SI Traceable Vicarious Calibration Test Sites and Methods

A. Białek

Submitted for the Degree of
Doctor of Philosophy
from the
University of Surrey

Surrey Space Centre
Faculty of Engineering and Physical Sciences
University of Surrey
Guildford, Surrey GU2 7XH, U.K.

August 2018

Summary

Earth Observation via satellite has been successfully used for several decades in many applications. Monitoring climate change is the most challenging one, as it requires highly accurate data to enable detection of small changes in naturally variable signals over different spacial and temporal scales. A measure used in metrology to assess the quality of the data is measurement uncertainty. However, to date, many satellite products still do not have uncertainties, the accuracy requirements are not defined precisely and even calibrations are performed without associated measurement uncertainty budgets. Thus is it often impossible to put an unbiased quality mark to the data that, by default, requires the highest levels of accuracy. This poses the risks of using poor quality data as the input to climate change models.

This research focuses on the "ground truth" measurement methodology called vicarious calibration. This is an independent post-launch satellite calibration technique based on a comparison of satellite readings with ground data and atmospheric modelling. Two test sites were selected as examples, land and ocean, to have uncertainty evaluated for their ground products following the Guide to the Expression of Uncertainty in Measurement (GUM) methodology.

A new radiometric calibration site, Gobabeb in the Namib Desert, was established for radiometric calibration of Top-of-Atmosphere (TOA) radiance/reflectance level 1 (L1) satellite products, and a campaign was conducted to measure the ground's reflectance. All instruments used during the initial characterisation were previously calibrated and characterised in optical laboratories. The *in situ* uncertainty budget was evaluated and validated by the comparison of the results to an alternative measurement source. The primary input of this research to the scientific community, apart from the new site, is a revised SI traceability chain for the ground reflectance field measurements. Hitherto, the reflectance reference standards used *in situ* had a calibration that did not match field illumination conditions. Although this problem was known, often it was not addressed or dealt with accurately. This study proposed a new field calibration value for the reflectance standard that combines direct and diffuse components weighted accordingly to the wavelength and atmospheric conditions during the measurement.

The work on the ocean site concentrated on the existing Boueé pour lacquisition de Séries Optiques á Long Terme (BOUSSOLE) site that is permanently deployed in the Ligurian Sea and provides Bottom of Atmosphere (BOA) water leaving radiance/reflectance level 2 (L2) Ocean Colour System Vicarious Calibration (SVC). This site had a preliminary uncertainty estimated as one generic number for all spectral channels and environmental conditions. A new uncertainty budget was developed by a detailed evaluation of each identified uncertainty component and these were combined by applying the Monte Carlo Method (MCM). As a result, a dynamic uncertainty evaluation for each measurement and the spectral band was produced addressing real measurement conditions and their effects on the quality of the relevant *in situ* products.

Key words: Vicarious calibration, uncertainty, SI traceability

Email: a.bialek@surrey.ac.uk
WWW: http://www.eps.surrey.ac.uk/

Acknowledgements

I would like to thank my employer, NPL, for providing the option of part-time funded studies, and to Dr Nigel Fox, the head of the Earth Observation, Climate and Optical Group, for approving my idea to do so. A big thank you to my colleagues who worked with me on the research relating to the thesis, and in particular, Claire Greenwell and Dr Maxime Lamare for their hard work and dedication into the Gobabeb site and its associated aspects. I would also like to extend my thanks to Dr Emma Woolliams for her support and understanding. I wish to acknowledge Dr Javier Gorroño, my PhD peer, for long hours of subject related conversations that often lead to interesting ideas and conclusions.

I would like to express my very great appreciation to my supervisor Professor Craig Underwood for his continuous efforts in reminding me to be an academic rather than a researcher delivering a project to a customer during my studies. I think that Karen Collar is probably the most helpful person in the world. Thank you, Karen, for your assistance with all the administrative related processes.

I would like to offer my special thanks to Professor David Antoine and Dr Vincenzo Vellucci for enabling me to work closely with them on ocean colour themes, and to all colleagues from LOV that I met during my secondment; this was a very fruitful time for me and I learnt a lot about oceanography.

This work was supported by the European Space Agency (ESA) funded ACTION project, and the European Union and European Metrology Programme through the European Metrology for Earth Observation and Climate Joint Research Project (MetEOC2). The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union.

Finally, I wish to thank my family for supporting me during all that time throughout the whole process.

Contents

List of Figures

List of Tables

Nomenclature

Roman Symbols

A	Area
Å	Ångström coefficient
c_{coef}	Calibration Coefficient
d	Ratio of the average and actual Sun−Earth distance
d	Fraction of the direct to the total radiant flux
E	Irradiance
f	Function
$f_r(\theta_i, \phi_i, \theta_r, \phi_r)$	BRDF
k	Coverage factor
K	Diffuse Attenuation Coefficient
L	Radiance
m	Optical air mass
n	Refraction index
Q	Radiant Energy
R	Reflectance Factor
r	Correlation coefficient
R^2	Coefficient of Determination
$\mathrm{R_{rs}}$	Remote Sensing Reflectance
t	Time
T	Temperature
u	Uncertainty

x	Input variable
y	Output variable
z	Depth

Greek Symbols

α	Absorbance
β	Radiance factor
∂	Partial derivative
Δ	Difference
λ	Wavelength
μ	Uncertainty
Ω	Projected solid angle
ω	Solid angle
Φ	Power
ϕ	Azimuth angle
π	Mathematical Constant 3.14
ρ	Reflectance
σ	Standard Deviation
τ	Transmittance
θ	Zenith angle

Subscripts

$_a$	atmospheric path
$_d$	downwelling
$_i$	incident
$_r$	reflected
$_s$	surface
$_u$	upwelling
$_w$	water-leaving
$_{WN}$	exact normalised water-leaving

Units

[Hz]	Hertz
$[J]$	Joule
$[m]$	meter
$[sr]$	steradian
$[W]$	Watt

Acronyms

6S	Second Simulation of a Satellite Signal in the Solar Spectrum
AERONET	AErosol RObotic NETwork
AERONET-OC	AERONET-Ocean Colour
AOD	Aerosol Optical Depth
AOE	Academy of Opto-Electronics
AOP	Apparent Optical Properties
ASTER	Advance Spaceborne Thermal Emission and Reflection radiometer
ASTM	American Society for Testing Materials
BIPM	Bureau International des Poids et Measures
BOA	Bottom of Atmosphere
BOUSSOLE	BOUée pour l'acquiSition d'une Série Optique á Long termE
BRDF	Bidirectional Reflectance Distribution Function
BRF	Bidirectional Reflectance Factor
CAS	Chinese Academy of Sciences
CCD	Charge-Coupled Device
CDOM	Coloured Dissolved Organic Matter
CEOS	Committee on Earth Observation Satellites
CHRF	Conical-Hemispherical Reflectance Factor
CLARREO	Climate Absolute Radiance and Refractivity Observatory
CNES	Centre National dEtudes Spatiales
CEOS-WGCV	CEOS Working Group on Calibration and Validation
EO	Earth Observation
ESA	European Space Agency

FOV	Field of View
GEO	Group on Earth Observations
GPS	Global Positioning System
GRASS	Gonio Radiometric Spectrometer System
GSFC	Goddard Space Flight Center
GSM	Global System for Mobile communications
GUM	Guide to the expression of Uncertainty in Measurement
GVR	Ground Viewing Radiometers
HDRF	Hemi-Directional Reflectance Factor
HPLC	High Precision Liquid Chromatography
InGaAs	Indium Gallium Arsenide
IOP	Inherent Optical Properties
SWIR	Short Wavelength Infra-red
ISO	International Organization for Standardization
IVOS	Infrared and Visible Optical Sensors
JRC	Joint Research Centre
L1	Level 1
L2	Level 2
LED	Light Emitting Diode
LES	Land Equipped Sites
libRadtran	library for Radiative transfer
MCM	Monte Carlo Methods
MERIS	Medium Resolution Imaging Spectrometer
MOBY	Marine Optical BuoY
MODIS	Moderate Resolution Imaging Spectrometer
MODTRAN	MODerate spectral resolution TRANsmittance
MOS	Marine Optical System
MSI	Multi Spectral Instrument
NASA	National Aeronautics and Space Administration

NIST	National Institute of Standards and Technology
NMI	National Measurements Institute
NOAA	National Oceanic and Atmospheric Administration
NPL	National Physical Laboratory
NRR	National Reference Reflectometer
OLCI	Ocean and Land Colour Instrument
OLI	Operational Land Imager
PACE	Plankton, Aerosol, Cloud ocean Ecosystem
PDF	Probability Distribution Function
PICS	Pseudo Invariant Calibration Site
QA4EO	Quality Assurance Framework for Earth Observation
RadCaTS	Radiometric Calibration Test Site
RadCalNet	Radiometric Calibration Network
ROSAS	RObotic Station for Atmosphere and Surface
RTC	Radiative Transfer Code
SAA	Solar Azimuth Angle
SeaPRISM	Sea Photometer Revision for Incident Surface Measurements
SEM	Scanning Electron Microscope
Si	Silicon
SI	Système International
SPOT	Satellite Pour l'Observation de la Terre
SRCA	Spectroradiometric Calibration Assemble
SRIPS	Spectral Radiance and Irradiance Primary Scales
SVC	System Vicarious Calibration
Sea-WiFS	Sea Viewing Wide Field of View Sensor
SZA	Solar Zenith Angle
TDR	Total Diffuse Reflectance
THRUTS	Traceable Radiometry Underpinning Terrestrial- and Helio- Studies
TOA	Top of Atmosphere

TSARS	Transfer Standard Absolute Radiance Source
UV	Ultraviolet
VAA	Viewing Azimuth Angle
VIIRS	Visible and Infrared Imaging Radiometer Suite
VIM	International Vocabulary of Metrology
VNIR	Visible and Near Infra-red
VZA	Viewing Zenith Angle
XRF	X-Ray Diffraction

Chapter 1

Introduction

Climate change is one of the most significant challenges for humankind in recent times and affects every human being. However, its impact and timescales, which are uncertain, is what leads to a vast number of debates and arguments at all levels of the community; starting with policy-makers and industrialists, followed by scientists, ecologists and finishing on a citizen personal view. Nevertheless, decisions about the mitigation of climate change are about to be taken.

Earth Observation (EO) via remote sensing provides the fundamental source of information necessary to understand and monitor climate change. Satellites continuously capture images of the Earths surface, collecting data that is fed into climate models to underpin our understanding of this change. The complexity of this phenomenon requires global Earth observations to be carried out at different timescales. Data from an individual satellite sensor needs to be merged with others, which have been acquired at different locations or times.

To ensure that the records from these various sources are compatible for merging they ought to have associated quality indicators that will contain information about their precision and accuracy. Quality assured data, ideally, should be linked with SI (Système International) and accompanied with uncertainty estimates. One of the methods to assure the quality of satellite products is their comparison to “ground truth”, which is the equivalent measurement at the Earth’s surface. The presented research focuses on test sites and methodology for such “ground truth” measurements.

1.1 Overview

All instruments require calibration to establish a relationship between their outputs and physical values that a particular device is built to measure. For optical sensors onboard satellites, this process links the digital numbers that result from the quantization of the signal detected by a sensor with the radiance values. However, in order to know the correct radiance values it is necessary to link them with a coherent system of measurement (SI). This linkage is called SI traceability and the official definition says: " Metrological traceability is 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" (JCGM200:2012 2012).

Pre-launch calibration is performed in various laboratories before the launch of a satellite sensor. Absolute radiometric calibration is conducted using a known source of spectral radiance that is traceable to SI units via one of the National Measurement Institutes (NMIs). Such a stable and calibrated light source would often be a big integrating sphere. Apart from the absolute calibration, characterisation tests such as spectral response, out-of-band spectral signal, spatial response and stray light levels, are also necessary. To mimic on-orbit conditions some of these tests, including the absolute calibration, take place in thermal vacuum conditions.

Some sensors, for example the Moderate Resolution Imaging Spectrometer (MODIS) (Salomonson et al. 1989), the Medium Resolution Imaging Spectrometer (MERIS) (Rast et al. 1999), the Sea Viewing Wide Field of View Sensor (SeaWiFS) (Hooker et al. 1992), the Operational Land Imager (OLI) (Knight & Kvaran 2014), the Multi Spectral Instrument (MSI) (Drusch et al. 2012) and the Ocean and Land Colour Instrument (OLCI) (Donlon et al. 2012), have built-in onboard calibration facilities. That allows for the detection of any changes occurred during launch and then monitoring long-term stability during the operational phase. However, they are potentially subject to changes as well.

Solar diffusers are most often used for onboard calibration. The solar irradiance reflected for the diffuser, which has known Bidirectional Reflectance Distribution Function (BRDF) properties, is a source of known radiance for calibration. The number of solar diffusers depends on the sensor design; some will have one diffuser on board (MSI) whereas others, up to three (MERIS, OLCI). If more than one diffuser is in use, then the second one will be a reference one used to monitor the stability of the working diffuser due to solar irradiance exposure. The third one, if present, detects the spectral changes.

Certain satellites have lamps as well as diffusers for onboard calibration, so in addition to solar irradiance, they carry their irradiance source (SeaWiFS, OLI). The MODIS sensor has an even more sophisticated device called Spectroradiometric Calibration Assemble (SRCA)(NASA) (Xiong & Barnes 2006) which contains lamps, an integrating sphere, and a monochromator and can be used simultaneously to the Earth's measurements.

Thus, in addition to the already mentioned calibration method, or as an alternative, "ground truth", also called vicarious calibration, is widely used to monitor in-flight satellite sensor performance from the ground.

For a wide range of commercial sensors that do not have onboard calibration capabilities (for example, Disaster Monitoring Constellation-2 (UK-DMC-2) (Lozano et al. 2012), Deimos-2 (Garca et al. 2008) and RapidEye (Tyc et al. 2005)) vicarious calibration is the sole method to monitor the sensor performance post-launch.

The word *"vicarious"* has its origin in Latin "vicārius" which means substituted. The dictionary definition of the word vicarious says *"performed, exercised, received, or suffered in place of another"* (*Dictionary*) and is often used in law as vicarious liability or punishment. The vicarious calibration term was introduced to mean *"in place of another"*, or in place of a laboratory calibration and was probably used, in this sense, for the first time in early eighties by Peter Koepke (Koepke 1982) as a method that allowed for absolute calibration of the satellite radiometers in orbit.

From the beginning, there were always two approaches to vicarious calibration.

1. Comparing radiometric counts of the satellites seeing the same target, where one of the sensors is a reference sensor.
2. Comparing radiometric counts of the satellite with radiances derived from ground measured (in-situ) data and atmospheric modelling.

Committee on Earth Observation Satellites (CEOS) defines: *"vicarious calibration refers to techniques that make use of natural or artificial sites on the surface of the Earth for the post-launch calibration of sensors. These targets are imaged in near-coincident fashion by the sensor to be calibrated and by one or more well-calibrated sensors from satellite or aircraft platforms or on the ground"* (CalVAlPortal).

New means of in-flight calibration have been proposed in missions, such as the Traceable Radiometry Underpinning Terrestrial- and Helio- Studies (THRUTS) (Fox et al. 2003) and Climate Absolute Radiance and Refractivity Observatory (CLARREO) (Wielicki et al. 2013), which will provide SI traceable measurements in space. Both missions aim to have a cryogenic radiometer as a primary radiometric standard on board, thus they will be able to redefine the radiometric scale on orbit post launch. The cryogenic radiometer is the most accurate radiometer and moreover less prompt to changes (see section 1.3.1 for more details). That ultimately becomes a standard in space which other satellites in orbit can be calibrated to. Both missions did not reach the commissions stage and are currently on hold. However, underpinning work on both of them continues while waiting for future funding that will allow for their progression into operation.

1.2 Research Motivation

EO data for climate change purposes requires the highest quality as only this allows distinguishing climate trends from the data variability. SI traceability is the most robust way to establish the quality of any measurement and is inseparable from the measurement uncertainty. Therefore, it is impossible to have SI traceability without an uncertainty budget for a given measurement. This research presents the practical implication of ensuring SI traceability for test sites that are then used as one of the methods to calibrate and validate satellite data.

Figure 1.1 graphically presents a link between the climate change, the EO data quality assurance and this research. The dotted arrow that connects the climate box directly with the satellite data box represents a weak link as the data, without additional processes represented by the remaining boxes, are not meaningful and trustful. Remote sensing data from numerous sensors cannot be used for climate change research directly.

The red arrow links the climate box with the Quality Assurance box because the data for these purposes are only applicable when they are quality assured. Quality assurance can be achieved by continuous calibration and validation of satellite sensors and products derived from them.

Calibrations and validations are performed at several stages of any mission and can be divided into pre-launch and post-launch. The latter can be branched further into on-board calibration, vicarious calibration and cross comparison. On the schematic (1.1) the pre-launch and on-board calibration boxes have blue frames because these will not be investigated during this research project. The vicarious calibration and cross calibration boxes are merged together as often a similar procedure is used in both of them. One could say that cross calibration is the next step that is performed after the vicarious calibration or that it uses the results of a single instrument vicarious calibration to conduct the cross comparison with other sensors. Both of them can be performed using several methods, such as test sites, Rayleigh scattering, sun glitter, clouds and the Moon.

The test site approach is examined further in-depth; however, the sites can be still subdivided into instrumented sites that are considered as SI traceable and pseudo-invariant calibration sites (PICS), which at present are not measured from the ground.

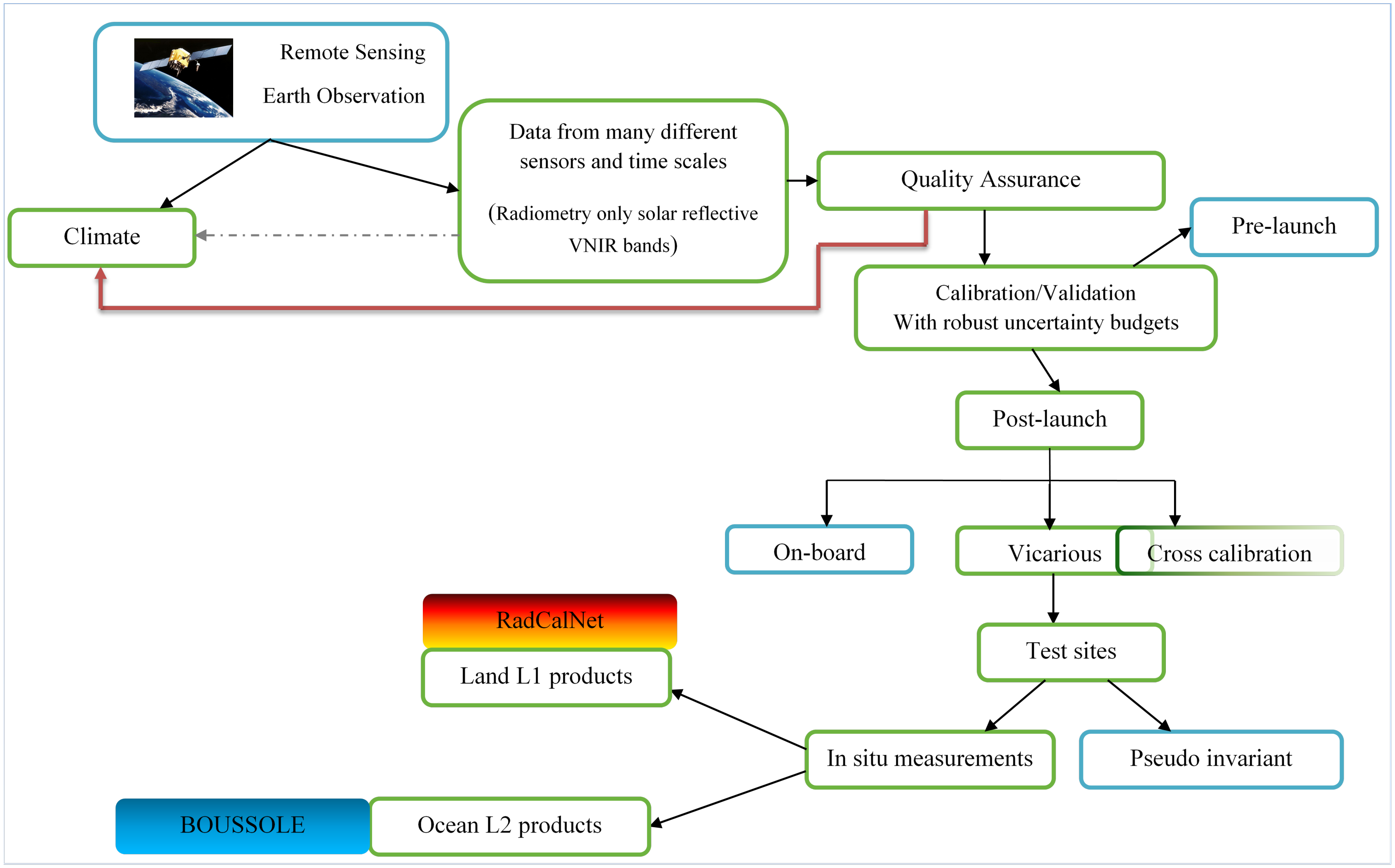


Figure 1.1: Schematic of the research project approach.

High reflectance sites like deserts are used for level 1 (L1) absolute radiometric calibrations, where top of atmosphere (TOA) radiance of a satellite is compared to the ground values propagated to TOA using a radiative transfer model. For an instrumented site the ground values are obtained from *in situ* measurements of the surface reflectance and atmosphere that are then fed into the Radiative Transfer Code (RTC). PICS do not have any inputs from the ground, thus the satellite data is used to predict the bottom of atmosphere (BOA) values and then the assumption of the site invariability enables comparison against repeated satellite acquisitions.

A slightly different approach is taken for an ocean site that is used for L2 ocean colour products. This is called System Vicarious Calibration (SVC) where the ground measurements are propagated to TAO using the same atmospheric inputs and models like the one used to derive L2 products. The atmospheric correction is seen as part of the system that is used to derive ocean colour L2 products and SVC provides vicarious adjustments to the whole system.

Both types of sites used for vicarious calibration of generic L1 and Ocean L2 products are used to ensure and monitor the quality of the satellite data. Even though vicarious calibrations have been commonly used for years, the uncertainty evaluation for both satellite and *in situ* products is often missing or is reported incompletely. Without robust uncertainty evaluation it is not possible to identity the biggest sources of error in the measurement process. Thus any improvements can be incorrectly focused on not significant elements rather than key uncertainty contributors. To improve quality of the satellite data and user confidence in them, especially for the climate change studies, every effort should be made to provide measurements with an appropriate quality level established by a link to SI and known uncertainty budget.

1.3 Definitions

Although remote sensing covers a broad range of measurements from different scientific areas, including fields such as optics, magnetism and acoustics, this particular research relates to optical radiation themes only; and in this section, the radiometric quantities are defined followed by the explanation of the SI traceability and uncertainty evaluation.

Radiometry is defined as the measurement of optical radiant energy. Optical radiant energy is a measurement of the energy of electromagnetic radiation. An electromagnetic waves energy is proportional to its squared amplitude, where the amplitude is the maximum field strength of the electric and magnetic fields in wave optics model. This is represented by the energy carried by a photon, that energy is proportional to the photons electromagnetic frequency in the quantum optics model.

Wavelengths from 10 nm to 1000 μm are defined as the optical part of the electromagnetic spectrum. This optical range is divided into the ultraviolet (UV), visible and near infrared (VNIR) and short wavelength infrared (SWIR) regions. The spectral range of interest for the research presented in this document includes the visible and near infrared region with wavelengths ranging from 400 nm to 1700 nm.

Table 1.1 presents a list of radiometric quantities. The wavelength dependence is not included in the equations for for sake of brevity. Irradiance, radiance and reflectance are of particular importance for this research and will frequently be used in this thesis.

Table 1.1: Radiometric quantities (Palmer & Grant. 2010).

Radiometric quantity	Equation and units	Definition
Radiant Energy	Q [J]	
Radiant Power (radiant flux)	$\Phi = \dfrac{dQ}{dt}$ [W]	Energy per unit time
Irradiance (radiance incidence)	$E = \dfrac{d\Phi}{dA_s}$ [Wm^{-2}]	Power per unit area that is incident on a surface Irradiance is measured at the detector
Solid angle	ω $[sr]$	The plane-angle concept extended to three-dimension
Radiance	$L = \dfrac{d^2\Phi}{dA_s d\Omega}$ [Wm^{-2}sr^{-1}]	Power per unit area and per unit projected solid angle
Radiometric properties of materials		
Reflectance	$\rho = \dfrac{\Phi_r}{\Phi_i}$	Ratio of reflected power to incident power
BRDF	$f_r(\theta_i, \phi_i, \theta_r, \phi_r) \equiv \dfrac{dL_r(\theta_r, \phi_r)}{dE_i(\theta_i, \phi_i)} = \dfrac{dL_r(\theta_r, \phi_r)}{L_i(\theta_i, \phi_i) d\Omega}$ [sr^{-1}]	Differential element of reflected radiance in a specified direction per unit differential element of irradiance, also in specified direction
Transmittance	$\tau = \dfrac{\Phi_t}{\Phi_i}$	Ratio of the total transmitted flux to the total incident flux
Absorbance	$\alpha = \dfrac{\Phi_a}{\Phi_i}$	Ration of absorbed power to incident power

1.3.1 Radiometric Scales

SI traceability is obtained by an unbroken chain of calibration or comparison against reference standards up to the primary standard that is realised at the National Measurements Institutes, for example, NPL in the UK. Bureau International des Poids et Measures (BIPM) (*BIPM* 2018), is an international coordinating organisation of the intentional system of measurements (SI).

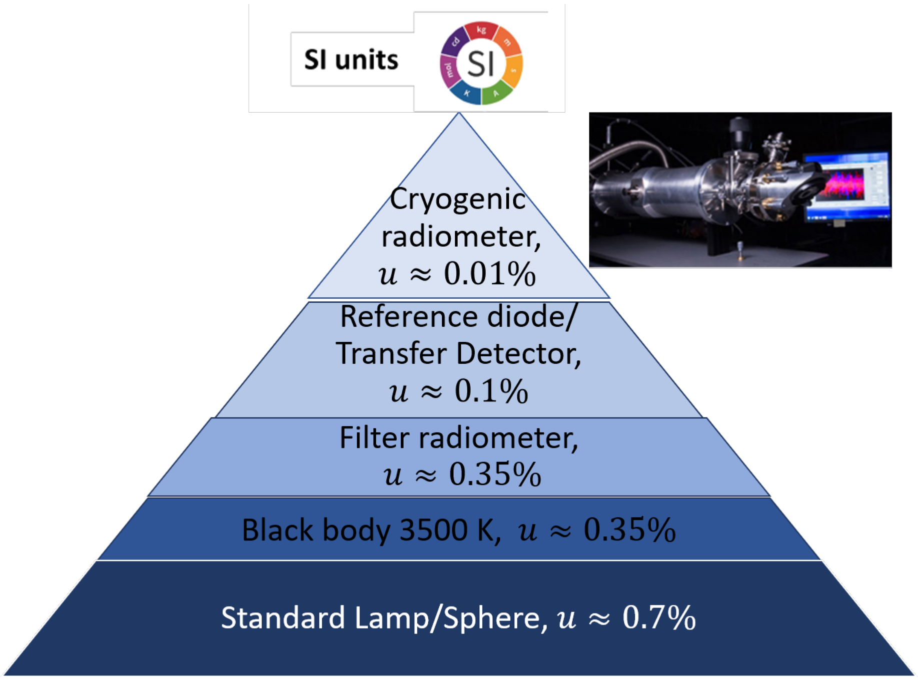


Figure 1.2: Schematic of NPL SI traceability chain for radiometric spectral scale u represents typical uncertainty values, (k=2).

The NPL SI traceability chain for the radiometric spectral scale is shown in Figure 1.2 where the cryogenic radiometer forms a primary standard and a direct link to SI units. Then at each level the scale is transferred down the chain, but measurement uncertainty increases at each step down the chain.

The principle of cryogenic radiometry is based on an old technology called electrical substitution radiometry (Gillham 1962), thus providing an independent means of defining optical power by direct comparison to electrical power.

Electrical substitution radiometry measurement consists of a black plate that is placed in front of a copper plate and when one illuminates this plate doublet, the black plate absorbs the optical power and causes a rise in the temperature of the copper plate. This temperature rise is measured, the light is then switched off, and the same temperature rise is obtained by electrical heating of the copper plate.

Thus,

$$\Delta T = \Delta T_{optical} = \Delta T_{electrical}, \tag{1.1}$$

and then power

$$\Phi_{optical} = \Phi_{electical}. \tag{1.2}$$

The knowledge of the electrical power that was used to create the same rise in the temperature defines the optical power.

For cryogenic radiometry, instead of a black plate, an absorbing cavity (to eliminate any reflection from the black plate) is used, and the back wall of this cavity is cooled to the cryogenic temperature (to reduce thermal noise), thus making the same measurement 1000 times more accurate (Martin et al. 1985). Cryogenic radiometers provide a direct link to an SI derived unit (Watt [W]) and create a basis for a radiometric optical scale traceability chain.

The second step in the chain uses a cryogenic radiometer to calibrate a trap detector (a photodiode-based standard of spectral responsivity). A stable laser source is measured by the cryogenic radiometer to establish the absolute optical power; then the same source is measured by a trap detector. The calibration from the trap detector is passed to a filter radiometer that is used to measure the radiance of a high-temperature blackbody at a given wavelength which allows it to determinate the temperature of the blackbody. The spectral radiance of the blackbody is derived from the knowledge of its temperature and Plancks law. The blackbody is used to transfer the spectral scale to the lamp and integrating sphere sources using the Spectral Radiance and Irradiance Primary Scales (SRIPS) facility (Woolliams et al. 2006).

The NPL reflectance scale is realised on the National Reference Reflectometer (NRR) facility (Chunnilall et al. 2003, Williams 1999) which uses the total flux method.

All goniometric optical scatter measurements have to be measured according to the standard practice ASTM E2387 (*ASTM E2387-05(2011)* 2011). There are three measurement set-ups for bidirectional reflectance factor (BRF) measurements, all consistent with each other, and between different NMIs within uncertainties (Yoon et al. 2009). These are total flux mode, radiance mode and irradiance mode. NMIs across the world make sure that their measurements are all in agreement and a set of comparisons between them defines the absolute reflectance scales. It is important to note that commonly used in remote sensing BRF expression in the metrology world is identified as R reflectance factor or β radiance factor, that under clearly specified angular conditions is the same. The section Reflectance Terminology in chapter 3 contains detailed information about this.

One of the standard measurement geometries for R and β is defined as 0:45, where the illumination is at 0 degrees relative to the sample normal, and the viewing angle is at 45 degrees. The azimuth term is omitted as it is undefined for zero angle illumination.

$$BRF(\theta_i, \theta_r, \lambda) = R(\theta_i, \theta_r, \lambda) = \beta(\theta_i, \theta_r, \lambda) = \lim_{\Omega \to 0} \frac{\pi \Phi_r(\theta_i, \theta_r, \lambda)}{\Omega \Phi_i(\lambda) \theta_r} \tag{1.3}$$

Where: $\Phi_i(\lambda)$ is total incident power (radiant flux) at a given wavelength, $\Phi_r(\theta_i, \theta_r, \lambda)$ is the reflected flux at the given direction and wavelength, Ω is the solid angle subtended by the detector aperture at the centre of the sample, and θ_r is the detection angle (with respect to the normal of the sample).

The NRR allows for measurements in the principal axis only; therefore, equation 1.3 does not contain the azimuth angle configuration as these are always the same, with relative azimuth equal to 180° or 0°.

1.3.2 Uncertainty Evaluation

The importance of uncertainty analysis and its presence in everyday life is clearly presented in (Willink & White 2011) using a biscuit manufacturer and its product weight as an example. The manufacturer ought to know precisely the uncertainty associated with its biscuits package weight to ensure that it will not sell underweight packs, and thus will not be prosecuted by a regulatory organisation, but at the same time does not want to systematically put on too much weight and lose profit. The same publication says:

"Effective cost-risk compromises can only be reached if those making the decisions have realistic (i.e., not optimistic or conservative) estimates of the uncertainties in the measurements on which the decisions are based." and then:

"uncertainty analysis is a tool for helping people manage the risks and costs associated with measurement-based decisions. To be useful and meaningful, the uncertainties reported on test reports and calibration certificates must accurately characterize real-world objects, processes and measuring instruments. That is, realistic estimates of uncertainty are the goal of uncertainty analysis, and our economies, our environment, and our lives depend on it."

Thus, on that simple example, the idea of uncertainty as a doubt around the measurement is explained by the estimate of the biscuits packaged weight and its implication to being under- or overweighted in the light of incorrect uncertainty estimations. This is an important aspect of the uncertainty analysis that it has to be realistic. In some ways it is easier to detect the under-estimated uncertainty by validation and intercomparison with other means. The overestimation so-called "conservative" approach might be harder to detect as the comparison results will agree.

In the field of metrology, a measurement is incomplete without a quality statement. The uncertainty is a relatively new concept that was initialised in the 1980s of the last century and replaced an old and well-known error analysis system where random errors were added in quadrature and systematic linearly. In the new approach all known sources of errors ought to be corrected, however, it is recognised that that correction is not perfect and there is a residual uncertainty associated with that correction, that needs to be propagated. There is no more linear addition of the systematic errors, all components are added in quadrature.

To have a universal method to calculate uncertainties is similar to having a universal way to define SI units; the basis of an uncertainty evaluation was defined in the GUM - Guide to the Expression of Uncertainty in measurement (JCGM100:2008). This document and its supplements state mathematical and statistical rules and definitions of the uncertainty evaluation and propagation through a measurement process.

Basic uncertainty terms from the GUM are briefly described here. A measurement is performed to determine the value of a measurand that is a particular quantity (i.e. radiance). The result of the measurement is only an estimate of the value of measurand and can be considered as completed when quoted with an uncertainty of that estimate. The term uncertainty is defined as a doubt around the estimate and ought to be quoted with an appropriate coverage factor that defines how confident one can be about our best estimate.

For example, a coverage factor of (k=1) or 1σ, defined as one standard deviation from the mean assuming a normal distribution function, expresses the confidence level at 67% that the estimated value is within its quoted uncertainty. Most of the measurements performed at NMI use coverage factor (k=2) which is equivalent to 95% of the measured values that will be within a quoted uncertainty range of the measurand. Some fields of science use a (k=3) coverage factor that is defined as 99% and is mainly for risk management and medical application where 99% confidence is essential for life-saving purposes.

Currently, uncertainty budgets for many Earth Observation measurements, including land and ocean *in situ* measurements, tend not to mention a coverage factor which makes the quoted uncertainty value incomplete.

The GUM defines Type A and B uncertainty evaluations. Type A is related to the evaluation of the uncertainty based on a statistical series of observations. Type B is used for all other means of uncertainty estimation that are not derived from statistics. Thus, the source of Type B uncertainty estimation might be a calibration certificate of an instrument or knowledge of detector sensitivity to temperature change.

The GUM stresses the difference between an error and uncertainty. These two terms refer to different concepts. An error is an imperfection in the measurement process and cannot be known precisely, but its effect on a measurement result can be estimated and expressed as uncertainty on the result due to this effect. An error has random and systematic components and can have additive or multiplicative natures. Random errors can be reduced by an increased number of measurements, whereas systematic errors are by application of a correction factor. There will always be uncertainty associated with systematic error corrections.

To derive a measurement output uncertainty all the individual input uncertainty components have to be established first and then combined according to the law of propagation of uncertainty. That is based on the Taylor series approximation given by equation 1.4:

$$u_c^2(y) = \sum_{i=1}^{N} \left(\frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j) \tag{1.4}$$

Where: $y = f(x_1, x_2, ..., x_n)$ is the output value and is a function of input parameters, $\frac{\partial f}{\partial x_i}$ are partial derivatives often called sensitivity coefficients, and $u(x_i)$ is a standard uncertainty of an input component (note this can be either uncertainty related to random or uncertainty on systematic effects correction). The second part of this equation is needed only if the input quantities are correlated; then the term $u(x_i, x_j)$ can be replaced by $u(x_i)u(x_j)r(x_i, x_j)$ where $r(x_i, x_j)$ is the correlation coefficient.

There are two methods to report uncertainty: absolute, where uncertainty is quoted in the unit of the measurand, for example the weight of the biscuits packaged in 200 g units with an uncertainty of 2 g, (k=2); whereas relative uncertainty is expressed as a percentage, so the same weight of 20 g has 1%, (k=2) uncertainty. Both methods are interchangeable and have exactly the same meaning. This is important to note that uncertainly is related to the unique measurand value, GUM point 1.2 says:

"This Guide is primarily concerned with the expression of uncertainty in the measurement of a well-defined physical quantity the measurand that can be characterized by an essentially unique value. If the phenomenon of interest can be represented only

as a distribution of values or is dependent on one or more parameters, such a time, then the measurands required for it description are the set of quantities describing that distribution or that dependence."

A given uncertainty value is assigned to the measurand value and cannot be transferred to other value of that measurand that was obtained at different time and at different conditions. A new uncertainty has to be evaluated in that case.

Uncertainty also defines the significant numbers that should be quoted with a value for the given example where is no point in quoting the weight as 200.245 g as the numbers after the stop do not provide any meaningful information because 2 g is the uncertainty value. Thus, the result is somewhere between the 198 g and 202 g.

The analytical method can become difficult to apply on 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 (JCGM101:2008). MCM is a numerical method that requires a distinct probability distribution function (PDF) of all input components; if input components are correlated then the joint probability function and the measurement equation are required. The MCM will then run a large number of numerical calculations of the measurement equation iteratively randomly choosing the input from the available range that is defined by the probability density function. The large number of output values calculated using different input values at each iteration provides the uncertainty of the output value with its PDF.

The analytical GUM approach is well suited for the laboratory-based part of this research. Thus, absolute radiometric calibration for any instrument that will be used is going to be accompanied with the uncertainty evaluation that follows the original GUM. However, the MCM approach offers an interesting alternative for a complex system with many inputs unknown or not easily convertible to a normal probability distribution. Thus, this method will be used for *in situ* measurements.

1.4 Research Objectives

This research aimed to establish an optimum strategy to transfer SI traceability from the measurement laboratory to test sites used for vicarious calibration and to determine their uncertainty limits. Detailed knowledge of *in situ* measurement uncertainty allows focus on the areas to improve in order to make these measurements more accurate in the future.

The following milestones ensure that the project meets the objectives:

1. Define SI traceability of a new radiometric calibration land site;
2. Reference standards and instruments calibration and preparation for the new site;
3. *in situ* characterisation measurements and their uncertainty evaluation;
4. An Ocean Colour site *in situ* uncertainty budget.

The Earth Observation community expects uncertainty per pixel that will come with the satellite products from a new generation of operational satellites (e.g. Sentinel series). To achieve this goal, uncertainty estimations are necessary from all calibration processes that are used to derive satellite data. This includes the vicarious calibration uncertainty that at present does not exist in a fully validated form.

An additional aim, that drives this research project, is to merge the world of metrology with the world of remote sensing community. A metrologist can fully understand the measurement process that is performed in controlled laboratory conditions and has means of testing a particular characteristic while keeping others at a constant level. This allows to establish a robust uncertainty budget that can be validated during inter-comparison with other laboratories. Remote sensing community measurements are performed in rough and dynamically changing environmental conditions and, due to their complexity, the guidance for uncertainty evaluation is not always followed. This thesis will form an example of applying metrology and in particular the uncertainty evaluation for these types of measurements.

1.5 Research Novelty

A new land test site in Gobabeb, Namibia was developed during this research. The site is now operational and is part of the prototype RadCalNet network of absolutely calibrated test sites. The site establishment was a joint effort of scientists from NPL, Magellium and CNES. The author was responsible for the NPL technical aspects of the research that comprised SI traceability for the new site and *in situ* site characterisation measurements, including multi-angular ground reflectance characterisation. The author personally worked on SI traceability and uncertainty budget aspects and supervised the work on multi-angular ground reflectance.

As a result of this research, the author proposed a novel method of using *in situ* reflectance standard to better match the real illumination conditions. Now, the RadCalNet best practice guidelines for new sites wanting to join the network contain this recommendation.

An uncertainty statement accompanies the site surface reflectance initial site characterisation measurements. The agreement of the measurement results comparison between two independent teams NPL and CNES validated this uncertainty. Thus, the well-known by the metrology community SI traceability concept, that starts from the traceability chain to the SI unit followed by measurements with uncertainty evaluations and finishes by its validation such as an independent comparison, is set here as a new example for the remote sensing community.

The Ocean Colour part of this research focused on an uncertainty budget for the radiometric products of the BOUSSOLE site (the European ocean colour site for system vicarious calibration of Sentinel 3 OC products). The author worked closely with a team of scientists that routinely operate the BOUSSOLE and provided a novel framework to evaluate *in situ* uncertainty.

This new uncertainty budget was developed using MCM and provides a novel concept of “dynamic” uncertainty assigned to an individual measurement from the BOUSSOLE site rather than one generic number. The condition on the site can change from one hour to another, and between different days. Thus, for the first time, the appropriate uncertainty value for a given measurement was evaluated, which provides a more accurate quality indicator. This will, in future, allow adding weight to an SVC process depending on the *in situ* measurement quality.

Both uncertainty budgets from land site characterisation and the operational ocean site form a basis to the uncertainty evaluation for other test sites.

The dissemination of the outcomes of this research included several conferences: SPIE Remote Sensing 2015 (Greenwell, Bialek, Marks, Woolliams, Berthelot, Meygret, Marcq, Bouvet & Fox 2015), ESA Living Planet Symposium 2016 (Lamare, Bialek, Greenwell, Woolliams, Lacherade, Marcq, Meygret, Bouvet, King & Fox. 2016), IGARSS 2016 (Bialek, Greenwell, Lamare, Meygret, Marcq, Lacherade, Woolliams, Berthelot, Bouvet, King, Underwood & Fox 2016a) and Ocean Optic 2016 (Bialek, Vellucci, Gentili, Antoine, Fox & Underwood 2016b).

1.6 Structure of the Thesis

The thesis consists of four chapters. The first chapter comprises background information and a literature review with a dedicated section to the reflectance terminology that clearly underlines the challenges in *in situ* reflectance SI traceability. Historical and current land and ocean test sites and networks of sites are described with methodologies for ground measurements.

Afterwards, the land and the ocean aspects both have dedicated chapters. Therefore, Chapter 3 focuses on the land test site and outlines the process of establishing a new radiometric calibration test site. Instrument calibration and characterisation in preparation for field measurements and test site characterisation are presented in this chapter. In particular, attention is brought to the SI traceability of ground reflectance measurement. The results with uncertainty estimates are demonstrated in the final part of this chapter.

Chapter 4 is dedicated to the European ocean site. The site is already operational; thus, the main focus is on its products uncertainty budget evaluation. In this chapter, the methodology of measurements and data processing is presented followed by a novel proposed uncertainty evaluation framework. Then all uncertainty components are identified, quantified and integrated into a model that allows evaluation of the uncertainty of the products from this site.

The final chapter contains a summary of the main findings and concludes from the research presented in this thesis. A plan for the future work that would benefit the area further is presented as are the closing remarks.

Chapter 2

Background and Literature Review

This chapter describes the methodologies used for the ground measurements of land and ocean including detailed descriptions of the reflectance terminology, followed by an explanation of the standard test site concept and a presentation of some such test sites. In particular, it highlights the evolution of *in situ* measurements towards permanent and autonomous instrumentation with associated benefits and challenges.

Then the atmospheric measurement and radiative transfer modelling are mentioned for the completeness of VC processes. However, the research does not examine these subjects further.

The methodology gaps in the ground measurements include a lack of correct ways of using reflectance standards *in situ*, evaluation of robust uncertainty budgets and a comparison of *in situ* results as a final step to validate them.

2.1 Introduction

Radiometric calibration is an essential step in obtaining accurate and meaningful physical measurements of the reflectivity of the Earth from satellite remote sensing data. With satellite products being derived from a multitude of sensors that have different spatial and spectral characteristics, cross-calibration between satellite sensors is necessary to provide a consistent long-term time series from multiple platforms. Furthermore, continuous calibration of the radiometric response of the sensors is needed to characterise temporal changes over the lifetime of the sensors.

Calibration of remote sensing instruments can be performed at three different stages (Dinguirard & Slater 1999, Teillet et al. 2001a): at the pre-flight stage; post-launch, with an on-board calibration system; or vicariously, using natural targets on the surface of the Earth. The latter is an ideal method for sensor-to-sensor, multi-temporal or multi-angular comparisons, and has been shown to work for sensors without on-board calibration systems (Dinguirard & Slater 1999, Govaerts & Clerici 2004).

Deserts are frequently used as vicarious calibration targets e.g. (Cabot et al. 2000, Govaerts & Clerici 2004, Heidinger et al. 2003, Holben et al. 1990, Kaufman & Holben 1993, Lacherade et al. 2013, Mishra et al. 2014, Staylor 1990, Teillet et al. 2001a), as they typically provide high reflectance, spatial uniformity, temporal stability and a low probability of clouds (Cosnefroy et al. 1996). However, knowledge of the reflectance of selected calibration targets either relies on radiative-transfer modelling based on lack of field measurements for PICS (Govaerts & Clerici 2004), or requires regular field campaigns for vicarious calibration sites.

PICS sites are generally inaccessible due to their remote location, such as in the middle of the desert or in some cases due to a political situation. For example, one of the most temporally stable and spectrally flat sites is in Libya. These sites are selected using satellite data; however, exactly the same characteristics as for instrumented sites are the key factors to define a good pseudo-invariant site. Thus, they exhibit high reflectance that is often spectrally flat and good temporal stability which allows the assumption that these sides are invariant. They can be used for sensor to sensor TOA comparison without an option of knowing the ground truth data as their characteristics are measured from space only.

Ocean colour measurements from space are extremely valuable for global monitoring of the marine biosphere. However, they are exceptionally challenging due to a small proportion of the marine signal in the total signal being measured by any satellite sensor. *In situ* ocean colour radiometry has been used for vicarious calibrations of satellite ocean colour missions since the successful post-launch calibration program (McCain et al. 2006, McClain et al. 1992) of the Sea Viewing Wide Field of View Sensor (Sea-WiFS) (Hooker et al. 1992). This is a special case of vicarious calibration (Franz et al. 2007, Gordon 1998, 1997) that provides a gain to the overall response of the sensor and the atmospheric correction algorithm, the so-called system vicarious calibration (SVC). At present SVC is the only available method to verify and maintain the uncertainty of satellite derived ocean colour products and has been applied for all consecutive ocean colour sensors including MODIS-A (Salomonson et al. 1989), MERIS (Rast et al. 1999), the Visible and Infrared Imaging Radiometer Suite (VIIRS) (Murphy et al. 2006), OLCI (Donlon et al. 2012) and is planned for the Plankton, Aerosol, Cloud ocean Ecosystem (PACE) (*NASA PACE*) mission.

2.2 Land Measurement Methodology

Historically, test site based vicarious calibration had two methods: reflectance and radiance-based (Slater et al. 1987). The radiance method was in use extensively at the beginning of the vicarious calibration “era” and involved measurements of the ground radiance from a plane at a height of at least 3 km to include the aerosol scattering. As the quantity of interest here was TOA radiance, the measurements at the high attitude contained already the signal with an atmospheric contribution. The instrument had to be calibrated and the calculation of TAO radiance was corrected for the residual scattering and absorption above the radiometer.

This method, classified initially as more accurate than the direct ground measurements (Dinguirard & Slater 1999, Slater et al. 1995), tended to be used less often over the years. This was due, firstly, to the costs of these measurements and complications in campaign organisation and, secondly, with the advances in RTC and atmosphere measurements from the ground, these measurements can now be more accurately propagated to the TOA. Finally, with the recent move towards permanently operated sites, the radiance method in its original form is not fit for purpose.

However, an adaptation of the radiance method is now used for permanent ground measurements, where radiometers measure radiance, and additional measurements are made of total downwelling irradiance to calculate ground reflectance.

The reflectance method is widely utilised and is based on the relative comparison measurement method to a reflectance standard. Thus, the instrument does not require, as such, absolute radiometric calibration. Nowadays, a well-known reference standard is Spectralon (a brand name that is a registered trademark of Labsphere (*Labshpere* 2018)), although other materials are available. In the early 1980s barium sulfate ($BaSO_4$) or pressed polytetrafluoroethylene called halon, TFE or PTFE were used (Biggar et al. 1988). It is worth noting that using "pre-spectalon" standards always included BRF calibration. Hence, the standard reflectance factor was matched to the SZA. Regrettably, with the introduction of almost "perfect Lambertian" material (i.e. Spectralon) it seems that the community started to use one reflectance value for the standard rather than its BRF. The error introduced by using a single value rather than BRF will vary depending on the SZA during the measurements but can reach even 5% for SZA of 25° (see Figure 3.3 in Chapter 3 for details).

A good standard needs to be calibrated and has the properties of an optically diffusing material, generally referred to as highly Lambertian. Thus, all ground measurements are ratioed to the measurements of the standard done at the same illumination and viewing geometry. This ratio is then multiplied by the known reflectance factor value of the standard to obtain an equivalent value for the measured ground. The surface reflectance is calculated according to equation 2.1.

$$\rho_{\text{ground}} = \frac{DN_{\text{ground}}}{DN_{\text{reference}}} \rho_{\text{reference}} \tag{2.1}$$

Where: ρ represents reflectance and DN is a digital number respectively for ground and reference panel. This is a simplified version of the equation as the angular representation as well as wavelength are omitted for clarity at this stage.

Sometimes, an average of two, before and after ground, panel readings are used in the denominator to average out any changes in illumination conditions occurring during the measurements. Alternatively simultaneous measurements of sun downwelling irradiance are used to normalise all of the readings for any short time changes in illumination conditions.

There was a third method that was an adaptation of the reflectance method called irradiance or the improved reflectance-based method (Biggar et al. 1991, 1990, Slater et al. 1995). That, in addition to the reflectance method, included the measurements of the ratio of diffuse to global downwelling irradiance. The measurements of the diffuse to global were performed in the reflectance mode as shown in Figure 2.1, where the diffuse component was estimated by shadowing the direct sun. The advantage of this method aimed to reduce the effects of various assumptions in the atmospheric inputs to RTC. THe irradiance method similarly to the original radiance method is not commonly used. The same reasons apply with the advances in atmospheric measurements and RTC the inputs obtained from *in situ* sun photometer measurements are used to extract the diffuse to the global ratio more accurately without the need for additional measurements. However, this is important to note that current RTCs do not have uncertainly estimation. The sensitivity studies and comparisons between then showed that their outputs agree with each other (see section 2.6 for details).

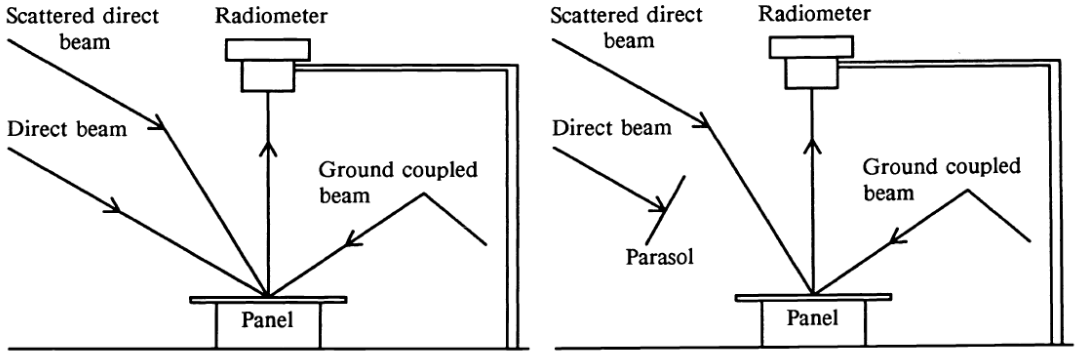


Figure 2.1: Schematic of the diffuse to global measurement from (Biggar et al. 1991).

At present, measurements on land sites include ground reflectance or radiance, solar irradiance and sky radiance. The vast majority of sun and sky measurements are performed as part of atmospheric measurements by a sun photometer. The main quantity measured *in situ* during a field campaign is the surface reflectance at nadir. Ground measurements are typically done using a spectroradiometer. For the last forty years instruments used *in situ* have evolved from multispectral to hyperspectral with advances in technology.

Currently, the majority of measurements are hyperspectral and are performed using, for example, the ASD FieldSpec spectroradiometer (MalvernPanalytical) (*ASD FieldSpec* 2018) 350 nm to 2500 nm interpolated to 1 nm steps. The actual spectral resolution might vary depending on the model of the instrument but the NPL owned ASD version has 3 nm for VNIR and 10 nm for SWIR spectral regions. The measurements are done in a relative mode using as a reference the Spectralon (*Labshpere* 2018) panel. The reference panel is measured and then several points at the ground are measured before the next reference measurement is taken.

The instruments are hand-held, sometimes attached to a trailer or even a car or mounted into some more permanent construction like special stands or tripods. The important factor is to carefully position an operator or a stand to minimise the shading effects on the measurement surface. GPS is connected to the instrument so that each point is localised.

However, there is a slight issue with the calibration of reflectance values for the standards and the following section presents an explanation of the reflectance terminology. Chapter 3 then details the reflectance standard calibration.

2.2.1 Reflectance Terminology

A quantity, commonly called “ground reflectance”, is often used to describe in-situ ground truth measurements. However, without a clear definition of illumination and viewing geometries this term is misleading and might lead to a misinterpretation of the results. Not only is there a discrepancy between clearly defined laboratory definitions and those in use by the remote sensing community, there is not always agreement within the community either (Schaepman-Strub et al. 2006). Nevertheless, the terms

BRDF and hemi-directional reflectance factor (HDRF) are used relatively loosely in the literature, not always referring to the same physical quantities measured *in situ*.

The definition of BRDF was first introduced by F.E. Nicodemus (Nicodemus et al. 1977). The nine kinds of reflectance quantities were defined in relationship to BRDF by integrating it and averaging over different incoming and reflected beam configurations, including directional ($d\omega$), conical (ω), and full hemispherical geometries (2π). Nine types were obtained from a combination of possible incident-reflected beam pairs. The geometrical configuration behind these definition is presented in Figure 2.2.

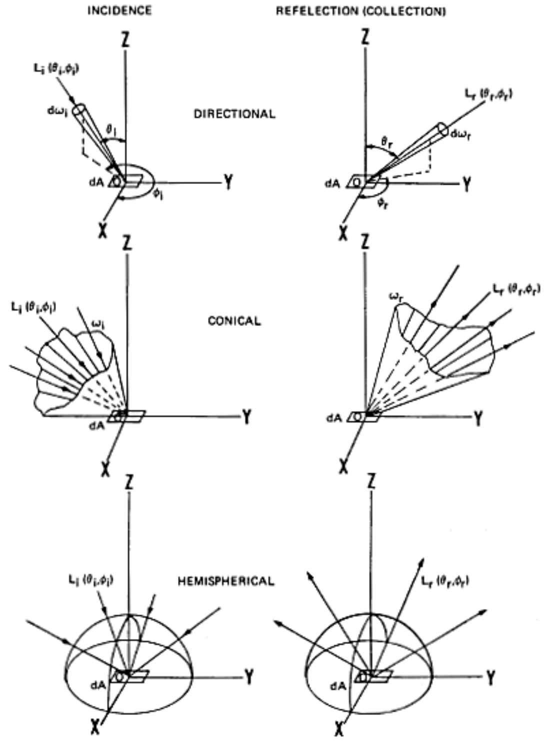
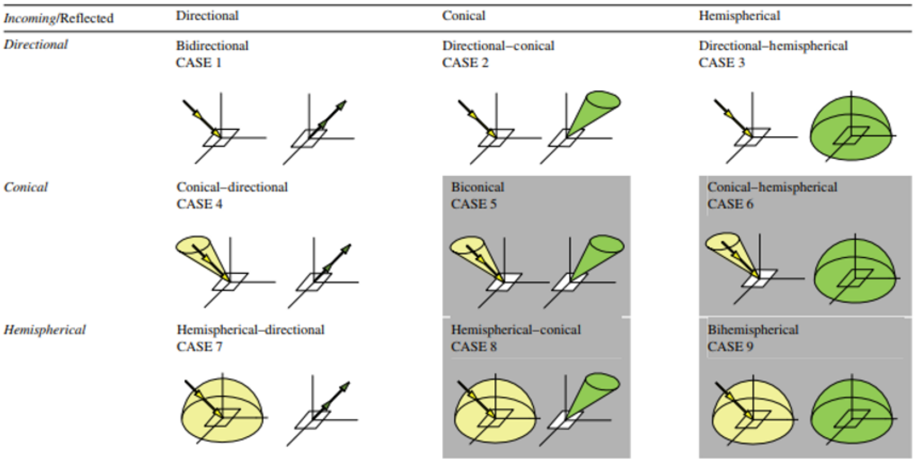


Figure 2.2: Classification for denoting incident and reflected beam geometry from (Nicodemus et al. 1977).

The nine reflectance definitions were developed theoretically and it is possible to create specified illuminations and viewing conditions in a laboratory. However, none of them takes into consideration the issues related to practical remote sensing measurements, such as the presence of both direct and diffuse illumination. Thus, Schaepman-Strub (Schaepman-Strub et al. 2006) proposed an updated set of definitions that allow accounting for more complex illumination configurations, these are presented in Figure 2.3. However, these remote sensing definitions are not supported by clearly defined methods for measuring these quantities. BRDF, according to Nicodemus, was proposed to simplify and unify surface scattering phenomena and is based on a number of assumptions and approximations, like many other theories in physics that allow the application of a simplified solution to a complex phenomenon and are usually sufficient, to a certain degree. The assumptions in Nicodemus's definitions include geometrical (ray) optics approximation, a flat surface that is uniformly irradiated, and incident radiance that depends only on direction. Further assumptions include uniform and isotropic scattering properties of the surface. All reflectance terms according to this nomenclature are represented by a Greek letter, ρ or $d\rho$, but must be accompanied by a pair of illumination and viewing directions.



Incoming/Reflected	Directional	Conical	Hemispherical
Directional	Bidirectional CASE 1	Directional–conical CASE 2	Directional–hemispherical CASE 3
Conical	Conical–directional CASE 4	Biconical CASE 5	Conical–hemispherical CASE 6
Hemispherical	Hemispherical–directional CASE 7	Hemispherical–conical CASE 8	Bihemispherical CASE 9

Figure 2.3: Classification for denoting incident and reflected beam geometry from (Schaepman-Strub et al. 2006).

Each of the nine reflectances have an equivalent reflectance factor, represented by a letter, R, and still requiring the angular information i.e. $R(\theta_i, \phi_i, \theta_r, \phi_r)$ is the Bidirectional Reflectance Factor, known in the remote sensing community as BRF, where: θ and ϕ represent zenith and azimuth angles for illumination beam and reflected beam respectively. A few other examples are presented here for clarity of the angular notation: $R(\omega_i, 2\pi)$ is the Conical-hemispherical Reflectance Factor (incoming light of solid angle ω and outgoing over the hemisphere), and $R(2\pi, \theta_r, \phi_r)$ is the Hemi-directional Reflectance Factor (incoming light even over the hemisphere, measured at the clearly specified θ_r, ϕ_r configuration).

Reflectance factors are defined in (Nicodemus et al. 1977) as

"the ratio of the radiance flux actually reflected by a sample surface to that which would be reflected into the same reflected-beam geometry by an ideal perfectly diffuse standard surface irradiated in exactly the same way as the sample".

There is one more quantity definition that can be seen in the literature, which is radiance factor, β. Three of the radiance factor quantities defined by Judd (1967) are equivalent to the Nicodemus's reflectance factors if the viewing configuration is directional. Thus, the commonly used term BRF for a given directional configuration is equivalent to $R(\theta_i, \phi_i, \theta_r, \phi_r)$ and $\beta(\theta_i, \phi_i, \theta_r, \phi_r)$.

However, all these configurations proposed by Nicodemus do not represent field illumination conditions which are actually present during *in situ* measurements. To address this issue Schaepman-Strub (Schaepman-Strub et al. 2006) created a definition of *in situ* HDRF (which takes into account complex illumination conditions) which is mainly directional, but includes some hemispherical diffuse illuminations to represent the sky radiance. They proposed to address this by including additional information to the equation:

$$HDRF = R(\theta_i, \phi_i, 2\pi; \theta_r, \phi_r) = R(\theta_i, \phi_i; \theta_r, \phi_r)d + R(2\pi; \theta_r, \phi_r)(1-d) \tag{2.2}$$

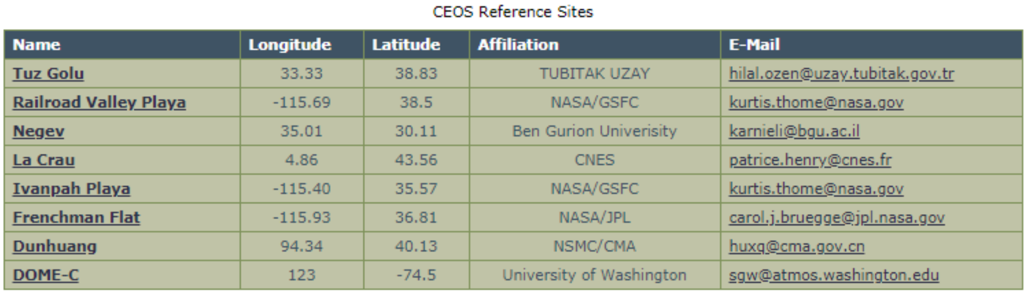
Where: (θ_i, ϕ_i) represents the directional (zenith, azimuth) component of Sun illumination, 2π represents the hemisphere of the diffuse sky radiance and d is the fraction of the direct to the total radiant flux.

Note the difference between angular notation of the Hemi-directional Reflectance Factor $R(2\pi, \theta_r, \phi_r)$ defined by Nicodemus (Nicodemus et al. 1977) under total diffuse illumination only, and Schaepman-Stub HDRF $R(\theta_i, \phi_i, 2\pi, \theta_r, \phi_r)$, where reflectance is a composite of the reflectance due to the directional Sun irradiance and that due to diffuse sky radiance, weighted in the same proportions as the directional/diffuse illumination.

2.3 Land Sites

The White Sands in New Mexico in the US was the first test site used for Thematic Mapper on Landsat 4 (Castle et al. 1984) and then for vicarious calibration of SPOT-1 and 2 (Biggar et al. 1991, Gellman et al. 1993), Landsat-5 Thematic Mapper (Thome et al. 1993) and the Advanced Spaceborne Thermal Emission and Reflection radiometer (ASTER) Slater et al. (1995). The site is flat, highly reflective and its elevation is 1200 m and thus has low aerosol loading. The concentration of aerosols (apart from volcano dust) significantly decreases with latitude.

With an increased number of satellite sensors and proven advantages of vicarious calibration, the subgroup Infrared and Visible Optical Sensors (IVOS) of the CEOS working group on calibration and validation (CEOS-WGCV) started to set reference standard test sites for VC purposes. These are called LANDNET sites and are Land Equipped Sites (LES) thus instrumented sites providing ground reflectance and atmospheric data. Figure 2.4 presents the list of the LANDNET sites.



CEOS Reference Sites

Name	Longitude	Latitude	Affiliation	E-Mail
Tuz Golu	33.33	38.83	TUBITAK UZAY	hilal.ozen@uzay.tubitak.gov.tr
Railroad Valley Playa	-115.69	38.5	NASA/GSFC	kurtis.thome@nasa.gov
Negev	35.01	30.11	Ben Gurion Univerisity	karnieli@bgu.ac.il
La Crau	4.86	43.56	CNES	patrice.henry@cnes.fr
Ivanpah Playa	-115.40	35.57	NASA/GSFC	kurtis.thome@nasa.gov
Frenchman Flat	-115.93	36.81	NASA/JPL	carol.j.bruegge@jpl.nasa.gov
Dunhuang	94.34	40.13	NSMC/CMA	huxq@cma.gov.cn
DOME-C	123	-74.5	University of Washington	sgw@atmos.washington.edu

Figure 2.4: The screen shot from (CalVAlPortal) which contains a table with CEOS reference sites.

There are many more test sites used for satellite calibration or validation purposes around the world; however, they are not classified as standard sites. In order to be accepted for the CEOS standard site, a site owner has to ensure that the site meets a set of requirements, such as site homogeneity, radiometric invariance, and spectral flatness. Size of a test site depends on the target satellite sensor resolution and should include at least few pixels, however, there are no stick rules. Thus for high resolution satellites like Lansat and Sentinel 2 the site can be much smaller than for medium resolution sensors like MODIS. These requirements are outlined in the guidelines of the Quality Assurance Framework for Earth Observation (QA4EO) (*QA4EO* 2018). A candidate site has to be characterised following the CEOS procedure for new site establishment and the records of the test site characterisation and following measurement campaigns must be available to the CEOS community.

In August 2010, CEOS Land Comparison was performed on the Tuz Gölü site (Özen et al. 2011) involving ten international teams. Most participants used an ASD field spectroradiometer and a Spectralon reflectance panel (30 cm x 30 cm) for surface reflectance measurements. The combined standard uncertainty (k=1) for ground reflectance measurements during the comparison for all field teams that took part was in the range 2.5% - 3.5% over the VNIR. This uncertainty should be considered as a nominal uncertainty level for ground reflectance *in situ* measurements using a traditional method of absolute calibration and field measurement protocols. It is important to note that surface reflectance uncertainty is only one of the components of the total uncertainty for vicarious calibration using a test site method.

Tuz Gölü is a salt lake located in the central plateau of Anatolia in Turkey, which is dry during the summer months. During this period, it makes up one of the LANDNET sites. Field campaigns were performed in August in 2008, 2009 and 2010[1] conducting ground reflectance measurements with traceability to SI through the National Physical Laboratory (NPL), in the UK, and additional atmospheric and meteorological measurements were collected. The average surface reflectance factor was between 0.4 and 0.6 in the VNIR and about 0.2 in the SWIR, making this site suitable for the VNIR spectral region. In this region, the site is spatially uniform to within 2% to 4% of the normalised standard deviation and has large homogeneous areas which are suitable for the calibration of high to medium resolution optical sensors.

The involvement of NPL in the Tuz Gölü field campaigns brought BRF calibration of the reference standards back to the attention of the community. Even Spectralon material does not match the theoretical lambertian diffuser and exhibits some deviation from

---

[1] The author did not participate in any of Tuz Gölü field campaigns

a perfect diffuser BRF, which has a constant value of $1/\pi$, regardless of illumination and viewing geometry. Therefore, the solar zenith angle is important and the reference panel calibration should match the illumination conditions on the site. For example, for Tuz Gölü the majority of satellite overpasses happened for SZA around 30°, thus the panel was calibrated for 30°/0° (30° SZA and 0° VZA) geometry.

There are a number of sampling strategies that have been developed to match big and small footprint sensors. Each site will have a custom sampling strategy and often this strategy is driven by some practical constraints rather than a best sampling solution. Figure 2.5 presents examples of sampling strategies applied in field campaigns for high resolution sensors (30 m pixel size) (Slater et al. 1987) and large footprint sensors where the sample size is 1 km x 1 km (Thome et al. 2004).

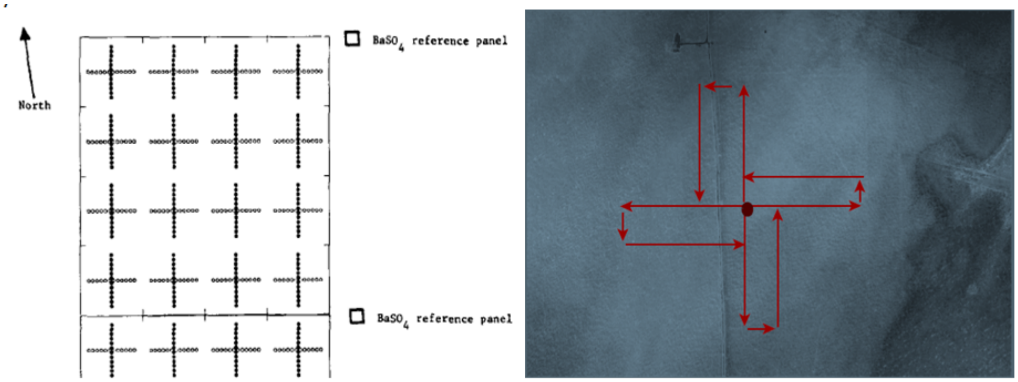


Figure 2.5: Example of sampling strategies for high (Slater et al. 1987) and low resolution sensors (Thome et al. 2004).

Field campaigns are expensive and can last for a very limited amount of time, usually a week, so site characteristics outside the measurements period are unknown. Therefore, over the last ten years a lot of effort has been put into autonomous test site developments which allow for continuous monitoring of the atmosphere and the surface reflectance on the site. A new network called RadCalNet is being established by the CEOS community. The idea behind RadCalNet is to establish autonomous sites that will provide ground reflectance and atmosphere all year round just as the AErosol RObotic NETwork (AERONET) or AERONET-OC (Holben et al. 1990), (Zibordi et al. 2009) networks are doing for the atmosphere and ocean colour.

2.3.1 RadCalNet

Field campaigns are labour intensive and calibrations are limited to the number of overpassing sensors at the time (Teillet et al. 2007). Therefore, the development of a network of fully automated instrumented calibration sites was proposed (Teillet et al. 2001b) to provide *in-situ* measurements on a routine basis. RadCalNet is an initiative from the CEOS-WGCV IVOS to establish a globally distributed network of autonomous instrumented radiometric calibration sites. RadCalNet is currently composed of four existing reference sites located at: Railroad Valley Playa, Arizona, USA (Czapla-Myers et al. 2015); Baotou, China (Li et al. 2015); La Crau, France (Meygret et al. 2011); and Gobabeb in Namibia (Bialek et al. 2016a), graphically presented in Figure 2.6. The network is about to be opened to the public so that new sites can join the network and satellite operators can access the data for vicarious calibration purposes.

Figure 2.6: Locations of the RadCalNet sites.

The surface properties of the RadCalNet sites have been characterised through field campaigns and have permanent instrumentation continuously monitoring the surface and atmospheric conditions. There is no specified type of instrument to be used *in situ*, which is a unique approach in comparison to other networks, such as AERONET, AERONET-OC, that uses exactly the same instrument and measurement protocols. Ground measurements and data from the detailed site characterisation campaigns have to be combined and processed by the site owner to generate BOA products. This is ground reflectance at the nadir view in 10 nm steps covering the spectral range from

400 nm to minimum 1700 nm at 30-minute intervals from 9 am to 3 pm local time, plus selected atmospheric parameters, such as the Ångström coefficient (Å), the 550 nm Aerosol Optical Depth (AOD), and the Aerosol type.

These inputs are fitted to the centralised RadCalNet processing chain that executes RTC (in this case MORTRAN is used) and provides TOA spectral reflectance at nadir for all member sites to facilitate Level 1 radiance absolute calibration and comparison between different optical satellite sensors.

2.3.1.1 Railroad Valley, U.S.

This is a site operated by the University of Arizona Remote Sensing Group. It is a dry lake bed, and is large in area, arid and remote (Czapla-Myers et al. 2008) located at N38.5°, W115.7°. Despite being mostly dry, the moisture content does change with the time of year and a layer of salt can appear on the surface. Either of these would have an effect on the reflectance. An autonomous Radiometric Calibration Test Site (RadCaTS) was set up there, involving several Ground Viewing Radiometers (GVRs).

The first version of a GVR was constructed out of a red LED (used as a wavelength selective detector) in a PVC tube, and was placed on the site, taking measurements over summer and winter to test the robustness of the design. After this was successful, another version involving three LEDs (green, red and NIR) in three separate tubes was tested. The LED based radiometers were found not accurate enough. The RadCaTS site was updated in 2011 with more accurate radiometers that replaced the LED approach with a traditional filter radiometer technology (Anderson et al. 2013).

Over time scientists estimated that for appropriate levels of spatial characterisation on this test site four instruments were enough and the high accuracy and quality of measurements had higher priority than better spatial coverage. The GVRs are arranged so that there is a core site, which gives spatial and spectral detail for high spatial resolution sensors, and node sites, which give spatial information for low spatial resolution sensors. The placement of the GVRs was investigated (Czapla-Myers et al. 2008) by comparing it to satellite measurements, and to the current method of calibrating the test sites, which involves walking around with a portable ASD FieldSpec spectroradiometer. The sun and sky measurements are carried out with the current AERONET set up, and measure the same things: aerosol properties, ozone, water vapour, and irradiance. There is also a meteorological station to provide data on the temperature, pressure and precipitation at the site. The site reflectance is calculated with the equation presented below:

$$\rho = \frac{\pi C_{\text{GVR}} V_{\text{GVR}}}{\frac{E_0}{d^2}\tau_{\text{solar}} cos\theta + E_{\text{sky}}} \tag{2.3}$$

Where: ρ is the surface reflectance, V_{GVR} is the voltage response of the GVR, C_{GVR} is the calibration coefficient, E_0 is the top of atmosphere solar irradiance, d is the Sun-Earth distance, τ_{solar} is the direct solar transmission, θ is the solar zenith angle, and E_{sky} is the diffuse sky irradiance.

Czapla-Myers et al. (2015) presents the uncertainty budget for this site and quoted at (1σ) equivalent to (k=1) for TOA Spectral radiance at the level of 4.1% for blue, and infrared channels and 3.1% for red band for typical clear-sky conditions and a solar zenith angle of 45°.

2.3.1.2 La Crau, France

The site is operated by the CNES RObotic Station for Atmosphere and Surface characterisation dedicated to on-orbit calibrations (ROSAS) (Meygret et al. 2011). It is based on a similar system to AERONET, but modified to also measure ground radiance, and to take measurements of short wave infrared lights. The system used at the La Crau test site N43.6°, E4.9° (an area in France approximately 20 km in diameter, covered in white pebbles) consists of a CIMEL sun photometer placed on a 10 m high pole. It measures solar extinction and sky radiance to allow the calculation of the atmosphere's optical properties.

Figure 2.7 presents a schematic of the measurements sequence for the CIMEL (*Cimel Electronique*) instrument on the La Crau site. Where the instrument that contains an optical head, a collimator and the robot is mounted on the top of a mast it is represented as the grey shape. The solid line below the instrument is a mast and the dotted line above the instrument represents the zenith. Four measurement scenarios are shown, starting from the left-hand side with the direct sun measurements, followed by the principal plane, where the curved line with arrows at both ends shows the movement of the optical head. Then there is the sky almucantar scenario where the optical head measures the sky radiance in all azimuth directions for a given SZA, and the last scenario is the ground radiance measurement, where the instrument head is directed to the ground.

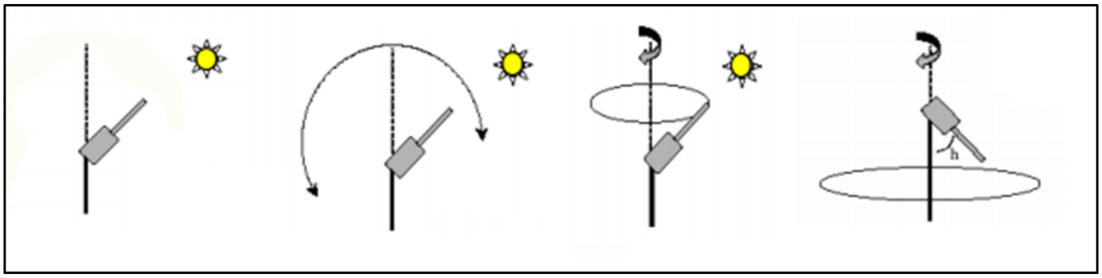


Figure 2.7: ROSAS schematic of the sun radiance, principal plane, almucantar and ground radiance measurements (Meygret et al. 2011).

Measurements were taken over nine narrow spectral bands, between 380 nm and 1600 nm. The instrument has been updated to a new twelve band version in December 2014. One of these channels is a vegetation red edge channel to detect dry grass, one is for water vapour, one for aerosols, one for absolute radiance calibration, and the rest are for actual measurements.

The following measurement sequence is run for when the air mass[1] is less than five and stopped when it is greater than five:

- Direct sun irradiance over all bands.
- Sky radiance in the principal plane over all bands except the water vapour.
- Sky radiance in the almucantar over all bands.

---

[1] $m = 1/cos(\theta_S)$ is a first order approximation of an optical air mas. A value of 5 is equivalent to 78.5° SZA

- Ground radiance at 12 zenith angles up to 60° and 72 azimuth angles, over all bands except the atmospheric and calibration bands.

The sequence starts with a sun measurement, then the almucantar and principal planes are measured. Each ground elevation is then done, alternated with another sun measurement until all twelve elevations have been measured. This whole cycle takes approximately ninety minutes.

The system is autonomously "calibrated"/monitored on site; the irradiance calibration is taken from the classical extinction formula, the Langley-Bouguer principle (Shaw 1983), and the radiance calibration is from the short wavelength molecular scattering. The site was characterised using ASD spectroradiometer. Daily CIMEL measurements are used as inputs to the Roujean BRDF model (Roujean et al. 1997) for each spectral band. To meet RadCalNet objectives the data from this site will have to be spectrally interpolated into 10 nm bands. Moreover this system does not measure nadir geometry, thus this geometry will have to be derived from the BRDF models. Currently there is no information about uncertainty associated with measurements from the La Crau site, although the site is used for SPOT vicarious calibration purposes.

2.3.1.3 Baotou, China

This is a relatively new site and its surface is covered with an artificial target. The site is run by the Academy of Opto-Electronics (AOE), the Chinese Academy of Sciences (CAS) and is located in Wulate Qianqi, Bayan Nur City, Inner Mongolia, with latitude N40.72° and longitude E108.65°. The artificial surface is made from gravel and consists of four squares with three different reflectance levels (see Figure 2.8). A natural sandy surface is in close proximity to the artificial site location.

This site became operational as AERONET in 2013. This site provides reference hyperspectral ground reflectance, as the only one in RadCalNet uses the SVC HR-1024 field spectrometer (*Spectra Vista Corporation* 2018) and the CIMEL sun photometer for atmosphere measurements. In addition there is a custom built multi-angular system for BRDF measurements.

A hyperspectral autonomous ground reflectance system is under development and will be based on the commercially available USB2000+VIS-NIR (*Ocean Optics* 2018) spectrometer with a wavelength range of 350 nm to 1000 nm and a spectral resolution of 1.5 nm.

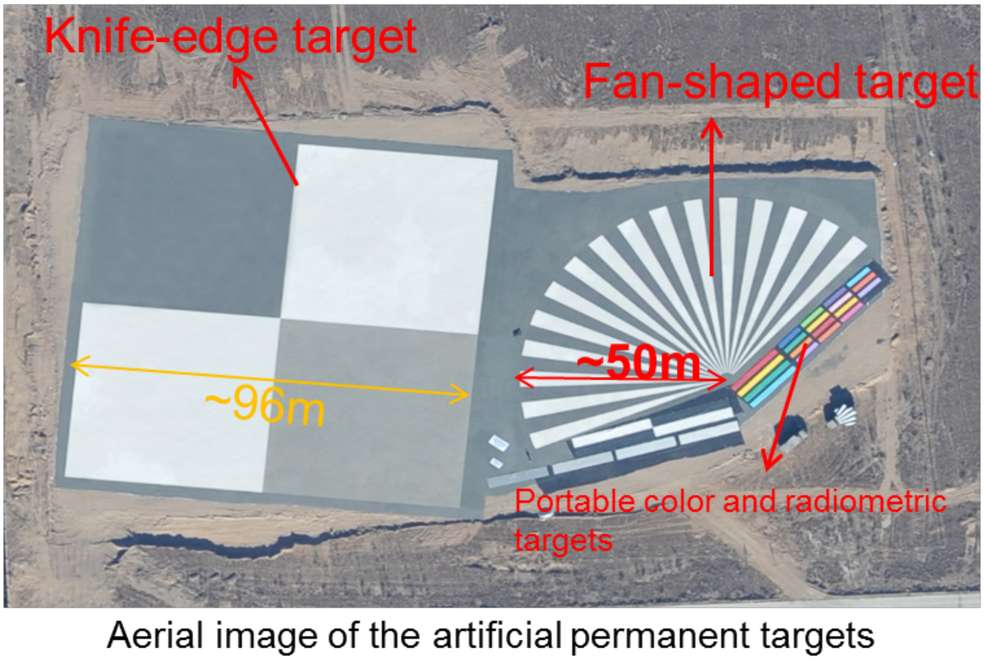


Figure 2.8: Aerial image of Baotou site.

2.3.1.4 New Site Gobabeb, Namibia

A new ESA site was developed during this research. NPL together with Magelium and CNES have been appointed by the ESA to establish the new test site and a selection of preferable site locations have been identified by scientists from Magelium.

Good test site criteria requirements include high reflectance and spectral flatness in the VNIR and SWIR regions, homogeneity, radiometric stability, low precipitation, low number of overcast days, and a low aerosol content. A subset of ideal candidates has been identified including places in Chile, Namibia, Australia, and Saudi Arabia. The final location has been chosen at the Gobabeb Namib desert in Namibia, latitude S23°36′72″ and longitude E15°7′10″.

Instrumentation preparation and characterisation was conducted as part of this study as is depicted in detail in Chapter 3. The new site has permanently installed a new version of the CIMEL sun photometer that has twelve filters and uses the same measurement sequence as the ROSAS instrument described before. This instrument provides continuous atmosphere and ground reflectance data sets. It is accompanied with a weather station. Before the operation stage the site was characterised using a hyperspectral spectroradiometer (ASD FieldSpec) and site HDRF was derived from the GRASS (Gonio Radiometric Spectrometer System) (Pegrum-Browning et al. 2008) instrument measurements. Before the *in situ* campaign, all instruments were absolutely calibrated and characterised at NPL.

2.4 Ocean Measurements

An ocean colour measurement is significantly different from a land measurement due to the nature of the ocean and the difference in the reflectance signal actually measured. A terrestrial product of interest is an actual surface reflectance, whereas in the case of ocean, the surface reflected part of the signal is removed. Sunlight that passes through sea water interacts with the various suspended particles it contains and remerges with useful information about them. However, since only a fraction of the sun's light is refracted below the sea surface measurements must be sensitive enough to detect small changes in whatever little light comes from the water. Therefore, this section begins with an introduction to ocean colour products and *in situ* measurement methods before continuing with an actual ocean sites description. The ocean is a dynamic medium and accessibility to most of it is very difficult and restricted. Ocean colour vicarious calibration sites have to be near the land to maintain the instruments on a buoy or a measurement tower, and the description of the site is related to actual instrumentation and its assembly at sea and sea characteristics.

Waters are classified as case-1 and case-2 waters, where case-1 refers to waters whose optical properties are determined primarily by phytoplankton and related coloured dissolved organic matter (CDOM); and case-2 waters whose optical properties are significantly influenced by other constituents, such as mineral particles, CDOM, or microbubbles, whose concentrations do not co-vary with the phytoplankton concentration (Morel 1988). Most of waters will be a mixture of case-1 and case-2 with a different contribution of particles depending of the time of the year (for example case-1 water change to case-2 during spring bloom). Nevertheless, this distinction is still in use for ocean colour science, and with a simplistic approach, one can say open waters are considered as case-1 and turbid coastal waters as case-2 waters.

2.4.1 Ocean Measurement Methodology

The primary ocean colour product is the spectral remote sensing reflectance $\mathrm{R_{rs}}$, which is used to generate higher-level products such as chlorophyll-a concentration. This is necessary to meet the OC missions science objectives: estimation of global primary production (production of organic compounds from atmospheric or aquatic carbon dioxide) and quantification of ocean biological variability on global scales for a time series (5 years at least).

The main challenge moving forward is to generate a long-term climate data series that can be formed by using data merged from different sensors. Moreover, increasing applications in coastal water management, such as harmful algal bloom monitoring, means detection requires a higher level of accuracy for local products. To achieve these requirements, there is a need to establish and further reduce uncertainties in ocean colour products.

Current requirements for ocean colour mission products (IOCCG 2010, 2012) state the "accuracy" at the level of 30% (k value unspecified) for Chlorophyll-a concentration for open case-1 water. To meet this requirement, it is necessary to obtain water leaving radiance accuracy at the 5% level at 443 nm (k value unspecified, but assumed to be 1), and there is a desire to reduce this uncertainty to 3% for future missions. It is important to note that water-leaving radiance contributes to less than 15% of TOA radiance; therefore, a current requirement for TOA radiance uncertainty at 0.5%, is achievable only after system vicarious calibration adjustment to the remotely sensed data.

The complexity of ocean measurements is presented in Figure 2.9 from (Mobley et al.). The quantity of interest for ocean colour is only a tiny part of the TOA signal that contains solely the part emerged from the water body and not the part that is reflected from the water surface and atmospheric scaterring. This is shown in Figure 2.9 as the red and green rays, whereas only the green rays are of interest.

There are two methods to derive *in-situ* $\mathrm{L_w}$, the water-leaving radiance: in-water and above-water measurements. In-water field radiometry measurands include the upwelling radiance, $\mathrm{L_u}(z, \lambda)$, the downward irradiance, $\mathrm{E_d}(z, \lambda)$, and the upward irradiance, $\mathrm{E_u}(z, \lambda)$. The measurements are taken at several depths, z, and are all normalised by above-water downward irradiance, $\mathrm{E_d}(0^+, \lambda)$, to compensate for differences in sun illumination during the measurements. Measured values are used to derive sub-surface

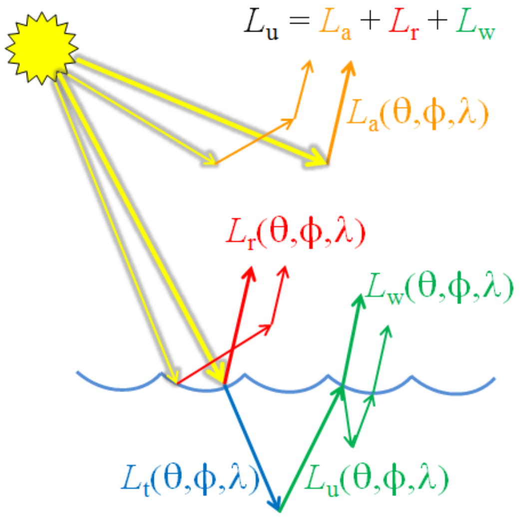


Figure 2.9: Contributions to the total upwelling radiance above the sea surface, (Mobley et al.), $\mathrm{L_u}$. Yellow arrows are the Sun's unscattered beam; orange arrows are atmospheric path radiance $\mathrm{L_a}$; red is surface-reflected radiance $\mathrm{L_r}$; and green is water-leaving radiance $\mathrm{L_w}$. Thick arrows represent single-scattering contributions; thin arrows illustrate multiple scattering contributions.

quantities using the least-square linear regression of $ln\Im_0(z,\lambda)$ as a function of depth, where $\Im_0$ stands for $\mathrm{L_u}(0^-,\lambda)$, $\mathrm{E_u}(0^-,\lambda)$ or $\mathrm{E_d}(0^-,\lambda)$. Water-leaving radiance from in-water measurement is defined as:

$$\mathrm{L_w} = 0.543\mathrm{L_u}(0^-,\lambda) \tag{2.4}$$

The constant 0.543 is to account for the reduction in radiance from below to above the water surface caused by the change of the refractive index between two different optical mediums (in this case water and air). The number is calculated under the assumption that the refractive index of sea water is not wavelength dependent (Austin 1974).

The above water water-leaving radiance equation is expressed as:

$$\mathrm{L_w}(\theta,\phi,\lambda) = \mathrm{L_T}(\theta,\phi,\lambda) - \rho(\theta,\phi,\theta_0,W)\mathrm{L_i}(\theta',\phi,\lambda) \tag{2.5}$$

Where: $\rho(\theta,\phi,\theta_0,W)$, the sea surface reflectance is derived from a theoretical model of the viewing geometry (θ, ϕ, θ_0) and the sea state, represented by the wind speed W, $\mathrm{L_T}(\theta, \phi, \lambda)$ is the total radiance from the sea and $\mathrm{L_i}(\theta', \phi, \lambda)$ is the sky radiance. The measured values need to be corrected for illumination changes over time using the downward irradiance, exactly the same as in the case of in-water measurements. The spectral remote sensing reflectance is given as:

$$\mathrm{R_{rs}}(\lambda) = \frac{\mathrm{L_w}(\lambda)}{\mathrm{E_d}(0^+, \lambda)}[sr^{-1}] \tag{2.6}$$

Where: $\mathrm{L_w}$ is water-leaving radiance in $[\mathrm{mWcm^{-2}\mu m^{-1}sr^{-1}}]$ and $\mathrm{E_d}(0^+)$ is downward sun irradiance just above the surface $[\mathrm{mWcm^{-2}\mu m^{-1}}]$. Spectral remote sensing reflectance, is then normalised to a common set of viewing conditions (solar-zenith angle, sensor zenith angle and relative azimuth angle) and corrected for Sun-Earth distance and atmospheric transmittance. The normalised water-leaving radiance L_{wn} is given by equation 2.7 or 2.8. However, the second equation is applicable to above-water measurement methods only.

$$\mathrm{L_{wn}} = \mathrm{R_{rs}}(\lambda) E_0(\lambda) \tag{2.7}$$

$$\mathrm{L_{wn}} = \mathrm{L_w}(\theta, \phi\, \lambda) \frac{E_0(\lambda)}{E_d(0^+, \lambda)} C_{\Re Q}(\theta, \phi, \theta_0, \lambda, \tau_a, IOP, W) \tag{2.8}$$

$E_0(\lambda)$ is the average extra-atmospheric solar irradiance (Thuillier et al. 2003), and the term

$$C_{\Re Q}(\theta, \phi, \theta_0, \lambda, \tau_a, IOP, W) = \frac{\Re_0}{\Re(\theta, W)} \frac{Q(\theta, \phi, \theta_0, \lambda, \tau_a, IOP)}{Q_n(\theta_0, \lambda, \tau_a, IOP)} \tag{2.9}$$

compensates for viewing angle dependence of above-water water-leaving radiance $\mathrm{L_w}(\theta, \phi\, \lambda)$. $\Re(\theta, W)$ and $\Re_0$ (at nadir, $\theta = 0$) specify the sea reflectance and refraction and depend on viewing angle θ and wind W speed. $Q(\theta, \phi, \theta_0, \lambda, \tau_a, IOP)$ and $Q_n(\theta_0, \lambda, \tau_a, IOP)$ are Q-factors at viewing angle θ and at nadir respectively, account for the anisotropic distribution of the in-water radiance field depending on viewing angle, sun azimuth, aerosol optical thickness and seawater inherent optical properties (Zibordi & Voss 2010). To account for BRDF effects in water an additional normalisation of water-leaving radiance was introduced (Morel & Gentili 1996). The new term is called the exact normalised water-leaving radiance, L_{WN}, and is given by:

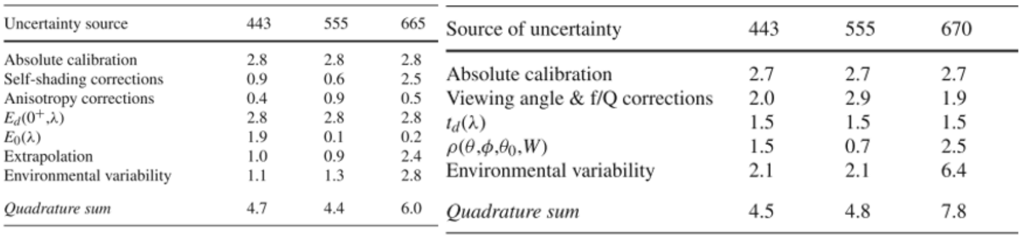
$$\mathrm{L_{WN}}(\lambda) = \mathrm{L_{wn}}(\theta) C_{f/Q}(\theta_0, \lambda, \tau_a, IOP) \tag{2.10}$$

Where:

$$C_{f/Q}(\theta_0, \lambda, \tau_a, IOP) = \frac{f_0(\lambda, \tau_a, IOP)}{Q_0(\lambda, \tau_a, IOP)} \left[\frac{f(\theta_0, \lambda, \tau_a, IOP)}{Q_n(\theta_0, \lambda, \tau_a, IOP)} \right]^{-1} \tag{2.11}$$

$f_0(\lambda, \tau_a, IOP)$is $f(\theta_0, \lambda, \tau_a, IOP)$ at $\theta = 0$ and these terms correspond to the relationship between the irradiance reflectance (AOP) and the seawater backscattering to absorption ratio (IOP). Chlorophyll a concentration, *Chla*, is used to predict this BRDF effect dependence on IOP but only for case-1 waters. For more optically complex coastal waters the f factor has to be replaced by its local equivalent f' (Morel et al. 2002).

Although the methods of measuring water-leaving radiance vary the instruments used to perform these measurements are similar and their primary purpose is to provide radiance (or irradiance) values. To convert the digital numbers to radiance values in physical units, an absolute radiometric calibration is necessary and is done using the same principles for both in-water and above-water instruments. The quality of this process is expressed by its uncertainty. To know how reliable *in situ* measurements are it is crucial to establish an uncertainty budget. The components of a budget will vary depending on the measurement method, i.e. in-water or above-water systems. Additionally, each measurement platform or system can have different features, and they should be accounted for.



Uncertainty source	443	555	665
Absolute calibration	2.8	2.8	2.8
Self-shading corrections	0.9	0.6	2.5
Anisotropy corrections	0.4	0.9	0.5
$E_d(0^+,\lambda)$	2.8	2.8	2.8
$E_0(\lambda)$	1.9	0.1	0.2
Extrapolation	1.0	0.9	2.4
Environmental variability	1.1	1.3	2.8
Quadrature sum	4.7	4.4	6.0

Source of uncertainty	443	555	670
Absolute calibration	2.7	2.7	2.7
Viewing angle & f/Q corrections	2.0	2.9	1.9
$t_d(\lambda)$	1.5	1.5	1.5
$\rho(\theta,\phi,\theta_0,W)$	1.5	0.7	2.5
Environmental variability	2.1	2.1	6.4
Quadrature sum	4.5	4.8	7.8

Figure 2.10: Uncertainty expressed in % for $\mathrm{L_{WN}}$ for in-water (left) and above-water (right) data from coastal water (Zibordi & Voss 2010)

Examples of uncertainty budgets for in-water and above-water measurements are presented in Figure 2.10. Absolute radiometric calibration uncertainties are always present in uncertainty budgets of ocean colour *in situ* measurements and although they seem to be easier to address compared to some of the environmental aspects, they still have a major contribution to the overall uncertainty budget.

The subchapters below present information about instrumentation and operation of the ocean sites that are used for *in situ* satellite system vicarious calibration and validation.

2.4.2 Examples of Ocean Sites

2.4.2.1 MOBY

The Marine Optical BuoY (MOBY) is located in case-1 waters 20 km of Lanai, Hawaii, Pacific Ocean (geographic coordinates N20°49′ and W157°11′). MOBY was designed and built to support vicarious calibration of ocean colour data from SeaWIFS and MODIS Terra and Aqua satellite sensors (Brown et al. 2007) and has been in operation since 1995. Water depth below the buoy is 1200 m.

The instruments on the buoy measure upward radiance and downward irradiance at different depths with defaults depths of 1 m, 5 m and 9 m. Additionally solar irradiance is measured above the water to correct for illumination changes during the measurements. The entire measuring system is called the Marine Optical System (MOS) and is placed in the instrument bay at the base of the buoy at 12 m depth and consists of two spectrographs with different spectral ranges of 340 nm - 640 nm and 550 nm - 955 nm and two thermoelectrically cooled CCD detectors that detect the signal from remote collectors at different depths. The collectors are connected to 1 mm core diameter fibre-optic cables and feed into the spectrographs via a fibre-optic rotary selector (multiplexer).

An individual buoy is in operation for three months and is then replaced by a second buoy with the same MOS system. The buoy replacement will include simultaneous operation of both buoys for several measurement runs to establish crossover calibration between the two instruments. While the buoy is not deployed it is being recalibrated and refurbished.

From the very beginning of the MOBY programme, much effort was put into SI traceability. NOAA, NASA and NIST worked together to develop the instrument and ensure its high-quality radiometric calibration and estimate its uncertainty budget. All spectral irradiance sources are re-calibrated directly by NIST at 50-hour lamp burn intervals to get lower uncertainties than secondary laboratories can provide. MOBY has in-water internal calibration systems with blue and red LEDs and incandescent lamps. Internal measurements are taken at the same time as the up-welling radiance measurement.

After nearly ten years of tests with developed data corrections for stay light and self-shading applied, and including only the data from very good environmental conditions standard combined uncertainty of MOBY up-welling radiance of the top arm (at 1 m depth) in 2007 was established at 3% for the blue and green part of the spectrum, and exceeding 3% for red wavelengths. One should note that the end product of ocean colour for vicarious calibration is normalised water-leaving radiance or remote sense reflectance and upwelling radiance is a component of the higher product, and their uncertainties might be higher than 3%. There is no information about them in the publication from 2007.

2.4.2.2 BOUSSOLE

BOUSSOLE is a French acronym “BOUée pour l’acquiSition d’une Série Optique á Long termE” and stands for the buoy for the acquisition of a long-term optical time series (Antoine et al. 2008b). The BOUSSOLE is moored in the Ligurian sea, one of the sub-basins of the Western Mediterranean Sea around 60 km from Nice (geographic coordinates N44°22′ and E7°54′) and has been in operation since 2002. The water depth underneath the buoy is in the range of 2350 m - 2500 m.

The BOUSSOLE project consists of three elements: a monthly cruise program, the buoy and the coastal AERONET station. The operational objective is vicarious calibration of Ocean Colour satellite observations and validation of level-2 geophysical products. The project science objective is the monitoring of short-term, seasonal and long-term

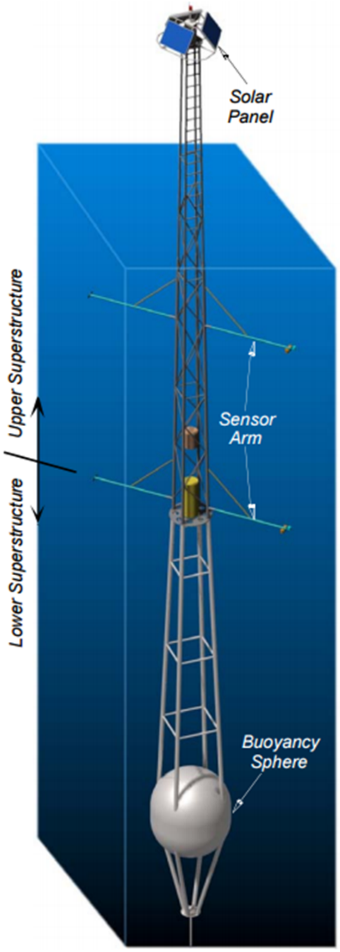


Figure 2.11: Schematic of BOUSSOLE buoy (Antoine et al. 2008a).

trends in inherent optical properties (IOPs) of the medium such as absorption and AOPs which are related to sun illumination and other external factors and hence biophysical parameters.

The instruments on the buoy measure upward radiance, upward and downward irradiance at different depths (4 m and 9 m) and wavelengths. Measurements are taken by the acquisition of a 1-minute sequence that is averaged; such measurements are repeated every fifteen minutes, day and night. Night measurements are used as dark measurements for multispectral radiometers as they do not have internal shutters (hyperspectral instruments take dark readings). All radiometers and spectroradiometers used on the buoy are manufactured by Satlantic (*Sea-Bird Scientific* 2018).

Hyperspectral radiometers were introduced on BOUSSOLE with the intention to replace multispectral instruments. However, due to the difference in readings between the old type of multispectral and the new hyperspectral radiometers, the old instruments are still in operation. To address that discrepancy a compete instrument characterisation would be required to correct for all differences between both instuments types, for example different detector non-linearity response. All radiometers are calibrated on a regular basis by the manufacturer. BOUSSOLE data are used to create a time series. The 10-year average of the annual cycle for the mixed layer and chlorophyll concentration shows seasonal changes between winter and summer in nutrients brought into the surface layer.

The estimated uncertainty is around 6% and assumes uncertainty in the absolute radiometric calibration of radiometers of 3% (Antoine et al. 2008b). The BOUSSOLE uncertainty is probably lower in the blue end on the spectrum and higher in the red end.

2.4.2.3 AERONET-OC

AERONET-OC (Zibordi et al. 2009) is an extension of the AERONET network (that is described in more details in the section Atmosphere Measurements 2.5), used to measure ocean colour for satellite vicarious validation. The AERONET-OC uses already existing structures, such as oceanographic towers, lighthouses or old oil platforms to hold instruments thus, their positions are even closer to the land than the previously discussed buoy systems. This allows for validation of local products only, i.e. specific for the Baltic Sea or the Mediterranean Sea only. Water types for an AERONET-OC sites will typically vary between the open sea and coastal waters.

Measurements of the normalised water-leaving radiance are taken using above water methods. Like AERONET, it consists of a network of standardised radiometer systems, removed for calibration at JRC or GSFC every six to twelve months. The measuring system is called the Sea Photometer Revision for Incident Surface Measurements (SeaPRISM), and this is a CIMEL CE-318 (*Cimel Electronique*) autonomous sun photometer, which is adapted to measure, in addition to the sky, sea radiance at a variety of angles relative to the Sun. The direct sun irradiance is also measured, and this data is used to cancel the measurement sequence if cloud cover reduces the atmospheric transmittance too much. Measurements are taken over five different wavelengths, with two additional ones used as a quality check. The measurement sequence is run every thirty minutes within four hours of local noon:

- Sun irradiance on all spectral bands.
- Sea radiance as a function of wavelength, azimuthal angle, and angle from the sun.
- Sky radiance as a function of wavelength, azimuthal angle, and angle from the sun.

Figure 2.12 presents the schematic view of the SeaPRIMS measurements, where a rectangle schematically presents the instrument head with a collimator, another thin rectangle attached to it. On the "side view" panel, on the left-hand site, the dotted line represents the zenith-nadir axis. The optical head is shown in three positions: E_s the Sun irradiance measurement, L_i the sky radiance measurement and the L_t the sea radiance measurement. The right panel "to view" shows the same from the top view. The emphasis here is on the fact that the azimuth angle, marked as ϕ on the Figure 2.12, is the same for the sea and sky radiance measurement.

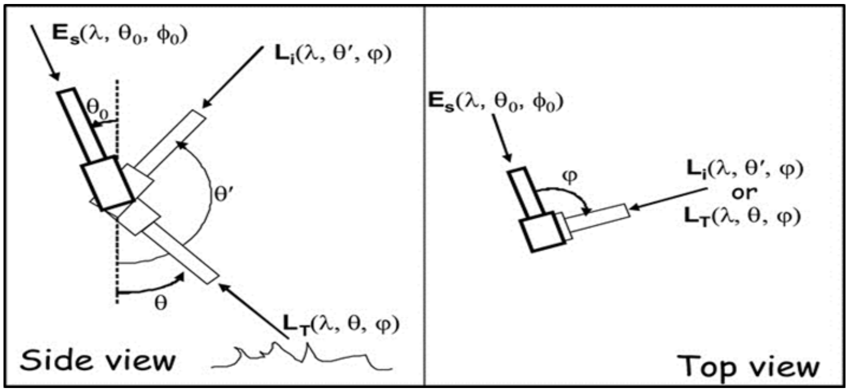


Figure 2.12: SeaPRISM measurement geometry (Zibordi et al. 2009).

The number of measurements of sea and sky radiance can be adjusted separately. Sea radiance is generally noisier, so more values are taken. An azimuthal angle of 40° was chosen to give the best results. The angle relative to the Sun must be less than 90°, or shadows from the structure can interfere with the results. AERONET-OC uncertainties are estimated to be at 5% for the blue end of the spectrum and increase to almost 8% for the longer wavelength as there is very little water-leaving radiance in the red where there is high signal to noise ratio.

2.5 Atmosphere Measurements

As already mentioned, the AERONET network (Holben et al. 1990) was established in the 1990s to provide ground-based atmospheric aerosol measurements for remote sensing and climate monitoring data validation. Measurements of the sun irradiance and sky radiance are collected by a sun photometer and total optical depth is extracted according to the Beer-Lambert-Bouguer law:

$$V_\lambda = V_{0\lambda} d^2 exp(\tau_\lambda m) t_y \tag{2.12}$$

Where: V_λ is digital voltage from the Sun irradiance measurement ,$V_{0\lambda}$ is the extraterresstional voltage, d is the ratio of the average and actual Sun/Earth distance, τ_λ total optical depth, m optical air mass and t_y is transmission of absorbing gasses.

All sites use CIMEL sun photometers and all instruments are regularly calibrated for sky channels in radiance mode using calibrated integrating sphere. The sun channels irradiance mode are calibrated using reference Langley plots and are traceable to the NOAA Mauna Lao Observatory in Hawaii. The Langley plot is a log of digital numbers taken in the morning hours plotted against the optical air mass. The observatory in Hawaii is isolated from the majority of aerosol sources and located at high altitude and provides a stable aerosol and irradiance condition in the morning. The atmosphere above Hawaii Islands is the primary reference standard for all aerosol optical depth measurements, so actually these measurements are not directly linked with SI units. The secondary standard is atmosphere above Canary Islands where the instrument that is installed on the new Gobabeb test site was calibrated.

The sun photometer is a multispectral radiometer where spectral channels are defined by a set of interference filters. There are a few versions of CIMEL sun-photometers with different set of filters depending on their application. An optical head traditionally has two separate silicon detectors, one for sun and one for sky measurements. Both have the same narrow field of view of around 1.2°, but the sky one has a larger aperture-lens system. Collimators are attached to the optical head for stray-light rejection. The model that is used on the new ESA site instead of the separate detector for sun and sky have a silicon and InGaAs detectors and each of them can be in sky and sun mode. In addition to an optical head, the sun-photometers have robots that enable autonomous movements of the radiometer in predefined directions and have solar panels. They have a built-in wet sensor so that when it is raining the measurements are stopped and the optical head is moved to park position pointing at the nadir to prevent the optical windows of the instrument from contamination. All AERONET sites use the same measurements sequence and processing to retrieve a set of parameters from sun and sky measurements. These include size distribution, phase function aerosol optical thickness, wavelength exponent, Langley plots, water vapour and sky radiance. All measurements are transmitted via satellite links, quality controlled and made available at the AERONET website.

Measurement sequences include direct sun and sky scenarios. Sun triplet, for example, is a direct sun measurement repeated three times in thirty second intervals. This scenario is performed in the morning and afternoon hours, every fifteen minutes, as the Langley calibration. After the sun measurements, Langley sky is conducted to ensure the stability of the Langley plot; this is a sky measurement taken at 20° from the Sun. The remaining sky sequences are the almucantar and principal plane (see Figure 2.7) and allow the retrieval of size distribution, phase function and aerosol optical thickness. Aureole measurements are taken with the sun collimator directed 3° from the Sun.

Data derived from AERONET, for vicarious calibration applications, are used as inputs to radiative transfer codes to model the atmosphere above the ground where reflectance was measured.

2.6 Atmosphere Modelling / Radiative Transfer Code

The next component of vicarious calibration is a radiative transfer model that is used to propagate the ground reflectance and ground aerosol measurements to the top of the atmosphere via the atmosphere. Satellite sensors measure TOA radiance that apart from the actual ground radiance contains an additional contribution from atmosphere absorption and scattering. There are a number of computer programmes that are commonly called RT codes that enable the propagation of the radiance or reflectance from the bottom of the atmosphere to the top and in reverse direction.

To obtain satellite level L2 products, such as ground reflectance, L1 product TOA radiance has to be propagated down to the bottom of the atmosphere. Similarly to complete vicarious calibration or validation processes, ground truth measurements have to be propagated upward through the atmosphere to be compared with TOA sensor readings.

The MODTRAN MODerate spectral resolution TRANsmittance code is commercially available software that was first developed, over twenty-five years ago, as the US Air Force standard. The newest version MODTRAN6 (Berk et al. 2014) is currently available. Other codes often show a comparison to the MODTRAN as a validation. MODTRAN5 will be used to provide operational RadCalNet TOA data; this processing will be centralised to ensure that the ground data is processed in the same way and is done by NASA.

6S - Second Simulation of a Satellite Signal in the Solar Spectrum Vector is a basic RT code and is a freeware (Vermote et al. 1997). This model has a vector base and accounts for polarisation effects. It was shown (Kotchenova & Vermote 2006, 2007) that 6S ,when used in its default accuracy setting compared to the benchmark codes (e.g. MODTRAN) within 0.4%-0.6%. This code has several ground BRDF models built in so a user can choose between using the Lambertian surface, a predefined model within the software, or BRDF measurements from the test site.

The library for Radiative transfer (LibRadtran) is another example of a free software package (Emde et al. 2016, Mayer & Kylling 2005) which contains several tools to calculate radiative transfer. It seems that there are more options for the bottom of the atmosphere retrieval and processing including 3-D ground modelling using Monte Carlo methods called the MISTIC package (Mayer et al. 2010).

2.7 Discussion

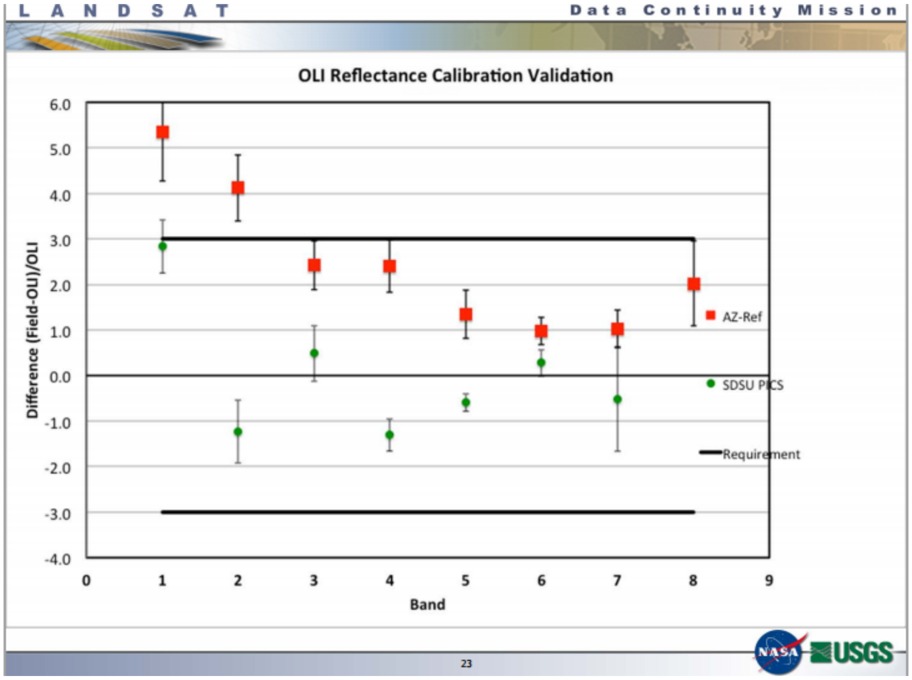
Vicarious calibration of the satellite sensors using test sites started in the early 1980s. At the same time, statisticians and metrologists started to work on a new concept of uncertainty in a measurement. The ISO published the first version of GUM in 1993. Dissemination of a new concept takes time, thus the first VCs were conducted before the "uncertainty" era. The idea of measurement accuracy and precision was known before, and the error budgets accompanied the first VC results.

The uncertainty analysis was very well accepted in the world of metrology as it unified the evaluation of a doubt about the measurement result. It has become an integral component of any calibration and indeed it is not possible to establish SI traceability without information about measurement uncertainty at each step in the traceability chain.

Spreading the GUM to other scientific communities was much slower, thus even up to now, error and uncertainty can be seen used interchangeably. Moreover, very rarely the uncertainty is accompanied with the appropriate level of details for a reader to be able to reconstruct the evaluation. The coverage factors are often omitted.

At the beginning of this project back in 2013, the OLI sensor on board Landsat 8 was the newest launched satellite of Landsat series. The state of the art can be nicely represented by interpretation of its vicarious calibration validation results presented in Figure 2.13. The black lines on this plot represent the 3% mission requirement in L1 TAO radiance/reflectance. The green series on this graph represents pseudo-invariant sites and the red series the instrumented sites. The error bars are the standard deviation of the readings. The difference between the field data and satellite data are expressed in percentages.

The goal to achieve would be to have a similar plot but presented in the form of the comparison of field and TOA reflectances including uncertainty associated with field data and L1 products. Then an indication of where the difference between them is higher than their combined uncertainties would inform about the areas that require further investigation.



LANDSAT

Data Continuity Mission

OLI Reflectance Calibration Validation

23

Figure 2.13: OLI vicarious calibration validation results red series 4 instrumented sites, green series pseudo-invariant Libya 4 site (Helder et al. 2013).

Nevertheless, VC has been successfully used for a few decades now, and it is an essential component of the satellite products quality assurance. Continuous development of the technology allows now to establish permanent autonomously operated sites and networks with wireless data transmission. *In situ* data are collected over years without the time limitations of short field campaigns. SI traceability comes from laboratory calibrations of the instruments before their installation on the site. They are in operation for a set period, depending on the site environment from three months as this is done for MOBY, to one year for some sites such as La Crau and Gobabeb.

Field campaigns are still essential to provide fully spectrally resolved information about the ground reflectance for the site where permanent instrumentation is multi-spectral. The SI traceability for the reflectance mode measurement is obtained from the calibrated reflectance panel. However, as already mentioned in section 2.2.1, the calibration of the panel does not match the illumination condition that is present *in situ*, thus introducing bias into measurements. Thus, a novel method of reference reflectance standard calibration for the field used is proposed in this research to close the gap between laboratory and field illumination conditions. The proposed method will improve uncertainty in the *in situ* measurements especially in the short wavelength range, where the contribution from the atmosphere is more significant.

The complete step by step uncertainty evaluation for the various measurements following GUM and a metrologist approach with attention to coverage factors is presented in this thesis. The examples include laboratory-based calibrations, on the land site characterisation, and ocean site operational data.

The concept of SI traceability with an inseparable uncertainty statement will be reiterated through the chapters of this thesis.

Chapter 3

Characterisation of the Reflectance of the Gobabeb Site

The new radiometric calibration site located in Gobabeb, on the edge of the Namib desert in Namibia, is the main focus of this chapter. The methodology for the site characterisation and its permanent operation is presented.

A novel approach to SI traceability for the site characterisation using the reflectance-based method was proposed and tested. Issues related to the stability of the reflectance standards during long field campaigns are discussed. Uncertainty components related to the operator were carefully studied by measurements of the large area of uniform tarpaulins.

The *in situ* measurement results are shown. The evaluated uncertainties at the level below 3% (k=1) were validated by the comparison of the *in situ* results with an independent measurement performed by colleagues form CNES. The potential for uncertainty reduction in the future was identified.

3.1 Introduction

A consortium formed by NPL, CNES and Magellium was appointed by the ESA to establish a new European radiometric calibration site. NPL's role was to provide SI traceability, install permanent instrumentation on the site and maintain the site operations; CNES is responsible for delivering operational BOA products during the operational phase. In December 2015, NPL and CNES conducted the initial site characterisation that led to the selection of the permanent site location. Magellium runs the RadCalNet portal and in the first phase of the project undertook the global site location search.

Gobabeb, Namibia and the Atacama Desert in Chile were the two shortlisted candidates for which the negotiations with site owners commenced. Agreement was reached with the Gobabeb Research and Training Centre and permission to install a new site granted. Originally the Gobabeb site was supposed to have the same equipment as the La Crau site including the sun photometer and the mast. Gobabeb is located in the Namib-Naukluft National Park, and the permit excludes any damage to the ground. The La Crau mast has a foundation dug into the ground. Thus, the search for a new guyed mast supplier started. The sun photometer needs to be installed on the top of the mast to perform sky, sun and ground measurements. For measurements of the Sun, the instrument has to point at the sun and having a small FOV of 1.2° this means the top of the mast has to be stable with the average winds present on the site to 0.6°.

In the meantime, the preparation for the site establishment was ongoing including planning of the site characterisation measurement protocols and calibration of reflectance standards and instruments.

3.2 Methodology

Two methods with different SI traceability chains are used for the ground reflectance measurements at Gobabeb. The reflectance-based is applicable only during the field campaigns, and for this method, an appropriately calibrated reference standard holds the link to the SI via the reflectance scale. The ground measurements are conducted as the relative comparison to the reference standard. However, there is a slight issue with the calibration of the reflectance standard due to different illumination conditions *in situ*, as already explained in Chapter 2 and a proposed solution is presented and implemented in this research.

The radiance method is used by the instrument that is permanently installed in Gobabeb and measures autonomously every day. The sun photometer is absolutely calibrated against known radiance sources. SI traceability is provided by the spectral radiance scale that can be traced back to the cryogenic radiometer, the primary standard. Then during an operational phase its calibration can be monitored by *in situ* calibration against the sun. That is not SI traceable but provides information about instrument stability. All the pre-deployment calibration and characterisation tests of the sun photometer are presented in this chapter.

At present the radiometer is back at NPL after its first year of autonomous operation, for recalibration and re-characterisation. The results of these tests will allow a through evaluation of any stability changes in its response.

3.2.1 Reflectance SI Traceability

The distinction between laboratory and remote sensing terminology is particularly important when reference reflectance targets are used for ground reflectance measurements, as their calibration values are always aligned with Nicodemus terminology, and can contain only one type of illumination conditions, not the combination of directional and diffuse that is present outside. Moreover, the majority of calibration laboratories provide only two default calibration options. The most popular one, commonly called reflectance is an 8° Hemispherical Reflectance Factor, which means that the sample is illuminated at 8° zenith angle and the reflected signal from the entire hemisphere is collected. This is the case 3 from the Schaepman-Strub et al. (2006) classification presented in Figure 3.1 (b). To follow the reflectance factor notation explained in Chapter 2 this means $R(\theta_i = 8°; 2\pi)$. The second option is the so-called 0:45 geometry, that is a bidirectional configuration, but only for one illumination and one viewing angle, $R(\theta_i = 0°; \theta_r = 45°)$. This option is illustrated in Figure 3.1 (a).

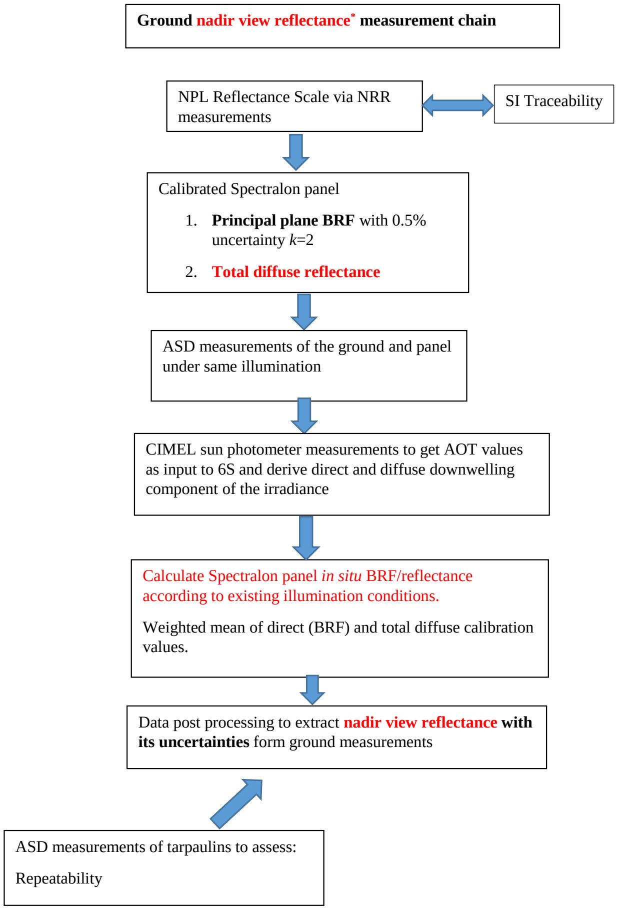
Figure 3.1: Measurements geometries from (Schaepman-Strub et al. 2006) (a) CASE 1 Bi-directional, (b) CASE 3 Directional-hemispherical.

As already mentioned in Chapter 2 the perfect diffuser does not exist. Thus for meaningful results from *in situ* reflectance measurements, BRF calibration for a range of illumination zenith angles that will occur during field measurements is needed. Nevertheless, using BRF calibration values for the reference standard does not provide the correct ground values because the reference panel is illuminated differently to the calibration condition has different reflectance factor values.

It is worth reviewing the definition of calibration from the International Vocabulary of Metrology (VIM) (JCGM200:2012) that states calibration as 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”*.

The part, “under specified conditions”, is not met when the illumination condition is changed.

The laboratory calibration of the reflectance panel cannot include a mixture of direct and diffuse (hemispherical) reflectance values; however, the direct and total (hemispher-



*Final term should be defined and accepted by RadCalNet community

Figure 3.2: Proposed *in situ* reflectance measurements SI traceability.

ical) diffuse reflectance can be measured separately. Both of these quantities have fully SI traceable measurement procedures. The novel approach to the calibration comprises both of these calibration values via a composite of the reflectance due to the directional Sun irradiance and, due to diffuse sky radiance. It will be weighted in the same proportions as the directional/diffuse illumination (for each wavelength and solar zenith angle) at the local atmospheric conditions, as measured by an *in situ* sun photometer. This actually corresponds to the Schaepman-Strub (Schaepman-Strub et al. 2006) definition of *in situ* HDRF $R(\theta_i, 2\pi; \theta_r)$ as shown here again in equation 3.1 (which takes into account complex illumination conditions).

$$HDRF = R(\theta_i, \phi_i, 2\pi; \theta_r, \phi_r) = R(\theta_i, \phi_i; \theta_r, \phi_r)d + R(2\pi; \theta_r, \phi_r)(1-d) \tag{3.1}$$

Thus, to calculate more a realistic ground reflectance that is a function of the SZA and sky radiance, the new improved equation becomes:

$$HDRF_{ground}(\theta_{sun}, 2\pi_{sky}; \theta_r = nadir) = \frac{BRF_{ref}(\theta_{sun})d + TDR_{ref}(1-d)DN_{ground}}{1/2(DN_{ref}(t_1) + DN_{ref}(t_2))} \tag{3.2}$$

Where: d is the fraction of the direct to total radiant flux and TDR is the purely diffuse illumination hemispherical directional reflectance factor called Total Diffuse Reflectance. This is the name of the laboratory calibration quantity.

The new SI traceability route is schematically presented in Figure 3.2. The text marked in the red font of that Figure highlights the new components needed for improved SI traceability. SI traceability of the reference panel is from the NPL Reflectance Scale, where in addition to BRF measurements a TDR is added. Then the *in situ* measurements are conducted as normal, but to estimate the ratio of the direct to total downwelling irradiance the sun photometer measurements are required. They form a part of *in situ* measurements for the atmospheric data. Then the RTC with measured values of AOT needs to be used to estimate the direct and diffuse ratios. Knowing these ratios, the ground reflectance can be calculated using equation 3.2.

If some of these calibration values are not accessible, an error due to incorrect reflectance calibration has to be determined and accounted for in the final uncertainty evaluation.

In preparation for the Gobabeb site characterisation, two Spectralon panels (NPL and CNES owned) were calibrated for spectral BRF. Figure 3.3 shows the plot with the results of that calibration for selected wavelengths. The x-axis is the detector angle

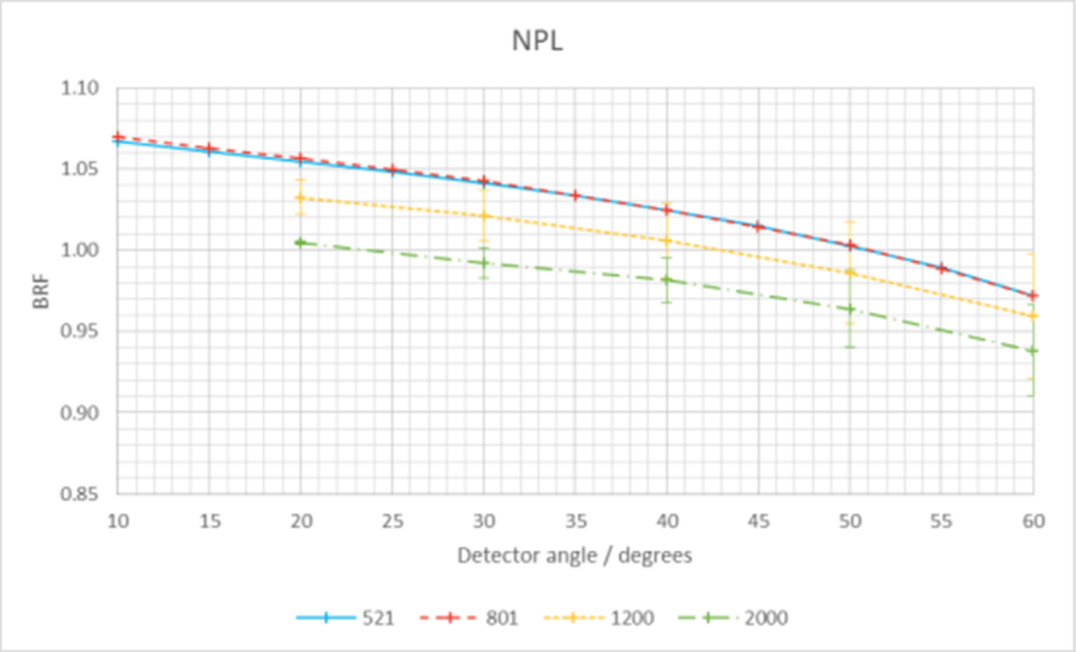


Figure 3.3: NPL Spectralon panel BRF calibration values for selected wavelengths.

that *in situ* is considered as SZA. This set of measurements allows use of actual BRF values during fieldwork, as according to the reciprocity rule for a Lambertian diffuser ($R(0°, 45°) = R(45°, 0°)$); for all measurements at nadir a value for the panel calibration can be found for any solar zenith angle. Also, the same range of viewing geometries was measured for illumination at 5 and 10 degrees to estimate the effect of not having the panel perfectly horizontal, or the instrument perfectly vertical, during the field measurements. The uncertainty in BRF measurement is 0.5%, (k=2) for wavelengths up to 1000 nm, increasing to 2% in the infrared region.

The NPL Spectralon panel was also calibrated for total diffuse (hemispherical) $0/d$ reflectance factor. These measurements are performed on the NRR facility with illumination at zero degrees and scattering angles varying from -85° to 85° in 5° intervals that are integrated to provide one hemispherical reflectance value. The total diffuse reflectance is calculated according to equation 3.3.

$$TDR = R(\theta_i = 0, 2\pi) = \frac{\int_0^{\frac{\pi}{2}} BRF(\theta) \sin 2\theta d\theta}{\int_0^{\frac{\pi}{2}} \sin 2\theta d\theta} \tag{3.3}$$

Where: $BRF(\theta)$ is the bidirectional reflectance factor for a given scattering (viewing) angle.

For diffuse reflectors such as Spectralon the diffuse reflectance with nadir illumination is equivalent to diffuse illumination and nadir viewing (Budde 1976). Therefore, this value can be used to account for the diffuse component of the reflectance during the field measurements.

3.2.2 Instruments Calibration and Characterisation

This section presents the results of laboratory-based tests on an ASD FiledSpec portable spectral radiometer and a CIMEL CE 318 BRDF 12 filter Sun photometer. The tests were performed in preparation for permanent installation at a new Gobabeb site.

The ASD spectroradiometer has a wavelength range of 350 nm - 2500 nm and contains three separate detectors. The VNIR (350 nm - 1000 nm) detector is a CCD array with a grating that is used to disperse the light. In addition, there are two SWIR detectors covering 1001 nm - 1800 nm and 1801 nm - 2500 nm respectively. Both of these have a single detector and an oscillation grating, thus a spectrum is acquired in sequential mode. Firstly, the SWIR detector is thermally controlled to reduce the noise floor. A fibre optic is used to feed the light in to the detectors. The fore optics can be changed as it includes a bare fibre, a lens that defines FOV or a cosine diffuser for irradiance measurements. The ASD was used during the site characterisation phase for ground reflectance measurements.

The CIMEL CE 318 BRDF 12 filter Sun Photometer is a new version of the sun photometer that is used for sun, sky and ground measurements. It contains twelve spectral bands from 414 nm to 1640 nm and has two detectors: a silicon detector for the channels 414 nm, 440 nm, 500 nm, 555 nm, 675 nm, 702 nm, 740 nm, 782 nm, 870 nm, and 1020 nm and InGaAs for wavelengths of 1020 nm and 1640 nm. The detectors are located in two separate collimators. The instrument is capable of recording on three separate gain settings: Sun, Aureole, and Sky (relative gain 1, 128 and 4096 respectively).

The sun photometer instrument was permanently installed on the top of a 10 m mast telescopic at the Gobabeb site in summer 2017. This is now in operation and transmits daily data via the GSM network and the Gobabeb site is about to be officially accepted to the RadCalNet network. The sun photometer was already swapped in June 2018 and at present is back at NPL waiting for a second round of its calibration and characterisation tests after one year of operation *in situ*.

3.2.2.1 ASD Spectroradiometer Laboratory Tests

In order to prepare this instrument for the field campaign the following tests were carried out: absolute radiometric calibration, wavelength calibration, “stability while battery powered” test, especially up to run down time, and the temperature dependence. The first two tests were undertaken to confirm the manufacturer specification and although *in situ* SI traceability is provided from a reference panel, the information about absolute calibration coefficients and their change in time is a good way to monitor instrument stability.

The battery run stability test was of particular importance for field application and this test was repeated a few times showing that the battery status, especially its lower voltage at the time that the battery was due for the recharge, would not affect the measurement data. In addition, this test confirmed quite long warm-up time so that the instrument has to be powered at least one hour before the measurements in order to achieve stability.

The temperature dependence results agreed with previous tests (Hueni & Bialek 2017), that indicated a temperature dependence for some wavelengths, in particular, the long wavelength range of the silicone detector, that does not have an internal temperature stabilisation.

3.2.2.2 Cimel Sun-Photometer Laboratory Tests

Absolute radiometric calibration of CIMEL sun photometers is provided by the manufacturer in two modes, irradiance and radiance, according to the procedure developed for AERONET (Holben et al. 1990). For Sun (irradiance mode) calibration Langley plots are used to determine the extra-terrestrial digital counts as the intercept for an air mass equal to zero. The NOAAs Mauna Loa Observatory is considered to be the community agreed primary standard and has been used since 1993 for AERONET calibration. The sun-photometer was calibrated in irradiance mode by CIMEL at the Canary Islands. This calibration method requires a high altitude and clear stable atmosphere. NPL could provide absolute irradiance calibration but not the extra-terrestial irradiance that is required as AERONET calibration procedure. Thus irradiance calibration was not be repeated at NPL.

Sky and Aureole (radiance mode) calibration is performed using an integrating sphere or other source of SI traceable spectral radiance. CIMEL provided calibration coefficients that were obtained from measurements at Laboratoire d'Optique Atmospherique, Lille, using five radiance levels. According to the AERONET procedure, the five radiance flux levels are measured with the sun photometer, then calibration coefficients are derived as gains from a straight line fit with the assumption that with dark current subtraction the line offset is removed and thus passes through the origin. A 5% uncertainty on radiance calibration is quoted with these measurements (Holben et al. 1990), but the coverage factor is not clearly specified, assumption is made that this refer to (k=1).

At NPL the radiance mode calibration set up consisted of two integrating spheres and a lamp-tile method to compare the manufacturer calibration coefficients and to reduce its uncertainties by using NMI primary standards. Four radiant flux levels were used; however, the radiance levels did not closely match all of the Lille Laboratory ones.

The first radiance source that was used is a lamp-tile system. The radiance is calculated from the spectral irradiance of a known source (in this case a lamp) and the radiance factor of the reflectance standard (the tile). The measurement set-up uses an 18-inch Spectralon panel that is illuminated by a 1000 W tungsten halogen lamp. Both standards have been previously calibrated at NPL. The lamp on the Spectral Radiance and SRIPS (Woolliams et al. 2006) facility has calibration uncertainties (except the short wavelengths) below 1% (k=2). The panel was calibrated on the NRR facility (Chunnilall et al. 2003) for a reflectance factor at 0:45 illumination and viewing geometry with 0.5% (k=2). The CIMEL sun photometer was aligned at a 45° viewing angle to the Spectralon and the lamp was placed firstly at 500 mm, the nominal irradiance calibration distance, and secondly at 1000 mm from the panel, to obtain different radiance levels. The laboratory set up is presented in Figure 3.4, where the sun photometer is pointed as the Spectralon panel.

A second, higher intensity system consisting of a small calibrated (in radiance units) reference integrating sphere was also used. This so called "radiance gauge" was calibrated

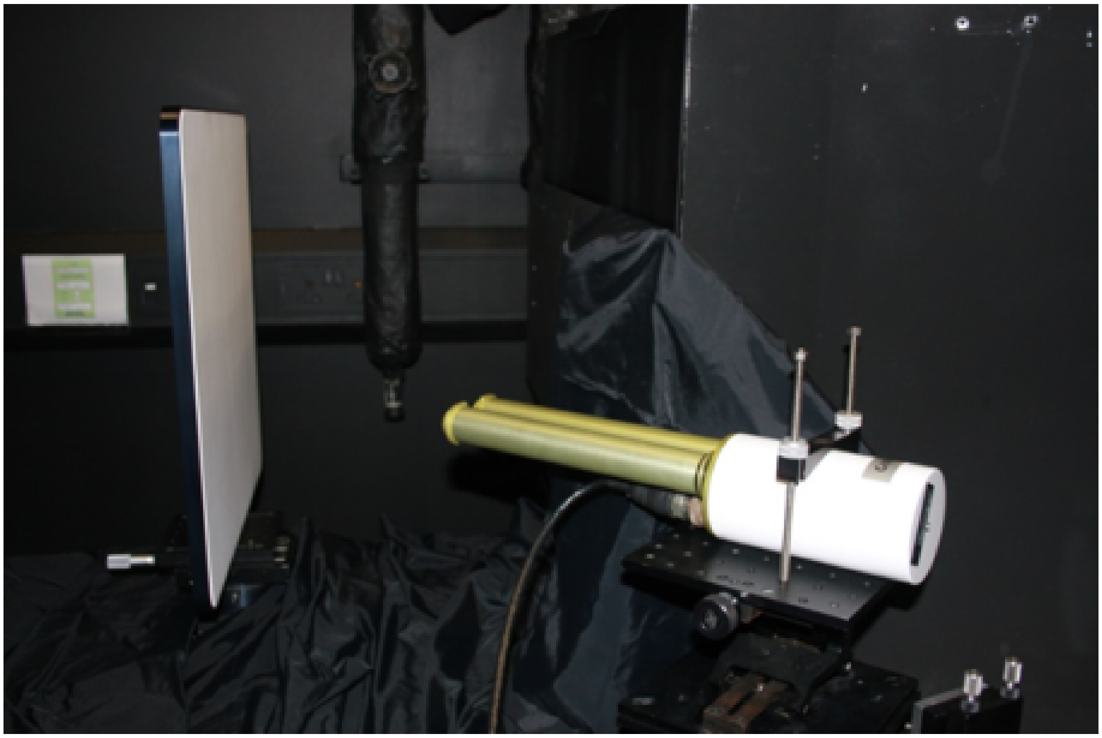


Figure 3.4: Radiance calibration laboratory experiment set up.

on the SRIPS facility, with total expanded uncertainty (k=2) around 2% for the blue end of the spectrum decreasing to 1% for the remaining wavelength range of interest in this case. The sun-photometer was aligned separately for each sensor/collimator so that its field of view was entirely filled by the sphere port. In addition, measurements from the Transfer Standard Absolute Radiance Source (TSARS) (Pegrum et al. 2004) were used for calibration purposes to provide another radiance level point. However, during the timescale of these tests there was no possibility of absolutely calibrating the TSARS sphere on the SRIPS facility. The absolute values were transferred from the radiance gauge via a transfer radiometer, an NPL ASD FieldSpec spectroradiometer. Additional uncertainty components are included to account for any ASD radiometer related effects such as non-linearity that could affect the scale transfer. Thus, the TSARS sphere was measured by an ASD spectroradiometer to obtain its radiance and then by the CIMEL sun photometer.

Radiance measurements were performed with four different radiance sources (two integrating spheres TSARS and "Radiance gauge" and two lamp - tile set ups) on different days, as the laboratory settings had to be changed for each source. The instrument was realigned before each radiance calibration. The measurement sequences included a few BLK (Black) readings to monitor internal instrument noise levels followed by several SKY measurements.

At a later stage of the measurements CIMEL provided an updated software procedure that allowed users to set a number of readings and time intervals between them. Since the software upgrade, 100 light measurements were taken, followed by a shield set. A black coated shield was placed in front of the radiance source and several SKY scenario measurements were recorded to evaluate the level of unwanted room stray light caused by inter-reflections from walls and ceiling.

These room stray light readings were negligible for the sphere sources; below 0.04% for the reference integrating sphere, "radiance gauge", (the brightest source). The same applied for the TSARS sphere, except for the 414 nm channel, for which this value was 0.13%. The lamp tile set-up produces more unwanted light in the room, and in addition the light level was very low for the CIMEL instrument. Due to this, for the lamp measurement, the percentage of dark signal to light was higher, especially for the first four channels. For the 500 mm lamp-tile distance, the amount of light detected during a shield measurement, compared to a full light measurement, was on average 0.12% for channels above 555 nm and 0.24% for 555 nm, 0.5% for 500 nm, 0.9% for 440 nm and 2.8% for 414 nm. For the lower radiance source (1000 mm lamp-tile distance) this was 0.5% for channels above 555 nm, increasing for the remaining channels to 10% at 414 nm, indicating that the signal to noise ratio for this radiance source is lower than desired.

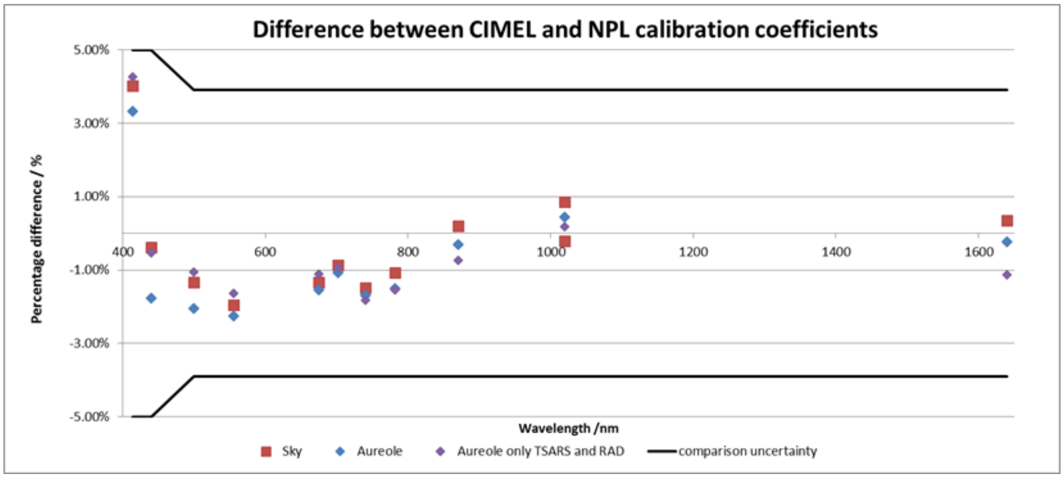


Figure 3.5: Absolute radiometric calibration of the sun photometer, NPL results.

Calibration coefficients were calculated using a straight-line calibration method according to ISO Technical Specification-Determination and use of straight-line calibration functions (*DD ISO/TS 28037:2010*). Cause 7 was applied where uncertainties associated with x and y values are included in the straight-line fit. This is generally a weighted least square approach to derive calibration coefficients that will be applicable to any radiance flux level. Following an assumption from the AERONET radiometric calibration, the line was forced to start at zero, including no intercept in the straight line equation. The slope of the line is the calibration coefficient. Signal uncertainty $u(x)$ was calculated for each measurement point according to the following equation:

$$u(x) = u(x_{\text{light}}) + u(x_{\text{dark}}) \tag{3.4}$$

Where: $u(x_{\text{light}}) = \delta(x_{\text{light}})$ uncertainty in light signal was estimated as a standard deviation of its readings, the same applied to the dark (shield) measurements. Uncertainty associated with radiance was estimated from a calibration certificate of a given standard combined with additional components due to the measurement set up, for example, alignment accuracy and distance uncertainty for lamp tile set up.

Figure 3.5 presents a comparison of NPL and CIMEL calibration coefficients. Generally, all channels, except for 414 nm, are in good agreement, with the difference between them less than 2%. For the 414 nm channel the difference is at the level of 4%. Mostly the disparity for both Aureole and Sky gains are similar for each wavelength, except for wavelengths below 555 nm. At NPL most of the radiance sources used were of a much lower flux than those used by CIMEL for their calibration. Therefore, an additional series is presented on this plot for aureole gain, where only data from the two integrating sphere measurements (the higher radiance level sources) were used to attempt to reduce error. The black line on this Figure represents an overall comparison uncertainty that was combined from the NPL and CIMEL calibration uncertainties. Although 5% is officially quoted by CIMEL calibration, B. Darmini stated in a personal communication that recently they had achieved 3% difference between five integrating spheres. Hence, a value of 3% was used, combined with the NPL calibration uncertainties presented in Table 3.1.

Table 3.1: Sun photometer NPL absolute radiometric calibration relative uncertainty.

Channels	NPL integrating sphere	Additional component due to different radiance sources	Instrument signal	Combined standard uncertainty (k=1)	Expanded uncertainty (k=2)
414	0.93%	0.97%	1.02%	1.68%	3.4%
440	0.92%	0.79%	1.24%	1.73%	3.5%
500	0.56%	0.84%	0.33%	1.07%	2.1%
555	0.53%	0.94%	0.38%	1.14%	2.3%
675	0.47%	0.93%	0.30%	1.08%	2.2%
702	0.46%	0.88%	0.16%	1.01%	2.0%
740	0.45%	0.95%	0.16%	1.07%	2.1%
782	0.43%	1.15%	0.19%	1.24%	2.5%
870	0.42%	1.16%	0.13%	1.24%	2.5%
1020	0.51%	0.97%	0.12%	1.10%	2.2%
1020	0.51%	0.95%	0.22%	1.10%	2.2%
1640	0.42%	1.26%	0.09%	1.33%	2.7%

The measurement equation for calculating a calibration coefficient from a single radiance source is given by:

$$c_{coef}(\lambda) = \frac{L(\lambda)}{x(\lambda)} \tag{3.5}$$

Where: $c_{coef}(\lambda)$ is the calibration coefficient for a given channel, $L(\lambda)$ is the radiance of the reference source at the central wavelength of the given channel, and $x(\lambda) is$ the instruments signal in digital numbers (DN) recorded for a channel. Then, according to the GUM (JCGM100:2008), the calibration coefficient uncertainties for each wavelength were calculated as:

$$\left[\frac{u(c_{coef})}{c_{coef}}\right]^2 = \frac{[u^2(x_{\text{light}}) + u^2(x_{\text{dark}})]}{[x_{\text{light}} - x_{\text{dark}}]^2} + \left[\frac{u(L)}{L}\right]^2 \tag{3.6}$$

To simplify the uncertainty budget calculation, the reference integrating sphere was used as the main radiance source. This has a lower uncertainty, so more weight in the least squares fit. To account for the fact that other radiance sources were used, an additional component was added to the uncertainty budget. This uncertainty was estimated from the calibration coefficients that were calculated for each radiance source separately. The standard deviation between them was used as an indication of this uncertainty. A good agreement between the CIMEL and NPL absolute calibrations was found. NPL sources were less bright. For future sun photometer calibrations, as the lamp tile methods proved to be not bright enough, other brighter radiance sources will be used.

Temperature sensitivity tests were carried out on the CIMEL instrument to ascertain how dependent the response of the instrument is on ambient temperatures, and to additionally assess any wavelength dependence in the temperature response. To control and vary the ambient temperature around the CIMEL instrument it was wrapped in tubing through which temperature-controlled water from a water bath was passed. The instrument was then insulated with foam and bubble wrap and mounted on a stand. One detector/collimator was then aligned to the front of the TSARS radiance source as shown in Figure 3.6.

The temperature of the water bath was set to 5°C, 15°C, 25°C and 35°C although the temperature of the CIMEL instrument, recorded by an internal thermistor, was for

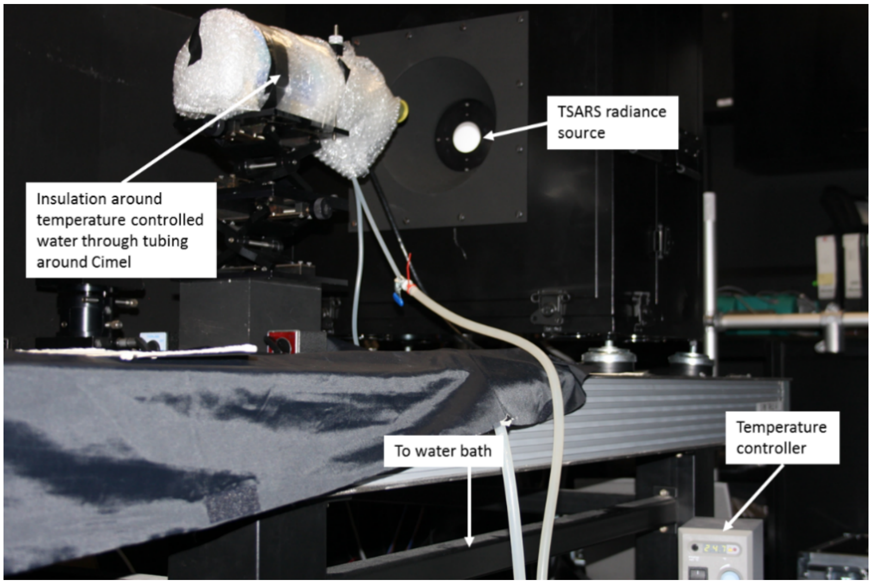


Figure 3.6: Set up of CIMEL instrument for temperature stability tests. CIMEL is mounted and aligned in front of TSARS.

each set point a few degrees closer to room temperature, this latter temperature of the radiometer was used for all calculations. At each temperature 100 readings were taken using the Sky gain setting followed by three readings with the Blk setting and three shield readings (Sky gain scenario, with a shutter between the CIMEL and TSARS). The procedure was then repeated going through the temperatures in reverse order. The detector aligned to the centre of TSARS was then switched over (from silicon to InGaAs) and the procedure repeated.

The previous reports (Holben et al. 1990) and manufacturer data tested only the Silicon detector, and established that only the 1020 nm channel had a temperature dependence at the level of 0.25%/°C±0.05%/°C. The results obtained in this research show a slightly higher (0.36%/°C) change with temperature for this channel, and generally confirmed that the remaining channels were not temperature sensitive, with the exception of the 870 nm channel, which exhibited a very small change. Generally, all channels presented slightly lower signals for the lowest tested temperature. However, only for the 870 nm channel is this change as large as 0.77% and higher than the worst case stability variation, which was recorded at the level of 0.5%. This indicates some small temperature effects, estimated from the polynomial fit to the measure points at a maximum of 0.08%/°C for low temperatures. This effect will be negligible for the Gobabeb site, as the average temperature during day light hours should not drop below 10°C.

The InGaAs detector is a relatively new addition to the CIMEL sun photometer series, and no previous tests have been reported. Temperature dependence for both infrared channels was observed. Figure 3.7 shows the percentage difference in signal compared to a temperature of 25°C for the InGaAs detector wavelengths of 1020 nm and 1640 nm. The change is negligible for temperatures above 25°C, but is significant for the temperatures below, where the change is 0.45%/°C. This result is indeed not expected. InGaAs detectors do not tend to exhibit temperature dependence and they are temperature stabilised to reduce the noise of the detector. The interference filters used to define spectral bands might change with temperature and this is related to spectral shift in their transmittance. However, all glass filter manufacturers provide the temperature characteristics for temperatures of 25°C and above. No records were found for the lower temperature. This test will be repeated this year and investigated further with the instrument's manufacturer.

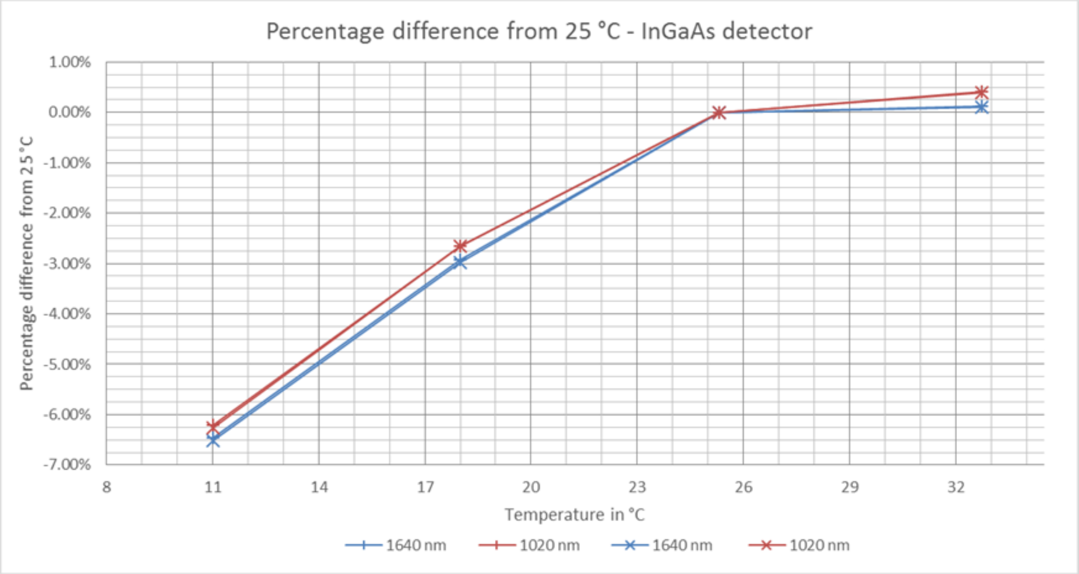


Figure 3.7: Percentage difference in signal at a given temperature compared to the signal at 25°C for wavelengths of the InGaAs detector.

The uncertainty in temperature sensitivity results was estimated by modelling. Two separated lines were fitted to the data with minimum observed temperatures for each nominal temperature measurement point during the tests and the second with the maximum temperatures. The difference in estimated temperatures using these two lines allowed the estimation of the uncertainty of these measurements. This uncertainty was estimated as 0.04%/°C.

To derive a percentage change per degree, a line was fitted to channels that exhibited temperature dependence. The linear fit was used for the InGaAs detector; however, two separate lines were fitted: one for temperatures below 25°C and a second one for higher temperatures. For the silicon 1020 nm channel, linear and third degree polynomial were fitted and for the 870 nm channel just a 3rd degree polynomial line. For all fits R^2 values were above 0.98, however none of these fitted data were validated, thus they are presented here for information purposes but do not contain accurate quantified values.

Nevertheless, we can see that for lower temperatures there might be a significant change in the signal levels for all infrared channels. This particular instrument will be installed at Gobabeb with moderate temperature conditions. We do not expect to operate in temperatures lower than 10°C or much higher than 30°C. For this temperature range the correction for 870 nm channel is not necessary, the 1020 nm silicon channel requires a temperature correction. Both InGaAs channels for temperature below 25°C need to be corrected as well.

3.3 Site Description

The Gobabeb calibration site S23°36′72″ and E15°7′10 416″ is located in the north of the Namib desert in Namibia. The site is located at 505 m above sea level. The closest road is 1.8 km away and the Gobabeb Training and Research Centre is 10 km away from the site. The surface is mainly composed of loose gravel on sand, without vegetation. A sand sample collected during the field campaign was analysed by Maxime Lamare at the Royal Holloway University of London. An SEM (Scanning Electron Microscope) image analysis of random samples taken across the site during the campaign show a typical grain size ranging from 10 μm to 100 μm. An XRF (X-Ray Diffraction) analysis of the samples indicates a large presence of Quartz and Feldspar, with traces of Silicates, Mica and Calcitep.

Figure 3.8: Photograph of the Gobabeb site.

A specific target area measuring 300 m x 300 m was selected from the predetermined zones for characterisation campaign based on a visual assessment and the GSM coverage check. As the whole area has a patchy 2G coverage it was essential to ensure the GSM coverage before conducting tests, as the permanent instruments require that to transmit the data.

Figure 3.9: Photograph of the Gobabeb site (zoomed in).

3.4 Measurements

Prior to the Gobabeb campaign, a trial field trip to La Crau was conducted. This allowed to test the measurement protocols. The main finding from this trip was that a hand-held method gave a higher standard deviation of measurement than a method with the instrument fore optic mounted on an arm attached to a tripod. In addition, up to a 1% difference was observed in the reference panel measurement between the tripod mounted and hand-held methods that were mainly attributed to the effects related to the shadowing caused by the operator.

Thus, in the Gobabeb both teams were using a similar measurement protocol that consists of static ground nadir view measurements. Two ASD FieldSpec spectroradiometers and two 18-inch Spectralon panels were used. The optical lens (the fore optics element of the spectroradiometer, which defines its field of view) was mounted on an arm attached to a tripod, which during measurements was only supported by a kneeling operator as a counter balance for the weight of the fore optics and arm. The length of the arm depends on the length on the optical fibre, and was 2 m for the CNES team and 1.2 m for the NPL team.

The measurement of a single point consisted of a short sequence that took approximately five minutes: a Spectralon measurement, four ground measurements (eight for CNES) at slightly different positions, and then a final Spectralon measurement. For the NPL method, each of these included five readings from the ASD at each position (each of 10 scans), which were then averaged. For the CNES method, each point position included twenty scans which were averaged by ASDs software. The operators then move the reference standard to the next point, and carried out the same measurement procedure there.

The NPL team applied a gimbal mount for the lens that allowed it to maintain a nadir viewing angle by default, reducing measurement uncertainty related to the viewing geometry. However, the ASD instrument did not have an optical fibre long enough to fully take advantage of the gimbal mount. The gimbal mount following the Gobabeb trip was successfully adapted by the CNES team for their future measurements. One quality check that was used to assess whether a measurement point was valid was to check the difference between the reference standard measurement at the start and end, and to remove that point if they differed more than 1% for the wavelength range 500 nm to 1200 nm. Figure 3.10 presents the difference between two panel measurements for all ground measurements, the blue series indicates the quality assured data, whereas for the red series the measurements that were discarded as the difference between them was bigger than 1%.

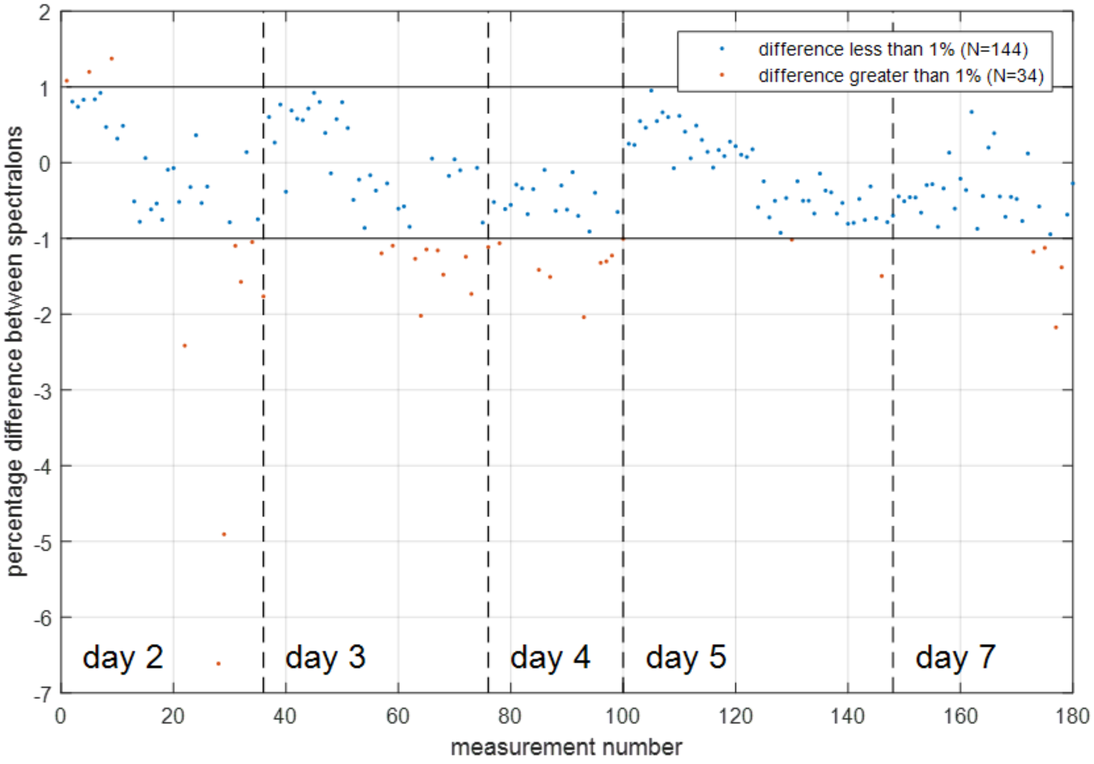


Figure 3.10: The differences between Spectralon panel measurements before and after each measurement point.

The repeatability of measurements was assessed using a set of three large (2 m by 10 m), uniform reference tarpaulins which were brought to the site. These were manufactured to have Lambertian reflectance properties and three different grey scale levels. They were previously calibrated for total diffuse reflectance and 0°:45° radiance factor at NPL. In the field they were used to provide a large uniform area over which we could perform comparisons of different measurement procedures, and investigate how much effect changing the ASD operator had on the results.

3.4.1 Sampling Strategy

The sampling idea is based on a ten-meter-long square with four marked corners that forms the basic measurement unit. Measurements are taken at each corner of the square in the following sequence:

- Reference panel;
- Several ground measurements covering an area of 1 m^2;
- Reference panel.

The measurement sequence lasted for about five minutes. Then the panel is moved to the next corner and the measurement sequence is repeated. When the 10 m by 10 m square measurement is completed the operators move to a next basic unit of 10 m by 10 m. This method was used to take measurements over wide areas of the site. The NPL team took repeated measurements over the same small set of points to obtain an understanding of repeatability and stability, while CNES took measurements over a much wider area to understand surface homogeneity on different scales.

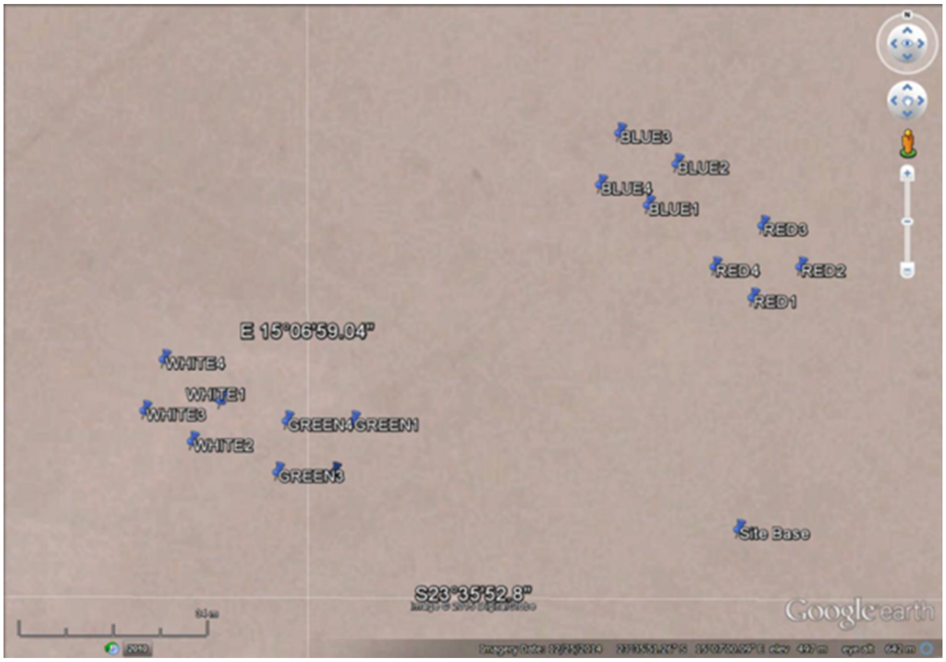


Figure 3.11: Gobabeb, location of NPL *in situ* measurements.

NPL area characterisation was performed over a set of four squares designated Red, Blue, Green and White, with four points each, shown in Figure 3.11. While CNES took measurements of the base squares 100 m away repeating the same procedure to obtain a 100 m by 100 m area measured in all corners. In addition, a 1.5 km transect was characterised with a series of eight reflectance measurements acquired every 25 m over 1.5 km in a direction roughly perpendicular to the area covered by previous measurements.

Figure 3.11 shows the locations of the measurement sites. Site Base (S23.60°,E15.12°) is the location of the CNES CIMEL instrument by the side of the access road. This instrument remained in position taking measurements for the duration of the campaign.

In addition, NPL conducted the BRF measurements using GRASS instruments. The structure has three arms separated by 30° in azimuth with 15 fore optics mounted in 10° zenith angle steps, from 10° to 50°. Each fore optic consists of a 12° (full angle) collimating lens coupled by an optical fibre to a multiplexer that sends a single output to the spectroradiometer. An additional fore optic is located at the nadir viewing position, measuring a circular footprint with a diameter of 4 cm. A 24-inch Spectralon reference diffuser was used as a standard.

However, its calibration did not cover any of the geometries apart from the principal axis (i.e. illumination and viewing azimuths in the same plane). Thus, these measurements are considered as relative knowledge about directional properties of the site, but not SI traceable. Maxim Lamare, who was then a PhD student of the Royal Holloway University of London and NPL, was responsible for the GRASS measurement.

The site surface degradation due to *in situ* measurements can be seen from space, as is presented in Figure 3.12. This is a PLEIADES image taken on 18th December 2015 about two weeks after the field campaign. The darker circular shapes are the marks after the GRASS instrument and the straight lines are the waking paths. Less visible are NPL measurement squares that are located above and below the left-hand site GRASS marks.

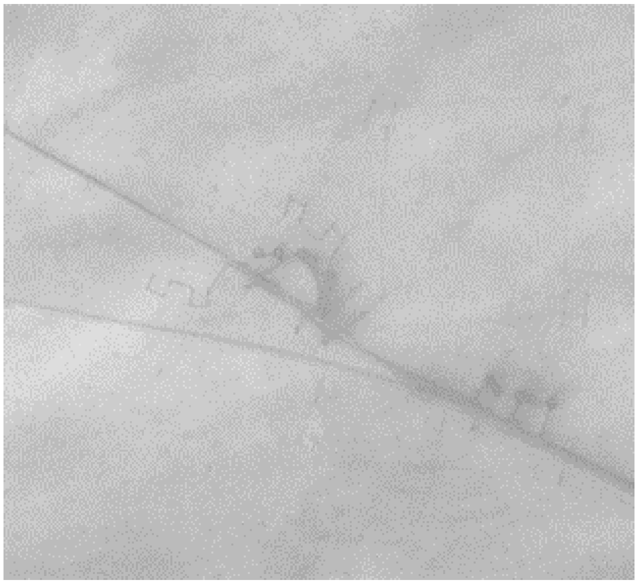


Figure 3.12: PLEIADES 70 cm panchromatic images of the site: taken 18th December 2015, approximately two weeks after the field campaign (darkened areas represent the surface damaged due to foot traffic and darkened circles GRASS positions).(Copyright CNES, Distribution Airbus Defence and Space).

3.4.2 Reflectance Standard Degradation

On arrival in Gobabeb the NPL team discovered that their Spectralon box had been damaged during the shipment. It was not possible to fully repair the damage, so during the campaign the NPL reference panel did not have a fully enclosed box to shelter it from the elements. Spectralon stability from day to day became a big challenge. Due to windy conditions during the measurements and the very dry ground surface, the Spectralon panels were rapidly covered with dust and sand particles. To assess the effects on Spectralon reflectance, from day four comparison measurements with a small reference panel were performed daily. The same reference panel was used only for this short inter-comparison test and was exposed as little as possible to the environmental conditions. After the campaign, all panels (CNES, NPL, GRASS and the small reference panel) were rechecked at NPL for the 0:45 reflectance factor. Figure 3.13 presents the percentage difference between the 0:45 reflectance factors for the panels before and after the campaign. The small reference panel is a 10 cm diameter and

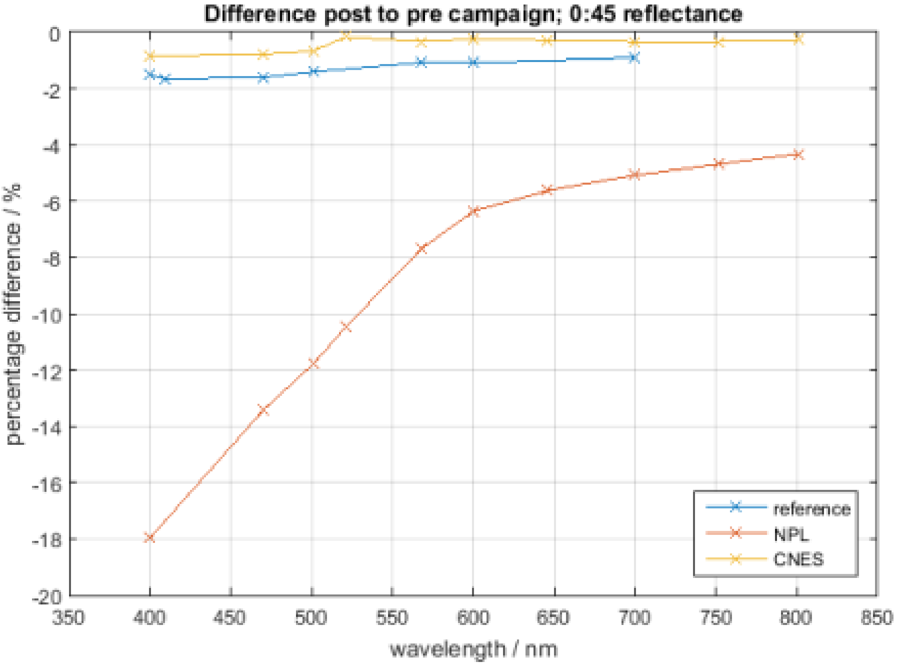


Figure 3.13: Percentage difference of reflectance factor 0:45 values before and after the Gobabeb campaign.

5 mm thick panel that comes with ADS instrument purchase as an accessory. The change in the reflectance factor was the lowest for the CNES panel presented as yellow series in Figure 3.13, which suggests that this panel was already contaminated and only very slightly changed further. The reference panel (blue series in Figure 3.13) that was used comparatively rarely still changed more than the CNES panel. The NPL panel exhibited the biggest change in its reflectance values (red series in Figure 3.13). It seems that new panels change significantly during their initial exposure to outside environmental conditions, even if this exposure is kept to a minimum (e.g. the reference Spectralon). If a panel was already used in the field (i.e. the CNES panel), its reflectance has already changed, so further change during the field measurements is minimised. Nevertheless protection of the panel from elements *in situ* is necessary to avoid any further contamination.

The NPL panel demonstrated the worst-case scenario, as a new panel was brought to the field and, in addition, its sheltering box was damaged and not able to fully protect the panel when it was not being used. To avoid similar issues in the future, the ageing of

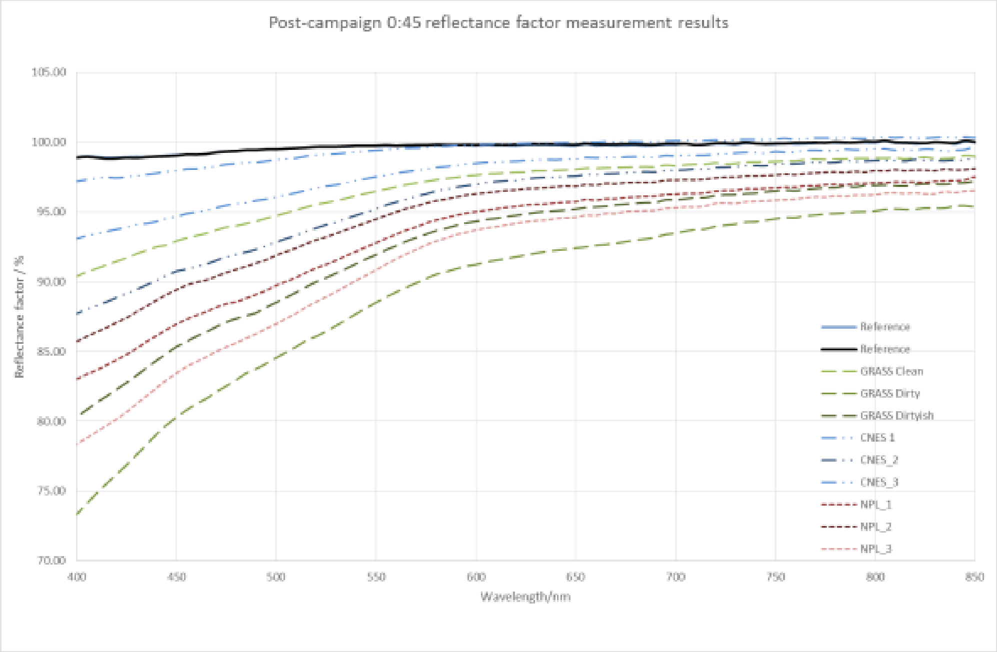


Figure 3.14: Post-campaign reflectance factor of Spectralon panels measured at different points on each panel.

new panels with direct sun illumination and outside environmental conditions, followed by a careful calibration prior to the field campaign, could minimise the changes in the panel reflectance.

In addition, the panels were not contaminated uniformly. Figure 3.14 presents the post campaign reflectance factor for three selected areas on each panel that were defined as clean, dirtyish and dirty. NPL laboratory measurements examine a relatively small area with a beam diameter of 1 cm, so the different results from each small area can be averaged for ASD measurements as the FOV is much bigger (depending on actual setting with diameter around 9.8 cm). Thus some of this non-uniformity will be reduced for *in situ* measurements.

In order to address the Spectralon degradation issue the following steps are performed. Firstly, the verification that the degradation in panel reflectance is the same for all geometries. This was checked with post campaign NRR measurements and confirmed the same dimming (degradation in its reflectance) at all illumination angles. Secondly, the Spectralon degradation *in situ* was traced by the ASD radiance readings of the

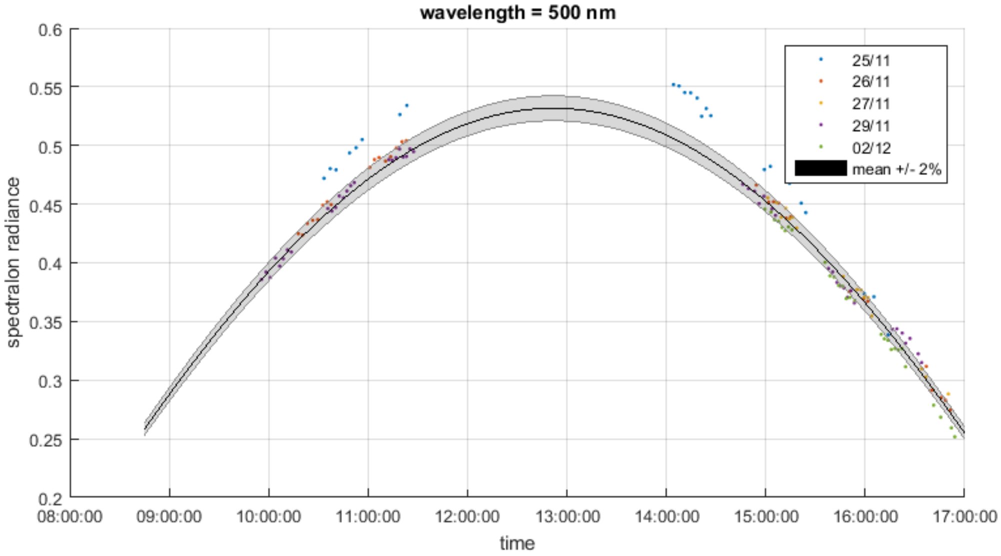


Figure 3.15: Spectralon radiance measured *in situ* over several different days.

Spectralon panel throughout each day. This is only valid with the assumption that the atmospheric conditions are consistent between each day. Figure 3.15 presents the results of the ASD Spectralon radiance measurements over the course of the field campaign. All measurements (except day 2 - 25th November) are in agreement within 2%, except at larger SZAs. This gives us an indication that the Spectralon degradation was quickest within the first two days of measurements on the 24th, when tests of the measurement procedures were conducted, and the 25th, during the first day of characterisation tests. For the data analysis the post campaign Spectralon calibration values are the correct ones, and measurements from the 25th November are excluded from the final data, as the Spectralon reflectance values for that day are not known accurately enough.

3.5 Results

Five days of measurements provided in total 24 runs of the NPL data. The run is defined as one set of measurements of a base unit (a 10 m long square, at each corner). During each measurement day the SZA changed from 18° to 56.5°. The appropriate SZA was used for calculation of the ground reflectance data according to the time of the *in situ* measurements and the interpolated values of Spectralon BRF from the post campaign calibration.

Figure 3.16 presents calculated ground BRF using post campaign NPL Spectralon calibration BRF values for a given SZA during the measurement. The diffuse reflectance component is not included in this plot. The red series represents the average of all measurements at the Red 4 location (see map in Figure 3.11) with their standard deviations, whereas the blue series represents the mean of all 16 points. Each point was measured on at least three different days and in the morning, as well as in the afternoon.

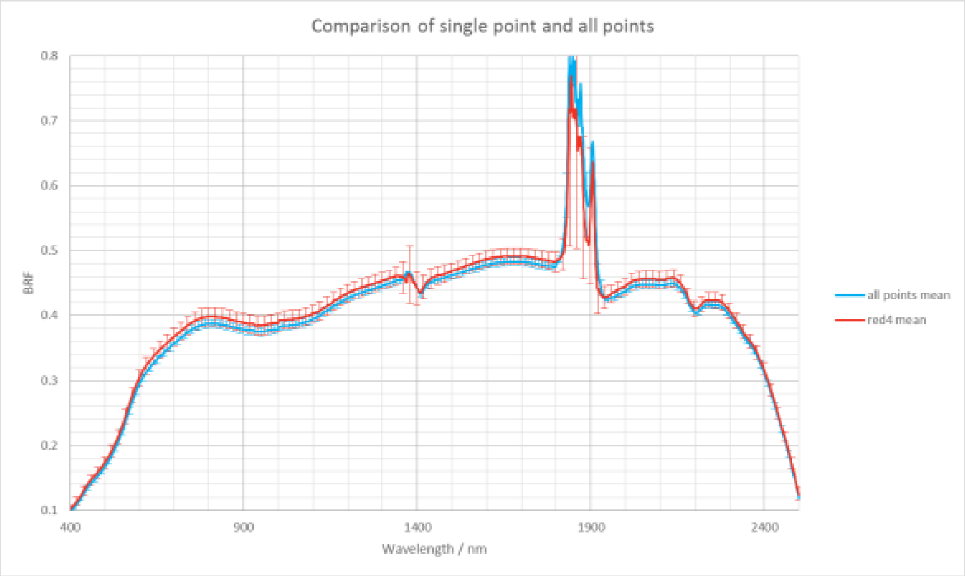


Figure 3.16: Average BRF($\theta_{sun}; \theta_r = nadir$) values for every point, at every time, and average for a single point, at every time, with standard deviations.

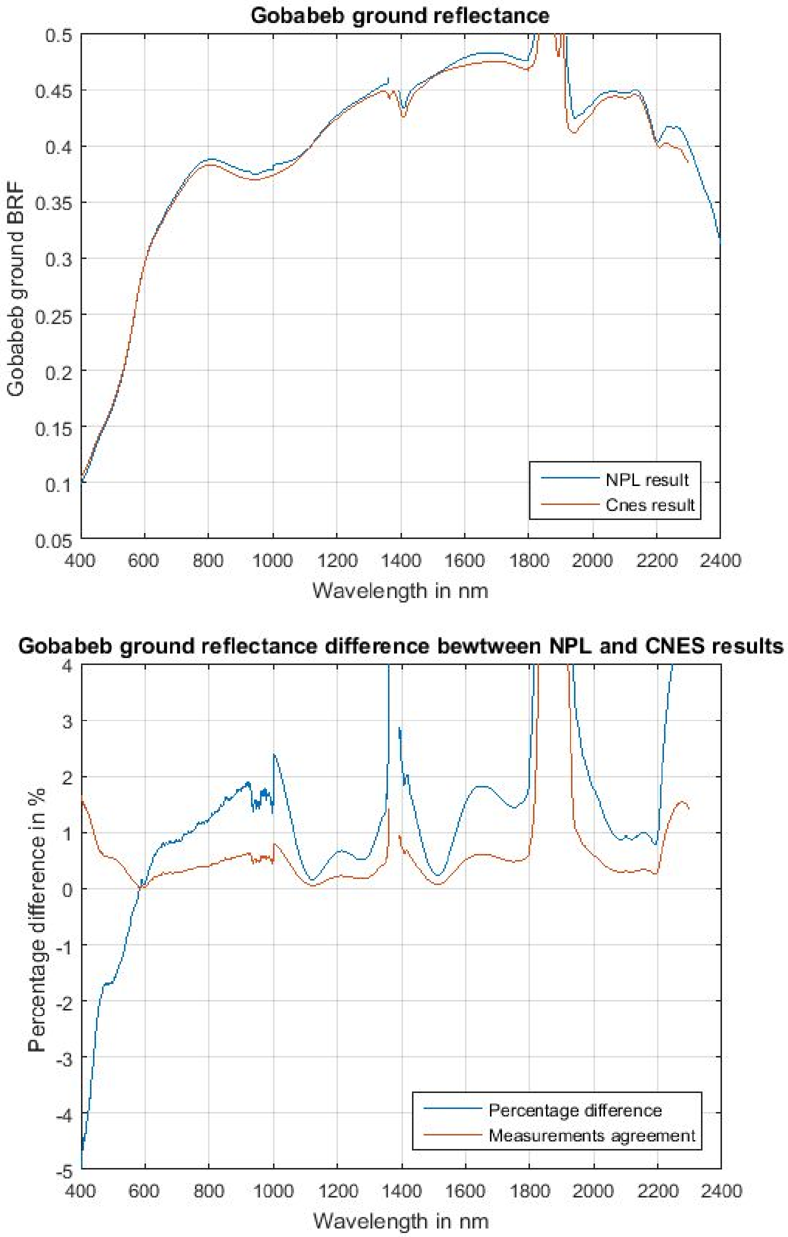
The highest standard deviation can be seen for measurements of individual points. These measurements include measurements of four different ground areas of roughly 9.8 cm in diameter in close proximity to each of the square comers. The average standard deviation of all four corners is lower than for the individual points.

The overall ground reflectance results combining all measurement points from NPL and CNES are shown in the top panel (a) of Figure 3.17 and panel (b), presents the percentage difference between the final results for the two groups and a measure of the measurement agreement, which was calculated as:

$$E = \frac{|NPL - CNES|}{\sqrt{u^2(NPL) + u^2(CNES)}} \tag{3.7}$$

Where: E indicates the measurement agreement, the absolute percentage difference between NPL and CNES results is divided by the sum of squares of their *in situ* measurement related relative uncertainties. Thus, without absolute calibration uncertainty components as both panels were calibrated at NPL, or any effects from the direct/diffuse illumination conditions, as both teams used direct illumination only for the final calculation. The results are therefore affected by these factors in a similar way (see Table 3.2 for a full uncertainty budget). A value of E, which is less than one, indicates that the results agree within their uncertainties. One uncertainty value was used for the entire wavelength range: 2.6% for NPL, and an estimated value of 1.5% for CNES (a lower value, due to much smaller issues with Spectralon degradation, and a longer ASD fibre).

The results from both teams are in agreement within their *in situ* measurement uncertainties at the (k=1) level, with the exception of wavelengths below 435 nm where the E factor is greater than one. For this wavelength range we should actually increase the uncertainty for the *in situ* measurements, as the short wavelengths are most affected by Spectralon changes due to exposure, and the elimination check for Spectralon measurement consistency was performed only on the wavelength range 500 nm to 1200 nm.



(a)

(b)

Figure 3.17: (a) Gobabeb ground reflectance as the average results from all NPL and CNES measurements. (b) Difference between NPL and CNES *in situ* ground reflectance final results. The blue series presents the percentage difference between the overall NPL and CNES results, and the red series represents measurement agreement which accounts for differences in the size of the uncertainty components associated with the measurements and the principle cause of the difference between the NPL and CNES data.

The multi-angular measurements performed using the GRASS instrument gave the very first sets of information about angular HDRF properties of the Gobabeb gravel plains. These results shown in 3.18 illustrated very good uniformity in the forward scattering direction, and a peak in the backscatter direction. The surface does not exhibit significant angular HDRF features, so is well suited for a radiometric calibration site.

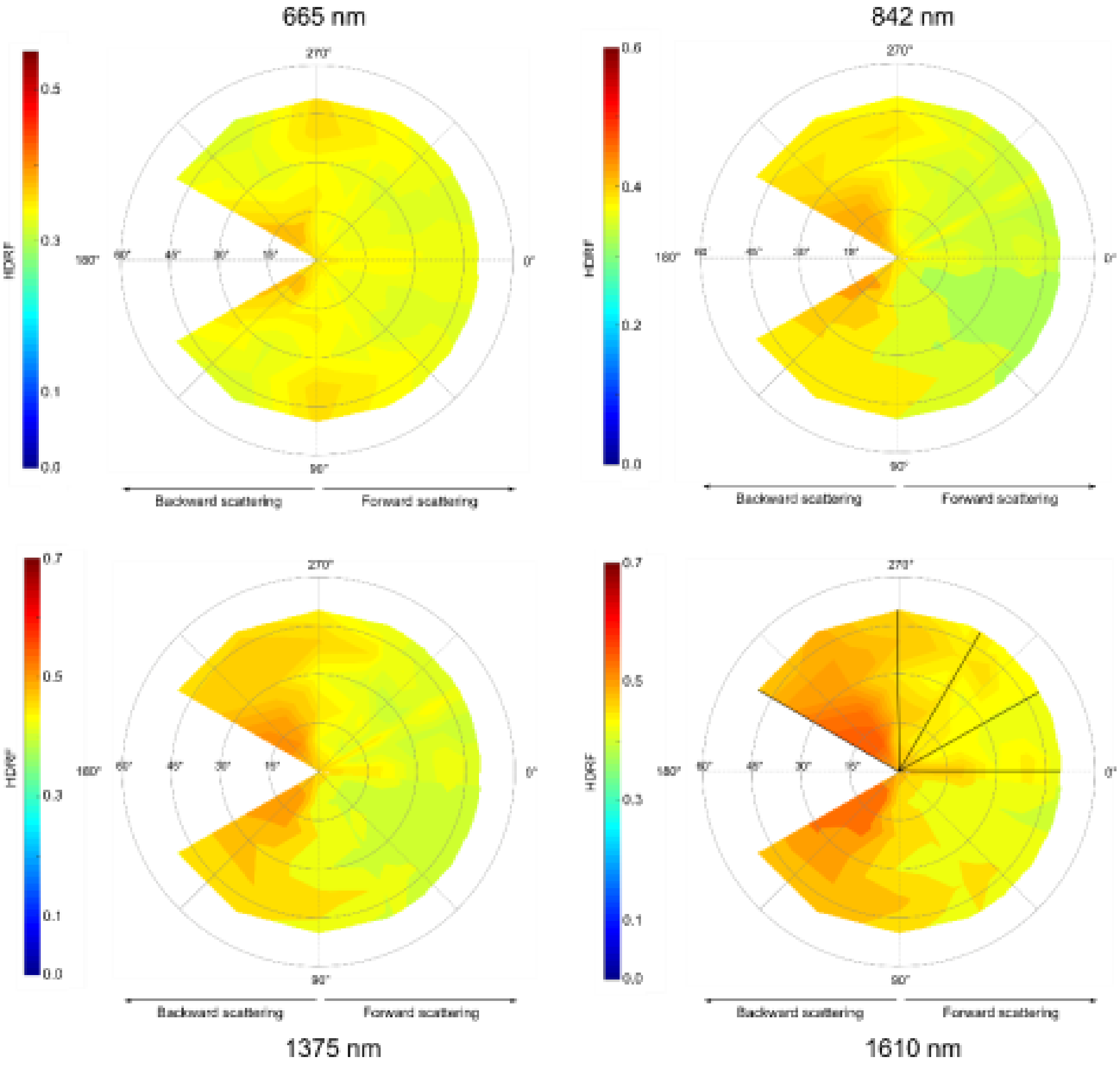


Figure 3.18: Polar plots of the mean angular distribution of reflectance over the two sites at 4 wavelengths matching the centre of Sentinel 2 bands. The solar azimuth angle was fixed at $\phi_i = 180°$. The solar zenith angle varied over a range of $\theta_i = 25° - 35°$ during the acquisitions. From (Lamare et al. 2016)

3.5.1 Uncertainty Evaluation

To evaluate the uncertainty associated with *in situ* measurements sources of uncertainty contributors need to be identified. The measurement equation is simple, with multiplication and division only, thus the sensitivity coefficient for all of the equation components are equal to 1. The uncertainty associated with each component of the equation are combined in a simple sum of squares from (JCGM100:2008).

$$u_{R_{ground_rel}} = \sqrt{u^2_{R_{ref_rel}} + u^2_{DN_{ground_rel}} + u^2_{DN_{ref_rel}}} \tag{3.8}$$

Where: $u_{R_{ground_rel}}$ is the standard relative (k=1) uncertainty of ground reflectance measurement, $u_{R_{ref_rel}}$ is the standard relative uncertainty of the reference panel reflectance, $u_{DN_{ground_rel}}$ is the standard uncertainty of the instrument ground reading, $u_{DN_{ref_rel}}$ is the standard uncertainty on the instrument reading during the panel readings.

However, each of these components listed in equation 3.8 contains a few subcomponents that are listed in Table 3.2.

The reference standard reflectance uncertainty contains its calibration uncertainty and due to move from the calibration facility to *in situ* a change in its calibration occurs (see section 3.4.2). Moreover, the area of the panel seen during calibration is different to the FOV of the instruments used in field therefore uncertainty due to the uniformity in FOV is needed.

The area covered by each of the laboratory-based measurements is just less than 1 cm^2, whereas *in situ* measurements are performed with an ASD instrument that sees a much bigger area of the panel. To estimate a correct calibration value for the panel, an average from all four post campaign measurement points is taken. The measurements are treated as four separate measurements of the panel, which need to be combined into one value, and the uncertainties need to be treated accordingly. The results from all four 0:45 points (one from the NRR and three from the commercial instrument) were averaged to get a single value for the Spectralon panel reflectance.

The standard deviation between these four points was calculated and included in the uncertainty calculation as a representation of the non-uniformity of the panel over the centimetre scale. Each of four measurement points has the absolute random and systematic uncertainties, as during measurements some sources of uncertainty are systematic, thus regardless how many time the measurement is repeated these are always present at the same level. Whereas other are random (for example noise in the detectors) these random components can be reduced increasing a number of measurements and when combined can be divided by a square root of a number of repeats. Therefore, for each point random and systematic components were combined separately in quadrature. The random uncertainty from each measurement, and the standard deviation between all the points were combined into one representative random uncertainty for 0:45, which contains the effect of the Spectralon non-uniformity over the centimetre scale.

The final uncertainty related to the panel reflectance for the measurements at NPL, i.e. across a spot of less than 1 cm^2. Its value mainly depended on panel non-uniformity and the fact that for a non-uniform surface a small measurement area is not very representative. Panel reflectances measured in the field are essentially averaged over the footprint of the ASD, which comes to about 75 cm^2. The measurement of a panel can therefore be treated as 75 points of the NPL measured panel reflectance, with associated uncertainty. Systematic uncertainty remains the same, but random uncertainties including non-uniformity effects are combined, taking into account the number of points that would fit within ASDs FOV, so are lower.

Figure 3.19 shows the uncertainties due to non-uniformity of the panel for different measurement scales as a percentage across all wavelengths. The value 0.22% is used in addition to calibration uncertainty to account for panel non-uniformity due to the contamination *in situ*.

To distinguish site homogeneity variation from the measurement variation, the results of tarpaulin tests were used to estimate instrument repeatability and measurement reproducibility.

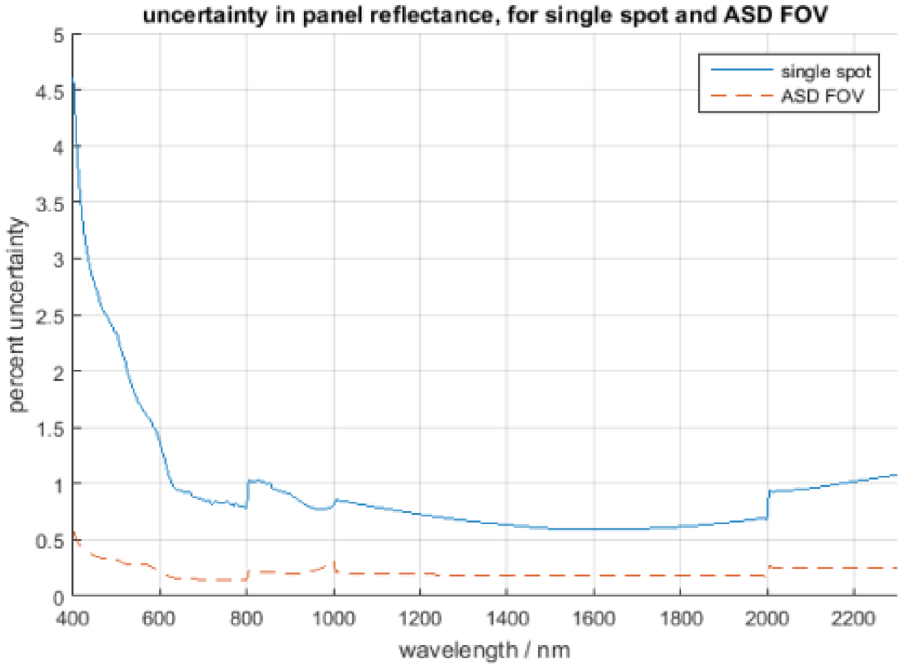


Figure 3.19: Estimates of the percentage uncertainty in panel measurements due to non-uniformity over the NPL measurement spot, and the ASD FOV in the field.

Table 3.2: Uncertainty budget for site characterisation measurements at Gobabeb

Equation components	Sources of uncertainty	Estimated values ($k = 1$)
$u_{R_{ref_{rel}}}$	Absolute BRF and total diffuse calibration	Below 1000 nm 0.25% Above 1000 nm 1%
	Change since calibration (Estimated from post calibration and daily Spectralon radiance consistency)	1.7%, 1% for CNES
	Uniformity within instrument FOV	0.22%
	Diffuse sky illumination (if not included in the calculation)	1.5% short wavelengths decreasing to 0.5% for longer wavelengths
$u_{DN_{ref_{rel}}}$	Instrument repeatability (short time effects)	0.2%
	Reference panel measurement reproducibility (includes pointing accuracy and operator effects for a given measurement protocol)	0.46%
$u_{DN_{ground_{rel}}}$	Instrument repeatability (short time effects)	0.2%
	Ground measurement reproducibility (includes pointing accuracy and operator effect for a given measurement protocol)	1.85 % 0.98% for CNES
Total $u_{R_{ground_{rel}}}$		2.8% - 3.0% (depending on wavelength) 1.9% - 2.1% (depending on wavelength for CNES)

A set of three large (2 m x 10 m), uniform reference tarpaulins was brought to the site. These were manufactured to have Lambertian reflectance properties and three different grey scale levels: light gray, medium gray, and dark gray with the values of diffuse hemispherical reflectance at 55%, 22% and 12% respectively. They were previously calibrated for total diffuse reflectance and 0°:45° radiance factor at NPL. In the field they were used to provide a large uniform area over which we could perform comparisons of different measurement procedures, and investigate how much effect changing the ASD operator had on the results. The tarpaulins were laid down next to each other during the morning hours when the wind was generally much weaker than in the afternoons.

The tarpaulin tests enabled estimation of the instrument repeatability and measurement reproducibility, as they provided uniform surfaces. Standard deviations from different readings were used to estimate operator related uncertainties. Four runs were performed on each of the three tarpaulins by both NPL operators in the same manner as the ground measurements.

The standard deviations of those measurements were then used to estimate operator effects during ground reflectance measurements. All the tests were performed during one morning within a short time. Thus an assumption was made that all other measurement conditions were constant during these measurements and the only effect on the final results (other than instrument effects) was operator influence.

The instrument repeatability is estimated by the standard deviation of 10 ASD readings taken in one position, pointing at each of the tarpaulins and the Spectralon. This test shows the random noise of the instrument over a short time scale and the static pointing accuracy when an operator holds the tripod and ASD is pointed at the same point on the target. Figure 3.20 presents these results for 16 independent runs of ASD readings. The plots show very low standard deviation, generally at the level of 0.1%. However, often one or two runs present significantly higher standard deviation. This could be caused by some movements, for example the operator was not able to stay still during the recording, or there was a slight time misalignment between the actual instrument recording and operator positioning.

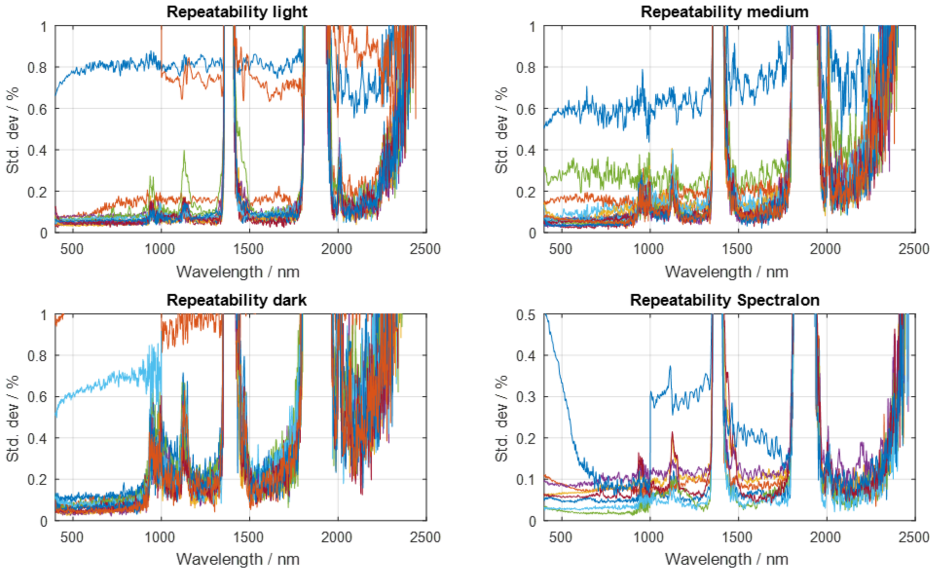


Figure 3.20: Tarpaulins tests results. Standard deviation of all individual runs containing 10 ASD readings.

Some data filtering could be applied to remove the runs with much higher standard deviations. However, including those runs only increased the final value to 0.2%, which is relatively low in comparison to other uncertainty components. In conclusion, it was decided that it was unnecessary to remove any runs. An interesting feature on the repeatability for the dark tarpaulin (and to a lesser extent the medium grey tarpaulin as well) was a significant increase in the standard deviation for NIR and SWIR regions. This could potentially be an effect related to the rapid change of tarpaulin surface temperatures, as the dark one became very hot and by the end of the tests it was burning hot. More tests to fully understand this effect will have to follow. Another possible source of this error is the much lower signal in these regions.

To estimate the operator effect on the measurements, the standard deviations of one measurement series are compared, where one series consists of four rotations of the ASD arm mounted on the tripod. Each rotation records 10 ASD measurements. Each rotation points at a different area on the tarpaulin surface. Basically this exercise mimics the area characterisation tests, so the difference between standard deviation of one of these two sets will allow us to see any operator related effect, and separate the actual site small scale variability. The bottom panel from Figure 3.21 presents the results from these tests. This test was repeated two times by each of the operators, so four runs of measurements for each tarpaulin are averaged. Spectralon measurements were taken at the beginning and the end of each measurement run, including all three tarpaulin measurements. Consequently, the time difference between two Spectralon measurements was roughly three times longer than the time difference that occurred during the area characterisation. The standard deviation here is from the radiance readings so are not corrected for SZA change.

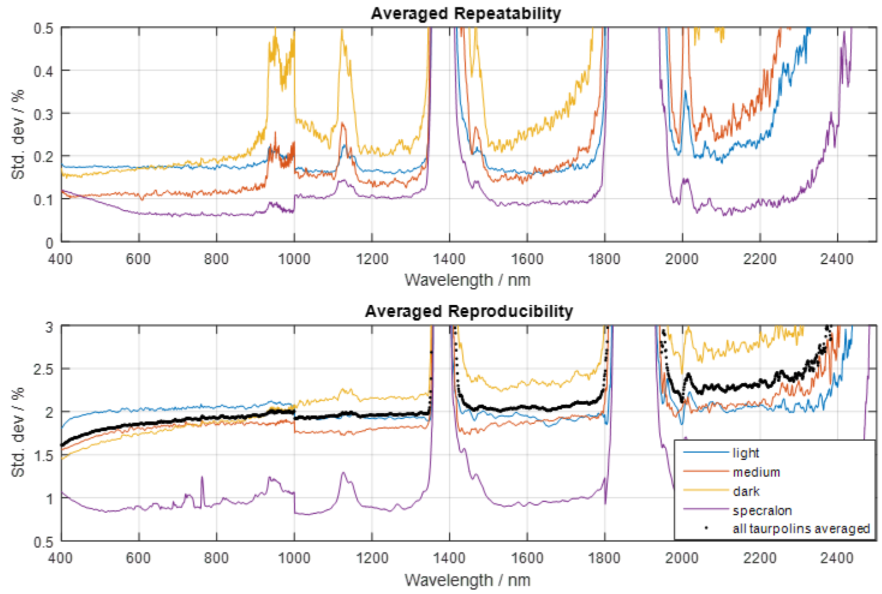


Figure 3.21: Tarpaulins tests results. Top panel averaged instrument repeatability, bottom panel measurement reproducibility.

Figure 3.22 shows the results from the Spectralon difference when measured at the default time scale of about two minutes, and the six minute gap from the tarpaulin tests. It is important to note that for the Spectralon measurements the operator aims to always measure the same middle part, whereas for the tarpaulins, measurements specifically pointed at different parts of the tarpaulin. The tarpaulin uniformity is unknown but in a desert environment you could assume that some dust and sand could affect it, as it does the Spectralon panel.

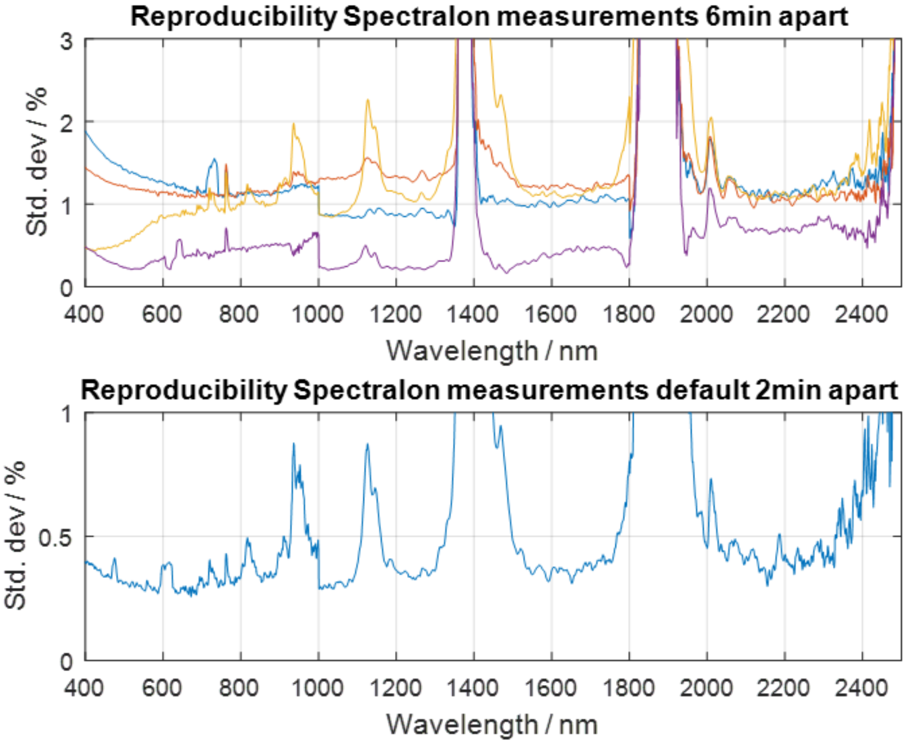


Figure 3.22: Spectralon panel measurement reproducibility at two time-scales. Top panel measurements done 6 min apart and the bottom panel measurements done 2 min apart.

The tarpaulin tests show that over a short time instrument repeatability is not a significant contributor as, even with unfiltered data, using some occasional measurements with higher standard deviation will not greatly influence the overall results. The uncertainty component for unfiltered data are estimated at the level of 0.2%.

The operator/measurement protocol related effect to Spectralon measurements is never higher than 1% as otherwise the data is rejected by the quality control measure. This value already contains the short term repeatability effect of 0.2%, so to calculate only the reproducibility of Spectralon measurements the short term repeatability effect should be subtracted in quadrature from 1 (see Table 3.2). The operator effect related to ground measurements is estimated from the standard deviation of the tarpaulin measurements. It does not seem to be the case that this effect is significantly wavelength dependent, thus 1.9% for the entire measurement wavelength range. The measurements of the tarpaulins show slightly higher standard deviation in SWIR, but this is attributed to the tarpaulin temperature change, and does not expect the ground to have such a strong response. This value again contains the instrument repeatability effect and tarpaulin non-uniformity.

The tarpaulin uniformity is related to inherent fabric uniformity, any fold marks due to storage and transport, the flatness of the ground below, and any sand or dust contamination. This effect is currently unknown, but disregarding it will cause an overestimation of the operator related uncertainty. Currently, a value of 0.4% is used to estimate this effect. This is based on a knowledge of other types of reflectance standards, such as ceramic tiles batch uniformity (estimated as 0.25%). Accounting for some additional effects a value of 0.4% is used for the time being, until further confirmation becomes available.

3.6 Discussion

From the perspective of NPL the Gobabeb campaign was successful. In particular, this experience led to the clarification of ground reflectance terminology that is required for SI-traceability. The proposed method of combining two different reflectance panel calibrations: BRF and total diffuse, provides an SI-traceable solution for measurements with complex illumination conditions that are traditionally not addressed by reflectance calibration methods and procedures. However, the results are reported using only Spectralon BRF values (not HDRF), and thus contain errors. These errors are estimated and included in the uncertainty budget.

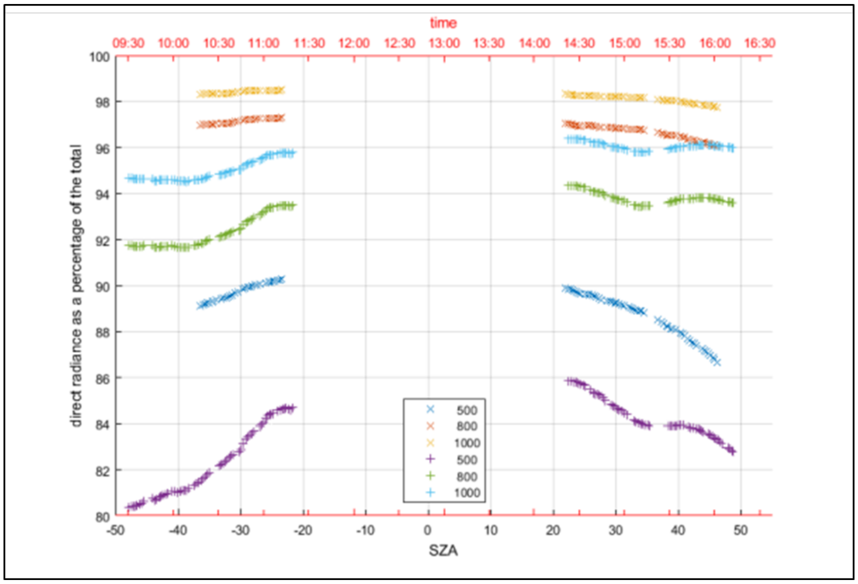


Figure 3.23: The proportion of direct solar illumination in the total for the 25th and 26th November

The CIMEL sun photometer installed on the site during the field campaign did not work properly, thus it was not possible to extract the local atmospheric data. The closest AERONET station is located at Gobabeb Research and Training Centre 10 km away. The AOD data from that station was used in RTC calculate the estimated the fraction of the direct solar irradiance in the total downwelling irradiance. Figure 3.23 presents the ratio of direct signal as a percentage of the total doweling irradiance for three wavelengths 500 nm, 800 nm and 1000 nm on two different days of the field campaign. This ratio increases with the increase of the wavelengths, thus the influence of the diffuse sky component is higher for the shorter wavelengths where the atmospheric scattering occurs, and almost negligible for the NIR wavelengths.

The difference between the ground reflectance values obtained using the BRF rather than HDRF are shown in Figure 3.24 for the same three selected wavelengths 500 nm, 800 nm and 1000 nm and the same days as the ratios of direct to total downwelling irradiance in Figure 3.23. The errors introduced by diffuse part of the donwwelling irradiance are higher for the short wavelength, where the is a bigger contribution of the diffuse component and can reach 2% . For the 1000 nm this error decreases to 0.2%.

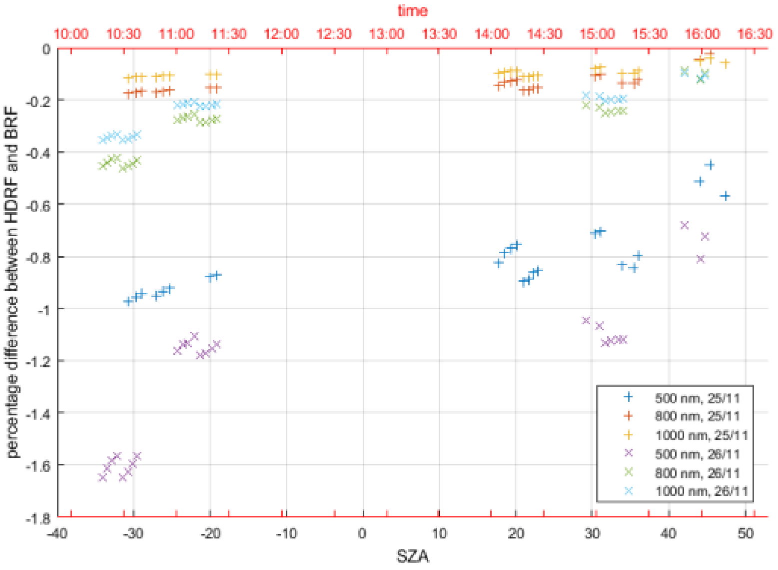


Figure 3.24: The difference in ground reflectance when calculated as HDRF$(\theta_{sun}, 2\pi_{sky}; \theta_r = nadir)$ compared to as BRF$(\theta_{sun}; \theta_r = nadir)$.

The main challenge during the measurement campaign was the Spectralon panel stability, which in particular affected the NPL team due to the case being broken in transit. Consequently reduced protection from the elements during the field work, and the fact that the Spectralon was brand new, meant that a rapid step change in response occurred due to exposure. The second challenge during the measurements was the surface degradation. Very careful walking paths were set after the first few days of the campaign to reduce the impact on the surface. This will be of particular importance during the permanent instrumentation installation to ensure minimal impact on the surface, which will be measured from then on.

NPLs area characterisation tests focused on repeatability of measurements of the same small area over several days, and at different times of the day, so that the ground was measured under different illumination conditions. Similar levels of standard deviation for these measurements from CNES at equivalent spatial scales were found. Uncertainty analysis shows that at least 1.85% of the measurement standard deviation could be attributed to operator effects during the process, which therefore implied that the site homogeneity values are better than what the basic measurements would indicate. NPL measurement uncertainty is estimated at the level of 3%, (k=1) and will vary slightly depending on wavelength. The *in situ* measurement uncertainty can be reduced to a level below 2%, if the diffuse reflectance component is included in the calculation, and the Spectralon reflectance change either reduced or better monitored on a daily, or even hourly, basis. For NPL measurements the ASD should have an extended fibre length, as currently it is too short and affects the measurement reproducibility.

3.7 Summary

This chapter presents the results from the site characterisation of the new RadCalNet site in Gobabeb, Namibia. Particular attention was given to the SI traceability of the site characterisation measurements. The reflectance terminology was reviewed and proposed a new way of deploying reflectance standard calibration values for *in situ* purposes as a better approximation of the actual measurement conditions. Thus using both calibration values for the reflectance standard: the BRF and the TDR in combination, the ratio of which depends on the actual fraction of the direct to diffuse illumination components. This fraction changes with wavelength and SZA. If the proper use of the reflectance standards is not possible it is still recommended to use BRF values for the current SZA during the *in situ* measurements and include the contribution from diffuse light in the uncertainty budget, as the authors did with the final results presented here.

The measurement protocol was strictly followed throughout the campaign by both groups, which led to very consistent final results.

The measurement uncertainty was evaluated and validated by the comparison of both groups' results.

The results of the site characterisation confirmed that Gobabeb is a very good location for the RadCalNet station due to its high homogeneity.

Chapter 4

Uncertainty Budget for Ocean Site

A new framework that enables evaluation of the *in situ* OC measurement uncertainty budget is depicted in this chapter. The BOUSSOLE operational data were screened according to pre-defined criteria to be suitable for SVC purposes. The uncertainty is evaluated then for each component of the *in situ* measurement and data processing step that leads to deriving the final site product which is $\mathrm{R_{rs}}$.

The MCM method is used to handle the data complexity in an efficient manner and derives a default uncertainty value per measurement for given conditions. A single summary value for the BOUSSOLE radiometry is calculated from the statistics of the individual uncertainties per each spectral channel. This summary value meets the current Ocean Colour mission requirement in the blue and green region as is below 5% (k=1) and increases in the red channels to exceed 5%.

4.1 Introduction

The BOUSSOLE buoy (Antoine et al. 2008b,a) is the only European SVC site that was used jointly with the data from the MOBY buoy for MERIS and now is used for Sentinel 3 SVC calibration. The BOUSSOLE buoy is permanently deployed in the Ligurian Sea where the water depth is 2440 m meters. The buoy structure was designed to achieve optimal horizontal positioning of the instruments and minimalise the structure shading effects on them. The buoy has a large sphere at a depth of 17 m that provides the buoyancy and is attached by a cable to the weight placed on the seafloor. A tubular structure of the buoy is fixed to the sphere. In the water, part of the structure has two arms at depths of around 4 m and 9 m that hold in water instruments. The structure above water is 4.5 m high, and on the top, the irradiance sensors and solar panels are installed. The more detailed description is presented in (Antoine et al. 2008a).

Currently, two types of radiometers are in use on the buoy multispectral instruments on one end of the arm and hyperspectral instruments on the other end of the arm. The multispectral instruments were in use from the beginning of the buoy's operation on both arms' ends. Hyperspectral instruments were installed on the other side several years ago. This study focuses in detail on the the multispectral instruments. The uncertainty budget was evaluated for this type of instrument. However, it will be easy to reapply the framework for the hyperspectral instrument in the future. Some of the uncertainty components will change due to the nature of the instrument, but the generic processing will be similar.

4.2 Methodology

The general in water radiometric measurement methodology was mentioned in Chapter 1. This is reiterated here for clarity with more details that are BOUSSOLE measurement specific.

The main product of interest for the SVC needs is the remote sensing reflectance $\mathrm{R_{rs}}$, that is derived from in water measurements of upwelling radiance and above water downwelling irradiance. The radiometric data from the two in water instruments are used to calculate the diffuse attenuation coefficient $\mathrm{K_{L_u}}$. The wavelength dependence is not included in the equations for sake of brevity. However, these effects are important, as water absorption significantly decreases for the red wavelengths, and are fully considered in the subsequent sections of this chapter.

$$\mathrm{K_{L_u}} = -\frac{\ln(\mathrm{L_{u_9}}/\mathrm{L_{u_4}})}{(z_9 - z_4)} \tag{4.1}$$

Where: $\mathrm{L_{u_9}}$ and $\mathrm{L_{u_4}}$ are the upwelling radiance from 4 and 9 meters respectively and $(z_9 - z_4)$ is the difference in depth between both instruments. Knowing the attenuation coefficient it is possible to extrapolate the shallowest upwelling radiance measurements to the level of just beneath the surface $\mathrm{L_u}(0^-)$.

$$\mathrm{L_u}(0^-) = \mathrm{L_{u_4}} \exp(z_4 \mathrm{K_{L_u}}) \tag{4.2}$$

The water-leaving radiance $\mathrm{L_w}$ is calculated according to the following equation,

$$\mathrm{L_w} = \mathrm{L_u}(0^-)\frac{1-\rho}{n^2} \tag{4.3}$$

Where: ρ is the Fresnel reflection coefficient for the water-air interface and n is the refractive index of seawater. This ratio is called water-air interface constant and its value is 0.543. Finally, the remote sensing reflectance $\mathrm{R_{rs}}$ is given by:

$$\mathrm{R_{rs}} = \frac{\mathrm{L_w}}{\mathrm{E_s}} \tag{4.4}$$

Where: E_s is the downwelling irradiance measured above water by the radiometer placed on the top of the buoy structure.

To evaluate the uncertainty related to the measurements of R_{rs} the GUM (JCGM100:2008) methodology is followed. However, due to the complexity of the measurements performed at the BOUSSOLE buoy the MCM, that is known as Supplement 1 to GUM (JCGM101:2008), has been chosen as the calculation method. The MCM uses the PDF of each input component, not its uncertainty value, as the traditional GUM. The measurements model is run a large number of times randomly drawing individual inputs from their PDFs, thus, in fact, conducting the calculation of different input values every time. The output of the model has its PDF, and the characteristics of that function are used to evaluate the best estimate and its associated uncertainty value.

It is important to note that the uncertainty evaluation is based on the assumption that models used in the data processing chain are in principle correct. Any unknown biases that might be present due to the limitation in the current knowledge will not be detected. Ideally, the inter-comparison with other models and methods that would agree within the uncertainty limits would be the final validation of the budget.

4.3 Data Processing Chain

Three multispectral instruments, two radiance and one irradiance of the Satlantic 200 series, are used to obtain the R_{rs} values. They all have seven spectral channels with a 10 nm bandwidth that covers those relevant for ocean colour radiometry wavelengths from 412 nm to 683 nm. The measurements are taken simultaneously through the day and night for one minute at 6 Hz every fifteen minutes. Also, other instruments on the buoy record ancillary data, such as salinity, temperature, depth and tilt at two axes, and the buoy heading. Meteorological inputs including wind direction and speed, atmospheric pressure are from a meteorological buoy that is two nautical miles away from the BOUSSOLE site.

Night measurements are used as dark readings for the multispectral radiometers as they do not have shutters, in addition, they allow the monitoring of instrument noise levels and detect any drift present during a deployment.

Every measurement from the one-minute acquisition is dark corrected using the night measurements, and then transformed into proper physical units (radiance or irradiance) by application of the calibration coefficient; the in-water measurements, in addition, have immersion factor correction applied. The one-minute series provides 360 individual acquisitions and they need to be converted into one number representing a relevant radiometric value for this one-minute measurement sequence. Median was chosen as this metric allows to eliminate occasional outliers that can be present in the one-minute series and would affect the mean value. The standard deviation of the one-minute series is kept for further processing and quality control purposes.

The ancillary data from the buoy that is acquired simultaneously with radiometric data is used to calculate the depth and tilt of the instruments during the measurement and their position in reference to the Sun. Although the radiometers are permanently attached to the buoy's arms, the whole buoy structure moves and tilts thus, the actual depth of the sensors change with time and the z_4 and z_9 values from the equations are not constant. They are continuously calculated for each data acquisition to include the real depth in any further calculation.

To calculate any of the quantities presented in equations 4.1 to 4.4 additional corrections are applied, and these are described in the following subsection.

4.3.1 Corrections Applied to the Data

The buoy structure and the instrument itself causes self-shading effects (Gordon & Ding 1992). This effect, if left uncorrected, would bias the in water measurements. BOUSSOLE's shading was modelled by E. Leymarie in his software called SimulO 3D Monte Carlo Code (Leymarie 2005). This model is used firstly, in the quality control procedure to reject any data where the shading contribution is more than 5% and to correct the shadowing effects in the remaining data. This correction is based on MC modelling and takes into account the solar geometry, and the buoy's arm position in reference to the Sun and the chlorophyll concentration. The shading correction is calculated for each depth and wavelength separately.

The measurements from nominally 4 and 9 meters are used to calculate $\mathrm{K}_{\mathrm{L_u}}$, and the measurements from 4 metres are extrapolated just below the surface to $\mathrm{L_u}(0^-)$ using equation 4.2. However, this extrapolation applies only to the wavelengths below 600 nm for which sunlight is transmitted though upper water column. Water absorption is much higher for longer wavelengths, and in addition, shorter wavelengths inelastically scatter the light to the red wavelengths (Raman scaterring and CDOM fluorescence), thus above 600 nm radiances do not attenuate with depth following this simple exponential function. To correct for that fact, especially for the wavelength above 600 nm and seawaters with higher chlorophyll concentrations, the Hydrolight software (Mobley 1994) is used to model that interpolation and to apply the correction. The software inputs are SZA, that is known, and daily chlorophyll values, which are a combination of chlorophyll concentrations database from the BOUSSOLE site using the High Precision Liquid Chromatography (HPLC) method and the satellite-derived products for the days without *in situ* data.

In a still system, the buoy would be perpendicular to the seamless sea surface. Thus all radiometers would record the readings at nadir for upwelling or zenith for downwelling views. *In situ* the buoy moves with waves and currents. The ancillary data from the buoy such as tilt and heading are used to derive the tilt correction. As the buoy moves, the actual correction for the position of the $\mathrm{E_s}$ sensor in reference to the SZA is necessary and this is applied to the direct part of the total downwelling irradiance. Additionally, tilt information is used for the actual depth calculation.

Many more corrections might be required to the other lower quality datasets, for example, bio-fouling corrections; however, this is beyond the scope of this study as the dataset used here was carefully screened for the best quality data.

The equation to calculate the $\mathrm{R_{rs}}$ becomes much more complicated than its form presented as the equation 4.4 from the Methodology section at 4.2.

$$\mathrm{R_{rs}} = \frac{\mathrm{L_{u_4}} f_{s_4} \exp[z_4 \frac{\ln(\mathrm{L_{u_9}} f_{s_9} / \mathrm{L_{u_4}} f_{s_4})}{(z_9 - z_4)}] f_h C_{\rho n}}{\mathrm{E_s} f_{tilt} f_{dir} + (1 - f_{dir}) \mathrm{E_s}} \tag{4.5}$$

Where: $\mathrm{L_{u_4}}$, $\mathrm{L_{u_9}}$ and $\mathrm{E_s}$ are median values in physical units of 1-minute measurements, for more details see section 4.4.2.1 Instrument readings, f_{s_4} and f_{s_9} are shading corrections applied at the depth of 4 m and 9 m, f_h is Hydrolight correction applied to the surface extrapolation, $C_{\rho n}$ is the water-air constant, f_{tilt} is a buoy tilt correction and f_{dir} is the fraction of the direct to total solar irradiance.

4.3.2 Quality Control

The data used for SVC purposes requires special quality control and a set of selection criteria was defined and these are presented in Table 4.1.

The first condition ensures lack of any rapidly changing clouds, like cirrus occurring during the measurements, plus screens for birds or boats being near the buoy affecting the measurements by shading the sensors.

The second one is a generally clear sky test where the measured downwelling irradiance is compared to the theoretical one that would be present at given atmospheric conditions. The difference higher than 10% indicates an overcast sky. The tilt condition apart from the of nadir/zenith buoy position eliminate any measurements when the wind speed is greater than 10 m/s at the same time. This wind speed value is a limit for the SCV measurement as for higher winds the sea state exhibits white caps that are not wanted.

The depth less than 11 m ensures no significant current that could pull down the above water buoy's structure to affect the downwelling irradiance measurements. This condition is never present if criteria 1 and 2 are met. The bio-fouling test is used for the moment, but in the future this criteria might be relaxed to allow for the data affected with an additional uncertainty component. The inter calibrations issue can be seen for some deployments (Antoine et al. 2008a) where the whole dataset tends to not agree with that from other years. This might happen due to some instrument calibrations issues.

The data set used in this study covers one summer deployment from June until August 2008 and after the data screening included 1090 individual observations.

Table 4.1: BOUSSOLE high quality data set selection criteria.

Selection criteria	
One minute readings stability	$< 2\%$
Clear sky test	$0.9 < \& > 1.1$
Tilt	$< 10°$
SZA	$< 75°$
Depth	< 11 m
Shading	$< 5\%$
Bio fouling	N
Screening for inter calibration issue	Passed

4.4 Uncertainty Budget

Equation 4.5 has been adapted further to include all the components that will have to be considered in the uncertainty budget and becomes a measurement equation with all currently defined and addressed uncertainty contributions.

$$\mathrm{R_{rs}} = \frac{\mathrm{L_{u_4}} k_{cal} f_{s_4} \exp[z_4 \frac{\ln(\mathrm{L_{u_9}} k_{cal} f_{s_9} / \mathrm{L_{u_4}} k_{cal} f_{s_4})}{(z_9 - z_4)}] f_h C_{\rho n}}{\mathrm{E_s} k_{cal} k_{cos} f_{tilt} f_{dir} + (1 - f_{dir}) \mathrm{E_s} k_{cal} k_{cos_h}} \tag{4.6}$$

Where: k_{cal} represents an uncertainty in absolute radiometric calibration, k_{cos} and k_{cos_h} are uncertainties due to the cosine response of the irradiance sensor diffuser affecting the direct Sun irradiance and the total diffuse irradiance integrated over the hemisphere respectively are added to already defined terms in equation 4.5

This equation is used in the MCM and was run 10^5 times, randomly drawing from the PDF of each component. The PDF of the output values is used to evaluate the best estimate and its standard uncertainty ($k = 1$). In the case when the output PDFs are close to Gaussian, the mean and standard deviation are suitable means of expressing wanted values.

The model is run for every individual observation from the study dataset. Thus each measurement has its uncertainty value that depends on the environmental conditions that were present during its acquisition. The model for every data point from the dataset provi des an individual “dynamic” uncertainty value. A summary of all observation points provides a general result that reveals the probability distribution and the most likely uncertainty value for the whole data set. The R programming language was used to develop this framework as this is a language used for operational BOUSSOLE data processing. Thus, the uncertainty evaluation framework can be integrated into the default data processing chain.

The following sections present in detail the PDFs of each component in equation 4.6 and the ways of defining them. It is important to note that PDFs are derived based on available information about the uncertainty associated with each component. Their values are associated with each input and not its effect on the final remote sensing reflectance value. All components are split into three groups related to the instrument, environmental and modelling effects. This classification is not always crystal clear as some components might be a mixture of two groups. Nevertheless, such a classification was made for the sake of clarity.

4.4.1 Instrument Related Effects

These are related to instrument design and performance. The knowledge about them comes from laboratory tests and theirs uncertainty is derived using traditional GUM (JCGM100:2008). They will have Gaussian distribution, as this is a default output from the GUM evaluation approach. Thus, any uncertainty from a calibration certificate has a normal distribution. Rectangular distribution is assigned for the cases when the actual GUM derived uncertainly is not present, and there is no detail knowledge about the behaviour of given component. Thus, it is assumed that all values from a predefined range are equally probable.

4.4.1.1 Absolute Radiometric Calibration

Absolute radiometric calibration is undertaken, on average, every six months at the manufacturer’s site (Satlantic) (*Sea-Bird Scientific* 2018). During that time any necessary instrument repairs are carried out. The absolute radiometric calibration coefficients are provided by the manufacturer, however they are not accompanied by the specified uncertainty value.

One set of the BOUSSOLE instruments was sent to NPL for calibration and characterisation tests in 2012. That time the radiometric calibration uncertainty was calculated and the coefficients were compared with those provided by Satlantic. Thus the uncertainty value for this term is derived from the NPL laboratory calibration uncertainty values plus the component due to Satlantic laboratory calibration that was estimated as a difference between NPL and Satlantic calibration coefficients from 2012. These derived uncertainty values are presented in Table 4.2 per spectral channel for irradiance and radiance. In that table, the uncertainties associated with the standards are shown first and then the uncertainty associated with using these standards for an absolute radiometric calibration at the NPL laboratory. It is important to note that although it is beyond the scope of this thesis, the full uncertainty budget for NPL calibration was evaluated including all effects that are associated with the calibration process, such as the lamp current, ageing and uniformity effects as well as realignment and instrument reading stability.

The radiance calibration was derived from measurements of one spectral channel only due to lack of large enough reflectance standard at the time at NPL and large field of view (FOV) of the Satlantic instrument. Radiance uncertainties are estimated based on one spectral channel measurement and the theoretical knowledge.

Table 4.2: BOUSSOLE radiometers absolute radiometric calibration uncertainties (k=1).

Wavelength [nm]	NPL irradiance standard	NPL irradiance calibration	Estimated Satlantic irradiance calibration	NPL reflectance standard	NPL radiance calibration	Estimated Satlantic radiance calibration
412	0.54%	0.84%	1.6%	0.50%	1.00%	2.3%
443	0.48%	0.59%	1.3%	0.50%	0.80%	2.1%
490	0.40%	0.54%	1.2%	0.50%	0.80%	2.1%
510	0.39%	0.73%	1.3%	0.50%	0.90%	2.1%
560	0.38%	0.55%	1.2%	0.50%	0.80%	2.0%
670	0.36%	0.53%	1.1%	0.50%	0.80%	2.0%
683	0.36%	0.49%	1.1%	0.50%	0.80%	2.0%

Until new calibration data becomes available the values from the "Estimated Satlantic calibration ..." columns are used as the uncertainty values associated with the radiometric calibration.

Another important component of absolute radiometric calibration is its stability during the deployment. The data used in this study is from the first three months of the instrument's operations and none of the quality control checks indicated any issues with the instruments within this time-frame. They were left in operation for a few more months and it was when the signal issues appeared that the recalibration values for the post deployment calibrations were not meaningful stability estimators for the first three months.

The k_{cal} values that are used in the uncertainty evaluation have normal distribution with the mean equal to 1, as actual calibration coefficient values were applied to the data in the previous processing step, thus this element now only carries information related to calibration uncertainty. The standard deviation of this distribution is equal to absolute calibration uncertainty combined with uncertainty related to the stability of this calibration estimated as 1%. The stability has a rectangular distribution and before being combined with the calibration uncertainty, is divided by the square root of three to provide a value corresponding to normal distribution.

4.4.1.2 Other Instrumental Characteristics

The instruments were tested for temperature dependence and detector linearity at NPL back in 2012. The multispectral instruments showed very good detector linearity, therefore this effect is considered negligible. Figure 4.1 shows the results of the linearity test for one radiance channel and a wide range of output voltage signal. The value of the linearity factor equal to 1 indicated linear response of the detector, thus the data presented on this Figure confirms that the detector has a very good linearity.

Similarly, temperature dependence for the dark readings was tested in the temperature range that is expected at BOUSSOLE site and varies form 5°C to 30°C. The dark readings were taken as this was not possible to ensure a stable lighting condition in the chamber. Measurements were taken at four separate temperatures (2°C, 12°C, 22°C,

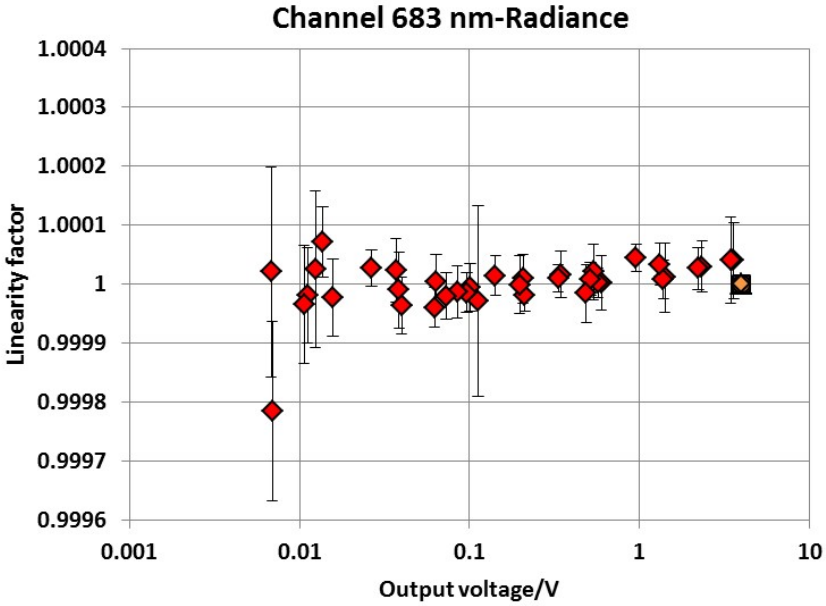


Figure 4.1: Multispectral radiometer Lu_4 683 nm channel detector linearity.

and 32°C) and each of these was repeated twice. The change in the dark signal was less than 0.05% for the entire measured temperature range. The thermal stability of the dark readings ensures that the night measurements used for dark correction were not affected by the temperature change between the day and night measurements.

The thermal stability of light readings is checked *in situ* on E_s data. The check is done using the ratio of measured E_s to the modelled values and the correlation of this ratio with the temperature during the measurement. The assumption is that modelled downwelling irradiance is insensitive to the instrument changes thus any change in their response due to temperature would affect the ratio and be detectable. The correlations calculated between that ratio and the temperature for each channel are small: -0.12 for 412 nm, -0.09 for 670 nm, and for the remaining channels below -0.04. The temperature range for the study dataset is within the range of 19°C-26°C thus any temperature and linearity corrections are considered negligible in the current uncertainty evaluation.

The immersion factor is provided by the instrument manufacturer and is derived from a simple model based on the reflective index of the medium and optical window. The information found in the literature (Zibordi 2006) for this particular type of instruments states a small bias between the improved model, laboratory measurements and the simple model. This bias was estimated as 0.4% with the uncertainty in the measurements at the level of 0.19%. The uncertainty related to immersion factor is not included in this budget as it is considered to have minor effect on radiance value.

The cosine response affects the irradiance sensor and its specification is provided by the instrument manufacturer as it agrees to 3% with a perfect cosine in the angular range ±60° and for angles above 60° increase to 10%. Ideally each instrument would have a cosine correction response characterised in order to correct for that effect and then only a residual uncertainty in that correction would be propagated. At the time of this budget evaluation, such information was not available from the Manufacturer, so cosine response tests were performed to check if the instruments met these requirements. The test results presented in Figure 4.2 show better performances than the aforementioned values.

Total downwelling irradiance is composed with the direct Sun irradiance and diffuse sky part. A non-perfect cosine response of the instrument's diffuser affects both but not in the same manner (Zibordi & Bulgarelli 2007). For the direct irradiance the cosine response is related to SZA and the difference between the perfect diffuser and the response of the instrument is for a given incident angle. For the diffuse part the cosine response for all angles is integrated over the whole hemisphere, calculated according to this equation:

$$k_{cos_h} = \int_0^{\pi/2} k_{cos}(\theta) sin(2\theta) d\theta \tag{4.7}$$

Where: the k_{cos_h} is the integrated cosine response over the full hemisphere while $k_{cos}(\theta)$ is a cosine response for a given incident illumination angle.

The values assigned to the k_{cos} were slightly modified from the manufacture specification, according to the test results presented at Figure 4.2. Thus, for angles below 20° a value of 2% is used, then for angles from 20° to 60° 3%, from 60° to 70° 5% and 10% for angles above that. These updated values provide the value of $k_{cos_h} = 3.5\%$. These numbers represent biases rather than uncertainty. However, as they are not corrected they are inserted into the model as uncertainty contributors.

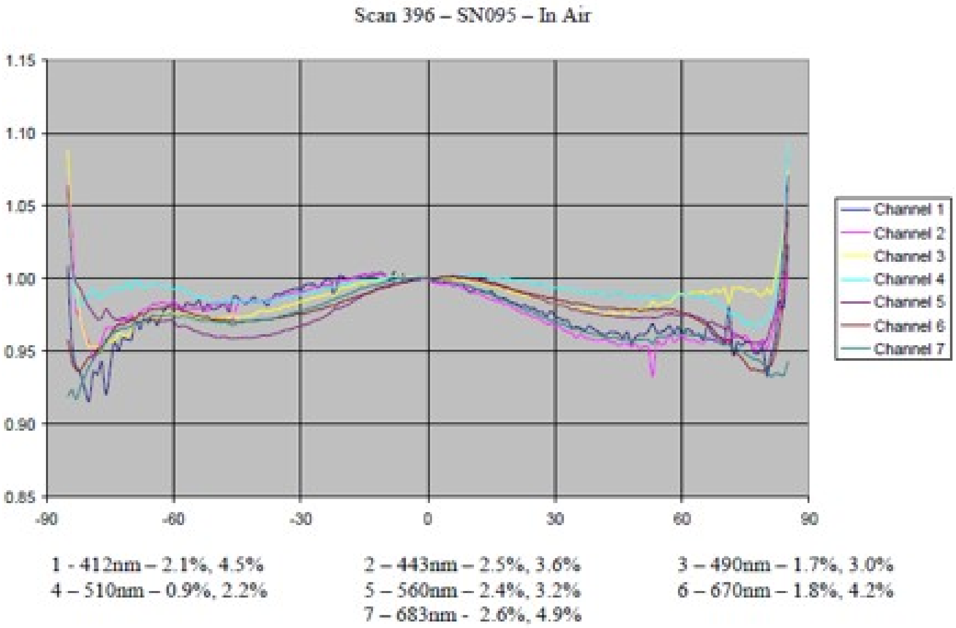


Figure 4.2: Irradiance instrument cosine response characterisation test results. X- axis incidence illumination angle, y axis ratio to the perfect cosine response. Graph obtained with the instruments' manufacturer test results.

The real diffuser response is almost always lower than the perfect cosine apart from the values at the 90°, which are not significant in the diffuse component calculation as they are weighted by the values of $sin(2\theta)$ which is literally zero. However, to not introduce a change in the current data processing the PDF assigned to cosine responses has to have a middle value equal to 1. Although in an ideal case this PDF would have values bigger then 1 and any possible deviation from that value would go only in one sense direction toward lower values. In the MCM the $k_{cos}(\theta)$ has a rectangular distribution and with the upper limit of 1 plus relevant for a given angle cosine response value and the lower limit defined as 1 minus relevant for a given angle cosine response value. A similar approach is chosen for k_{cos_h}, in this the lower limit is always 0.965 and the upper limit is 1.035.

4.4.2 Environmental Effects

During the buoy's operations environmental conditions are constantly changing. The most rapid changes are excluded from the dataset by the selection criteria of 2% standard deviation in one minute of E_s signal. Nevertheless environmental conditions are inseparable from the BOUSSOLE site and the majority of them affect the measurements. These include winds, waves and currents that very much influence the buoy position, but in addition water composition expressed by chlorophyll concentration that varies during the year will affect those factors such as shading and depth extrapolation.

4.4.2.1 Instrument Readings

Dispersion of instrument signals within a one-minute series is a measure of environmental conditions, such as changing sky illumination or wave focusing. The instrument's noise in laboratory conditions expressed as a standard deviation of the mean of 360 individual measurements taken within one-minute acquisition times, is below 0.1% for all spectral channels. The values obtained at the laboratory can be considered as negligible *in situ*, and a standard deviation of the mean is used as a measure of changes in the signal caused by environmental variation occurred during the one-minute measurements. The median is the best estimate for 360 dark corrected readings converted to physical units. This value rather than the mean is selected as occasionally the instrument's signal might have several outlier readings clearly affected by some instrument delay at the beginning of the acquisition (see Figure 4.3), or something shortly thereafter obscuring the instruments FOV. Median is a quick and efficient method to not include these biased readings in the further processing, which mostly effects E_s.

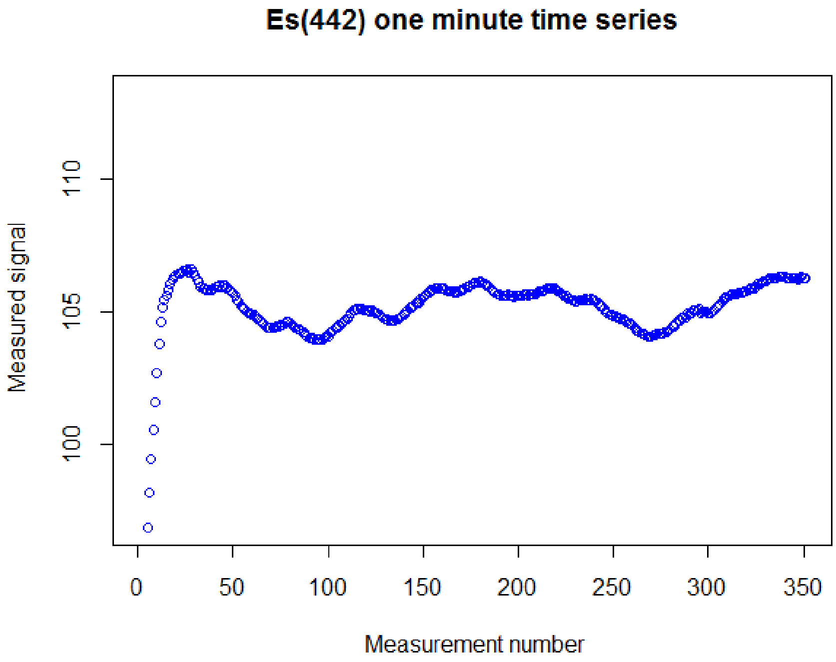
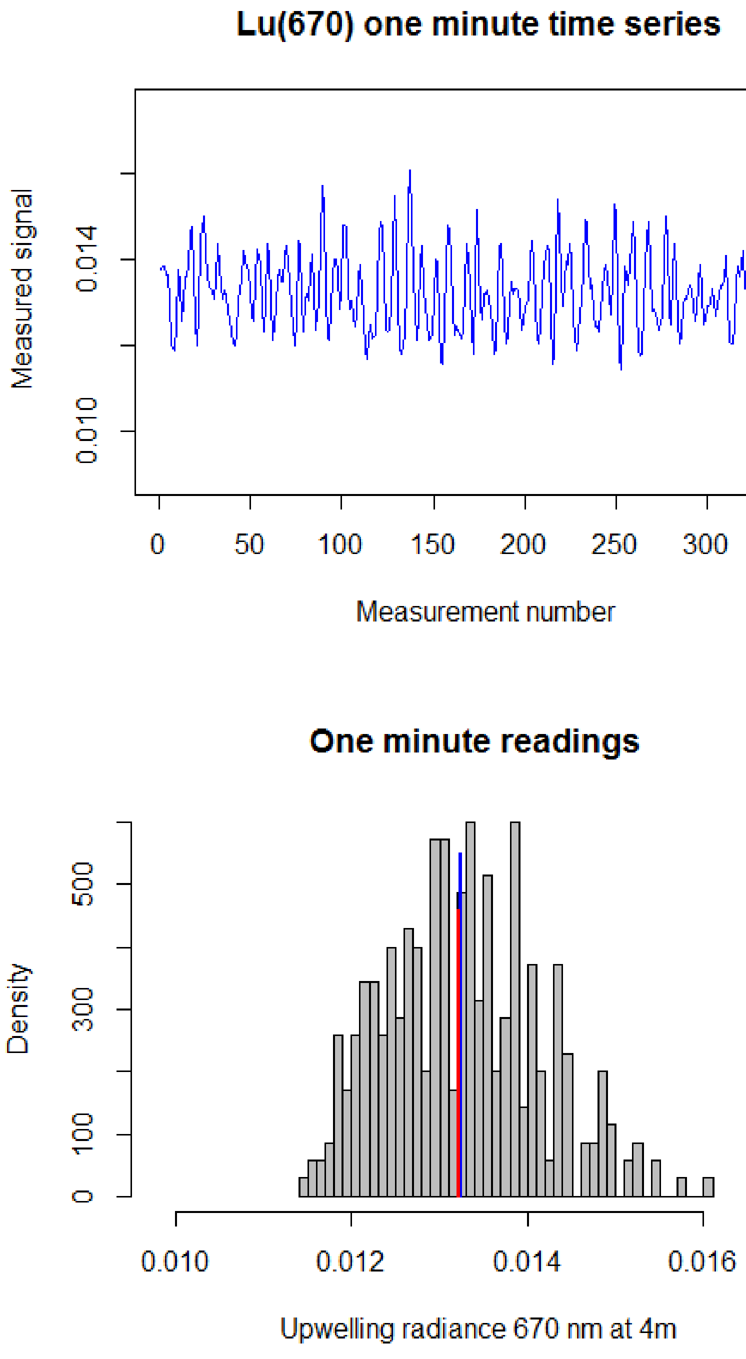


Figure 4.3: BOUSSOLE example of one-minute E_s readings, where the beginning of the acquisition is clearly biased.

The majority of L_{u_4} and L_{u_9} are affected by wave focusing and defocusing cycles. They can be seen in one-minute series; however, the distribution is very close to Gaussian and an averaging of the signal reduces the wave focussing effects. Figure 4.4 shows the characteristics of an example one-minute series of the L_{u_4}, where the wave focusing can be seen as the repeating cycle in the one-minute series, but the same data presented in a histogram form shows the Gaussian like distribution with the mean value represented by a red line and median as a blue. In the MC model the instrument signal is represented by normal distribution with a mean equal to the median value of one-minute readings and standard deviation equal to standard deviation of the mean of the same signal series. For the one-minute signal presented in Figure 4.4 the best estimate of 0.0132 and the standard uncertainty 0.35% are used in further processing. Thus, any spread in the instrument readings that is dependent on the current environmental conditions is included in the further processing as the standard deviation of the mean for each individual measurement point is carried in to the model.


Lu(670) one minute time series

One minute readings

(a)

(b)

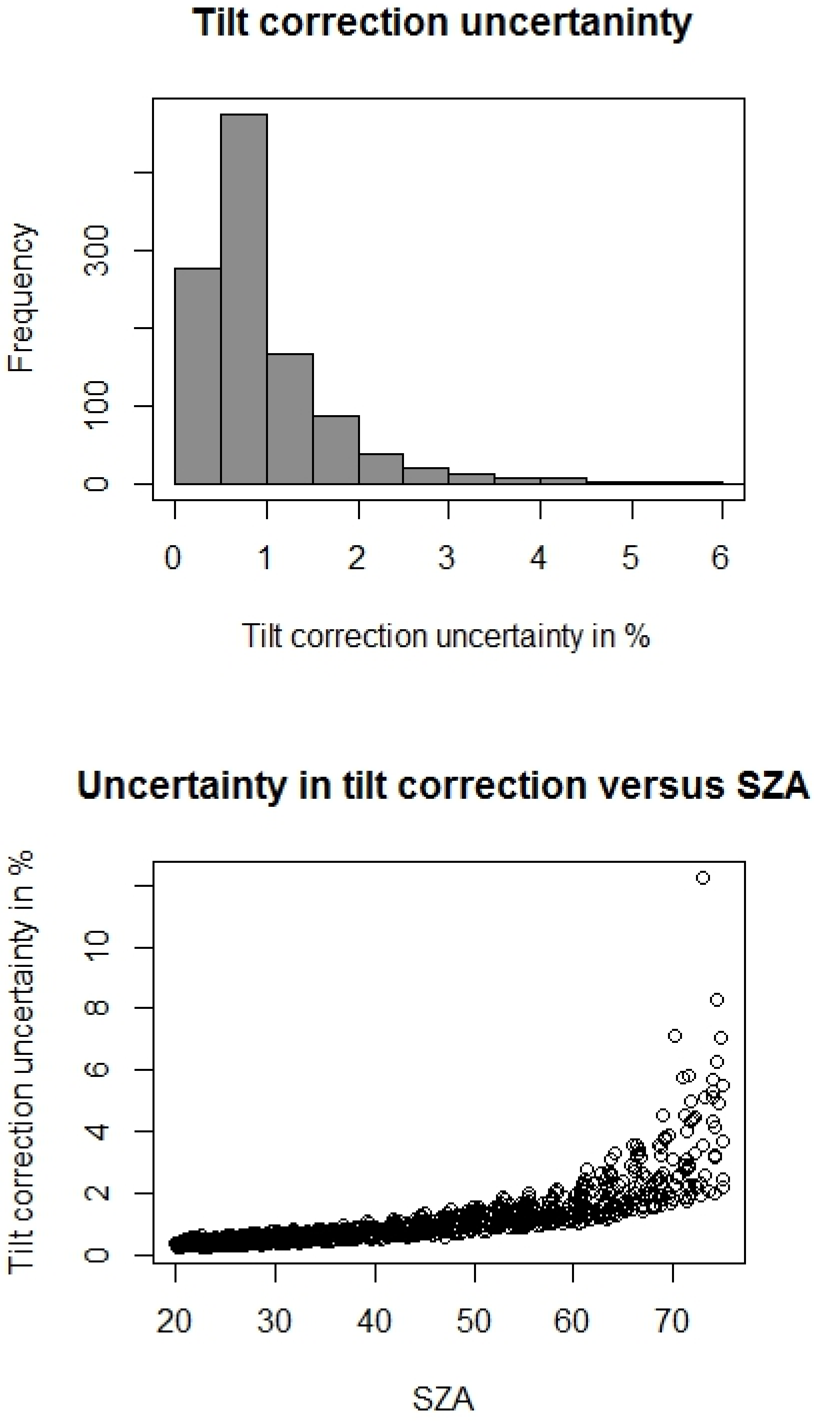
Figure 4.4: (a) Example of one-minute L_u(670 nm) readings, where the wave focusing effects are visible) readings from 4 meters, wind speed 2.5 m/s wave period 4, mean 0.0132 , median 0.0132, st.dev 0.0009 (6.6%). (b) Example of one-minute L_u readings histogram, where the mean (red line) and median (blue line) values are almost identical.

4.4.2.2 Tilt Correction

The data acquired by EZCompass-dive sensors include two axes tilt and the buoy headings are used to calculate the buoy θ and ϕ at every measurement. Including the actual SZA and SAA the cosine of the sun for a given buoy's position is then calculated. The tilt correction is a ratio of the real cos(SZA) and the calculated cosine.

An uncertainty in the tilt correction was estimated by running the same calculation using normal PDFs of all input components defined from the instrument readings. The mean is equal to the median value of one-minute acquisitions and the standard deviation is the standard deviation of the mean combined with an additional uncertainty in the sensor accuracy. This accuracy was stated by the manufacturer as 0.5° for the azimuth (heading sensor) and due to the lack of such information for the tilt, the same value is used.

A summary of the tilt correction uncertainties calculated for the dataset is presented in Figure 4.5 panel (a). For the majority of the observations this uncertainty is lower than 1%, and the tilt correction uncertainty increased with SZA this relationship is shown in Figure 4.5, panel (b). The outputs from this model are fitted to the main MC model, thus each observation will have an uncertainty in the tilt correction evaluation based on the condition which occurred during that acquisition.



(a)

(b)

Figure 4.5: (a) BOUSSOLE buoy tilt correction uncertainty. (b) Tilt correction uncertainty plotted versus SZA.

4.4.2.3 Instruments' Depth

The depth is recorded at the level of the lower arm by the 37-SI CTD sensor that measures conductivity, temperature and pressure. The pressure accuracy is stated by the manufacturer (SeaBird) as 0.1% and the stability 0.05%, the combined uncertainty of those two gives a value of 0.11%. Before a buoy deployment the distance is measured between the CTD sensor and both arms, plus the distance between the instruments mounted on the arm and the main structure, and the distance between the two arms; for all these measurements the uncertainty of 5 mm is applied with the rectangular PDF.

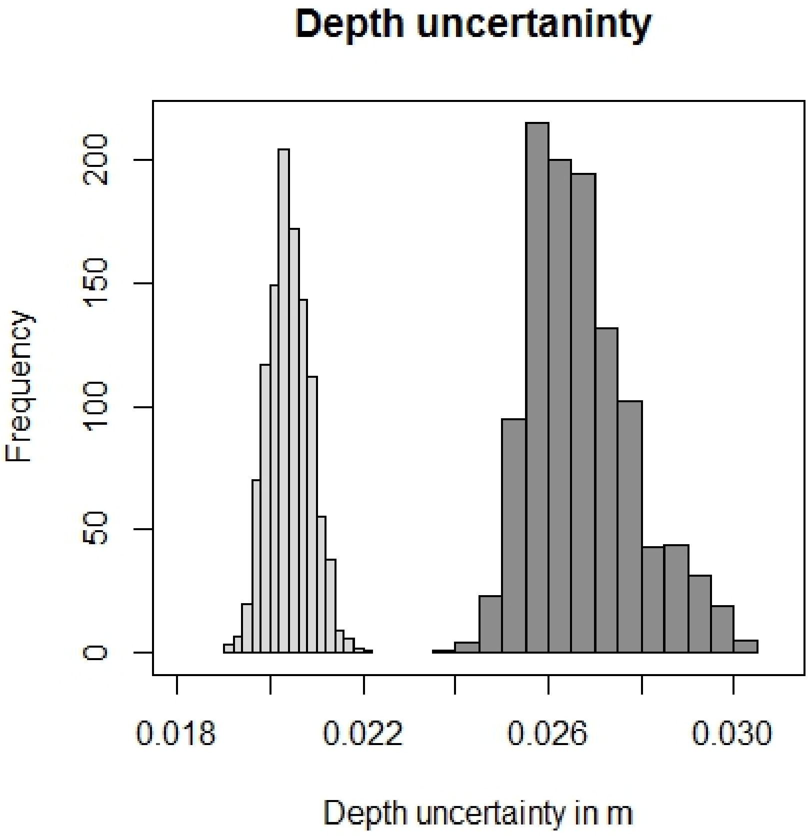


Figure 4.6: Depth uncertainty for the instrument mounted on the lower arm left and upper arm right.

The instrument depth is calculated for each measurement including pre-deployment data, actual measurement depth and the buoy tilt. A separated MC model was run for these depth calculations to estimate instrument depth uncertainty. Figure 4.6 presents histograms composed of each measurement uncertainty value for lower and upper arms. Generally, the majority of points in the data set have an uncertainty of around 2 cm for the instrument attached to the lower arm, increasing to 2.6 cm - 2.7 cm for the upper instrument. Higher uncertainties for the instrument located at the upper arm is due to the longer arm length, thus this instrument is farther away from the buoy's main structure and more sensitive to the buoy tilt. The output of this model is fitted to the general BOUSSOLE uncertainty MC model, thus each measurement point has an associated depth uncertainty that is expected for the given environmental conditions.

4.4.2.4 Shading Effects

Shading effects and their corrections are hard to clearly classify to one of the groups as the effects are related to the environmental conditions but the correction was derived from the model, nevertheless we decided to keep shading corrections in the environmental part. The uncertainty in shading corrections is estimated from a model validation exercise that compared the outputs from the SimlO (Leymarie 2005) to (Piskozub 2004) model. The difference found between them is 2%. Although this value is not a true measure of the uncertainty in the shading correction, as it represents the difference between the two models, this can be used as an indication until a better solution is found.

The shading correction is derived for each spectral band at two depths for every measurement in the data series. In the MC model, that shading correction value is propagated in the form of a rectangular distribution, where the lower and upper limits are defined as the actual shading correction value ±2%. The 683 nm band is the only exception where, due to very low number of photons the SimlO cannot provide a solution. In the processing a shading correction from the preceding band (670 nm) is used with an increased to ±3% uncertainty value.

4.4.3 Modelling Related Uncertainties

Several models are used to derive the $\mathrm{R_{rs}}$ value. They include an atmospheric model (Gregg & Carder 1990) that is used to estimate the direct and diffuse component of the downwelling solar irradiance, the Hydrolight software that is used to improve the extrapolation of the $\mathrm{L_{u_4}}$ value to the surface, and finally the theoretical model of light propagation through the water-air interface measurements need to be propagated through water air interface and for that a constant is used that is calculated theoretically.

The aim was to find evidence to assign an uncertainty value to the outputs of well-known models that are used in the processing. To achieve this aim, information about the performance of a given model were gathered using the literature or by running simulation and model sensitivity studies. However, studies that would lead to a verification of those models were not performed, thus any undefined bias in the model is currently not identified in the uncertainty budget.

4.4.3.1 Atmospheric Modelling

This is necessary to estimate an amount of the direct and diffuse illumination in the $\mathrm{E_s}$ because tilt correction is applied to the direct Sun part of the total downwelling illumination. The model of Gregg & Carder (1990) is used to find direct to total irradiance fraction. The model uses inputs from the METEO buoy which includes wind speed, relative humidity, atmospheric pressure, total ozone (O_3), and precipitable water. The model was found to be most sensitive to visibility, but this is very good and stable at the BOUSSOLE site. It is reported by Gregg & Carder (1990) that the model agrees within the 6.2% root mean square (RMS) value with spectral irradiance measurements for the wavelength range being 400 nm - 700 nm. Similarly to the shading correction case, here the RMS value is not a real uncertainty on the model, but this is a good indication of the model capabilities. Actual uncertainty in the model might be lower as the RMS value includes some of the uncertainties that are associated with spectral irradiance measurements. Nevertheless, the current MC model assigns a rectangular PDF to the direct to total irradiance fraction f_{dir} with the limits of ± 6.2%.

4.4.3.2 Extrapolation to the Surface

The sensitivity of Hydrolight correction to the chlorophyll concentrations and SZA was tested for the values observed during the study time-frame. The chlorophyll concentration range was 0.1-0.6 mg m^3 and SZA 20-60. A 20% error, as a rough approximation, was assigned to the chlorophyll concentration database that contains a combination of HPLC values for the days when water samples were collected (during monthly cruises) and values derived from satellite data for the remaining days. The aim here was to see the effect of that error on the Hydrolight model and it was run for a given SZA but changing the chlorophyll concentration ±20% (k=1). The results presented in Figure 4.7 shows a marginal impact in the blue and green spectral range and a significant change for red channels.

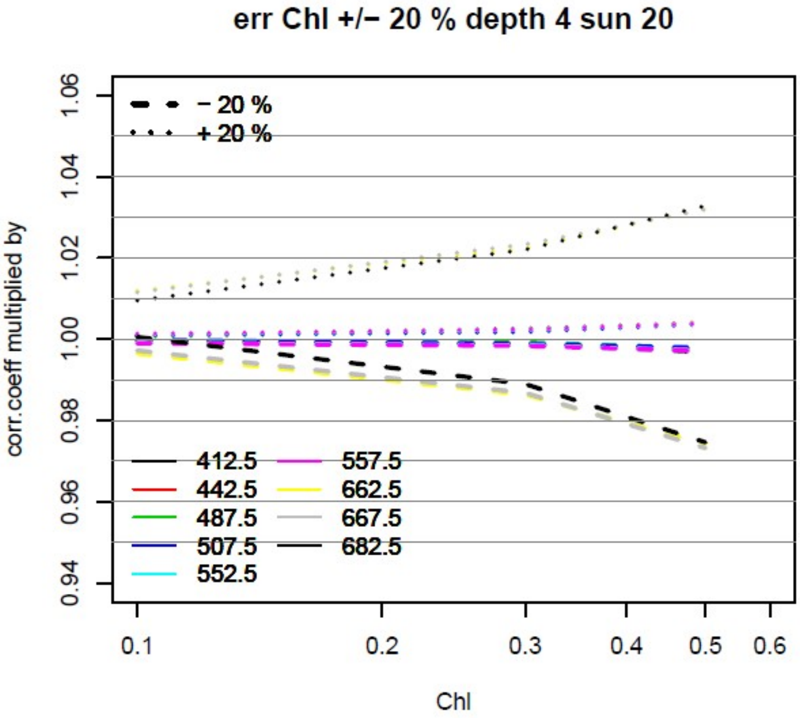


Figure 4.7: Changes in the Hydrolight correction model due to chlorophyll error.

The uncertainty of the Hydrolight correction is based on these sensitivity studies, one value of 0.5% is assigned for all channels below 600 nm. For the red bands the value strongly depends on the chlorophyll concentration, the effects of SZA are minimal. Thus, this uncertainty for wavelengths above 600 nm and chlorophyll concentration below 0.25 mg m^{-3} is estimated as 2% and for higher concentration but not exceeding 0.6 mg m^{-3} 3%.

4.4.3.3 Sea - Air Interface

Although the so-called sea-air interface constant in the equation 4.5 is equal to 0.543, an uncertainty to that value is assigned. The main source of information about this number and its calculation are from (Austin 1974) and (Austin & Halikas 1976). More recently (Wei et al. 2015) tried to confirm the theoretical value with *in situ* measurements and found in principal good agreements between the two, but the constant value used there was 0.54 and the level of its accuracy of 10% is far lower than the SVC needs. Therefore the information from old publications are used for the further evaluations of the sea-air constant uncertainty.

$$C_{\rho n} = \frac{(1-\rho)}{n^2} \tag{4.8}$$

All components of this equation were previously defined for equation 4.3 in the Methodology section at 4.2. The relative uncertainty in the constant using the traditional GUM approach is given by:

$$\left(\frac{u(C_{\rho n})}{C_{\rho n}}\right)^2 = \left(\frac{u(\rho)}{(1-\rho)}\right)^2 + \left(\frac{2u(n)}{n}\right)^2 \tag{4.9}$$

Where: $\frac{u(\rho)}{(1-\rho)}$ is a relative uncertainty in the nominator of equation 4.8 related purely to the uncertainty in the Fresnel reflection coefficient and $\frac{2u(n)}{(n)}$ is a relative uncertainty in the denominator of the same equation related to refractive index of seawater and due to its power 2 in the original equation, the sensitivity coefficient assigned to uncertainty in n is 2. Thus, the refractive index uncertainty contributes more to the constant's uncertainty.

In the index of the refraction of seawater (Austin & Halikas 1976) stated that n depends on water salinity, temperature, pressure and varies spectrally. The tables from that report were used to verify the range of the n changes. Table 4.3 presents values for the first three variables recorded during this BOUSSOLE deployment. There is very little change in the salinity and atmospheric pressure for the data records in the current data set, thus these two factors are not considered further. The temperature range is around 6°C, according to the data from the Table 4-2 in (Austin & Halikas 1976) the difference in n at 20°C and 25°C is 0.04%, thus very small and considered as negligible.

The change in the refractive index between the 412 nm and 683 nm was calculated using the data from Table 4-2 in (Austin & Halikas 1976) for the salinity 34.99 and the atmospheric pressure 0 kg/cm^2. A second order polynomial was fitted to the data from the report to estimate the refractive index values. The refractive index estimated for 412 nm is 1.34894, and for 510 nm is 1.34199, that provides an exact 0.543 constant value and 683 nm is 1.3362. These spectral changes in n are the major contributors to the overall uncertainty and for the two border wavelengths this effect on $C_{\rho n}$ calculated using only the second component of the right-hand side of equation 4.9 is 0.89%, decreasing to 0.53% and for the wavelengths in between.

Table 4.3: Summary values of salinity, temperature and atmospheric pressure during the BOUSSOLE deployment under studies.

Observed values	Mean	Median	Standard Deviation σ	Actual percentage coverage within 1 σ limits	Min- Max
Salinity in ‰	38.4	38.4	0.11	90%	37.8-39.6
Temperature in °C	23.5	23.8	1.7	67%	19.7-26.3
Atmospheric Pressure in mbar	1014.1	1014	3.2	77%	1007-1024.6

For the investigation into ρ value validity (Austin 1974) data was used where the ρ for winds speeds from 0 to 10 m/s^2 and for a viewing angle up to 10°, (this is a limit in the SVC dataset) varies from 0.0211 to 0.0218. Uncertainty in $C_{\rho n}$ due to ρ, calculated using only the first component of the right-hand site of equation 4.9 is 0.042%, that when combined with the second part of the same equation does not change the final result.

Thus, the uncertainty of 0.89% is assigned to 412 nm and red wavelengths, and then for the remaining bands a value of 0.53% is used. The PDF for this component has a normal distribution with a mean value 0.543 and one standard deviation assigned depending on the wavelengths as stated above. This is actually a bias not an uncertainty and according to the GUM this bias should be corrected and then any residual uncertainty related to that correction can be propagated. This is not done in this case and the present uncertainty budget is calculated for the existing processing and currently there is no correction for that. For relatively small biases the GUM allows to propagate them as an uncertainty and that is what has been done in this case.

4.5 Results

The results are presented at different processing steps to show the uncertainty value evolution from the measurement at a single depth through calculation of attenuation coefficient from two radiance readings at different depths following all the computations to the final product $\mathrm{R_{rs}}$.

Firstly, the uncertainty in $\mathrm{L_{u_4}}$ measurement *in situ* is presented. It comprises the signal statistics, instrument related uncertainties contained in the calibration component and the instrument shading $\mathrm{L_{u_4}}, k_{cal}, f_{s_4}$. The uncertainties calculated for each measurement in the data set are presented in Figure 4.8, where the data series with the numbers from 1 up to 7 represents spectral channels from 412 nm to 683 nm respectively.

The uncertainty for upwelling radiance measurements at 4 m is mostly driven by the instrument related uncertainty (i.e. absolute calibration, etc.). What is clearly seen for channel 4 is that it has a higher absolute calibration uncertainty value due to absolute radiometric source non-uniformity most badly affecting this channel during calibration. Channel 7 (683 nm) has considerably higher uncertainty as the uncertainty in the shading correction is higher for this channel. The environmental uncertainties evaluated for the signal statistics start to be visible for both red channels 6 (670 nm) and 7 (683 nm) by the sinusoidal like structure to the data series. Both red channels

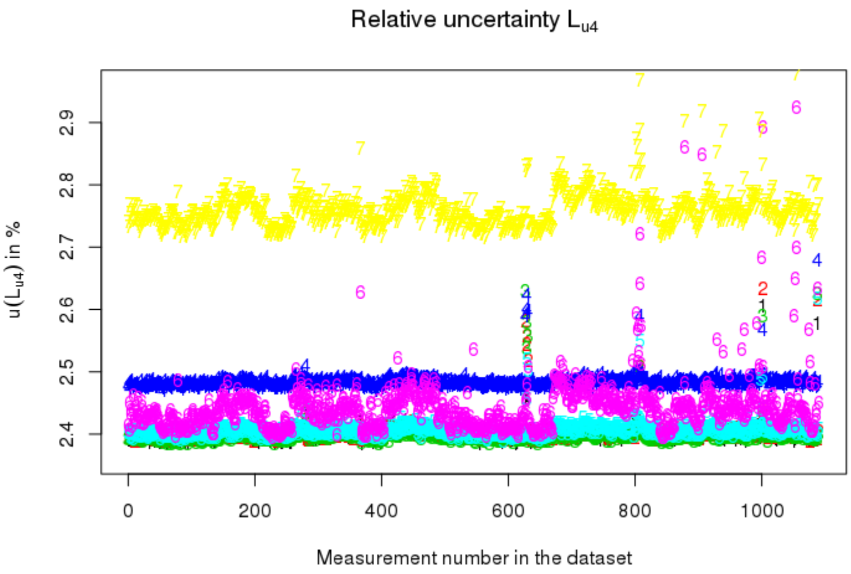


Figure 4.8: Relative uncertainty in L_{u_4} (k=1).

have signal to noise ratios significantly lower than the shorter wavelength channels, for which the environmental effects are not noticeable.

Figure 4.9 presents upwelling radiance uncertainty expressed in absolute terms (radiance units). The signal magnitude changes through the day with the changes of SZA. The signal in blue part of the spectrum is 100 times higher than for the red channels, thus absolute uncertainty values are plotted on logarithmic scale. To evaluate the variability attributed to the environmental effects, one observation from the data set was normalised (divided by its maximum value). The standard deviation of the normalised signal for red channel 7 (683 nm) is higher than the standard deviation of the blue channel 3 (490 nm). This normalised signal variability for the red channel is caused by environmental effects as during laboratory measurements, the opposite trend is seen (red channel signals have smaller standard deviations).

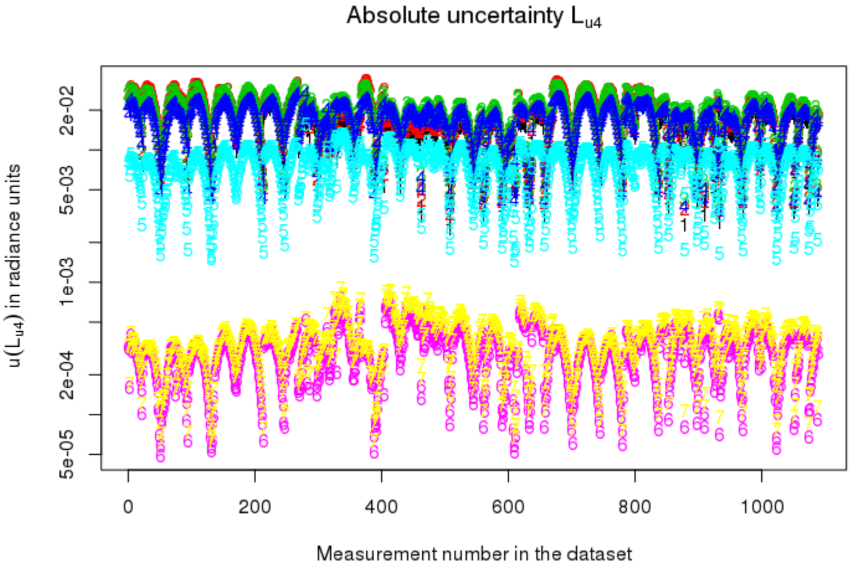


Figure 4.9: Absolute uncertainty in L_{u_4} (k=1).

There are some outliers with uncertainty values much higher than the expected range. The presence of these points indicates imperfect data screening that still needs human interaction to fully validate the QC of the data.

The attenuation coefficient, K_{lu}, is a bit more complex as here, radiometric measurements from 4 and 9 meters are included with all aforementioned uncertainty contributors that affect them and, in addition uncertainty in the depth measurements. This uncertainty is presented as absolute uncertainty thus as m^{-1} in Figure 4.10. The spectral bands are represented as data series.

The highest uncertainties in the attenuation coefficient are observed for channel 7 (681 nm). The remaining spectral channels have uncertainty at similar levels, with channel 6 (670 nm) exhibiting slightly higher values. The uncertainty for channel 4 (510 nm) is at the same level as the other blue/green spectral bands. The calibration uncertainty that pulled up this value for L_{u_4} is no longer so dominant as the calibration uncertainty for the radiometers at two depths are strongly correlated.

The next processing step includes extrapolation of the upwelling radiance measurements from the shallowest depth to just beneath the surface 0^-, so called L_{u0^-}. At this stage

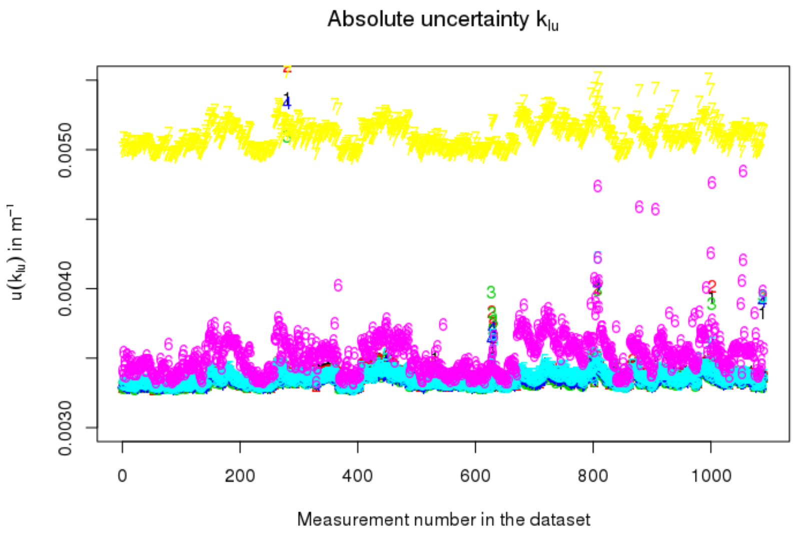


Figure 4.10: Absolute uncertainty in K_{lu} (k=1).

the uncertainty in Hydrolight correction is added to the modelling. The uncertainty in L_{u0^-} is strongly dependant on the depth of the instrument as this depth, together with the attenuation coefficients is used in the exponential expression used for the extrapolation of the measurements to the surface. Thus the shorter extrapolation distance the more certain the value of the upwelling radiance just beneath the surface would be. If the traditional GUM approach using the Law of Propagation of Uncertainty was used, the relative uncertainty in L_{u0^-} would be expressed as equation 4.10.

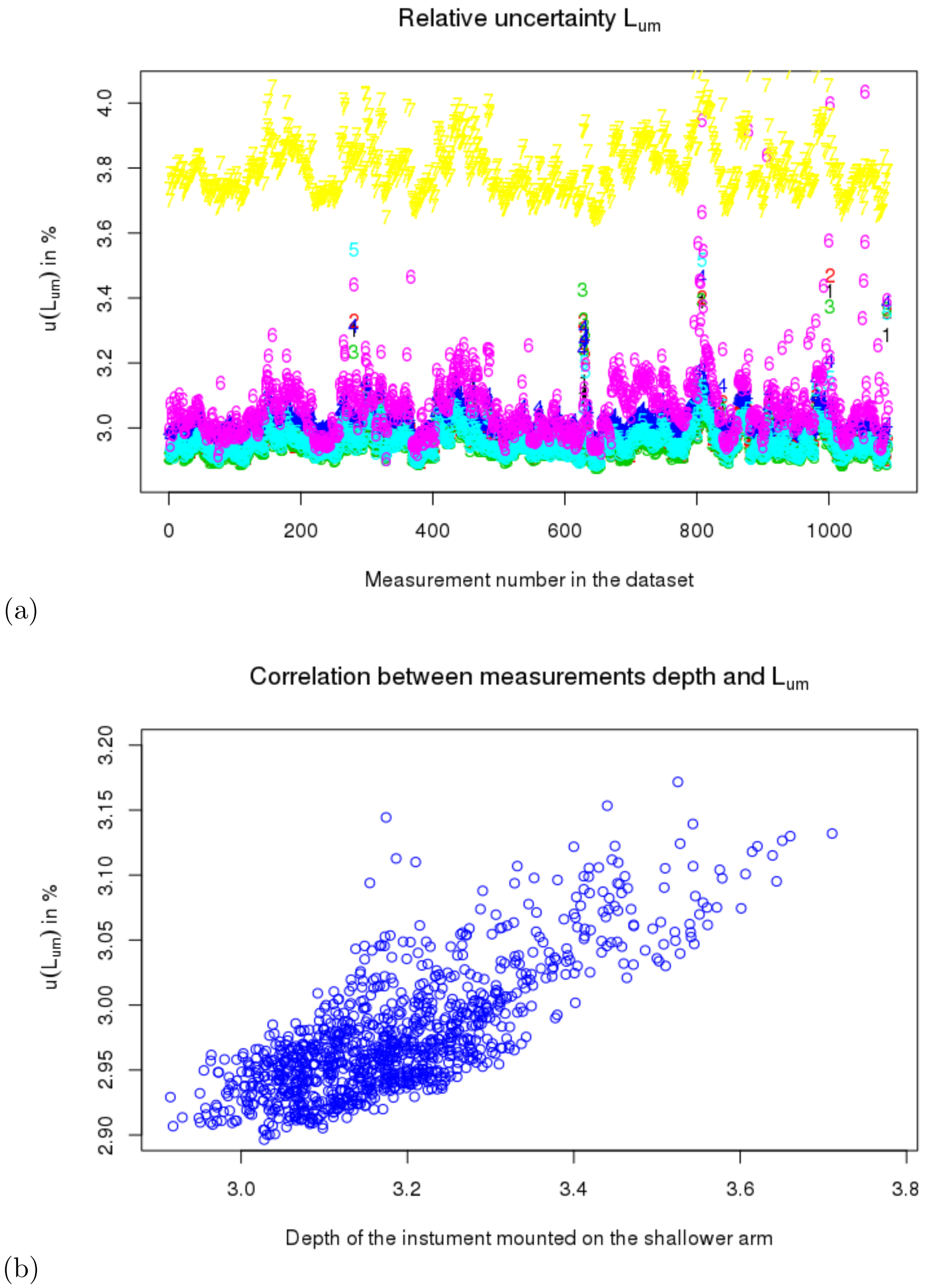


Figure 4.11: (a) Relative uncertainty in $\mathrm{L_{u0-}}$ (k=1). (b) Correlation between instrument depth and $\mathrm{L_{u0-}}$, shown on this plot for 412 nm spectral band.

$$\frac{u^2(\mathrm{L_{u0-}}))}{(\mathrm{L_{u0-}}))^2} = \frac{u^2(\mathrm{L_{u_4}})}{(\mathrm{L_{u_4}})^2} + z^2 u^2(\mathrm{K_{lu}}) + \mathrm{K_{lu}}^2 u^2(z) \tag{4.10}$$

Thus, the depth of the instruments has a role of sensitivity coefficient for the attenuation coefficients uncertainty. This effect is presented in Figure 4.11, panel (b), where the correlation between the actual depth of the radiometer and the uncertainty in L_{u0^-} is clearly visible.

To obtain water-leaving radiance L_w the L_{u0^-} has to be multiplied by the sea-air constant. This constant has a higher uncertainty value associated with the shortest and longest wavelengths. The results of relative uncertainty in L_w are presented in Figure 4.12.

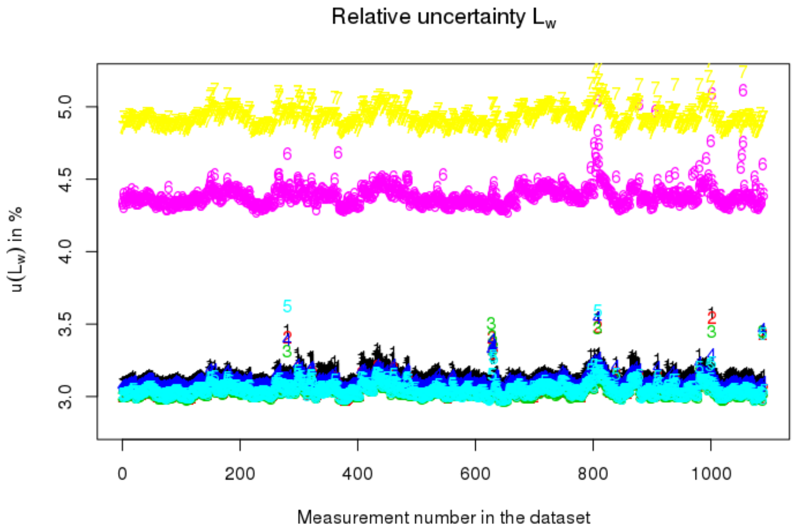


Figure 4.12: Relative uncertainty in L_w (k=1).

Total downwelling irradiance measured above water is used as a denominator in the calculation equation 4.4. This uncertainty is mostly affected by the tilt of the buoy and SZA. The relationship between an uncertainty in downwelling irradiance for all channels in relationship to SZA is presented in Figure 4.13. The same data series numeric convention is used as for the plots presented before, where series 1 represents first spectral bans 412 nm and 7 the last 681 nm one.

E_s is presented for each wavelength against the SZA as this is the factor that influences the most uncertainty values. The clear step change in the uncertainty values at the SZA 60° is caused by a step change in the direct part of the cosine diffuser uncertainty, this value then tends to decrease due to the decreasing ratio of the dirrect Sun irradiance

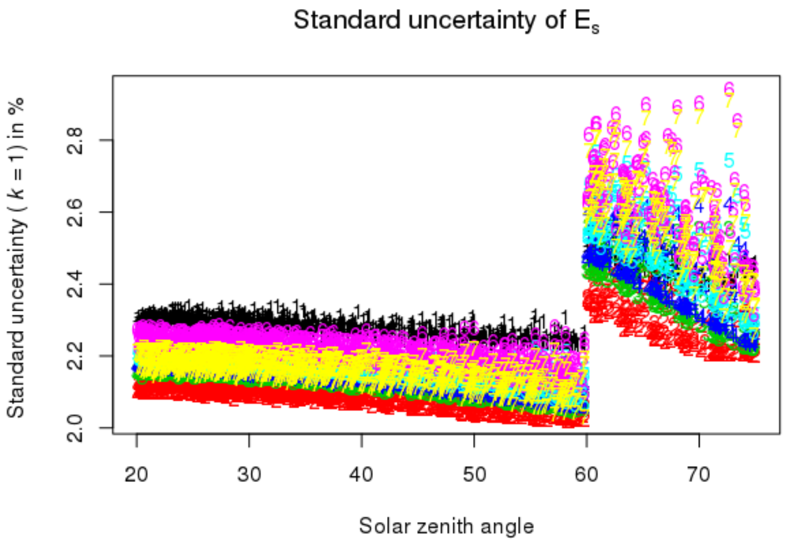
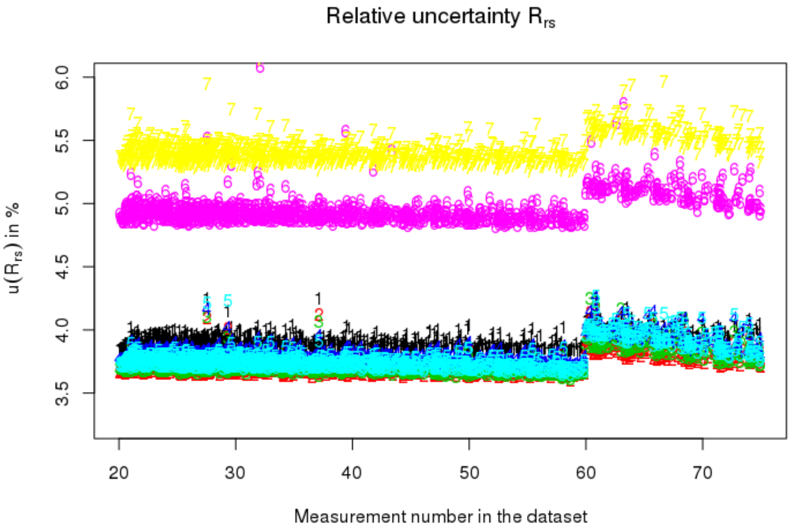


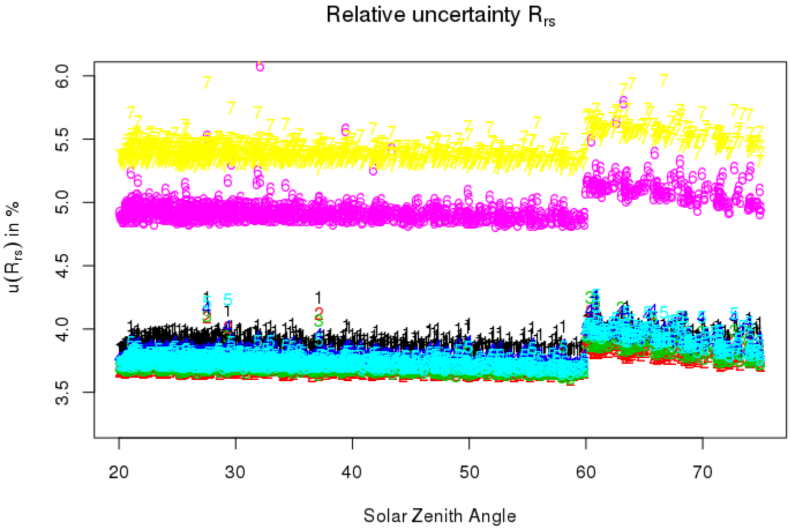
Figure 4.13: Relative uncertainty in E_s (k=1) as a function of SZA.

with the increase of SZA.

The uncertainty in the final product R_{rs} for all the points in the data set is presented in Figure 4.14 panel (a) and the same uncertainty plotted as a function of SZA on panel (b). The two red channels exhibit the highest uncertainty values as they are driven by the environmental conditions, such as signal to noise ratio and shading, and well as modelling aspects related to the Hydrolight corrections and air-sea constant. For this wavelength the instrument related uncertainties are lower and the absolute radiometric sources have lower uncertainties for this wavelength range. The blue channels have the higher instrumental related uncertainties but *in situ* environment does affect them less than the red channels, thus their overall uncertainty is below 4%.



(a)



(b)

Figure 4.14: (a) Relative uncertainty in R_{rs} (k=1). (b) Relative uncertainty in R_{rs} (k=1) as a function of SZA.

A summary table with all wavelengths and a generic value for uncertainty from this study data set is presented in Table 4.4. The columns show different quantities that are used to obtain the final product R_{rs}. This universal value is calculated as the highest density probability from all uncertainties calculated for each data point per spectral band. The examples of the histogram for individual spectral bands and given quality are shown in Figure 4.15. These values are used as a general idea of uncertainties that are most likely to be achieved. However, as presented in Figure 4.15 both lower and higher values can be observed depending on the actual conditions at the site, and they can be extracted for every individual measurement which is useful for SVC and uncertainties associated with each match up.

The stability of the model results was tested by running the same simulation several times and comparing the summary output values. For the runs with 10^5 repetition the model's outputs vary less than 0.1%.

Table 4.4: BOUSSOLE uncertainty budget (k=1).

Wavelength [nm]	E_s	L_{u_4}	L_{u0^-}	L_w	R_{rs}
412	2.29%	2.40%	2.96%	3.14%	3.86%
443	2.11%	2.40%	2.95%	3.04%	3.68%
490	2.16%	2.40%	2.94%	3.02%	3.70%
510	2.17%	2.49%	3.01%	3.09%	3.77%
560	2.20%	2.40%	2.93%	3.02%	3.73%
670	2.23%	2.43%	3.03%	4.38%	4.88%
683	2.17%	2.78%	3.78%	4.90%	5.35%

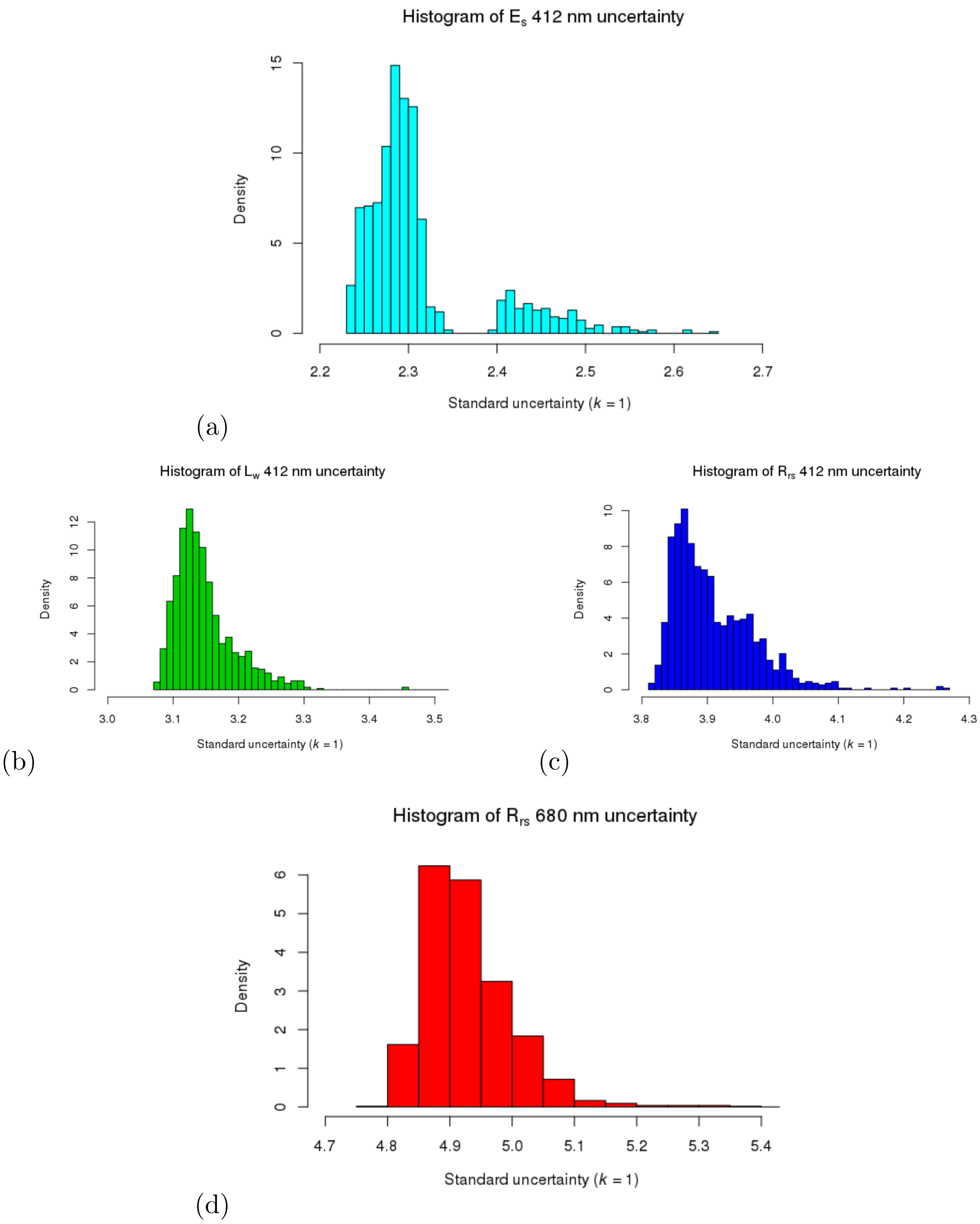


Figure 4.15: Histogram of relative uncertainty in one spectral band for selected quantities (a) E_s(412 nm), (b) L_w(412 nm), (c) R_{rs}(412 nm) and (d) R_{rs}(680 nm).

4.6 Discussion

BOUSSOLE is an excellent asset for the European OC community and has provided important *in situ* radiometric data for over a decade now. The SI traceability needs to be enhanced, as presently the regular manufacturer's calibration comes without uncertainty budgets. One-off calibration for one set of instruments was provided by NPL back in 2012 and to derive this uncertainty budget the uncertainty of that calibration and comparison to the manufacturer's calibration form a basis to estimate the radiometric calibration uncertainty. There is a significant potential for improvement and reduction of this value. Regular six-monthly calibrations with full, robust SI traceability link and with an uncertainty budget should become a default option in the future. The reduction in that uncertainty component will decrease uncertainties in the blue and green spectral channel, and the value of 3% could be achieved. Improvements in absolute calibration will not have such a significant effect on red channels since other uncertainty components related to *in situ* environmental factors, and the data processing chain are substantial and will still be there.

The MCM used to evaluate uncertainties addresses correlations in the input components. This is achieved by selective sampling of the PDFs of each input component. Thus, for example, the absolute calibration part is highly correlated for both upwelling radiance readings from 4 and 9 meters. Hence, the same draw from k_{cal} PDF is used in the calculation related to them. This effect is visible in a change in the relative uncertainty value between $\mathrm{L_{u_4}}$ for channel 4 490 nm. This channel has a higher absolute radiometric calibration uncertainty assigned (see Figure 4.8) where the $\mathrm{L_{u_4}}$ uncertainties are higher than for channel 3 and 5. However, this relation is not observed for $\mathrm{K_{lu}}$ (see Figure 4.10) as to calculate $\mathrm{K_{lu}}$ a ratio of two upwelling radiances is used.

As expected, an explicit correction between the depth of the shallower instrument and $\mathrm{L_{u0^-}}$ is observed as that depth represents the length of data exportation. The shorter the distance, the less uncertainty due to the attenuation coefficient in $\mathrm{L_{u0^-}}$. To minimise this uncertainty component, the instruments should be placed as close to the sea surface as possible. However, then a much stronger wave focusing effect will affect other uncertainty components. Therefore, the overall uncertainty has to always be calculated for the whole system as individual components separately might affect different aspects of the measurements.

The uncertainty in $\mathrm{E_s}$ is affected by the non-perfect cosine response of the diffuser and due to that SZA during the measurements. The recent recommendation report (Mazeran et al. 2017) stated that *in situ* final product should be considered as $\mathrm{L_w}$ or $\mathrm{L_{WN}}$ with a modelled value of $\mathrm{E_s}$, using the same model as the satellite data processing chain, rather than measurement*in situ* to improve the accuracy of the SVC. Thus, the $\mathrm{E_s}$ *in situ* might be used as ancillary information for quality control and cloud screening purposes in the future. If this will become an official approach taken by all data providers, then it seems not fit for purpose to invest in better quality $\mathrm{E_s}$ instruments and their characterisation.

The uncertainty of the red channels will always be very challenging due to the small signal levels. As the final quantity is close to zero it is very easy to have high relative uncertainty. For example, a typical value of $\mathrm{L_w}$ at 440 nm is 0.8458 and relative uncertainty 3.11%, with absolute value of 0.0226 in the radiance units. For the red channel 680 nm, the $\mathrm{L_w}$ is 0.0187 with relative uncertainty 4.90% but absolute 0.0008. Currently, the requirements for 5% with the aim of reaching 3% in the future are defined for blue and green wavelengths recognising the signal issue for the red spectral range. The future requirements might set the expected threshold in the uncertainty value in absolute units.

Ongoing work on defining the uncertainty budget is taking place for the MOBY site simultaneously. The recent publications include looking at uncertainty in air-sea constant and depth extrapolation (Voss & Flora 2017a, Voss et al. 2017b) shows good agreements with approaches presented in this study. In Voss et al. (2017b) MOBY team looked at the air-sea constant wavelength dependence and come to the same conclusion that this value 0.543 should actually change with wavelength. The second publication (Voss & Flora 2017a) presents the method to improve upwelling radiance extrapolation to just beneath the surface similarly to what is here called Hydrolight correction. Although, MOBY team used different modeling approach (Gordon 1979) they validated it with the Hydrolight used it this study.

4.7 Summary and Conclusions

This chapter presented an updated uncertainty budget for the BOUSSOLE radiometric measurements that are used for SVC applications. The new budget is evaluated per spectral channel for multispectral instruments operated at the site. The MCM was used to evaluate the final values. The main advantage of this method is the uncertainty evaluation for a given observation rather than one generic value for the BOUSSOLE site. These uncertainty values will alter due to changes in environmental conditions during the measurements, thus each day or at different times during the same day can have bespoke uncertainty values associated with the measurement. It is then up to a data user to decide whether a particular measurement with associated uncertainty is suitable for further use or not.

This budget was prepared for the existing data processing route and indicated the values of each contribution. It is now easier to justify efforts in reducing some of them as the results can be clearly seen. For example, in future processing, the air-water constant may vary with wavelengths, especially in the case of hyperspectral instruments that cover wider spectral ranges to avoid easily reducible bias in the data. It is almost definite that continuously increased efforts in the absolute radiometric calibration and proven track of its stability would considerably lower the BOUSSOLE product uncertainties. More effort put in to the modelling of the BRDF effects that would allow for correction is another example of possible future improvements.

Further work will apply the same framework to hyperspectral instruments that are in operation on the BOUSSOLE buoy. These radiometers will have more uncertainty components related to its characteristics and differences in the operational mode, such as varying integration time and simultaneous dark readings.

Chapter 5

Conclusions and Future Work

The work presented in this thesis focused on measurement uncertainty estimation and highlighted the importance and implications of this. To ensure that the measurements are meaningful and trustful they have to be traceable, ideally to the SI unit, and be accompanied by a quality indicator. Uncertainty is a good candidate for such an unbiased and quantitative indicator. Only the measurement accompanied with this information can be considered as completed. Uncertainties were put into the context of the test sites that are used for vicarious calibration of optical satellite sensors.

Two test sites were presented, one on land and one on the ocean. The land site is used for radiometric calibrations of L1 satellite products. The new Gobabeb site was established during this study. The primary input of this research is revised traceability of the ground reflectance measurements. At the beginning of this research, the field scientists tended to say that the reference standard reflectance or BRF was the main source of uncertainty during *in situ* measurements which seemed unrealistic as this was one of the surest things that could be measured as a standard in a laboratory to a very high accuracy. Very soon after first field campaign it became clear that firstly that bright white reference standard become dirty very quickly and secondly, that the illumination in the field was nowhere near the strictly defined laboratory conditions.

As a result of this experience, a detailed description of the revised reflectance terminology that highlighted the discrepancies between the laboratory and the field defined quantities known as BRF or HRDF was provided. Then, an updated version of the reflectance standard calibration that would combine direct and diffuse components and allows having SI traceable measurements *in situ* proposed.

The main site characterisation outcome, apart from the new ground results that show it is a very bright and spectrally flat site, is the agreement in measurement between two independent teams. The uncertainty budget for *in situ* measurements is described step by step and can be used as an example for other scientists less familiar with the uncertainty evaluation. The agreement between the team validates the uncertainty estimation and confirms that following strict measurement protocols is essential to obtaining repeatable and reproducible results.

The issue of the reference standard degradation *in situ* is another interesting aspect of this research. The first lesson learnt was to not bring a new unconditioned panel *in situ*, especially for sandy surfaces and windy environments. The severe standard degradation did not ruin the site measurements. The post-campaign calibration is an effective way to address the changes and the use of calibrated instruments, although not needed for reflectance-based methods that added a bonus of ability to measure panel radiance when it is necessary for ancillary data use. In this case, to monitor daily panel changes.

The permanent instrument was truly calibrated and characterised before installation at Gobabeb and is now back at NPL for the next set of laboratory tests. The NPL calibration provided lower uncertainty then the manufacturer calibration, but most importantly reached an excellent agreement with the manufacturer calibration. That allows them to reconsider the default uncertainty budget provided and re-evaluate it to become less conservative and more realistic. An unexpected result of the temperature dependence for the InGaAs detector for the sun-photometer will require further investigation.

In the course of this study, a new RadCalNet site was established and it is now operating. This was a joint effort of teams from the NPL and CNES supported by the mast manufacturer Clarks Mast and the sun photometer manufacturer CIMEL engineer. Many practical and administrative issues had to be addressed to accomplish the task and sometimes these seemed to be more challenging to sort out than the scientific challenges.

The work on the ocean site was related to the existing BOUSSOLE site that is permanently deployed in the Ligurian sea and is used for the L2 Ocean Colour System Vicarious Calibration. The instruments on this site provide data almost constantly through the years of its operation. The SVC is based on the match up with the satellite sensors; however, here the environmental conditions can vary from one match up to another thus the quality of *in situ* measurements can vary highly. Currently, all match-ups that are used in SVC have the same weight. A dynamic uncertainty budget for the radiometric products of that site was developed. This means that an uncertainty is calculated for each measurement and is not a constant value assigned to this system. Thus, in future, “ground truth” measurements can be weighted according to their quality and the SVC will not be biased by poor quality measurements.

The biggest challenge for the BOUSSOLE buoy is the access to high quality absolute radiometric calibrations on a regular basis and to establish an accurate method to track instrument stability. Currently, only a single NPL calibration back in 2012 is the closest link to SI and is not sufficient to maintain the SI traceability. The regular radiometric calibration provided by the manufacturer does not have uncertainty budgets. This is hopefully going to change in the near future as the whole SVC community have worked very hard in the recent year on SI traceability. The ESA funded FRM4SOC project involved the worldwide inter-comparison of many interested laboratories involved in the radiometric calibration of the ocean colour instruments, and the participants are working with NPL on their uncertainty budget. Satlantic, the BOUSSOLE buoy instrument manufacturer, is one of them.

Prospectus work plans for both land and ocean aspects look extremely exciting. Next year, a second field campaign will repeat the measurements from November 2015 and a new portable transfer radiometer will be used so that it can be sent, in turn, to all RadCalNet sites to provide site inter-comparison as a validation of the SI traceability uncertainty evaluation. The next site characterisation will be performed using the new reference standard calibration method as the instrument operating on the site is measuring the required atmospheric inputs. The second set of laboratory calibrations after a year of operation will allow to begin to monitor instrument stability by an alternative method to the so-called *in situ* calibration with the Sun, thus giving the option of comparison and validation. Temperature sensitivity tests especially for the InGaAs detector will be repeated and investigated further.

Longer terms plans for the Gobabeb site are to be expanded and become a part of a new HYPERNETS (*HYPETNETS*) network with additional instruments. HYPERNETS is a project that started in 2018 and has four years to establish a new network of land and water test sites with hyperspectral instruments performing multi-angular measurements. The work on a prototype instrument is ongoing, and NPL is preparing plans for the land site operational aspects. The advantage of the HYPETNETS instrument is the ability to make nadir ground measurements as well as multi-angular, whereas currently at Gobabeb the permanent instrument starts the measurements from VZA$= 10^\circ$. The combination of these two instruments in proximity will enable extension of the measurements capabilities of each of them and will provide a daily data enabling comparison.

A new ocean colour site will be established in Europe in the coming years using modern technology of the sensors and calibration sources. All those will require uncertainty budget evaluations. The new instruments are likely to have integrated internal stability sources and a dedicated *in situ* source to monitor the stability while in operation, which will be the major improvement to the existing buoy. This will allow monitoring of the stability *in situ* continually, firstly to accurately estimate it, and secondly to apply proper corrections and reduce uncertainty in the buoy final product.

Before that the hyperspectral instruments currently in operation at BOUSSOLE will have their uncertainty budget evaluated. A few effects that were not addressed in the first budget, such as polarisation and BRDF, are on the agenda for the second iteration. The primary outcome from the existing uncertainty budget shows that a very simple aspect of the air-sea constant can be quickly improved by only using better constrained numbers. This will already enhance the blue and the red ends of spectrum which is in particular necessary for hyperspectral instruments as they have extended spectral range in comparison to multispectral.

SI traceability and the reduction of absolute radiometric calibrations will benefit all site measurements. The calibration needs to be fit for purpose, and a new tuneable laser based radiometric facilities offer a big step in the uncertainty reduction. Matching even more the laboratory conditions to the field one, or adding laboratory tests to estimate better the difference between laboratory calibration and instruments operation is the next step in improving these types of measurements. For example, a tuneable laser facility might allow for reducing the spectral mismatch for ocean sites as currently the spectral shape of the calibration source is opposite to what is presented *in situ*.

The next big challenge for VC is to start evaluating uncertainties in RTC. The atmosphere is a significant contributor to the TOA signal and very often now the models do not have any uncertainties to assign.

Bibliography

Anderson, N., Czapla-Myers, J., Leisso, N., Biggar, S., Burkhart, C., Kingston, R. & Thome, K. (2013), 'Design and calibration of field deployable ground-viewing radiometers', *Appl. Opt.* **52**(2), 231–240.

Antoine, D., d'Ortenzio, F., Hooker, S. B., Becu, G., Gentili, B., Tailliez, D. & Scott, A. J. (2008b), 'Assessment of uncertainty in the ocean reflectance determined by three satellite ocean color sensors (MERIS, SeaWiFS and MODIS-A) at an offshore site in the Mediterranean sea (BOUSSOLE project)', *JOURNAL OF GEOPHYSICAL RESEARCH-OCEANS* **113**(C7).

Antoine, D., Guevel, P., Deste, J.-F., Becu, G., Louis, F., Scott, A. J. & Bardey, P. (2008a), 'The "BOUSSOLE" buoy - A new transparent-to-swell taut mooring dedicated to marine optics: Design, tests, and performance at sea', *JOURNAL OF ATMOSPHERIC AND OCEANIC TECHNOLOGY* **25**(6), 968–989.

ASD FieldSpec (2018). visited on 22/06/2018.
URL: *https://www.malvernpanalytical.com/en/products/product-range/asd-range/fieldspec-range/fieldspec-4-hi-res-ng-spectroradiometer*

ASTM E2387-05(2011) (2011). Standard Practice for Goniometric Optical Scatter Measurements.

Austin, R. & Halikas, G. (1976), The Index of Refraction of Seawater, Scripps Institution of Oceanography AD-A024 800, Naval Supply Center, U.S. DEPARTMENT OF COMMERCE National Technical Information Service.

Austin, R. W. (1974), The remote sensing of spectral radiance from below the ocean surface, *in* N. G. Jerlov & E. Steemann-Nielsen, eds, 'Optical Aspects of Oceanography', Elsevier, New York., p. 317 344.

Berk, A., Conforti, P., Kennett, R., Perkins, T., Hawes, F. & van den Bosch, J. (2014), Modtran6: A major upgrade of the MODTRAN radiative transfer code, *in* '2014 6th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)', pp. 1–4.

Bialek, A., Greenwell, C., Lamare, M., Meygret, A., Marcq, S., Lacherade, S., Woolliams, E., Berthelot, B., Bouvet, M., King, M., Underwood, C. & Fox, N. (2016a), New radiometric calibration site located at Gobabeb, Namib desert, *in* '2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)', pp. 6094–6097.

Bialek, A., Vellucci, V., Gentili, B., Antoine, D., Fox, N. & Underwood, C. (2016b), 'An uncertainty budget for the BOUSSOLE radiometry, as derived using a Monte Carlo Method'. Ocean Optics XXV, Victoria, Canada.

Biggar, S. F., Dinguirard, M. C., Gellman, D. I., Henry, P. J., Jackson, R. D., Moran, M. S. & Slater, P. N. (1991), Radiometric calibration of SPOT 2 HRV: a comparison of three methods, Vol. 1493, pp. 1493 – 1493 – 8.

URL: *https://doi.org/10.1117/12.46693*

Biggar, S. F., Labed, J., Santer, R. P., Slater, P. N., Jackson, R. D. & Moran, M. S. (1988), Laboratory Calibration Of Field Reflectance Panels, Vol. 0924, pp. 0924 – 0924 – 9.

URL: *https://doi.org/10.1117/12.945691*

Biggar, S. F., Santer, R. P. & Slater, P. N. (1990), Irradiance-based Calibration Of Imaging Sensors, *in* '10th Annual International Symposium on Geoscience and Remote Sensing', pp. 507–510.

BIPM (2018). visited on 22/06/2018.

URL: *https://www.bipm.org/en/about-us/*

Brown, S., Flora, S., Feinholz, M., Yarbrough, M., Houlihan, T., Peters, D., Kim, Y., Mueller, J., Johnson, B. & Clark, D. (2007), The marine optical buoy (MOBY) radiometric calibration and uncertainty budget for ocean color satellite sensor vicarious calibration, Vol. 6744, pp. 6744 – 6744 – 12.
URL: *https://doi.org/10.1117/12.737400*

Budde, W. (1976), 'Calibration of Reflectance Standards', *Journal of Research of the National Bureau of Standards Section A-physics and Chemistry* **80**(4), 585–595.

Cabot, F., Hagolle, O. & Henry, P. (2000), Relative and multitemporal calibration of AVHRR, SeaWiFS, and VEGETATION using POLDER characterization of desert sites, *in* 'IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. InProceddings (Cat. No.00CH37120)', Vol. 5, pp. 2188–2190 vol.5.

CalVAlPortal (2018), 'Vicarious Calibration - Cal/Val wiki - CalValPortal'. visited on 22/06/2018.
URL: *http://calvalportal.ceos.org/cal/val-wiki/-/wiki/CalVal+Wiki/Vicarious+Calibration*

Castle, K. R., Holm, R. G., Kastner, C. J., Palmer, J. M., Slater, P. N., Dinguirard, M., Ezra, C. E., Jackson, R. D. & Savage, R. K. (1984), 'In-Flight Absolute Radiometric Calibration of the Thematic Mapper', *IEEE Transactions on Geoscience and Remote Sensing* **GE-22**(3), 251–255.

Chunnilall, C., Deadman, A., Crane, L. & Usadi, E. (2003), 'NPL scales for radiance factor and total diffuse reflectance', *METROLOGIA* **40**(1, SI), S192–S195. 8th International Conference on New Developments in Optical Radiometry (NEWRAD 2002), NATL INST STANDARDS & TECHNOL, GAITHERSBURG, MD, MAY 20-24, 2002.

Cimel Electronique (2018). visited on 22/06/2018.
URL: *https://www.cimel.fr/?lang=en*

Cosnefroy, H., Leroy, M. & Briottet, X. (1996), 'Selection and characterization of Saharan and Arabian desert sites for the calibration of optical satellite sensors', *Remote Sensing of Environment* **58**(1), 101 – 114.

Czapla-Myers, J., McCorkel, J., Anderson, N., Thome, K., Biggar, S., Helder, D., Aaron, D., Leigh, L. & Mishra, N. (2015), 'The Ground-Based Absolute Radiometric Calibration of Landsat 8 OLI', *REMOTE SENSING* **7**(1), 600–626.

Czapla-Myers, J. S., Thome, K. J., Cocilovo, B. R., McCorkel, J. T. & Buchanan, J. H. (2008), Temporal, spectral, and spatial study of the automated vicarious calibration test site at Railroad Valley, Nevada, Vol. 7081, pp. 7081 – 7081 – 9.

DD ISO/TS 28037:2010 (2011). Determination and use of straight-line calibration functions.

Dictionary (2018). visited on 23/06/2018.
URL: *http://www.dictionary.com/browse/vicarious*

Dinguirard, M. & Slater, P. N. (1999), 'Calibration of Space-Multispectral Imaging Sensors: A Review', *Remote Sensing of Environment* **68**(3), 194 – 205.

Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M.-H., Fmnias, P., Frerick, J., Goryl, P., Klein, U., Laur, H., Mavrocordatos, C., Nieke, J., Rebhan, H., Seitz, B., Stroede, J. & Sciarra, R. (2012), 'The Global Monitoring for Environment and Security (GMES) Sentinel-3 mission', *Remote Sensing of Environment* **120**, 37 – 57. The Sentinel Missions - New Opportunities for Science.
URL: *http://www.sciencedirect.com/science/article/pii/S0034425712000685*

Drusch, M., Bello, U. D., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F. & Bargellini, P. (2012), 'Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services', *Remote Sensing of Environment* **120**, 25 – 36. The Sentinel Missions - New Opportunities for Science.
URL: *http://www.sciencedirect.com/science/article/pii/S0034425712000636*

Emde, C., Buras-Schnell, R., Kylling, A., Mayer, B., Gasteiger, J., Hamann, U., Kylling, J., Richter, B., Pause, C., Dowling, T. & Bugliaro, L. (2016), 'The libRadtran software package for radiative transfer calculations (version 2.0.1)', *Geoscientific Model Development* **9**(5), 1647–1672.

URL: *https://www.geosci-model-dev.net/9/1647/2016/*

Fox, N. P., Aiken, J., Barnett, J. J., Briottet, X., Carvell, R., Froehlich, C., Olivier Hagolle, S. B. G., Haigh, J. D., Lean, J., Pollock, D. B., Kieffer, H. H., Quinn, John, T., Sandford, M. C. W., Schaepman, M. E., Shine, K. P., Schmutz, W. K., Teillet, P. M., Thome, K. J., Verstraete, M. M. & Zalewski, E. F. (2003), Traceable radiometry underpinning terrestrial- and helio-studies (TRUTHS), Vol. 4881, pp. 4881 – 4881 – 12.

URL: *https://doi.org/10.1117/12.462438*

Franz, B. A., Bailey, S. W., Werdell, P. J. & McClain, C. R. (2007), 'Sensor-independent approach to the vicarious calibration of satellite ocean color radiometry', *Appl. Opt.* **46**(22), 5068–5082.

URL: *http://ao.osa.org/abstract.cfm?URI=ao-46-22-5068*

Garca, A. F., da Silva Curiel, A., Davies, P., Penson, J., de Groot, Z., Hodgson, D. & Stephens, P. (2008), Deimos-1 - The Next Generation of Commercial DMC Spacecraft, *in* 'In Proceddings of the IAA Symposium on Small Satellite Systems and Services (4S)'.

Gellman, D. I., Biggar, S. F., Dinguirard, M. C., Henry, P. J., Moran, M. S., Thome, K. J. & Slater, P. N. (1993), Review of SPOT-1 and -2 calibrations at White Sands from launch to the present, Vol. 1938, pp. 1938 – 1938 – 8.

URL: *https://doi.org/10.1117/12.161536*

Gillham, E. (1962), 'Recent investigations in absolute radiometry', *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **269**(1337), 249–276.

URL: *http://rspa.royalsocietypublishing.org/content/269/1337/249*

Gordon, H. (1998), 'In-Orbit Calibration Strategy for Ocean Color Sensors', *Remote Sensing of Environment* **63**(3), 265 – 278.

URL: *http://www.sciencedirect.com/science/article/pii/S0034425797001636*

Gordon, H. R. (1979), 'Diffuse reflectance of the ocean: the theory of its augmentation by chlorophyll a fluorescence at 685 nm', *Appl. Opt.* **18**(8), 1161–1166.
URL: *http://ao.osa.org/abstract.cfm?URI=ao-18-8-1161*

Gordon, H. R. (1997), 'Atmospheric correction of ocean color imagery in the Earth Observing System era', *Journal of Geophysical Research: Atmospheres* **102**(D14), 17081–17106.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JD02443*

Gordon, H. R. & Ding, K. (1992), 'Self-shading of in-water optical instruments', *Limnology and Oceanography* **37**(3), 491–500.
URL: *http://dx.doi.org/10.4319/lo.1992.37.3.0491*

Govaerts, Y. M. & Clerici, M. (2004), 'Evaluation of radiative transfer simulations over bright desert calibration sites', *IEEE Transactions on Geoscience and Remote Sensing* **42**(1), 176–187.

Greenwell, C., Bialek, A., Marks, A., Woolliams, E., Berthelot, B., Meygret, A., Marcq, S., Bouvet, M. & Fox, N. (2015), Preparation of a new autonomous instrumented radiometric calibration site: Gobabeb, Namib Desert, Vol. 9639, pp. 963919–963919–14.
URL: *http://dx.doi.org/10.1117/12.2194885*

Gregg, W. & Carder, K. (1990), 'A Simple Spectral Solar Irradiance Model for Cloudless Maritime Atmospheres', *LIMNOLOGY AND OCEANOGRAPHY* **35**(8), 1657–1675.

Heidinger, A. K., Sullivan, J. T. & Rao, C. R. N. (2003), 'Calibration of visible and near-infrared channels of the NOAA-12 AVHRR using time series of observations over deserts', *International Journal of Remote Sensing* **24**(18), 3635–3649.

Helder, D., Pesta, F., Brinkmann, J., Leigh, L., Aaron, D., Markham, B., Julia Barsi, R. M. & Czapla-Myers, J. (2013), 'Landsat-8 OLI: On-Orbit Spatial Uniformity, Absolute Calibration and Stability.'. Conference on Characterization and Radiometric Calibration for Remote Sensing (CALCON).

Holben, B. N., Kaufman, Y. J. & Kendall, J. D. (1990), 'NOAA-11 AVHRR visible and near-IR inflight calibration', *International Journal of Remote Sensing* **11**(8), 1511–1519.

Hooker, S. B., Esaias, W. E., Feldman, G. C., Gregg, W. W. & Mc Clain, C. R. (1992), An overview of SeaWiFS and ocean colour in Seawifs Technical Report Series, NASA Tech. Memo 104566, NASA Goddard Space Flight Centre.

Hueni, A. & Bialek, A. (2017), 'Cause, Effect, and Correction of Field Spectroradiometer Interchannel Radiometric Steps', *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **10**(4), 1542–1551.

HYPETNETS (2018). visited on 16/07/2018.
URL: *http://www.hypernets.eu*

IOCCG (2010), IOCCG Report 10 Atmospheric Correction for Remotely-Sensed Ocean-Colour Products, Reports and Monographs of the International Ocean-Colour Coordinating Group, IOCCG.

IOCCG (2012), IOCCG Report 13 Mission Requirements for Future Ocean-Colour Sensors, Reports and Monographs of the International Ocean-Colour Coordinating Group, IOCCG.

JCGM100:2008 (2008), Evaluation of measurement data - Guide to the expression of uncertainty in measurement, Guidance document, BIPM.

JCGM101:2008 (2008), Evaluation of measurement data - Supplement 1 to the "Guide to the expression of uncertainty in measurement" - Propagation of distributions using a Monte Carlo method, Guidance document, BIPM.

JCGM200:2012 (2012), International vocabulary of metrology Basic and general concepts and associated terms (VIM), Guidance document, BIPM.

Judd, D. B. (1967), 'Terms, Definitions and Symbols in Reflectometry', *Journal of the Optical Society of America* **57**(4), 445–&.

Kaufman, Y. J. & Holben, B. N. (1993), ‘Calibration of the AVHRR visible and near-IR bands by atmospheric scattering, ocean glint and desert reflection’, *International Journal of Remote Sensing* **14**(1), 21–52.

Knight, E. J. & Kvaran, G. (2014), ‘Landsat-8 Operational land Imager Design, Characterization and Performance’, *Remote Sensing* **6**(11), 10286–10305.
URL: *http://www.mdpi.com/2072-4292/6/11/10286*

Koepke, P. (1982), ‘Vicarious satellite calibration in the solar spectral range by means of calculated radiances and its application to meteosat’, *Appl. Opt.* **21**(15), 2845–2854.

Kotchenova, S. Y. & Vermote, E. F. (2006), ‘Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. part i: path radiance’, *Appl. Opt* pp. 6762–6774.

Kotchenova, S. Y. & Vermote, E. F. (2007), ‘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**(20), 4455–4464.
URL: *http://ao.osa.org/abstract.cfm?URI=ao-46-20-4455*

Kriebel, K. (1978), ‘Measured spectral bidirectional reflection properties of four vegetated surfaces’, *Appl. Opt.* **17**(2), 253–259.
URL: *http://ao.osa.org/abstract.cfm?URI=ao-17-2-253*

Kuusk, J., Ansko, I., Bialek, A., Vendt, R. & Fox, N. (2018), ‘Implication of Illumination Beam Geometry on Stray Light and Bandpass Characteristics of Diode Array Spectrometer’, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* pp. 1–8.

Labshpere (2018). visited on 22/06/2018.
URL: *https://www.labsphere.com/labsphere-products-solutions/materials-coatings-2/targets-standards/test-child/*

Lacherade, S., Fougnie, B., Henry, P. & Gamet, P. (2013), ‘Cross Calibration Over Desert Sites: Description, Methodology, and Operational Implementation’, *IEEE Transactions on Geoscience and Remote Sensing* **51**(3), 1098–1113.

Lamare, M., Bialek, A., Greenwell, C., Woolliams, E., Lacherade, S., Marcq, S., Meygret, A., Bouvet, M., King, M. & Fox., N. (2016), Characterising the HDRF of a new autonomous instrumented radiometric calibration site: Gobabeb, Namib Desert, *in* '2016 European Space Agency Living Planet Symposium'.

Leymarie, E. (2005), 'Simulo'. visited on 20/20/2017.
URL: *http://omtab.obs-vlfr.fr/SimulO/index.htm*

Li, C., Ma, L., Gao, C., Tang, L., Wang, N., Liu, Y., Zhao, Y., Dou, S., Zhang, D. & Li, X. (2015), 'Permanent target for optical payload performance and data quality assessment: spectral characterization and a case study for calibration', *Journal of Applied Remote Sensing* **8**, 8 – 8 – 14.
URL: *https://doi.org/10.1117/1.JRS.8.083498*

Lozano, F. J., Romo, A., Moclan, C., Gil, J. & Pirondini, F. (2012), The DEIMOS-1 mission: Absolute and relative calibration activities and radiometric optimisation, *in* '2012 IEEE International Geoscience and Remote Sensing Symposium', pp. 4754–4757.

Martin, J. E., Fox, N. P. & Key, P. J. (1985), 'A Cryogenic Radiometer for Absolute Radiometric Measurements', *Metrologia* **21**(3), 147.
URL: *http://stacks.iop.org/0026-1394/21/i=3/a=007*

Mayer, B., Hoch, S. W. & Whiteman, C. D. (2010), 'Validating the MYSTIC three-dimensional radiative transfer model with observations from the complex topography of Arizona's Meteor Crater', *Atmospheric Chemistry and Physics* **10**(18), 8685–8696.
URL: *https://www.atmos-chem-phys.net/10/8685/2010/*

Mayer, B. & Kylling, A. (2005), 'Technical note: The libRadtran software package for radiative transfer calculations - description and examples of use', *Atmospheric Chemistry and Physics* **5**(7), 1855–1877.
URL: *https://www.atmos-chem-phys.net/5/1855/2005/*

Mazeran, C., Brockmann, C., Ruddick, K., Voss, K. & Zagolski, F. (2017), Requirements for Copernicus Ocean Colour Vicarious Calibration Infrastructure, Technical Report REF: SOLVO/EUM/16/VCA/D8, EUMETSAT.

McCain, C., Hooker, S., Feldman, G. & Bontempi, P. (2006), ‘Satellite data for ocean biology, biogeochemistry, and climate research’, *Eos, Transactions American Geophysical Union* **87**(34), 337–343.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006EO340002*

McClain, C. R., Esaias, W. E., Barnes, W., Guenther, B., Endres, D., Hooker, S., Mitchell, G. & Barnes, R. (1992), Calibration and Validation Plan for SeaWiFS, *in* S. B.Hooker & E. R. Firestone, eds, ‘NASA Tech. Memo. 104566’, Vol. vol.3, NASA Goddard Space Flight Center,Greenbelt, Md., pp. 503–514.

Meygret, A., Santer, R. P. & Berthelot, B. (2011), ROSAS: a robotic station for atmosphere and surface characterization dedicated to on-orbit calibration, Vol. 8153, pp. 815311–815311–12.
URL: *http://dx.doi.org/10.1117/12.892759*

Mishra, N., Helder, D., Angal, A., Choi, J. & Xiong, X. (2014), ‘Absolute Calibration of Optical Satellite Sensors Using Libya 4 Pseudo Invariant Calibration Site’, *Remote Sensing* **6**(2), 1327–1346.

Mobley, C., Boss, E. & Roesler, C. (n.d.), ‘Ocean Optics Web Book’. visited on 22/06/2018.
URL: *http://www.oceanopticsbook.info/*

Mobley, C. D. (1994), *Light and water: radiative transfer in natural waters*, Academic press.

Morel, A. (1988), ‘Optical modeling of the upper ocean in relation to its biogenous matter content (case I waters)’, *Journal of Geophysical Research: Oceans* **93**(C9), 10749–10768.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JC093iC09p10749*

Morel, A., Antoine, D. & Gentili, B. (2002), ‘Bidirectional reflectance of oceanic waters: accounting for Raman emission and varying particle scattering phase function’, *Appl. Opt.* **41**(30), 6289–6306.
URL: *http://ao.osa.org/abstract.cfm?URI=ao-41-30-6289*

Morel, A. & Gentili, B. (1996), 'Diffuse reflectance of oceanic waters. III. implication of bidirectionality for the remote-sensing problem', *Appl. Opt.* **35**(24), 4850–4862.
URL: *http://ao.osa.org/abstract.cfm?URI=ao-35-24-4850*

Murphy, R. E., Ardanuy, P., DeLuccia, F. J., Clement, J. E. & Schueler, C. F. (2006), The Visible Infrared Imaging Radiometer (VIIRS), *in* J. Qu, W. Gao, M. Kafatos, R. Murphy & V. Salomonson, eds, 'Earth Science Satellite Remote Sensing', Springer and Tsinghua Univ. Press, pp. 199–223.

NASA PACE (2018). visited on 05/07/2018.
URL: *https://pace.gsfc.nasa.gov/*

Nicodemus, F. E., Richmond, J., Hsia, J., I.W., G. & Limperis, T. (1977), *Geometrical Considerations and Momenclature for Reflectance*, Vol. 160, US Department of Commerce, National Bureau of Standards.

Ocean Optics (2018). visited on 22/06/2018.
URL: *https://oceanoptics.com/*

Özen, H., Fox, N., Leloglu, U. M., Behnert, I. & Deadman, A. (2011), The 2010 Tuz gölü field campaign ;An overview, *in* '2011 IEEE International Geoscience and Remote Sensing Symposium', pp. 3867–3870.

Palmer, J. M. & Grant., B. G. (2010), *Art of radiometry*, SPIE.

Pegrum-Browning, H., Fox, N. & Milton, E. (2008), The NPL Gonio RAdiometric Spectrometer System (GRASS)., *in* 'Proceedings of the Remote Sensing and Photogrammetry Society Conference 2008, 'Measuring change in the Earth system','.

Pegrum, H., Woolliams, E., Fox, N., van Riel, L., Otter, G. & Kowalewski, M. (2004), Calibration of the NPL transfer standard absolute radiance source (TSARS) and its use with GOME 2-FM3 spectral radiance measurements., *in* Meynart, R and Neeck, SP and Shimoda, H, ed., 'SENSORS, SYSTEMS, AND NEXT-GENERATION SATELLITES VIII', Vol. 5570 of *PROCEEDINGS OF THE SOCIETY OF PHOTO-OPTICAL INSTRUMENTATION ENGINEERS (SPIE)*, SPIE; Soc Espanola Opt;

Natl Aeronaut & Space Adm; European Opt Soc, pp. 503–514. Conference on Sensors, Systems, and Next-Generation Satellites VIII, Maspalomas, SPAIN, SEP 13-15, 2004.

Piskozub, J. (2004), 'Effect of 3-D instrument casing shape on the self-shading of in-water upwelling irradiance', *Opt. Express* **12**(14), 3144–3148.

QA4EO (2018). visited on 22/06/2018.
URL: *http://qa4eo.org/*

Rast, M., Bezy, J. L. & Bruzzi, S. (1999), 'The ESA Medium Resolution Imaging Spectrometer MERIS a review of the instrument and its mission', *International Journal of Remote Sensing* **20**(9), 1681–1702.
URL: *https://doi.org/10.1080/014311699212416*

Roujean, J. L., Leroy, M. & Deschamps, P. Y. (1997), 'A bidirectional reflectance model of the Earth's surface for the correction of remote sensing data', *Journal of Geophysical Research: Atmospheres* **97**(D18), 20455–20468.

Salomonson, V. V., Barnes, W. L., Maymon, P. W., Montgomery, H. E. & Ostrow, H. (1989), 'MODIs: advanced facility instrument for studies of the Earth as a system', *IEEE Transactions on Geoscience and Remote Sensing* **27**(2), 145–153.

Schaepman-Strub, G., Schaepman, M. E., Painter, T. H., Dangel, S. & Martonchik, J. V. (2006), 'Reflectance quantities in optical remote sensing-definitions and case studies', *REMOTE SENSING OF ENVIRONMENT* **103**(1), 27–42.

Sea-Bird Scientific (2018). visited on 23/06/2018.
URL: *http://www.seabird.com/*

Shaw, G. E. (1983), 'Sun Photometry', *Bulletin of the American Meteorological Society* **64**(1), 4–10.

Slater, P. N., Biggar, S. F., Holm, R. G., Jackson, R. D., Mao, Y., Moran, M. S., Palmer, J. M. & Yuan, B. (1987), 'Reflectance- and radiance-based methods for the in-flight absolute calibration of multispectral sensors', *Remote Sensing of Environment* **22**(1), 11–37.

Slater, P. N., Thome, S. F. B. K. J., Gellman, D. I. & Spyak, P. R. (1995), In-flight radiometric calibration of ASTER by reference to well-characterized scenes, Vol. 2317, pp. 2317 – 2317 – 12.
URL: *https://doi.org/10.1117/12.198956*

Smith, J., Thome, K., Crowther, B. & Biggar, S. (1998), Field evaluation of a diffuse to global irradiance meter for vicarious calibration, *in* 'Geoscience and Remote Sensing Symposium Proceedings, 1998. IGARSS '98. 1998 IEEE International', Vol. 2, pp. 663–665 vol.2.

Spectra Vista Corporation (2018). visited on 23/06/2018.
URL: *https://www.spectravista.com/hr-1024i/*

Staylor, F. (1990), 'Degradation Rates of the AVHRR Visible Channel for the NOAA 6, 7, and 9 Spacecraft', *J. Atmos. Oceanic Technol.* **7**, 411–423.

Teillet, P., Fedosejevs, G., Gauthier, R., O'Neill, N., Thome, K., Biggar., S., Ripley, H. & Meygret, A. (2001a), 'A generalized approach to the vicarious calibration of multiple Earth observation sensors using hyperspectral data', *Remote Sensing of Environment* **77**(3), 304 – 327.

Teillet, P., Fedosejevs, G., Thome, K. & Barker, J. L. (2007), 'Impacts of spectral band difference effects on radiometric cross-calibration between satellite sensors in the solar-reflective spectral domain', *Remote Sensing of Environment* **110**(3), 393–409.

Teillet, P. M., Thome, K. J., Fox, N. P. & Morisette, J. T. (2001b), Earth observation sensor calibration using a global instrumented and automated network of test sites (GIANTS), *in* 'International Symposium on Remote Sensing', International Society for Optics and Photonics, pp. 246–254.

Thome, K. (2001*a*), 'Absolute radiometric calibration of landsat 7 ETM+ using the reflectance-based method', *Remote Sensing of Environment* **78**(1), 27 – 38.

Thome, K. (2001*b*), 'Absolute radiometric calibration of Landsat 7 ETM+ using the reflectance-based method', *REMOTE SENSING OF ENVIRONMENT* **78**(1-2, SI), 27–38.

Thome, K., Helder, D., Aaron, D. & Dewald, J. (2004), 'Landsat-5 TM and Landsat-7 ETM+ absolute radiometric calibration using the reflectance-based method', *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING* **42**(12), 2777–2785.

Thome, K. J., Gellman, D. I., Parada, R. J., Biggar, S. F., Slater, P. N. & Moran, M. S. (1993), In-flight radiometric calibration of Landsat-5 Thematic Mapper from 1984 to the present, Vol. 1938, pp. 1938 – 1938 – 5.
URL: *https://doi.org/10.1117/12.161537*

Thuillier, G., Hersé, M., Labs, D.and Foujols, T., Peetermans, W., Gillotay, D., Simon, P. & Mandel, H. (2003), 'The Solar Spectral Irradiance from 200 to 2400 nm as Measured by the SOLSPEC Spectrometer from the atlas and Eureca Missions', *Solar Physics* **214**(1), 1–22.

Tyc, G., Tulip, J., Schulten, D., Krischke, M. & Oxfort, M. (2005), 'The RapidEye mission design', *Acta Astronautica* **56**(1), 213 – 219. 4th IAA International Symposium on Small Satellites for Earth Observation.
URL: *http://www.sciencedirect.com/science/article/pii/S0094576504003170*

Vermote, E. F., Tanre, D., Deuze, J. L., Herman, M. & jacques Morcrette, J. (1997), 'Second simulation of the satellite signal in the solar spectrum, 6S'.

Voss, K. & Flora, S. (2017a), 'Spectral dependence of the seawater-air radiance transmission coefficient', *Journal of Atmospheric and Oceanic Technology* **34**(6), 1203–1205.

Voss, K., Gordon, H., Flora, S., Carol Johnson, B., Yarbrough, M., Feinholz, M. & Houlihan, T. (2017b), 'A method to extrapolate the diffuse upwelling radiance attenuation coefficient to the surface as applied to the Marine Optical Buoy (MOBY)', *Journal of Atmospheric and Oceanic Technology* **34**(7), 1423–1432.

Wei, J., Lee, Z., Lewis, M., Pahlevan, N., Ondrusek, M. & Armstrong, R. (2015), 'Radiance transmittance measured at the ocean surface', *Opt. Express* **23**(9), 11826–11837.

Wielicki, B. A., Young, D. F., Mlynczak, M. G., Thome, K. J., Leroy, S., Corliss, J., Anderson, J. G., Ao, C. O., Bantges, R., Best, F., Bowman, K., Brindley, H., Butler, J. J., Collins, W., Dykema, J. A., Doelling, D. R., Feldman, D. R., Fox, N., Huang, X., Holz, R., Huang, Y., Jin, Z., Jennings, D., Johnson, D. G., Jucks, K., Kato, S., Kirk-Davidoff, D. B., Knuteson, R., Kopp, G., Kratz, D. P., Liu, X., Lukashin, C., Mannucci, A. J., Phojanamongkolkij, N., Pilewskie, P., Ramaswamy, V., Revercomb, H., Rice, J., Roberts, Y., Roithmayr, C. M., Rose, F., Sandford, S., Shirley, E. L., Smith, W. L., Soden, B., Speth, P. W., Sun, W., Taylor, P. C., Tobin, D. & Xiong, X. (2013), 'Achieving climate change absolute accuracy in orbit', *Bulletin of the American Meteorological Society* **94**(10), 1519–1539.

Williams, D. (1999), 'Establishment of absolute diffuse reflectance scales using the NPL Reference Reflectometer', *ANALYTICA CHIMICA ACTA* **380**(2-3), 165–172. visited on3rd Conference on Optical Spectrometry - Applications and Instrumentation into the 21st-Century, EGHAM, ENGLAND, JUN 29-JUL 01, 1998.

Willink, R. & White, R. (2011), 'Disentangling Classical and Bayesian Approaches to Uncertainty Analysis'. visited on 05/07/2018.

Woolliams, E. R., Fox, N. P., Cox, M. G., Harris, P. M. & Harrison, N. J. (2006), 'The CCPR K1-a key comparison of spectral irradiance from 250 nm to 2500 nm: measurements, analysis and results', *Metrologia* **43**(2), S98.
URL: *http://stacks.iop.org/0026-1394/43/i=2/a=S20*

Xiong, X. & Barnes, W. (2006), 'An overview of MODIS radiometric calibration and characterization', *Advances in Atmospheric Sciences* **23**(1), 69–79.
URL: *https://doi.org/10.1007/s00376-006-0008-3*

Yoon, H. W., Allen, D. W., Eppeldauer, G. P. & Tsai, B. K. (2009), The extension of the NIST BRDF scale from 1100 nm to 2500 nm, Vol. 7452, pp. 745204–745204–12.
URL: *http://dx.doi.org/10.1117/12.827293*

Zibordi, G. (2006), 'Immersion factor of in-water radiance sensors: Assessment for a class of radiometers', *JOURNAL OF ATMOSPHERIC AND OCEANIC TECHNOLOGY* **23**(2), 302–313.

Zibordi, G. & Bulgarelli, B. (2007), ‘Effects of cosine error in irradiance measurements from field ocean color radiometers’, *Appl. Opt.* **46**(22), 5529–5538.

Zibordi, G., Mlin, F., Berthon, J.-F., Holben, B., Slutsker, I., Giles, D., DAlimonte, D., Vandemark, D., Feng, H., Schuster, G., Fabbri, B. E., Kaitala, S. & Seppl, J. (2009), ‘AERONET-OC: A Network for the Validation of Ocean Color Primary Products’, *Journal of Atmospheric and Oceanic Technology* **26**(8), 1634–1651.

Zibordi, G. & Voss, K. (2010), *Field radiometry and ocean color remote sensing*, Springer Netherlands, pp. 307–334.