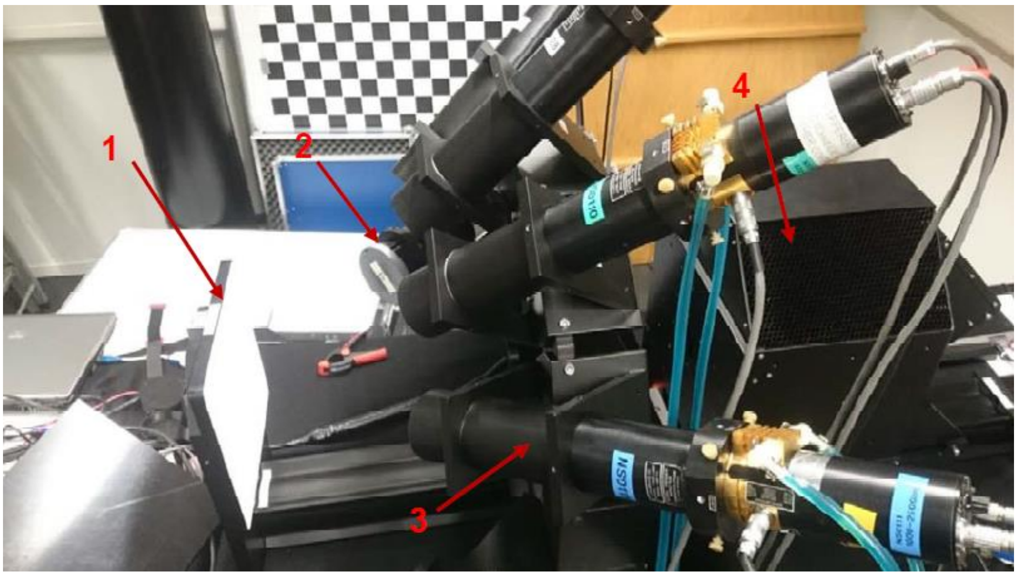


Figure 19. OCR-200 Radiometer mounted on a custom build radiance system. 1: Spectralon Panel, 2: OCR-200 Radiometer, 3: Reference radiometers, 4: Lamp housing.

Figure 20 and Figure 21 show examples of different illumination patch sizes that depend on the size, shape and position of the aperture in the light shield.

Figure 20. Example of laboratory set up during the radiance measurement. Top left the rectangular aperture that creates the illumination patch size equal to the reflectance panel, top right the illumination patch is circular and fills up the reflectance panel, the bottom left the illumination patch is bigger than the size of the panel (see the shadowing on the wall behind the panel)

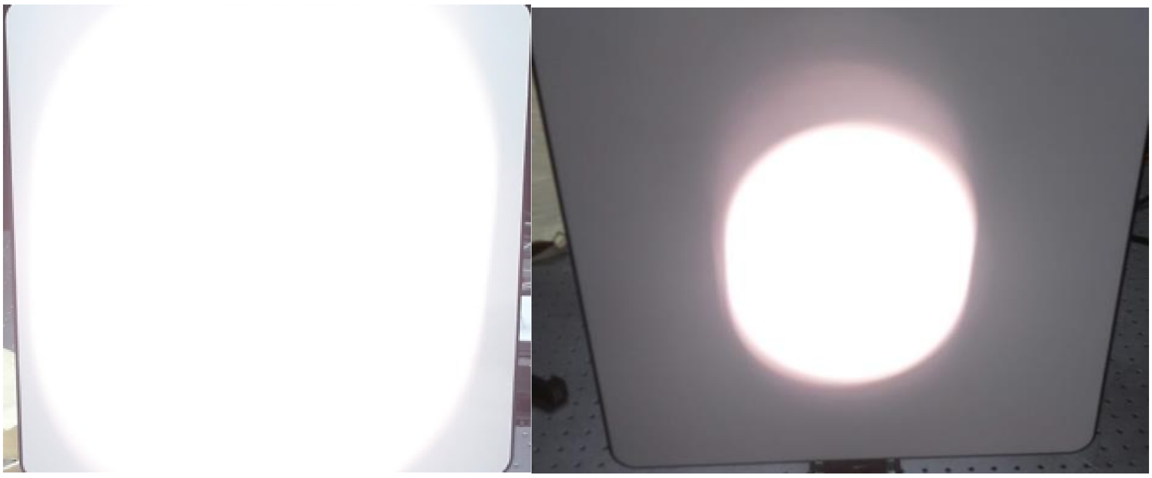


Figure 21. An example of panel illumination for one participant that due to the aperture shape has oval shape at two different distance settings. Left 1000 mm distance to the lamp and right 500 mm.

14.4 Results from the First FRM4SOC Reference Radiance and Irradiance Source Verification Laboratory Calibration Experiment Campaign" [D-120]

Technical Report TR-4 "Results from the First FRM4SOC Reference Radiance Source Verification Laboratory Calibration Experiment Campaign" [D-120], [152] addresses the requirements of the FRM4SOC SOW [7] to:

- Report in full the activities of the LCE-1 and the results obtained;
- Conclude with a set of activities to improve future LCE and any actions required to bring reference radiance sources used for OCR satellite validation up to FRM standards.

14.4.1 Irradiance sources comparison

The irradiance values are reported at the following wavelengths (350, 360, 370, 380, 390, 400, 450, 500, 600, 700, 800, 900) nm. This wavelength selection was dictated by the wavelengths reported in the calibration certificates from the participants. Although currently there are no OC missions that provide data below 400 nm, we wanted to include the ultraviolet spectral region in the comparison. There is a scientific interest to cover shorter wavelength and this is planned for the PACE mission [162].

The radiance values are reported at the transfer radiometer spectral bands values (412, 443, 491, 510, 560, 667, 684) nm. The overall summary result of the irradiance comparison is presented in Figure 22. The data series for each lamp used in the comparison are marked as Lamp A to Lamp N. The black dash line indicates the mean ratio to NPL.

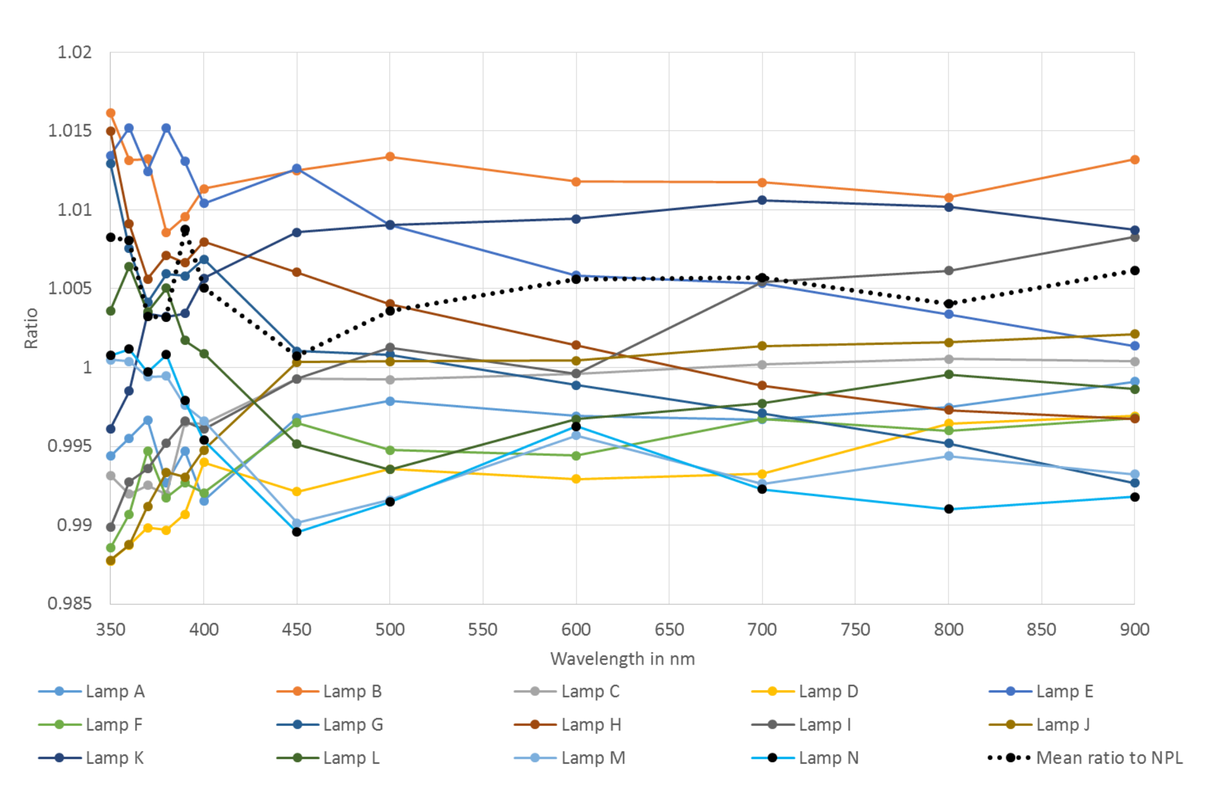


Figure 22. Summary result of all lamps used in irradiance comparison.

The results show an agreement, consistent with their uncertainties and that of the comparison between all measured lamps as all data series above 400 nm lay within the 0.99-1.013 range.

The spread in the results is higher for shorter wavelengths, as expected, due to the higher measurements uncertainty presented in the absolute radiometric calibration for this spectral range.

14.4.2 Uncertainty in irradiance measurement

The uncertainty for the individual lamp ratio is expressed by (11) and was calculated, according to the GUM [163], combining several uncertainty components. All components are relative, thus expressed in percentage form.

$$u(E_r) = \sqrt{(u^2\left(\frac{c_{\text{cer}}}{2}\right) + u^2(s_{\text{NPL}}) + u^2(c_{\text{n}}) + u^2(c_{\lambda}) + u^2(c_{\text{sl}}) + u^2(c_{\text{cur}}) + u^2(c_{\text{age}}) + u^2(c_{\text{alig}}),} \quad (11)$$

where the $u(E_r)$ is the uncertainty in the ratio of the irradiance values from the lamp calibration certificate to measurements performed at NPL, the $u\left(\frac{c_{\text{cef}}}{2}\right)$ is the uncertainty from the lamp calibration certificate, note that the uncertainty value is divided by two to convert it to a standard uncertainty from a coverage factor $k = 2$ used in the certificate.

The term $u(s_{\text{NPL}})$ is the NPL scale uncertainty and the additional components related to the measurements performed at NPL which included noise $u(c_{\text{n}})$, wavelength setting accuracy $u(c_{\lambda})$ room stray light $u(c_{\text{sl}})$, lamp current $u(c_{\text{cur}})$, ageing of the lamp $u(c_{\text{age}})$ and the lamp alignment $u(c_{\text{alig}})$.

The typical values for the ratio uncertainty expressed in percent are given in Table 8.

Table 8. Example of an FEL lamp's comparison uncertainty values.

Wavelength (nm)	Example of a lamp ratio uncertainty (*k*=1)
350	1.6 %
360	1.7 %
370	1.7 %
380	1.7 %
390	1.5 %
400	1.5 %
450	1.1 %
500	1.1 %
600	1.0 %
700	1.0 %
800	1.0 %
900	0.9 %

14.4.3 Radiance sources comparison

The overall summary result of the radiance comparison is presented in Figure 23 for the radiometer with serial number 051 and in Figure 24 for the radiometer with serial number 110. The colour triangles represent the seven spectral bands of the radiometers and the participants are marked as letters from A to M. Please note that we present here 13 entries to the summary results that came from 10 participating institutes. The number of entries is higher because some organisations provided results at two different distance settings between the lamp and the reflectance tile, or two different measurement set-ups for example radiometers alignment to the central spectral channel versus alignment to each spectral channel in turn.

The results for both instruments show the same trends. The difference for any individual participant from the mean participant value is up to 4 %, which is slightly higher than expected. A clear division into two separate groups can be seen. Most results form a group located at around 0 % and below on the y axis values. A second group of 4 entries is located at the level of 3 % difference from the mean comparison value.

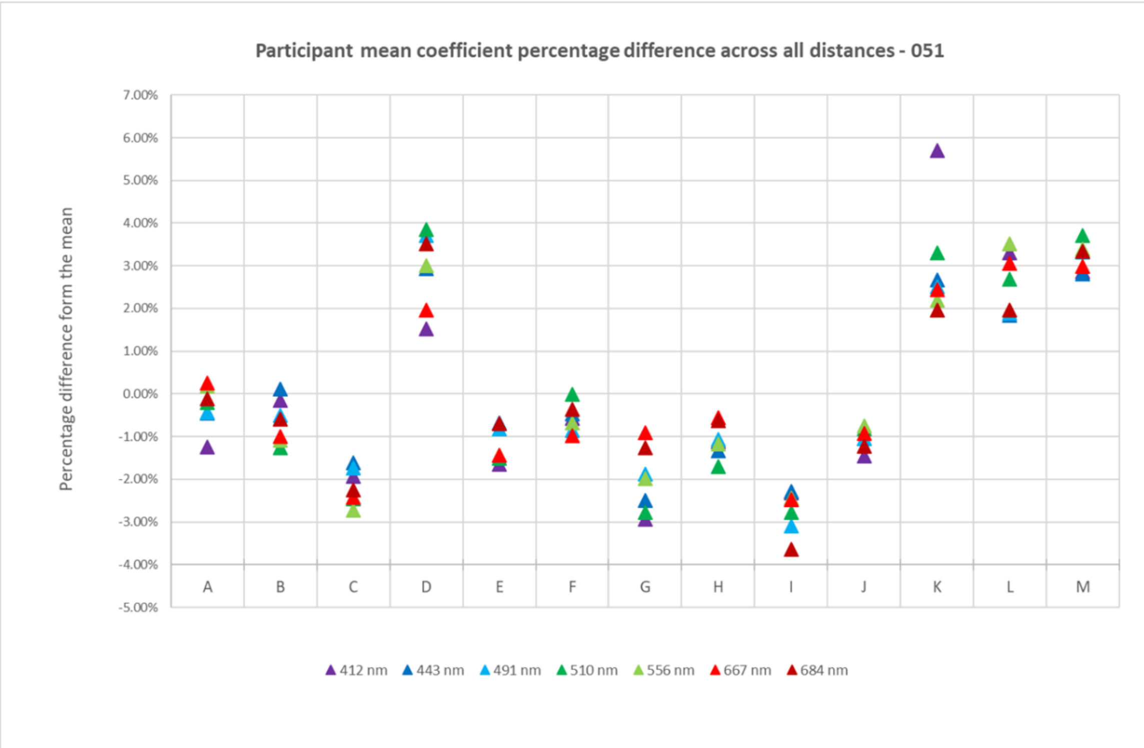


Figure 23. Summary of all radiance comparisons for radiometer number 051.

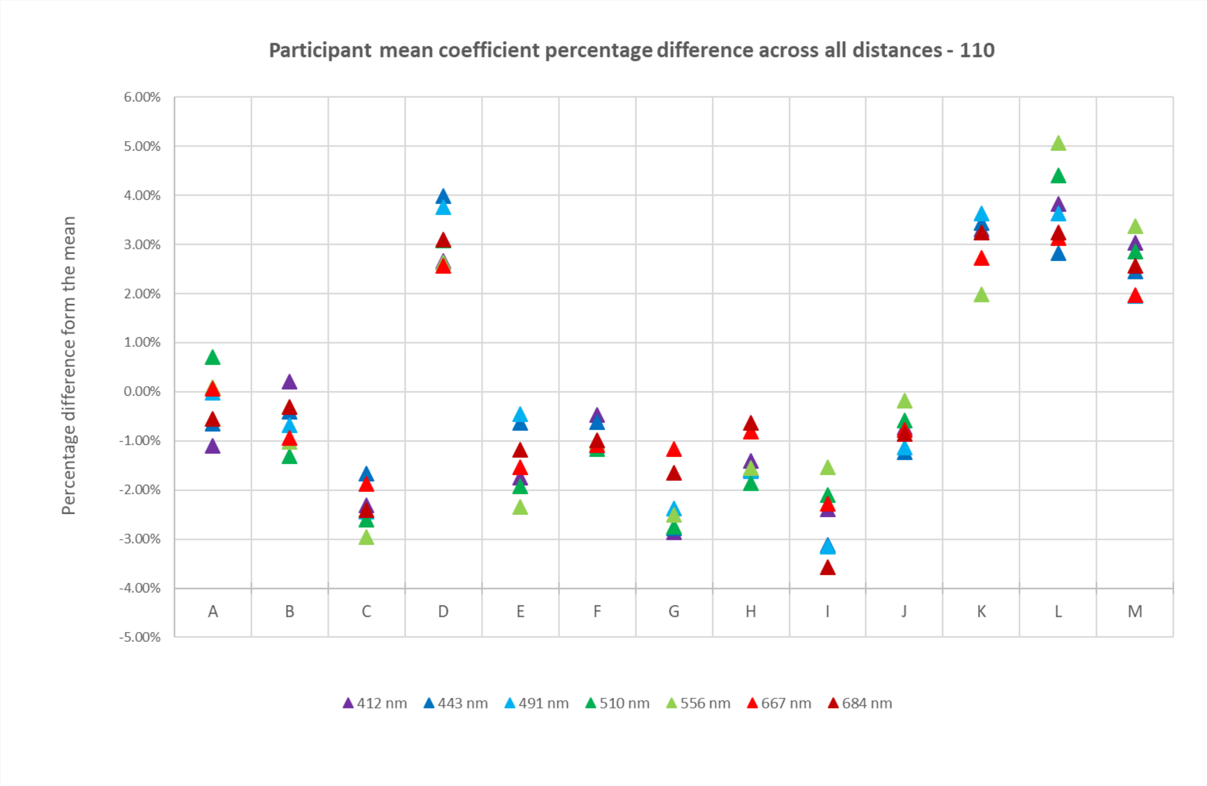


Figure 24. Summary of all radiance comparisons for radiometer number 110.

14.4.4 Uncertainty in radiance measurements

Each participant was asked to provide their radiance measurement's uncertainty budget. The uncertainty of radiance measurements is calculated according to equations (12) and (13). All components are expressed in percentage form, apart from radiometer's readings during the radiance calibration expressed in digital numbers DN_l and DN_d, where index $_\mathrm{l}$ stands for light and $_\mathrm{d}$ for dark reading respectively. Relative combined variance of the radiance calibration coefficient $u^2(L_\mathrm{cal})$ is expressed as

$$u^2(L_\mathrm{cal}) = \left(\frac{u^2(DN_\mathrm{l}) + u^2(DN_\mathrm{d})}{(DN_\mathrm{l} - DN_\mathrm{d})^2}\right) + u^2(c_\mathrm{sl}) + u^2(L), \qquad (12)$$

where $u(DN_\mathrm{l})$ and $u(DN_\mathrm{d})$ are the absolute uncertainties of the radiometer's readings, $u(c_\mathrm{sl})$ is the room stray light uncertainty, expressed as percentage of the radiometer signal after the dark reading subtraction, and $u(L)$ is the relative uncertainty of the radiance source. It is necessary to use $u(DN_\mathrm{l})$ and $u(DN_\mathrm{d})$ in absolute form in order to address correctly the subtraction of the dark reading in the measurement equation. The combined uncertainty $u(L_\mathrm{cal})$ is calculated as square root of the variance. The radiance source uncertainty $u(L)$ components are combined in (13).

$$u(L) = \sqrt{(u^2(E) + 2^2 u^2(d) + u^2(\rho_{0:45}) + u^2(c_\mathrm{cur}) + u^2(c_\mathrm{age}) + u^2(c_\mathrm{alig}) + u^2(c_\mathrm{unif}),} \qquad (13)$$

where $u(E)$ is lamp irradiance absolute calibration uncertainty converted to k = 1 from the certificate values, $u(d)$ is the uncertainty in the distance setting, note that this component has a sensitivity coefficient equal to 2 (from the inverse square law) and hence in (13) there is a term 2^2 just before it. In addition, for all measurements at distances different from 500 mm, the participants were requested to include a filament-offset uncertainty component to account for the difference in the plane of the distance setting and actual lamp filament position. The $u(\rho_{0:45})$ is the relative uncertainty of the reflectance standard calibration. Please note we use the reflectance factor calibration at 0° : 45° geometry at k = 1; for the case where a diffuse reflectance calibration value is corrected to 0° : 45° geometry, an additional uncertainty of that correction has to be included in the equation, NPL recommends a 0.5 % value. The $u(c_{\mathrm{cur}})$ is the lamp current uncertainty, $u(c_{\mathrm{age}})$ is the lamp ageing uncertainty, $u(c_{\mathrm{alig}})$ is the uncertainty in the alignment of the lamp and the reflectance panel at 0° : 45° configuration, and $u(c_{\mathrm{unif}})$ is the uncertainty due to the reflectance target illumination non-uniformity.

A few examples of participants' radiance measurements uncertainty expressed in percent are presented in Table 9.

Table 9. Examples of participants' radiance calibration relative uncertainty, (k = 2).

Band (nm)	Participant with low u	Participant with medium u	Participant with high u
412	2.0 %	2.4 %	3.1 %
443	1.8 %	2.4 %	2.9 %
491	1.8 %	2.2 %	2.7 %
510	1.8 %	2.3 %	2.7 %
556	1.8 %	2.2 %	2.5 %
667	1.8 %	2.1 %	2.5 %
684	1.8 %	2.1 %	2.5 %

14.4.5 Key findings of the LCE-1

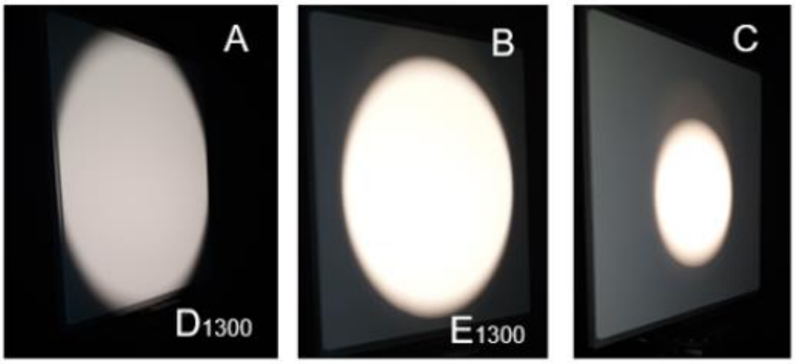
The irradiance comparison results showed a good agreement within the expected uncertainty range. Higher differences were observed for wavelengths below 400 nm as this is a more challenging spectral region for radiometric measurements. This indicates that, for future satellite missions the absolute calibration will have higher uncertainties for these new wavelengths. It is important to note that the same lamps (irradiance standards) used elsewhere in a different laboratory environment, using a different power supply or being aligned less carefully could produce different results. Therefore, for any measurements using the lamp as the irradiance source it is essential to consider all uncertainty components related to the measurement, as was shown in equation (11) in addition to the absolute calibration uncertainty values from the calibration certificate.

The radiance comparison results with the higher discrepancy led NPL to perform further investigations to explain the cause of the difference. We tried to identify common features for the smaller group of the results that agree well with each other but have a significantly higher percentage difference to the rest (the main group).

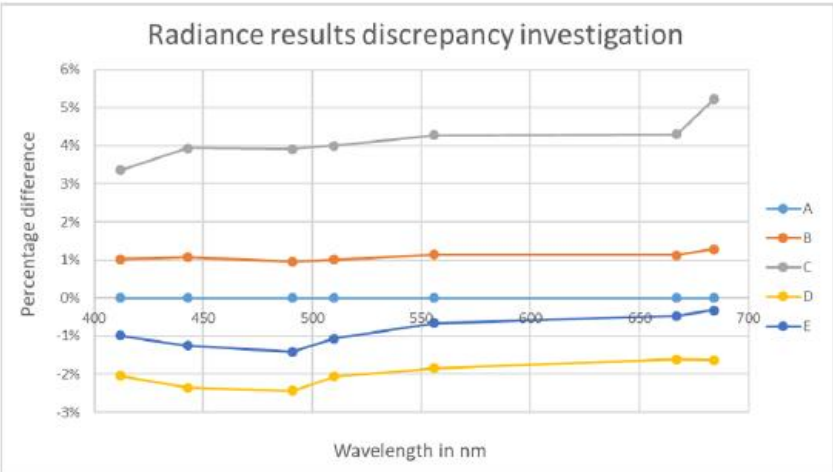
The first common feature for the small group is the measurement distance between the lamp and the reflectance tile set at 500 mm, however, the main group do have results from that distance as well. Thus, the distance setting is not the only cause of the difference. The second common feature was the size of the illuminated patch on the reflectance target from the lamp. This size was influenced by the distance between the diffuser and lamp and the choice of light shields and other baffles in a particular laboratory set up.

14.4.6 Additional measurements

NPL repeated a set of measurements to accommodate various conditions that participants may have in their own labs. The additional investigation was performed using an 18" Spectralon panel that was illuminated by the lamp at a distance of 500 mm and 1300 mm. The second distance was chosen as this was the longest distance used during the comparison. The size of the illuminated patch on the panel was varied from the fully illuminated panel, via a patch size with the diameter of around 23 cm to the small patch size of around 15 cm. The top panel in Figure 25 presents the photographs taken for the three different illumination patch sizes. The bottom panel in Figure 25 presents the percentage difference in the calibration coefficients obtained from five scenarios plotted as the data series from A to E. The series A, B and C represent the measurements at 500 mm distance for the fully illuminated panel, patch size 23 cm and the patch size 15 cm, respectively. The series D and E were performed at a 1300 mm distance for the fully illuminated and 23 cm patch size. Series A is set as the reference in this data set, therefore, the percentage difference for this series is 0 %.



(a) Photographs of illumination patch sizes.



(b) Plot with percentage difference between measurements for different patch size and lamp-tile distance setting.

Figure 25. Difference in the radiance measurement due to the distance setting and the size of the source.

A clearly visible positive bias can be seen for the measurements performed at 500 mm with a smaller size of the source. In addition, a small negative bias can be seen for the measurements performed at the higher distance. All participants from the small group had their measurements done at the 500 mm distance with a relatively small size of their radiance source. Although the radiometers have a FOV defined as 3°, this is a FHWM value and with a Gaussian shape to the FOV, rather than a top-hat. That 3° FOV, when plotted on a logarithmic scale rather than a linear scale, (see Figure 26) shows that there is still light detected at the level of 10^{-3} at 5°.

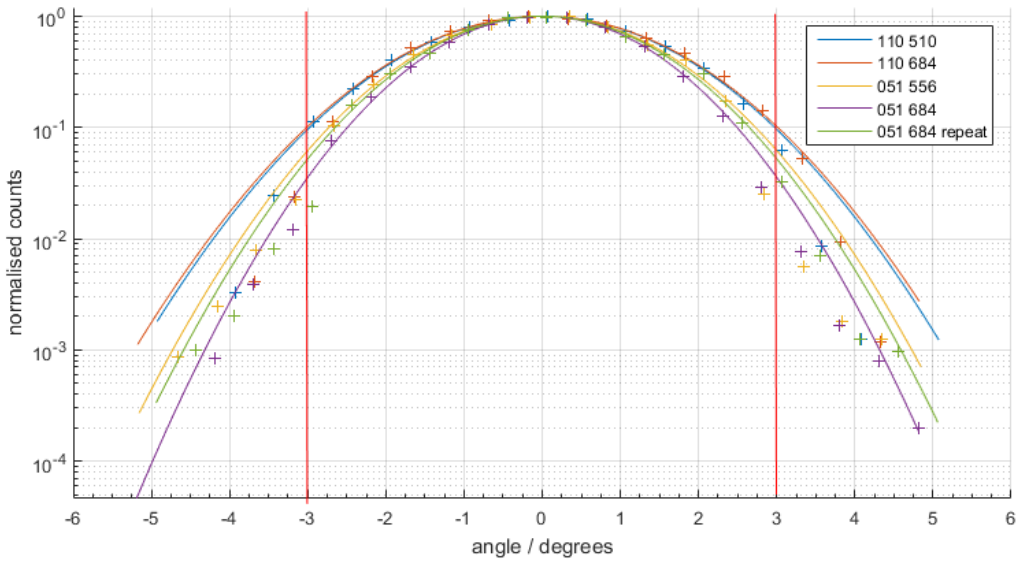


Figure 26. Measurement results to confirm the FOV of the transfer radiometers being used in the FRM4SOC LCE-1 radiance round robin presented on a logarithmic scale.

Thus, the smaller size of the source affects the results for this particular instrument set up. This effect is not as strong at longer distances, as can be seen in Figure 25, series E, which has a smaller patch size but did not show a positive bias. Thus, here the effect of the size of the source is compensated by a negative bias introduced by the distance setting. We analysed the data of all participants according to their sensitivity to distance. The results of this analysis are presented in Figure 27.

The four data series represent the averaged comparison results for different distances of 500 mm, 750 mm, 1000 mm and 1300 mm. Please note 750 mm and 1300 mm had only one entry to the comparison thus these are not averaged. The 500 mm series contains only the results from the main group and was set as the reference distance for that exercise, thus this data series has 0 % difference. The negative bias can be seen with the distance increase for the 1000 mm and 1300 mm distance. For the 750 mm distance this is not so obvious, however, this might be due to the fact that this particular participant has the radiance calibration values provided in radiance units for the whole system, rather than calculated from a lamp irradiance calibrated at 500 nm and a reflectance factor.

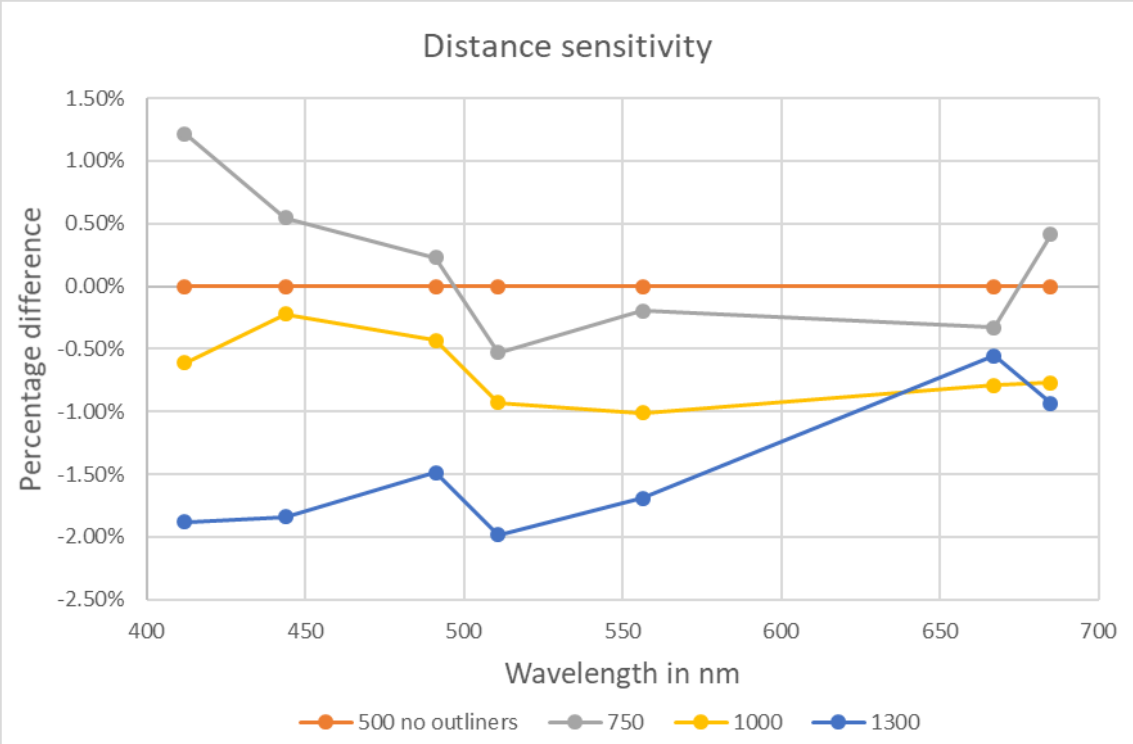


Figure 27. Difference in the radiance measurement due distance setting.

NPL recommended an uncertainty component be added for set-ups with a different distance to 500 mm to account for the filament offset, but we did not include a correction for the distance setting. It might be advisable to include the filament offset correction into the measurement equation to reduce further the discrepancies between the measurement results. Although the distance bias is observed in the main group, this would consolidate that group even more.

14.4.7 Lessons learnt

The transportation of the FEL lamps and transfer radiometers can be challenging. NPL recommends to hand carry the FEL lamps. However, the hand-carrying poses a risk of customs officers mishandling the artefact. This indeed happened to one of the lamps taking part in the LCE-1 comparison where one lamp was broken by a customs officer. It was shown in a previous formal metrology key comparison [164] that one in three lamps is likely to change its characteristic even with careful transportation, independent of any interference by customs officers. The indicative rule of a minimum three lamps to be enrolled in any comparison allows all participants to have some redundancy in case of lamp changes. The transfer radiometers were sent around the world in a robust peli-case and were always carefully packed and protected. We did not notice any change in the radiometer responsivity; thus, the instruments when transported in this manner can be considered very robust. However, at each transportation journey (between participants) stress gauges placed inside and outside the case were broken indicating that the parcel was not handled properly according to the request for fragile equipment handling!

More than one distance between the lamp and the reference panel for the radiance measurements was allowed to accommodate the capabilities of participating laboratories and indeed was important as the comparison was devised to evaluate differences between how participants assign calibration values and not necessarily how they can follow a defined procedure. This freedom in the distance setting led to the discrepancy in the radiance results and thus the identification of a source of error. The recommended 500 mm distance from the lamp is not frequently used in the ocean colour community due to the relatively large FOV of the instruments and the increased panel illumination non-uniformity in comparison to the longer distances [165]. However, a longer distance introduces an additional uncertainty component due to the filament offset (the difference between the actual position of the lamp filament and the plane from which the distance measurement is taken). This is an important lesson that shows how different laboratory settings influence the overall agreement in results and of course may impact differences when used in the field.

14.4.8 Conclusions

The irradiance comparison was run at NPL, where all participating lamps were measured against NPL standards. The results of that comparison were reported as the difference between the spectral irradiance values measured by each participant and the mean spectral irradiance values measured by all participants to show the degree to which the individual measurements agree with one another, without introducing a bias toward the NPL scale. The irradiance comparison values showed good agreements between all lamps.

The radiance comparison took the form of a round robin where two radiometers were sent in turn to each participant for the radiance measurements in house. The results were collated in the form of calibration coefficients obtained by each participant and expressed as the percentage difference to the participants' mean value. The results showed the discrepancies between the participants at the level of 4 % and two separate groups with measurement agreement (see Figure 23 and Figure 24). Additional investigation showed that the reason for this difference was caused by a combination of the size of source effect and instrument effective FOV that affected the results of the smaller group. If these effects could be corrected for, or the measurements repeated at a different setting, we would expect to see all measurements agreeing within 2.5 %.

The secondary objective of the comparison exercise was to increase the community awareness of measurement uncertainty evaluation using the GUM methodology. This was achieved via the training course that was provided for the participants being present at NPL during the irradiance comparison exercise week. The participants were given instruction in how to derive the uncertainty components related to their radiance measurements in house and all the round robin radiance results were reported accompanied by the uncertainty budgets.

14.5 Data package LCE-1 DATA [D-110]

All data collected during the laboratory comparison experiment LCE-1 (including raw, traceability, auxiliary and processed data) has been compiled into a data package file FRM4SOC-D-110-LCE-1-DATA.zip and handed over to ESA.

15 SI-traceable Laboratory Comparison Experiment (LCE-2) for verification of Fiducial Reference Measurement (FRM) Ocean Colour Radiometers (OCR)

15.1 Introduction

In order to meet Objective 4 of the FRM4SOC project – design, document measurement protocols and procedures, and implement a laboratory-based comparison experiment to verify the performance of FRM field OCR used for satellite validation – the following documents were prepared:

- "Protocols and Procedures to Verify the Performance of Fiducial Reference Measurement (FRM) Field Ocean Colour Radiometers (OCR) used for Satellite Validation" [D-130], [166];
- "LCE 2 Implementation Plan (LCE-2-IP)" [D-140] [145].

Following the guidelines as provided by these two documents

- a laboratory comparison experiment (LCE-2) for verification of FRM OCR was organised [D-150].

The results of the LCE-2 are presented in the report

- "Results from the First FRM4SOC Field Ocean Colour Radiometer Verification Round Robin Campaign" [D-170] [135]

and also in the peer-reviewed papers [167,168] published in the FRM4SOC special issue of the MDPI journal Remote Sensing. **For citation of the chapter 15**, the papers [167,168] should be considered as preferable references.

All data collected during the comparison experiment LCE-2 has been collected into the

- data package LCE-2 DATA [D-160].

15.2 Protocols and Procedures to Verify the Performance of Fiducial Reference Measurement (FRM) Field Ocean Colour Radiometers (OCR) used for Satellite Validation [D-130]

The document [D-130], [166] addresses the requirements of the FRM4SOC SOW [3] to

- be written as a definitive handbook for those wishing to perform future LCE of this nature;
- build on the experience and lessons learned from previous activities;
- follow the principles and guidelines settled by QA4EO [10,159,160,169];
- critically review the exact methodology used by teams to practically verify the calibration of FRM OCR using external reference SI traceable calibration sources and/or other approaches;
- establish and document protocols and good practice to practically verify the performance of FRM OCR using external reference SI traceable calibration sources and/or other approaches;
- define how to establish, present and maintain uncertainty budgets for FRM OCR.

The principles described in [D-130] are presented in detail in Section 15.3 and 15.4.

15.3 The Comparison Experiment (LCE-2) for verification of Fiducial Reference Measurement (FRM) Ocean Colour Radiometers (OCR) [D-150]

Following the established protocols and procedures [D-130], [166] the SI-traceable Laboratory Comparison Experiment (LCE-2) for verification of FRM OCR was organised from 8 to 13 May 2017 at TO, Estonia [D-150], [167,168] with the aim to:

1. establish the traceability of measurement results to the units of SI by calibration;
2. document and review critically the methodologies used by teams for OCR FRM;
3. evaluate and compile end-to-end uncertainty budgets;
4. validate the methods and uncertainty budgets in comparison experiments and draw conclusions for good practice.

The LCE-2 also served as a preparation stage for the FICE-AAOT field intercomparison exercise [170]. Technical support and dedicated training was provided to the participants of the comparison in order to assure that all participants would be fully aware of the relevant measurement procedures, protocols and uncertainty evaluation.

Figure 28. Participants and organizers of the LCE-2 intercomparison at TO.

There were 11 institutions participating in the LCE-2 (Figure 28 and Table 10) [135,167]. In order to plan and manage organisational issues of the comparison event (e.g. overview of the activity, dates, times, locations, customs and shipping aspects, hotels and travel details, visa requirements, etc.), the LCE2 Implementation Plan (LCE-2-IP) [D-140], [171] was prepared and followed.

Altogether 44 radiometric sensors from five different manufacturers were involved (Figure 12). The list of radiometers reflects the typical selection of instruments used for shipborne validation of satellite-derived water reflectance ("ocean colour validation"). However, the number of each type of instrument is not necessarily representative of total validation data usage, since the SeaPRISM instrument is used by a multi-site network of autonomous systems [1], thus providing very significant quantities of validation data. As denoted by the combination "(2L, 1E)" in Table 10, most of the participating teams use an above-water field measurement protocol with three radiometers: two radiance sensors, for upwelling (water) and downwelling (sky) radiances, respectively, and an irradiance sensor, measuring downwelling irradiance.

Table 10. Institutes and instruments participating in the LCE-2 intercomparison.

Participant	Country	L-radiance; E-irradiance sensor
Tartu Observatory (pilot)[7]	Estonia	RAMSES (2 L, 1 E) WISP-3 (2 L, 1 E)
Alfred Wegener Institute	Germany	RAMSES (2 L, 2 E)
Royal Belgian Institute of Natural Sciences	Belgium	RAMSES (7 L, 4 E)
National Research Council of Italy	Italy	SR-3500 (1 L, 1 E) WISP-3 (2 L, 1 E)
University of Algarve	Portugal	RAMSES (2 L, 1 E)
University of Victoria	Canada	OCR-3000[(8)] (2 L, 1 E)
Satlantic (Sea-Bird Scientific)	Canada	HyperOCR (2 L, 1 E)
Plymouth Marine Laboratory	UK	HyperOCR (2 L, 1 E)
Helmholtz-Zentrum Geesthacht	Germany	RAMSES (2 L, 1 E)
Estonian Marine Institute, University of Tartu	Estonia	RAMSES (1 L, 1 E)
Cimel Electronique S.A.S	France	SeaPRISM (1 L)

For the RAMSES and HyperOCR this is achieved by 3 separate devices, while the WISP-3 contains three spectroradiometers integrated into a single device, and the SR-3500 uses a single spectrometer equipped with interchangeable entrance optics for irradiance and radiance (and can, like all radiance sensors, be used sequentially to measure both upwelling and downwelling radiance). The SeaPRISM estimates irradiance (E) from direct sun radiance (L) – see [1]. In the scope of laboratory measurements, the multiple entrance optics of SR-3500 and WISP-3 were treated as separate radiometers. Technical parameters of the participating radiometers [135,167] are given in Table 11.

Water reflectance can also be measured from underwater radiometers deployed either at fixed depths or during vertical profiles. Indeed the RAMSES and HyperOCR designs (but not WISP-3, SR-3500, SeaPRISM) may also be used underwater. The present study is fully relevant for the calibration aspects of such radiometers in underwater applications, although extra characteristics, particularly immersion coefficients to transfer in-air calibrations to in-water [30] must also be studied.

[7] At the time of LCE-2 Tatu Observatory was an independent institute. Tartu Observatory joined with University of Tartu since 01.01.2018.
[8] OCR-3000 is the predecessor of HyperOCR

The LCE-2 was divided into three sub-tasks:

1) Provision of SI-traceable radiometric calibration for all radiometers participating in the comparison exercise. (02 – 07 May 2017)

All participating radiometers were calibrated prior to the comparison measurements under similar conditions at the same time and at the same laboratory (TO) in order to guarantee that differences in comparison results would not be primarily due to various calibration sources and/or calibration times. Traceability to SI for the radiance and irradiance standards was provided by a national metrology institute (NPL), see Figure 29.

2) Laboratory intercomparison of radiometers used for satellite validation in the (400…900) nm wavelength range. (09 – 10 May 2017)

The method of direct comparison to the group's median and an external reference were used. Measurements were performed by the owner/operator of the radiometer following prescribed procedures under supervision of the organising laboratory staff (TO).

3) Field intercomparison of radiometers used for satellite validation in the (400…900) nm wavelength range. (11 – 12 May 2017)

Table 11. Technical parameters of the participating radiometers.

Parameter	RAMSES	HyperOCR	WISP-3	SR-3500	SeaPRISM
Field of View (L/E)	7°/cos	6° [(9)] or 23°/cos	3°/cos	5°/cos	1.2°/NA
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
Number 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

A direct comparison between calibrated OCR instruments with a reference value combined from the measured values. Two types of outdoor exercises were planned.

i) *Primary comparison*, where all instruments are pointed at the same physical object and data are acquired synchronously. The instruments in this case are mounted and

[9] According to the manufacturer, the HyperOCR radiance sensors 444 and 445 have 6° FOV. After the report [D-170] and the Remote Sensing special issue paper [168] were published, it was confirmed with additional characterisation, that the rest of the HyperOCR and OCR-3000 instruments participating in LCE-2 have also 6° FOV.

directed by an organiser while the data is collected by instrument owners/operators. Instrument gains and integration times are selected according to everyone's own usual fieldwork practice. The physical quantities selected for the comparison exercise are absolute spectral irradiance and radiance.

ii) *Secondary comparison*, where participants set up their usual fieldwork configuration pointing towards the collectively agreed direction. Measurement directions and time frames are agreed collectively during the exercise while each participant is responsible to collect all the data (if supported by the instruments) needed for comparison of the physical quantities: 1) water-leaving absolute spectral radiance; 2) downward global absolute spectral irradiance; 3) remote sensing reflectance of water. Participants also collect weather and metadata (except water samples).

The measurement results and information about measurement parameters were reported back to the pilot laboratory by participants (Table 10) 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 GUM [18].

15.4 Results from the First FRM4SOC Field Ocean Colour Radiometer Verification Round Robin Campaign". [D-170]

The Technical Report TR-6 "Results from the First FRM4SOC Field Ocean Colour Radiometer Verification Round Robin Campaign" [D-170], [135] addresses the requirements of the FRM4SOC SOW [3] 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.

15.4.1 Calibration and characterisation of the participating radiometers

Traceability of the in situ measurements to SI units is established by regular calibration of field radiometers. Thus, immediately before the comparison, TO performed consistent calibration of all participating radiometers in order to guarantee that differences in comparison results will not be primarily due to various calibration sources and/or calibration times [167]. For determination of spectral responsivity of a radiometer, it is usually calibrated against a known source placed at a specified distance from the entrance optics of the radiometer. Such a calibration procedure is well established and validated [172–178].

The traceability scheme of the intercomparison is presented in Figure 29. A FEL type 1000 W quartz tungsten halogen spectral irradiance standard lamp was used for radiometric calibration of the radiometers. NPL provided two Gigahertz-Optik BN9101-2 FEL-type 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 340 nm to 980 nm determined using these lamps with a precision filter radiometer and a 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, the arithmetic mean of the responsivity measured by the two lamps was used.

The lamp was powered by a Newport/Oriel 69935 stabilized radiometric power supply 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.

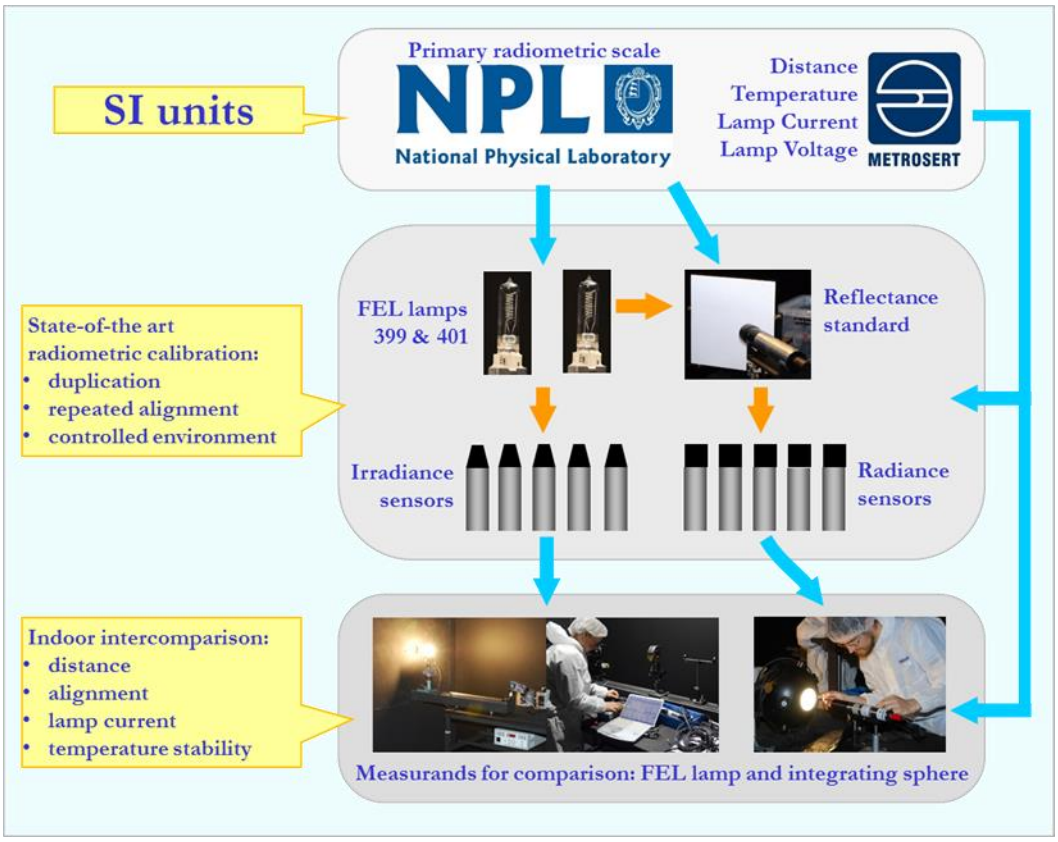
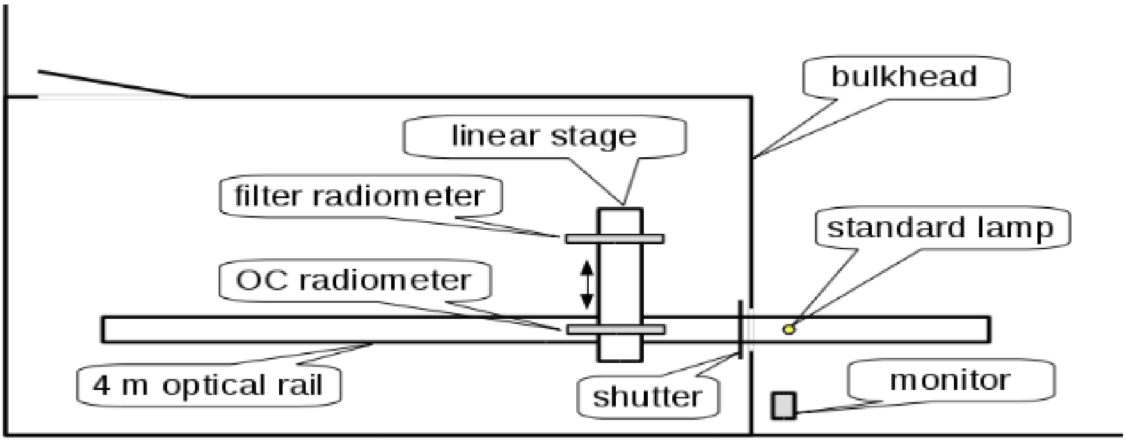


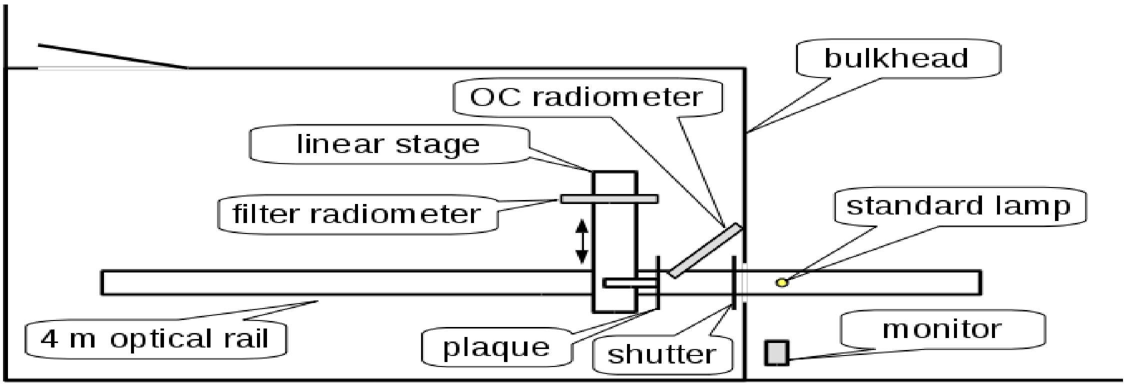
Figure 29. Traceability scheme of the LCE-2 [167].

The lamp and OC radiometer under calibration were mounted on an optical rail that passed through a bulkhead, which separated the lamp and radiometer during calibration [D-170], [135] (Figure 30). 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 (Sea-Bird Scientific) HyperOCR group) and all the instruments within the group had an identical outside diameter. This allowed the use of 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 an appropriate position was attached to the sensor. During calibration radiometers of the same type were swapped without turning off the lamp. Placing the clamp against the end of the V-block ensured the proper distance between the lamp and the radiometer during calibration.



(a)



(b)

Figure 30. Calibration setup for irradiance (a) and radiance (b).

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 a 120° angle (see 6 in Figure **32**). 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. A 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 monitoring 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 WISP-3 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 was collected, in order to estimate the dark signal and ambient stray light. All measurements were repeated at least twice, including readjustment of the lamp and the sensor.

The radiance sensor calibration setup (Figure 28 b) was based on the lamp/plaque method and utilized the same components as the irradiance sensor calibration setup and a Sphere Optics sg3151 (200×200) mm calibrated white reflectance standard. 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 the 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.

Additionally, a large number of the sensors involved in the comparisons were recalibrated at TO a year later for the FRM4SOC field intercomparison on the Aqua Alta Oceanographic Tower (AAOT) in the Gulf of Venice (see below) allowing the evaluation of the stability of the sensors. Most of these sensors (over 80 %) changed during this year less than ±1 %. Nevertheless, the change of the responsivity of some radiometers was unexpectedly large: more than ±10 %. Due to such changes, intermediate checks between regular calibrations during field campaigns are strongly advisable.

According to [19] calibration is an operation that, under specified conditions, in a first step, establishes a relation between the quantity values with measurement uncertainties provided by measurement standards and corresponding indications with associated measurement uncertainties and, in a second step, uses this information to establish a relation for obtaining a measurement result from an indication. Unfortunately, specified conditions during the calibration may be quite different from the varying conditions that may prevail during later use of the instrument. For radiometric sensors, there can be significant differences between calibration and later use in the field, regarding operating temperature, spectral variation of the target (causing different spectral stray light effects), angular variation of the light field (especially for irradiance sensors), and the intensity of the measured radiation. Each of these factors may interact with instrument imperfections to add further uncertainties when an instrument is used in the field and estimation of such uncertainties requires instrument characterisation in addition to the well-established absolute radiometric calibration.

This instrument characterisation, which can lead to corrections to reduce uncertainties, should include determination of:

- thermal effects,
- non-linearity,

- spectral stray light effects,
- wavelength calibration,
- angular response, and
- polarization effects.

Procedures for determination of corrections, including measurement of all relevant influencing factors are much less studied and, for some of these characteristics, for some instruments, often corrections might be not available. For applying corrections, individual testing of radiometers for each effect considered is indispensable. For most of the corrections, tests may be more time consuming than the radiometric calibration. Generally, the corrections should be applied both for calibration spectra and for field spectra calculated using the calibration coefficients, increasing critically the impact of data handling. Fortunately, these individual tests are carried out usually only once in the lifetime of an instrument unit (i.e. a sensor from an instrument family/design with a unique serial number) while radiometric calibration has to be performed on a regular basis at least once a year. Methods for correcting temperature effects [141,166,179–183], nonlinearity effects [166,181,183,184], polarization effects [185], and spectral stray light effects [166,186–190] are rather well studied. Nevertheless, difficulties may arise during the use of a calibrated instrument when some parameters influencing correction should be determined. Some radiometers do not have internal temperature sensors, and therefore, for these instruments the accuracy of temperature corrections is limited even when external temperature sensors are applied during the calibration and later use [179]. Non-linearity effects present in calibration spectra can be determined rather satisfactorily, but it can be much more difficult to account for non-linearity when unstable natural radiation sources are measured. Effects due to the response error of cosine collectors [191,192] can be satisfactorily accounted for with a well-known radiation source, but in the field conditions the angular distribution of radiation is often not known accurately enough for efficient correction of the cosine error.

15.4.2 Indoor intercomparison

15.4.2.1 Measurements

The indoor experiment took place in the optical radiometry laboratory of TO within a few days after the radiometric calibration [D-170], [135] (Figure 31). The radiance and irradiance comparison experiments were set up in neighbouring rooms and simultaneously carried out by project participants under the supervision of TO's personnel during two days.

The irradiance setup is presented in Figure 32. An FEL lamp was used as a stable irradiance source for the indoor intercomparison. The power supply, current feedback unit, monitoring 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 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 WISP-3 instruments two series (including re-alignment) of 10 readings were recorded and one series with the shutter closed.

Figure 31. The indoor intercomparison exercise.

The radiance setup for indoor intercomparison is depicted in Figure 33. 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 a 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 a 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 monitoring detector. The monitor detector current was recorded using an Agilent 3458A multimeter, and the 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 a 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.

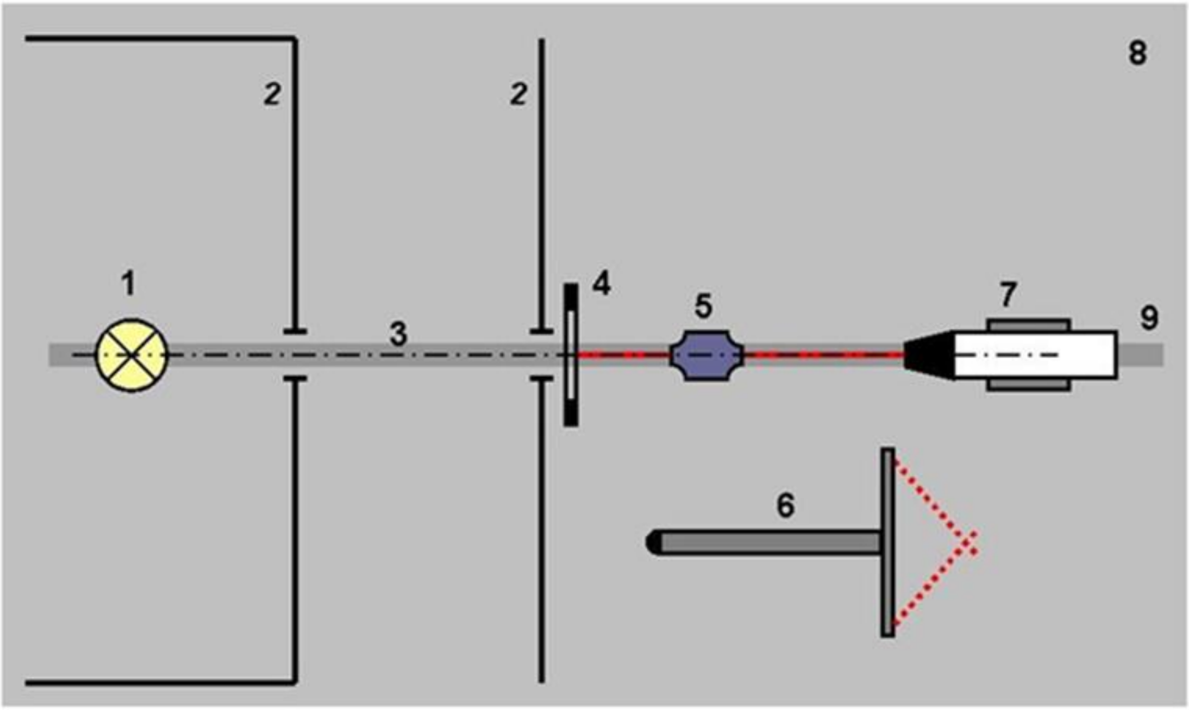


Figure 32. Indoor irradiance comparison. 1 - FEL lamp; 2 - baffles; 3 - main optical axis; 4 - alignment jig; 5 - alignment laser; 6 - distance tool; 7 - radiometer on the support; 8 - optical table; 9 - optical rail [D-170], [135].

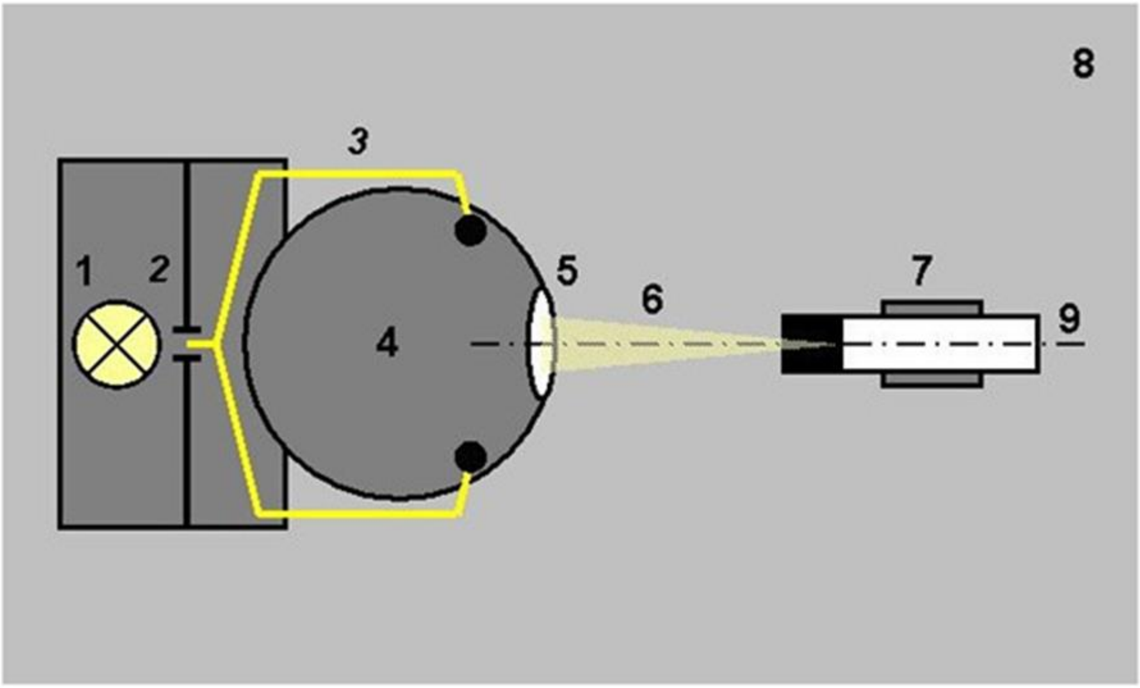


Figure 33. Indoor radiance comparison. 1 - quartz tungsten halogen lamp; 2 - variable slit; 3 - optical fibre; 4 - integrating sphere; 5 - output port; 6 - FOV of the radiometer; 7 - radiometer on the support; 8 - optical table; 9 - main optical axis [D-170], [135].

15.4.2.2 Data handling

The measurement results, including measurement uncertainty and information about measurement parameters, were reported back to the pilot laboratory in the form of spreadsheet files by most of the participants (for 33 out of the 44 sensors involved). For the rest, the pilot carried out the data analysis based on the raw instrument data. In the case of discrepancies, the 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.

The participants were encouraged to perform the data processing for their radiometers and report the radiance/irradiance values with uncertainties. However, in a few cases, TO repeated the calculations for some participants as well. Data processing for the RAMSES, HyperOCR and WISP-3 instruments was fully automated at TO by purpose-designed computer software freely available for the participants.

Data processing of each instrument was performed independently and included the following steps:

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

15.4.2.3 Device-specific issues

TriOS RAMSES series instruments include both the radiance (ARC) and irradiance (ACC) sensors. The raw spectra are stored in American Standard Code for Information Interchange (ASCII) and/or Microsoft ACCESS database files. Data processing for these radiometers is fully unified based on the measured data (2-byte integer numbers) and calibration files provided by the manufacturer and TO. The detailed procedure to derive the calibrated results is described in [193]. RAMSES instruments are equipped with black-painted pixels on the photodiode array used to compensate for the dark signal and electronic 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 according to the responsivity coefficients, linearity correction was applied, see Section 15.4.2.7.

Satlantic HyperOCR/OCR3000 series instruments include also both the radiance and irradiance sensors with a similar data processing chain. The raw spectra stored in binary files 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 RAMSES procedure. The HyperOCR radiometers are equipped with an internal mechanical shutter, deployed automatically after every fifth target spectrum. The shutter measurements were detected in the data files and the closest shutter measurement was subtracted from each raw spectrum before the next steps.

Water Insight WISP-3 contains a three-channel Ocean Optics JAZ module spectrometer and computer. Two of the input channels are connected to the radiance inputs while the third is attached to the irradiance adaptor. Users can start the acquisition of the spectra by pressing a button, the internal computer is setting the measurement sequence, determining the integration times, and storing the data. All three channels are acquired simultaneously and the data are stored into a single ASCII file. The spectrometers have painted detector array pixels like the RAMSES radiometers. The internal dark signal is subtracted automatically and resulting data are 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 Section 15.4.2.7 was not used.

Spectral Evolution 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 ASCII files and can contain both the raw and radiometrically calibrated results based on the internal calibration coefficients. 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 adjustment factors related to the existing coefficients were derived. The calibrated data in the files was multiplied by these factors, and finally, the linearity correction as described in Section 15.4.2.7 was used.

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

15.4.2.4 Calculation of Sentinel-3/OLCI band values

As the final step of data processing, the radiance and irradiance values were re-calculated for the OLCI spectral bands for each radiometer except for the multispectral SeaPRISM, in which case the initial band values were used. Based on the given CWL of the spectroradiometer λ_n, and the OLCI band definition $O_i(\lambda)$ [194], the weight factors were found for each pixel:

$$C_i(n) = O_i(\lambda_n), \qquad K_i(n) = \frac{C_i(n)}{\sum_k C_i(k)} \tag{14}$$

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

$$I_i = \sum_n I(n) \cdot K_i(n), \tag{15}$$

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

15.4.2.5 Consensus and reference values used for the analysis

Consensus values were calculated as median [195] of all presented comparison values. Reference values were applicable only for the indoor irradiance measurements (Figure 36), when the measurand used for this exercise was also measured during the comparison using the precision filter radiometer serving as a reference.

15.4.2.6 Results of indoor experiment

The comparison results are presented as relative deviations from the consensus value. Despite the 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 (Figure 34 - Figure 37) with no outliers present. In these figures, blue dashed lines show the expanded uncertainty covering 95 % of all data points on the right graphs. Solid lines represent RAMSES sensors, dashed lines - HyperOCR sensors, double line – SR-3500, and dotted lines – WISP-3 sensors.

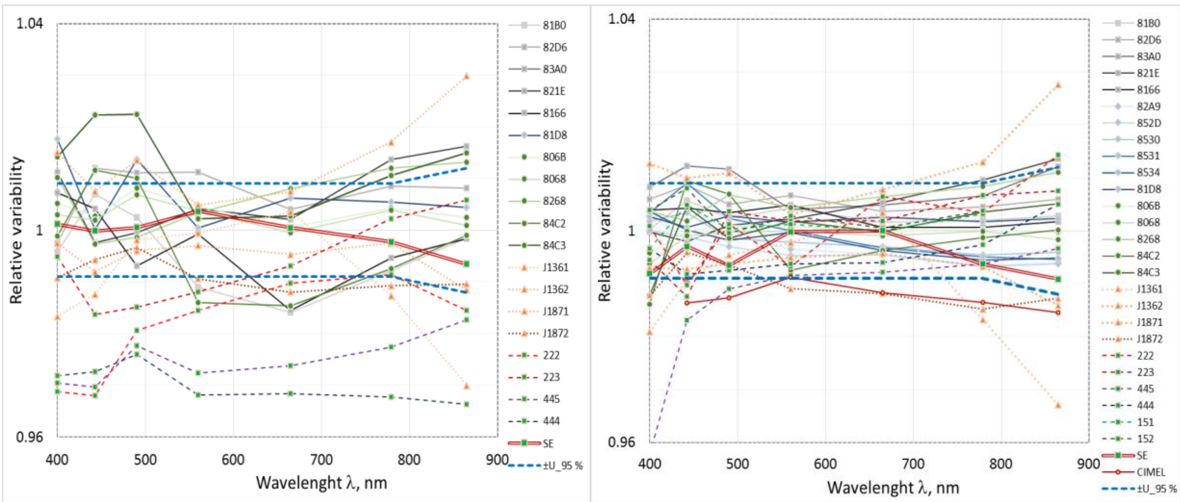


Figure 34. Low intensity radiance; agreement using data from participants (left), and after data was reviewed by pilot (right). Blue dashed lines - expanded uncertainty covering 95 % of all data points on the right graph. Solid lines – RAMSES sensors; dashed lines - HyperOCR sensors; double line – SR-3500; dotted lines - WISP-3 sensors.

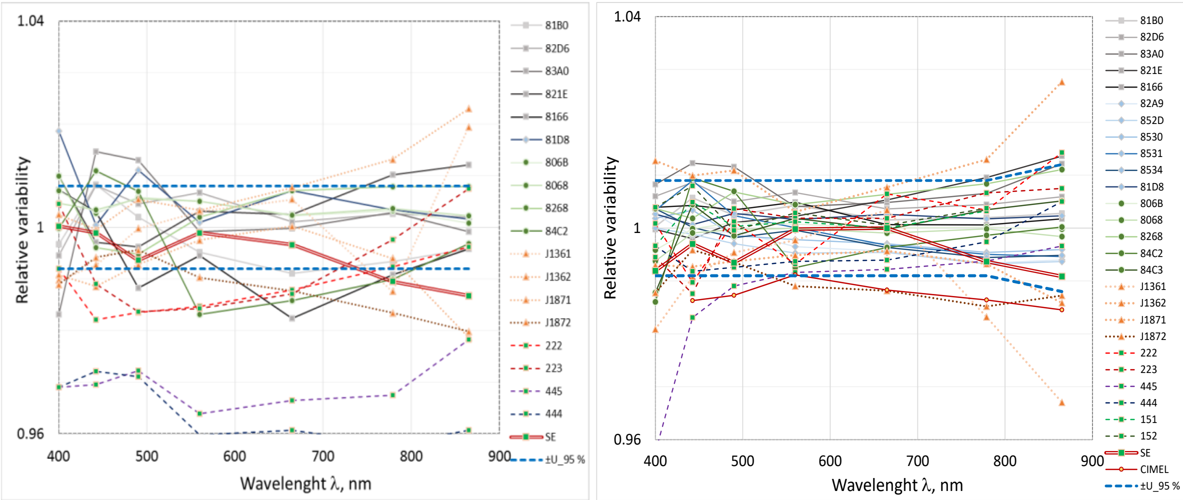


Figure 35. High intensity radiance; agreement using data from participants (left), and after data was reviewed by pilot (right).

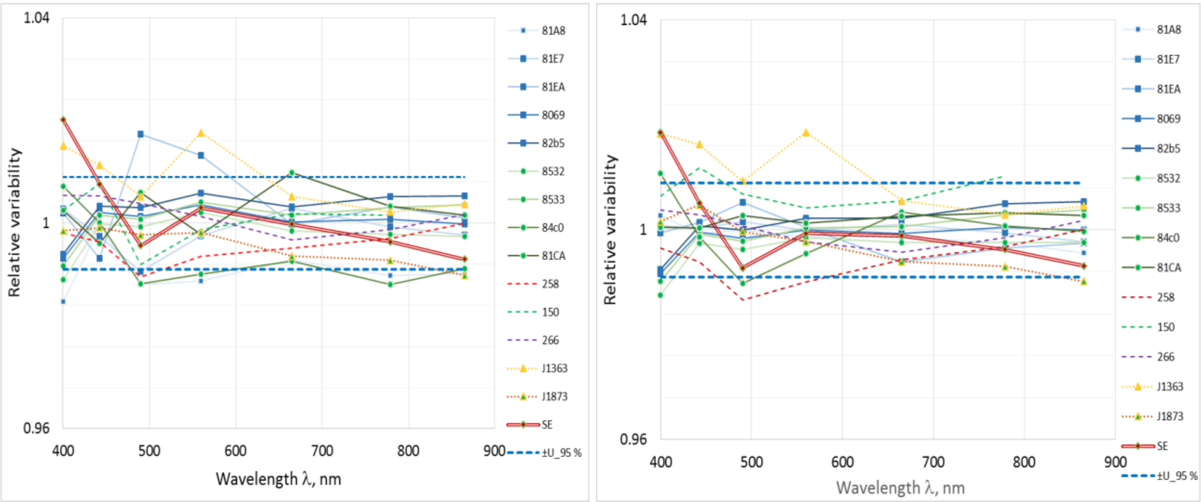


Figure 36. Irradiance sensors; agreement using data from participants (left), and after data was reviewed by pilot (right).

The larger variability in the results initially reported by participants was caused by applying out-of-date calibration coefficients, by diversely applying or not applying the non-linearity correction or by calculating the OLCI band values differently. For unified data handling carried out by the pilot and described in 15.4.2.2 to 15.4.2.4, 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. After unified data handling, agreement between the comparison results was significantly improved for the radiance sensors (Figure 34 - Figure 36). There was almost no improvement in the case of the irradiance sensors in Figure 36.

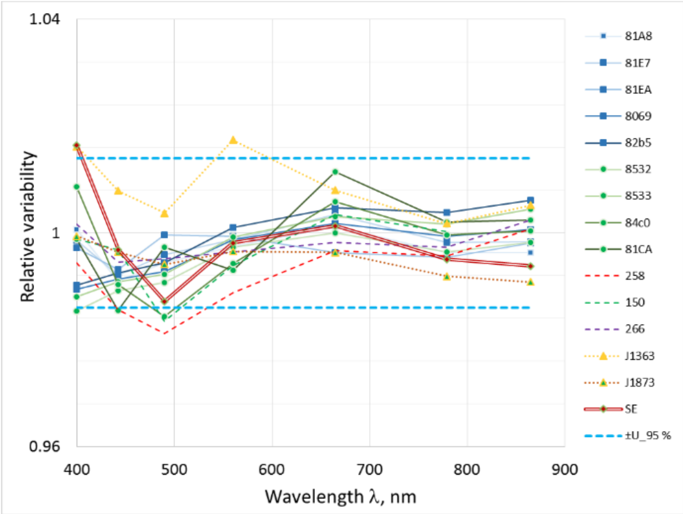


Figure 37. Irradiance sensors; agreement with reference values of the filter radiometer. Blue dashed lines - expanded uncertainty covering 95 % of all data points. Uncertainty of radiometric calibration is included.

15.4.2.7 Effects causing variability of the results

15.4.2.7.1 State 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 on the same set of radiometers by TO one year later (before the FRM4SOC FICE-AAOT intercomparison [167]) demonstrates the importance of radiometric calibration for SI traceable results and reveals interesting information about instability of the sensors. Some uncertainty contributions characteristic of calibration can also be estimated.

The 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 is depicted in Figure 38. All the radiometers were calibrated before LCE-2 using the same pair of lamps (section 17.4.1). Two sets of calibration coefficients were obtained for each sensor and the difference between the lamps was presented as the ratio of these coefficients. The curves in Figure 38 are calculated as standard deviations from the ratios of a whole set of calibration coefficients determined by using the two standard lamps. The systematic difference between lamps (due to small difference in traceability to SI) is neglected and only the other uncertainty components related to individual setting up and measurement of radiometers are accounted for by using the standard deviation. Data in Figure 38 include calibration of more than 25 sensors for the LCE-2 intercomparison and for the FICE-AAOT intercomparison [167] one year later when a different pair of lamps was used. Remarkable in Figure 38 is the rapid increase of variability between sensors in the UV region.

Figure 39 shows the 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 39 are calculated similarly to Figure 38 as the standard deviations of the ratios of previous and the last calibration coefficients. It has to be noted, however, that in this case standard deviation is characterizing dispersion between previous calibrations as these were performed by using various standards and conditions. 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 [167]. Those sensors were radiometrically calibrated again 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 calibration standard lamps were exchanged since LCE-2. Nevertheless, the L_1 yr and E_1 yr curves in Figure 39 obtained as standard deviations of the ratios of the calibrations coefficients one year apart exclude the systematic differences between lamps. The two calibrations done in the same lab one year apart showed that over 80 % of the sensors have changed less than ±1 %. Thus, the inherent long-term stability of the sensors is likely better than 5 % to 10 % revealed from the previous calibration history, where the differences were likely caused by other factors such as different calibration standards, environmental conditions, calibration setups and methodologies, etc. However, rapid 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 quick changes were observed for the RAMSES radiance sensors, however, even after omitting outliers from the stability data of irradiance sensors, the stability of RAMSES radiance sensors is still better.

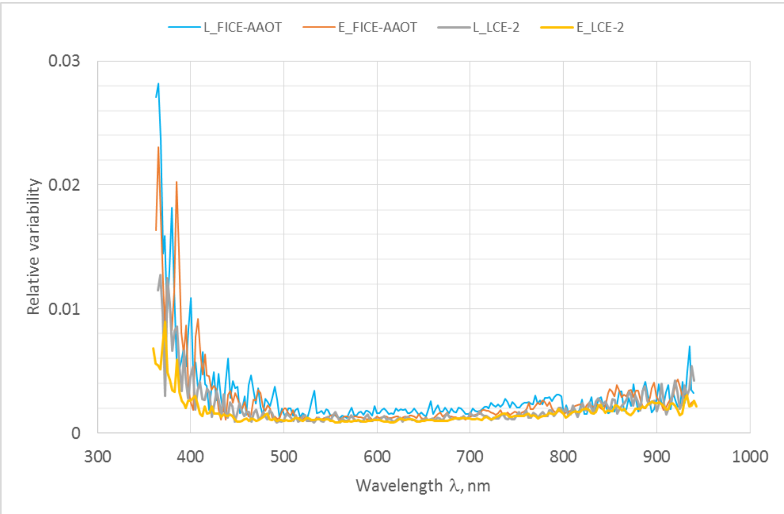


Figure 38. Relative variability of calibration coefficients of radiance (L) and irradiance (E) sensors with two different lamps used for calibration before LCE-2 and a year later before FICE-AAOT.

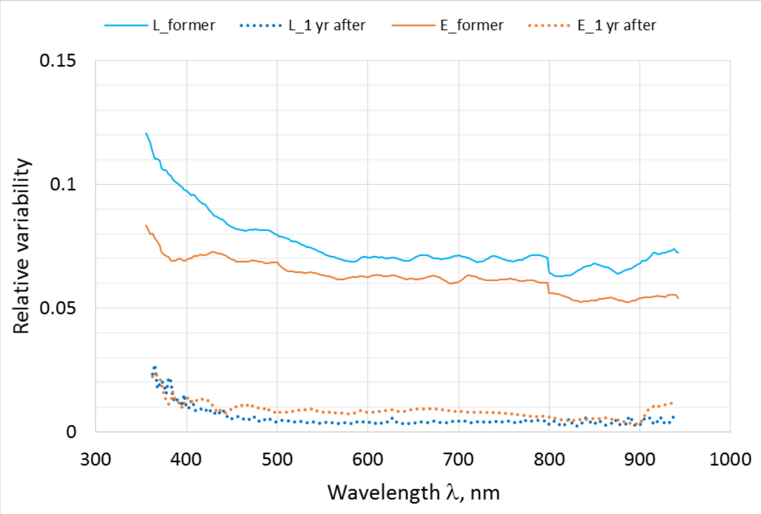


Figure 39. Relative variability of calibration coefficients of radiance (*L*) and irradiance (*E*) sensors: former - difference of previous known calibrations and results of LCE-2 calibration; 1 year after - changes during one year after LCE-2 calibrations, some extra-large changes excluded.

15.4.2.7.2 Abrupt changes of responsivity

Factors causing the variability in the responsivity of radiometers were listed in [166]. During the calibration, the uncertainty of the radiation source is the dominant component in the uncertainty budget, assuming that usually the ambient temperature will be within ±1 °C. Based on the experience from LCE-2 and the following FICE activities, differences smaller than ±2 % in the wavelength range of (350...900) nm can be observed between different sources used for calibration. Nevertheless, in some cases sharp changes in the responsivity of radiometers were detected, substantially exceeding all possible effects which can cause variability during calibration like the radiation source, alignment of instruments, contamination of foreoptics, temperature effects, etc. Relative change of the spectral irradiance responsivity of the TriOS RAMSES SAM_8329 calibrated ten times during the previous eight years period is depicted in Figure 40. Each calibration frpm 2016 – 2018 consisted of three repetitions conducted over a short time.

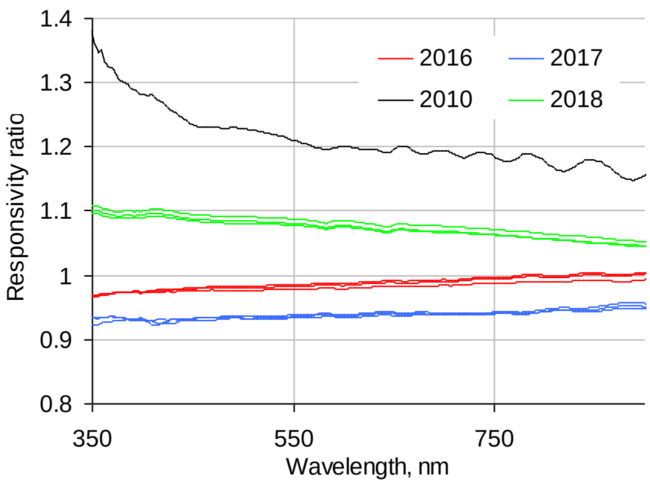


Figure 40. Relative change of responsivity of the SAM 8329. Year of the radiometric calibration is shown with colour: 2010 black, 2016 red, 2017, blue, 2018 green.

15.4.2.7.3 Temperature effects

Individual variation of the calibration coefficients as a function of temperature for each radiometer was not determined because of the limited time schedule of LCE-2 and FRM4SOC. Temperature effects for the TriOS RAMSES radiometers were evaluated based on [180], see Figure 41.

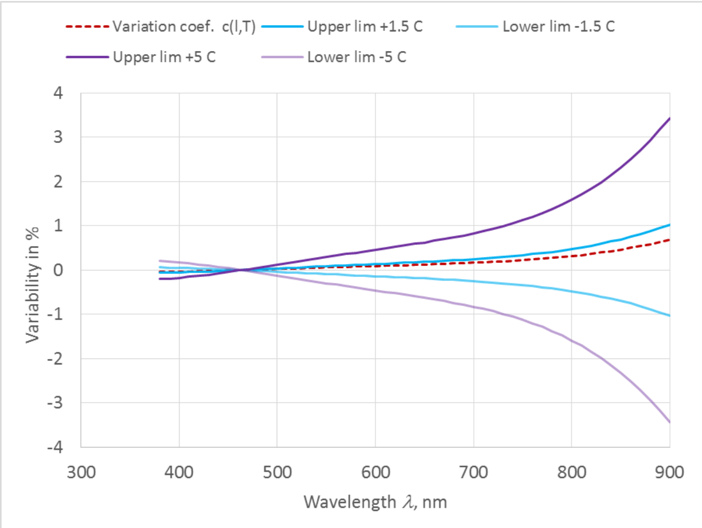


Figure 41. Relative variability of calibration coefficients due to temperature deviations from the reference temperature 21.5 °C.

15.4.2.7.4 Non-linearity due to the integration time

The maximum relative non-linearity effect due to integration times determined from calibration spectra of TriOS RAMSES radiometers remained in the range of (1.5...3.5) % (Figure 42). Variability between the instruments due to this effect, if not corrected, will mostly be in the range of ±1 %. The effect can be corrected down to 0.1 % for certain types of radiometers by using the special formula, if there are at least two spectra with different integration times available for the same source. Correction formula is based on the following assumptions:

i) the non-linearity effect is zero for the dark signal;

ii) the effect is proportional to the recorded signal;

iii) the effect is wavelength dependent, and

iv) the size of the corrected signal does not depend on the initial pair of spectra used for estimation i.e. it should be the same for all possible pairs used with equation (16)

Linearity corrected raw spectrum $S_{1,2}(\lambda)$ is calculated as

$$S_{1,2}(\lambda) = \left[1 - \left(\frac{S_2(\lambda)}{S_1(\lambda)} - 1\right)\left(\frac{1}{t_2/t_1 - 1}\right)\right] S_1(\lambda). \tag{16}$$

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 non-linearity correction is not needed. Uncertainty of the corrected spectrum is predominantly determined using the uncertainty of the initial spectrum measured with a smaller integration time. Therefore, the smallest uncertainty of the corrected spectrum will be obtained if the initial spectra with the largest and with the second-largest non-saturating integration times are used for estimation.

The formula has been found to perform quite effectively for TriOS RAMSES and Satlantic HyperOCR radiometers in the range of (400…800) nm. This non-linearity correction method is not recommended for outdoor measurements, as due to temporal variability of the natural radiation consecutive measurements with different integration times may lead to the uncertainty of the corrected spectrum much larger than the acceptable 0.2 %.

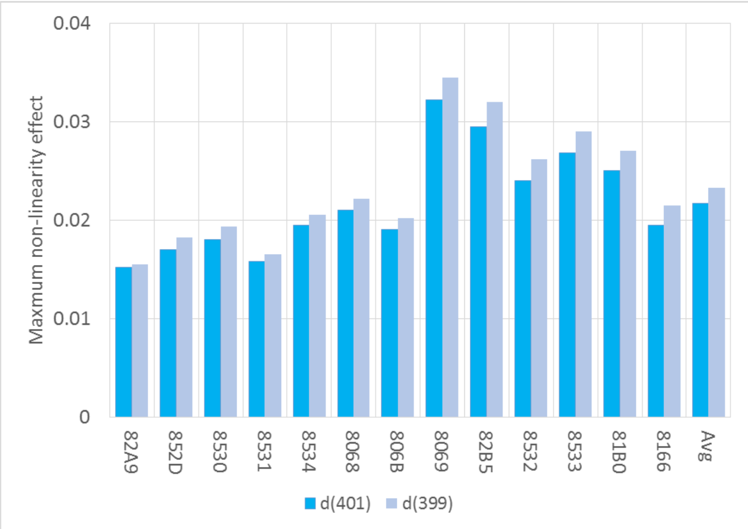


Figure 42. Maximum relative non-linearity effect determined for 14 RAMSES sensors (both radiance and irradiance) from calibration spectra using FEL lamps 399 and 401.

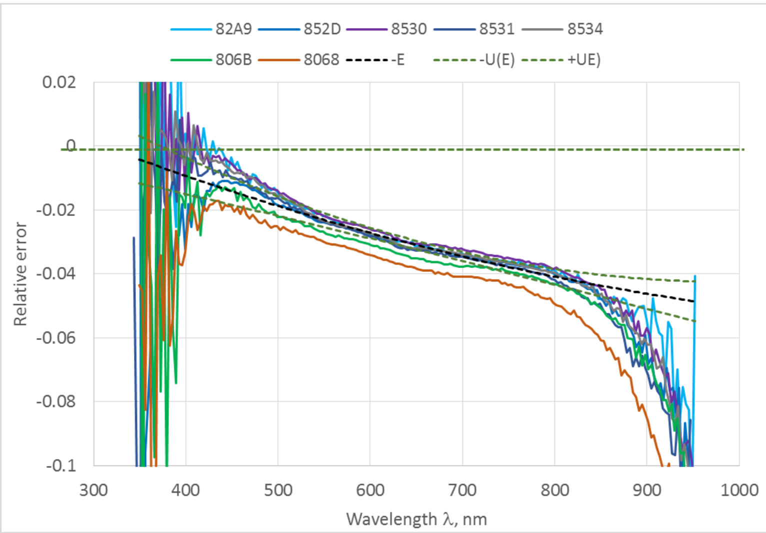


Figure 43. Non-linearity errors of different radiance sensors scaled to the full-range. Dashed lines are the fitted model with uncertainty.

From the analysis of the non-linearity data obtained by the two-spectra formula (16), it became evident that the non-linearity errors of different radiance sensors behave as if re-calculated to a respective full-scale value. This serves as a basis for derivation of non-linearity correction applicable to a single spectrum, and thus, can be used for the field measurements, see [D-170], [135].

The relative non-linearity correction for the full scale signal $\delta x_{\max}(\lambda)$ is:

$$\delta x_{\max}(\lambda) = -5.1 \cdot 10^{-8} \lambda^2 + 0.00014 \cdot \lambda - 0.0355. \quad (17)$$

The relative non-linearity correction $\delta x(\lambda)$ for the signal $x(\lambda)$ is:

$$\delta x(\lambda) = \frac{x}{x_{\max}} \delta x_{\max}(\lambda). \quad (18)$$

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

$$x_{\mathrm{cor}}(\lambda) = x(\lambda) \left[1 + \frac{x}{x_{\max}} \delta x_{\max}(\lambda)\right]. \quad (19)$$

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

15.4.2.7.5 Spectral stray light effects

For many measurements, spectral stray light can lead to significant distortion of the measured signal and become a significant source of uncertainty [46,186]. An iterative technique [186,196] can be used for the simultaneous correction of bandpass and stray light effects. When the spectral stray light matrix (SLM) of a spectrometer is known, the stray light contribution can be removed from the measured signal and the original source spectrum restored. The stray light correction for a remote sensing reflectance measurement made by a common 3-radiometer above-water system means that altogether six raw spectra have to be corrected - two for each radiometer, because stray light correction needs to be applied also for the standard source spectrum during the radiometric calibration.

The SLM was known for some radiometers from previous characterisation such as for RAMSES sensors of TO, and for HyperOCR sensors of Plymouth Marine Laboratory (PML). Figure 44 presents the impact of stray light correction, evaluated for indoor measurements. The indoor radiance and irradiance sources were spectrally similar to the calibration sources; therefore, the stray light correction has a relatively small impact. WISP-3, SR-3500, and SeaPRISM have different optical designs, thus, their spectral stray light properties can be different compared to the data presented in Figure 44.

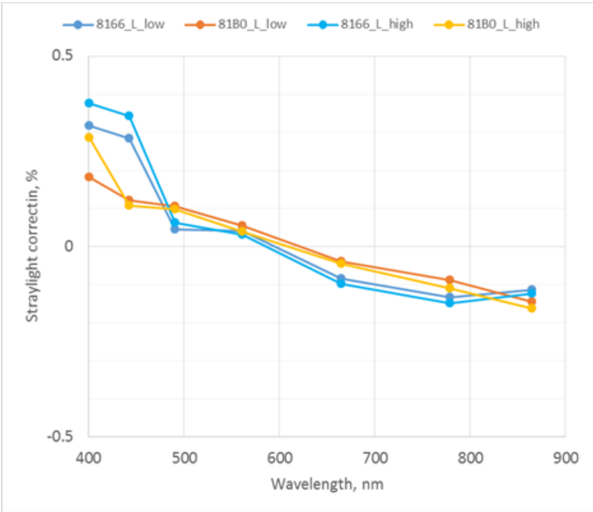


Figure 44. Relative stray light effects for indoor radiance measurements. Two RAMSES radiance sensors measuring sphere radiance at high and low intensity.

15.4.2.8 Measurement uncertainty

The uncertainty analysis has been carried out according to the GUM [18]. The evaluation is based on the measurement model, which describes the output quantity Y as a function f of input quantities X_i: $Y = f(X_1, X_2, X_3 \ . \ . \ .)$. For every input quantity X_i, respectively, the estimate x_i and standard uncertainty $u(x_i)$ are evaluated, which are considered as parameters of the probability distribution describing the X_i. The combined standard uncertainty $u_c(y)$ for the output estimate is calculated from the standard uncertainties associated with each input estimate x_i, using a first-order Taylor series of $y = f(x_1, x_2, x_3, \ldots)$. There are two types of standard uncertainties: Type A is of statistical origin; Type B is determined by any other means. Both types of uncertainties are indicated as standard deviations and denoted by s and u respectively. The uncertainty component arising from averaging a large number of repeatedly measured spectra of radiation sources by array spectrometers 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 all input quantities, relative standard uncertainties are estimated. The relative combined standard uncertainty of the output quantity is calculated by combining the relative standard uncertainty of each input estimate by using equation (12) in [18]. Uncertainty of the result is given as a relative expanded uncertainty with a coverage factor $k = 2$.

Uncertainty of the irradiance and radiance measurements of TriOS RAMSES sensors is analysed in laboratory conditions. Uncertainty for irradiance of an FEL source measured at approximately 1 m distance is given in Table 12, and for radiance of an integrating sphere, in Table 13.

All the uncertainty estimates of RAMSES sensors, besides experimental data, are based on information from [1,46,154,185,197–199]. For the other radiometer models that took part in the intercomparison, very little information about instrument characteristics that influence the measurement results is publicly available [54]. In addition, only the RAMSES sensor was sufficiently numerous for statistical analysis.

The uncertainty budgets (Table 12- Table 13) describe variability between individual sensors, while uncertainty of radiometric calibration and other contributions, which influence all the instruments in a similar way, are not fully accounted for. Nevertheless, for traceability to SI units, these contributions are relevant.

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 contributions to the spectral irradiance and/or radiance due to the setting and measurement of the lamp current, 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 the averaged signal are included.

15.4.2.8.1 Calibration certificate

The calibration certificate of a radiometer provides calibration points of radiometric responsivity following the individual wavelength scale of the radiometer. The radiometric calibration uncertainty in Table 12 and Table 13 is presented for reference only and is not added to the combined uncertainty because all participating radiometers were calibrated using common standards shortly before the intercomparison, and the contribution from calibration when using unified data handling does not affect the relative differences of participants to each other.

15.4.2.8.2 Interpolation

Interpolation of radiometer data is needed due to differences between individual wavelength scales of the radiometers. Therefore, measured values were transferred to a common scale basis (Sentinel-3/OLCI bands) for comparison, (see section 15.4.2.4). 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. In Table 12 and Table 13 the interpolation components include the contribution of wavelength scale uncertainty estimated from data presented in Figure 60.

15.4.2.8.3 Temporal instability of radiometer

Instability of the radiometric responsivity can be estimated from data of repeated radiometric calibrations. For LCE-2, the instruments were calibrated just before the comparisons and only short-term instability relevant for the time needed for the measurements has to be considered. The values are derived from the data collected in the calibration sessions of LCE-2 and FICE-AAOT a year later, Figure 38. The variability over two weeks was interpolated from the yearly variability data. In addition to instability of the sensors, the data shown in Figure 38 includes other uncertainty components related to the calibration setup (e.g. alignment, short-term lamp instability, etc.).

15.4.2.8.4 Back-reflection

Back-reflection from the radiometer into the integrating sphere was estimated using different distances between the sphere and the radiometer as the contribution of radiation reflecting from the radiance sensor back into the integrating sphere.

15.4.2.8.5 Polarization

The polarization effect was estimated by indoor irradiance measurements, repeating the cast after the radiometer was rotated 90° around its main optical axis, and thus revealing the combined effect of alignment and polarization. According to [200] the FEL emission is polarized by about 3 %. As reported in [185], the polarization sensitivity of RAMSES irradiance sensors is varying from (0.05…0.3) % at 400 nm to (0.3…0.6) % at 750 nm. Due to the depolarizing nature of the cosine collector this effect is smaller than the polarization sensitivity of RAMSES radiance sensors. Therefore, the observed differences with rotated sensors are mostly caused by other effects like alignment, instability of the measured source, etc., and therefore the polarization component is omitted from the indoor irradiance uncertainty budget. Polarization is also not included in the indoor radiance uncertainty budget as the integrating sphere is a strong depolarizer.

15.4.2.8.6 Alignment

Evaluation of alignment errors includes determination of the distance between the source and the reference plane of the cosine collector, measured along the optical axis. Alignment includes also position errors of the lamp source across optical axes, rotation errors of the lamp [201], and positioning errors of the input optics of the radiometer. Combined alignment and positioning errors are included in variability data of radiometers calibrated with two different lamps (Figure 38).

15.4.2.8.7 Non-linearity

Due to non-linearity, some hyperspectral radiometers, measuring at different integration times may show relative differences of up to 4 %, see Figure 42. According to recommendations, the non-linearity effects of good sensors should be correctable to less than 0.1 %. The non-linearity correction (16) was applied to both calibration and measurement spectra, with residues expected to be less than 0.2 %.

15.4.2.8.8 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. The value is estimated from Figure 44, and from [190,196].

15.4.2.8.9 Temperature

For array spectroradiometers with silicon detectors, the present estimate of 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) [180].

15.4.2.8.10 Temporal instability of radiation source

The short-term instability of the source is relevant for the indoor measurements as they were not made simultaneously by all the participants. Thus, the time needed for intercomparison measurements, including power cycling the source between the two days of the 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 drift of the irradiance values (at 500 nm) measured during the calibration campaign was ~0.1 %, (Section 15.4.1).

15.4.2.8.11 Stray light in the laboratory

A source of stray light is associated with the stray light in the laboratory during the indoor experiment. This component has been estimated in previous experiments made in the TO.

15.4.2.8.12 Type A uncertainty of repeated measurements

For Type A uncertainty of the time series of indoor measurements, white noise in the measured series can often be expected. The analysis has indicated that sometimes the measurements are not completely independent and the autocorrelation of time series has been accounted for. If there is autocorrelation in the time series, the effective number of independent measurements has to be considered instead of the actual number of points n_t in the series [202]:

$$n_e \approx n_t \frac{1-r_1}{1+r_1}, \tag{20}$$

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

15.4.2.9 Conclusions from the LCE-2 indoor intercomparison

Altogether 44 radiometric sensors from 11 institutions were involved: 16 RAMSES, 2 OCR3000, 4 HyperOCR, 4 WISP-3 and 1 SE-3500 radiance sensors, and 10 RAMSES, 1 OCR3000, 2 HyperOCR, 2 WISP-3 and 1 SE-3500 irradiance sensors participated in the LCE-2 indoor intercomparison. Additionally, many of the sensors involved in LCE-2 were recalibrated at TO a year later (for FICE-AAOT) giving an estimate of their long-term stability. More than 80 % of the sensors changed during one year less than ±1 %.

Agreement between the radiometers is mostly affected by the calibration state of the sensor. For example, factory calibrations made at different times can cause differences exceeding ±10 %. Former calibrations in different laboratories 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) % for each factor. The best agreement of (0.5…0.8) % between participants has been achieved when measurements were carried out just after calibration and where data handling unified procedures have been used, including application of non-linearity correction and the same algorithm for calculation of OLCI band values.

Dependence of the calibration coefficients on temperature can also cause significant deviation in SI-traceable results, especially in the NIR spectral region. For a maximum temperature difference of about 20 °C between calibration and later measurements (typically between 0°C and 40°C) a responsivity change more than 10 % is possible [143,179]. For laboratory measurements in a controlled environment, the temperature effect is expected to be within (0.1…0.5) %.

The effect of stray light correction evaluated for indoor measurements in the range (400 … 700) nm has been shown to be less than 0.5 %. Though, outside the range of (400 … 700) nm the relative uncertainty may increase substantially if correction is not applied.

The maximum value of the non-linearity effect due to integration times determined from calibration spectra of TriOS RAMSES radiometers for a group of 15 radiometers was in the range of (1.5…4) %. At the same time, the variability between the instruments due to this effect if not corrected, remained within ±1 % due to the systematic nature of the nonlinear behaviour affecting all the instruments in a similar manner. During laboratory measurements the non-linearity correction was applied to both calibration and measurement spectra, with residuals expected to be less than 0.2 %.

Table 12. Relative uncertainty budget for the irradiance (in percent), based on the spread of individual sensors measuring the same lamp during the indoor comparison. Data highlighted in green are not used for combined and expanded uncertainties.

	400 nm	442.5 nm	490 nm	560 nm	665 nm	778.8 nm	865 nm
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
Non-linearity	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
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
Signal, type A	0.11	0.04	0.02	0.02	0.01	0.02	0.04
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 13. Relative uncertainty budget for the radiance (in percent) based on the spread of individual sensors measuring the same integrating sphere during the indoor comparison. Data highlighted in green are not used for combined and expanded uncertainties.

	400 nm	442.5 nm	490 nm	560 nm	665 nm	778.8 nm	865 nm
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
Non-linearity	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
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
Signal, type A	0.12	0.07	0.04	0.02	0.03	0.03	0.06
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

Field of view and cosine responsivity effects can significantly depend on the limits of error set based on the specifications of radiometers and on the results of individual tests showing how large a deviation there is from the specified values. In the laboratory, the cosine responsivity error of the sensor during calibration was close to the error during the intercomparison measurements due to a similar illumination geometry, and therefore, the resulting systematic error is insignificant.

Through the indoor experiment, when conditions for later measurements and conditions specified for calibration were quite similar, a high effectiveness of the SI-traceable radiometric calibration has been demonstrated, as the large group of different type radiometers operated by different scientists achieved satisfactory consistency between results showing low standard deviations between radiance (27 in total) or irradiance (15 in total) results ($s < 1$ %). This is provided that some unification of measurement and data processing is applied: alignment of sensors, structure of collected data, application of unified wavelength bands and non-linearity corrections. Nevertheless, variability between sensors may be insufficient for the complete quantification of uncertainties in the measurements. For example, standard deviation of individual non-linearity estimates (Figure 42) cannot reveal an average error of all instruments, not to mention the error limits. Therefore, all radiometers should be individually tested for all significant systematic effects, which may affect the results, as this is the only way to get a full estimate of the effects degrading traceability to the SI scales.

15.4.3 Outdoor intercomparison

15.4.3.1 Venue and measurement setup

The outdoor exercise took place at Lake Kääriku, Estonia, 58° 0' 5" N, 26° 23' 55" E on 11 – 12 May 2017. Lake Kääriku is a small eutrophic lake with 0.2 km² surface area. Maximum depth is 5.9 m, with an average of 2.6 m. The water colour is greenish-yellow with a measured transparency (Secchi disk depth) of 2.6 m. The average chlorophyll content, Chl_a = 7.3 mg m^{-3}, total suspended matter content, TSM = 3.9 g m^{-3}, absorption of the coloured dissolved organic matter, a_{CDOM}(442 nm) = 1.7 m^{-1}, diffuse attenuation coefficient of downward irradiance K_d(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 45).

Figure 45. Pier and diving platform at the southern coast of Lake Kääriku.

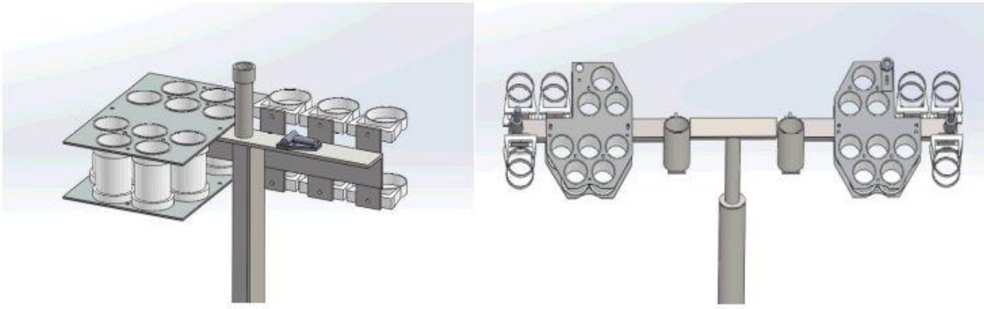


Figure 46. Mounting frame for the irradiance sensors (left) and for the radiance sensors (right) used during the outdoor experiment, shown as 3D CAD drawings.

Figure 47. 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 instruments were located roughly 7.5 m above the water surface. Depth of the water around the diving platform is 2.6 m to 3.6 m and the bottom was not visible to observers. 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 46 - Figure 47). The irradiance sensors were mounted in a fixed frame ensuring 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.

15.4.3.2 Environmental conditions and selection of casts

The environmental conditions during the outdoor experiment were not ideal, mainly due to the scattered cumulus clouds. The aerosol content was low, average daily aerosol optical depth at 500 nm (AOD500) was 0.077 on May 11 and 0.071 on May 12 (measured at Tõravere AERONET station, 30 km north of Lake Kääriku [203]). 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}.

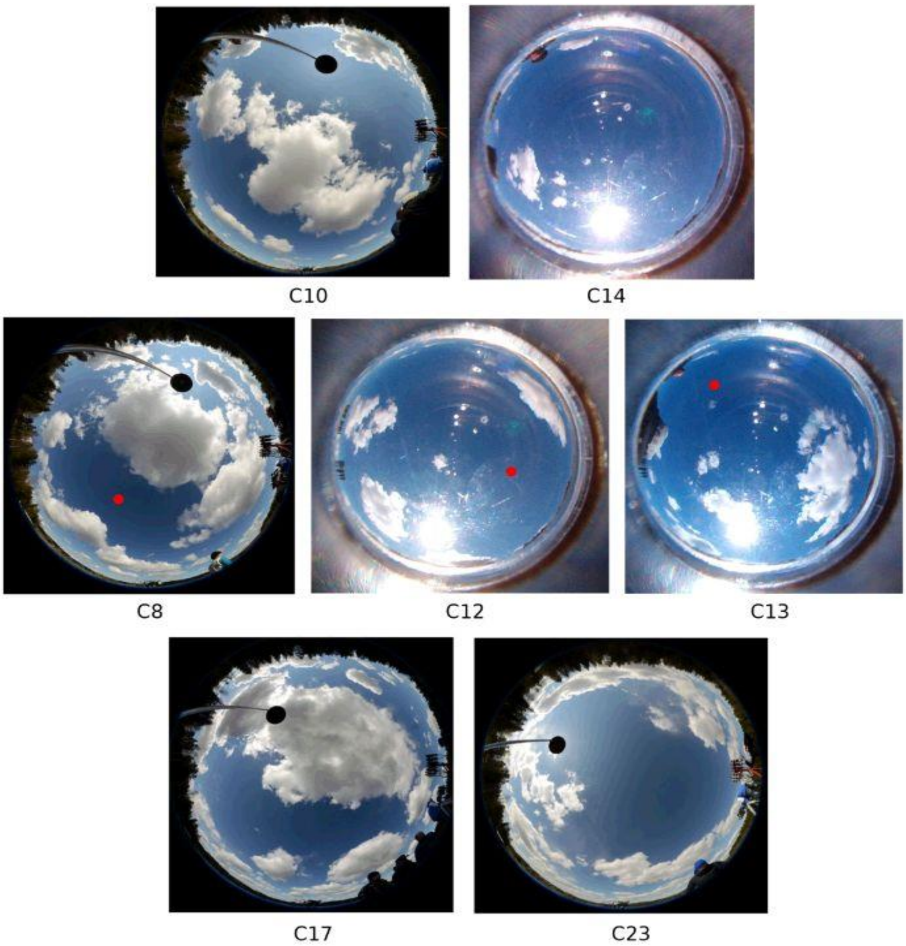


Figure 48. All-sky camera images captured in the middle of the casts used in the analysis. Irradiance - C10, C12, C13, C14; blue sky radiance - C8, C12, C13; water radiance - C17, C23. Red dots in C8, C12, C13 indicate approximate view direction of the radiance sensors.

The outdoor measurements were performed in 5-min casts, with the exception of the 25-min 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 analysis. The selection of casts was based on the time series of the 550 nm spectral band data. The coordinating laboratory 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 48.

The casts used in the analysis of LCE-2 outdoor intercomparison are listed in Table 14. Four casts (C10, C12, C13, and C14) were chosen for irradiance, all recorded with direct sunlight, although with some clouds in the sky away from the sun. Five casts were chosen for radiance: three casts (C8, C12, and C13) recorded with blue sky as the target, one (C17) measurement of the water surface in cloud shadow, and one (C23) measurement of sunlit water. Measurement C17 was made at a zenith angle suggested in the protocols for above-water radiometry, while measurement C23 is made at a slightly more oblique angle. These measurements were made for azimuth angles 107° and 143° with respect to the sun, which avoid sunglint and direct shadow from the platform. 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 49.

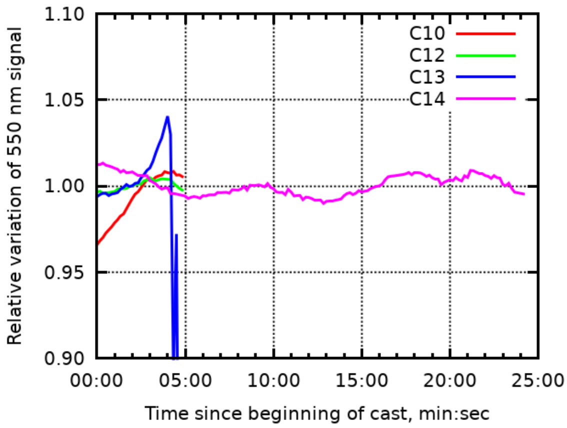


Figure 49. Relative variation of 550 nm signal of one RAMSES sensor during irradiance (left) and radiance (right; C8, C12, C13 blue sky; C17 water in cloud shadow; C23 sunlit water) casts selected for intercomparison analysis.

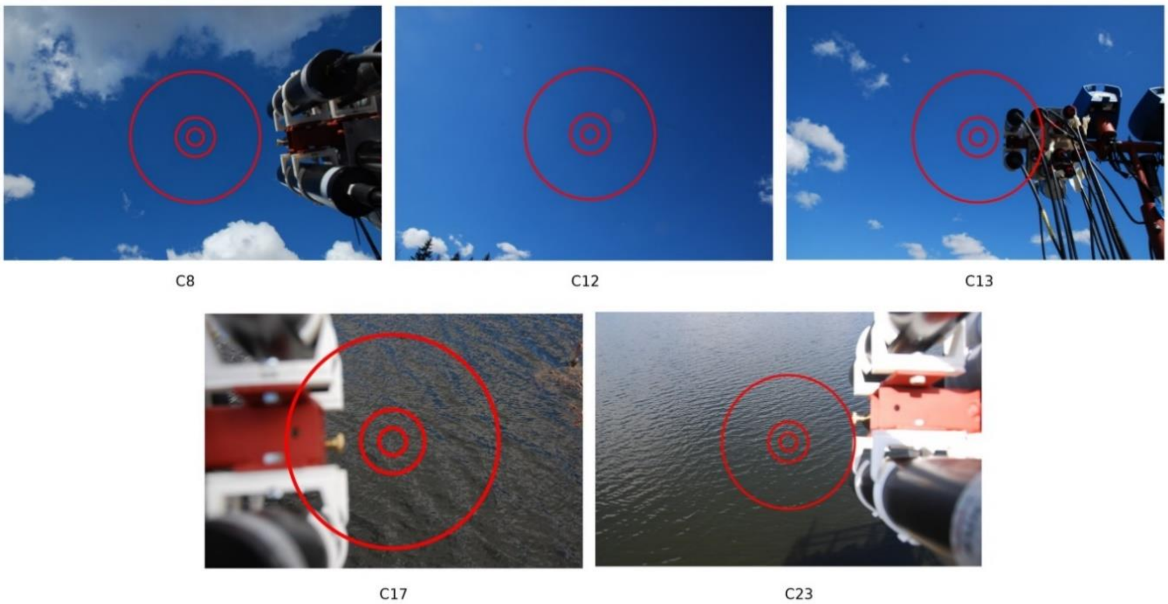


Figure 50. Photographs of radiance targets used in the intercomparison analysis. The circles denote approximate FOV of WISP-3 (smallest), RAMSES, and HyperOCR (largest)[10].

The initial cast start and stop times were adjusted based on Figure 49 to exclude the intervals with a high temporal variability. Photographs of the radiance targets can be seen in Figure 50. Approximate FOV footprints for WISP-3 (3°), RAMSES (7°), and HyperOCR (23°) are also shown in Figure 50. Although by the time of writing this final report it has been

[10] According to the manufacturer, the HyperOCR radiance sensors 444 and 445 have 6° FOV. After the report [D-170] and the Remote Sensing special issue paper [168] were published, it was confirmed with additional characterisation, that the rest of the HyperOCR and OCR-3000 instruments participating in LCE-2 have also 6° FOV. Thus, the FOV of those sensors is slightly smaller compared to the TriOS RAMSES sensors having 7° FOV.

confirmed that all the participating HyperOCR and OCR-3000 radiance sensors were the narrower FOV variants having 6° FOV. Therefore, the FOV footprints of those radiance sensors is slightly smaller compared to that of RAMSES (7°) in Figure 50. The images were taken with a handheld Nikon D40X digital single-lens reflex (DSLR) camera equipped with a Nikkor 18-200 mm zoom lens. According to the Exchangeable image file format (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 in line with the radiometers, its collinearity with the radiometers is uncertain and the actual FOVs of the radiometers may slightly differ from the circles shown in Figure 50.

Table 14. Casts used in the 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°	NA
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	NA
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}

List of Abbreviations and symbols used in the Table 14

UTC - coordinated universal time
SZA - solar zenith angle
SAA - solar azimuth angle
VAA - view azimuth angle
VZA - view zenith angle
NA - not applicable
L_d - downwelling sky radiance
L_u – total upwelling water radiance
E_d - downward irradiance

15.4.3.3 Outdoor experiment of the LCE-2

Initially [193], the outdoor intercomparison was planned in two phases: (I) direct intercomparison of the downward irradiance E_d, the downwelling sky radiance L_d, and the total upwelling water radiance L_u; (II) intercomparison of the remote sensing reflectance R_{rs} and the water-leaving radiance L_w derived from simultaneously measured E_d, L_d, and L_u. 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 the 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 [198].

On the first day of the outdoor measurements, seven casts of simultaneous E_d, L_d, and L_u measurements using the typical above-water 3-radiometer configuration were recorded but due to cumulus clouds causing rather unsteady illumination conditions, none of the casts was considered suitable for the analysis. On the second day of the outdoor experiment, priority was given to the phase (I) measurements and all the radiance sensors were simultaneously measuring either L_u or L_d.

15.4.3.4 Data processing

In total, data for 40 out of 44 radiometers were reported back to the pilot. For the rest, the pilot carried out the data handling using the provided raw files. The data processing details are described in [204]. The outdoor data processing chain contained the following steps:

- separation of the raw datafiles based on the casts' start and stop timestamps;
- subtraction of the dark signal;
- division by the radiometric responsivity;
- interpolation/convolution of spectra into the OLCI bands.

15.4.3.5 Consensus value used for the analysis

The group median was used as the consensus value. 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 sensors and one radiance sensor were not accounted for in the variability estimate, because they had extremely large deviations from the group median.

15.4.3.6 Accuracy of the sensor adjustment

The collinearity of groups of radiance sensors on the left and right frame was set by visual observation from the side of the frame and was better than ±1°. Due to the flexibility of the plastic clamps used to fix the HyperOCR radiometers, a slight deflection from collinearity of HyperOCR and RAMSES sensors within the groups was noticed during the experiment (visually much larger than misalignment between the groups).

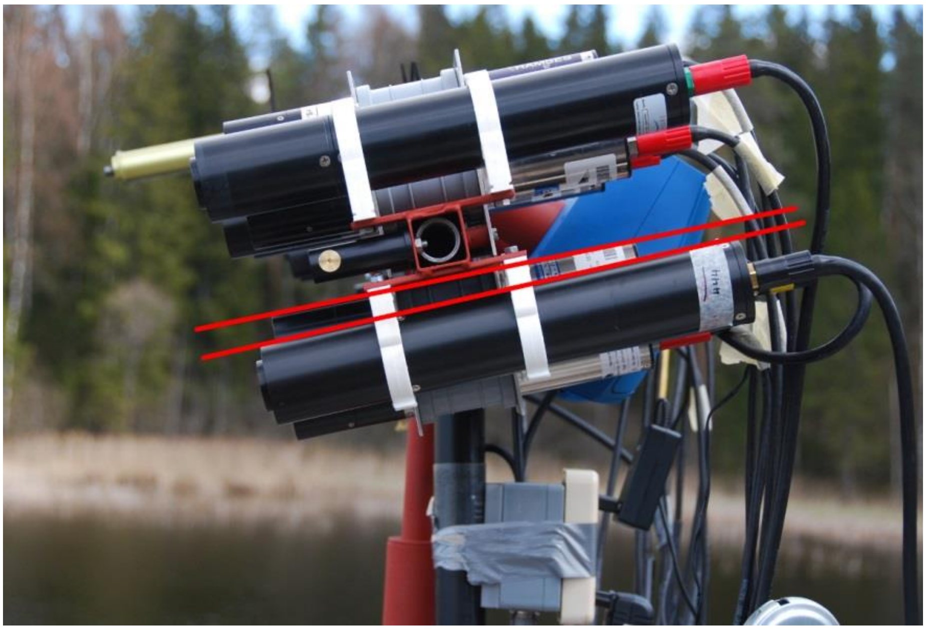


Figure 51. The angle between red lines marking the directions of HyperOCR and RAMSES sensors was measured to be 1.3° from this image.

Using Figure 51, the angle between HyperOCR and RAMSES sensors was measured to be 1.3°, the HyperOCR sensors were pointing lower than the RAMSES instruments. An 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 maximum angle between the HyperOCR instruments on the frames could have been about 2.5°. Although this is less than half of the FOV of the HyperOCR instruments, it can have a significant impact when measuring spatially heterogeneous targets.

15.4.3.7 Results

Compared to the indoor experiment [204], much larger variability between radiometric sensors can be expected in the outdoor experiment due to much larger differences in target signal and environmental temperature with respect to the radiometric calibration conditions.

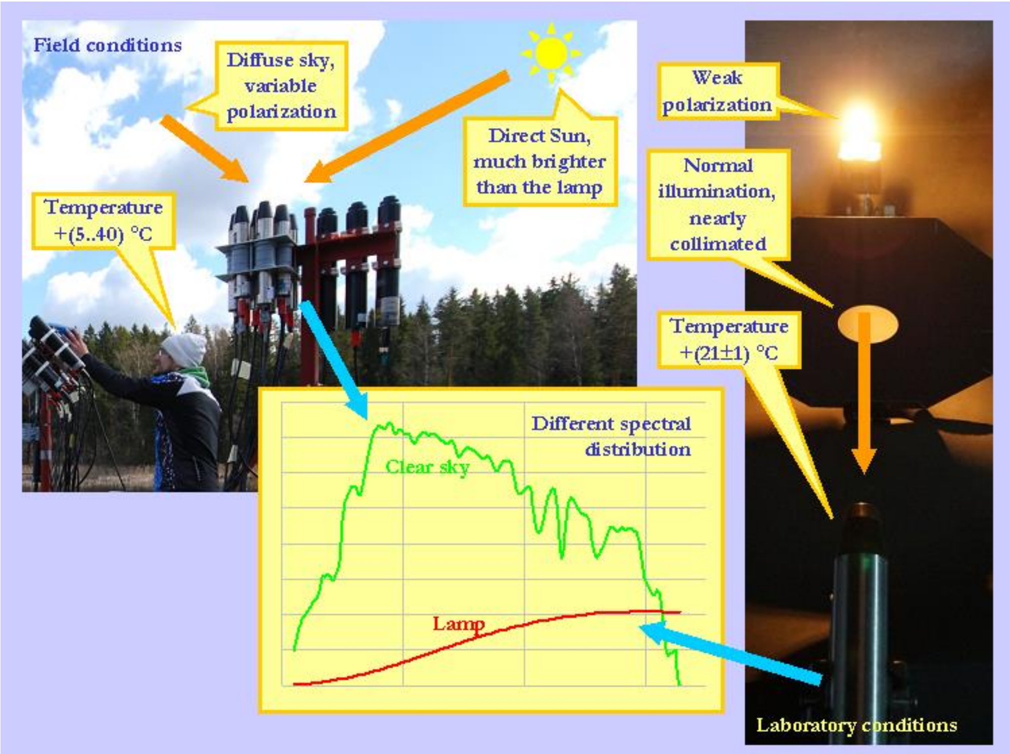


Figure 52. Main differences between the field and laboratory measurements of LCE-2 causing substantial increase in uncertainty of the field measurements.

The analysis of field measurement data is more complicated than in the case of indoor intercomparison [204]. The main differences in field and laboratory measurements of LCE-2, causing substantial increase in uncertainty of the field measurements, are shown in Figure 52. The spectral composition and intensity of radiation from the target being measured (sky, water) is significantly different from the incandescent source used as the radiometric calibration standard. The angular distribution of downward irradiance is also very different from the nearly collimated radiation source used during radiometric calibration. Ambient temperature in the field can differ from the stable laboratory temperature during the radiometric calibration by more than ±15 °C. The stray light effect may be an order of magnitude larger due to different shapes of the calibration and field spectra. Strong autocorrelation in recorded time series implies that statistical analysis of intercomparison results should be suitably rearranged.

Due to non-ideal performance of radiometers (temperature dependence, deviation from ideal cosine response for irradiance sensors, non-linearity, spectral stray light, etc.) all the differences between conditions during radiometric calibration and field measurements can contribute to the bias between radiometers and increase the measurement uncertainty. The known measurement errors should be corrected and the unknown or residual errors have to be assessed and accounted for in the uncertainty budget. Unfortunately, information needed for these corrections is often available only through highly time- and resource-consuming individual tests of radiometers and it is often necessary to make such corrections based on characterisation of an instrument from the same family.

15.4.3.8 Results of outdoor comparison

The consensus spectra for the irradiance and radiance targets are presented in Figure 53. The difference between the casts of radiance sensors measuring the sky and water is evident. Radiation from water with blue sky gave the smallest signal.

The measurement results for the field casts are presented as relative deviations from the consensus value in Figure 54. Rather different behaviour of RAMSES and HyperOCR sensor groups became evident. For the irradiance measurements, the deviation of HyperOCR sensors from the consensus value was very small, and the group of RAMSES sensors caused the increase of mean variability, see Figure 54. Conversely, variability of the radiance sensors during the indoor and outdoor exercises were almost at the same level for the RAMSES group, and the increase of the outdoor variability was caused largely by the HyperOCR sensors, see Figure 55.

All the irradiance casts in Figure 54 were measured with direct sunshine and no big difference between casts can be observed for the consensus irradiance spectra (Figure 53). The group of HyperOCR sensors shown in Figure 54 with dashed lines is more consistent with the consensus value than the sensors of the RAMSES group shown with solid lines. Remarkable is the much higher variability across the sensors of the RAMSES group. Interestingly, the intra-sensor variability of irradiance is almost wavelength-independent, except at 400 nm.

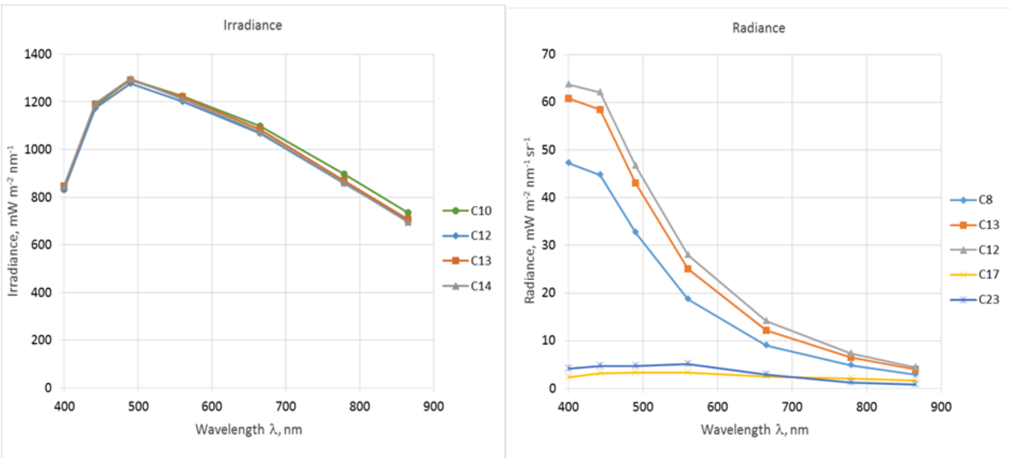


Figure 53. Irradiance and radiance consensus values in the outdoor experiment. C8, C10, C12, C13, C14 - blue sky (radiance) or direct sunshine (irradiance); C17 – water in cloud shadow; C23 – sunlit water.

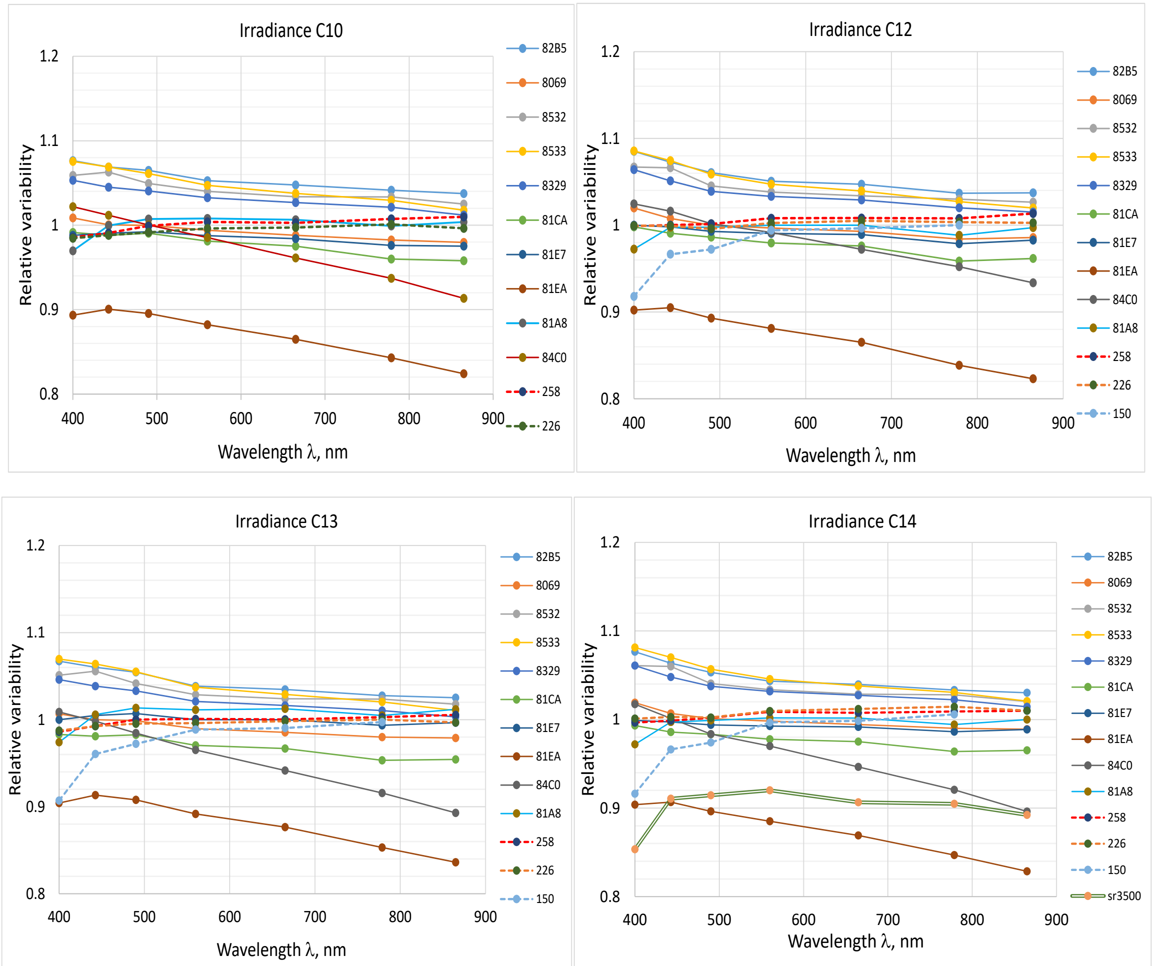


Figure 54. Irradiance sensors compared to the consensus value. Solid lines – RAMSES sensors; dashed lines - HyperOCR sensors; double line – SR-3500.

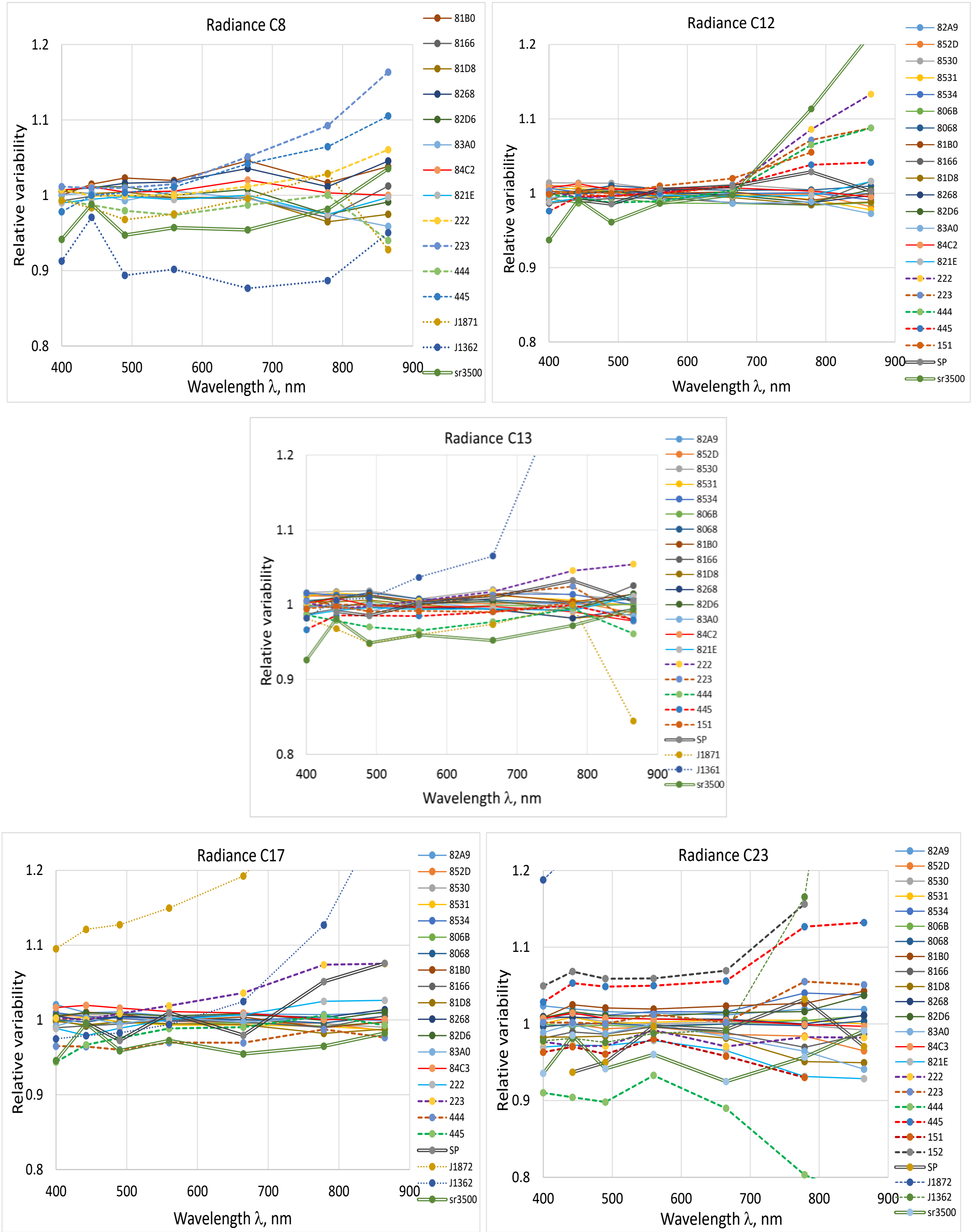


Figure 55. Radiance sensors compared to the consensus value in the outdoor experiment. C8, C12, C13 - blue sky; C17 – water in cloud shadow at 139° VZA; C23 – sunlit water at 130° VZA. Solid lines – RAMSES sensors; dashed lines - HyperOCR sensors; double lines – SeaPRISM (SP) and SR-3500; dotted lines – WISP-3.

The comparison of different radiance sensors (Figure 55) did show a very good agreement to within 1.2 % across the full spectrum for all RAMSES sensors for casts C12 and C13 - the most homogeneous blue sky targets. Higher variability between all sensors, and particularly the HyperOCR radiance sensors, is seen for the obliquely viewed water target C23 (Figure 55). This is probably caused by spatial heterogeneity of the target (C23 in Figure 50), and by slight bias from collinearity of the sensors (Figure 51). This assumption is supported by the fact that radiometers 151, 222, and 444, which were mounted on the left frame are below the consensus value in Figure 55, and radiometers 152, 223, and 445, which were mounted on the right frame, all remain above the consensus value. The water-viewing measurement C17 has better spatial heterogeneity and is more representative due to the more suitable zenith angle normally used for water reflectance measurements because the angular variability of the Fresnel reflection coefficient for 41° angle of incidence (cast C17) is smaller than for 50° (cast C23), and hence gives less spatial variability of skylight reflection.

In Figure 55, the SeaPRISM shows good agreement with the consensus value of LCE-2, while SR-3500 is through all casts biased towards somewhat smaller values. WISP-3 sensors show above average scattering of results, partly because their alignment to the frame in line with the other radiometers was difficult due to the ergonomic shape of these handheld instruments without straight surfaces suitable for exact alignment. It is not possible to conclude which sensor(s) showed the best agreement with SI due to the lack of a well-characterized SI-traceable reference radiometer involved simultaneously in the comparison.

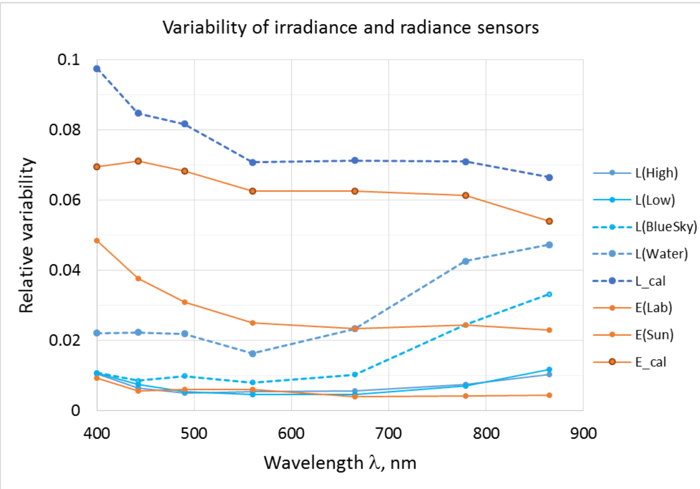


Figure 56. Variability between irradiance and radiance sensors. E_cal and L_cal – due to calibration state; E(Lab), L(Low) and L(High) – variability in laboratory intercomparison; E(Sun), L(BlueSky) and L(Water) variability in field.

The spread of irradiance and radiance results in the LCE-2 in comparison with differences between sensors due to their calibration state before the experiment is summarized in Figure 56. All standard deviations of laboratory measurements are smaller than 1 %. Standard deviations of the field results are substantially higher (1 % - 5 %), but still much smaller than the variability due to the calibration state of sensors before the experiment (5 % - 10 %) i.e. the calibration that each participant would have used if the radiometers were not freshly calibrated just before the start of the LCE-2 intercomparison exercise. It must be noted, however, that some instruments had not been used in recent years, thus, their previous calibration coefficients were several years old.

15.4.3.9 Measurements after the end of LCE-2 comparison

Large variability between irradiance sensors of the RAMSES group during the outdoor exercise cannot be fully explained by the poor stability of sensors, or by factors such as temperature dependence (which is rather similar for the whole RAMSES group [180]), non-linearity (which would be stronger for wavelengths with high digital counts), and stray light (which would show more spectral features). Most likely, the main reason for differences between RAMSES and HyperOCR irradiance sensors comes from the different properties of the entrance optics (angular response). The results of [191] for six RAMSES irradiance sensors suggest a cosine error within ±2 % for sun zenith angles lower than 50° when radiometric calibration is conducted using a 20° tilted sensor with respect to the incident irradiance. For the "conventional" calibration procedure at normal illumination a somewhat larger cosine error may be expected. Therefore, after the end of LCE-2, in January 2019 the in-air cosine response error of five RAMSES irradiance sensors was measured, see Figure 57. One new sensor number 8598, that was measured had not been involved in LCE-2.

Dependence of the cosine error on the zenith angle varies from radiometer to radiometer significantly with values ranging from –16 % up to +9 % at ±65°. Deviation from the ideal cosine response is irregular and does not always show monotonic increase with the incident angle. This is in agreement with the results of [191]. For one sensor, 8329, significant asymmetry is evident. The best of the characterized sensors, 81A8, demonstrated in the outdoor experiment irradiance results very close to the consensus value (Figure 54), whereas the sensor 81EA with the largest cosine error, at the same time, had a deviation from the consensus value of about -10 % to -15 %, depending on wavelength.

The manufacturer's specification of the HyperOCR [205] states that the cosine root mean square (RMS) error is within 3 % at 0° … 60°, and within 10 % at 60° … 85° incidence angles. For RAMSES [206], accuracy is stated to be better than 6 % … 10 % depending on the spectral range. Respective specification in [207] are: for E_d measurement, the response to a collimated source should vary as cosθ within less than 2 % for angles 0°< θ< 65° and 10 % for angles 65° < θ < 90°. For easier comparison of different sensors the deviation from an ideal cosine response was quantified as the integral of azimuth-independent absolute values of the cosine error for θ in the 0° to 85° interval, see Figure 58.

Increased variability between the RAMSES sensors in comparison with HyperOCR sensors in Figure 54 can be reasonably explained by a too tolerant specification of the cosine error, as departures from cos θ will translate directly to approximately equal errors in E_d in the case of direct sunlight [207]. Although the majority of the RAMSES sensors meet the present specification, differences revealed during the field measurements may render the specification unsatisfactory for users, unless laboratory characterisation data and an indication of the angular variation of the downwelling radiance field, e.g. direct/diffuse ratio, is available to correct for the imperfect cosine response.

Following the 20° "offsetting" calibration method suggested in [191], by using the cosine response characterisation results, the comparison data of Figure 54 were recalculated for two sensors. The effect of calibration with the sensor tilted to 20° with respect to the incident irradiance is shown in Figure 59. Improvement is evident for both sensors, but for 81EA the residual error is still large.

Thus, rather large cosine errors of RAMSES irradiance sensors can be considered to be the main reason for the differences between irradiance sensors during the LCE-2 outdoor measurements.

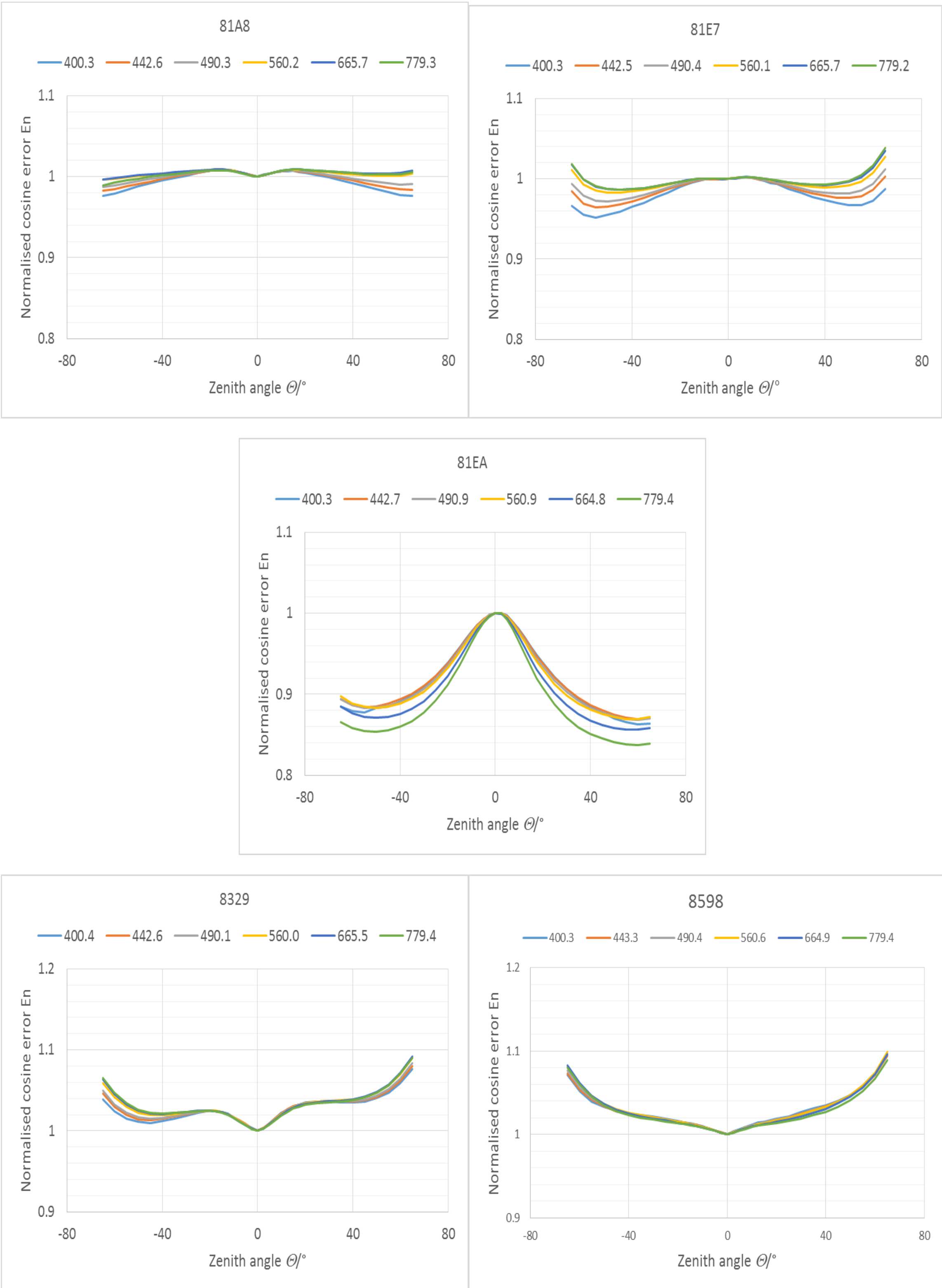


Figure 57. Normalized cosine response error of five RAMSES sensors.

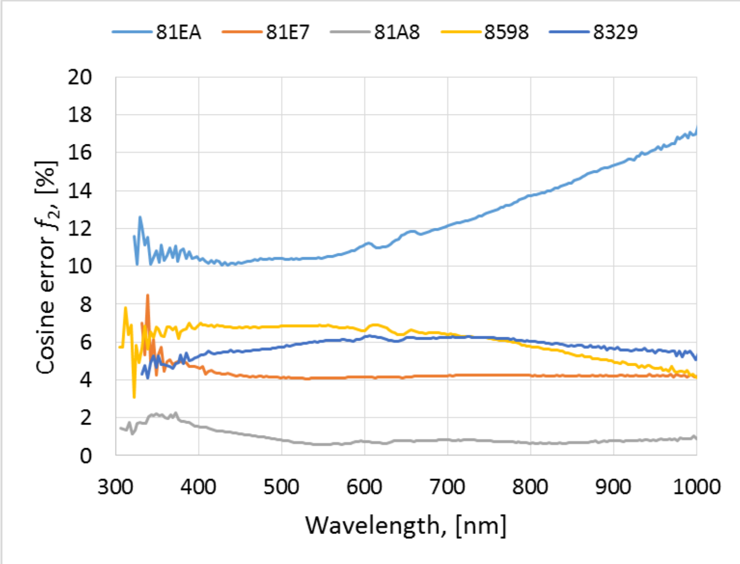


Figure 58. Integrated cosine error of the five RAMSES radiometers.

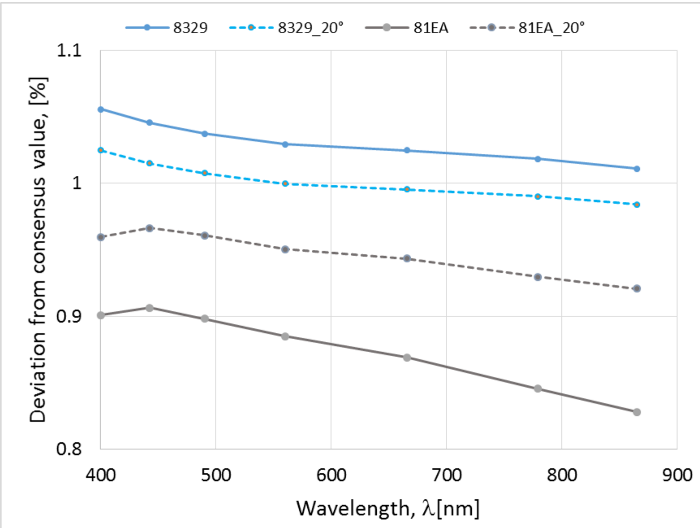


Figure 59. Effect of calibration with the sensor tilted to 20° with respect to the incident irradiance.

15.4.3.10 Uncertainty budgets of outdoor comparisons

An uncertainty analysis according to [18,72] was carried out for the outdoor measurements with the aim to see which contributions explain the observed variability between sensors. The outdoor downward irradiance uncertainty estimates are presented in Table 15; Table 16 corresponds to the blue sky radiance, and Table 17 to the radiance of sunlit water. All the uncertainty estimations in Table 15 - Table 17 are based on experimental variability data of TriOS RAMSES sensors and information from [1,154,185,197–199,208]. For the other radiometer models that took part in the intercomparison, very little information about instrument characteristics that influence the measurement results is publicly available [54]. In addition, only the RAMSES sensor was sufficiently numerous for a robust statistical analysis.

In general, the uncertainty is calculated from the contributions originating from 1) the spectral responsivity of the radiometer, including data from the calibration certificate; 2) interpolation of the spectral responsivity values to the designated wavelengths and/or spectral bands; 3) temporal instability of the radiometer; 4) contribution caused by polarization sensitivity; 5) non-linearity effects; 6) effect of spectral stray light; 7) temperature effects; 8) error of the cosine collector; 9) type A component of the recorded signal; 10) alignment and FOV effects.

The calibration uncertainty is most relevant for traceability to the SI units. The rest of the uncertainty sources in Table 15 - Table 17 describe variability between the sensors while overlooking possible systematic effects that can influence all the instruments in similar way. Moreover, there was no fully characterized reference instrument involved during the LCE-2 outdoor exercise. Thus, the uncertainty analysis presented here is not sufficient to link the measurements to the SI units.

Table 15. Relative uncertainty budget for the downward irradiance (in percent), based on the spread of individual sensors measuring the same target during the outdoor comparison. Data highlighted in green are not used for combined and expanded uncertainties. Last row: relative experimental variability of sensors evaluated from the results of the field comparisons.

	400 nm	442.5 nm	490 nm	560 nm	665 nm	778.8 nm	865 nm
Certificate	0.88	0.68	0.65	0.62	0.59	0.62	0.56
Interpolation	0.5	0.3	0.3	0.3	0.3	0.3	0.3
Instability (sensor)	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
Non-linearity	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	4.8	3.7	3	2.4	2.2	2.2	2
Signal, type A	0.01	0.01	0.01	0.01	0.01	0.02	0.02
Combined (k=1)	4.9	3.8	3.1	2.5	2.3	2.4	2.3
Expanded (k=2)	9.8	7.6	6.2	5	4.6	4.8	4.6
Variability (k=2)	9.7	7.6	6.2	5	4.7	4.9	4.6

Table 16. Relative uncertainty budget for the radiance of blue sky (in percent), based on the spread of individual sensors pointing at the same target during the outdoor comparison. Data highlighted in green are not used for combined and expanded uncertainties. Last row: relative experimental variability of sensors evaluated from the results of the field comparisons.

	400 nm	442.5 nm	490 nm	560 nm	665 nm	778.8 nm	865 nm
Certificate	1.2	0.78	0.76	0.73	0.71	0.73	1.35
Interpolation	0.5	0.3	0.3	0.3	0.3	0.3	0.3
Instability (sensor)	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
Non-linearity	0.4	0.4	0.5	0.5	0.5	0.6	0.6
Stray light	0.8	0.6	0.2	0.2	0.5	0.9	1
Temperature	0.4	0.2	0.2	0.2	0.2	0.4	0.8
Alignment, FOV	0.3	0.4	0.6	0.6	0.5	2	2.9
Signal, type A	0.01	0.01	0.01	0.01	0.02	0.11	0.2
Combined (k=1)	1.1	0.9	0.9	0.9	1	2.4	3.3
Expanded (k=2)	2.2	1.8	1.8	1.8	2	4.8	6.6
Variability (k=2)	2.2	1.8	2	1.6	2	4.8	6.6

Table 17. Relative uncertainty budget for the radiance of sunlit water (in percent), based on the spread of individual sensors pointing at the same target during the outdoor comparison. Data highlighted in green are not used for combined and expanded uncertainties. Last row: relative experimental variability of sensors evaluated from the results of the field comparisons.

	400 nm	442.5 nm	490 nm	560 nm	665 nm	778.8 nm	865 nm
Certificate	1.2	0.78	0.76	0.73	0.71	0.73	1.35
Interpolation	0.6	0.3	0.3	0.3	0.3	0.30.2	0.3
Instability (sensor)	0.04	0.03	0.02	0.01	0.01	0.02	0.01
Polarisation	0.2	0.2	0.2	0.2	0.2	0.2	0.2
Non-linearity	0.7	0.8	0.9	1	1.1	1.2	1.3
Stray light	0.9	0.7	0.3	0.3	0.7	0.9	1
Temperature	0.4	0.2	0.2	0.2	0.2	0.4	0.8
Alignment, FOV	1.7	1.8	1.8	1.6	1.8	4	4.3
Signal, type A	0.04	0.07	0.11	0.11	0.21	0.55	0.72
Combined (k=1)	2.2	2.1	2.1	1.9	2.3	4.2	4.6
Expanded (k=2)	4.4	4.2	4.2	3.8	4.6	8.4	9.2
Variability (k=2)	4.4	4.4	4.4	3.2	4.6	8.6	9.4

For the RAMSES group, the variability of radiance sensors during indoor and outdoor exercises (Figure 40, except C8 and C23) was close. Therefore, variability due to the significant influencing factor - temperature, and the respective estimate used in the uncertainty budget, can be considered practically the same as a large systematic change is likely similar for all sensors [180]. For example, during outdoor measurements, the temperature was rather stable varying from 5 °C to 9 °C, a range fairly comparable with the variation of temperature during the indoor exercise which was from 21 °C to 24 °C. As the construction of radiance and irradiance sensors (except the input optics) is similar, the similar estimate is likely suitable also for the temperature caused variability between irradiance sensors. Some increase in variability may be expected due to non-linearity and spectral stray light in the outdoor results. Major differences in combined uncertainty estimates for the outdoor measurements are likely caused by different FOV of the sensors (including deviation from cosine response for irradiance instruments) and due to temporal variation and non-uniformity of the targets.

15.4.3.10.1 Calibration certificate

The calibration certificates of the radiometers provide calibration points following the individual wavelength scale of the radiometer. During the relatively short time needed for LCE-2 measurements, this uncertainty component normally is not contributing to the variability between radiometric sensors freshly calibrated at the same laboratory using the same calibration standards. Therefore, this component is presented only for reference and is not included in the combined and expanded uncertainties. At the same time, for the full uncertainty of SI traceable results, the radiometric calibration uncertainty should always be accounted for and included in the uncertainty budget.

15.4.3.10.2 Interpolation

Interpolation of radiometer's data is needed due to the differences between the individual wavelength scales of the radiometers. Therefore, measured values were transferred to a common scale basis (a selection of Sentinel-3/OLCI bands) for comparison. 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 interpolation component therefore includes interpolation as well as wavelength scale uncertainty contributions. Figure 45 shows the change of the OLCI band values of a measured spectrum as a function of the wavelength scale error of a radiometer, as determined for a single RAMSES radiance sensor for casts C8, C12, C17, and C23. Precision of the wavelength scale of the RAMSES instrument is stated by the manufacturer as 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.

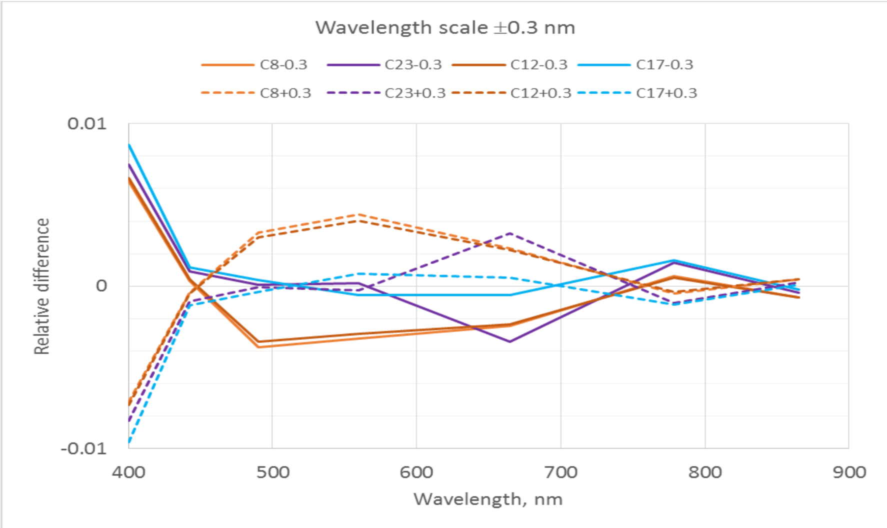


Figure 60. Relative variability due to wavelength error of ±0.3 nm of a radiance sensor.

15.4.3.10.3 Temporal instability of sensors

Instability of the radiometric responsivity can be estimated from data of repeated radiometric calibrations. For LCE-2, the instruments were calibrated just before the comparisons and only short-term instability relevant for the time needed for the comparison measurements, has to be considered. The values were derived from the data collected in the calibration sessions of LCE-2 and FICE-AAOT a year later, see Figure 38. The instability over two weeks was interpolated from the yearly variability data assuming only a smooth drift (no mechanical shocks, no abrupt changes). Besides the instability of the sensors, data in Figure 38 include other uncertainty components related to the calibration setup (e.g. alignment, short-term lamp instability, etc.).

15.4.3.10.4 Polarization

For the outdoor radiance measurements, the uncertainty contribution caused by polarization sensitivity is estimated using worst-case data in [185]. Evaluation of the polarization effect for the outdoor irradiance measurements is difficult as the degree of linear polarization (DoLP) depends on various factors such as wavelength, solar zenith angle (SZA), aerosol optical depth (AOD), amount and location of clouds, etc. In addition, the DoLP can strongly vary over the hemisphere, being due to Rayleigh scattering, with the largest at 90° from the sun, and for the direct solar flux decreasing to zero. However, according to [185] the polarization sensitivity of RAMSES irradiance sensors is rather small, hence, regardless of the DoLP value of downward irradiance the contribution of the polarization effect to the uncertainty budget is also small. The uncertainty component of solar irradiance associated with polarization is estimated to be less than 0.25 %.

15.4.3.10.5 Non-linearity

Non-linearity of the participating radiometers was evaluated by varying the integration time during the calibration. As an automatically adjusted optimal integration time is typically used in the field conditions, a class-specific method for the RAMSES instruments was developed and validated using the indoor results, see equations (17), (18) and (19). Variability between sensors due to non-linearity was evaluated by applying this equation to the different casts of the field spectra of five RAMSES sensors.

15.4.3.10.6 Spectral stray light

Figure 61 presents the impact of the stray light on the outdoor measurements. The effect is much stronger than in the indoor experiment due to the significantly different spectral shape of the target and calibration signals. General impact of the stray light correction is similar for RAMSES and HyperOCR radiometers. Variability between sensors and between different measurements targets for HyperOCR radiometers 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 of the actual impact of spectral stray light. The spectral stray light matrices of HyperOCR sensors used in the analysis had a somewhat higher noise level compared to the matrices of the RAMSES instruments. Data in Table 16 - Table 17 is estimated from Figure 61, and from [190,196].

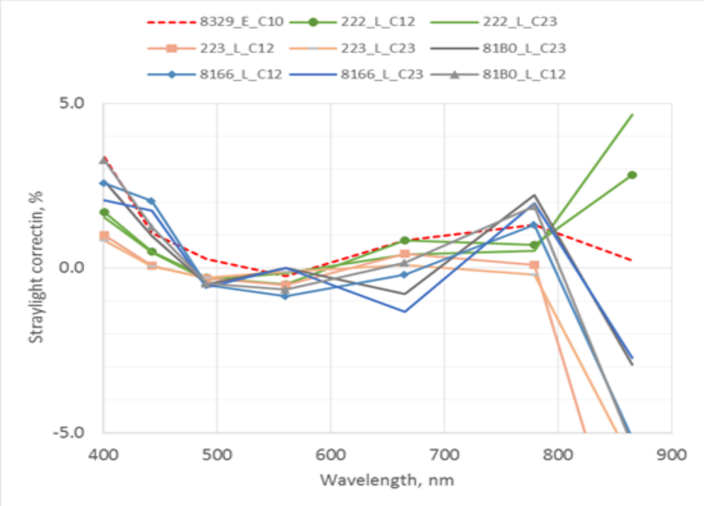


Figure 61. Stray light effects in the outdoor experiment. One RAMSES 8329 irradiance sensor - dashed line; two RAMSES and two HyperOCR radiance sensors: solid lines with markers – blue sky (C12), and solid lines without markers – sunlit water (C23).

15.4.3.10.7 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) [180]. In the case of outdoor measurements, the temperature differences between sensors were likely in the range of (±2 °C), so temperature contribution is slightly larger than for the indoor experiment. Nevertheless, outside air temperature between 5 °C and 9 °C was significantly lower than the calibration temperature contributing to systematic biases common to all the instruments and thus not accounted for in Table 15 - Table 17.

15.4.3.10.8 Cosine error

The irradiance sensors were calibrated using normal illumination, but during the outdoor solar irradiance measurements the radiation arriving from the hemisphere has to be measured with the angular dependence of responsivity ideally corresponding to the cosine of the incidence angle. Typical class-specific values of uncertainty related to deviation of the cosine response are derived from [191]. Measurements carried out after LCE-2 at TO (Figure 57 - Figure 58) have shown that RAMSES sensors can have rather large cosine errors around ±10 %. This may be a likely reason for excessive differences from +7 % up to -16 % evident for irradiance sensors during the LCE-2 outdoor measurements (Figure 54).

15.4.3.10.9 Type A uncertainty of repeated measurements

The type A uncertainty was estimated from the ratio of two RAMSES radiometers. While there is strong autocorrelation in individual time series due to the unstable nature of natural illumination, there was almost no correlation between individual ratios during one cast, and the effective number of measurements was close to the actual number of data points in the time series. The effective number of measurements was calculated by using the lag-1 autocorrelation coefficient as shown in [202]. The instability of the target signal during the outdoor measurements was significantly larger compared to the indoor experiment, however, all the instruments measured simultaneously and the impact of source temporal variability affected all the radiometers in a similar manner without causing differences between the sensors. This was verified by separately analysing some shorter and more stable sections of the selected casts, where no reduction of variability between the sensors was observed. Thus, the uncertainty due to temporal variability of the target is not included in Table 15 - Table 17.

15.4.3.10.10 Alignment and field-of-view

During the outdoor radiance measurements, the spatial non-uniformity of the target can substantially contribute to the uncertainty due to the different FOVs of the radiance sensors, and due to misalignment (Figure 50 - Figure 51).

15.4.4 Conclusions

This study aims to evaluate the effectiveness of SI-traceable radiometric calibration for consistency of OC field measurements, indicates some gaps presently revealed during LCE-2, and discusses techniques and procedures, which could be helpful for improvement of the traceability of OC field measurements.

For irradiance, the difference in cosine response is the main source of differences between sensor groups revealed during field the experiment. For radiance, the angular response and spatial non-uniformity of the targets is the main source of differences between sensor groups. In the case of a spatially heterogeneous target (sky with scattered clouds, water at an oblique viewing angle), the large differences of FOV of the different sensors will likely cause significant discrepancies between sensors. Without reliable data or individual testing of the input properties of all involved sensors, interpretation of measurement results may be strongly hindered. For field measurements, the variability between radiance sensors was about two times larger than during the indoor exercise. This can be explained, by among other things, larger effects due to outside influence factors like temperature, stray light and non-linearity, which all have not been corrected for during the field experiment.

Dependence of the calibration coefficients on temperature can cause a significant deviation from a low uncertainty SI-traceable result. For a maximum temperature difference of about 20 °C between calibration and later measurements (typically between 0 °C and 40 °C) a responsivity change of more than 10 % is possible [179,180]. The calibration procedure may be improved if its specified conditions cover all situations possible during the use of a calibrated instrument. For example, if it is known that the radiometer has a linear response with temperature [180], the responsivity of the radiometer can be adequately evaluated when the calibration is performed at three different temperatures covering the possible range of temperature variations during its later use.

Variability between irradiance sensors was about five times larger than during the indoor exercise. Rather large variability between sensors during the outdoor exercise cannot be explained by poor stability of sensors, as a stability check in laboratory conditions a year later has shown much smaller changes than during the outdoor measurements some days after calibration. Variability cannot be fully explained by factors such as temperature, non-linearity, and stray light either as one could expect smaller differences between radiance and irradiance sensors in this case. Most likely, the different behaviour of RAMSES and HyperOCR sensors is largely due to different construction of input optics of these sensors and hence imperfect cosine response. This hypothesis is supported by the angular response characterisation of five RAMSES irradiance sensors and comparing the integral cosine error values in Figure 57 - Figure 58 to the relative deviations from the consensus value in the outdoor experiment shown in Figure 54.

The different behaviour of RAMSES and HyperOCR sensor groups was clearly revealed during the 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 an increase in variability was caused mostly by the group of RAMSES sensors.

The indoor experiment has demonstrated a great effectiveness of radiometric calibration at the same laboratory just before intercomparison measurements [204] for obtaining consistent results. Nevertheless, a sufficient individual characterisation of radiometers by testing them for all significant systematic effects, besides regular radiometric calibration, is the shortest way to enable reduction of biases in outdoor intercomparisons, and thus smaller variability between measurements from different instruments, and more realistic and complete quantification of uncertainties in measurement.

15.4.5 Lessons learned for the design of future intercomparisons.

In order to help in interpretation of the results, the following suggestions are proposed for future outdoor intercomparison campaigns.

1. The number of involved radiometers should be around ten for each radiometer type in order to have a sufficiently representative group for robust statistical analysis.
2. Consistent calibration of the responsivity of all involved radiometers just before the campaign is indispensable.
3. Calibration history of each radiometer should be available to detect long-term instabilities.
4. Together with radiometric calibration, the angular response of all individual radiance and irradiance sensors should be measured if such information is not available from previous characterisations.
5. Before radiometric calibration, all instruments involved should be tested or be characterized for temperature, non-linearity, spectral stray light and wavelength scale effects. As these tests may be rather time consuming they should be performed well before the radiometric calibration.
6. Spectral responsivity should be calibrated at different ambient temperatures relevant to the campaign. Non-linearity and wavelength correction coefficients should also be available.
7. The usefulness of individual characterisation of the spectral stray light should be further proven by thorough field tests using an independent validation method based on a reference instrument less affected by stray light.

During outdoor campaign measurements, well-synchronized data acquisition for all instruments is strongly advised. A starting timer should be aligned better than within ±1 s;

setting exactly the same sampling interval for all sensors is indispensable, allowing comparison of individual spectral instead of cast averages. Data processing algorithms should be well defined and agreed between the participants. For that, sufficient calibration and test information should be available for each sensor in order to be able to apply all needed corrections similarly. Instruments' temperature should be recorded whenever possible. Use of a well-characterized additional reference instrument is highly recommended, as is using an aligned photo- or video camera to record the measurement scene during outdoor experiments simultaneously with radiometric sensors.

Metrological specifications of all involved radiometers whenever possible should be based on suitable international standards. Minimum requirements should be agreed between the participants, instruments involved should be tested to give evidence that all these minimum requirements are met.

15.5 Data package LCE-2 DATA [D-160]

All data collected during the laboratory comparison experiment LCE-2 (including raw, traceability, auxiliary and processed data) has been compiled into a data package file FRM4SOC-D-160-LCE-2-DATA.zip and handed over to ESA.

16 OC FRM Field Inter-comparison Experiments (FICE)

16.1 Introduction

In order to meet Objective 5 of the FRM4SOC project – design, document protocols and procedures and implement field inter-comparisons of FRM OCR radiometers and build a database of OCR field radiometer performance knowledge over several years – the following documents were prepared:

- "Protocols and Procedures for Field Inter-Comparisons of Fiducial Reference Measurement (FRM) Field Ocean Colour Radiometers (OCR) used for Satellite Validation" [D-190] , [209];
- "FICE Implementation Plan (FICE-IP)" [D-200], [210].

Following the guidelines as provided by these two documents OC FRM Field Inter-Comparison Experiments FICE-AMT and FICE-AAOT were organised.

The results of the of the FICE Inter-Comparison Experiments are presented in the report

- "Results from the First FRM4SOC Field Inter-Comparison Experiment (FICE) of Ocean Colour Radiometers" [D-220], [211]

and also in the peer-reviewed papers [109,212] published in the FRM4SOC special issue of the MDPI journal Remote Sensing. **For citation of the chapter 16**, the papers [109,212] should be considered as preferable references.

All data collected during the FICE experiments has been collected into the

- Field inter-comparison experiment database (FICE-DB). [D-210] [213].

16.2 Protocols and Procedures for Field Inter-Comparisons of Fiducial Reference Measurement (FRM) Field Ocean Colour Radiometers (OCR) used for Satellite Validation" [D-190]

The document [D-190], [209] addresses the requirements of the FRM4SOC SOW [3] to

- Be the master guide for the FRM4SOC team to implement side-by-side comparisons of OCR field radiometers. The intention is that [D-190] could be used for future campaigns.
- Critically review the methodology used to measure ocean colour parameters in the field using FRM OCR.
- Establish, by consensus, and document community best practises for OCR field deployments.
- Define procedures and protocols to maintain pre-deployment and post-deployment calibration verification of FRM OCR that are traceable to SI standards.
- Define good practice approaches and protocols to validate uncertainties for FRM OCR measurements made in field under a range of operational conditions and biogeochemical conditions (i.e. end-to-end).

The principles described in [D-190] are presented in detail in Sections 16.4 and 16.5.

16.3 "FICE Implementation Plan (FICE-IP)" [D-200].

In order to plan and manage organisational issues of the comparison event (e.g. overview of the activity, dates, times, locations, customs and shipping aspects, hotels and travel details, visa requirements, etc.), the "FICE Implementation Plan (FICE-IP)" [D-200], [210] was prepared and followed.

16.4 Results from the First FRM4SOC Field Inter-Comparison Experiment (FICE) of Ocean Colour Radiometers" [D-220] – The Atlantic Meridional Transect (FICE-AMT) cruise field intercomparison experiment

Plymouth Marine Laboratory (PML), in collaboration with the National Oceanography Centre (NOC) Southampton, has operated the AMT since 1995 [214]. The cruise is conducted between the UK and the sparsely sampled South Atlantic during the annual passage from October to November of a NERC ship (*RRS James Clark Ross, RRS James Cook* or *RRS Discovery*). The transect covers several ocean provinces where key physical and biogeochemical variables such as chlorophyll, primary production, nutrients, temperature, salinity and oxygen are measured. The stations sampled are principally in the North and South Atlantic Gyres, but also the productive waters of the Celtic Sea, Patagonian Shelf and Equatorial upwelling zone are visited, which therefore offered a wide range of variability in which to conduct field intercomparisons for FRM4SOC. The results from the AMT cruises have enabled the intercomparison of simultaneous measurements of water-leaving radiance and reflectance. The differences observed between these measurements form a key component of estimating errors and uncertainties resulting from environmental variability, as well as instrument deployment methodology, instrument specifications and calibration.

The main AMT comparison for FRM4SOC was conducted from 23rd September to 4 November 2017 from Southampton, UK to South Georgia and the Falkland Islands on AMT-27 to compare measurements of $L_w(\lambda)$, $E_d(\lambda)$ and normalised water leaving reflectance $[\rho_w(\lambda)]_N$ between radiometers of PML and TO [212]. Measurements were carried out in various solar zenith angle, water and weather conditions (e.g. -53.65 < latitude (°) < 48.93; -38.05 < longitude (°) < -7.62; 5.84 < sun zenith angle (°) < 60.54; 1.48 < wind speed (m·s–1) < 19.71). The ambient temperature varied from 1 °C to 28 °C. Altogether, data was collected from 32 stations, which enabled to compare the radiometric data slightly outside of the strict rules applied to produce validation datasets for satellites. This is important in order to show the reliability of the in situ measurements and study the behavior of existing radiometers close to (or even beyond) the specification limits in order to plan the next-generation systems. The measurement stations are listed in Table 18; Sentinel-3A OLCI images were available for stations with id-s No 22, 32, 46, 48, 56.

The AMT-27 cruise data consists of synchronized measurements of water-leaving reflectance with two sets of hyperspectral radiometers, both consisting of three radiometers in order to measure the upwelling radiance $L_u(\lambda)$, downwelling radiance from the sky $L_d(\lambda)$, and downward solar irradiance $E_d(\lambda)$. The PML set consisted of three Satlantic HyperSAS sensors and the TO set of three TriOS RAMSES sensors. Technical parameters [205,206] of the applied radiometers are given in Table 19.

All radiance and irradiance sensors were SI-traceably calibrated at the Tartu Observatory, University of Tartu before and after the campaign. All of these sensors were involved a year before in the LCE-2 (Section 14) and demonstrated differences less than ±1 % both for radiance and irradiance results during indoor measurements. However, during the outdoor exercise, the PML irradiance sensors showed up to 6 % higher values in the blue part of the spectrum, and the PML radiance sensors showed up to approximately 10 % higher values in the red and NIR parts of the spectrum when compared to the respective TO sensors. The radiance sensors $L_d(\lambda)$ and $L_u(\lambda)$ were mounted on a common steel frame positioned at the front of the ship side by side by using 40° and 120° angles from zenith, respectively. The downward irradiance sensors were mounted at higher level of the same mast in order to avoid any shadows. Positioning of sensors ensured nearly identical measurements conditions for both 3-sensor radiometric systems (see Figure 62).

Table 18. Overview of measurement conditions during the midday station at AMT-27

No	Station id	Date	Latitude	Longitude	Sun zenith angle	Wind speed (W, m·s⁻¹)	Temperature (t, °C)
1	1	24.09.2017	48.9°	-7.6°	52.37°	1.48	16.2
2	3	25.09.2017	46.7°	-12.0°	51.52°	7.23	17.3
3	6	27.09.2017	42.2°	-18.8°	46.31°	2.24	19.3
4	8	28.09.2017	39.4°	-22.7°	45.31°	5.94	23.0
5	10	30.09.2017	35.1°	-26.3°	38.87°	1.69	24.3
6	12	01.10.2017	31.8°	-27.2°	35.84°	5.69	23.5
7	16	03.10.2017	25.7°	-28.7°	30.52°	7.15	24.5
8	18	04.10.2017	22.3°	-29.5°	28.58°	1.69	25.7
9	20	05.10.2017	18.8°	-29.7°	26.4°	5.47	26.6
10	22	06.10.2017	15.5°	-28.8°	23.21°	4.31	27.8
11	24	07.10.2017	12.8°	-28.2°	20.4°	8.43	28.0
12	26	08.10.2017	9.9°	-27.4°	18.38°	6.89	28.3
13	28	09.10.2017	6.9°	-26.7°	15.41°	5.13	27.6
14	32	11.10.2017	1.5°	-25.4°	10.42°	6.34	26.0
15	34	12.10.2017	-1.8°	-25.0°	8.23°	8.44	25.9
16	36	13.10.2017	-4.6°	-25.0°	6.07°	10.74	25.7
17	38	14.10.2017	-7.1°	-25.0°	5.84°	6.8	25.5
18	40	15.10.2017	-10.5°	-25.1°	5.93°	6.63	25.1
19	42	16.10.2017	-13.7°	-25.1°	7.85°	8.12	23.8
20	43	17.10.2017	-16.0°	-25.1°	8.22°	8.07	22.9
21	46	19.10.2017	-21.8°	-25.1°	13°	8.76	21.7
22	48	20.10.2017	-25.1°	-25.0°	15.5°6	6.26	21.2
23	50	21.10.2017	-27.9°	-25.2°	17.7°8	3.88	20.7
24	52	22.10.2017	-31.3°	-26.2°	21.2°2	8.74	19.4
25	54	23.10.2017	-33.9°	-27.1°	26.0°5	6.17	17.3
26	56	24.10.2017	-37.0°	-28.3°	28.6°2	8.12	15.8
27	59	26.10.2017	-42.1°	-30.4°	34.2°2	7.92	10.0
28	61	27.10.2017	-45.2°	-31.7°	36.15°	16.25	8.8
29	62	28.10.2017	-47.1°	-32.6°	54.53°	8.03	6.4
30	64	29.10.2017	-50.4°	-34.2°	40.23°	11.63	1.6
31	66	30.10.2017	-52.9°	-35.7°	43.25°	9.25	0.9
32	67	01.11.2017	-53.7°	-38.1°	60.54°	19.71	2.0

Table 19. Technical parameters of radiometers used for comparison.

Parameter	RAMSES	HyperOCR
Field of View (L/E)	7°/cos	6°/cos
Adaptive integration time	yes	yes
Min. integration time, ms	4	4
Max. integration time, ms	4096	4096
Min. sampling interval, s	1	0.5
Recording dark signal	Masked pixels	Internal shutter
Number of channels	256	256
Wavelength range, nm	320…1050	320…1050
Wavelength step, nm	3.3	3.3
Spectral resolution, nm	10	10

The intercomparison allowed the analysis of the variability of responsivity between different types of freshly calibrated sensors, with respect to the environmental and illumination conditions. As an example, the difference in the results of downward irradiance between PML and TO, as a function of ambient temperature and solar zenith angle are shown in Figure **63**. With regard to ambient temperature, radiometric calibration of the sensors was performed in lab conditions at 21 °C and no temperature correction factors were applied for the field results. Responsivity change for both sensors is larger (and unknown) compared to the change of the signal ratio shown. The differences vary from approximately -5 % to +5 % in the temperature range from 1 °C to 30 °C. However, the sensors record similar irradiance values around 21 °C, which corresponds to the calibration temperature. This result clearly shows the need for characterisation of field radiometers for thermal effects.

Figure 62. The route of AMT-27 through the Atlantic and the position of the FRM4SOC radiometers on RRS Discovery in operation during AMT-27.

For solar zenith angle, the variation is in agreement with known or expected errors of the cosine collectors of compared sensors, evaluated to be within ±2 % [168]. The stray light correction effect is negligible and shown in Figure 63 for reference only.

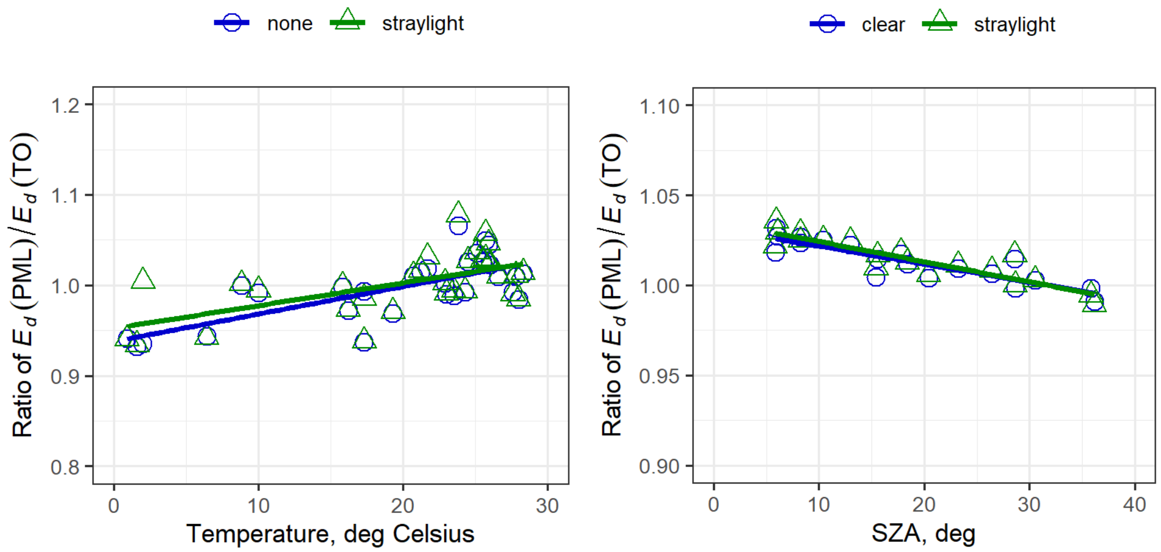


Figure 63. Difference in downward irradiance between PML and TO as a function of ambient temperature (left) and solar zenith angle (SZA, right).

The comparison of HyperSAS and RAMSES measured water-leaving reflectance after applying stray light correction showed a very high agreement over all wavelengths. The systematic biases were negligible (see Figure 64).

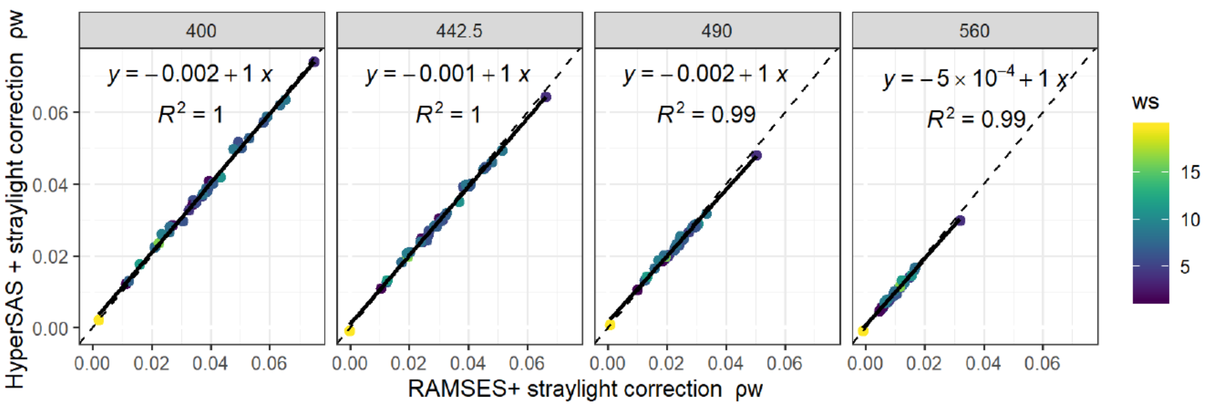


Figure 64. Correlation between HyperSAS and RAMSES measured water-leaving reflectance after stray light correction on selected wavelengths. Colour is wind speed (m s^{-1}) during the measurement.

The comparison between the derived from OLCI and in situ water-leaving reflectance, either by RAMSES (Figure 65 right) or HYPERSAS (Figure 65 left) showed a very good correlation in the blue to green wavelengths. For these wavelengths, the correlation with OLCI-derived water-leaving reflectance was even better after applying the NIR similarity correction [198,215] (Figure 65).

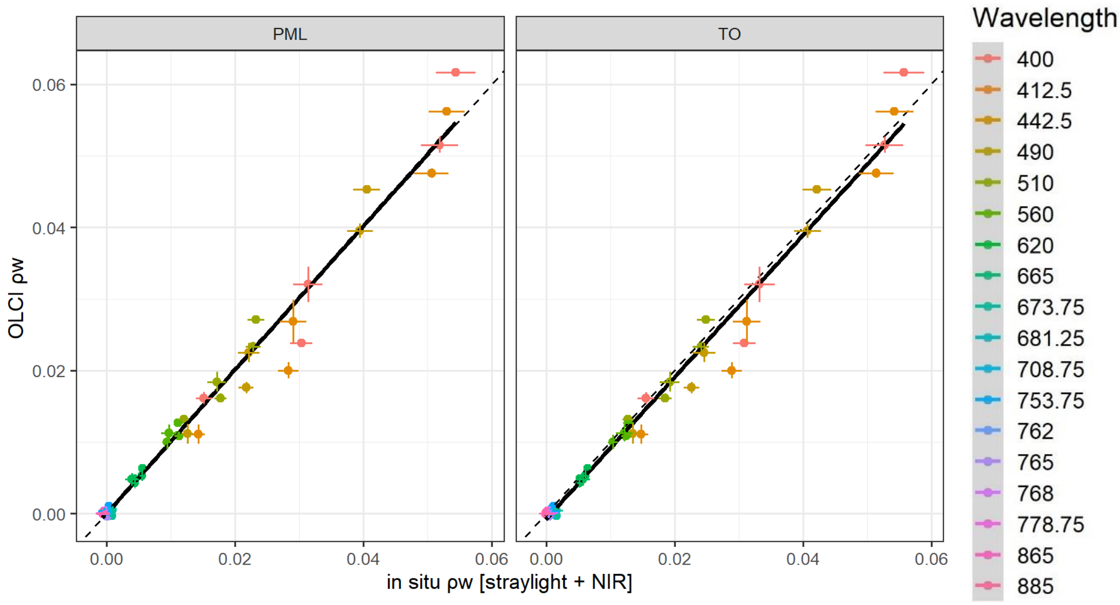


Figure 65. Correlation between OLCI to HyperSAS (left) and OLCI to RAMSES (right) water-leaving reflectances on selected wavelengths.

Comparison of the water-leaving reflectance after stray light correction (brown) and after stray light+NIR similarity correction (yellow) compared to OLCI water-leaving reflectance in two stations is shown in Figure 66.

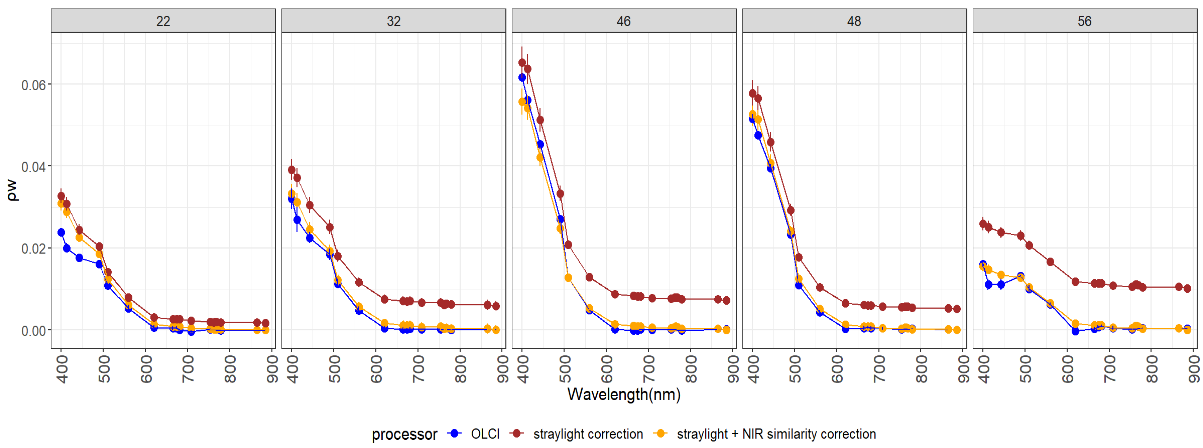


Figure 66. Comparison of RAMSES radiometer derived water-leaving reflectance after stray light correction (brown) and after stray light+NIR similarity correction (yellow) compared to OLCI water-leaving reflectance (blue) in five match-up stations.

The above summary analysis shows that by comparing results to ancillary instrument data, and during the cruise (with regards to environmental conditions), the sources of any differences can begin to be established. From these results, recommendations can be made to adjust processing methodology (e.g. applying appropriate filtering thresholds), future instrument deployment methodology, and calibration processes. Furthermore, these comparisons contribute to the Type B estimates in an uncertainty budget [17]. A wider comparison analysis including uncertainties is underway using data collected during the AMT-27 and these initial results are promising, especially given the large differences in environmental conditions experienced during the AMT cruise.

In general, the agreement between the two in situ systems during the whole comparison exercise was satisfactory, with up to a 5 % difference over visible wavelengths before corrections applied to ρ_w. In the range of (400…510) nm, the relative mean uncertainty of in situ data is close to the Sentinel-3 mission requirements of 5 %, but with an increase in wavelength beyond 500 nm, the relative uncertainty also increased mainly due to unstable targets, highly variable environmental conditions, and the low signal at red bands in oligotrophic waters. The consistency between the satellite data and in situ data with stray light and NIR similarity correction applied was also satisfactory. The consistency between the corrected in situ and OLCI radiometric data in the range of (400…510) nm was good and respective uncertainties less than 7 %. SI traceable calibration of radiometers before field campaigns with a reasonably small uncertainty is very important. Nevertheless, this may be insufficient and various additional individual tests of radiometers, like for temperature dependence, non-linearity, spectral stray light, etc., are also needed. At the same time, in order to correct different biases, and to improve uncertainties, measurement of environmental conditions during deployment is also highly relevant.

16.5 Results from the First FRM4SOC Field Inter-Comparison Experiment (FICE) of Ocean Colour Radiometers" [D-220] – The Acqua Alta Oceanographic Tower (FICE-AAOT) field intercomparison experiment.

The second FRM4SOC field intercomparison was conducted from 09 to 19 July 2018 at the Acqua Alta Oceanographic Tower (AAOT), which is located in the Gulf of Venice, Italy, in the northern Adriatic Sea at 45.31°N, 12.50°E. The AAOT is a purpose-built steel tower with a platform containing an instrument house to facilitate the measurement of ocean properties under exceptionably stable conditions (Figure 67). Nine international institutes participated in the intercomparison enabling the comparison of ten measurement systems comprising 29 radiometers, see Table 20.

Table 20. Field intercomparison measurement systems, sensors and institutes. All sensors are hyperspectral except C-OPS, which is multispectral.

	Method (identifier)	Radiometers	Institute
1	Above-water (RAMSES-A)	TriOS RAMSES	University of Algarve, Portugal
2	Above-water (RAMSES-B)	TriOS RAMSES	Tartu Observatory, University of Tartu, Estonia
3	Above-water (RAMSES-C)	TriOS RAMSES	Helmholtz-Zentrum Geesthacht, Germany
4	Above-water (RAMSES-D)	TriOS RAMSES	Alfred Wegener Institute, Germany
5	Above-water (RAMSES-E)	TriOS RAMSES	Royal Belgian Institute of Natural Sciences, Belgium
6	Above-water (HyperSAS-A)	Seabird	Plymouth Marine Laboratory, UK
7	Above-water (HyperSAS-B)	Seabird	University of Victoria, Canada
8	Above-water (PANTHYR)	TriOS RAMSES + pan and tilt	Flanders Marine Institute, Belgium
9	In-water C-OPS (in-water A)	Bio-spherical microradiometers	Institut de la Mer de Villefranche-sur-Mer, France
10	In-water (RAMSES-B)	TriOS RAMSES	Alfred Wegener Institute, Germany

The main aim of the AAOT intercomparison was to assess differences in radiometric quantities determined using a range of above-water and in-water radiometric systems as operated in the field (including both different instruments and processing protocols). Specifically, differences among the following were evaluated:

1. Hyperspectral (five above-water TriOS RAMSES, two Seabird-HyperSAS, one Pan-and-Tilt System with TriOS RAMSES sensors - PANTHYR, one in-water TriOS RAMSES system) and multispectral (one in-water Biospherical-C-OPS) sensors.
2. In-water and above-water measurement systems.

To rule out any differences arising from absolute radiometric calibration, all of the sensors used during the campaign were calibrated at the Tartu Observatory, University of Tartu, under the same conditions, within ~1 month of the campaign. Measurements were then performed at the AAOT under near ideal conditions, on the same deployment platform and frame, under clear sky conditions, relatively low sun zenith angles and moderately low sea state.

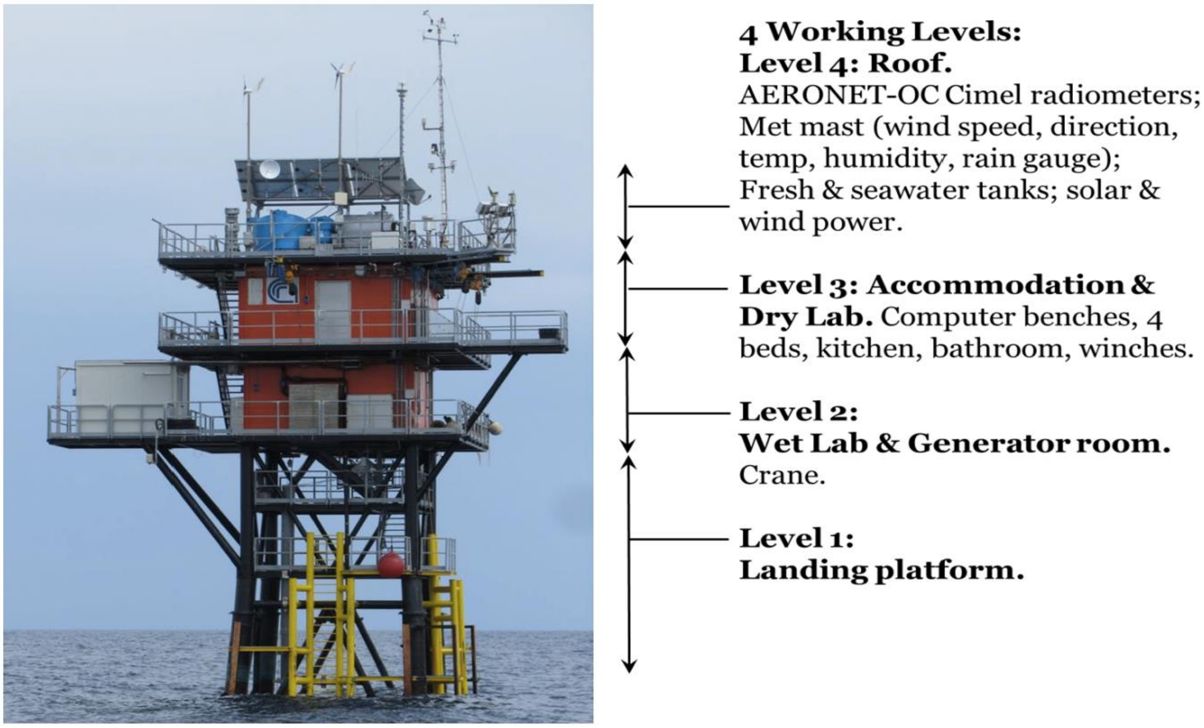


Figure 67. Layout of the Acqua Alta Oceanographic Tower (AAOT).

Configuration of the radiance and irradiance sensors is shown in Figure 68. All above-water radiometers except the PANTHYR system were located on the same purpose-built frames. The radiance sensors were located on the deployment platform on level 3 on a 6 m pole that situated them above the solar panels on level 4 (Figure 67). The frame was fabricated from aluminium to position the sensors side by side at 12 m from the sea surface (Figure 68 a). All L_{sky} [11] and L_t [12] sensors were installed on this frame with identical viewing zenith angles. The deployment frame was adjusted for each measurement sequence to reduce sunglint. The

[11] L_{sky} in previous also noted as L_d. See section 12.2.
[12] L_t in previous also noted as L_u. See section 12.2.

radiance mast was positioned at the same level as the SeaPRISM AERONET-OC system (Figure 68 b and c).

For irradiance measurements, a telescopic mast was used on level 4 to minimize interference from the tower super-structure and other overhead equipment (Figure 68 e and f). The mast and sensors were installed in the eastern corner of the platform at a height of 18.9 m above the sea surface (Figure 68 e).

All above-water measurements were conducted every 20 min from 08:00 to 13:00 GMT over a discrete measurement period of 5 min (known as casts). In-water C-OPS were also coordinated to these times and in water TriOS (AWI) were made directly after the above-water casts. Only casts with wind speeds < 5 m s^{-1} and clear skies (no cloud) were accepted. Using these criteria, 35 valid casts resulted from the campaign. Each institute used their standard processing to compute downward irradiance (E_d), sky radiance (L_{sky}), radiance from the water surface (L_t) and remote sensing reflectance (R_{rs}). Mean, median and standard deviation values of these parameters over each 5-minute cast were submitted. These were compared to the weighted mean of above-water systems that were submitted by the 'blind' submission date, and subsequently used as a reference. For the computation of the weighted mean, the mean of 3 x TriOS-RAMSES (RAMSES-A, -B & -C) systems was calculated, and the mean of 2 x Seabird-HyperSAS (HyperSAS-A, HyperSAS-B) systems was calculated. The data from these systems were used since final versions of these data sets were available on the 'blind' data submission date. From these, the mean (referred to hereafter as the weighted mean) of the TriOS-RAMSES and Seabird-HyperSAS was computed. In-water systems were excluded from the computation of reference values to allow a direct comparison with above-water systems and because of the lower number of comparable radiometric products.

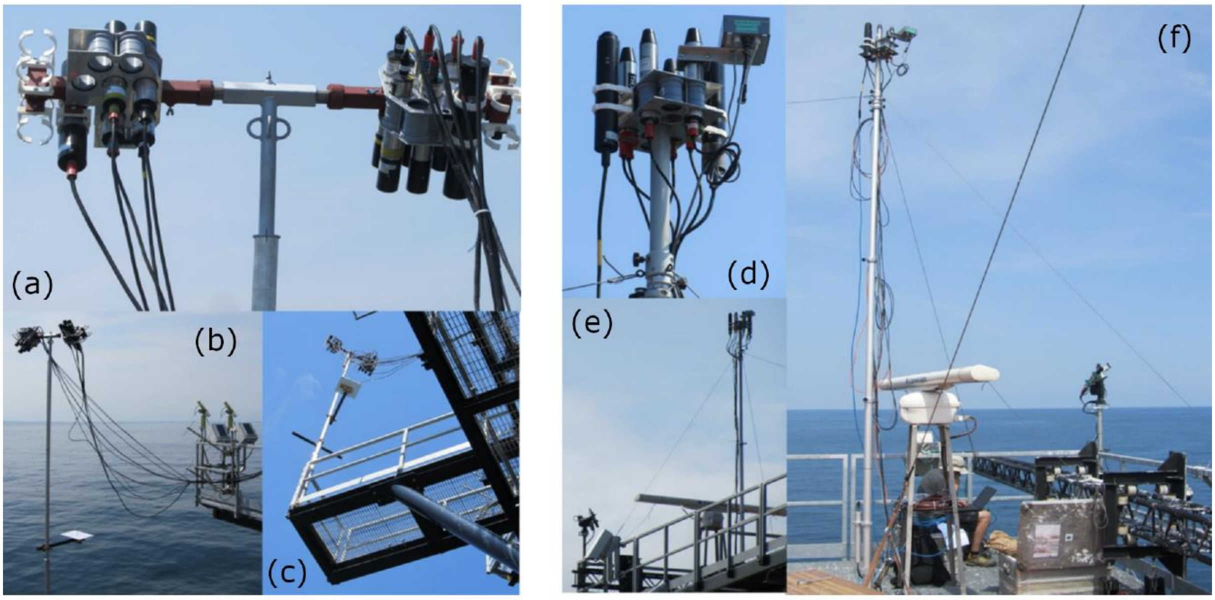


Figure 68. Configuration of the radiance (left) and irradiance sensors (right) showing (a) the mounting for L_{sky} and L_t radiometers, (b) location of the radiance sensors next to the AERONET-OC SeaPRISM, (c) location of the radiance sensors on level 3 of the AAOT, (d) location of the irradiance sensors on the mounting block, (e) telescopic mast with irradiance sensors at the eastern corner of the AAOT, (f) proximity of the telescopic mast with irradiance sensors and the PANTHYR system just above the railings below.

The in-water deployment of the AWI TriOS profiler was carried out using an extendable boom from level 4 of the tower, whereas the C-OPS-LOV in water system was deployed from the CNR Research Vessel Litus. Underwater optical light fields were measured with hyperspectral TriOS RAMSES radiometers, (Figure 69 D.) to obtain profiles of radiance, L_t, and irradiance, E_d, following the methods outlined in [216,217]. All measurements were collected with sensor-specific automatically adjusted integration times (between 4 ms and 8 s). The radiance and irradiance sensors were deployed from an extendable boom to 12 m off the southwestern corner of the AAOT (Figure 69 C). The height of the boom was 12 m above sea surface, and is designed to reduce shadow and scatter from the tower. The E_d sensor was equipped with an inclination and a pressure sensor. For this study, we only used the depth and inclination information from this sensor. During the intercomparison, the in-water inclination in either dimension was <6°.

Table 21. Differences between laboratories in the processing of data from E_d, L_t, L_{sky} to R_{rs}. Year (E_d, L_{sky}, L_t) is the year of manufacture of sensors; N are the number of replicates used for processing each cast; QC flag are quality control flags used; FOV is the radiance field of view; ρ is the Fresnel reflectance factor used to process the data.

Sensor type	Year (E_d, L_{sky}, L_t)	N E_d	N L_{sky}	N L_t	QC flag	FOV	ρ
RAMSES-A	2015, 2015, 2015	3-30	3-30	3-30	Visual QC	7°	[218]
RAMSES-B	2004, 2006, 2010	3-30	3-30	3-30	Visual QC	7°	[218]
RAMSES-C	2006, 2006, 2006	117-140	116-140	102-140	5 min scans	7°	[218–220]
RAMSES-D	2007, 2006, 2011	123-141	4-90	4-54	L_t< 1.5%; L_{sky}< 0.5% of min.	7°	[218]
RAMSES-E	2008, 2001, 2001	1st 5 QC	1st 5 QC	1st 5 QC	1st 5 scans	7°	[198]
HyperSAS-A	2006, 2006, 2006	280-345	284-398	93-198	5 min scans	6°	[218]
HyperSAS-B	2004, 2004, 2004	~130	~86	~86	lower 20%	6°	[218] [198]
PANTHYR	2016, 2016	2*3	2*3	11		7°	[198] [215]
In-water A	2010, N/A, 2010	3-4	N/A	‡3-4	Visual QC	N/A	[218]
In-water B	2007, 2010	150-200	N/A		Z extrapolation	7°	[221]

For all casts, the instruments were first lowered to just below the surface, at approximately 0.5 m, for 2 min to adapt them to the ambient water temperature. The frame was then lowered to approximately 14 m, with stops every 1 m for a period of 30 s each, to obtain representative average values at each depth.

Differences between the used radiometers (year of manufacture, FOV) and between methods and procedures used by different laboratories in the processing of data from E_d, L_t, L_{sky} to the final result R_{rs} are collected in Table 21. In columns 3 to 5, the number of replicates used for processing of each cast is shown. In the sixth column, quality control flags used by participants are listed. In the last column the references to the methods describing the use of the Fresnel reflectance factor in data processing are given.

The variability in $E_d(442)$, $L_{sky}(442)$, and $L_t(442)$ for the days and casts used in the intercomparison are shown in Figure 70. The coefficient of variation of $R_{rs}(442)$ represents temporal changes in both in water constituents and in the bi-directionality of the light field.

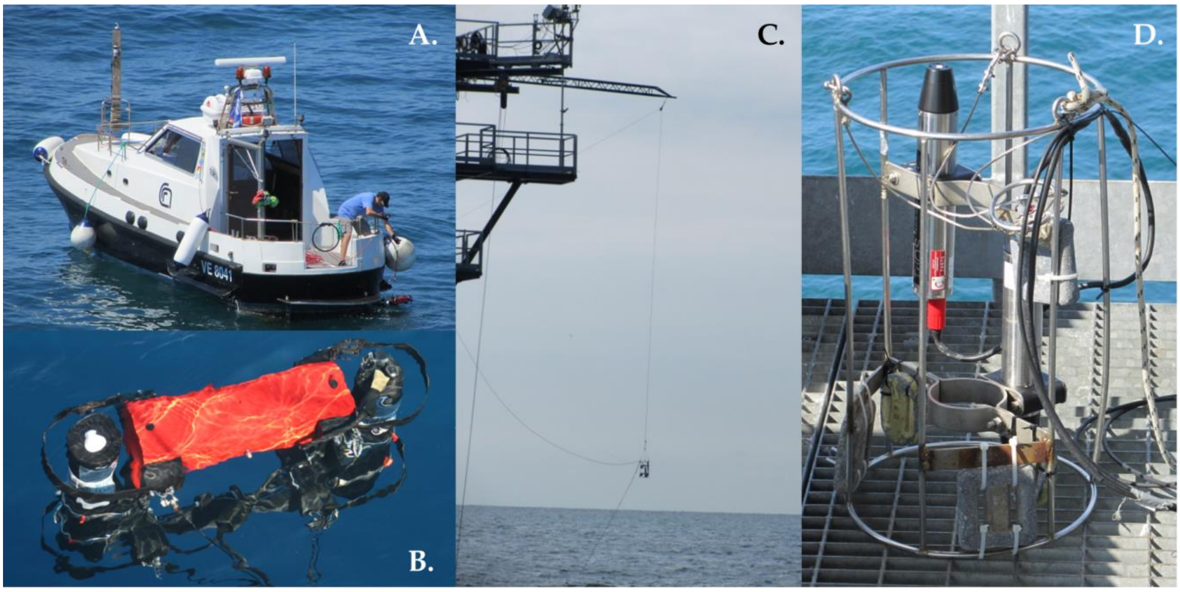


Figure 69. In-water sensors (a) C-OPS being deployed from RV Litus, (b) Positioning of C-OPS in-water, (c) in-water TriOS deployment from an extendable boom on the AAOT (d) TriOS in-water irradiance sensor in metal deployment frame.

For downward irradiance (E_d), there was generally good agreement between sensors with differences of <6 % for most of the sensors over the spectral range 400 nm – 665 nm. One sensor exhibited a systematic bias, of up to 11 %, due to poor cosine response. For L_{sky}, the spectrally averaged difference between optical systems was <2.5 %. For L_t, the difference was <3.5 %. Further details of these results are given in Tilstone et al. 2020 [109]. For R_{rs}, the differences between above-water TriOS RAMSES were <5 % at 443 nm and 560 nm, but were >10 % for some systems at 665 nm. Seabird HyperSAS sensors were on average within 7 % at 443 nm, 3 % at 560 nm, and 14.5 % at 665 nm (see Figure 71). Comparison results are giving an indication of the importance and need for similar regular comparisons in highlighting errors in or differences between sensor systems and methods and helping characterize possible uncertainties.

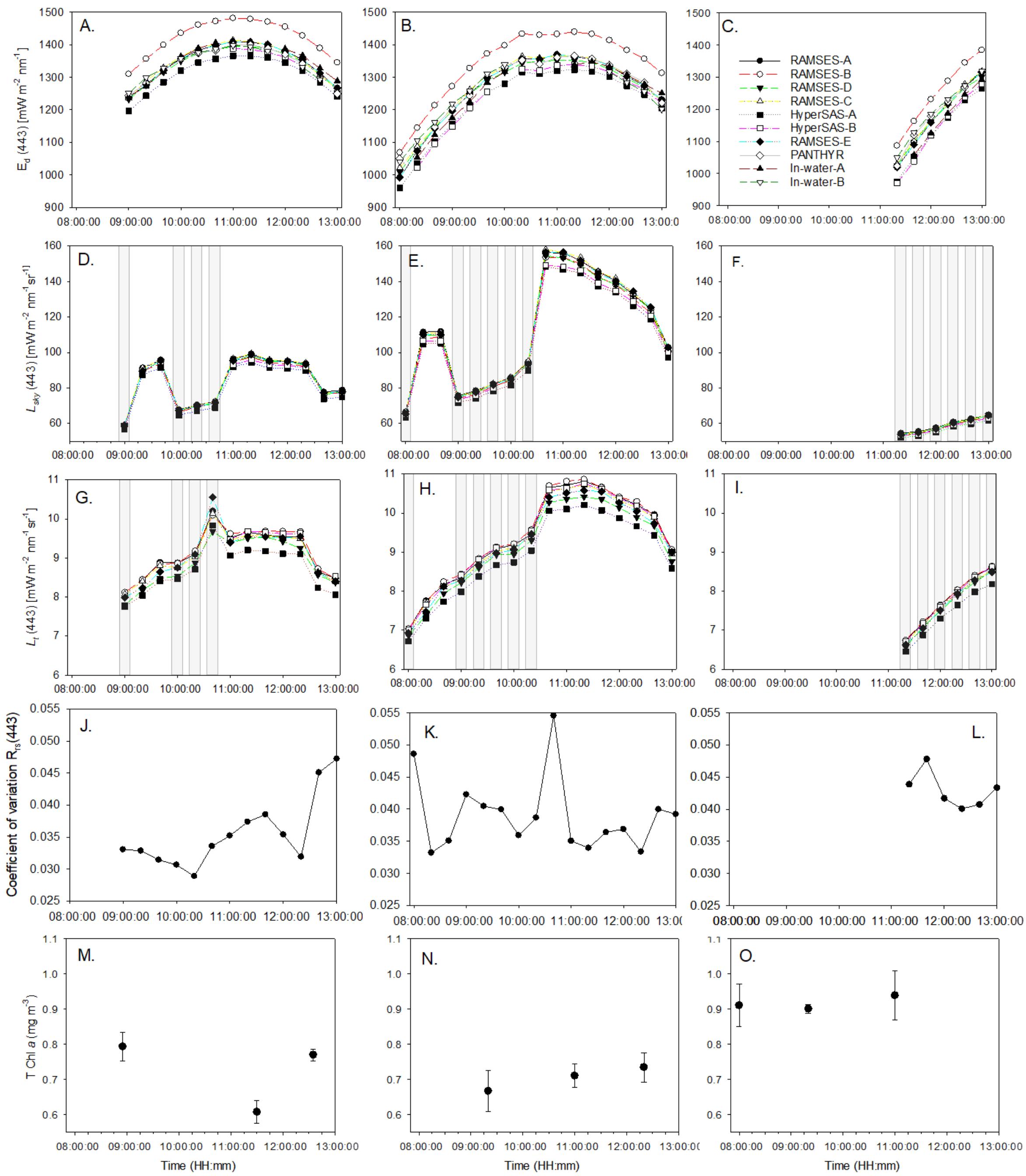


Figure 70. Variation in measurements used for the intercomparison for (A.) $E_d(443)$ on 13 July 2018, (B.) 14 July 2018, (C.) 17 July 2018; $L_{sky}(443)$ on (D.) 13 July 2018, (E.) 14 July 2018, (F.) 17 July 2018; $L_t(443)$ on (G.) 13 July 2018, (H.) 14 July 2018, (I.) 17 July 2018; (J.) Coefficient of variation in $R_{rs}(443)$ on 13 July 2018, (K.) 14 July 2018, (L.) 17 July 2018 and TChl a on (M.) 13 July 2018, (N.) 14 July 2018, (O.) 17 July 2018. Only above water sensor results are shown. L_{sky} and L_t were measured at 90 and 135° relative azimuth. Grey shaded bars represent measurements taken at 135° relative azimuth; the un-shaded area are measurements made at 90° relative azimuth.

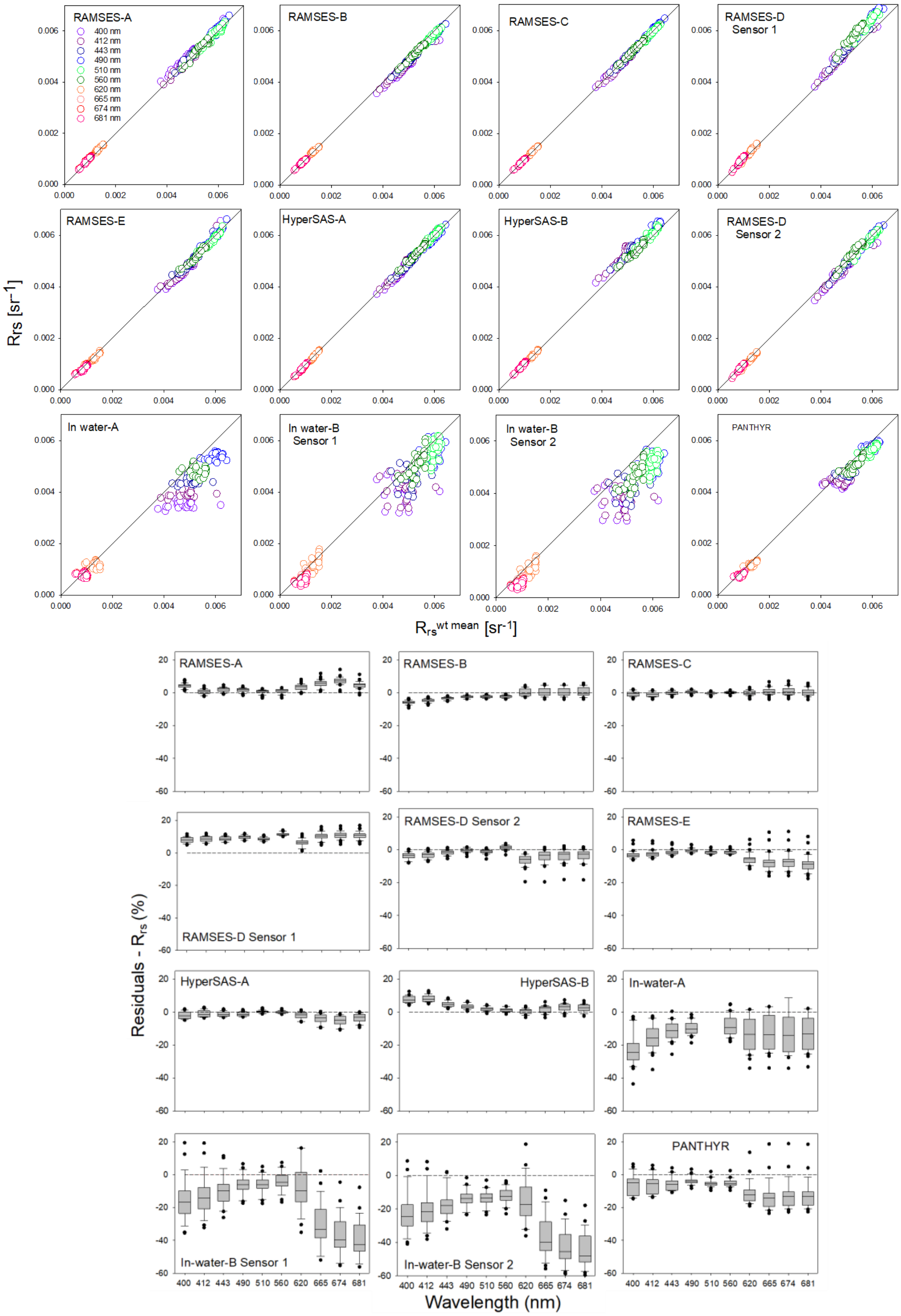


Figure 71. Top Panel: Scatter plots of R_{rs} from the different above- and in-water systems vs. weighted mean R_{rs} ($R_{rs}^{wt\ mean}$) from above-water systems (RAMSES-A, -B, -C, HyperSAS-A, -B). For RAMSES-D and in-water B, S1 is Sensor 1 and S2 is Sensor 2. **Bottom Panel:** Percent residuals of R_{rs} for the different above- and in-water systems. The residuals at each wavelength are calculated from each system as $[(R_{rs} - R_{rs}^{wt\ mean})/R_{rs}^{wt\ mean}]*100$.

16.6 Field inter-comparison experiment database (FICE-DB). [D-210] [213]

During the course of the project, PML designed and built a field intercomparison database for FRM4SOC; the overall design of the database is shown in Figure 72. Essentially this is a PostgreSQL database with a GIS web portal interface. It provides a web interface (https://frm4soc.eofrom.space/) to remotely sensed, modelled and in-situ data. Its functionality includes the ability to carry out simple analysis and plotting, as well as at all stages of analysis, the ability to download data for local processing if preferred.

The portal uses the Open Geospatial Consortium (OGC) Web Map Service (WMS) for displaying imagery data and the OGC Web Feature Service (WFS) and Sensor Observation Service (SOS) interface standards for interacting with in situ data. The analysis and plotting capabilities include: time series; latitude or longitude Hovmöller; scatter / regression; compositing; animations; match-ups from CSV file. Data from the AMT cruises and the AAOT experiment have been included along with the calibration and traceability information for the OCR radiometers that were used throughout the FRM4SOC intercomparisons.

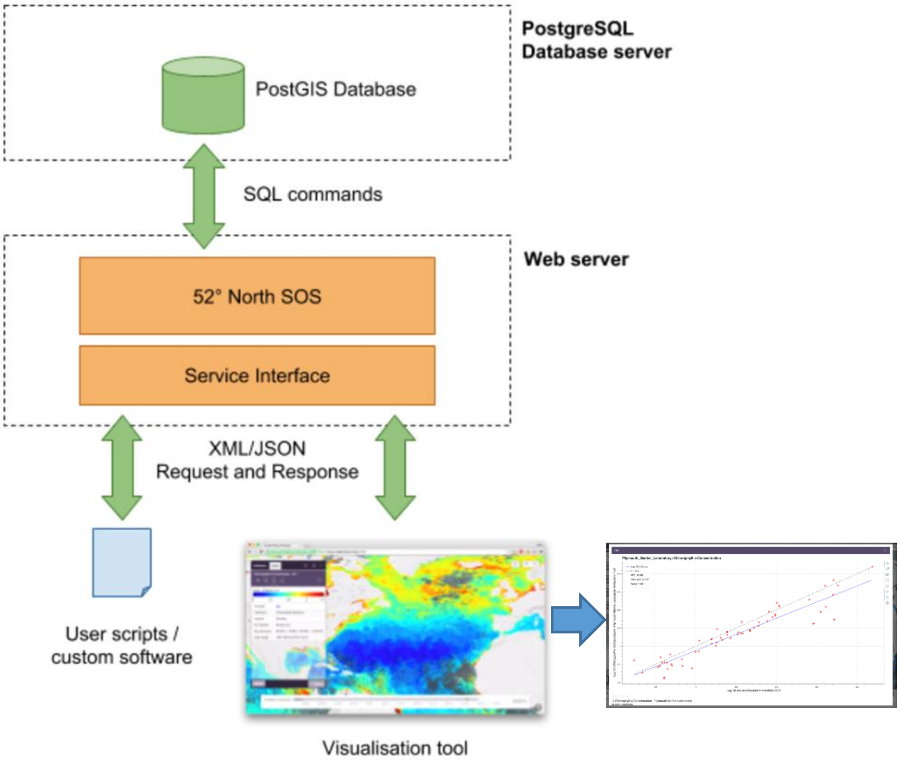


Figure 72. The architecture and functionality of the FRM4SOC field intercomparison database.

17 Uncertainty budgets for FRM OCR

17.1 Introduction

In order to meet Objective 6 of the FRM4SOC project – conduct a full data analysis, derivation and specification of uncertainty budgets for FRM OCR field measurements – the study on uncertainty budgets is presented in the

technical report "Uncertainty Budgets of FRM4SOC Fiducial Reference Measurement (FRM) Ocean Colour Radiometer (OCR) systems used to Validate Satellite OCR products" [D-180], [222] and

- peer-reviewed paper [223] published in the FRM4SOC special issue of the MDPI journal Remote Sensing. **For citation of the chapter 17**, the paper [223] should be considered as the preferable reference.

As required by the FRM4SOC SOW [3], the study

- follows the "Guide to the Expression of Uncertainty in Measurement (GUM) [17];
- describes the methodology used to establish uncertainty budgets for the end-to-end measurement process (Type A and Type B uncertainty) for FRM OCR systems.

17.2 Uncertainty Budgets of FRM4SOC Fiducial Reference Measurement (FRM) Ocean Colour Radiometer (OCR) systems used to Validate Satellite OCR products [D-180]

Having an uncertainty estimate for a measurement result is crucial for objectively and numerically gauging how much trust we can place in that measurement. Furthermore, an uncertainty estimate or budget for a field OCR measurement should be constructed and calculated from uncertainty estimates from an unbroken chain of calibrations back to a primary reference standard (preferably SI), in order for this measurement to be considered as an FRM. This concept of end-to-end uncertainty for FRM4SOC meant using NMI agreed protocols to conduct a derivation and specification of uncertainty budgets for FRM OCR field measurements used for satellite OCR validation and collected as part of FRM4SOC. NPL therefore developed a methodology that was based on the guide to the expression of uncertainty in measurement [17]. This was based on the Monte Carlo method of uncertainty evaluation GUM supplement [224] and calculated this uncertainty budget for three TriOS RAMSES instruments, one ACC-VIS measuring irradiance and two ARC-VIS measuring radiance, supplied by the Tartu Observatory, University of Tartu [225].

These radiometers were used throughout FRM4SOC, i.e. they were calibrated, characterized and used as part of the laboratory intercomparison measurements, the controlled outdoor intercomparison measurements and the FRM4SOC field intercomparison experiment at the Acqua Alta Oceanographic Tower (AAOT) in the Gulf of Venice (see previous sections). These AAOT measurements were used as the example where uncertainty is propagated from the preceding FRM4SOC calibrations and characterisations. Two sets of observations of irradiance and radiance were used from the AAOT, one from 13 July 2018 between 11:00 and 11:04 ('cast 1') and another from 14 July 2018 between 11:40 and 11:44 local time ('cast 2'). At these times, downward irradiance, downwelling radiance and upwelling radiance were all measured simultaneously. Measurements were performed at the AAOT under near ideal conditions, on the same deployment platform and frame (see previous section), under clear sky conditions, sun zenith angles of approximately 24° and moderately low sea state with wind speed of 3.1 m s^{-1} and 0.5 m s^{-1} for each cast. The average chlorophyll content was

Chl = 0.77 mg m^{-3} and absorption of the coloured dissolved organic matter was CDOM (442 nm) = 0.12 m^{-1}.

A Monte Carlo approach was chosen for this uncertainty propagation because the analytical method can become difficult to apply to complex functions with many correlated input parameters where the calculation of sensitivity coefficients is not straightforward. Monte Carlo Methods (MCM) for uncertainty estimation are recognised, accepted and summarised in the GUM supplement [224]. MCM is a numerical method that requires a distinct probability distribution function (PDF) for each of the input components; if input components are correlated then the joint PDF and the measurement equation are required. The MCM will then run a large number of numerical calculations of the measurement equation and with each iteration will use a random choice of each of the inputs from the available range defined by the relevant PDF. The large number of output values calculated using different input values at each iteration, provides the uncertainty of the output value with its PDF.

17.2.1 Uncertainty evaluation methodology

The uncertainties are calculated for the two in situ measurement products: downward irradiance, E_d, and water-leaving radiance, L_w that are convoluted to Sentinel-3 OLCI spectral bands as the final product of interest. The same in situ input data can be used for validation of other satellite sensors as they come from hyperspectral instruments, thus derived radiance and irradiance values can be convoluted with any spectral bands of interest. The wavelength dependence is addressed but omitted in the equations below for better readability. The measurement function for downward irradiance is

$$E_d(\theta_S) = f_{dirr} E_{OLCI}(\theta_S) f_{cos} + (1 - f_{dir}) E_{OLCI}(\theta_S) f_{cosh} + 0. \qquad (21)$$

This measurement equation for downward irradiance, E_d, at a given sun zenith angle, θ_S, is split into two components, one for direct solar irradiance and the other for diffuse sky irradiance. The first term includes the direct-to-total-fraction of irradiance, f_{dirr}, the cosine response for direct irradiance, f_{cos}, and the total measured irradiance, E_{OLCI}, which has already been convolved to OLCI bands and has had various correction factors applied to it. The second term contains, f_{cosh}, the cosine response correction for the full hemispherical diffuse irradiance. The fraction of direct-to-total irradiance, f_{dir}, is also shifted to provide the diffuse-to-total fraction instead. The term 0 is used as a placeholder for any currently undefined model error. The measurement function for water-leaving radiance is

$$L_w(\theta, \Delta\phi, \theta_S) = L_{OLCI,u}(\theta, \Delta\phi, \theta_S) - \rho(\theta, \Delta\phi, \theta_S, W) L_{OLCI,d}(\theta', \Delta\phi, \theta_S) + 0. \qquad (22)$$

For water-leaving radiance, L_w, the measurement setup consists of two radiometers, one pointing upwards towards the sky with the zenith angle, θ' = 140° and the other downwards towards the water, θ = 40°. They are both at the same azimuth angle i.e., the difference between the sun and the sensor ($\Delta\phi$ = 90° or $\Delta\phi$ = 135°). The upward-facing instrument measures the downwelling radiance from the sky (marked here with subscript d for downwelling) while the down-facing instrument measures the upwelling radiance from the water (upwelling, u). In the measurement equation, $L_{OLCI,u}$ is defined as the upwelling radiance from the water, which has had the same correction factors applied as the downward irradiance with the addition of a polarisation correction. Polarisation effects can be assumed to be negligible in irradiance sensors due to their cosine diffuser. The measured values have also been convolved to OLCI bands. $L_{OLCI,d}$, is the equivalent for downwelling radiance. The upwelling radiance includes both light from below the surface (water-leaving radiance) and light which is reflected from the surface of the water. The reflectance of the water is characterised by the Fresnel reflectance, ρ, that is a function of the sensor viewing geometry $(\theta, \Delta\phi)$, solar zenith angle, θ_S and wind speed, W.

The true value of a measurand fully consistent with the definition [19] can never be exactly known; only an estimate can be made which is as good as the instruments and methods used. Therefore, a bias will always exist between the measured value and the best estimate consistent with the definition. Figure 73 and Figure 74 illustrate the errors and respective uncertainty contributions in connection with the measurement equations for downward irradiance and water-leaving radiance respectively. The diagrams demonstrate through the measurement equations how the different uncertainty components contribute to the combined uncertainty of downward irradiance and water-leaving radiance. These diagrams were first designed in the Horizon 2020 FIDUCEO project [226] to show the sources of uncertainty from their origin through to the measurement equation. The outer labels describe the effects, which cause the corresponding uncertainty.

The colour coding of the contributions presented in Figure 73 and Figure 74, although not always straightforward, is a classification of the errors as those due to instrument, environment or applied modelling. All yellow boxes in both figures represent a class of instruments related contributors related to the instruments signal, S, absolute radiometric calibration, (c_{cal}), and the instruments characteristics such as temperature non-stability, (c_{T}), detector non-linearity, (c_{lin}), spectral stray light, (c_{stray}) etc. In addition, in Figure 73 pinkish boxes address the errors due to the angular response of the cosine diffuser. The turquoise box represents the convolution of spectral bands with the satellite spectral response function, thus classified as a modelling component. For the irradiance, the green box represent the fraction of the direct to total irradiance and this is an example where the environmental and modelling contributions are tangled together, as some model input includes environmental conditions such as the actual value of the aerosol optical depth (AOD). A similar situation applies to the red box in Figure 74, i.e. for water-leaving radiance estimating the Fresnel reflectance values. This is, again, a combination of environmental influence (wind speed) and modelling errors that are used to derive the reflectance. The blue box with the +0 term represents any other environmental effects that are not fully accounted for in the current version of the uncertainty budget. This can be the structure shadings effect, for example.

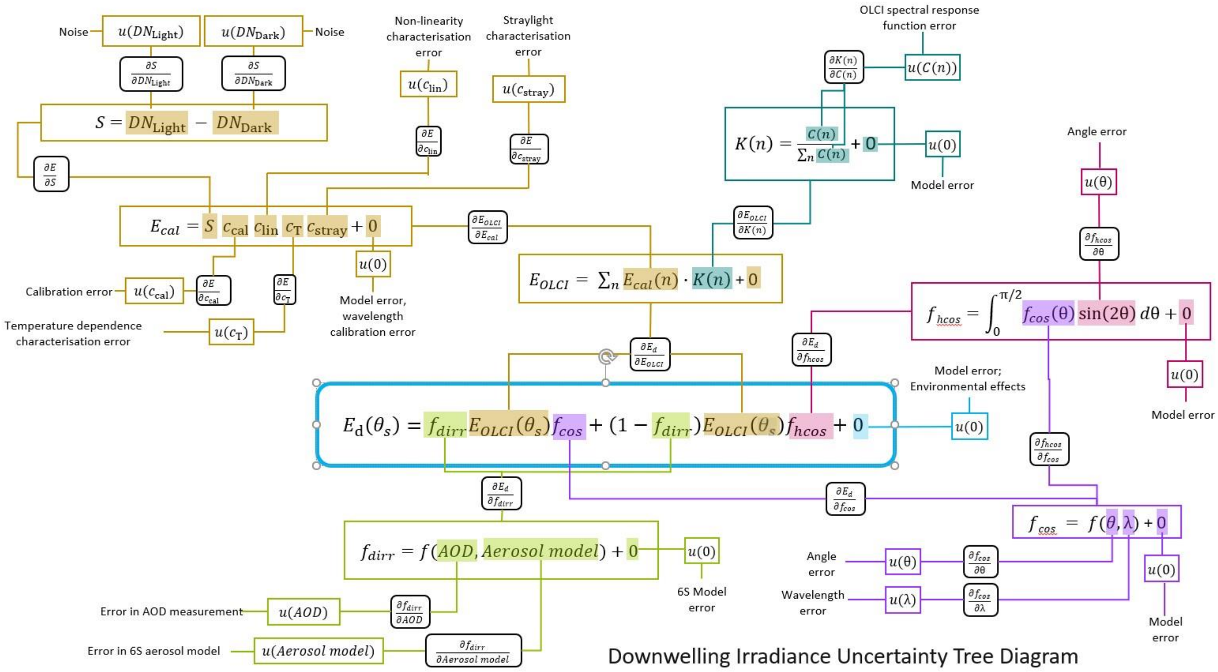


Figure 73. Uncertainty tree diagram for downward irradiance (E_{d}).

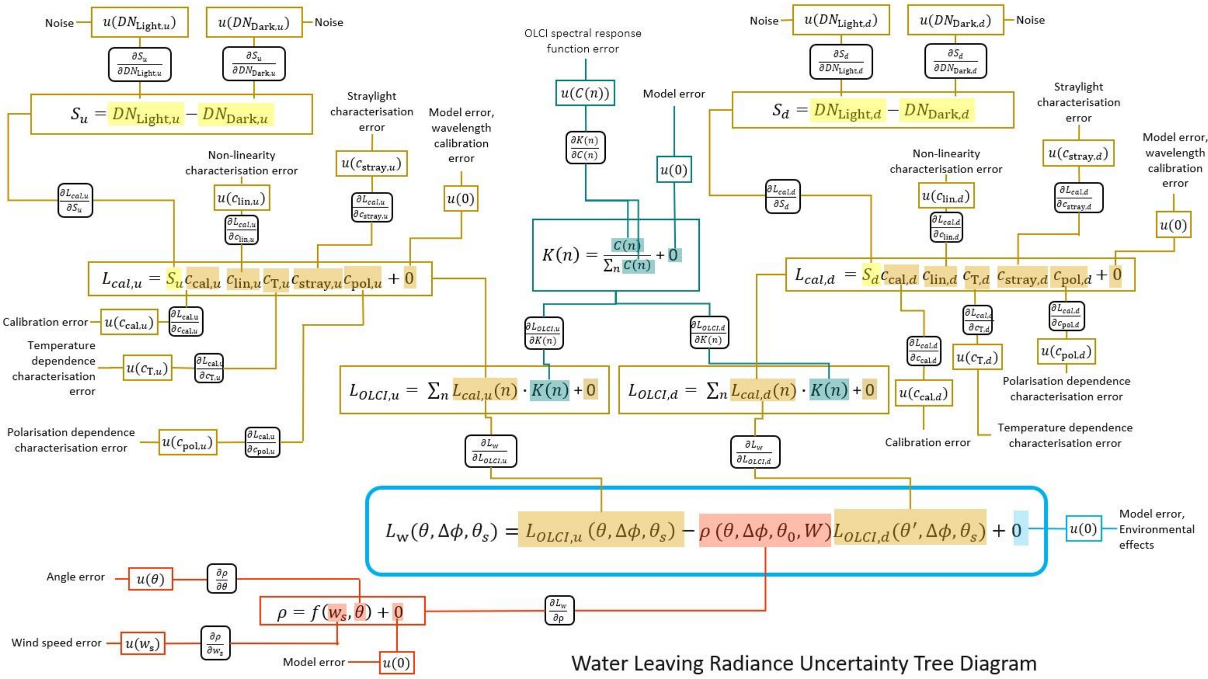


Figure 74. Uncertainty tree diagram for water-leaving radiance (L_{w}).

17.2.2 Uncertainty calculation: defining the PDF-s for some inputs.

17.2.2.1 Data Processing Steps

The processing follows the structure of the measurement equations in Figure 73 and Figure 74. Firstly, all parameters are assigned a PDF and a decision is made over whether any correlation should be assigned for this parameter. Following this, the correction factors (calibration, non-linearity, temperature and stray light and polarisation for radiance only) are applied to the signal. The next step involves a band integration, which is performed to convolve the hyperspectral instrument wavelengths with seven OLCI bands. The final step is to calculate the downward irradiance and water-leaving radiance respectively using the defined and calculated parameters in the final measurement equations (21) and (22). Each PDF is assigned 10,000 draws for this MCM process.

17.2.2.2 Instrument Signal (S)

The instrument signal is defined as the digital numbers when the sensor is exposed to light conditions (DN_{light}) subtracted by the digital numbers in dark conditions (DN_{dark}). The RAMSES radiometers do not have a mechanical internal shutter. Instead, black-painted pixels on the photodiode array are used to derive the dark signal and electronical drifts [190]. The values obtained by NPL for further analysis were already converted to radiometric values. Nevertheless, the statistics of masked pixels are used for assessment of the signal uncertainty.

The main steps in defining the PDF for the signal are firstly interpolating the signal to align with OLCI wavelengths and secondly deciding how best to approach the low number of repeated measurements (15 repetitions for cast 1 and 12 for cast 2). Here, we consider both procedures.

The three instruments used in this study have a spectral range from 320 to 1050 nm, however for this exercise we are interested in seven OLCI bands (400, 442.5, 490, 560, 665, 778.8, 865 nm) since this procedure is intended to inform comparisons of ground in situ OCR data with satellite sensors. Therefore, the signal is extracted from only the wavelengths overlapping with the OLCI SRF of the seven OLCI bands and then linearly interpolated to match the 200 wavelengths of the OLCI spectral response functions (SRF) [194] (see Figure 75). There are 200 OLCI SRF values for each band, which means the aforementioned signal interpolation to OLCI wavelengths produces 200 values for each repeated measurement. In the next step, the mean of the repeated measurements is calculated for each of the 200 wavelengths, producing 200 signal values per band. These mean values are assigned as the mean of 200 Gaussian distributions, which represents the PDF of the signal at each of these wavelengths. Later, through the band integration step these values are convolved with the OLCI bands to produce one value per band for E_{OLCI}.

The readings taken for DN_{light} and DN_{dark} only covered a 5-min window, meaning that only 15 repeat measurements were taken for cast 1 and 12 repeats for cast 2. This is a small number of repeated measurements and is far from enough to get a representative mean and standard deviation of the measurements. Since there is not a smooth distribution, it is difficult to decide whether to use the mean, median or mode of the repeated measurements. In this study, the typical standard uncertainty of the mean formula is used.

Signal

wvl	Repeat 1	Repeat 2	Repeat 3	...
382.93	0.172	0.164	0.170	
386.28	0.170	0.163	0.169	
389.62	0.177	0.170	0.175	
392.97	0.186	0.179	0.184	
396.32	0.213	0.205	0.211	
399.67	0.260	0.250	0.257	
403.01	0.299	0.289	0.296	
406.36	0.324	0.312	0.321	
409.71	0.347	0.336	0.344	
413.06	0.374	0.362	0.371	
416.41	0.399	0.386	0.396	
419.77	0.421	0.408	0.417	
423.12	0.438	0.425	0.434	
426.47	0.446	0.433	0.442	
429.82	0.460	0.447	0.456	
433.18	0.494	0.480	0.489	
436.53	0.535	0.521	0.530	

219

SRF

wvl	SRF
387.75	5.7E-08
387.86	1.1E-07
387.98	2.1E-07
388.10	3.9E-07
388.22	7.1E-07
388.34	1.3E-06
388.46	2.3E-06
388.57	4.1E-06
388.69	7.2E-06
388.81	1.2E-05
388.93	2.1E-05
389.05	3.6E-05
389.17	5.9E-05
389.28	9.6E-05
389.40	1.6E-04
389.52	2.5E-04
389.64	3.9E-04
...	...
410.94	7.1E-07
411.06	3.8E-07
411.18	2.0E-07
411.30	1.0E-07

200

Signal interpolated to SRF wvls

Repeat 1	Repeat 2	Repeat 3	...
0.173	0.166	0.166	
0.174	0.167	0.167	
0.174	0.167	0.167	
0.174	0.167	0.167	
0.174	0.167	0.167	
0.175	0.168	0.168	
0.175	0.168	0.168	
0.175	0.168	0.168	
0.175	0.168	0.168	
0.176	0.169	0.169	
0.176	0.169	0.169	
0.176	0.169	0.169	
0.176	0.169	0.169	
0.177	0.170	0.170	
0.177	0.170	0.170	
0.177	0.170	0.170	
0.177	0.170	0.170	
...	...	...	
0.357	0.345	0.345	
0.358	0.346	0.346	
0.359	0.347	0.347	
0.360	0.348	0.348	

200

Mean of repeats

Mean
0.169
0.169
0.169
0.169
0.170
0.170
0.170
0.170
0.171
0.171
0.171
0.171
0.172
0.172
0.172
0.172
0.173
...
0.349
0.350
0.351
0.352

200

Figure 75. Description aid for the process of interpolating the signal values to the SRF values and finding the mean of the repeats. Note that this is for one band only; each band will have its own set of tables. The blue number below each table shows the number of rows in that table.

17.2.2.3 Calibration Coefficients (c_{cal})

The calibration of each instrument was performed in the optical laboratory at the University of Tartu. The calibration coefficients and associated uncertainty values are incorporated into the Monte Carlo analysis through the appropriate PDF. The instrument readings are automatically adjusted for the calibration coefficients, thus in the MCM, the PDF is taken to be a Gaussian distribution with a mean equal to 1. The standard deviation is equal to the standard uncertainty associated with the calibration.

17.2.2.4 Non-Linearity Correction (c_{lin})

The integration time used when capturing the measured spectra can lead to non-linearity effects in the results. The maximum value of non-linearity effects was determined in the indoor calibration for the two radiance instruments, but not for the irradiance instrument [168]. As the principal aim of this study is to outline the method and only secondarily, to provide results, the radiance instrument corrections are used for the irradiance instrument. The non-linearity correction values are provided for all instrument wavelengths. Similarly, to the instrument signal, in order to align with the OLCI bands, the only non-linearity values used are those whose wavelengths overlap with the SRFs of each of the seven OLCI bands. These are then interpolated to match the 200 OLCI SRF wavelengths. The interpolation method chosen was linear interpolation due to the smoothness of the non-linearity correction curve. Each of the 200 resultant non-linearity correction values is assigned a PDF, which is a rectangular distribution with a mean value of 1 and half-width equal to the linearity correction value.

17.2.2.5 Temperature Correction (c_T)

The variation of the instrumental calibration coefficients due to temperature is based on a previous evaluation [180]. The variability in % at the seven central wavelengths of the bands of interest (400, 442.5, 490, 560, 665, 778.8, 865 nm) were selected. The seven PDFs which represent the temperature correction at each band were defined to be Gaussian distributions with a mean value of 1 and a standard deviation equal to the aforementioned variability.

17.2.2.6 Stray Light Correction (c_{stray})

Scattering or reflections in the radiometer optics cause light from one part of the spectrum to fall on pixels associated with light from another part of the spectrum. This effect is known as spectral stray light and is common in hyperspectral instruments and must be corrected for.

This study was intentionally planned so that all instruments would be well characterised, whereas typical campaigns have much less information about the instrument performance, including the stray light characterisation. Hence, we consider two scenarios for the stray light. The first case is for an ideal situation in which the stray light is corrected for based on the performed characterisation and we assign an uncertainty to the method described in point i) below; the second, non-ideal, case does not correct for stray light and instead we demonstrate the scale of uncertainty which will result from this in point ii).

i) The stray light characterisation provides correction values for each of the instrument wavelengths (see Figure 76). The correction values obtained are quite erratic. Due to the correction values being low in magnitude, it is possible that the erratic behaviour could be due to noise from the stray light measurement. The stray light correction values used were calculated by convoluting the stray light values with the OLCI SRF. The selected values are presented in Figure 76 as the green cross series.

For the ideal case, the PDFs assigned for stray light for each band is a Gaussian distribution with a mean equal to the stray light correction acquired from the polynomial. The stray light characterisation does have an uncertainty associated with it, but this is unknown. A value of 5% is chosen which accommodates for the true uncertainty and is negligible in comparison to the assigned uncertainty of other variables [168], meaning that the quantity will have little or no effect on the further results.

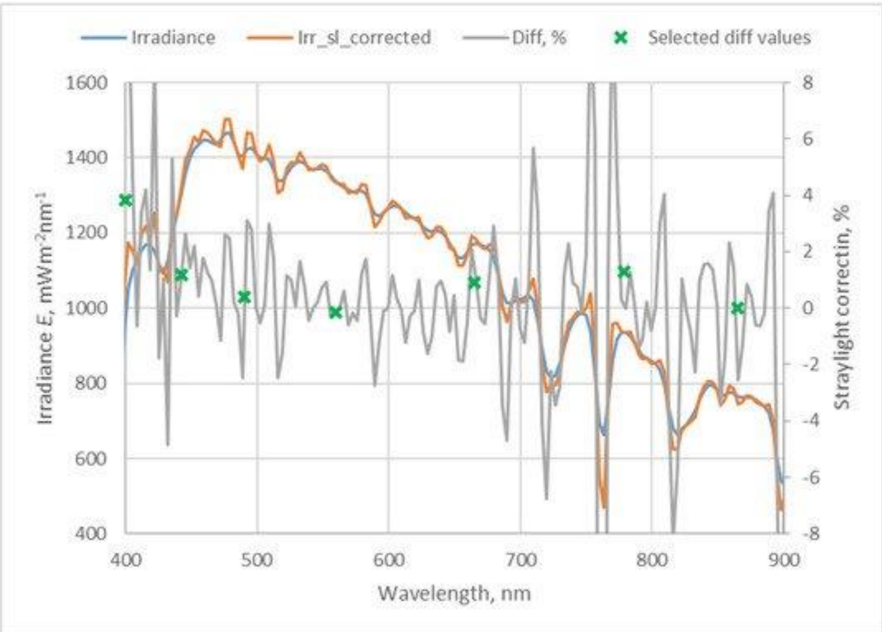


Figure 76. The stray light correction of the irradiance measurements and the percentage difference between non-corrected and corrected. The selected values are calculated using a weighted interpolation technique.

ii) For the non-ideal case, no correction is applied. The assigned distribution is a Gaussian distribution with mean equal to 1 and standard deviation equal to the fitted polynomial stray light correction (see Figure 76 and Table 22 at the seven OLCI bands of interest.

Table 22. Stray light correction values applied to the three radiometers in the ideal case and used as the standard deviation of a Gaussian distribution in the non-ideal case.

	Wavelength	400 nm	442.5 nm	490 nm	560 nm	665 nm	778.8 nm	865 nm
Stray light correction (%)	Downward irradiance	3.82	1.17	0.38	−0.17	0.89	1.31	0.03
	Upwelling radiance	2.40	0.84	−0.01	0.30	−1.01	−0.21	−15.2
	Downwelling radiance	2.77	1.92	−0.48	−0.81	0.27	1.45	−7.02

17.2.2.7 The fraction of direct to total irradiance (f_{dirr})

The fraction of direct to total irradiance is applicable to downward irradiance measurements and can be estimated using measurements of the aerosol optical depth, water vapour content and total column ozone in a radiative transfer model. This takes into account the atmospheric transmission of radiation for the conditions specified. In this study, an AERONET-OC [91] station at the observation site provided the atmospheric conditions at the time of data acquisition, and then the radiative transfer model, 6S [227], was used to estimate the direct to total ratio. The atmospheric parameters for both casts were the following: AOD at 550 nm 0.112 and 0.297; precipitable water 2.83 cm and 3.13 cm; ozone 330.1 DU and 329.5 DU. The uncertainty components of f_{dirr} consist of i) the accuracy of the 6S radiative transfer model, ii) the uncertainty of the inputs to 6S and iii) an error related to the designated atmosphere type.

i) 6S does not provide an estimate of its own accuracy; however a comparison of 6S with a highly accurate Monte Carlo radiative transfer yielded a maximum observed relative difference between the two methods of 0.79 % for a maritime atmosphere [228]. This value is used as an estimate of 6S uncertainty and is shown as the “6S model accuracy” row in Table 23.

ii) There are several inputs of the AERONET data to 6S, namely aerosol optical depth (AOD) at 550 nm, precipitable water and total column ozone. The only of the three variables that causes a change in f_{dirr}, is the AOD at 550 nm. Therefore, the only variable that we need to consider the uncertainty of, is AOD at 550 nm, which has an uncertainty of 0.01 according to [32]. The sensitivity of the resultant f_{dirr} in 6S with a change of ±0.01 in AOD has been calculated. The difference observed makes up the corresponding uncertainty values shown in the rows labelled “AOD 550 nm” in Table 23.

iii) The AAOT site is eight nautical miles from the coast of Venice; the atmosphere is between continental and maritime and should possibly be considered coastal.

However, the 6S model has several defined atmospheres for the user to select which does not include coastal, hence for this study the 'maritime' option will be used and the error in this assumption will be estimated by comparing the results of both 'maritime' and 'continental'. The differences between the two atmospheric types varies across wavelengths, hence the uncertainty due to this will also be wavelength dependent (see the row labelled "Atmosphere type assumptions" in Table 23).

Assuming that each contributor is independent, using the law of propagation of uncertainties, we can calculate the uncertainty of f_{dirr} (see Table 23). This uncertainty is applied as half the width of a rectangular distribution centered on a mean value of 1. Note that, here we assume Table 23 shows the uncertainty in terms of the width of a rectangular PDF, whereas in further processing this uncertainty is used as the standard deviation of the Gaussian PDF where the values used in MCM are quoted). This means that the values for f_{dirr} are not equal since the standard deviation of a rectangular distribution is smaller than half the width of a rectangular distribution.

Table 23. Uncertainty components for the model, input data (AOD 550 nm) and assumptions relating to the atmosphere specified for each band of interest. All values, except model accuracy, are written in terms of the uncertainty applied to the output, f_{dirr}.

	400 nm	442.5 nm	490 nm	560 nm	665 nm	778.8 nm	865 nm
6S model accuracy			0.79				
AOD 550 nm (cast 1)	0.0077	0.0081	0.0085	0.0083	0.0082	0.0077	0.0077
AOD 550 nm (cast 2)	0.0063	0.0066	0.0070	0.0069	0.0069	0.0068	0.0067
Assumed Atmosphere	0.008	0.001	0.010	0.0023	0.0041	0.058	0.067
f_{dirr} (cast 1)	0.66	0.73	0.78	0.83	0.86	0.88	0.89
f_{dirr} (cast 2)	0.54	0.60	0.64	0.68	0.72	0.74	0.75
Uncertainty f_{dirr} (cast 1)	0.012 (1.82 %)	0.010 (1.36 %)	0.014 (1.83 %)	0.025 (3.08 %)	0.042 (4.92 %)	0.059 (6.65 %)	0.068 (7.63 %)
Uncertainty f_{dirr} (cast 2)	0.011 (2.01 %)	0.008 (1.36 %)	0.013 (2.03 %)	0.025 (3.63 %)	0.042 (5.86 %)	0.058 (7.92 %)	0.068 (9.05 %)

17.2.2.8 Cosine response (f_{cos} and f_{cosh})

To capture irradiance over a full hemisphere, instruments are equipped with cosine diffusers. The ideal diffuser will transmit light in proportion with the cosine of the incident angle. However, instruments always differ from the theoretical ideal, hence the need for this to be characterised and corrected for, meaning the residual elements of uncertainty depend on the correction method.

Downward irradiance is made up of two components: direct solar irradiance and diffuse sky irradiance. These components are not affected the same way by a non-perfect cosine response and so are separated in the measurement equation. The cosine response term, f_{cos}, incorporates only the direct solar component thus relates to the SZA during the measurement and the error in the cosine response diffuser for this particular angle. Whereas, f_{cosh}, the cosine response over the full hemisphere, integrates the diffuse light component across the hemisphere, and integrates the deviation from the perfect diffuser across the whole hemisphere as well.

In this study, the cosine response has been fully characterised and is propagated as the ideal case below i). However, often the cosine response is unknown and thus not corrected for. This section demonstrates the impact of an additional scenario ii), in which the only information known about the cosine response is the manufacturer's quote of the uncertainty due to the cosine response. Here we choose the value of 3 % as this is similar to the cosine response error in this study for all bands and is also a typical value quoted by manufacturers for angles under 60° [205]. The ideal case is having the cosine response errors for a range of solar zenith angles and wavelengths. In this study, the diffuser was characterised at TO before the field comparison in May 2017 for 45 angles across the hemisphere and seven wavelengths. Figure 77 shows the results of the laboratory test for the instruments used in this study. The average of the four repeated measurements was linearly interpolated to the solar zenith angle and the central wavelength of each of the seven bands of interest. The cosine response correction is treated as a Gaussian distribution centred on the interpolated value. The standard deviation of this distribution is assigned from the standard deviation of the four repeated measurements.

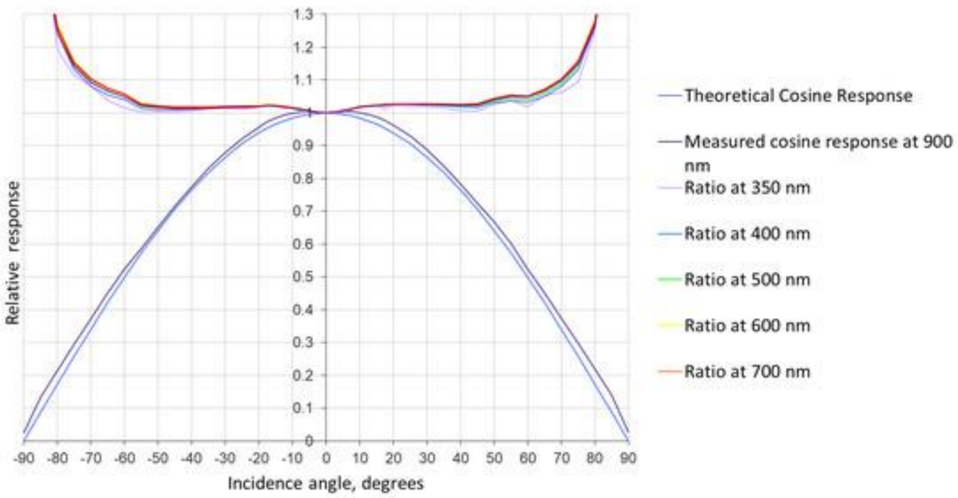


Figure 77. Diffuser's cosine response test results. Series marked as ratio at particular wavelengths are the ratios of the measured instrument response to the theoretical cosine response for a given angle and normalised to 0°.

The non-ideal case examines the typical scenario in which a manufacturer has quoted the cosine response to be equal to 3 % with no additional information. Typically, this is assigned as the standard deviation of a Gaussian PDF of mean 1 (i.e. no correction is applied, just an uncertainty). However, the cosine response is a systematic error and should shift the mean value of E_d as in the ideal scenario i), but this will not happen if parameterised as a PDF of mean 1, hence the mean value is incorrect. This should be accounted for by applying an additional measure of uncertainty based on the impact on E_d of not accounting for this bias.

The diffuse component requires all cosine responses to be integrated over the hemisphere. Assuming an isotropic sky radiance distribution, this is calculated using:

$$f_{cos_h} = \int_0^{2\pi} f_{cos}(\theta)\, sin\,(2\theta) d\theta, \qquad (23)$$

where f_{cosh} is the integrated cosine response over the full hemisphere, and $f_{cos}(\theta)$ is the cosine response for a given illumination angle. We again look at two similar scenarios, the ideal case i) demonstrates a scenario in which all cosine response errors are known, and the non-ideal ii) demonstrates a situation where the only information known regarding the response is that it is within 3 % for angles < 60° and 10 % for angles > 60°. The manufacturer in this study provided no values for the cosine response, but these values are typical for some manufacturers (e.g. Sea-Bird Scientific HyperOCR radiometer [205]). The isotropic sky is a simplification and clear-sky radiance distributions are not isotropic and normally show larger radiances for large zenith angles (e.g. [229]) and band circumsolar brightening (aureole) [230]. For measurements presented here with a SZA of approximately 24° the aureole effect is minimised by the small errors in cosine response of the diffuser for small incidence angles (see Figure 77). For the horizon brightening the $\sin(2\theta)$ factor in Equation (23) would however minimize their contribution to f_{cosh}. The diffuse component of the downward irradiance for clear skies is about 40 % … 30 % for short wavelengths and decreases at longer wavelengths (see Table 23 for the actual values observed during the field measurements), further reducing the impact of f_{cosh} on the E_d uncertainty evaluation. Obviously, the situation is different for hazier skies, or for high values of the solar zenith angle, which anyway for the satellite validation activities are not recommended. For the ideal case, the cosine response values are corrected for by taking an average over the four repeated measurements for f_{cos} at each angle and each wavelength. Then these can be integrated over all angles and interpolated to the wavelength of interest. The PDF assigned for the diffuse component is a Gaussian distribution with a mean equal to these integrated cosine values and a standard deviation originating from the standard deviation of the four repeated measurements of the cosine response. The non-ideal case does not correct for the cosine response, therefore the mean of the Gaussian distribution is equal to 1 and the standard deviation relates to 3 % for angles ≤60° and 10 % for angles >60° (which are typical for HyperOCR radiometer [205]).

However, the cosine response is a systematic error, which should be corrected for, hence an additional measure of uncertainty must be applied to take into account the lack of completing a correction. This is equal to the difference in the mean values of the resultant E_d calculated with and without a correction applied (i.e. the difference between the resultant E_d from the ideal and non-ideal cases).

17.2.3 Downward Irradiance

To propagate uncertainty for the measurands of interest for FRM4SOC (E_d and L_w) the following Monte Carlo approach was applied:

1. Measurement functions were defined based on the uncertainty tree diagrams that include all inputs defined as quantities that can have an influence on the measurand.
2. All inputs had their estimates in terms of a probability density function (PDF) with associated magnitudes (values) and shapes of the PDF (standard uncertainties).
3. The measurement equations were run with random inputs defined by the PDFs a large number of times (10^4 in this case).
4. The correlation between some input quantities (for example, the absolute radiometric calibration coefficients of the different instruments) was handled as systematic contributions, thus the draws from that distribution are not randomised.
5. The final estimate – magnitude and its uncertainty value – is derived from the resultant PDF.
6. All uncertainties are reported with a k = 1 coverage factor.

a)

	S	c_{cal}	c_{lin}	c_{temp}	c_{stray}	E_{cal}	K(n)	E_{OLCI}	f_{cos}	f_{hcos}	f_{dirr}	E_d
400	1070.0	1.0	1.0	1.0	1.04	1110.0	0.0084	1120.0	1.03	1.04	0.66	1150.0
442.5	1480.0	1.0	1.0	1.0	1.01	1500.0	0.0095	1500.0	1.03	1.04	0.73	1540.0
490	1590.0	1.0	1.0	1.0	1.0	1600.0	0.0094	1600.0	1.03	1.04	0.78	1650.0
560	1510.0	1.0	1.0	1.0	0.998	1510.0	0.0094	1510.0	1.03	1.05	0.83	1550.0
665	1320.0	1.0	1.0	1.0	1.01	1330.0	0.0094	1330.0	1.03	1.05	0.86	1370.0
778.8	1060.0	1.0	1.0	1.0	1.01	1070.0	0.008	1060.0	1.03	1.05	0.88	1090.0
865	873.0	1.0	1.0	1.0	1.0	873.0	0.0072	876.0	1.02	1.04	0.89	899.0

	S	c_{cal}	c_{lin}	c_{temp}	c_{stray}	E_{cal}	K(n)	E_{OLCI}	f_{cos}	f_{hcos}	f_{dirr}	E_d
400	0.0228	0.873	0.0171	0.203	0.19	0.913	2.2	0.913	0.619	0.312	1.0	1.0
442.5	0.0179	0.683	0.133	0.203	0.057	0.723	1.5	0.722	0.612	0.215	0.79	0.85
490	0.017	0.647	0.292	0.203	0.019	0.741	1.2	0.741	0.596	0.257	1.1	0.873
560	0.0142	0.619	0.902	0.305	0.0083	1.13	0.9	1.13	0.585	0.194	1.8	1.23
665	0.0147	0.599	0.957	0.711	0.044	1.35	0.78	1.35	0.554	0.188	2.8	1.43
778.8	0.0158	0.618	0.878	1.32	0.065	1.69	0.74	1.69	0.542	0.199	3.8	1.76
865	0.0246	0.557	0.508	2.74	0.0012	2.86	0.73	2.86	0.536	0.174	4.4	2.9

b)

	S	c_{cal}	c_{lin}	c_{temp}	c_{stray}	E_{cal}	K(n)	E_{OLCI}	f_{cos}	f_{hcos}	f_{dirr}	E_d
400	1070.0	1.0	1.0	1.0	1.0	1070.0	0.0084	1080.0	1.0	1.0	0.66	1080.0
442.5	1480.0	1.0	1.0	1.0	1.0	1480.0	0.0095	1480.0	1.0	1.0	0.73	1480.0
490	1590.0	1.0	1.0	1.0	1.0	1590.0	0.0094	1590.0	1.0	1.0	0.78	1590.0
560	1510.0	1.0	1.0	1.0	1.0	1510.0	0.0094	1510.0	1.0	1.0	0.83	1510.0
665	1320.0	1.0	1.0	1.0	1.0	1320.0	0.0094	1320.0	1.0	1.0	0.86	1320.0
778.8	1060.0	1.0	1.0	1.0	1.0	1060.0	0.008	1050.0	1.0	1.0	0.88	1050.0
865	873.0	1.0	1.0	1.0	1.0	872.0	0.0072	875.0	1.0	1.0	0.89	876.0

	S	c_{cal}	c_{lin}	c_{temp}	c_{stray}	E_{cal}	K(n)	E_{OLCI}	f_{cos}	f_{hcos}	f_{dirr}	E_d
400	0.0228	0.873	0.0171	0.203	3.9	3.96	2.3	3.96	3.01	1.34	1.1	4.44
442.5	0.018	0.683	0.133	0.203	1.2	1.38	1.5	1.38	2.99	1.35	0.78	2.62
490	0.0165	0.647	0.292	0.203	0.38	0.839	1.1	0.84	2.99	1.33	1.1	2.51
560	0.0141	0.619	0.902	0.305	0.16	1.14	0.91	1.14	2.99	1.36	1.8	2.73
665	0.015	0.599	0.957	0.711	0.89	1.62	0.79	1.62	2.98	1.33	2.8	3.04
778.8	0.016	0.618	0.878	1.32	1.3	2.13	0.74	2.13	3.0	1.32	3.8	3.41
865	0.025	0.557	0.508	2.74	0.025	2.86	0.74	2.86	2.96	1.34	4.4	3.9

Figure 78. Irradiance MCM outputs. (a) Ideal case results. The mean (top panel) and standard uncertainty as a percentage of the mean (second panel) of each variable. This is shown for all OLCI bands of interest and relates to cast 1. The colour code intensity indicates the magnitude. (b) Non-ideal case results. The mean (third panel) and standard uncertainty as a percentage of the mean (lower panel) of each variable.

The ideal scenario relates to the case in which the cosine correction and stray light were well characterised and were corrected for. As previously mentioned, the MCM requires each parameter in the measurement equations to be assigned a PDF based on collected data, or best knowledge, etc. These distributions, as shown in Figure 78, were then propagated through the measurement equation to find the downward irradiance, E_d. Figure 78 presents the summary values of the MCM simulation for cast 1. The value of several of the variables (S, c_{lin}, E_{cal}, $K(n)$) is chosen for just one wavelength out of 200. Most of these variables vary over the 200 wavelengths. Here we present the 99 value as an example for reference. In the calculations, all 200 values are propagated.

The results for the cases in which neither stray light nor the cosine error are corrected for look different (the two bottom panels in Figure 78). They were each assigned a mean value of 1 and an uncertainty relating to typical manufacturers' estimates (cosine) and the correction itself (stray light). It is possible to see the effects of stray light and the cosine error separately since each parameter exists in separate parts of the measurement equation. For example, it is clear that the mean of c_{stray} is set to 1 and the standard deviation is much higher in the non-ideal case than the ideal. Additionally, the changes in parameters can be tracked through up to E_{olci}. Beyond this, the correction is mixed with the cosine correction, which can be seen to dominate the standard deviation of the of E_d in majority of bands.

The uncertainty presented in the lower panel of Figure 78 is an under-representation of the true uncertainty that is applicable when corrections are not applied. This is because we have taken into account the uncertainty in the measurement but have not accounted for the lack of correction. Table 24 provides an overview of the values of interest in this correction. The mean values of the PDF function obtained in the MCM processing for downward irradiance are shown for the ideal and non- ideal case in irradiance units (column two and three in the Table 24). Column 4 contains the bias values, thus the difference between the two means. The columns 5 to 7 present relative standard uncertainties for the three different scenarios.

Table 24. The mean and standard uncertainty as a percentage of the mean of the downward irradiance, E_d, presented for the ideal and non-ideal cases.

	Mean E_d, mW m^{-2} nm^{-1}			Standard uncertainty $u(E_\mathrm{d})$, %		
Band ,nm	**Ideal**	**Non-ideal**	**Bias of E_d**	**Ideal**	**Non-ideal**	**Corr. non-ideal**
400	1150	1080	73.9	1.0	4.4	11
442.5	1540	1480	60.8	0.8	2.6	6.7
490	1650	1590	53.8	0.9	2.5	5.9
560	1550	1510	43.5	1.2	2.7	5.6
665	1370	1320	51.9	1.4	3.0	7
778.8	1090	1050	44.6	1.8	3.4	7.7
865	900	877	23.6	2.9	3.9	6.6

17.2.4 Water-Leaving Radiance

The same approach as for the downward irradiance case is used to present, in Figure 79, the water-leaving radiance results. The value of several of the variables (S_u, S_d, $c_{lin,u}$, $c_{lin,d}$, $c_{stray,u}$, $c_{stray,d}$, $L_{cal,u}$, $L_{cal,d}$, $K(n)$) presented in Figure 79 is chosen for just one wavelength out of 200. Most of these variables vary over the 200 wavelengths. Here we present the 99 value as an example for reference. In the calculations, all 200 values are propagated.

a)

	S_u	S_d	c_{cal}	c_{lin_u}	c_{lin_d}	c_{temp}	c_{stray_u}	c_{stray_d}	c_{pol_u}	c_{pol_d}	$L_{cal,u}$	$L_{cal,d}$	K(n)	$L_{OLCI,u}$	$L_{OLCI,d}$	rho	L_w
400	7.06	51.9	1.0	1.0	1.0	1.0	1.02	1.03	1.0	1.0	7.23	53.3	0.0084	7.28	53.6	0.026	5.9
442.5	9.62	95.1	1.0	1.0	1.0	1.0	1.01	1.02	1.0	1.0	9.7	96.9	0.0095	9.71	95.2	0.026	7.26
490	11.4	89.6	1.0	1.0	1.0	1.0	1.0	0.995	1.0	1.0	11.4	89.2	0.0094	11.4	88.9	0.026	9.1
560	9.31	76.8	1.0	1.0	1.0	1.0	1.0	0.992	1.0	1.0	9.34	76.2	0.0094	9.33	76.5	0.026	7.35
665	1.85	38.3	1.0	1.0	1.0	1.0	0.99	1.0	1.0	1.0	1.83	38.4	0.0094	1.83	38.4	0.026	0.836
778.8	0.583	24.5	1.0	1.0	1.0	1.0	0.998	1.01	1.0	1.0	0.582	24.8	0.008	0.573	24.6	0.026	0.0
865	0.369	13.9	1.0	1.0	1.0	1.0	0.848	0.93	1.0	1.0	0.313	13.0	0.0072	0.315	13.3	0.026	0.0

	S_u	S_d	c_{cal}	c_{lin_u}	c_{lin_d}	c_{temp}	c_{stray_u}	c_{stray_d}	c_{pol_u}	c_{pol_d}	$L_{cal,u}$	$L_{cal,d}$	K(n)	$L_{OLCI,u}$	$L_{OLCI,d}$	rho	L_w
400	0.361	0.132	1.19	0.0171	0.0332	0.203	0.12	0.14	0.058	0.0577	1.26	1.22	2.3	1.21	1.22	1.8	1.28
442.5	0.246	0.162	0.784	0.133	0.0984	0.203	0.041	0.095	0.058	0.0577	0.857	0.838	1.5	0.816	0.82	1.3	0.935
490	0.144	0.183	0.757	0.292	0.238	0.203	0.0005	0.024	0.0582	0.115	0.854	0.851	1.2	0.839	0.825	0.69	0.886
560	0.0999	0.193	0.729	0.902	0.55	0.305	0.015	0.04	0.0579	0.115	1.2	0.983	0.91	1.19	0.958	0.6	1.4
665	0.466	0.207	0.721	0.957	0.701	0.711	0.051	0.014	0.116	0.231	1.5	1.27	0.8	1.41	1.23	2.5	3.89
778.8	0.982	0.228	0.728	0.878	0.613	1.32	0.011	0.071	0.115	0.233	2.01	1.65	0.74	1.74	1.62	2.5	n/a
865	1.37	0.566	1.34	0.508	0.334	2.74	0.89	0.38	0.115	0.231	3.52	3.17	0.75	3.24	3.11	2.5	n/a

b)

	S_u	S_d	c_{cal}	c_{lin_u}	c_{lin_d}	c_{temp}	c_{stray_u}	c_{stray_d}	c_{pol_u}	c_{pol_d}	$L_{cal,u}$	$L_{cal,d}$	K(n)	$L_{OLCI,u}$	$L_{OLCI,d}$	rho	L_w
400	7.06	51.9	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	7.06	51.9	0.0084	7.11	52.1	0.026	5.77
442.5	9.62	95.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	9.62	95.1	0.0095	9.63	93.4	0.026	7.22
490	11.4	89.6	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	11.4	89.7	0.0094	11.4	89.3	0.026	9.09
560	9.31	76.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	9.31	76.8	0.0094	9.3	77.2	0.026	7.31
665	1.85	38.3	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.84	38.3	0.0094	1.85	38.3	0.026	0.858
778.8	0.583	24.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0.583	24.4	0.008	0.575	24.3	0.026	0.0
865	0.369	13.9	1.0	1.0	1.0	1.0	0.999	1.0	1.0	1.0	0.369	13.9	0.0072	0.371	14.3	0.026	0.0

	S_u	S_d	c_{cal}	c_{lin_u}	c_{lin_d}	c_{temp}	c_{stray_u}	c_{stray_d}	c_{pol_u}	c_{pol_d}	$L_{cal,u}$	$L_{cal,d}$	K(n)	$L_{OLCI,u}$	$L_{OLCI,d}$	rho	L_w
400	0.358	0.132	1.19	0.0171	0.0332	0.203	2.4	2.8	0.0572	0.058	2.75	3.04	2.2	2.72	3.04	1.8	3.33
442.5	0.245	0.161	0.784	0.133	0.0984	0.203	0.83	1.9	0.0575	0.0575	1.19	2.11	1.5	1.16	2.1	1.3	1.58
490	0.147	0.185	0.757	0.292	0.238	0.203	0.01	0.48	0.0579	0.115	0.851	0.985	1.2	0.839	0.959	0.69	0.893
560	0.0979	0.196	0.729	0.902	0.55	0.305	0.3	0.8	0.0578	0.116	1.24	1.27	0.91	1.23	1.25	0.6	1.48
665	0.462	0.207	0.721	0.957	0.701	0.711	1.0	0.27	0.116	0.232	1.79	1.3	0.79	1.73	1.26	2.5	4.37
778.8	0.982	0.229	0.728	0.878	0.613	1.32	0.21	1.4	0.115	0.23	2.02	2.17	0.74	1.75	2.16	2.5	n/a
865	1.37	0.555	1.34	0.508	0.334	2.74	15.0	7.0	0.117	0.231	15.5	7.68	0.73	15.5	7.67	2.5	n/a

Figure 79. Water-leaving radiance MCM outputs. (a) Ideal case results. The mean (top panel) and standard uncertainty as a percentage of the mean (second panel) of each variable. This is shown for all OLCI bands of interest and relates to cast 1. The colour code intensity indicates the magnitude. (b) Non-Ideal case results. The mean (third panel) and standard uncertainty as a percentage of the mean (lower panel) of each variable.

The ideal scenario relates to the case in which stray light was well characterised and corrected for. For the non-ideal scenario stray light is not corrected for and was assigned a mean value of 1 (i.e., no correction applied) and an uncertainty relating to the previously used correction values. From Figure 79 it is clear that the mean of the stray light correction factors is 1 and the standard deviation is larger as compared to the ideal case. These changes can be seen to affect downstream parameters in the measurement equation.

The 778.8 nm and 865 nm bands resulted in a negative water-leaving radiance. A negative value here is not theoretically possible, so is likely due to an error in the measurement model. However, we do not expect to observe any water-leaving radiance measurable with the used radiometers at AAOT for these wavelengths. We quote the water-leaving radiance as 0 mW m^{-2} nm^{-1} sr^{-1} with no meaningful uncertainty.

The uncertainty for the non-ideal case is an under-representation of the true uncertainty that is applicable when corrections are not applied. This is because we have taken into account the uncertainty in the measurement but have not accounted for the lack of correction. Hence, we add the difference between the ideal mean value and the non-ideal mean value and add this to the standard deviation of the non-ideal case. This exhibits a much wider standard deviation, which is the true standard deviation, and which should be used but cannot be calculated without the ideal characterisation. Table 25 provides an overview of the values of interest in this correction. The mean values of the PDF function obtained in the MCM processing for water-leaving radiance are shown for the ideal and non-ideal case in radiance units (column 2 and 3 in the Table 25). Column 4 contains the bias values, thus the difference between the two means. The columns 5 to 7 present relative standard uncertainties for the three different scenarios.

Table 25. The mean and standard uncertainty as a percentage of the mean of water-leaving radiance L_w presented for the ideal and non-ideal cases.

	Mean L_W, mW m^{-2} nm^{-1}sr^{-1}			Standard uncertainty $u(L_W)$, %		
Band, nm	**Ideal**	**Non-ideal**	**Bias of L_W**	**Ideal**	**Non-ideal**	**Corr. non-ideal**
400	5.9	5.77	0.134	1.2	3.3	5.7
442.5	7.26	7.22	0.0353	0.94	1.6	2.1
490	9.1	9.09	0.0097	0.89	0.9	1.0
560	7.35	7.31	0.0442	1.4	1.4	2.1
665	0.84	0.86	0.0218	3.9	4.4	6.9

17.2.5 Conclusions

This part of FRM4SOC has demonstrated how to conduct an end-to-end uncertainty analysis for in situ radiometers of ocean colour measurements. An evaluation of an uncertainty budget for above-water OCR measurements is performed, which demonstrates the importance of correcting for instrumental biases. The data from one participant (Tartu Observatory, University of Tartu) were used for the study, as the radiometers, in addition to common radiometric calibration, had a set of additional optical characterisations completed. This enabled an investigation of the three scenarios: ideal, non-ideal and corrected non-ideal. In an ideal case, the measurand is corrected for the known instrumental biases. In a non-ideal case, the instrumental biases are not corrected for. The most important case is, however, the corrected non-ideal scenario where the real uncertainty related to the uncorrected bias has been evaluated. In Figure 73 and Figure **74**, the measurement equations are presented using uncertainty tree diagrams that in graphical form show all the relationships between different uncertainty contributors.

The required data for remote sensing reflectance include downward irradiance, downwelling radiance and upwelling radiance as well as all correction factors, the Fresnel reflectance of the water surface, and the fraction of diffuse to direct radiation at the time of measurement. The resultant outputs of the uncertainty analysis are therefore for the ideal and non-ideal cases, as well as a corrected case where an extra correction is applied to show the true resultant uncertainty when not corrected. The uncertainty in irradiance and radiance measurement was evaluated by assigning PDFs to each contributor and propagating this through the measurement equations using the MCM method. The MCM for downward irradiance and water-leaving radiance was run over two casts and results are presented for the seven OLCI bands of interest (400, 442.5, 490, 560, 665, 778.8, 865 nm) in Figure 78, Figure 79 and in Table 24, and Table 25. It should be noted that environmental uncertainty is not included and this may be the limiting factor since it is likely to be larger than the absolute calibration uncertainty. An evaluation of how to correctly estimate environmental uncertainty for particular conditions prevailing during measurements is yet to be completed.

The results of the different scenarios highlight the importance and benefits of carrying out instrument characterisations before campaigns and performing instrument corrections in addition to absolute radiometric calibration. It is recommended that the sources of uncertainty that are likely to dominate over the absolute calibration uncertainty (or other more dominant uncertainty contributors, which cannot be corrected for) should be characterised before campaigns so that these can be corrected for. The most likely parameters that will need prior characterisations are stray light, cosine, temperature and non-linearity corrections. Full details can be found in [225] and following these guidelines will support compliance with the FRM requirements of in situ ocean colour measurements for use in satellite product validation.

17.2.6 Ancillary data

Meteorological and oceanographic data must be collected for quality control purposes. This includes temperature, wind speed and direction, atmospheric pressure, wave height etc.

18 Conclusions and outcomes of the FRM4SOC project

18.1 FRM4SOC Final Workshop [D-260]

The FRM4SOC Final Workshop was held at NPL on 4 and 5 October 2018. International experts gathered at NPL to see the outcomes of the FRM4SOC project and to discuss the future needs of Ocean Colour Radiometry to support calibration and validation of the current and planned ocean colour missions. Representatives from many agencies including ESA, EUMETSAT, NASA, NOAA and CMEMS presented their requirements and approaches to marine satellite product validation.

The presentations on the project activities were followed by fruitful discussions on all aspects of FRM strategy. The consensus was that FRM4SOC activities are extremely useful and necessary for the community. They should include even more global cooperation for comparisons in the laboratory as well as in situ measurements. The community needs to quickly prepare for the validation of hyperspectral sensors like PACE, and thus suitable specifications for validation instruments is also urgently needed by the manufacturers. Moreover, it is clear that a greater understanding of uncertainty in the community is needed.

The agenda and all presentations of the workshop can be downloaded from the project website at *https://frm4soc.org*.

18.2 Proceedings of the FRM4SOC Final Workshop [D-270]

The proceedings of the workshop are available as a book of abstracts (Figure 80) [231] at the project website at *https://frm4soc.org* and as a special issue of the *MDPI Journal Remote Sensing "Fiducial Reference Measurements for Ocean Colour"* (ISSN 2072-4292). https://www.mdpi.com/journal/remotesensing/special_issues/2nd_ocean_color_RS (Figure 81).

The results from the FRM4SOC project are published in the following papers of the special issue.

1. Banks, A.C.; Vendt, R.; Alikas, K.; Bialek, A.; Kuusk, J.; Lerebourg, C.; Ruddick, K.; Tilstone, G.; Vabson, V.; Donlon, C.; Casal, T. Fiducial "Reference Measurements for Satellite Ocean Colour (FRM4SOC)". Remote Sens. 2020, 12, 1322. [232]

2. Ruddick, K.G.; Voss, K.; Banks, A.C.; Boss, E.; Castagna, A.; Frouin, R.; Hieronymi, M.; Jamet, C.; Johnson, B.C.; Kuusk, J.; Lee, Z.; Ondrusek, M.; Vabson, V.; Vendt, R. "A Review of Protocols for Fiducial Reference Measurements of Downwelling Irradiance for the Validation of Satellite Remote Sensing Data over Water." Remote Sens. 2019, 11, 1742. [82]

3. Ruddick, K.G.; Voss, K.; Boss, E.; Castagna, A.; Frouin, R.; Gilerson, A.; Hieronymi, M.; Johnson, B.C.; Kuusk, J.; Lee, Z.; Ondrusek, M.; Vabson, V.; Vendt, R. "A Review of Protocols for Fiducial Reference Measurements of Water-Leaving Radiance for Validation of Satellite Remote-Sensing Data over Water." Remote Sens. 2019, 11, 2198. [83]

4. Vabson, V.; Kuusk, J.; Ansko, I.; Vendt, R.; Alikas, K.; Ruddick, K.; Ansper, A.; Bresciani, M.; Burmester, H.; Costa, M.; D'Alimonte, D.; Dall'Olmo, G.; Damiri, B.; Dinter, T.; Giardino, C.; Kangro, K.; Ligi, M.; Paavel, B.; Tilstone, G.; Van Dommelen, R.; Wiegmann, S.; Bracher, A.; Donlon, C.; Casal, T. "Laboratory Intercomparison of

Radiometers Used for Satellite Validation in the 400–900 nm Range." Remote Sens. 2019, 11, 1101. [167]

5. Vabson, V.; Kuusk, J.; Ansko, I.; Vendt, R.; Alikas, K.; Ruddick, K.; Ansper, A.; Bresciani, M.; Burmester, H.; Costa, M.; D'Alimonte, D.; Dall'Olmo, G.; Damiri, B.; Dinter, T.; Giardino, C.; Kangro, K.; Ligi, M.; Paavel, B.; Tilstone, G.; Van Dommelen, R.; Wiegmann, S.; Bracher, A.; Donlon, C.; Casal, T. "Field Intercomparison of Radiometers Used for Satellite Validation in the 400–900 nm Range." Remote Sens. 2019, 11, 1129. [168]

6. Bialek, A.; Douglas, S.; Kuusk, J.; Ansko, I.; Vabson, V.; Vendt, R.; Casal, "T. Example of Monte Carlo Method Uncertainty Evaluation for Above-Water Ocean Colour Radiometry." Remote Sens. 2020, 12, 780. [223]

7. Bialek, A.; Goodman, T.; Woolliams, E.; Brachmann, J.F.S.; Schwarzmaier, T.; Kuusk, J.; Ansko, I.; Vabson, V.; Lau, I.C.; MacLellan, C.; Marty, S.; Ondrusek, M.; Servantes, W.; Taylor, S.; Van Dommelen, R.; Barnard, A.; Vellucci, V.; Banks, A.C.; Fox, N.; Vendt, R.; Donlon, C.; Casal, T. Results from Verification of Reference Irradiance and Radiance Sources Laboratory Calibration Experiment Campaign. Remote Sens. 2020, 12, 2220. [153]

8. Alikas, K.; Vabson, V.; Ansko, I.; Tilstone, G.H.; Dall'Olmo, G.; Nencioli, F.; Vendt, R.; Donlon, C.; Casal, T. Comparison of Above-Water Seabird and TriOS Radiometers along an Atlantic Meridional Transect. Remote Sens. 2020, 12, 1669. [212]

9. Tilstone, G.; Dall'Olmo, G.; Hieronymi, M.; Ruddick, K.; Beck, M.; Ligi, M.; Costa, M.; D'Alimonte, D.; Vellucci, V.; Vansteenwegen, D.; Bracher, A.; Wiegmann, S.; Kuusk, J.; Vabson, V.; Ansko, I.; Vendt, R.; Donlon, C.; Casal, T. Field Intercomparison of Radiometer Measurements for Ocean Colour Validation. Remote Sens. 2020, 12, 1587. [109]

Figure 80. Proceedings of the FRM4SOC Final Workshop [187].

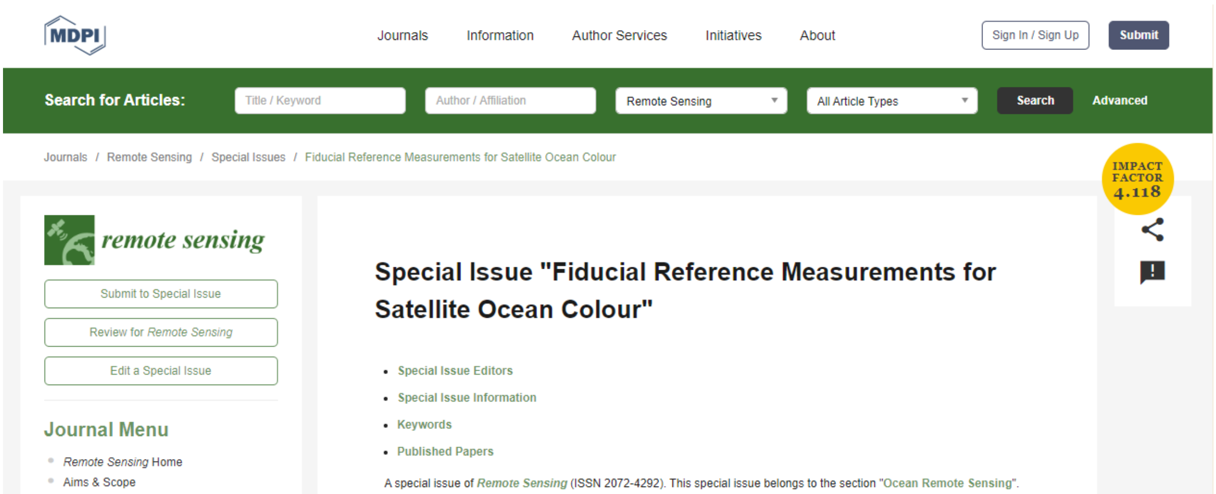


Figure 81. Outcomes of the FRM4SOC are published in the special issue of the *MDPI Journal Remote Sensing “Fiducial Reference Measurements for Ocean Colour”* (ISSN 2072-4292)

18.3 FRM4SOC Scientific and Operational Roadmap [D-280] [30]

The work and results of FRM4SOC have had (and continue to have) a significant impact on the earth observation and ocean colour community. In particular FRM4SOC played a prominent role in three Sentinel-3 validation team meetings at EUMETSAT and ESRIN/ESA [233–235], and the FRM4SOC international workshop report [D-240], [31] on ocean colour system vicarious calibration (OC-SVC) is being used as one of the main requirements reference documents for the future of Copernicus OC-SVC infrastructure. Considering that this continued effort is in support of ensuring high quality and accuracy Copernicus satellite mission data, in particular Sentinel-2 MSI and Sentinel-3 OLCI ocean colour products, and contributes directly to the work of ESA and EUMETSAT to ensure that these instruments are validated in orbit, FRM4SOC produced a scientific road map for the FRM based future of satellite ocean colour validation and vicarious calibration.

The FRM4SOC Scientific and Operational Roadmap (SOR) [30] provides a critical analysis of all the feedback from participants and institutions working in the project; identifies potential strategies for integrating the project outcomes into existing initiatives and operational institutions; defines a plan for fostering a transition of FRM4SOC outcomes from research to operational activities; and identifies priority areas to be addressed in potential future projects in support of OCR calibration and validation activities.

The document lists 24 main conclusions (C) and recommended actions (A) summarised from the reports and series of discussions on several events in the framework of the FRM4SOC project. A strategy and an implementation plan are proposed to integrate the FRM4SOC outcomes with existing initiatives, operational institutions and activities. [30]

The SOR also identifies the **priority areas** to be addressed in implementation and further development of FRM for satellite ocean colour.

1. Updating measurement protocols and uncertainty budgets
2. Development of a community processor for data handling and uncertainty evaluation
3. Providing examples and short practical guides for uncertainty evaluation
4. Training on implementation of measurement protocols and end-to-end uncertainty evaluation
5. Establishment of required specifications of FRM OCR
6. General plan for calibration and characterisation of OC radiometers
7. Development of metrology infrastructure for calibration and characterisation of new generation OC radiometers
8. Periodic calibration and characterisation of OC radiometers
9. Describing a global comparison strategy for FRM measurements
10. Organising periodic comparison measurements on all levels of the traceability chain.
11. Development of ocean colour system vicarious calibration infrastructure.

CONCLUSIONS (C) FROM THE FRM4SOC PROJECT AND THEIR RECOMMENDED ACTIONS (A)

1. IMPLEMENTING FRM

C1 Measurement results collected for EO data validation shall have metrological traceability to the units of SI with related uncertainty evaluation.

C2 Space agencies should: *i*) in the medium term, encourage and stimulate the adoption of FRM requirements, and *ii*) in the long term, when sufficient progress and consensus is achieved, use only FRM for the routine validation of satellite ocean colour data. In the near term, use of non-FRM quality data for satellite calibration or validation should only be done with great care.

C3 Space agencies and National Metrology Institutes should consider increased collaboration in order to harmonise approaches, methodologies and implement the principles of FRM worldwide.

C4 Financial support from ESA and other space agencies or entities shall be ensured for implementing the principles of FRM.

A1 International communication and agreement on establishment and implementing FRM requirements shall be encouraged.

CONCLUSIONS (C) FROM THE FRM4SOC PROJECT AND THEIR RECOMMENDED ACTIONS (A)

2. METHODS, PROTOCOLS, PROCEDURES & UNCERTAINTY BUDGETS

C5 International worldwide cooperation on all levels (e.g. agencies, research institutes, experts, etc.) is imperative in order to ensure high quality global climate data. Different protocols existing for OCR data validation all over the world shall be harmonised, understood and applied in a consistent manner to ensure global uniformity of measurements.

C6 Data (including appropriate metadata) and expertise collected over years by the international community shall be acknowledged, preserved and passed on to the next generations.

C7 Principles of good practice in performing measurements shall be documented and their application encouraged.

C8 Practical consolidated examples on compiling uncertainty budgets shall be provided.

C9 Established methods, principles of good practice, and uncertainty budgets shall be validated in comparison measurements.

C10 Definition, adoption and validation of the principles of good practice and uncertainty budgets shall be supported with appropriate funding from ESA and other space agencies or entities.

A2 International co-operation on all levels to:

a. document measurement protocols;
b. agree and establish principles of good practice in performing measurements;
c. identify, harmonise, and establish requirements for measurement and correction of gain and assess its uncertainty gained measurement uncertainty levels;
d. provide consolidated examples on compiling uncertainty budgets
e. provide training on good practice and building uncertainty budgets.

A3 Ensure appropriate funding to define, adopt and validate the principles of good practice and uncertainty budgets.

CONCLUSIONS (C) FROM THE FRM4SOC PROJECT AND THEIR RECOMMENDED ACTIONS (A)

3. PROPERTIES OF OCR

C11 Properties of OC radiometers must reflect the needed accuracy for satellite OCR data validation and correspond to requirements as identified and established by the international community in the field. Community consensus on practically feasible requirements is needed. However, the principles of metrology - SI traceability and acceptable uncertainty limits - must be followed.

C12 A document, setting minimum requirements for the most important properties of radiometric instruments used for satellite OCR validation, is needed. Preparation of such a document should be encouraged and funded by ESA and other space agencies or entities.

C13 Vital components and specifications for new generation (e.g. hyperspectral) instruments shall be identified and characterisation capabilities of required metrology infrastructure shall be developed accordingly.

C14 ESA and other space agencies or entities should encourage further development of OCR instruments, including a requirement that such developments provide FRM-compatible information on radiometer characterisation.

C15 Characterisation and regular calibration of OCR is needed in order to ensure traceability to the units of SI and evaluate the instrument related uncertainty contributions.

C16 ESA and other space agencies or entities should fund and encourage activities to test radiometers from all manufacturers according to a standardised methodology.

A4 Identify and document requirements and expected specifications (e.g. measurement range, maximum permissible errors, uncertainties, etc.) for Ocean Colour Radiometry (OCR) instruments to meet the requirements for validation of mission data (A2. c.)

A5 Identify, document, map existing and develop missing metrology infrastructure and its capabilities required for calibration and characterisation of OCR (incl. new generation e.g. hyperspectral) instruments.

A6 Identify, document and implement a recommended (standardised) plan for initial and periodic calibration and characterisation of OCR instruments.

A7 Establishment of regional reference laboratories for calibration and characterisation of OCR.

A8 Ensure appropriate funding to identify and document requirements for specifications of OCR instruments and their calibration and characterisation.

CONCLUSIONS (C) FROM THE FRM4SOC PROJECT AND THEIR RECOMMENDED ACTIONS (A)

4. COMPARISON EXPERIMENTS & DATABASE OF OCR FIELD RADIOMETER PERFORMANCE

C17 Periodic comparison experiments are needed for validation of established methods and uncertainty budgets on all levels of the traceability chain.

C18 Comparison experiments also serve the purpose of training, sharing experience, and support achievement of common understanding and interpretation of the measurement protocols.

C19 Application of unified data handling or a community processor will reduce overall uncertainty and improve agreement between individual datasets, although care not to limit innovation must be ensured.

C20 Worldwide international participation of agencies and research organisations in comparison exercises shall be aimed for.

C21 ESA and other space agencies or entities shall encourage and support implementing of comparison experiments with appropriate funding.

A9 Organise periodic comparison experiments on all levels of the traceability chain:

a. reference standards (NMI and OCR calibration laboratory level);
b. calibration and characterisation methods of OCR (calibration laboratory level);
c. in situ field measurements:
   - understanding, interpretation, and following established protocols;
   - competence and experience of personnel (all levels).

A10 Development and application of unified data handling/ community processor.

A11 Ensure appropriate funding to organise comparison experiments for validation of established methods and uncertainty budgets on all levels of the traceability chain.

CONCLUSIONS (C) FROM THE FRM4SOC PROJECT AND THEIR RECOMMENDED ACTIONS (A)

5. OPTIONS FOR LONG-TERM FUTURE EUROPEAN SATELLITE OCR VICARIOUS ADJUSTMENT

C22 Operational FRM infrastructures to underpin SVC with SI traceability, full uncertainty characterisation and the best possible accuracy and precision are mandatory. Such FRM infrastructure of the quality needed for SVC shall be redundant in order to ensure steady and sufficient data provision.

C23 BOUSSOLE as the existing unique SVC site in Europe must be maintained in the long term and upgraded to full operational status.

C24 Development and long term operation of a second new European infrastructure for OC-SVC in a suitable location to gain ideal SVC conditions and ensure operational redundancy is needed.

A12 Upgrade BOUSSOLE to fully operational status.

A13 Develop a new infrastructure based on MOBY-Net and/or new European technology in a suitable location, e.g. the Eastern Mediterranean near Crete.

A14 Involvement of National Metrological Institutes (NMIs) at all stages of development of an SVC infrastructure.

A15 Train a new group to operate a second SVC.

A16 Support long-term interaction of the different SVC operations groups.

A17 Support scientific and research activities on SVC sites.

A18 Ensure long-term investments for both SVC sites.

19 Communication, Outreach and Promotion

19.1 FRM4SOC web portal [D-10]

The FRM4SOC project web site at https://frm4soc.org was developed and operated in order to provide a public 'communications and study management' portal during the course of the project (Figure 82).

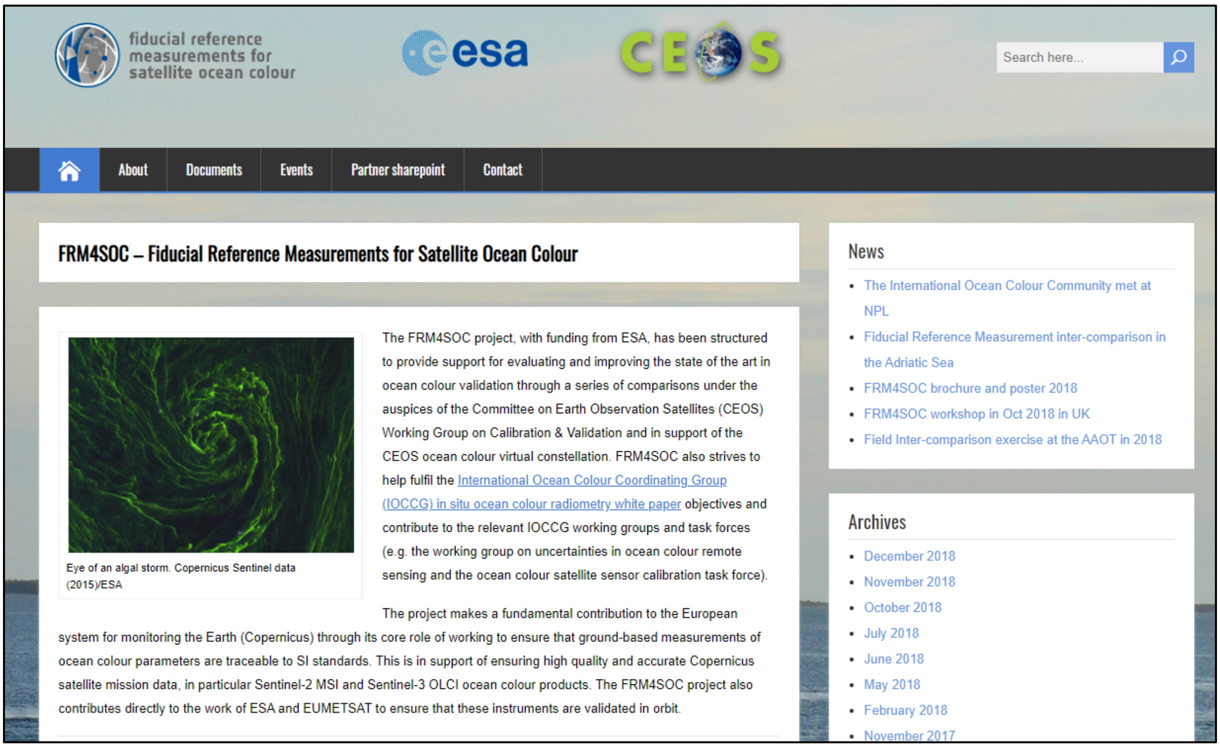


Figure 82. Front page of the FRM4SOC website at *https://frm4soc.org*.

The web portal includes the following features as required by the SOW [3]:

i) Designed to follow the ESA SPPA website template (https://earth.esa.int/web/sppa/home) in order to ensure that the content of the FRM4SOC web page includes all elements of the SPPA web site.

ii) ESA, FRM4SOC CEOS and partner logos.

iii) Description of the FRM4SOC project based on the SOW and contractor proposal.

iv) A Gantt chart and a public calendar for all project activities, meetings and events.

v) A public list of project deliverables.

vi) A project document library with on-line access to download project documents, reports, data, and presentations with cross-references to the SOW and contract deliverables.

vii) Indexed access to reference documents used by the project and a set of relevant links to the project and other useful resources.

viii) A secured password protected area where project management documents can be accessed.

Means for public users to provide feedback and comments to the project team were granted by published contact data on the project website and using social media tools such as Twitter, Facebook (Figure 88).

The website was updated along with the progress of the project with documents, presentations and news headlines on a monthly basis.

19.2 Project brochures [D-20] and [D-30]

Two glossy (4-8 pages) promotional brochures describing the activities of the FRM4SOC project were published (Figure 83).

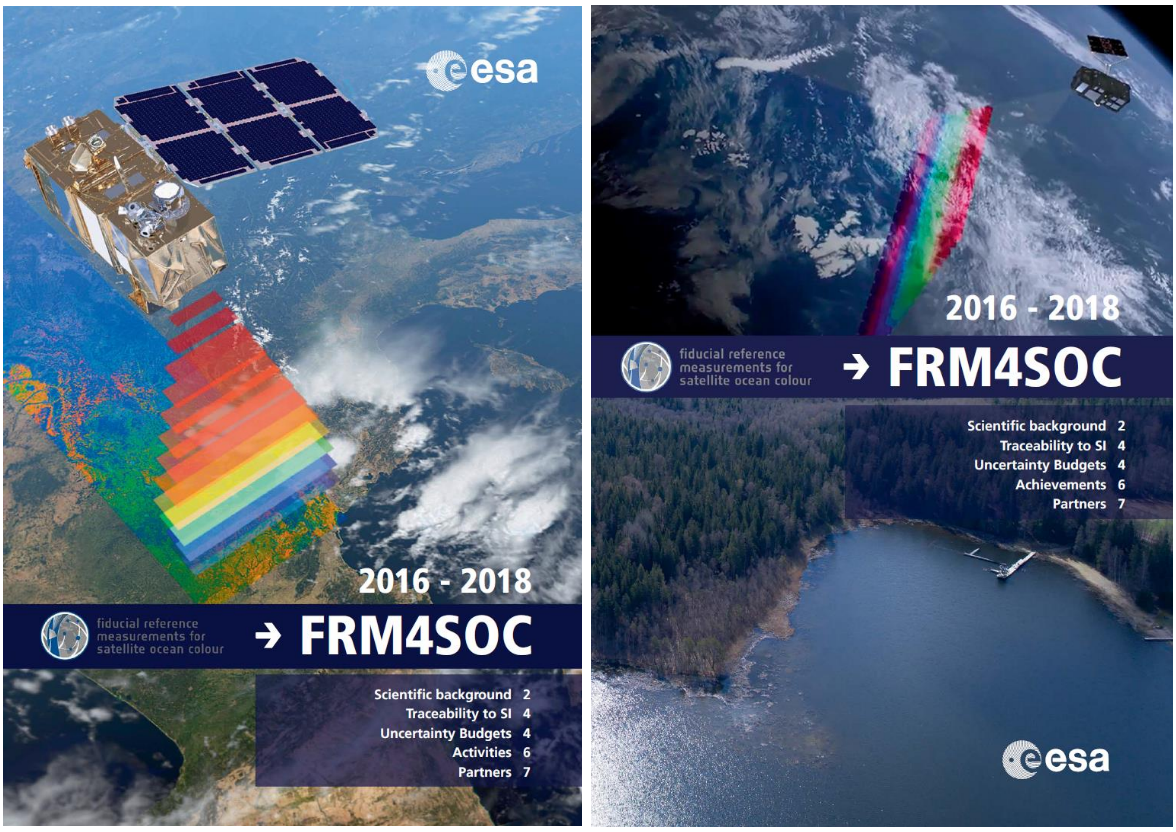


Figure 83. Title pages of the FRM4SOC brochures BRO-1 (left) and BRO-2 (right).

Both brochures introduce the scientific background of the project as well as need for action. Principles of metrological traceability to the units of SI with related uncertainty evaluation are explained. The first brochure (BRO-1) focuses on advertising the events and comparisons in the project while the second one (BRO-2) lists and provides reference to the achievements and results.

A number of 200 printed copies of each brochure were distributed at several events such as FRM4SOC workshops, Sentinel Validation Team (SVT) and other organisational meetings, conferences and dedicated visits. Both brochures are also shared online via the project website and social media channels.

19.3 High quality graphics [D-40]

A repository of high quality graphics and images to be used for promotion of the principles of FRM, project activities and results is stored at the password-protected area of the *https://frm4soc.org* website (Figure 84). The repository holds photos from the project events and activities, several graphs, figures and other illustrative materials.

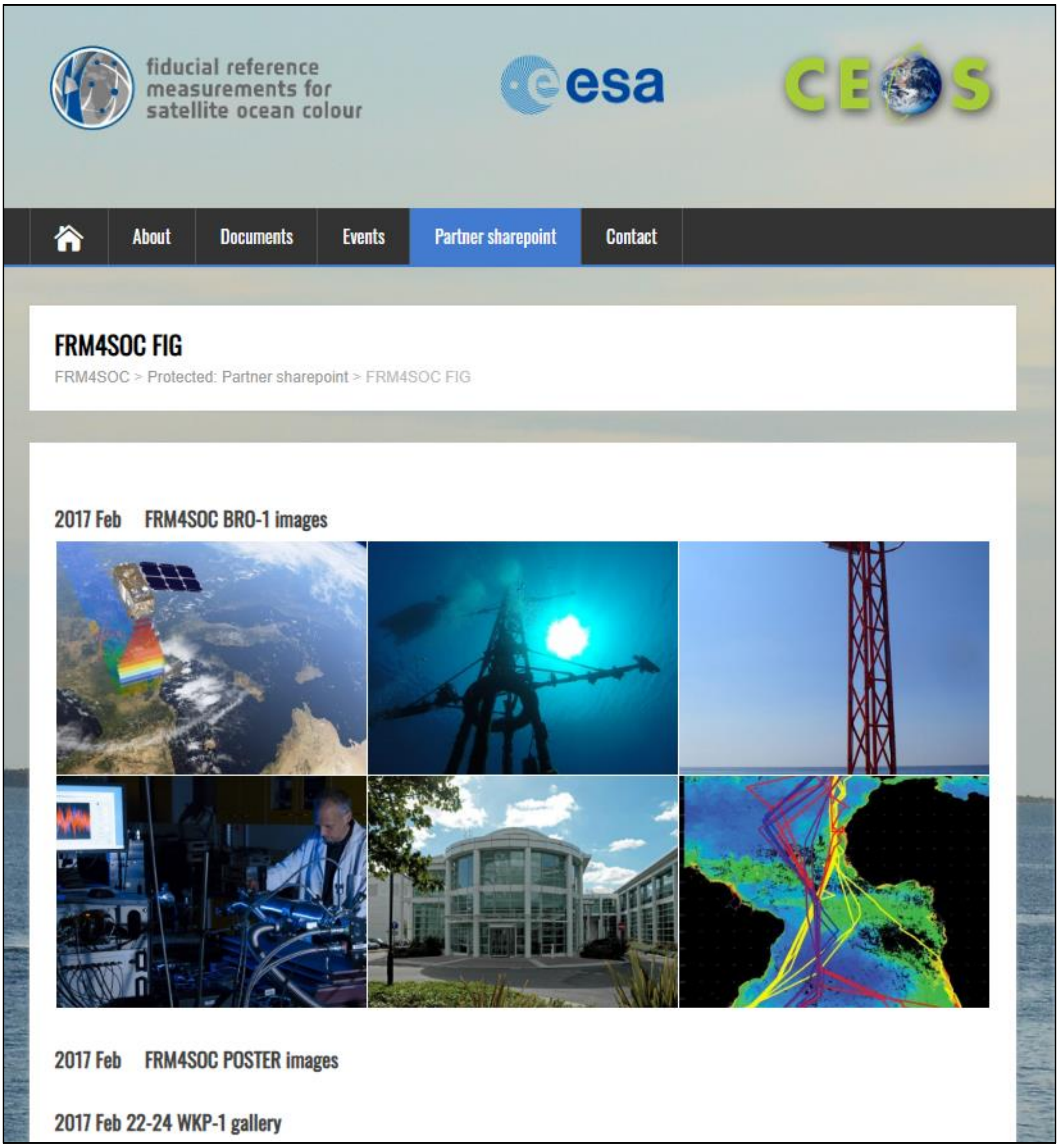


Figure 84. A repository of high quality graphics and images [D-40] in the password-protected area of the *https://frm4soc.org* website.

19.4 Web stories for the FRM4SOC website describing the activities of the FRM4SOC project [D-50]

All together 11 web stories describing activities of the FRM4SOC activities were published on the frm4soc.org website (Figure 85):

1. "Monitoring Ocean Change using Copernicus satellites" by Gavin Tilstone, PML, January 2017;
2. "World class experts gathered in ESRIN to make the best of Ocean Colour data from Sentinel era" by Christophe Lerebourg, ACRI-ST, March 2017;
3. "The Australian Integrated Marine Observing System's Radiometry Task Team", by David Antoine, Curtin University, April 2017;
4. "Calibration sources for satellite ocean colour radiometers successfully compared at NPL, UK (FRM4SOC LCE-1)" by Andrew Banks, NPL, April 2017;
5. "International SI traceable comparison exercise to verify the performance of Field Ocean Colour Radiometers (FRM4SOC-LCE2)" by Mari Allik, Joel Kuusk, Tiia Lillemaa and Riho Vendt, Tartu Observatory, May 2017;
6. "A Basin Scale Inter-comparison of Fiducial Reference Measurements" by Gavin Tilstone, PML, October 2017;
7. "Remote sensing of inland waters is a challenging task" by Krista Alikas, Tartu Observatory, November 2017;
8. "AMT4SENTINELFRM: The second voyage" by Gavin Tilstone, PML, November 2017;
9. "A Review of Commonly used Ocean Colour Radiometers (OCR) used for Satellite OCR Validation" by Riho Vendt, Tiia Lillemaa, Tartu Observatory of University of Tartu, and Kevin Ruddick, RBINS, May 2018;
10. "Fiducial Reference Measurement inter-comparison in the Adriatic Sea" by Gavin Tilstone, PML, October 2018,
11. "The International Ocean Colour Community met at NPL" by Agnieszka Bialek, NPL, October 2018.

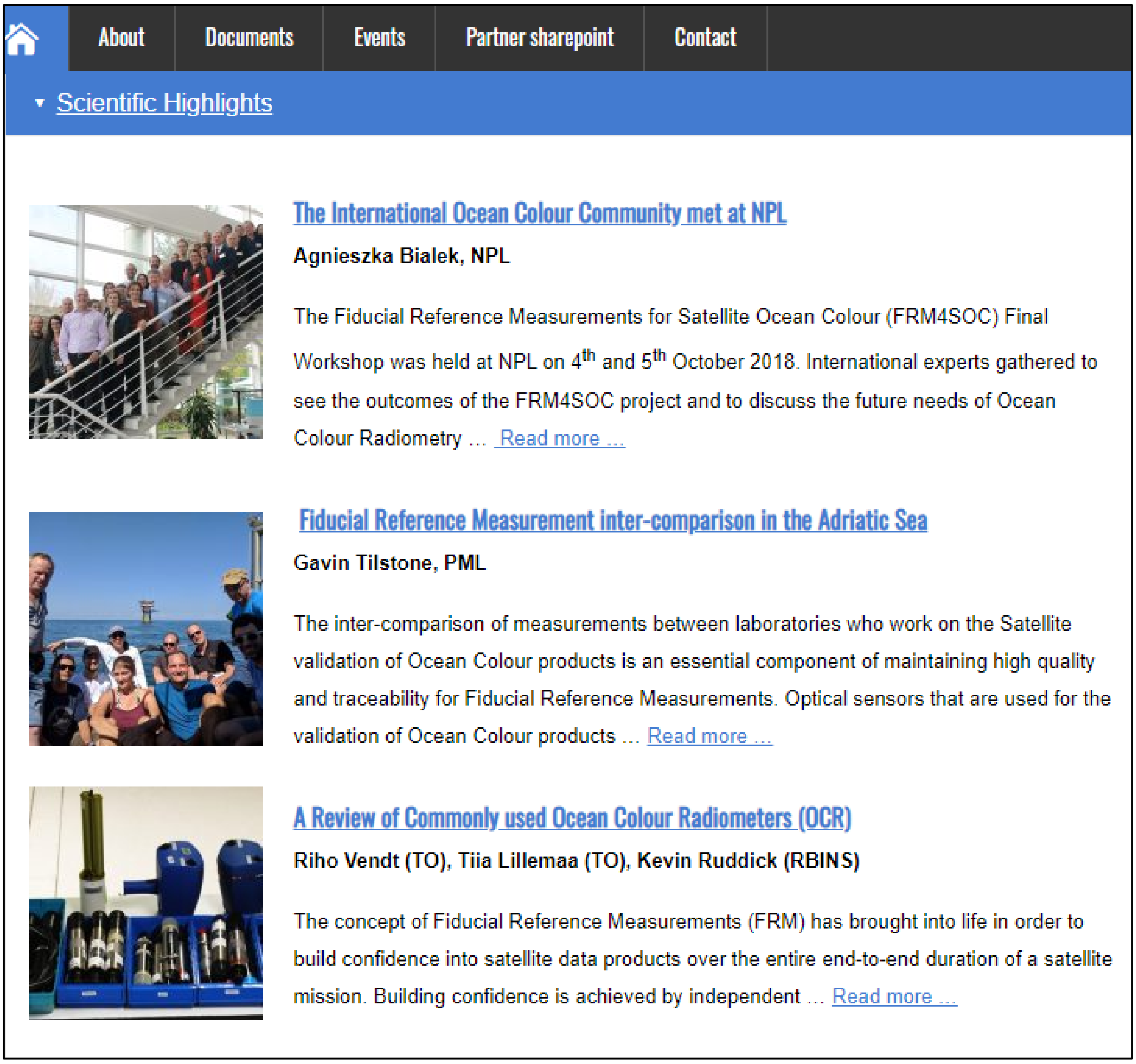


Figure 85. Eleven web stories describing the project activities were published at the *https://frm4soc.org* website.

19.5 Other outreach activities.

The principles of FRM, project activities and results have been actively promoted to scientific and user communities via several communication channels such as presentations on S3VT, IOCCG, IOCCG IOCS, CEOS-WGCV-IVOS and other meetings. Presentations on project workshops (e.g. AMT4SentinelFRM, FRM4ALT) and conferences (Ocean Optics, ESA Living Planet Symposium) were given.

Two different posters (Figure 86) were presented at several meetings and conferences. The first poster (left in Figure 86) focused on describing the scientific background of the project explaining the principles of metrological traceability to the units of SI with related uncertainty evaluation while the second one (right in Figure 86) provided reference to the achievements and results of the project.

Figure 86. Two posters on the FRM4SOC project activities and results were presented at several meetings and conferences.

Several leaflets (Figure 87) to promote the events of the project were distributed to potential participants.

Several manufacturers of ocean colour radiometers in Europe (CIMEL, TriOS, Water Insight) were visited. Other non-European manufacturers were contacted by WebEx teleconference and the first round table seminar of the manufacturers of OCR was held at ESTEC on 6th September 2017.

Regular news on the project events and achievements were posted on social media channels such as Facebook and Twitter (Figure 88).

Figure 87. FRM4SOC leaflet to promote the events of the project to potential participants.

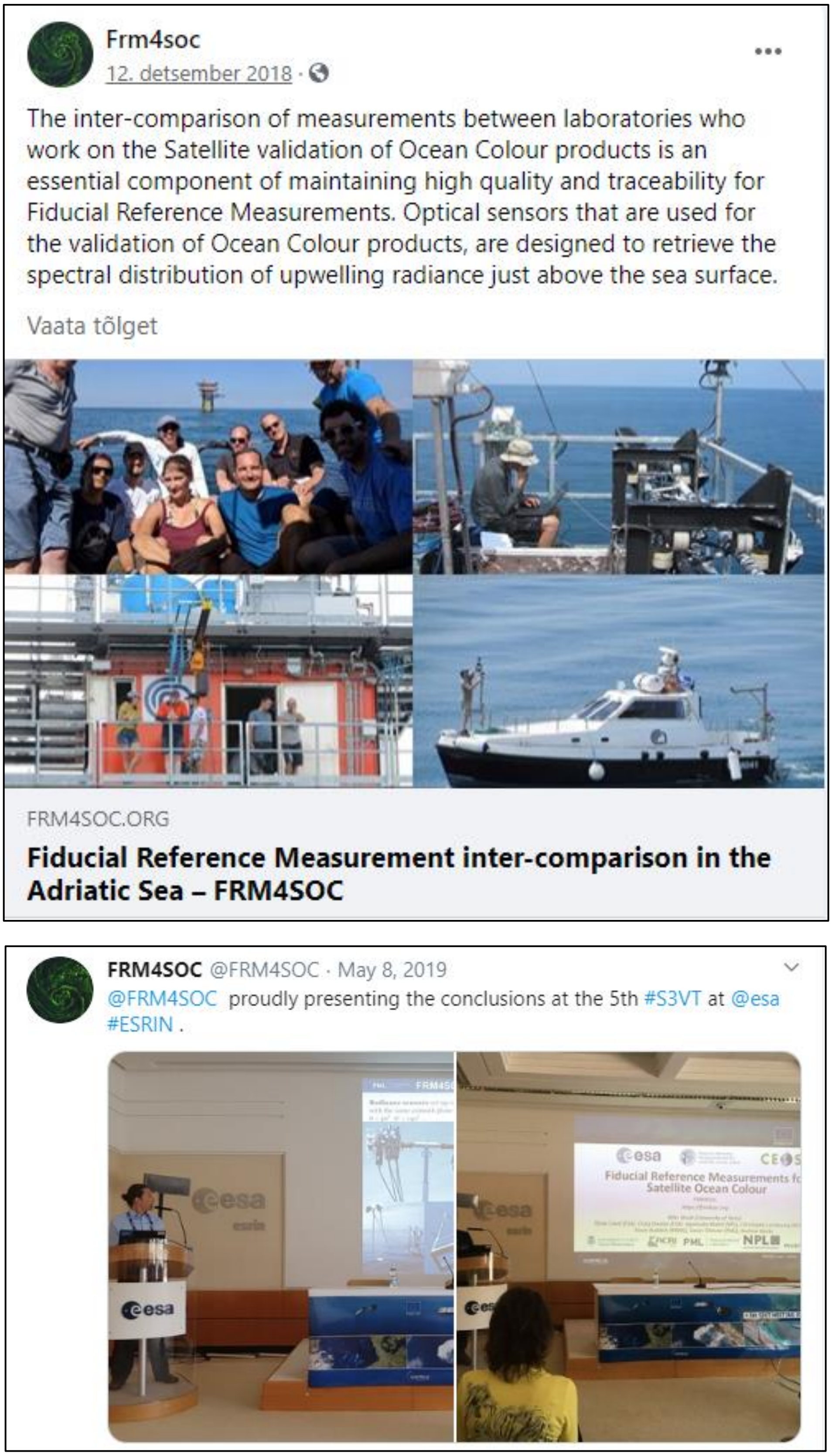


Figure 88. Examples of the FRM4SOC newsfeed on Facebook (top) and Twitter (below).

20 Reporting

The **"Project Management Plan and Schedule" [D-310]**, **"Executive monthly progress report and actions database" [D-320]** as well as minutes of meetings were submitted and updated in due course.

The **FRM4SOC Final Report (this document) [D-290]** highlights all of the activities conducted during the project (with reference to the deliverables of the contract) and the results obtained.

The **FRM4SOC Technical Data package TDP [D-300]** has been compiled from the final versions of all approved technical documents.

The **"Contract Closure Summary" [D-330]** is submitted according to the format as provided by the Appendix A of the FRM4SOC Contract between ESA and the Main Contractor at the end of the contract.

All reports and minutes of meetings are stored in the password protected area of the *https://frm4soc.org* portal.

21 Deliverables

Table 26. List of the FRM4SOC deliverables.

ID	Code	Title	Section
D-10	WWW	FRM4SOC web portal to be operated and updated for the duration of the Contract	19.1
D-20	BRO-1	Initial FRM4SOC Project Brochure	19.2
D-30	BRO-2	Final FRM4SOC Project Brochure	19.2
D-40	FIG	High quality graphics (FIG) that can be used by the FRM4SOC project and ESA to promote the outcomes of the project throughout the project.	19.3
D-50	WEBS	Web Stories for the FRM4SOC web site describing the interesting and innovative activities of the FRM4SOC project. (2 per year)	19.4
D-60	TR-1	Technical Report: "Measurement Requirements and Protocols when Operating Fiducial Reference Measurement (FRM) Ocean Colour Radiometers (OCR) for Satellite Validation"	12
D-70	TR-2	Technical Report: "A Review of Commonly used Fiducial Reference Measurement (FRM) Ocean Colour Radiometers (OCR) used for Satellite OCR Validation"	13
D-80	TR-3	Technical Report: "Protocols and Procedures to Verify the Performance of Reference Irradiance Sources used by Fiducial Reference Measurement Ocean Colour Radiometers for Satellite Validation"	14.2
D-90	LCE-1-IP	LCE-1 implementation plan	14.3
D-100	LCE-1	Following TR-3, implement a round robin LCE to verify the performance of reference radiance and irradiance sources (ie. lamps, plaques etc.) used to maintain the calibration of FRM OCR radiometers traceable to SI	14.3
D-110	LCE-1-DATA	Data package containing all data collected during LCE-1.	14.5
D-120	TR-4	Technical Report: "Results from the First FRM4SOC Reference Radiance and irradiance Source Verification Laboratory Calibration Experiment Campaign"	14.4
D-130	TR-5	Technical Report: "Protocols and Procedures to Verify the Performance of Fiducial Reference Measurement (FRM) Field Ocean Colour Radiometers (OCR) used for Satellite Validation"	15.2
D-140	LCE-2-IP	LCE-2 implementation plan	15.3
D-150	LCE-2	Following TR-5 implement a round robin LCE campaign to verify the performance of FRM OCR	15.3

ID	Code	Title	Section
		used for Satellite Validation traceable to SI	
D-160	LCE-2-DATA	Data package containing all data collected during LCE-2	15.5
D-170	TR-6	Technical Report "Results from the First FRM4SOC Field Ocean Colour Radiometer Verification Round Robin Campaign"	15.4
D-180	TR-7	Technical Report "Uncertainty Budgets of FRM4SOC Fiducial Reference Measurement (FRM) Ocean Colour Radiometer (OCR) systems used to Validate Satellite OCR products"	17
D-190	TR-8	Technical Report "Protocols and Procedures for Field Inter-Comparisons of Fiducial Reference Measurement (FRM) Field Ocean Colour Radiometers (OCR) used for Satellite Validation"	16.2
D-200	FICE-IP	Implementation plan for the FRM4SOC field inter-comparison experiments (FICE)	16.3
D-210	FICE-DB	Field inter-comparison experiment database (FICE-DB)	16.6
D-220	TR-9	Technical Report "Results from the First FRM4SOC Field Inter-Comparison Experiment (FICE) of Ocean Colour Radiometers"	16.4 16.5
D-230	WKP-1	International workshop "Options future infrastructure required for the long-term vicarious adjustment of the Sentinel-3 OLCI and Sentinel-2 MSI A/B/C and D instruments"	11.1
D-240	PROC-1	Proceedings of the workshop WKP-1	11.2
D-250	TR-10	Technical Report "Requirements and recommendations for infrastructure required for the long-term vicarious adjustment of the Sentinel-3 OLCI and Sentinel-2 MSI A/B/C and D instruments"	11.3
D-260	WKP-2	FRM4SOC Final Workshop	18.1
D-270	PROC-2	FRM4SOC workshop proceedings	18.2
D-280	SOR	FRM4SOC Scientific and Operational Roadmap	18.3
D-290	FR	FRM4SOC Final Report	Present document.
D-300	TDP	FRM4SOC Technical Data Package	20
D-310	PMP	Project Management Plan	20
D-320	MR	Executive monthly progress report and actions database	20
D-330	CCS	Contract Closure Summary	20

22 References

1. IOCCG, "International Network for Sensor Inter-comparison and Uncertainty assessment for Ocean Color Radiometry (INSITU-OCR)," (2012).
2. BIPM, "The International System of Units (SI), 9th edition," (2019).
3. European Space Agency (ESA), "Fiducial Reference Measurements for Satellite Ocean Colour (FRM4SOC), Statement of work," (2015).
4. UN General Assembly, "A/RES/70/1 Transforming our world: the 2030 Agenda for Sustainable Development," (2015).
5. GEO, "GEO Strategic Plan 2016-2025: Implementing GEOSS," (2015).
6. J. Aschbacher, "ESA's Earth Observation Strategy and Copernicus," in *Satellite Earth Observations and Their Impact on Society and Policy*, M. Onoda and O. R. Young, eds. (Springer Singapore, 2017), pp. 81–86.
7. "Copernicus programme," https://www.copernicus.eu.
8. R. Hollmann, C. J. Merchant, R. Saunders, C. Downy, M. Buchwitz, A. Cazenave, E. Chuvieco, P. Defourny, G. de Leeuw, R. Forsberg, T. Holzer-Popp, F. Paul, S. Sandven, S. Sathyendranath, M. van Roozendael, and W. Wagner, "The ESA Climate Change Initiative: Satellite Data Records for Essential Climate Variables," Bull. Am. Meteorol. Soc. **94**, 1541–1552 (2013).
9. C. J. Donlon and G. Zibordi, "Chapter 3 - In Situ Optical Radiometry," in *Experimental Methods in the Physical Sciences*, G. Zibordi, C. J. Donlon, and A. C. Parr, eds., Optical Radiometry for Ocean Climate Measurements (Academic Press, 2014), Vol. 47, pp. 245–246.
10. QA4EO, "A Quality Assurance Framework for Earth Observation: Principles," (2010).
11. GEO, "GEO Work Plan 2007 - 2009. Toward Convergence.," (2008).
12. GEO, "GEO Work Plan 2009 - 2011," (2008).
13. GEO, "GEO Work Plan 2012 - 2015," (2014).
14. CEOS, "CEOS 2018 - 2020 Work plan," (2018).
15. BIPM, "Comptes Rendus de la 20e reunion de la Conference Generale des Poids et Measures," (1996).
16. G. Ohring, B. Wielicki, R. Spencer, B. Emery, and R. Datla, "Satellite Instrument Calibration for Measuring Global Climate Change: Report of a Workshop," Bull. Am. Meteorol. Soc. **86**, 1303–1314 (2005).
17. "JCGM 100, Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement (GUM), First Edition, September 2008. Available online: http://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf," (2008).
18. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, and OIML, *Guide to the Expression of Uncertainty in Measurement*, 2nd ed. (International Organisation for Standardisation, 1995).
19. JCGM/WG2, *International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM)* (JCGM, 2008).
20. World Meteorological Organization (WMO), "Systematic Observation Requirements for Satellite-based Products for Climate Supplemental details to the satellite-based component of the Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC: 2011 update, GCOS-154," (2011).
21. "BIPM - Bureau International des Poids et Mesures," https://www.bipm.org.
22. N. Fox, "Community Workshop," presented at FRM4SOC Final Workshop, NPL, Teddington, UK (2018).
23. G. Zibordi and C. J. Donlon, "Chapter 5 - In Situ Measurement Strategies," in *Experimental Methods in the Physical Sciences*, G. Zibordi, C. J. Donlon, and A. C. Parr, eds., Optical Radiometry for Ocean Climate Measurements (Academic Press, 2014), Vol. 47, pp. 527–529.

24. ESA, "ESA Sensor Performance, Products and Algorithms (SPPA)," https://earth.esa.int/web/sppa/activities.
25. C. J. Donlon, W. Wimmer, I. Robinson, G. Fisher, M. Ferlet, T. Nightingale, and B. Bras, "A Second-Generation Blackbody System for the Calibration and Verification of Seagoing Infrared Radiometers," J. Atmospheric Ocean. Technol. **31**, 1104–1127 (2014).
26. "FRM4SOC," https://frm4soc.org/.
27. IOCCG, *Report, "Why Ocean Colour? The Societal Benefits of Ocean Colour Radiometry* (2008).
28. IOCCG, "International Ocean Colour Coordinating Group," http://ioccg.org/.
29. CEOS, "Ocean Colour Radiometry," http://ceos.org/ourwork/virtual-constellations/ocr/.
30. FRM4SOC, R. Vendt, and et. al., *D-280, FRM4SOC Scientific and Operational Roadmap (SOR)* (2019).
31. FRM4SOC, C. Lerebourg, and et. al., *D-240, Proceedings of WKP-1 (PROC-1). Report of the International Workshop* (2017), p. 108.
32. E. Leymarie, C. Penkerc'h, V. Vellucci, C. Lerebourg, D. Antoine, E. Boss, M. Lewis, F. D'Ortenzio, and H. Claustre, "ProVal: A New Autonomous Profiling Float for High Quality Radiometric Measurements," Front Mar Sci (2018).
33. G. P. Gerbi, E. Boss, P. J. Werdell, C. W. Proctor, N. Haëntjens, M. R. Lewis, K. Brown, D. Sorrentino, J. R. V. Zaneveld, A. H. Barnard, J. Koegler, H. Fargher, M. DeDonato, and W. Wallace, "Validation of Ocean Color Remote Sensing Reflectance Using Autonomous Floats," J. Atmospheric Ocean. Technol. **33**, 2331–2352 (2016).
34. FRM4SOC, C. Lerebourg, and et. al., *D-250, Technical Report TR-10, Requirements and Recommendations for Infrastructure Required for the Long-Term Vicarious Adjustment of the Sentinel-3 OLCI and Sentinel-2 MSI A/B/C and D Instruments* (2019).
35. International Standards Organisation (ISO), "Space Environment (Natural And Artificial)—Process for Determining Solar Irradiances. ISO Report 21348:2007," (2007).
36. H. R. Gordon and M. Wang, "Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: a preliminary algorithm," Appl. Opt. **33**, 443–452 (1994).
37. D. Antoine and A. Morel, "A multiple scattering algorithm for atmospheric correction of remotely sensed ocean colour (MERIS instrument): Principle and implementation for atmospheres carrying various aerosols including absorbing ones," Int. J. Remote Sens. **20**, 1875–1916 (1999).
38. K. G. Ruddick, F. Ovidio, and M. Rijkeboer, "Atmospheric correction of SeaWiFS imagery for turbid coastal and inland waters," Appl. Opt. **39**, 897–912 (2000).
39. D. A. Siegel, M. Wang, S. Maritorena, and W. Robinson, "Atmospheric correction of satellite ocean color imagery: the black pixel assumption," Appl. Opt. **39**, 3582–3591 (2000).
40. M. Wang and W. Shi, "The NIR-SWIR combined atmospheric correction approach for MODIS ocean color data processing," Opt. Express **15**, 15722–15733 (2007).
41. C. Brockmann, R. Doerffer, M. Peters, S. Kerstin, S. Embacher, and A. Ruescas, "Evolution of the C2RCC Neural Network for Sentinel 2 and 3 for the Retrieval of Ocean Colour Products in Normal and Extreme Optically Complex Waters," Living Planet Symp. **740**, 54 (2016).
42. F. Steinmetz, P.-Y. Deschamps, and D. Ramon, "Atmospheric correction in presence of sun glint: application to MERIS," Opt. Express **19**, 9783–9800 (2011).
43. I. S. Robinson, *Measuring the Oceans from Space: The Principles and Methods of Satellite Oceanography*, Geophysical Sciences (Springer-Verlag, 2004).

44. IOCCG, *Mission Requirements for Future Ocean-Colour Sensors*, Reports and Monographs of the International Ocean-Colour Coordinating Group (2012).
45. G. Zibordi, F. Mélin, K. J. Voss, B. C. Johnson, B. A. Franz, E. Kwiatkowska, J.-P. Huot, M. Wang, and D. Antoine, "System vicarious calibration for ocean color climate change applications: Requirements for in situ data," Remote Sens. Environ. **159**, 361–369 (2015).
46. G. Zibordi and K. J. Voss, "Chapter 3.1 - In situ Optical Radiometry in the Visible and Near Infrared," in *Experimental Methods in the Physical Sciences*, G. Zibordi, C. J. Donlon, and A. C. Parr, eds., Optical Radiometry for Ocean Climate Measurements (Academic Press, 2014), Vol. 47, pp. 247–304.
47. H. R. Gordon and D. K. Clark, "Clear water radiances for atmospheric correction of coastal zone color scanner imagery," Appl. Opt. **20**, 4175–4180 (1981).
48. H. R. Gordon, D. K. Clark, J. W. Brown, O. B. Brown, R. H. Evans, and W. W. Broenkow, "Phytoplankton pigment concentrations in the Middle Atlantic Bight: comparison of ship determinations and CZCS estimates," Appl. Opt. **22**, 20–36 (1983).
49. H. R. Gordon, "Calibration requirements and methodology for remote sensors viewing the ocean in the visible," Remote Sens. Environ. **22**, 103–126 (1987).
50. S. B. Hooker, W. B. Esaias, G. C. Feldman, W. W. Gregg, and C. R. McClain, *SeaWiFS Technical Report Series: An Overview of SeaWiFS and Ocean Color* (National Aeronautics and Space Administration, 1992), Vol. 1.
51. World Meteorological Organization (WMO), S. and C. O. (UNESCO) United Nations Educational, (IOC) Intergovernmental Oceanographic Commission, (UNEP) United Nations Environment Programme, (ICSU) International Council of Scientific Unions, and World Meteorological Organization (WMO), *GCOS, 200. The Global Observing System for Climate : Implementation Needs* (WMO, 2016).
52. H. R. Gordon, "In-Orbit Calibration Strategy for Ocean Color Sensors," Remote Sens. Environ. **63**, 265–278 (1998).
53. K. von Schuckmann, M. Drévillon, and P.-Y. Le Traon, "Multi Year Product Strategy Plan (CMEMS-MYP)," (2016).
54. FRM4SOC, K. Ruddick, and et. al., *D-70, Technical Report TR-2, A Review of Commonly Used Fiducial Reference Measurement (FRM) Ocean Colour Radiometers (OCR) Used for Satellite OCR Validation* (2018).
55. D. Antoine, P. Guevel, J.-F. Desté, G. Bécu, F. Louis, A. J. Scott, and P. Bardey, "The “BOUSSOLE” Buoy—A New Transparent-to-Swell Taut Mooring Dedicated to Marine Optics: Design, Tests, and Performance at Sea," J. Atmospheric Ocean. Technol. **25**, 968–989 (2008).
56. "BOUSSOLE," http://www.obs-vlfr.fr/Boussole/html/project/introduction.php.
57. "Marine Optical BuoY (MOBY) – Provides vicarious calibration of ocean color satellites," https://www.mlml.calstate.edu/moby/.
58. "MOBY, A Radiometric Buoy for Performance Monitoring and Vicarious Calibration of Satellite Ocean Color Sensors: Measurement and Data Analysis Protocols," https://www.researchgate.net/publication/24293242_MOBY_A_Radiometric_Buoy_for_Performance_Monitoring_and_Vicarious_Calibration_of_Satellite_Ocean_Color_Sensors_Measurement_and_Data_Analysis_Protocols.
59. S. W. Brown, S. J. Flora, M. E. Feinholz, M. A. Yarbrough, T. Houlihan, D. Peters, Y. S. Kim, J. L. Mueller, B. C. Johnson, and D. K. Clark, "The marine optical buoy (MOBY) radiometric calibration and uncertainty budget for ocean color satellite sensor vicarious calibration," in *Sensors, Systems, and Next-Generation Satellites XI* (International Society for Optics and Photonics, 2007), Vol. 6744, p. 67441M.
60. D. K. Clark, M. Feinholz, M. Yarbrough, B. C. Johnson, S. W. Brown, Y. S. Kim, and R. A. Barnes, "Overview of the radiometric calibration of MOBY," in *Earth Observing Systems VI* (International Society for Optics and Photonics, 2002), Vol. 4483, pp. 64–76.

61. S. J. Flora, S. W. Brown, and B. C. Johnson, "White paper on MOBY/AHAB wavelength resolution," presented at MOBY/AHAB Review Meeting (July 18, 2006).
62. M. Feinholz, B. C. Johnson, K. Voss, M. Yarbrough, and S. Flora, "Immersion Coefficient for the Marine Optical Buoy (MOBY) Radiance Collectors.," J. Res. Natl. Inst. Stand. Technol. **122**, (2017).
63. J. L. Mueller, *Self-Shading Corrections for MOBY Upwelling Radiance Measurements* (2007).
64. K. J. Voss, B. C. Johnson, M. A. Yarbrough, and A. Gleason, "An ocean colour vicarious calibration system," presented at Earth Science Technology Forum (ESTF), Pasadena (CA), USA (June 2015).
65. K. J. Voss and N. Souaidia, "POLRADS: polarization radiance distribution measurement system," Opt. Express **18**, 19672–19680 (2010).
66. K. J. Voss, S. McLean, M. Lewis, C. Johnson, S. Flora, M. Feinholz, M. Yarbrough, C. Trees, M. Twardowski, and D. Clark, "An Example Crossover Experiment for Testing New Vicarious Calibration Techniques for Satellite Ocean Color Radiometry," J. Atmospheric Ocean. Technol. **27**, 1747–1759 (2010).
67. K. Voss and S. Flora, "Spectral dependence of the seawater-air radiance transmission coefficient," J. Atmospheric Ocean. Technol. **34**, 1203–1205 (2017).
68. K. Voss, H. R. Gordon, S. Flora, B. C. Johnson, M. Yarbrough, M. Feinholz, and T. Houlihan, "A Method to extrapolate the diffuse upwelling radiance attenuation coefficient to the surface as applied to the Marine Optical Buoy (MOBY)," J. Atmospheric Ocean. Technol. **34**, 1423–1432 (2017).
69. D. Antoine, M. Chami, H. Claustre, F. d'Ortenzio, A. Morel, G. Bécu, B. Gentili, F. Louis, J. Ras, E. Roussier, A. J. Scott, D. Tailliez, S. B. Hooker, P. Guevel, J.-F. Desté, C. Dempsey, and D. Adams, "BOUSSOLE: a joint CNRS-INSU, ESA, CNES and NASA ocean colour calibration and validationactivity," NASA Tech. Memo. N° 2006-214147 61 (2006).
70. V. Velluci, E. Leymarie, B. Gentili, and D. Antoine, "Shadowing corrections of BOUSSOLE radiometric measurements," in *Proceedings of SPIE, Ocean Optics XXII Conference* (2014), pp. 27–31.
71. A. Białek, V. Vellucci, B. Gentil, D. Antoine, J. Gorroño, N. Fox, and C. Underwood, "Monte Carlo–Based Quantification of Uncertainties in Determining Ocean Remote Sensing Reflectance from Underwater Fixed-Depth Radiometry Measurements," J. Atmospheric Ocean. Technol. **37**, 177–196 (2020).
72. C. Mazeran, C. Brockmann, K. Ruddick, K. Voss, and F. Zagolski, "Requirements for Copernicus Ocean Colour Vicarious Calibration Infrastructure. Technical report," (2017).
73. G. Zibordi, M. Talone, K. J. Voss, and B. C. Johnson, "Impact of spectral resolution of in situ ocean color radiometric data in satellite matchups analyses," Opt. Express **25**, A798–A812 (2017).
74. O. Burggraaff and O. Burggraaff, "Biases from incorrect reflectance convolution," Opt. Express **28**, 13801–13816 (2020).
75. G. Zibordi and F. Mélin, "An evaluation of marine regions relevant for ocean color system vicarious calibration," Remote Sens. Environ. **190**, 122–136 (2017).
76. G. Zibordi, F. Mélin, and M. Talone, "JRC Technical Report EU 28433 EN: System Vicarious Calibration for Copernicus Ocean Colour Missions: Requirements and Recommendations for a European Site," (2017).
77. K. J. Voss, B. C. Johnson, M. A. Yarbrough, S. J. Flora, M. E. Feinholz, T. Houlihan, D. Peters, and A. Gleason, "Overview and status of MOBY-Net concept.," presented at International Ocean Colour Symposium, Lisbon, Portugal (2017).
78. S. W. Bailey and P. J. Werdell, "A multi-sensor approach for the on-orbit validation of ocean color satellite data products," Remote Sens. Environ. **102**, 12–23 (2006).

79. B. A. Franz, S. W. Bailey, P. J. Werdell, and C. R. McClain, "Sensor-independent approach to the vicarious calibration of satellite ocean color radiometry," Appl. Opt. **46**, 5068–5082 (2007).
80. D. Vansteenwegen, K. Ruddick, A. Cattrijsse, Q. Vanhellemont, and M. Beck, "The Pan-and-Tilt Hyperspectral Radiometer System (PANTHYR) for Autonomous Satellite Validation Measurements—Prototype Design and Testing," Remote Sens. **11**, 1360 (2019).
81. FRM4SOC, K. Ruddick, and et. al., *D-60, Technical, Report TR-1, Measurement Requirements and Protocols When Operating Fiducial Reference Measurement (FRM) Ocean Colour Radiometers (OCR) Used for Satellite Validation* (2018).
82. K. G. Ruddick, K. Voss, A. C. Banks, E. Boss, A. Castagna, R. Frouin, M. Hieronymi, C. Jamet, B. C. Johnson, J. Kuusk, Z. Lee, M. Ondrusek, V. Vabson, and R. Vendt, "A Review of Protocols for Fiducial Reference Measurements of Downwelling Irradiance for the Validation of Satellite Remote Sensing Data over Water," Remote Sens. **11**, 1742 (2019).
83. K. G. Ruddick, K. Voss, E. Boss, A. Castagna, R. Frouin, A. Gilerson, M. Hieronymi, B. C. Johnson, J. Kuusk, Z. Lee, M. Ondrusek, V. Vabson, and R. Vendt, "A Review of Protocols for Fiducial Reference Measurements of Water-Leaving Radiance for Validation of Satellite Remote-Sensing Data over Water," Remote Sens. **11**, 2198 (2019).
84. C. D. Mobley, *Light and Water: Radiative Transfer in Natural Waters* (Academic Press, 1994).
85. C. D. Mobley, "Ocean Optics Web Book," http://www.oceanopticsbook.info/view/overview_of_optical_oceanography/reflectances.
86. H. R. Gordon and D. K. Clark, "Clear Water Radiances for Atmospheric Correction of Coastal Zone Color Scanner Imagery," Appl. Opt. **20**, 4175–4180 (1981).
87. J. L. Mueller, "Shadow corrections to in-water upwelled radiance measurments: a status review," in *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 5, Volume VI: Special Topics in Ocean Optics Protocols, Part 2*, NASA/TM (2004), pp. 1–7.
88. G. Thuillier, M. Herse, D. Labs, T. Foujols, W. Peetermans, D. Gillotay, P. C. Simon, and H. Mandel, "The solar spectral irradiance from 200 to 2400 nm as measured by the SOLSPEC spectrometer from the ATLAS and EURECA missions," Sol. Phys. **214**, 1–22 (2003).
89. A. Morel and J. L. Mueller, "Chapter 4. Normalized water-leaving radiance and remote sensing reflectance: bidirectional reflectance and other factors," in *Ocean Optics Protocols For Satellite Ocean Color Sensor Validation, Revision 4, Volume III: Radiometric Measuremens and Data Analysis Protocols* (2003), pp. 32–59.
90. J. L. Mueller, G. S. Fargion, C. R. McClain, J. L. Mueller, S. W. Brown, D. K. Clark, B. C. Johnson, H. Yoon, K. R. Lykke, S. J. Flora, N. Souaidia, C. Pietras, T. C. Stone, M. A. Yarbrough, Y. S. Kim, and R. A. Barnes, *Ocean Optics Protocols For Satellite Ocean Color Sensor Validation*, Technical Memorandum TM-2003-21621/Revision 5 (NASA Goddard Space Flight Space Center, 2004).
91. G. Zibordi, F. Mélin, J.-F. Berthon, B. Holben, I. Slutsker, D. Giles, D. D'Alimonte, D. Vandemark, H. Feng, G. Schuster, B. E. Fabbri, S. Kaitala, and J. Seppälä, "AERONET-OC: A Network for the Validation of Ocean Color Primary Products," J. Atmospheric Ocean. Technol. **26**, 1634–1651 (2009).
92. IOCCG, *Bio-Optical Sensors on Argo Floats.* (International Ocean Colour Coordinating Group (IOCCG), 2011).
93. IOCCG, G. Zibordi, and K. J. Voss, "Protocols for Satellite Ocean Color Data Validation: In situ Optical Radiometry," (2019).

94. EC, FP5-EESD, "Regional validation of meris chlorophyll products in north sea coastal waters (REVAMP Project EVG1-CT-2001-00049)," (2002).
95. G. Tilstone, G. Moore, K. Sorensen, R. Doerffer, R. Rottgers, K. G. Ruddick, R. Pasterkamp, and P. V. Jorgensen, "REVAMP Regional Validation of MERIS Chlorophyll products in North Sea coastal waters: Protocols document," in (European Space Agency, 2003), Vol. WPP-233.
96. R. Doerffer, "Protocols for the Validation of MERIS Water Products, Doc. no. PO-TN-MEL-GS-0043, Issue 1, Revision 3.," (2002).
97. EC, "Coastal region long-term measurements for colour remote sensing development and validation, Project MAS-CT97-0087," (1997).
98. K. Barker, C. Mazeran, C. Lerebourg, M. Bouvet, D. Antoine, M. Ondrusek, G. Zibordi, and S. Lavender, "MERMAID: the MEris MAtchup In-situ Database," in *2nd MERIS/(A)ATSR User Workshop Proceedings* (2008).
99. Tartu Observatory, SYKE, CNR, and Water Insight, *Global Lake Sentinel Services (GLASS) Technical Report about Measurement Protocols* (2015).
100. A. Hommersom, S. Kratzer, M. Laanen, I. Ansko, M. Ligi, M. Bresciani, C. Giardino, J. M. Beltran-Abaunza, G. Moore, M. R. Wernand, and S. W. Peters, "Intercomparison in the field between the new WISP-3 and other radiometers (TriOS Ramses, ASD FieldSpec, and TACCS)," J. Appl. Remote Sens. **6**, 063615 (2012).
101. S. G. H. Simis and J. Olsson, "Unattended processing of shipborne hyperspectral reflectance measurements," Remote Sens. Environ. **135**, 202–212 (2013).
102. K. G. Ruddick, V. D. Cauwer, Y.-J. Park, and G. Moore, "Seaborne measurements of near infrared water-leaving reflectance: The similarity spectrum for turbid waters," Limnol. Oceanogr. **51**, 1167–1179 (2006).
103. K. Kallio, S. Koponen, A. Ruiz-Verdú, T. Heege, K. Sørensen, T. Pyhälahti, and R. Doerffer, *Development of MERIS Lake Water Algorithms Document Title: Validation Protocol (WP 2.4)* (2007).
104. C. R. McClain, M. L. Cleave, G. C. Feldman, W. W. Gregg, S. B. Hooker, and N. Kuring, "Science quality SeaWiFS data for global biosphere research," Sea Technol. **9**, 10–16 (1998).
105. C. B. Mouw, S. Greb, D. Aurin, P. M. DiGiacomo, Z. Lee, M. Twardowski, C. Binding, C. Hu, R. Ma, T. Moore, and others, "Aquatic color radiometry remote sensing of coastal and inland waters: Challenges and recommendations for future satellite missions," Remote Sens. Environ. **160**, 15–30 (2015).
106. S. B. Hooker, G. Zibordi, and S. Maritorena, "The Second SeaWiFs Ocean Optics DARR (DARR-00)," in *Results of the Second SeaWiFS Data Analysis Round Robin, March 2000 (DARR-00)*, SeaWiFS Postlaunch Technical Report Series Volume 15 No. NASA Technical Memorandum 2001-206892 (2001), Vol. 15, pp. 4–45.
107. G. Zibordi, K. Ruddick, I. Ansko, G. Moore, S. Kratzer, J. Icely, and A. Reinart, "In situ determination of the remote sensing reflectance: an inter-comparison," OCEAN Sci. **8**, 567–586 (2012).
108. M. Ondrusek, V. P. Lance, R. Arnone, S. Ladner, W. Goode, R. Vandermeulen, S. Freeman, J. E. Chaves, A. Mannino, A. Gilerson, S. Ahmed, C. Carrizo, A. El-Habashi, R. Foster, M. Ottaviani, J. I. Goes, H. D. R. Gomes, K. McKee, C. Hu, C. Kovach, D. English, J. Cannizzaro, B. C. Johnson, Z. Lee, J. Wei, Q. Wang, J. Lin, N. Tufillaro, J. Nahorniak, C. O. Davis, K. J. Voss, E. Stengel, and M. Wang, *Report for Dedicated JPSS VIIRS Ocean Color December 2015 Calibration/Validation Cruise.* (U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service., 2016).
109. G. Tilstone, G. Dall'Olmo, M. Hieronymi, K. Ruddick, M. Beck, M. Ligi, M. Costa, D. D'Alimonte, V. Velluci, D. Vansteenwegen, A. Bracher, S. Wiegmann, J. Kuusk, V. Vabson, I. Ansko, R. Vendt, C. Donlon, and T. Casal, "Field intercomparison of radiometer measurements for ocean colour validation," Remote Sens. **12**, (2020).

110. A. Morel, "In-water and remote measurements of ocean colour," Bound. Layer Meteorol. **18**, 177–201 (1980).
111. W. Joseph Rhea and C. O. Davis, "A comparison of the SeaWiFS chlorophyll and CZCS pigment algorithms using optical data from the 1992 JGOFS Equatorial Pacific Time Series," Deep Sea Res. Part II Top. Stud. Oceanogr. **44**, 1907–1925 (1997).
112. K. L. Carder and R. G. Steward, "A remote-sensing reflectance model of a red-tide dinoflagellate off west Florida," Limnol. Oceanogr. **30**, 286–298 (1985).
113. E. J. Milton, M. E. Schaepman, K. Anderson, M. Kneubühler, and N. Fox, "Progress in field spectroscopy," Remote Sens. Environ. **113**, S92–S109 (2009).
114. B. Fougnie, R. Frouin, P. Lecomte, and P.-Y. Deschamps, "Reduction of skylight reflection effects in the above-water measurement of diffuse marine reflectance," Appl. Opt. **38**, 3844–3856 (1999).
115. P.-Y. Deschamps, B. Fougnie, R. Frouin, P. Lecomte, and C. Verwaerde, "SIMBAD: a field radiometer for satellite ocean-color validation," Appl. Opt. **43**, 4055–4069 (2004).
116. J. L. Mueller, R. Frouin, C. Davis, R. Arnone, K. Carder, R. G. Steward, S. Hooker, C. D. Mobley, and S. McLean, "Chapter 3. Above-Water Radiance and Remote Sensing Reflectance Measurement and Analysis Protocols," in *Ocean Optics Protocols For Satellite Ocean Color Sensor Validation, Revision 4, Volume III: Radiometric Measurements and Data Analysis Protocols* (2003), pp. 21–31.
117. J. L. A. Mueller, G. S. Fargion, and C. R. McClain, "Ocean Optics Protocols For Satellite Ocean Color Sensor Validation, Revision 4, Volume I: Introduction, Background and Conventions," (2003).
118. J. L. Mueller, G. S. Fargion, C. R. McClain, J. L. Mueller, A. Morel, R. Frouin, C. Davis, R. Arnone, K. Carder, R. G. Steward, S. Hooker, C. D. Mobley, S. McLean, B. Holben, C. Pietras, K. D. Knobelspiesse, G. S. Fargion, and J. Porter, "Ocean Optics Protocols For Satellite Ocean Color Sensor Validation, Revision 4, Volume III: Radiometric Measurements and Data Analysis Protocols," 84 (2003).
119. Z. Lee, Y.-H. Ahn, C. Mobley, and R. Arnone, "Removal of surface-reflected light for the measurement of remote-sensing reflectance from an above-surface platform," Opt. Express **18**, 26313 (2010).
120. Z. Lee, N. Pahlevan, Y.-H. Ahn, S. Greb, and D. O'Donnell, "Robust approach to directly measuring water-leaving radiance in the field," Appl. Opt. **52**, 1693 (2013).
121. H. Lin, H. Lin, Z. Lee, G. Lin, X. Yu, and X. Yu, "Experimental evaluation of the self-shadow and its correction for on-water measurements of water-leaving radiance," Appl. Opt. **59**, 5325–5334 (2020).
122. E. Knaeps, A. I. Dogliotti, D. Raymaekers, K. Ruddick, and S. Sterckx, "In situ evidence of non-zero reflectance in the OLCI 1020nm band for a turbid estuary," Remote Sens. Environ. **120**, 133–144 (2012).
123. A. I. Dogliotti, J. I. Gossn, Q. Vanhellemont, and K. G. Ruddick, "Detecting and Quantifying a Massive Invasion of Floating Aquatic Plants in the Río de la Plata Turbid Waters Using High Spatial Resolution Ocean Color Imagery," Remote Sens. **10**, 1140 (2018).
124. S. Mekaoui and G. Zibordi, "Cosine error for a class of hyperspectral irradiance sensors," Metrologia **50**, 187–199 (2013).
125. S. G. R. Salim, N. P. Fox, E. Theocharous, T. Sun, and K. T. V. Grattan, "Temperature and nonlinearity corrections for a photodiode array spectrometer used in the field," Appl. Opt. **50**, 866–875 (2011).
126. H. W. Yoon, J. J. Butler, T. C. Larason, and G. P. Eppeldauer, "Linearity of InGaAs photodiodes," Metrologia **40**, S154–S158 (2003).
127. W. W. Gregg and K. L. Carder, "A simple spectral solar irradiance model for cloudless maritime atmospheres," Limnol. Oceanogr. **35**, 1657–1675 (1990).

128. A. Castagna, B. C. Johnson, K. Voss, H. M. Dierssen, H. Patrick, T. A. Germer, K. Sabbe, and W. Vyverman, "Uncertainty in global downwelling plane irradiance estimates from sintered polytetrafluoroethylene plaque radiance measurements," Appl. Opt. **58**, 4497–4511 (2019).
129. C. D. Mobley, "Estimation of the remote-sensing reflectance from above-surface measurements," Appl. Opt. **38**, 7442–7455 (1999).
130. D. Vansteenwegen, K. Ruddick, A. Cattrijsse, Q. Vanhellemont, and M. Beck, "The Pan-and-Tilt Hyperspectral Radiometer System (PANTHYR) for Autonomous Satellite Validation Measurements—Prototype Design and Testing," Remote Sens. **11**, 1360 (2019).
131. S. P. Garaba, J. Schulz, M. R. Wernand, and O. Zielinski, "Sunglint Detection for Unmanned and Automated Platforms," Sensors **12**, 12545–12561 (2012).
132. C. Carrizo, A. Gilerson, R. Foster, A. Golovin, and A. El-Habashi, "Characterization of radiance from the ocean surface by hyperspectral imaging," Opt. Express **27**, 1750–1768 (2019).
133. A. Gilerson, C. Carrizo, R. Foster, T. Harmel, A. Golovin, A. El-Habashi, E. Herrera, and T. Wright, "Total and polarized radiance from the ocean surface from hyperspectral polarimetric imaging," in *Ocean Sensing and Monitoring XI* (International Society for Optics and Photonics, 2019), Vol. 11014, p. 110140F.
134. R. Frouin and R. T. Pinker, "Estimating Photosynthetically Active Radiation (PAR) at the earth's surface from satellite observations," Remote Sens. Environ. **51**, 98–107 (1995).
135. FRM4SOC, J. Kuusk, and et. al., *D-170, Technical Report TR-6, Results from the First FRM4SOC Field Ocean Colour Radiometer Verification Round Robin Campaign* (2018).
136. FRM4SOC, "Radiometers used to make Fiducial Reference Measurements for Ocean Colour Satellite Validation," in *Minutes of the FRM4SOC Seminar with Manufacturers of OCR* (2017).
137. J. L. Mueller and R. W. Austin, "Chapter 3. Characterization of Oceanographic and Atmospheric Radiometers," in *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 4, Volume II: Instrument Specifications, Characterization and Calibration* (2003), pp. 17–33.
138. E. Aas, *On Submarine Irradiance Measurements* (1969).
139. T. Ohde and H. Siegel, "Derivation of immersion factors for the hyperspectral TriOS radiance sensor," J. Opt. Pure Appl. Opt. **5**, L12–L14 (2003).
140. G. Zibordi and M. Darecki, "Immersion factors for the RAMSES series of hyper-spectral underwater radiometers," J. Opt. Pure Appl. Opt. **8**, 252–258 (2006).
141. J. Kuusk, "Dark Signal Temperature Dependence Correction Method for Miniature Spectrometer Modules," https://www.hindawi.com/journals/js/2011/608157/.
142. S. B. Hooker, G. Bernhard, J. H. Morrow, C. R. Booth, T. Comer, R. N. Lind, and V. Quang, *Optical Sensors for Planetary Radiant Energy (OSPREy): Calibration and Validation of Current and Next-Generation NASA Missions* (2012).
143. G. Zibordi, M. Talone, and L. Jankowski, "Response to Temperature of a Class of In Situ Hyperspectral Radiometers," J. Atmospheric Ocean. Technol. **34**, 1795–1805 (2017).
144. S. B. Hooker, S. McLean, and M. Small, "Rotation and polarization uncertainties (Chapter 7)," in *The Seventh SeaWiFS Intercalibration Round-Robin Experiment (SIRREX-7), March 1999*, NASA/TM SeaWiFS Postlaunch Technical Report Series No. Volume 17 (2002).
145. S. Mekaoui and G. Zibordi, "Cosine error for a class of hyperspectral irradiance sensors," Metrologia **50**, 187–199 (2013).
146. S. W. Brown, B. C. Johnson, H. W. Yoon, J. J. Butler, R. Barnes, S. F. Biggar, P. R. Spyak, K. J. Thome, E. F. Zalewski, M. Helmlinger, C. J. Bruegge, S. Schiller, G.

Fedosejevs, R. Gauthier, S. Tsuchida, and S. Machida, "Radiometric Characterization of Field Radiometers in Support of the 1997 Lunar Lake, Nevada, Experiment to Determine Surface Reflectance and Top-of-Atmosphere Radiance," Int. Symp. Remote Sens. Environ. **77**, (2001).

147. S. Nevas, A. Sperling, and B. Oderkerk, "Transferability of stray light corrections among array spectroradiometers," AIP Conf. Proc. **1531**, 821–824 (2013).
148. S. B. Hooker, G. Bernhard, J. Morrow, C. Booth, T. Comer, R. N. Lind, and V. Quang, "Optical Sensors for Planetary Radiant Energy (OSPREy): Calibration and Validation of Current and Next-Generation NASA Missions, NASA/TM–2012–215872," (2012).
149. FRM4SOC, A. C. Banks, and et. al., *D-80a, Technical Report TR-3 Part I, Protocols and Procedures to Verify the Performance of Reference Irradiance Sources Used by Fiducial Reference Measurement Ocean Colour Radiometers for Satellite Validation* (2017).
150. FRM4SOC, A. C. Banks, and et. al., *D-80b, Technical Report TR-3 Part II, Protocols and Procedures to Verify the Performance of Reference Radiance Sources Used by Fiducial Reference Measurement Ocean Colour Radiometers for Satellite Validation* (2017).
151. FRM4SOC, A. C. Banks, and et. al., *D-90, FRM4SOC Laboratory Calibration Exercise 1: Implementation Plan for LCE-12 (LCE-IP)* (2016).
152. FRM4SOC, A. Bialek, and et. al., *D-120, Technical Report TR-4, Results from the First FRM4SOC Reference Radiance and Irradiance Source Verification Laboratory Calibration Experiment Campaign* (2019).
153. A. Białek, T. Goodman, E. Woolliams, J. F. S. Brachmann, T. Schwarzmaier, J. Kuusk, I. Ansko, V. Vabson, I. C. Lau, C. MacLellan, S. Marty, M. Ondrusek, W. Servantes, S. Taylor, R. Van Dommelen, A. Barnard, V. Vellucci, A. C. Banks, N. Fox, R. Vendt, C. Donlon, and T. Casal, "Results from Verification of Reference Irradiance and Radiance Sources Laboratory Calibration Experiment Campaign," Remote Sens. **12**, 2220 (2020).
154. M. Gergely and G. Zibordi, "Assessment of AERONET-OC L WN uncertainties," Metrologia **51**, 40 (2014).
155. G. Meister, P. Abel, R. Barnes, J. Cooper, C. Davis, M. Godin, D. Goebel, G. Fargion, R. Frouin, D. Korwan, R. Maffione, C. McClain, S. McLean, D. Menzies, A. Poteau, J. Robertson, and J. Sherman, "The First SIMBIOS Radiometric Intercomparison (SIMRIC-1), April-September 2001," NASA Tech. Memo. TM-2002-210006 (2002).
156. G. Meister, P. Abel, K. Carder, A. Chapin, D. Clark, J. Cooper, C. Davis, D. English, G. Fargion, M. Feinholz, R. Frouin, F. Hoge, D. Korwan, G. Lazin, C. McClain, S. McLean, D. Menzies, A. Poteau, J. Robertson, J. Sherman, K. Voss, and J. Yungel, "The second SIMBIOS Radiometric Intercomparison (SIMRIC-2), March-November 2002," NASA Tech. Memo. **Vol. II**, (2003).
157. IOCCG, *Report, "Atmospheric Correction for Remotely-Sensed Ocean-Colour Products"* (2020).
158. N. Fox and M. C. Greening, "A guide to comparisons – organisation, operation and analysis to establish measurement equivalence to underpin the Quality Assurance requirements of GEO, versio-4, QA4EO-QAEO-GEN-DQK-004," (2010).
159. QA4EO, "The Guide," (n.d.).
160. "CEOS Working Group on Calibration and Validation (WGCV)," (n.d.).
161. H. W. Yoon, D. W. Allen, G. P. Eppeldauer, and B. K. Tsai, "The extension of the NIST BRDF scale from 1100 nm to 2500 nm," in *Proc. SPIE* (2009), Vol. 7452, pp. 745204–745212.
162. "NASA PACE," (2018).
163. JCGM100:2008, *Evaluation of Measurement Data - Guide to the Expression of Uncertainty in Measurement* (2008).

164. E. R. Woolliams, N. P. Fox, M. G. Cox, P. M. Harris, and N. J. Harrison, "The {CCPR}{K}1-a key comparison of spectral irradiance from 250~nm to 2500~nm: measurements, analysis and results," Metrologia **43**, S98 (2006).
165. S. B. Hooker, S. McLean, J. Sherman, M. Small, G. Lazin, G. Zibordi, and J. W. Brown, *The Seventh SeaWiFS Intercalibration Round-Robin Experiment (SIRREX-7)* (2002).
166. FRM4SOC, J. Kuusk, and et. al., *D-130, Technical Report TR-5, Protocols and Procedures to Verify the Performance of Fiducial Reference Measurement (FRM) Field Ocean Colour Radiometers (OCR) Used for Satellite Validation* (2017).
167. V. Vabson, J. Kuusk, I. Ansko, R. Vendt, K. Alikas, K. Ruddick, A. Ansper, M. Bresciani, H. Burmester, M. Costa, D. D'Alimonte, G. Dall'Olmo, B. Damiri, T. Dinter, C. Giardino, K. Kangro, M. Ligi, B. Paavel, G. Tilstone, R. Van Dommelen, S. Wiegmann, A. Bracher, C. Donlon, and T. Casal, "Laboratory Intercomparison of Radiometers Used for Satellite Validation in the 400–900 nm Range," Remote Sens. **11**, 1101 (2019).
168. V. Vabson, J. Kuusk, I. Ansko, R. Vendt, K. Alikas, K. Ruddick, A. Ansper, M. Bresciani, H. Burmester, M. Costa, D. D'Alimonte, G. Dall'Olmo, B. Damiri, T. Dinter, C. Giardino, K. Kangro, M. Ligi, B. Paavel, G. Tilstone, R. Van Dommelen, S. Wiegmann, A. Bracher, C. Donlon, and T. Casal, "Field Intercomparison of Radiometers Used for Satellite Validation in the 400–900 nm Range," Remote Sens. **11**, 1129 (2019).
169. N. Fox and M. C. Greening, "A guide to comparisons – organisation, operation and analysis to establish measurement equivalence to underpin the Quality Assurance requirements of GEO," (2010).
170. FRM4SOC, "Field Inter-Comparison Experiment (FICE) the Acqua Alta Oceanographic Tower (AAOT)," https://frm4soc.org/index.php/activities/fice/.
171. FRM4SOC, J. Kuusk, and et. al., *D-140, FRM4SOC Laboratory Calibration Exercise 2: Implementation Plan for LCE-2 (LCE-IP)* (2016).
172. B. C. Johnson, H. W. Yoon, J. P. Rice, and A. C. Parr, "Chapter 1.2 - Principles of Optical Radiometry and Measurement Uncertainty," in *Experimental Methods in the Physical Sciences*, G. Zibordi, C. J. Donlon, and A. C. Parr, eds., Optical Radiometry for Ocean Climate Measurements (Academic Press, 2014), Vol. 47, pp. 13–67.
173. J. L. Mueller, G. S. Fargion, and C. R. McClain, "Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 4, Volume II: Instrument Specifications, Characterization and Calibration," NASA Tech. Memo. TM-2003-21621Rev-Vol II (2003).
174. S. G. R. Salim, N. P. Fox, W. S. Hartree, E. R. Woolliams, T. Sun, and K. T. V. Grattan, "Stray light correction for diode-array-based spectrometers using a monochromator," Appl. Opt. **50**, 5130–5138 (2011).
175. S. B. Hooker, E. R. Firestone, S. McLean, J. Sherman, M. Small, G. Lazin, G. Zibordi, J. W. Brown, and C. R. McClain, "The Seventh SeaWiFS Intercalibration Round-Robin Experiment (SIRREX-7), TM-2003- 206892, vol. 17, NASA Goddard Space Flight Center, Greenbelt," (2002).
176. S. G. R. Salim, E. R. Woolliams, and N. P. Fox, "Calibration of a Photodiode Array Spectrometer Against the Copper Point," Int. J. Thermophys. **35**, 504–515 (2014).
177. L. Ylianttila, R. Visuri, L. Huurto, and K. Jokela, "Evaluation of a Single-monochromator Diode Array Spectroradiometer for Sunbed UV-radiation Measurements¶," Photochem. Photobiol. **81**, 333–341 (2005).
178. G. Seckmeyer, "Instruments to Measure Solar Ultraviolet Radiation Part 4: Array Spectroradiometers (lead author: G. Seckmeyer) (WMO/TD No. 1538). 44 pp. November 2010.," (2010).
179. D. Antoine, T. Schroeder, M. Slivkoff, W. Klonowski, M. Doblin, J. Lovell, D. Boadle, B. Baker, E. Botha, C. Robinson, E. King, P. Fearns, N. Hardman-Montford, R. Johnson,

N. Cherukuru, A. Dekker, T. Malthus, R. Mitchell, P. Thompson, and P. Van Ruth, *IMOS Radiometry Task Team* (2017).

180. G. Zibordi, M. Talone, and L. Jankowski, "Response to Temperature of a Class of In Situ Hyperspectral Radiometers," J. Atmospheric Ocean. Technol. **34**, 1795–1805 (2017).
181. S. G. R. Salim, N. P. Fox, E. Theocharous, T. Sun, and K. T. V. Grattan, "Temperature and nonlinearity corrections for a photodiode array spectrometer used in the field," Appl. Opt. **50**, 866–875 (2011).
182. L. L. A. Price, R. J. Hooke, and M. Khazova, "Effects of ambient temperature on the performance of CCD array spectroradiometers and practical implications for field measurements," J. Radiol. Prot. **34**, 655 (2014).
183. L. Li, C. Dai, Z. Wu, and Y. Wang, "Temperature and nonlinearity correction methods for commercial CCD array spectrometers used in field," in *AOPC 2017: Space Optics and Earth Imaging and Space Navigation, Proc. of SPIE* (International Society for Optics and Photonics, 2017), Vol. 10463, p. 104631K.
184. M. Talone and G. Zibordi, "Non-linear response of a class of hyper-spectral radiometers," Metrologia **55**, 747 (2018).
185. M. Talone and G. Zibordi, "Polarimetric characteristics of a class of hyperspectral radiometers," Appl. Opt. **55**, 10092–10104 (2016).
186. H. Kostkowski, *Reliable Spectroradiometry* (Spectroradiometry Consulting, 1997).
187. Y. Zong, S. W. Brown, B. C. Johnson, K. R. Lykke, and Y. Ohno, "Simple spectral stray light correction method for array spectroradiometers," Appl. Opt. **45**, 1111–1119 (2006).
188. Y. Zong, S. W. Brown, G. Meister, R. A. Barnes, and K. R. Lykke, "Characterization and correction of stray light in optical instruments," in (2007), Vol. 6744, pp. 67441L-67441L–11.
189. S. Nevas, G. Wübbeler, A. Sperling, C. Elster, and A. Teuber, "Simultaneous correction of bandpass and stray-light effects in array spectroradiometer data," Metrologia **49**, S43 (2012).
190. 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. **55**, 3966–3977 (2016).
191. S. Mekaoui and G. Zibordi, "Cosine error for a class of hyperspectral irradiance sensors," Metrologia **50**, 187 (2013).
192. R. R. Cordero, G. Seckmeyer, and F. Labbe, "Cosine error influence on ground-based spectral UV irradiance measurements," Metrologia **45**, 406 (2008).
193. 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, 2017).
194. "Spectral Response Function Data," https://sentinel.esa.int/web/sentinel/technical - guides/sentinel-3-olci/olci-instrument/spectral-response-function-data.
195. J. W. Müller, "Possible Advantages of a Robust Evaluation of Comparisons," J. Res. Natl. Inst. Stand. Technol. **105**, 551–555 (2000).
196. V. Vabson, I. Ansko, K. Alikas, J. Kuusk, R. Vendt, and A. Reinat, "Improving comparability of radiometric in situ measurements with Sentinel-3A/OLCI data," in (2017).
197. G. Zibordi, K. Ruddick, I. Ansko, M. Gerald, S. Kratzer, J. Icely, and A. Reinart, "In situ determination of the remote sensing reflectance: an inter-comparison," Ocean Sci. **8**, 567–586 (2012).
198. 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. **51**, 1167–1179 (2006).

199. K. Alikas, I. Ansko, V. Vabson, A. Ansper, K. Kangro, M. Randla, K. Uudenberg, M. Ligi, J. Kuusk, R. Randoja, E. Asuküll, B. Paavel, and A. Reinat, "Validation of Sentinel-3A/OLCI data over Estonian inland waters," in (2017).
200. K. J. Voss and L. B. da Costa, "Polarization properties of FEL lamps as applied to radiometric calibration," Appl. Opt. **55**, 8829–8832 (2016).
201. G. Bernhard and G. Seckmeyer, "Uncertainty of measurements of spectral solar UV irradiance," J. Geophys. Res. Atmospheres **104**, 14321–14345 (1999).
202. B. D. Santer, T. M. L. Wigley, J. S. Boyle, D. J. Gaffen, J. J. Hnilo, D. Nychka, D. E. Parker, and K. E. Taylor, "Statistical significance of trends and trend differences in layer-average atmospheric temperature time series," J. Geophys. Res. Atmospheres **105**, 7337–7356 (2000).
203. "Aerosol Robotic Network (AERONET) Homepage," https://aeronet.gsfc.nasa.gov/.
204. V. Vabson, J. Kuusk, I. Ansko, R. Vendt, K. Alikas, A. Ansper, M. Bresciani, H. Burmeister, M. Costa, D. D'Alimonte, G. Dall'Olmo, B. Damiri, T. Dinter, C. Giardino, K. Kangro, M. Ligi, B. Paavel, K. Ruddick, G. Tilstone, R. Van Dommelen, S. Wiegmann, C. Donlon, and T. Casal, "Laboratory intercomparison of radiometers used for satellite validation in the 400 - 900 nm range," Remote Sens. Special Issue "Fiducial Reference Measurements for Satellite Ocean Colour," (2019).
205. Sea-Bird Scientific, "Specifications for HyperOCR Radiometer," https://www.seabird.com/hyperspectral-radiometers/hyperocr-radiometer/family?productCategoryId=54627869935.
206. TriOS, "RAMSES Technische Spezifikationen," https://www.trios.de/ramses.html.
207. J. L. Mueller, G. S. Fargion, and Charles. R. McClain, eds., "Ocean optics protocols for satellite ocean color sensor validation: Instrument specifications, characterization, and calibration.," (2003).
208. 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*, C. J. D. and A. C. P. Giuseppe Zibordi, ed., Experimental Methods in the Physical Sciences, Vol. 47 (Academic Press, 2014), Vol. 47, pp. 247–304.
209. FRM4SOC, G. Tilstone, and et. al., *D-190, Technical Report TR-8, Protocols and Procedures for Field Inter-Comparisons of Fiducial Reference Measurement (FRM) Field Ocean Colour Radiometers (OCR) Used for Satellite Validation.* (2018).
210. FRM4SOC, G. Tilstone, G. Dall'Olmo, and R. Brewin, *D-200, FRM4SOC Field Inter-Comparison Exercise (FICE) Implementation Plan* (2016).
211. FRM4SOC, G. H. Tilstone, and et. al., *D-220, Technical Report TR-9, Results from the First FRM4SOC Field Inter-Comparison Experiment (FICE) of Ocean Colour Radiometers at the Atlantic Meridional Transect and the Acqua Alta Oceanographic Tower (AAOT)* (2018).
212. K. Alikas, V. Vabson, I. Ansko, G. H. Tilstone, G. Dall'Olmo, F. Nencioli, R. Vendt, C. Donlon, and T. Casal, "Comparison of Above-Water Seabird and TriOS Radiometers along an Atlantic Meridional Transect," Remote Sens. **12**, 1669 (2020).
213. FRM4SOC, G. H. Tilstone, and B. Carlton, *D-210, Field Inter-Comparison Experiment Database (FICE-DB)* (2018).
214. https://www.amt-uk.org/Home.
215. K. Ruddick, V. De Cauwer, and B. Van Mol, "Use of the near infrared similarity reflectance spectrum for the quality control of remote sensing data," in (2005), Vol. 5885, pp. 588501-588501–12.
216. B. B. Taylor, E. Torrecilla, A. Bernhardt, M. H. Taylor, I. Peeken, R. Röttgers, J. Piera, and A. Bracher, "Bio-optical provinces in the eastern Atlantic Ocean and their biogeographical relevance," Biogeosciences **8**, 3609–3629 (2011).
217. A. Bracher, M. H. Taylor, B. Taylor, T. Dinter, R. Röttgers, and F. Steinmetz, "Using empirical orthogonal functions derived from remote-sensing reflectance for the prediction of phytoplankton pigment concentrations," Ocean Sci. **11**, 139–158 (2015).

218. C. D. Mobley, "Estimation of the remote-sensing reflectance from above-surface measurements," Appl. Opt. **38**, 7442–7455 (1999).
219. M. Hieronymi, "Polarized reflectance and transmittance distribution functions of the ocean surface," Opt. Express **24**, A1045–A1068 (2016).
220. C. D. Mobley, "Polarized reflectance and transmittance properties of windblown sea surfaces," Appl. Opt. **54**, 4828–4849 (2015).
221. D. Stramski, R. A. Reynolds, M. Babin, S. Kaczmarek, M. R. Lewis, R. Röttgers, A. Sciandra, M. Stramska, M. S. Twardowski, B. A. Franz, and H. Claustre, "Relationships between the surface concentration of particulate organic carbon and optical properties in the eastern South Pacific and eastern Atlantic Oceans," Biogeosciences **5**, 171–201 (2008).
222. FRM4SOC, A. Bialek, and et. al., *D-180, Technical Report TR-7, Uncertainty Budgets of FRM4SOC Fiducial Reference Measurement (FRM) Ocean Colour Radiometer (OCR) Systems Used to Validate Satellite OCR Products* (2019).
223. A. Bialek, S. Douglas, J. Kuusk, I. Ansko, V. Vabson, R. Vendt, and T. Casal, "Example of Monte Carlo Method Uncertainty Evaluation for Above-Water Ocean Colour Radiometry," Remote Sens. **12**, 780 (2020).
224. "JCGM 101, Evaluation of measurement data - Supplement 1 to the “Guide to the expression of uncertainty in measurement” - Propagation of distributions using a Monte Carlo method. JCGM guidance document, First Edition, 2008. Available online: http://www.bipm.org/utils/common/documents/jcgm/JCGM_101_2008_E.pdf," (2008).
225. A. Białek, S. Douglas, J. Kuusk, I. Ansko, V. Vabson, R. Vendt, and and T. Casal, "Example of Monte Carlo Method Uncertainty Evaluation for Above-Water Ocean Colour Radiometry," Remote Sens. **12**, 780 (2020).
226. "Fidelity and uncertainty in climate data records from Earth Observations," http://www.fiduceo.eu/.
227. E. F. Vermote, D. Tanre, J. L. Deuze, M. Herman, and J.-J. Morcette, "Second Simulation of the Satellite Signal in the Solar Spectrum, 6S: an overview," IEEE Trans. Geosci. Remote Sens. **35**, 675–686 (1997).
228. S. Y. Kotchenova and E. F. Vermote, "Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part II. Homogeneous Lambertian and anisotropic surfaces," Appl. Opt. **46**, 4455–4464 (2007).
229. G. Zibordi and K. J. Voss, "Geometrical and spectral distribution of sky radiance: Comparison between simulations and field measurements," Remote Sens. Environ. **27**, 343–358 (1989).
230. C. A. Coombes and A. W. Harrison, "Calibration of a three-component angular distribution model of sky radiance," Atmosphere-Ocean **26**, 183–192 (1988).
231. FRM4SOC, *FRM4SOC International Workshop, 4-5 October 2018. Proceedings*. (2018).
232. A. C. Banks, R. Vendt, K. Alikas, A. Bialek, J. Kuusk, C. Lerebourg, K. Ruddick, G. Tilstone, V. Vabson, C. Donlon, and T. Casal, "Fiducial Reference Measurements for Satellite Ocean Colour (FRM4SOC)," Remote Sens. **12**, 1322 (2020).
233. EUMETSAT and ESA, "Sentinel-3 Validation Team (S3VT) meeting 2017," https://sentinel.esa.int/web/sentinel/events/-/article/sentinel-3-validation-team-meeting.
234. EUMETSAT and ESA, "Sentinel-3 Validation Team (S3VT) meeting 2018," https://www.eumetsat.int/website/home/News/ConferencesandEvents/DAT_3645214.html.
235. EUMETSAT and ESA, "Sentinel-3 Validation Team (S3VT) meeting 2019," https://s3vt.org.