Universiteit
Leiden
The Netherlands

Accessible remote sensing of water

Burggraaff, O.

Citation
Burggraaff, O. (2022, December 13). *Accessible remote sensing of water*. Retrieved from https://hdl.handle.net/1887/3497379

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3497379

Note: To cite this publication please use the final published version (if applicable).

Accessible remote sensing of water

Olivier Burggraaff

© Olivier Burggraaff, 2022

Accessible remote sensing of water
PhD thesis, Universiteit Leiden, Leiden, Netherlands
ISBN: 978-94-6421-952-4

Cover art: Stylised image of water in six spectral bands, corresponding to colours from the Forel-Ule scale for natural waters.
Cover design by Sanjana Panchagnula.

Printed by Ipskamp Printing.
An electronic copy of this thesis can be found at https://scholarlypublications.universiteitleiden.nl/.

Accessible remote sensing of water

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof. dr. ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 13 december 2022
klokke 12.30 uur

door

Olivier Burggraaff

geboren te Zoetermeer
in 1996

Promotores:	Prof. dr. C. U. Keller
	Prof. dr. ing. M. G. Vijver
Co-promotor:	Dr. ir. F. Snik
Promotiecommissie:	Prof. dr. H. J. A. Röttgering
	Prof. dr. ir. P. M. van Bodegom
	Prof. dr. E. S. Boss (University of Maine)
	Dr. S. G. H. Simis (Plymouth Marine Laboratory)
	Dr. D. M. Stam (Technische Universiteit Delft)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 776480 (MONOCLE).

Contents

1 General introduction **1**
- 1.1 Water . . . 2
  - 1.1.1 Inorganic constituents . . . 2
  - 1.1.2 Organic constituents . . . 4
  - 1.1.3 Pollution . . . 6
  - 1.1.4 Challenges and opportunities . . . 6
- 1.2 Detection of light . . . 8
  - 1.2.1 Radiometry . . . 9
  - 1.2.2 Spectroradiometry . . . 11
  - 1.2.3 Spectropolarimetry . . . 15
- 1.3 Remote sensing of water . . . 19
  - 1.3.1 Inherent optical properties . . . 20
  - 1.3.2 Reflectance . . . 22
  - 1.3.3 Constituent retrieval and inversion . . . 25
  - 1.3.4 Validation . . . 28
- 1.4 Citizen science . . . 29
- 1.5 This thesis: Accessible remote sensing of water . . . 32

2 Citizen science with colour blindness: A case study on the Forel-Ule scale **35**
- 2.1 Introduction . . . 36
- 2.2 Methods . . . 38
  - 2.2.1 Forel-Ule scale . . . 38
  - 2.2.2 Simulation of colour blindness . . . 38
  - 2.2.3 Colour discrimination . . . 40
- 2.3 Results . . . 40
  - 2.3.1 Colour blindness simulation . . . 40
  - 2.3.2 Colour discrimination . . . 43
  - 2.3.3 Practical consequences . . . 44
- 2.4 Discussion . . . 44
- 2.5 Conclusions & recommendations . . . 46

3 Standardised spectral and radiometric calibration of consumer cameras **47**
- 3.1 Introduction . . . 48
- 3.2 Trends in consumer cameras . . . 50
  - 3.2.1 Smartphones . . . 52
- 3.3 Methods . . . 52
  - 3.3.1 Experimental setup . . . 53
  - 3.3.2 General properties . . . 54
  - 3.3.3 Linearity . . . 54
  - 3.3.4 Bias & read-out noise . . . 55
  - 3.3.5 Dark current . . . 55
  - 3.3.6 ISO speed . . . 55
  - 3.3.7 Gain variations . . . 55

3.3.8 Flat-field correction . . . 56
3.3.9 Spectral response . . . 57
3.3.10 Relative radiometric calibration . . . 58
3.4 Results . . . 59
3.4.1 General properties . . . 59
3.4.2 Linearity . . . 60
3.4.3 Bias & read noise . . . 62
3.4.4 Dark current . . . 63
3.4.5 ISO speed . . . 64
3.4.6 Gain . . . 64
3.4.7 Flat-field correction . . . 66
3.4.8 Spectral response . . . 68
3.4.9 SPECTACLE database . . . 70
3.5 Discussion & conclusions . . . 70

4 Accuracy and reproducibility of above-water radiometry with calibrated smartphone cameras using RAW data 75
4.1 Introduction . . . 76
4.2 Methods . . . 77
4.2.1 Data acquisition . . . 78
4.2.2 Reference data processing . . . 79
4.2.3 Smartphone data processing . . . 80
4.2.4 Colour . . . 82
4.2.5 Replicate analysis . . . 83
4.2.6 Match-up analysis . . . 83
4.3 Results . . . 85
4.3.1 Replicate analysis . . . 85
4.3.2 Smartphone comparison . . . 86
4.3.3 Smartphone vs. reference comparison . . . 89
4.3.4 JPEG data . . . 93
4.4 Discussion . . . 94
4.4.1 Uncertainty . . . 94
4.4.2 Reproducibility . . . 95
4.4.3 Accuracy . . . 96
4.4.4 Recommendations . . . 98
4.4.5 Outlook . . . 99
4.5 Conclusions . . . 100
4.A Uncertainty propagation . . . 100
4.A.1 Covariance and correlation . . . 100
4.A.2 Averaging the G and G_2 channels . . . 102
4.A.3 Remote sensing reflectance . . . 102
4.A.4 Band ratios . . . 103
4.A.5 Chromaticity and hue angle . . . 104
4.B Smartphone RGB to XYZ transformation . . . 105

5 Biases from incorrect reflectance convolution **109**
5.1 Introduction . . . 110
5.2 Theoretical background . . . 111
5.2.1 Reflectance . . . 111
5.2.2 Spectral convolution . . . 112
5.2.3 Reflectance convolution . . . 112
5.2.4 General rule . . . 114
5.3 Methods . . . 114
5.3.1 Radiometric data . . . 114
5.3.2 Spectral convolution . . . 115
5.3.3 Spectral response functions . . . 116
5.3.4 Retrieval algorithm propagation . . . 117
5.4 Results . . . 117
5.4.1 Simulated instruments . . . 117
5.4.2 Real instruments . . . 118
5.4.3 Retrieval algorithms . . . 119
5.5 Discussion & conclusions . . . 119
5.A Validity of spectral convolution . . . 122
5.B Radiometric data . . . 123

6 A universal smartphone add-on for portable spectroscopy and polarimetry: iSPEX 2 **127**
6.1 Introduction . . . 128
6.2 Working principle . . . 129
6.2.1 Definitions . . . 129
6.2.2 SPEX polarisation modulation optics . . . 130
6.3 Add-on design . . . 131
6.3.1 Optical tube . . . 131
6.3.2 Smartphone clip . . . 134
6.3.3 Smartphone backplate . . . 135
6.4 Production . . . 136
6.4.1 Multi-order retarder foils . . . 136
6.4.2 Injection moulding . . . 136
6.4.3 3D printing . . . 136
6.4.4 Assembly . . . 137
6.5 Data acquisition and processing . . . 138
6.5.1 Smartphone app . . . 138
6.5.2 Data pre-processing . . . 139
6.5.3 Wavelength calibration . . . 140
6.5.4 Spectral response calibration . . . 141
6.5.5 Polarimetric demodulation . . . 141
6.6 Future perspective . . . 143
6.6.1 Calibration & validation . . . 143
6.6.2 Scientific applications . . . 145
6.6.3 Future opportunities . . . 145

7 General discussion and future outlook **147**
7.1 Uncertainty in remote sensing of water 148
7.2 Accessibility of water research 150
7.3 Spectropolarimetry of floating debris 152
7.4 Conclusions and recommendations 159

Bibliography **161**

English summary **185**

Nederlandstalige samenvatting **189**

List of publications **193**

Curriculum vitae **195**

Acknowledgements **197**

1 | General introduction

Water is all around us. We use it to drink, wash, play, fish, sail, and much more. Natural waters, like streams, rivers, lakes, seas, and oceans, are full of life and interesting chemistry. Because of its importance, we need to study water intensively, measuring the various compounds and life forms that inhabit it. This way, we can better understand the world around us, our impact on it, and its impact on us.

Light is a powerful tool for studying water. Humans and other animals can intuitively infer some properties of water from its colour and clarity. Using many different wavelengths and other properties of light provides a wide spectrum of information. In remote sensing, we aim to determine the chemical and biological makeup of water by measuring how it reflects light. These measurements can be done on a global scale over long periods of time, without needing to physically sample the water. Automated remote sensing instruments can also provide data on more local scales quickly and with consistent quality.

Because water is so important to everyone, participation in water research should be widely accessible. Citizen science involves non-professionals in taking measurements, interpreting results, and thinking of new research. Citizens provide many data and local knowledge that, in collaboration with professional researchers, can lead to new possibilities and shared insights. At the same time, the citizen scientists learn more about their environment by conducting their own research and they are socially and politically empowered as stakeholders.

This introduction provides a general overview of the field. Section 1.1 introduces the various constituents of global surface waters. In Section 1.2, we look at light and how it is measured. Section 1.3 combines the two by discussing remote sensing of water. Section 1.4 is about citizen science in water research. Each section discusses the current state of research, challenges, and future opportunities. Finally, Section 1.5 provides an overview of the main chapters in this thesis, which focus on various aspects of accessibility and uncertainty in remote sensing.

Accessibility here means the degree to which people can create, use, and interpret data. In Chapter 2, we investigate and improve the ability of people with a disability, namely colour blindness, to participate in water research by performing colour measurements. In Chapters 3 and 4, we develop and validate a method for using consumer cameras, especially on smartphones, to perform professional-grade measurements at a low cost. In Chapter 6, we present iSPEX 2, a new smartphone add-on for remote sensing measurements. The ability to use consumer cameras opens up remote sensing to people without expensive professional instruments, for example people in low-income settings and in citizen science.

Uncertainty here refers to the spread in measured values caused by random effects and to errors caused by known or unknown systematic effects. To use data to their fullest potential, it is crucial to determine, understand, and minimise the associated uncertainty. In Chapter 2, we investigate the uncertainty in water colour measurements by citizen scientists due to colour blindness. In Chapters 3 and 4, we determine and minimise the uncertainty in measurements from consumer cameras by characterising and calibrating their response. Lastly, in Chapter 5, we determine the impact of a common mathematical error in comparing data from different instruments.

1.1 Water

Water is one of the most important resources on Earth and it plays a central part in all aspects of our lives [1, 2]. Clean drinking water is crucial to human health and wellbeing, and many cultures and religions attribute cleansing properties to water itself. Recreation, such as water sports, provides further social value. Economic value lies in the use of water in agriculture, energy generation, fishery, industry, and infrastructure. Of course, humans are not alone on Earth, and water is equally important to all other life forms. In fact, the detection of liquid water is seen as an essential step in the search for extraterrestrial life [3]. Given the many important roles water plays in our lives, it is imperative that we study its biological and chemical constituents and their interactions well.

Social and political discussions on water often focus on the concept of *water quality* [4–8], but this term is misleading. The meaning of water quality is inherently subjective and varies by subject and by stakeholder. For example, most people naturally judge the quality of water for recreation by its aesthetics and think of clear, blue waters as having a high water quality [9]. However, such blue waters are often lacking in nutrients and many algae would consider them low quality [10], if algae could speak. Discussing natural waters in terms of their *quality* thus introduces a human-centric bias and narrows our field of view. It is better to consider nature from a broader perspective that includes humans as part of a greater whole [11]. Even in the United Nations Sustainable Development Goals (SDGs), biodiversity and ecosystem health are primarily considered means towards human-centric ends such as economic growth, rather than goals unto themselves [12].

Scientifically, we are interested in the many processes, substances, and organisms in surface waters, together referred to as *biogeochemistry*. Studying the properties of a water body provides an understanding of its ecosystem and physics. This knowledge is valuable in itself, and further scientific and social value is derived by, for example, improving our understanding of the role of the oceans in climate change adaptation [2] and detecting events that affect human and animal health, such as harmful algal blooms [13].

Environmental regulations typically consist of chemical measures, such as maximum concentrations for pollutants, and ecological indicators, such as targets for population size, species richness, biodiversity indices, and population viability. Monitoring of these variables is performed by all levels of governmental and private parties through extensive sampling and laboratory analysis. Within the water management community, there is a desire to shift the focus of monitoring programmes from individual concentrations to a full understanding of water bodies through systems analysis [7]. Biodiversity is seen as an important proxy for ecosystem health and as a means to improve the resilience of ecosystems against change [4,5]. High biodiversity is also a goal in itself, as it provides an inherent existence value [1].

This thesis focuses on the characterisation of surface waters through optical observations from above the surface. What follows in Sections 1.1.1–1.1.3 is a brief overview of surface water biogeochemistry, separated into inorganic constituents, organic constituents, and pollution. Lastly, Section 1.1.4 discusses challenges relating to the observation of water properties through traditional measurement methods.

1.1.1 Inorganic constituents

Oceans and inland waters contain a wide variety of dissolved and particulate inorganic material. The distinction between dissolved and particulate is typically drawn at a particle size

of 0.2 µm or 0.7 µm. While each substance contributes to the optical and biogeochemical properties of a water body in its on way, most fit into one of several major categories, which are discussed below [14].

Important quantities relating to dissolved inorganic material include salinity, pH, and concentrations of oxygen and trace elements [2, 5]. Salinity is the concentration of dissolved ions and is commonly measured in situ by lowering a conductivity-temperature-depth (CTD) probe into the water to obtain a vertical profile [15]. Seawater has a typical salinity of ~35 g kg^{-1}, consisting primarily of sodium chloride (Na^+ and Cl^-) [16], and a slightly alkaline pH of ~8.1 [17]. Most inland waters are fresh, meaning the salinity is near-zero, with a typical pH of 6.0–8.5 [18], influenced by surrounding land cover, evaporation rates, and other environmental factors [15, 19]. Increasing absorption of atmospheric CO_2 is decreasing ocean pH, a process called *acidification* [20], causing decreased growth rates of organisms with calcite shells like corals and coccolithophorid algae [20, 21]. Some dissolved trace chemicals play a role in biological processes, such as iron ions in photosynthesis and cell structure [21, 22], while others like mercury are toxic [23]. Trace chemical concentrations are traditionally determined from field samples, requiring extensive sample preparation and expensive laboratory equipment and causing a delay between sampling and results [23]. Lastly, dissolved oxygen (O_2) concentrations range from near-zero to $\gtrsim$360 µmol kg^{-1} [24]. Animals and other aerobic organisms require oxygen to function, and insufficient dissolved oxygen levels lead to mass death. Large algal blooms resulting from a sudden influx of nutrients, such as iron from a volcanic eruption, can decrease oxygen levels below habitable conditions and kill millions of animals within days [2, 13]. Dissolved oxygen is measured using in-situ optical and electrochemical sensors or through laboratory analysis of samples [25, 26].

Particulate inorganic material consists of suspended sediments and minerals [14, 27], ranging in size from several µm to >10 mm [28, 29]. The concentration of particulate inorganic material is described by several quantities with different, partially overlapping definitions, some of which also include particulate organic material (Section 1.1.2). Commonly reported quantities include mineral suspended sediment (MSS), suspended particulate matter (SPM), total suspended matter (TSM), total suspended solids (TSS), detritus, and turbidity [14, 27, 30–34]. The quantity of choice depends on the specifics of the measurement site, application, and available methods and materials. Suspended inorganic matter is supplied by erosion of the shoreline and bottom, in the water body itself or in supplying rivers and tributaries, and by deposition of atmospheric aerosols [28, 33, 35]. Particulate matter that has settled can be resuspended by water currents, particularly in shallow waters with high wind speeds and in tidal areas [27, 29, 33]. The dynamic nature of the processes supplying suspended matter can lead to highly variable concentrations. For example, SPM concentrations ranging from <1–4000 g m^{-3} have been observed in the Wadden Sea [29]. Because many sediments scatter light strongly (Section 1.3.1), variations in SPM induce variations in the amount of photosynthetically available radiation and thus affect phytoplankton and vegetation populations [35].

Suspended matter concentrations are determined in the laboratory by filtering and optionally drying water samples [35, 36]. The variability between replicate measurements is typically 1–5%, with outliers of ≥10% [37]. The particle size distribution can be inferred from the forward scattering of light (Section 1.3.1), for example using a Sequoia LISST instrument [14, 28]. Portable turbidimeters are relatively low-cost and easy to use, and provide in-situ measurements of turbidity with an uncertainty of 1–3% [37]. TSM or SPM can be derived from turbidity with a mean bias as low as 2% and root-mean-square (RMS) error

of 25% using a well-tuned algorithm, but due to spatial and temporal variability in particle properties, these errors can be ≥50% in some sites [38, 39]. In general, variability between replicate in-situ measurements is caused by both measurement uncertainty and inherent variability [40].

1.1.2 Organic constituents

The ocean and other surface waters are host to a wide variety of organisms and organic materials, ranging in size from dissolved proteins to blue whales [14]. While each contributes to the ecosystem in its own way, here we will focus on phytoplankton and coloured dissolved organic matter (CDOM). Both are important biogeochemical components of virtually all surface waters and are commonly studied through visible-light remote sensing [2, 41].

Phytoplankton is a broad category of free-floating microscopic organisms that perform primary production of organic compounds through photosynthesis. The most prominent taxonomical groups are green algae, coccolithophorids, diatoms, dinoflagellates, and cyanobacteria [14]. Species may alternatively be classified by their role in the ecosystem or *functional type* to better understand nutrient cycles. Important functional types include nitrogen fixers, which bind dissolved nitrogen into nutrients; calcifiers, which produce calcium carbonate shells; and silicifiers, which produce structures of silica. Carbonate and silica structures aggregate and sink, exporting nutrients from the surface to deeper layers. Diversity is necessary to provide and distribute all nutrients required in the ecosystem and thus to maintain population levels, biodiversity, and resilience to change [42]. Phytoplankton are responsible for half of the global primary production of organic matter and oxygen, contribute significantly to the global carbon cycle, and form the basis of the marine food web [14,43,44]. This high ecological importance makes phytoplankton biomass and taxonomic as well as functional diversity essential biodiversity and climate variables [43, 45].

Phytoplankton has fast growth and reproduction cycles, so it responds rapidly to changes in environmental conditions and nutrient availability, on a time scale of hours or days. These changes can cause algal blooms, during which one species rapidly increases in biomass and dominates the local ecosystem [13, 46]. For example, number concentrations of *Phaeocystis* algae of up to 100 000 cm^{-3} were observed during the 1994 spring and summer blooms in the Marsdiep, Netherlands, while almost none were found between blooms [46]. Changes in phytoplankton biomass and diversity serve as a proxy for overall changes in the ecosystem and rapid detection of algal blooms is required to mitigate harmful effects such as toxins released into the water [13]. Many inland waters, particularly in the Netherlands, are currently undergoing a strong increase in phytoplankton contents (*eutrophication*) and harmful algal bloom frequency due to pollution with nutrients such as nitrogen and phosphorus. These trends are causing major social and economic problems by decreasing oxygen concentrations and introducing toxins into the water, in turn disrupting the food chain, killing aquatic animals, and preventing the use of surface waters for drinking, irrigation, and recreation [1, 4, 5, 47]. For example, cyanobacterial blooms induced by emission of phosphorus are dangerous to swimmers, since exposure to cyanobacteria can cause skin irritation, gastrointestinal disease, and even death [48].

Chlorophyll-a (*chl-a*) and other chemical pigments are used by phytoplankton for photosynthesis. Different types of phytoplankton use different pigments, based on niches in the absorption spectrum of water [49]. Pigment concentrations are used to quantify phytoplankton biomass and differentiate between species. Chl-a is the most important pigment [42] and

ranges in concentration from 0.01–100 mg m^{-3} in the ocean and most inland waters [50, 51]. Chl-a is so dominant that its concentration alone is enough to describe the optical properties of most of the open ocean (*Case I* waters) since all other constituents largely covary with it [52]. Aside from chl-a, important pigments include other types of chlorophyll, carotenoids, and the cyanobacteria-specific pigments phycocyanin and phycoerythrin [49, 53]. Photosynthesis is performed using visible light with wavelengths of 400–700 nm, also termed photosynthetically active radiation (PAR) [49].

There are several common methods for in-situ detection and characterisation of phytoplankton. High-performance liquid chromatography (HPLC) is used to accurately measure pigment concentrations, and thus infer the biomass of different phytoplankton types, and is the standard for calibration and validation of other methods [42, 50]. A major drawback of HPLC is its high cost and labour intensity. Intercomparisons of HPLC measurements of the same samples by different teams have yielded reproduction uncertainties of 5–40% in chl-a concentration, which were ascribed to differences in HPLC protocol [54]. Many pigments fluoresce (Section 1.3.1) and the intensity of fluorescence at different wavelengths is a proxy for pigment concentrations [52, 53]. Fluorescence measurements can be done at a low cost using a light emitting diode (LED) or laser light source [55, 56] and typically agree well with HPLC [50]. For example, a comparison between 402 simultaneous HPLC and fluorometric chl-a concentrations from five oceanic sites yielded a median absolute difference of 11% [57]. However, the conversion from intensity to pigment concentration is complicated by the distribution of pigments inside the algal cells and sensitivity to fluctuating light levels [42]. Determination of pigment concentration from other inherent optical properties is discussed in Section 1.3.3. Imaging techniques such as microscopy and flow cytometry detect and characterise individual cells and other suspended matter. These methods are labour-, equipment-, and compute-intensive, but this is rapidly improving and machine learning classification algorithms with >80% accuracy (precision and recall) have been demonstrated [43]. Lastly, genomics can accurately distinguish between hundreds of species based on sampled (meta-)genomes [58].

Dissolved organic matter (DOM) is created by biological processes in a water body or its sources, or is introduced into the water body from the bottom or from surrounding land [59]. DOM consists of thousands of different organic compounds. Some of these compounds strongly absorb light and thus change the appearance of the water (Section 1.3.1) – together, these compounds are called *chromophoric* or *coloured dissolved organic matter* (CDOM) [25]. The terms *yellow substance* and *gelbstoff* are also used [60]. Absorption by CDOM gives many peat lakes in the Netherlands, such as the Kagerplassen near Leiden, their characteristic dark brown colour by removing blue–green light from the reflectance spectrum, leaving only yellow–red light (Section 1.3.3). CDOM is a tracer for DOM, but the exact relation depends on the concentrations of specific chemicals and varies by location and over time [14, 25]. Often, one is specifically interested in the amount of carbon dissolved in the water in various compounds, for example when studying the carbon cycle or climate change. In that case, the closely related quantity *dissolved organic carbon* (DOC) is used [25]. The DOC concentration is 0.2–2.0 g m^{-3} in the ocean [61] and ranges from near-zero to >20 g m^{-3} in lakes and rivers [25, 62, 63]. DOM, CDOM, and DOC are typically studied with optical techniques based on fluorescence or absorption [60, 64], discussed further in Section 1.3.1, or with catalytic oxidation systems [25, 65].

1.1.3 Pollution

In addition to their natural constituents, most surface waters contain some degree of pollution by humans. This pollution comes in many forms, from dissolved molecules to metre-scale debris, and from many sources. Faecal matter from humans and other animals is a major threat to drinking water safety in many parts of the world [1, 66], including historically in the Netherlands [67]. Excreta are also a major vector for pollution with pharmaceuticals, which can have severe impacts on aquatic animals [5, 68]. Agricultural runoff can induce eutrophication and algal blooms by adding nutrients to the water (Section 1.1.2), or alternatively decimate ecosystems by unintentional emission of pesticide residue through various routes such as runoff, leaching, drift, and aerosol deposition [69]. These effects cascade through the food web and affect all members of an ecosystem [69]. The forms of pollution described above are detected through chemical monitoring and ecological surveillance data [5, 18, 69]. Within the European Union, the establishment of programmes for water monitoring and surveillance is required by the Water Framework Directive [8].

Solid debris comes in many forms, including wood, food waste, cloth and fabric, metal, glass, and plastics [70]. Plastic pollution is particularly commonly researched and discussed currently because of its close relation to human activities, its broad and persistent distribution across global waters [71–74], and its far-reaching, albeit still debated, effects on individual organisms and ecosystems [5, 73]. Plastic particles range in size from the nano- and micrometre-scale (*nanoplastics* and *microplastics*) to large objects like shopping bags (*macroplastics*). Smaller particles are released through weathering and breakdown of large objects, but also aggregate into larger clusters through the creation of biofilms and integration into organic matter [73]. Microplastic surface densities in the ocean are typically 1 g km^{-2}–1 kg km^{-2}, but can reach up to 100 kg km^{-2} in the Pacific garbage patch [75]. Plastics and other debris are detected through in-situ soil and water samples and photographs from ships and other platforms, but protocols are not sufficiently standardised and different methods often yield different results [76, 77]. An integrated observation system with remote sensors (Section 1.3) is required to monitor pollution with debris on all spatial and temporal scales with high accuracy [75].

1.1.4 Challenges and opportunities

Characterisation of surface water constituents through in-situ sampling faces three main challenges, namely cost of equipment and labour; availability of data on different spatial and temporal scales; and reproducibility between researchers, sites, and methodologies.

As discussed in Sections 1.1.1–1.1.3, many constituents require trained personnel and expensive laboratory equipment to measure. For example, an HPLC measurement of chl-a in a lake requires researchers to sail out on a ship, take and properly store a water sample, bring the sample to a laboratory for filtering and extraction, run it through an HPLC instrument costing several thousand euros, and analyse the results using specialised software. Even with economies of scale, such as measuring multiple parameters per sample and taking many samples per cruise, the cost per HPLC chl-a measurement is on the order of hundreds of euros. Other methods like fluorometric analysis have a lower cost per sample, but come with greater uncertainties. Cost and labour requirements limit the scope of constituent monitoring at large spatial scales and high frequencies, and disproportionately limit the accessibility of data in low-income settings, which often bear the brunt of environmental problems caused by pollution and climate change [1, 11].

In general, in-situ measurements are difficult to scale up spatially and temporally. Global coverage is necessary to understand global processes like climate change, but the funds and personnel required to physically sample every water body regularly are not available. Research into local processes, for instance in individual lakes and other inland waters, is similarly constrained by the lack of scaleability of in-situ measurements. Moreover, while early detection of rapid events such as harmful algal blooms saves money and lives, laboratory analysis cannot provide the required high-frequency measurements and rapid responses [2].

Lastly, standardisation in sampling and analysis protocols between researchers, both locally and internationally, is limited [78]. Chl-a concentrations measured simultaneously by different teams differ by as much as 5–40% (Section 1.1.2). Due to differences in measurement protocol and reporting of results, it is nearly impossible to quantitatively compare microplastic loads across the world [77]. This lack of reproducibility between data from different periods or areas creates spurious trends and is difficult to detect or correct post hoc, as samples cannot be stored and re-used indefinitely. The issue is further exacerbated by scant reporting of uncertainty.

Three important opportunities towards meeting these challenges are automation, remote sensing, and citizen science. Automation enables large-scale, high-frequency measurements with reproducible results. Remote sensing reduces the cost of sample processing, can be performed on all spatial and temporal scales, and is highly reproducible. Citizen science can significantly decrease the cost per measurement and increase both the scale and quality of measurements by including more people in the scientific process and incorporating local knowledge. Each of the three opportunities is discussed in more detail below.

In recent years, autonomous platforms have increasingly been used to provide high-quality, reproducible data at a relatively fine spatial and temporal resolution. For example, the Argo array consists of nearly 4 000 robotic floats, which in the last 20 years have provided over 2 million CTD profiles covering most of the ocean [79]. Many Argo floats are now also equipped with bio-optical sensors to measure chl-a and other biogeochemical parameters [80]. Similarly, underwater gliders can autonomously perform horizontal and vertical transects for up to six months at a time [81]. Gliders and floats have been used together to measure variations in density, chl-a concentration, and primary production with sub-kilometre spatial resolution [82]. These autonomous platforms provide high-quality data, but they are limited by high purchase and use costs and are less suitable for shallow and inland waters. In-line flow-through systems installed on research vessels and ships of opportunity use water pumped into the vessel to automatically perform bio-optical measurements with high frequency and reliability. By combining professional instruments such as the aforementioned LISST, Sea-Bird AC-s transmissometer, and Sea-Bird ECO series scattering meters with an automated pumping system and data logger, the data volume can be massively increased and the cost per measurement decreased [83, 84]. While automation is not discussed in detail in this thesis, some of our findings can be applied or extrapolated to automated sensors, as discussed in the individual chapters.

Remote sensing is a measuring technique that does not require physically sampling a water body, instead relying on reflected light to carry information. Satellite remote sensing of water started in earnest in the 1970s and has evolved continually since then [85], with the forthcoming Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission [86] as the most prominent new development (Sections 1.2.2, 1.2.3, and 1.3.3). Operational remote sensing platforms include satellites, aeroplanes and uncrewed aerial vehicles (UAVs or drones), ships, and terrestrial instruments [31, 87]. Remote sensing provides data with wide spatial

coverage over long periods, and with sufficient resolution, precision, and accuracy for many practical applications [2, 88]. For example, satellite data are used to rapidly detect and warn against potentially harmful algal blooms anywhere in the global ocean and many inland waters, with a speed and coverage that cannot be achieved with traditional sampling-based techniques [13]. Terrestrial and shipborne remote sensing provides high-frequency data with fast response times and high reproducibility, since data are recorded electronically and can thus be processed automatically and near-instantly [89, 90]. The uncertainty in measurements of reflected light is typically $\lesssim$5% for a single instrument [88], as is the reproducibility between different instruments [91, 92]. The physics of measuring light are discussed in Section 1.2 and the implementation and application of remote sensing in Section 1.3.

Citizen science, the involvement of non-professionals in the scientific process, provides further opportunities to reduce cost and improve scalability by increasing the accessibility of water research. Citizens are important stakeholders in water management and thus offer valuable *bottom-up* perspectives and insights to research aims and methods [93]. Through *crowd-sourcing*, citizens massively increase data collection capabilities [94], on large spatial scales and over long periods as well as on the local level and with fast response times. There are various degrees of citizen involvement, ranging from data collection for a *top-down* research project to co-creation of an entirely new project [95]. Citizen science is discussed in more detail in Section 1.4.

1.2 Detection of light

To understand remote sensing, it is useful to first understand the physics of light itself. Fundamentally, light consists of electromagnetic (EM) waves, quantised into discrete packets of energy called *photons*. EM waves are generated by various processes involving the acceleration of electrically charged particles and propagate outward at $c = 299\,792\,458$ m s^{-1}. For EM waves in free space, the electric and magnetic field vectors oscillate orthogonally to the direction of propagation and to each other. Defining properties of individual photons include wavelength, spin angular momentum or polarisation, and orbital angular momentum. Remote sensing instruments gather many photons within one exposure, so this thesis focuses on the bulk properties of light fields consisting of many photons. For a comprehensive description of remote sensing from basic physical principles, the reader is referred to [14, 96].

The fundamental observable quantity of remote sensing is spectral radiance L, the radiant power emitted or received by a surface, as defined in Equation (1.1) and illustrated in Figure 1.1. $\vec{L}$ is the radiance at a position $\vec{x}$ and time t, from a direction $\hat{\xi}$, at a nominal wavelength λ; ΔQ is the energy emitted or received by a surface ΔA over a time Δt, within a wavelength range $\Delta\lambda$ and a solid angle $\Delta\Omega$ [14]. Radiance is written as a vector $\vec{L}$ when polarisation is considered or as a scalar L when polarisation is neglected. Radiance has units of W m^{-2} nm^{-1} sr^{-1}.

$$\vec{L}(\vec{x}, t, \lambda, \hat{\xi}) = \frac{\Delta Q}{\Delta A\, \Delta t\, \Delta\lambda\, \Delta\Omega} \tag{1.1}$$

Another important quantity is planar irradiance E, which is the total radiance integrated over a hemisphere, as shown in Equation (1.2). For example, the downwelling irradiance E_d represents the total radiant power received by a surface, such as a patch of sea, from all upwards directions [97]. Irradiance has units of W m^{-2} nm^{-1}.

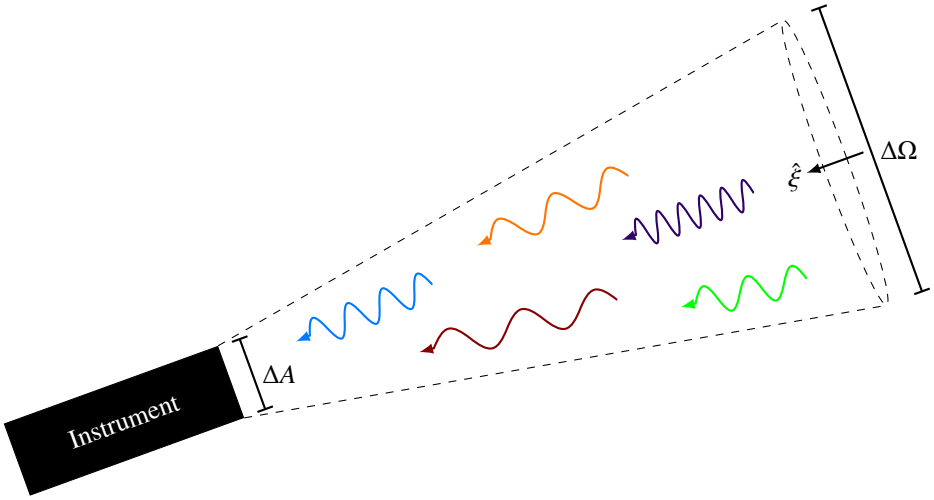


Figure 1.1: Schematic representation of a radiometric instrument detecting incoming light, represented by several photons with different wavelengths. ΔA is the collecting area of the instrument, $\hat{\xi}$ the observing direction, and $\Delta\Omega$ the solid angle subtended by the field of view, indicated by dashed lines. Note that photons travel in straight lines, as indicated by the arrowheads; the waves indicate the direction of the oscillating electric field.

$$E(\vec{x}, t, \lambda) = \frac{\Delta Q}{\Delta A \, \Delta t \, \Delta \lambda} \tag{1.2}$$

The following subsections provide a brief description of the measurement of radiance in general (1.2.1), as a function of wavelength (1.2.2), and considering polarisation (1.2.3). The use of radiometry in remote sensing is discussed in Section 1.3.

1.2.1 Radiometry

Currently, most scientific instruments and consumer cameras are based on charge-coupled device (CCD) or complementary metal-oxide-semiconductor (CMOS) sensors [14, 98, 99]. CCD and CMOS sensors use the photo-electric effect to detect incoming photons that strike a layer of semiconductor and, if they have enough energy, free one or more electrons [100], which are called *photo-electrons*. The number of photo-electrons generated per photon is the quantum efficiency η, which is highly wavelength-dependent [101]. CCD and CMOS sensors typically contain millions of µm-sized pixels for high-resolution imaging and spectrometry. Other sensor types of historical relevance include bolometers [99] and photographic plates and film [102, 103]. This thesis mostly deals with CMOS sensors, which are by far the most common in consumer cameras [98].

The collected photo-electrons are converted to a voltage, which is then amplified and digitised [101]. CMOS sensors include amplification and digitisation electronics on individual pixels, enabling pixel-level control. For example, the amplifier gain, the conversion factor from photo-electrons to volts, can be varied between pixels to compensate for differences in sensitivity (Section 3.4.6). Many CMOS amplifiers have built-in compensation for non-linearity and other sources of uncertainty [104], although additional corrections are sometimes necessary [105]. An analogue-to-digital converter (ADC) is used to digitise the ampli-

fied voltage into analogue-digital units (ADU). The ADC bit depth determines the number of possible digital values, which was as low as 128 (7-bit) for early satellite instruments [96], but is now more typically in the range of 4 096–65 536 possible values (12–16-bit).

The raw data in ADU are converted to radiant energy ΔQ through a radiometric calibration [106]. First, the optical properties of the instrument, including the sensor and any fore-optics, are characterised (Chapter 3). Next, an absolute radiometric calibration is performed, in which the sensor observes a light source with known characteristics to determine the conversion factor from ADU to physical units [107, 108]. Ideally, the materials used in the calibration and characterisation are traceable to standards such as those maintained by the National Institute of Standards and Technology (NIST, USA) or the National Physical Laboratory (NPL, UK) [106, 108].

Professional field-going and satellite instruments are calibrated and characterised by their manufacturers prior to delivery. Field-going instruments are certified to sub-percent radiometric accuracy at most wavelengths and the reproducibility between different instruments under realistic conditions is normally 1–5% [92, 109, 110], as discussed further in Chapter 4. Satellite instruments require sub-percent accuracy and uncertainty, or equivalently a signal-to-noise ratio (SNR) $\geq$100, in some cases $\geq$1000, to deliver percent-level accuracy and uncertainty on ground-level radiance, reflectance, and derived parameters [45, 111, 112]. Calibration is repeated regularly during sensor lifetimes to account for changes in performance. The sensitivity of field-going spectroradiometers typically drifts by $\leq$5% per year, often more strongly in their first years and $\leq$1% later on, although individual instruments may drift by as much as 40% year-to-year [108]. Changes in sensitivity tend to be wavelength-dependent. For satellite instruments, regular vicarious calibration using well-known targets such as the MOBY buoy site and the Moon is used to maintain the desired accuracy and uncertainty of $\sim$5% in reflectance [113]. For consumer cameras, the calibration process is left to the user (Chapter 3).

Common sources of uncertainty include stray light; non-zero sensitivity to wavelengths outside the desired spectral range; sensitivity to polarisation for unpolarised radiance measurements; changes in response due to fluctuations in temperature, humidity, and other environmental factors; degradation of transmission due to fouling by dust, oils, biofilms, and sea spray; shot noise due to the quantum nature of photons; and uncertainties in the calibration materials themselves [108, 110, 114–117]. The relative contribution of each component is highly variable between different instruments and varies by wavelength and over time.

Observations are done over a limited exposure time Δt to obtain the radiant power $\frac{\Delta Q}{\Delta t}$. The exposure time is ideally as long as possible before the sensor saturates. Individual exposures are realised by physically blocking the sensor from incoming light with a mechanical shutter or by electronically powering the sensor on and off [99]. An advantage of mechanical shutters over electronic shutters is the ability to take fully dark exposures for calibration purposes; a disadvantage is the reliance on moving parts, which eventually malfunction or break. In either case, it is important to characterise the accuracy and reproducibility of the shutter to ensure accurate radiance measurements [118, 119].

A well-collimated radiometer detects light from a viewing direction $\hat{\xi} = (\theta, \phi)$. The simplest well-collimated instrument is the Gershun tube, a long tube with multiple baffles to select only light from within a narrow field of view around the desired $\hat{\xi}$ [14]. The instrument can be moved or rotated to observe in multiple directions sequentially, which is called *pushbroom* or *whiskbroom* sensing. Many instruments use lenses to focus incoming light onto a sensor array, in which each pixel corresponds to a viewing direction. Field radiome-

ters are typically aligned on a mount which is fixed in, or swivels to, the desired viewing direction [120]. Some systems must be aligned by eye [121, 122], others are fully automated [123, 124]. The required pointing accuracy and stability depend on the application, but are typically 0.5°–5° for instruments mounted on stationary observing platforms [124, 125].

The radiometer fore-optics also determine its collecting area ΔA and field of view (FOV) $\Delta\Omega$. The collecting area is ideally as large as possible to collect more light and thus improve the SNR, but is usually limited by physical restrictions. ΔA is on the order of mm^2–cm^2 for field radiometers [126] and consumer cameras (Section 3.4.1), and up to hundreds of cm^2 for satellite instruments [14]. The desired value of $\Delta\Omega$ depends on the application, particularly on the degree of isotropy in the target. Most spot radiometers for above-water radiometry have 1°–7° FOVs [105, 108, 126], equivalent to a solid angle $\Delta\Omega$ of 10^{-4}–10^{-2} sr. For imaging systems, it is important to geometrically calibrate the FOV for individual pixels, which may vary due to distortions induced by the fore-optics [99, 107].

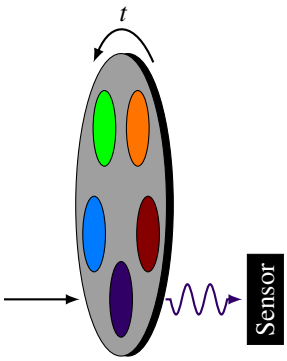
1.2.2 Spectroradiometry

Spectroradiometry is the process of measuring the distribution (*spectrum*) of photons or radiance across different wavelengths λ. Absorption and scattering alter the spectrum of light as it travels through a medium (Section 1.3.1). The full spectrum of light includes gamma rays, x-rays, ultraviolet (UV), visible light, infrared (IR), microwave, and radio waves. The same physics apply to light at all wavelengths, although with differences in sensor design and performance, and all are used for remote sensing. This thesis focuses on visible light, with λ between 390–700 nm. This spectral range covers the sensitivity of the human eye and most consumer cameras (Chapters 2 and 3), and contains all photosynthetically active radiation (PAR) [49].

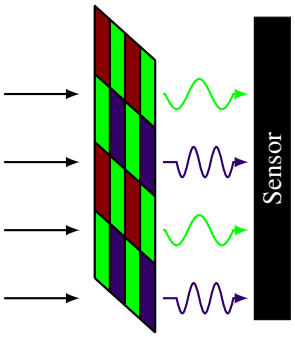
The wavelength of a photon is inversely proportional to its frequency and energy, which is $\sim 10^{-19}$ J for visible-light photons. For illustration, the typical visible-range clear-sky solar irradiance of ~400 W m^{-2} corresponds to $\sim 10^{21}$ photons s^{-1} m^{-2} [14]. In a medium, the speed and wavelength of light are decreased by a factor n, the *refractive index*. Typical values of n are $1.000\,293 \approx 1$ in air and 1.33–1.35 in water, depending on its constituents; n itself also varies by wavelength [14, 127].

CCD and CMOS sensors detect photons with different wavelengths, but cannot distinguish between wavelengths. Therefore, performing spectroradiometry requires additional optical elements. The most common techniques are moveable or tuneable filters, patterned filters, and dispersion with a grating or prism (Figure 1.2). Each method is briefly described below; for a more general overview, the reader is referred to [128, 129].

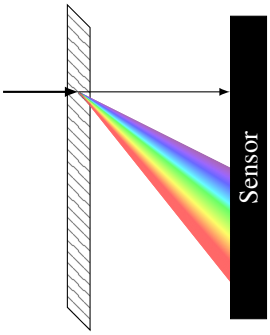
Before proceeding, it is important to define some terms. *Spectral filters* are optical elements that preferentially transmit certain wavelengths. A spectral filter or *band* B has a *spectral response function* (SRF), $S_B(\lambda)$, describing its transmittance at each wavelength λ. The SRF of an instrument is the product of its filter SRFs, the transmission efficiency of other optical elements, and the sensor quantum efficiency (Section 1.2.1). SRFs are often defined in terms of their central wavelength λ_c and bandwidth or *spectral resolution* $\Delta\lambda$. Typically, λ_c is taken to be the wavelength with the highest transmittance or the SRF-weighted average wavelength, and $\Delta\lambda$ is the SRF full width at half maximum (FWHM). A *multispectral* instrument measures in discrete bands, while a *hyperspectral* instrument covers a continuous wavelength range (Figure 1.3). Following the Shannon-Nyquist theorem, the *spectral sampling*, which is the difference between successive wavelengths in hyperspectral data, needs to



(a) Spectral filters in a rotating filter wheel. Incoming broad-spectrum light (black arrow) passes through one filter at a time, transmitting only specific wavelengths (purple arrow) to the sensor.



(b) Spectral filters in a Bayer pattern. Incoming broad-spectrum light (black arrows) passes through one filter for each pixel, transmitting the corresponding wavelengths (coloured arrows) to the sensor. For clarity, only four light rays are shown.



(c) Dispersive spectrometer setup. Incoming broad-spectrum light (black arrow) is dispersed by a grating or prism and transmitted to the sensor (black arrow, spectrum). For clarity, only a single incoming light ray is shown here.

Figure 1.2: Schematic representation of three techniques for spectroradiometry.

be half the spectral resolution or finer. The terms spectral resolution and spectral sampling are often confused in the literature, but they are fundamentally different quantities. A commonly used criterion for the *hyperspectral* label is $\Delta\lambda \lesssim 10$ nm in >20 continuous bands, but the distinction with multispectral is ultimately subjective [14, 87]. Methods for comparing and combining the two are discussed in Chapter 5.

The simplest method for spectroradiometry is a rotating wheel with individual filters corresponding to spectral bands of interest, mounted in front of a sensor (Figure 1.2a). A major advantage of this method is the use of the same sensor for each measurement, removing the need for many calibration steps when performing relative measurements such as reflectance band ratios (Section 1.3.2). A major disadvantage is the delay between measurements in successive bands, which makes this method unsuitable for targets that move or

(a) Multispectral data cover a few discrete spectral bands.

(b) Hyperspectral data cover a continuous wavelength range.

Figure 1.3: Multispectral and hyperspectral data.

vary over time scales comparable to the rotation period of the filter wheel, and sensitive to vibrations in the instrument [130]. Rotating filter wheels are used in satellite instruments including GOCI [131] and POLDER 1–3 [132] and in many ground-based spectroradiometers [117,119,133,134]. A variation on the technique, commonly used in satellite instruments, is using fixed filters on different pixels in one sensor, which observe the target sequentially (pushbroom) due to the orbital movement of the satellite [135]. Finally, some instruments instead employ a single filter with a tuneable SRF, such as a Fabry-Pérot etalon [136] or variable liquid crystal retarder between polarising filters [137].

Virtually all consumer cameras employ a patterned Bayer filter [138] with red-green-blue (RGB) bands (Section 3.2). These have wide SRFs (Section 3.4.8) and are arranged in a checkerboard pattern on top of the sensor, with two green filters for every red or blue one (Figure 1.2b). Each pixel observes in a single band and a *demosaicking* algorithm is used to interpolate the results and form a combined multispectral image [139, 140]. The main advantage of patterned filters is the ability to take snapshot multispectral images, with no delay between filters or between parts of the image. Disadvantages include the need to calibrate the response of each individual pixel before band ratios can be calculated and uncertainties introduced by the demosaicking algorithm, especially for targets with sharp features, which can create spurious edge effects when comparing the different channels. New variations on the method mitigate these disadvantages through additional colour filters and different filter arrangements [139, 141].

The third major technique is spectral dispersion through a prism or grating (Figure 1.2c). Many variations exist, including transmission and reflection gratings and integrated lens-grating or grating-prism systems [128, 142]. Because the entire spectrum is dispersed across the sensor, data can be obtained hyperspectrally rather than being limited to hand-picked filters as in the previous two methods. The dispersion takes up one dimension of the sensor array, requiring incoming light to first pass through a point aperture, one-dimensional slit, or integral-field unit to reduce the spatial dimensionality. Imaging is performed through whisk- or pushbroom sensing. The properties of the dispersive element and the width of the entrance pupil or slit determine the spectral resolution, while the size and number of pixels in the sensor determine the spectral sampling (Section 6.5.3). Spectral dispersion is highly suitable for hyperspectral spot radiometers and satellite instruments [86, 116, 126]. A dispersive element can also be mounted onto an existing camera [94, 143]; the design and performance of such an add-on, iSPEX 2, are discussed in Chapter 6.

The optimal design choices for a spectroradiometer depend on the application and technical constraints. On paper, measuring at more wavelengths λ and with a narrower bandwidth $\Delta\lambda$ provides more information about the target and is thus desirable. However, the wavelength range is limited by technical factors such as the quantum efficiency and size of the sensor. Furthermore, for a given aperture and exposure time, narrower bandwidths lead to a lower SNR, decreasing the effective information content of a signal [144]. Finally, the increase in information content with a finer sampling or wider range of wavelengths may be smaller than expected due to correlations in the spectrum at different wavelengths [145, 146]. Thus, instruments are best designed with their desired application in mind. For example, AERONET-OC was built to validate radiometric data from pre-existing multispectral satellite instruments, meaning only a small number of pre-determined spectral bands were needed and a relatively simple rotating filter wheel design could be used [117]. When studying a substance with distinct, well-known spectral features, only a handful of spectral bands covering those features and any applicable calibrations are necessary.

The most commonly used instruments for above-water spectroradiometry are the TriOS RAMSES and Sea-Bird HyperOCR [92, 114]. They cover wavelengths of 320–950 nm and 350–800 nm, respectively, both with a bandwidth of ∼10 nm and spectral sampling of 3.3 nm [147, 148]. Also commonly used are the Malvern Panalytical ASD FieldSpec, which covers 350–2500 nm with a 3 nm bandwidth [149], and the Water Insight WISP-3, which covers 380–800 nm with a 5 nm bandwidth [126]. Each of these instruments covers the entire visible range, which includes the spectral features of important constituents like chl-a, CDOM, and suspended particles (Section 1.3.1). The additional near-infrared (NIR) coverage of the RAMSES and ASD enables cross-calibration between NIR wavelengths, where no signal is expected, and visible wavelengths [90]. The typical accuracy and reproducibility of these instruments is on the percent level (Section 1.2.1). The WISP-3 suffers from normalisation issues when compared to other instruments [126], as discussed in detail in Chapter 4. Other field-going spectroradiometers of note include the DALEC [109], PANTHYR [124], WISPstation [89], and the SeaPRISM instrument used for AERONET-OC [117].

At the core of satellite instrumentation design is the trade-off between extending capabilities and maintaining a sufficient SNR. For example, the Operational Land Imager (OLI) on Landsat 8 and 9 has a fine spatial sampling of 30 m per pixel, but is consequently limited to relatively wide spectral bands, with bandwidths of ≥40 nm for the blue, green, red, and NIR bands [135]. Conversely, the Ocean and Land Colour Instrument (OLCI) on the Sentinel-3 satellites features 21 bands from 400–1020 nm with a narrow 10 nm bandwidth, and a much coarser spatial sampling of 300 m per pixel, at a sensitivity of $\sim 10^{-3}$ W m^{-2} sr^{-1} nm^{-1} [150]. The MultiSpectral Imager (MSI) on Sentinel-2 combines both approaches, having 9 bands in the visible-NIR range with bandwidths of 15–115 nm and spatial resolutions of 10–60 m [151]. Other satellite instruments of note include the Sea-Viewing Wide Field-of-View Sensor (SeaWiFS), Moderate Resolution Imaging Spectroradiometer (MODIS), Visible Infrared Imaging Radiometer Suite (VIIRS), and Geostationary Ocean Color Imager (GOCI) [131, 152]. Successive generations of satellite instruments have been built with continuity in mind, for example by maintaining certain spectral bands [152]. Finally, 2024 will see the launch of the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) satellite, which primarily consists of the Ocean Color Instrument (OCI). OCI is a hyperspectral instrument that provides continuous coverage of wavelengths of 340–890 nm with 5 nm bandwidth [86]. PACE will also carry the SPEXone and HARP-2 spectropolarimeters (Section 1.2.3). For a broader overview of satellite instruments used in remote sensing of water, the reader is referred to [2, 41, 131, 153].

The human eye is essentially a patterned multispectral radiometer and can thus be used as a scientific instrument. The retina contains two types of photosensitive cells, namely rods and cones. Rods are extremely sensitive to light at all visible wavelengths and are highly useful in low-light conditions, but cannot distinguish between wavelengths. The cones contain groups of proteins sensitive to either long (L), medium (M), or short (S) wavelengths (Figure 1.4). The LMS cones are spread around the retina and by comparing their measurements, the brain estimates the spectrum of the incoming radiance. We perceive this as colour [101, 154]. Many animals' eyes function similarly but with different numbers of cones and different SRFs. A number of scientific techniques use the human eye as a sensor to perceive brightness, colour, distance, and other quantities (Section 1.4). A few percent of the human population have some degree of *colour deficiency* or *colour blindness*, meaning their perception of colour is different to that of the average human [155]. Remote sensing of water by citizen scientists with colour blindness is explored in Chapter 2.

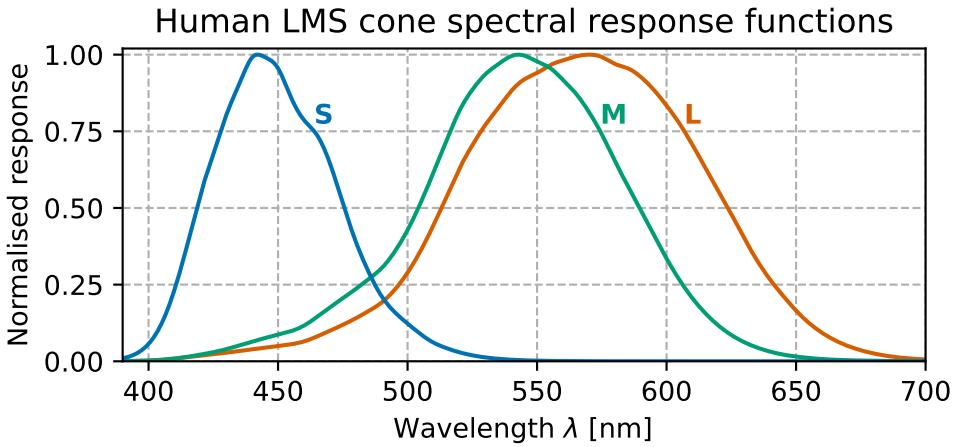


Figure 1.4: Normalised spectral response functions of the long-, medium-, and short-wavelength (LMS) cones in the average human eye. Data obtained from http://www.cvrl.org/.

1.2.3 Spectropolarimetry

Spectropolarimetry is the process of measuring radiance and polarisation at different wavelengths. Polarisation is a wave property of light, relating to the direction of oscillation of the electric[1] field vector $\vec{E}$ as light propagates. For light propagating in the z direction, $\vec{E}$ has components E_x, E_y in the (x, y) plane orthogonal to z.

Light exhibits various polarisation states. These are most easily understood by considering coherent light, for instance from a laser, which behaves like a single wave. As the electric field oscillates, $\vec{E}$ traces a shape in the (x, y) plane, determined by the phase delay δ between the oscillations in E_x and E_y. For example, if E_x and E_y oscillate in phase ($\delta = 0°$), $\vec{E}$ traces a straight line in the (x, y) plane and the light is *linearly polarised*. If E_x and E_y oscillate exactly out of phase ($\delta = 180°$), the light is linearly polarised at the supplementary angle. If E_x and E_y have the same amplitude and oscillate with a delay of $\delta = 90°$, $\vec{E}$ traces a circle in the (x, y) plane and the light is *circularly polarised*. In the general case, $\vec{E}$ traces an ellipse and the light is *elliptically polarised*. The polarisation state of coherent light is described through Jones calculus, combining E_x and E_y into a vector, the Jones vector, and representing optical elements as linear transformations of this vector [116, 156].

Natural light is an incoherent sum of many individual light waves, each with its own polarisation state, resulting in a bulk polarisation state of the light field. When the polarisation states of the constituent EM waves are uniformly randomly distributed, the light is *unpolarised*. If the distribution instead has a preferential direction or handedness, the light is *partially polarised*. Finally, despite being incoherent, the EM waves in *fully polarised* light all have the same polarisation state. For a further physical description of polarisation in natural light, the reader is referred to [157].

This thesis deals with linear polarisation of incoherent light in remote sensing of water (Chapter 6 and Section 7.3). For discussions on remote sensing of the atmosphere with linear polarisation and of vegetation with circular polarisation, the reader is referred to [116, 158].

[1] Because the magnetic field vector is always perpendicular to the electric field vector, it need not be considered separately.

Polarisation of incoherent light is described through Mueller calculus, which generalises Jones calculus but disregards coherent phenomena. Polarised radiance is quantified through the Stokes vector, as in Equation (1.3). $\vec{L}$ is the radiance Stokes vector, I the total radiance, Q the difference in radiance between the horizontal and vertical polarisation states ($I_{\leftrightarrow} - I_{\updownarrow}$), U between the two diagonal polarisation states, and V between right- and left-handed circular polarisation[2]. Any direction can be chosen as the reference *horizontal* $+Q$ direction. For most experiments and models, there is a natural choice, such as the direction of a particular polarising filter, the sky, or the sea surface. The vertical and diagonal directions are defined relative to the horizontal [156]. I is given by $I = I_{\leftrightarrow} + I_{\updownarrow} = I_{\nearrow} + I_{\nwarrow} = I_{\circlearrowleft} + I_{\circlearrowright}$. In cases where it is more intuitive to discuss polarisation in relative terms, the Stokes vector is normalised to the fractional polarisation $q, u, v = \frac{Q}{I}, \frac{U}{I}, \frac{V}{I}$.

$$\vec{L} = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \begin{bmatrix} I_{\leftrightarrow} + I_{\updownarrow} \\ I_{\leftrightarrow} - I_{\updownarrow} \\ I_{\nearrow} - I_{\nwarrow} \\ I_{\circlearrowleft} - I_{\circlearrowright} \end{bmatrix} \tag{1.3}$$

Linearly polarised radiance is often described in terms of the *degree of linear polarisation* (DoLP, P_L) and *angle of linear polarisation* (AoLP, ϕ_L), as defined in Equation (1.4). P_L ranges from 0 (no net polarisation) to 1 (fully polarised), ϕ_L from 0° (horizontal or $+Q$) through 90° (vertical or $-Q$) to 180° (horizontal or $+Q$), or equivalently from 0 to π radians. Because of symmetry (the 180° ambiguity), $\phi_L = 185°$ is equivalent to $\phi_L = 5°$.

$$P_L = \frac{\sqrt{Q^2 + U^2}}{I} \qquad \phi_L = \frac{1}{2} \arctan2\,(U, Q) \tag{1.4}$$

Direct sunlight is unpolarised, for all practical purposes, with $P_L < 10^{-6}$ [159], while diffuse skylight is partially linearly polarised by single scattering of sunlight in the atmosphere [160]. For a clean sky without any aerosols, viewed at 90° from the Sun, $P_L \approx 0.7$ at visible wavelengths. Multiple Rayleigh scattering by air molecules prevents P_L from reaching 1.0, while multiple scattering by aerosols can reduce P_L further, down to 0.5 on a hazy day and as low as 0.2 when smoke is present [161]. Sky polarisation is also decreased by reflections from clouds and the ground. For example, an effective ground albedo of 0.8 decreases P_L by as much as 0.3 compared to an albedo of 0 [162]. Polarisation also occurs due to scattering by particles in water [163, 164], specular reflection off surfaces [165], and interaction with *birefringent* materials such as coccolithophores [21].

In birefringent materials, the refractive index n depends on the polarisation state, being smaller along the *fast axis* and greater along the *slow axis*. This difference induces a phase delay δ as polarised light travels through the material. Optical elements designed to induce a specific δ are called *retarders* or *wave plates*. For example, the optical path through a $\frac{\lambda}{4}$ retarder or quarter-wave plate (QWP) is extended by $\frac{\lambda}{4}$ for light polarised along the slow axis, inducing a phase delay of $\delta_{\mathrm{QWP}} = 90°$ between the fast and slow axes. A QWP with a vertical fast axis converts fully right-handed circularly polarised light ($q = u = 0, v = 1$) to linearly polarised light with $\phi_L = 135°$ ($q = 0, u = -1, v = 0$).

[2] I, Q, U, V are sometimes referred to as S_0, S_1, S_2, S_3, and alternative sign conventions for Equation (1.3) are sometimes used.

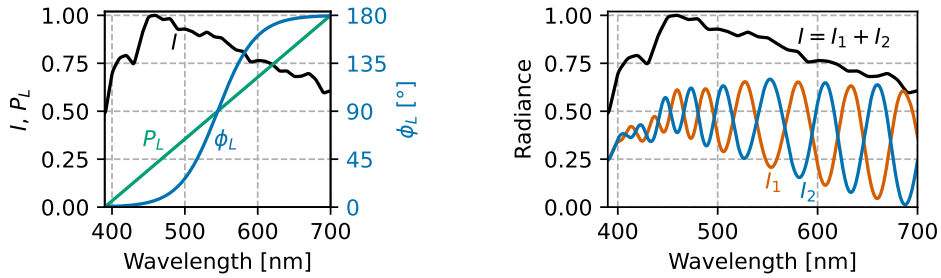
The interaction between polarised light and a medium or object is described by the Mueller matrix **M**, as defined in Equation (1.5). The Stokes vector of the outgoing radiance is the product of the incoming Stokes vector with the Mueller matrix, $\vec{L}_{\text{out}} = \mathbf{M}\vec{L}_{\text{in}}$. By determining the elements of **M**, which often vary with wavelength and viewing geometry [166, 167], the properties of a target can be inferred [168].

$$\mathbf{M} = \begin{bmatrix} M_{I\to I} & M_{Q\to I} & M_{U\to I} & M_{V\to I} \\ M_{I\to Q} & M_{Q\to Q} & M_{U\to Q} & M_{V\to Q} \\ M_{I\to U} & M_{Q\to U} & M_{U\to U} & M_{V\to U} \\ M_{I\to V} & M_{Q\to V} & M_{U\to V} & M_{V\to V} \end{bmatrix} \tag{1.5}$$

In the same way that wavelength increases the dimensionality of remote sensing data, and spectroradiometry on a two-dimensional sensor requires additional optics (Figure 1.2), so does measuring polarisation. Measurements techniques for spectropolarimetry fall into three categories, namely temporal, spatial, and spectral modulation [169, 170]. Temporal modulation involves sequential measurements at different angles, similarly to the use of spectral filters (Section 1.2.2). The linear polarisation state can be determined by measuring at 0°, 60°, and 120°; using more angles reduces the uncertainty. Multiple filters are arranged in a filter wheel (Figure 1.2a) or a single filter is rotated or tuned [132, 171]. The time delay between measurements in different polarisation states can create spurious signals when the target moves or changes on a similar time scale. Spatial modulation involves simultaneous measurements of different polarisation directions on different parts of one sensor, through a patterned filter (Figure 1.2b) or by projecting multiple images on the sensor [172, 173]. Alternatively, multiple sensors can be used, either in completely separate, co-aligned instruments or in one instrument using shared fore-optics and a polarising beam splitter [160, 163, 174]. A major challenge for spatial modulation is the appearance of false signals due to misalignment or intercalibration issues between sensors or pixels [160, 172].

Most relevant to this thesis is spectral modulation (sometimes called channelled polarimetry) through the SPEX technique. SPEX encodes the polarisation state into the observed spectrum using a highly chromatic multi-order retarder (MOR) [175]. Because the phase delay or *retardance* δ induced by the MOR varies strongly with λ, the polarisation state of light is modified with a known, wavelength-dependent modulation. When the light is passed through an analysing polarisation filter and dispersed, the resulting spectrum displays a sinusoidal modulation corresponding to the resulting polarisation state. The amplitude of the modulation corresponds to P_L and its phase to ϕ_L (Figure 1.5). By measuring in two channels I_1, I_2 corresponding to orthogonal analysing polarisers, for example through a polarising beamsplitter or a double aperture (Chapter 6), the full linear polarisation state can be retrieved using a demodulation algorithm. The total radiance I is recovered from $I = I_1 + I_2$, while P_L and ϕ_L are recovered by fitting a sinusoidal modulation to the normalised signal $\frac{I_1 - I_2}{I_1 + I_2}$ [116]. Because there are no moving parts and the light is spectrally dispersed, SPEX enables snapshot hyperspectral polarimetry. Existing implementations include groundSPEX [116, 125, 176], iSPEX [94], SPEX Airborne [177], and SPEXone, which will soon fly on the PACE mission [86, 112]. In Chapter 6, we present iSPEX 2 and discuss its design, manufacturing, and data processing. Results from a recent experiment with groundSPEX are discussed in Section 7.3.

Instruments are assessed on their polarimetric sensitivity and accuracy. The *polarimetric sensitivity* of an instrument is the smallest polarisation signal it can detect, and is usually limited by the uncertainty in the individual measurement channels and by spurious signals



(a) Input spectrum with I (black), P_L (green), and ϕ_L (blue).

(b) Spectrum observed by SPEX instrument, with I_1 in red and I_2 in blue, and the total radiance $I = I_1 + I_2$ in black.

Figure 1.5: Simulated SPEX observations (right) of polarised light (left). A normalised average European daylight spectrum (CIE D65) was used for the total radiance I. The P_L and ϕ_L spectra were constructed to illustrate the technique and are not physical. On the right, the sinusoidal modulation is clearly visible in both SPEX channels, I_1 and I_2, which correspond to orthogonal polarising filters (0° and 90°, respectively) and thus display a 180° modulation phase difference. As P_L increases, so does the modulation amplitude. The shift in ϕ_L from 500–600 nm is reflected by a change in modulation phase in the SPEX data.

due to instrumental or environmental causes. For detectable signals, the *polarimetric accuracy* expresses how well the measured polarisation state describes the true polarisation state of the incoming radiance, accounting for instrumental polarisation and depolarisation as well as cross-talk between the polarisation channels. The accuracy is affected by the physical characteristics of the instrument as well as the data processing and analysis. The observed polarisation state is particularly sensitive to instrumental polarisation caused by systematic errors in the individual measurement channels [116, 169]. Polarimetric accuracy and sensitivity are often conflated in the literature as the *polarimetric uncertainty*.

Compared to unpolarised spectroradiometry, there are relatively few instruments for spectropolarimetry of water. Research has historically been focused on the polarisation of the underwater light field [178]. Many studies have characterised the effects of various constituents on the Mueller matrix of water, for example by measuring the Stokes vector of scattered laser light at varying angles for phytoplankton and suspended particles [164, 179]. Spectropolarimetry above the surface is often conducted by adding polarising filters to spectroradiometers like the HyperOCR [180, 181]. The Mantis instrument, built by Polaris, measures linear polarisation hyperspectrally at 382–1017 nm with an uncertainty of ≤2.5% in radiance, ≤0.5% in P_L, and ≤0.5° in ϕ_L [171]. The mantis is in use by the US Naval Research Laboratory for spectropolarimetry of water [182]. Sony recently released their IMX250MYR and IMX250MZR sensors, featuring four polarisation filters (0°, 45°, 90°, 135°) in a Bayer pattern with an optional additional layer of RGB pixels. Initial results from observing floating debris with one of these sensors are presented in Section 7.3. For a broader review of spectropolarimetric remote sensing of water, the reader is referred to [165].

The most prominent satellite-based spectropolarimeters have been the POLarization and Directionality of the Earth's Reflectances (POLDER) instruments on the ADEOS and PARASOL satellites [132, 165, 180, 183]. The latest, POLDER-3, flew in 2004–2013 and measured

polarised radiance at 490 nm, 670 nm, and 865 nm, in addition to 7 unpolarised bands at 443–1020 nm, at 16 viewing angles between −51° and +51° relative to nadir and with a spatial resolution of ∼6 km per pixel [132]. As discussed previously, POLDER 1–3 measured polarised radiance using a filter wheel rotated to 0°, 60°, and 120° sequentially, with a polarimetric accuracy of 8.5×10^{-4} in P_L [180].

The upcoming PACE mission will feature the SPEXone and HARP-2 spectropolarimeters [86]. SPEXone, designed and built in the Netherlands, uses the SPEX technique to measure radiance and linear polarisation hyperspectrally at 385–770 nm with a 2 nm resolution in radiance and 20–40 nm resolution in P_L. It observes at five viewing angles (0°, ±20°, and ±57° from nadir) and has a spatial sampling of 2.3×2.7 km^2 per pixel. Lastly, the expected uncertainty in SPEXone data is 2% in radiance and 0.0025 in P_L [112,184]. Unlike SPEXone, HARP-2 is multispectral and hyperangular, measuring at 10 viewing angles in three spectral bands (440, 550, and 870 nm) and at 60 viewing angles at 670 nm, spanning 114°. Its ground pixel size is similar to that of SPEXone, at 3 km, and its expected polarimetric accuracy is <0.005 in P_L [86, 185]. SPEXone and HARP-2 are primarily intended for observations of atmospheric aerosol [185], but will also be used in conjunction with OCI for remote sensing of water [186–188]. Other satellite spectropolarimeters of note include the Directional Polarimetric Camera (DPC) on the GaoFen-5 mission [107, 189], and the upcoming PolCube cubesat [190]. For a broader review of satellite spectropolarimetry, though with a focus on aerosols, the reader is referred to [132].

Some animals, including humans, can also sense polarisation. For example, some insects navigate using sky polarisation [191] and mantis shrimp famously see linear and circular polarisation [192]. With training, humans can see linear polarisation as Haidinger's and Boehm's brushes [193, 194].

1.3 Remote sensing of water

Remote sensing is a technique for observing a target without physically sampling it. This thesis is focused on optical remote sensing, using light as the information carrier, of surface waters. Alternative approaches exist, such as acoustic sensing, but are not discussed here. Furthermore, we focus on *passive remote sensing*, using reflected sun- and skylight. The alternative, *active remote sensing*, involves the use of a custom, controllable light source, for example for lidar measurements.

Current platforms for remote sensing of water include in-situ spectroradiometers that are hand-held or mounted on a platform, uncrewed aerial vehicles (UAVs or drones), aeroplanes, and satellites. Each platform has an associated cost, spectral and spatial coverage and resolution, revisit time, and sensitivity. This thesis focuses primarily on the validation and application of smartphone cameras as in-situ spectroradiometers (Chapters 3, 4, and 6), along with the human eye (Chapter 2) and the principles behind comparing different instruments and platforms (Chapter 5).

Remote sensing provides solutions to many of the challenges posed in Section 1.1.4. Autonomous instruments like satellites and in-situ radiometers significantly reduce the expense and labour requirements of continuous measurements. Earth observation of aquatic ecosystems with satellite sensors enables global measurements of water constituents, which in turn enables research on global issues such as climate change. Reducing the dependence on physical sampling and human labour improves the consistency over time and between sensors,

facilitating decade-long time series with a high degree of reproducibility. Lastly, since data are obtained and processed electronically, remote sensing can provide near-instant response times. Since remote sensing does not directly probe the same parameters as in-situ sampling, but instead probes the proxies discussed below, it is not a one-to-one replacement, but rather a complementary method.

In this section, we discuss the physics and practice of remote sensing of water. First, the inherent optical properties of water constituents are discussed (Section 1.3.1), followed by the apparent optical property of reflectance (Section 1.3.2), from which optical parameters and constituents are retrieved (Section 1.3.3). Lastly, Section 1.3.4 discussed the practical aspect of validation.

1.3.1 Inherent optical properties

The interaction between light and water is described by the *inherent optical properties* (IOPs) of the water column [14, 195]. The IOPs are the coefficients of *absorption* a, *scattering* b, and *beam attenuation* c, all in m^{-1}. This c is not to be confused with the speed of light. Absorption converts light into chemical or thermal energy, while scattering shifts energy into different directions of propagation and, in inelastic scattering, wavelengths. Beam attenuation is the sum of the two, as in Equation (1.6), and describes the exponential decay in radiant power as light passes through water, following Lambert's law as in Equation (1.7). Here $L(x, \lambda)$ is the radiance at wavelength λ after travelling a distance x through the water column. For illustration, $c = 0.7\ \mathrm{m}^{-1}$ means the radiance is halved for every metre travelled.

$$a + b = c \tag{1.6}$$

$$L(x, \lambda) = L(0, \lambda)\, e^{-c(\lambda)x} \tag{1.7}$$

Absorption is highly wavelength-dependent and many constituents of natural waters (Section 1.1) have characteristic absorption spectra (Figure 1.6). Adding the constituent absorption spectra together gives the total absorption spectrum in the water column. The major absorbing constituents are phytoplankton, CDOM, and non-algal particles (NAP). NAP refers to particles that do not contain extractable pigments [14]. Water itself is highly absorbing at longer wavelengths, hence its blue appearance [49, 196].

Phytoplankton primarily absorbs through pigments with characteristic spectral signatures. The average phytoplankton absorption spectrum is dominated by a narrow chl-a peak around 665 nm and a broad peak around 442 nm (Figure 1.6), primarily caused by chl-a and broadened by other chlorophylls such as chl-b and chl-c as well as by carotenoids such as β-carotene, fucoxanthin, and zeaxanthin [14, 199]. The cyanobacterial pigment complexes phycocyanin and phycoerythrin can cause additional absorption peaks or shoulders around 495 nm, 550–570 nm, and 630 nm [49]. Phytoplankton absorption spectra are broadened and flattened by formation of pigment-protein complexes and packing of pigment inside cellular structures, both of which change the vibration energies of molecular bonds. The precise shape and magnitude of the spectrum depend on the species present and on cell characteristics that vary over the diurnal cycle [14, 88]. For a detailed review of pigments found in various classes of phytoplankton, the reader is referred to [200].

The absorption spectra of CDOM and NAP also depend on their composition, but can be approximated by an exponential function in λ, as shown in Equation (1.8). Here $a(\lambda_0)$ is the

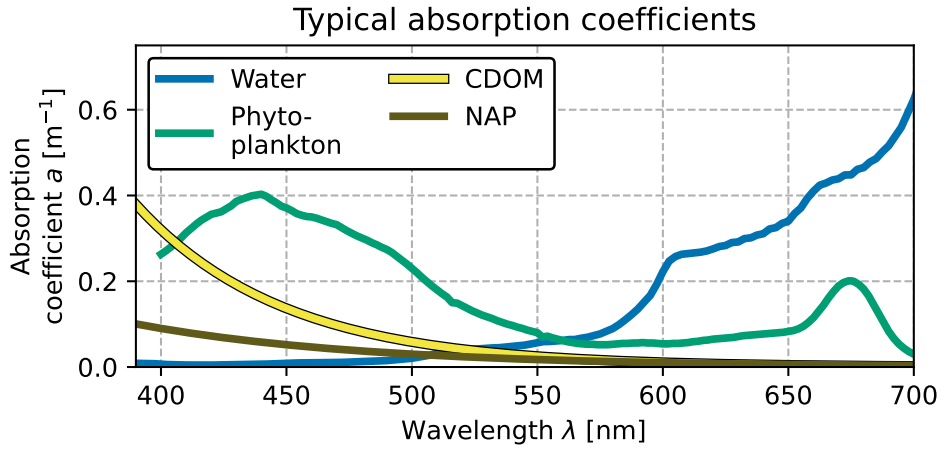


Figure 1.6: Typical spectral absorption coefficients $a(\lambda)$ of water, phytoplankton (at 10 mg m^{-3} chl-a), CDOM ($a(400\text{ nm}) = 0.32\text{ m}^{-1}$, $S = 0.017\text{ nm}^{-1}$), and non-algal particles (NAP, $a(400\text{ nm}) = 0.09\text{ m}^{-1}$, $S = 0.011\text{ nm}^{-1}$). Based on data and models from [14, 196–198].

absorption at a reference wavelength λ_0 and S is the spectral slope, typically 0.006–0.013 nm^{-1} for NAP and 0.012–0.022 nm^{-1} for CDOM, depending on the composition [14, 198]. For CDOM, the exponentially-shaped spectrum is thought to arise from a superposition of different bonds within the long organic molecules. Short bonds absorb at shorter wavelengths, while resonances of multiple bonds absorb at longer wavelengths. Short bonds are more common and thus the absorption decreases with λ [14].

$$a(\lambda) = a(\lambda_0)\, e^{-S(\lambda-\lambda_0)} \tag{1.8}$$

Elastic scattering by particles is relatively spectrally flat, but highly directional [201]. The scattering coefficient decreases slightly with wavelength, with a spectral slope related to the average particle size [32, 202]. Rayleigh scattering by water molecules scales with λ^{-4} and is only significant at blue and UV wavelengths, since $b_{\text{water}}(\lambda > 350\text{ nm}) < 0.02\text{ m}^{-1}$ [203]. The *volume scattering function* $\beta(\hat{\xi})$ (VSF, in $\text{m}^{-1}\text{ sr}^{-1}$) describes the distribution of light across different scattering angles. $\beta(\hat{\xi})$ is the product of b (magnitude, as before) and the phase function $\tilde{\beta}(\hat{\xi})$: $\beta(\hat{\xi}) = b\tilde{\beta}(\hat{\xi})$.

For unpolarised light, scattering in water is typically azimuthally symmetric and β varies only with the scattering angle ψ [14], as in Figure 1.7. The angular dependence of β is dictated by the morphology and internal structure of particles, and measuring at multiple angles probes these properties [164, 204]. In remote sensing, only light scattered back out of the water is observed, which is described by the *backscattering coefficient* b_b. Mathematically, b_b is the integral of $\beta(\psi)$ over the angles $90° \leq \psi \leq 180°$. Natural waters tend to be highly forward-scattering, with relative backscattering coefficients $\frac{b_b}{b} \approx 0.01$ [204, 205]. Elastic scattering induces linear polarisation, particularly at $\psi \approx 90°$ where $M_{I\to Q} \geq 0.5$ for pure sea water and many types of sediment and phytoplankton [164, 167].

Unlike elastic scattering, inelastic processes like fluorescence and Raman scattering do not conserve wavelength. A photon of wavelength λ is absorbed and another is re-emitted with a wavelength λ'. In the visible range, ~2% of the solar irradiance is re-emitted through

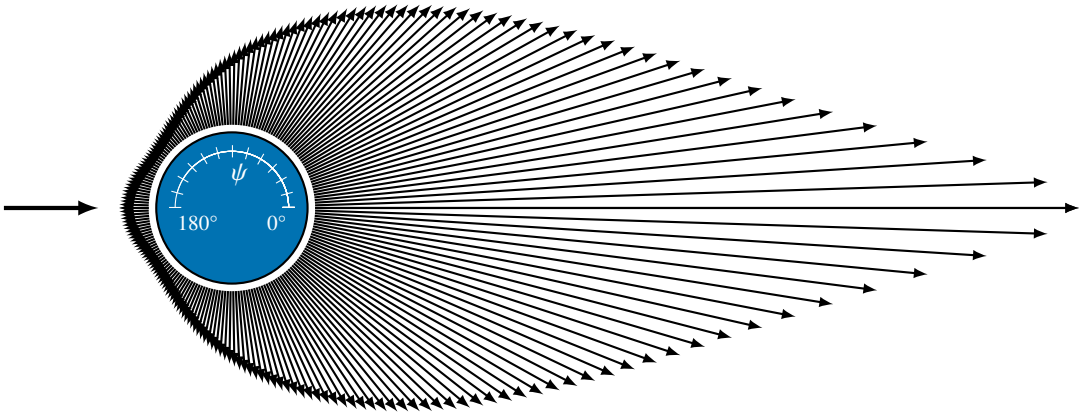


Figure 1.7: Scattering of light in water. Light coming in from the left gets scattered in all directions. The scattered arrow length scales logarithmically with the phase function $\tilde{\beta}(\psi)$ at scattering angles ψ. A parameterised version of the empirical Petzold turbid phase function from San Diego harbour was used [204, 205], which is azimuthally symmetric. The three-dimensional phase function can be visualised by rotating this image about the horizontal axis.

fluorescence, primarily by CDOM and phytoplankton pigments [165]. Phytoplankton species can be distinguished by their distinctive re-emission spectra, for example using the ratio between re-emission at 660 nm vs. at 684 nm when illuminated with 445 nm light [53]. Fluorescence is one of the most common techniques for in-situ measurement of chl-a (Section 1.1.2). Raman scattering by water contributes up to several percent of the total light field, varying with wavelength [96]. Fluorescence and Raman scattering are depolarising, which is used to distinguish light emitted through these processes from elastic scattering [165].

Constituent concentrations and properties can be derived from IOP measurements. For example, measurements of the a(676 nm) peak height (Figure 1.6) can be inverted to obtain phytoplankton concentrations [52]. Flow-through instruments like the LISST and AC-s provide high-quality in-situ IOP data for these purposes [28, 40]. However, sampling is often laborious and expensive (Section 1.1.4). Remote sensors can complement IOP measurements by measuring the *apparent optical properties* (AOPs) of water, the results of the interaction between incoming light and the IOPs [14]. As described below, AOP measurements can be used to estimate the IOPs and constituent concentrations.

1.3.2 Reflectance

The most important apparent optical property in remote sensing is *reflectance*, the ratio of water-leaving over downwelling (ir)radiance. Light impinging on water is directly reflected off the surface or penetrates into the water column. The latter component is then scattered and absorbed by the water and its constituents. A fraction of this light is scattered back out of the water, either directly by a single backscattering event or indirectly through multiple scattering events. This component is called the *water-leaving radiance* L_w and holds information on the IOPs. Directly reflected light (*glint*) does not hold information on the IOPs and is not part of L_w, but acts as a source of uncertainty. The relative contribution of glint to the total upwelling radiance, which includes glint and L_w, ranges from near-zero to near-100% depending on viewing conditions, IOPs, and wavelength [206]. The main challenge in remote sensing of

water is to accurately measure L_w, and from it determine the reflectance [120].

Modern remote sensing of water uses the *remote sensing reflectance* R_{rs}, which is the ratio of L_w over the downwelling irradiance E_d, as in Equation (1.9). R_{rs} is most commonly defined to be in air, just above the water surface, and has units of sr^{-1}. Alternative combinations of (ir)radiances exist but are less popular (Section 5.2). L_w is a directional quantity $L_w(\lambda, \hat{\xi})$, and thus so is R_{rs}. However, R_{rs} is usually only evaluated in a single viewing direction with a narrow field of view (1°–7° diameter), so this directionality is left out [14]. Spectropolarimetric versions of R_{rs} include the reflectance Mueller matrix and the ratio of polarised L_w (defined as $\sqrt{Q^2 + U^2}$) over unpolarised E_d [165].

$$R_{rs}(\lambda) = \frac{L_w(\lambda)}{E_d(\lambda)} \tag{1.9}$$

Strictly speaking, R_{rs} has two wavelength dependencies, namely λ for E_d and λ' for L_w, because of inelastic scattering. For example, fluorescence can shift energy from $E_d(445\,\mathrm{nm})$ to $L_w(685\,\mathrm{nm})$. However, because field measurements are performed using natural broad-spectrum E_d from the Sun and sky, it is not possible to actually determine R_{rs} as a two-dimensional quantity, and the one-dimensional approximation in Equation (1.9) must be used. In edge cases where fluorescence or Raman scattering is the dominant process at a particular wavelength, this approximation can make the interpretation of results less intuitive.

Satellite instruments determine R_{rs} from the top-of-atmosphere (TOA) E_d and L_w. The TOA E_d is well-known from measurements of the extraterrestrial solar irradiance, which is essentially constant over time. The TOA E_d at a given time depends primarily on the current Sun-Earth distance, which is easily calculated from orbital models [97]. The TOA total upwelling radiance is measured at a chosen angle. An atmospheric correction algorithm is used to determine the corresponding surface-level E_d and L_w. Scattering in the atmosphere provides the bulk of the TOA upwelling radiance, with L_w accounting for ≤20% at blue wavelengths. While the contribution of scattering by air is relatively constant over time, absorption and scattering by aerosols is highly variable. Aerosols come in many flavours, including dust, pollen, sea salt, and soot, and accordingly have a wide variety of optical properties. Atmospheric correction algorithms remove the contributions of air, aerosols, and trace gases like ozone to the TOA radiance based on empirical concentrations and optical properties and on models of radiative transfer and aerosol transport through the atmosphere [207]. Atmospheric correction is a high-dimensional problem and thus includes many uncertainties, which are propagated into L_w five- to tenfold and are the primary contributor to uncertainty in satellite-based R_{rs} measurements [186, 208]. In inland and coastal waters, additional uncertainty is introduced by the adjacency effect, where radiance from targets outside the field of view (FOV) of a pixel, such as clouds or land, is scattered into the FOV by the atmosphere [188]. Currently, the typical uncertainty on satellite measurements of R_{rs} is ~5% at blue and green wavelengths, although the reproducibility between different instruments is worse [88].

Terrestrial and airborne instruments determine R_{rs} from (near-)simultaneous measurements of L_w and E_d. The technique most relevant to this thesis is terrestrial above-water radiometry (Figure 1.8), in which measurements of the total upwelling radiance L_u and sky radiance L_{sky} are used to determine L_w [209], as in Equation (1.10). For a full review of terrestrial techniques, the reader is referred to [97, 120].

$$R_{rs}(\lambda) = \frac{L_u(\lambda) - \rho L_{sky}(\lambda)}{E_d(\lambda)} \tag{1.10}$$

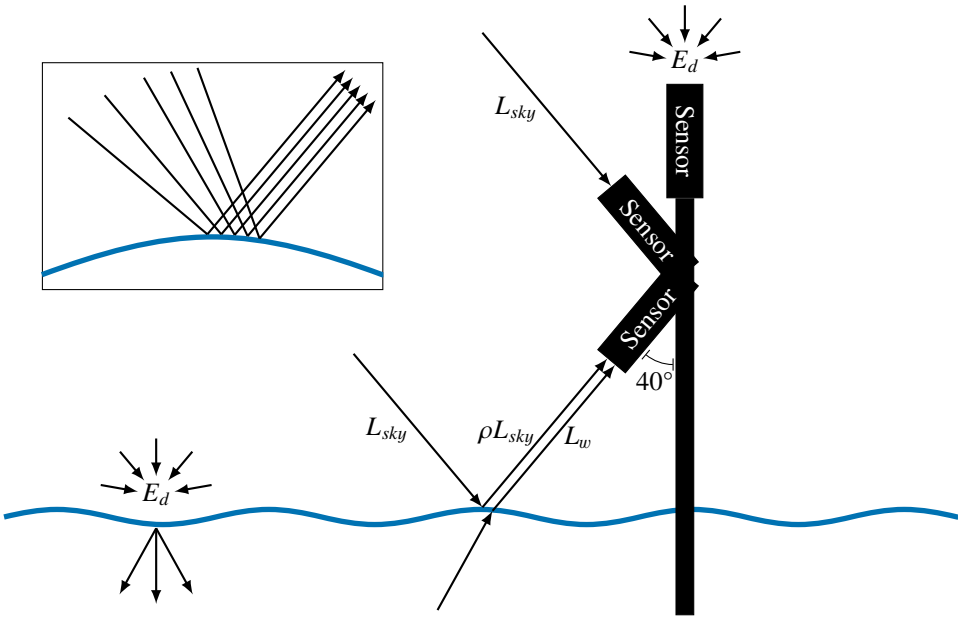


Figure 1.8: Schematic representation of an above-water radiometry setup. Downwelling irradiance E_d impinges on the water surface and propagates into the water column (bottom left), while simultaneously being measured by an E_d sensor (top right). A sensor looking downwards measures the upwelling radiance L_u, which is the sum of the water-leaving radiance L_w and the directly reflected sky radiance ρL_{sky}, with ρ the surface reflectivity. A third sensor, looking upwards, measures L_{sky} directly. As shown in the inset, the reflected sky radiance ρL_{sky} originates from different locations on the sky due to specular reflectance from different wave facets.

The surface reflectivity ρ expresses the proportion of L_{sky} that is directly reflected into the sensor field of view, typically 0.025–0.030 but highly dependent on viewing conditions [209, 210]. For pure, still water, L_{sky} is directly reflected from only a single direction and the Fresnel reflectance can be used. However, real water surfaces have waves. As illustrated in the inset of Figure 1.8, each facet of a wave corresponds to a different reflection angle, and wavy surfaces thus reflect light from a wide area on the sky, depending on the exact viewing geometry and wave conditions. The variability of ρ with wave conditions, specifically wind speed, is minimised by observing at $\theta = 40°$ from nadir or zenith (in elevation) and $\phi = 135°$ in either direction from the Sun (in azimuth) [209, 211].

Distinguishing between water-leaving and directly reflected radiance remains a major challenge [120, 209]. The accuracy of above-water measurements is improved by empirically determining ρ for each measurement and by subtracting an additional offset Δ from the right-hand side of Equation (1.10). Δ and ρ can be determined from imaging of wave conditions and from spectral analysis of L_u at wavelengths with known atmospheric features. Using these methods, the typical error on R_{rs} induced by glint can be reduced to 6–8% [210]. Above-water radiometry is discussed in more detail in Chapter 4.

1.3.3 Constituent retrieval and inversion

R_{rs} is proportional to the ratio of b_b over a, or some empirical function thereof, as shown in Equation (1.11). Traditionally, the former relation is used in phytoplankton-dominated Case I waters, and the latter in Case II waters, but this distinction is controversial [212]. The relation can be understood intuitively – highly absorbing waters (large a) appear dark because few photons make it out, whereas highly scattering waters (large b_b) appear bright because photons are quickly scattered back out. The proportionality constants between R_{rs} and the IOP ratios vary slightly due to environmental conditions, constituent concentrations, and viewing geometry [212]. Sometimes, a polynomial function is used instead of a linear relation. An exact inversion from R_{rs} to a, b_b requires a full radiative transfer simulation accounting for all radiative processes including detailed knowledge of the IOPs and their dependence on viewing angle and depth, but the relations shown in Equation (1.11) provide an adequate approximation [14, 96]. It is important to note that passive remote sensing is only sensitive to the top layer of the water column, since few photons penetrate deep into the water, and of those that do, even fewer make it back out [213]. The penetration depth ranges from centimetres in highly turbid or absorbing waters to ~50 m in clear ocean waters and is highly wavelength-dependent [96].

$$R_{rs}(\lambda) \propto \frac{b_b(\lambda)}{a(\lambda)} \qquad R_{rs}(\lambda) \propto \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \tag{1.11}$$

The spectral signatures of different constituents in the IOP spectra propagate into R_{rs} (Figure 1.9), and are exploited to retrieve constituent IOPs and concentrations. Strong absorption from water decreases R_{rs} with wavelength for $\lambda > 600$ nm in clear waters with low chl-a and SPM concentrations. In CDOM-rich waters, R_{rs} at blue wavelengths ($\lambda < 500$ nm) is depressed by CDOM[3] and NAP absorption; S_{CDOM} can be estimated from the spectral slope in R_{rs} at these wavelengths. Phytoplankton lowers R_{rs} at wavelengths where chl-a absorbs strongly (Figure 1.6), particularly around 443 nm and 665–675 nm [213]. Fluorescence can increase R_{rs} in specific bands, such as around 680–690 nm in some of our data from around Lake Balaton (Section 4.2.2). Different pigments have specific absorption and fluorescence bands, which are used to distinguish phytoplankton species and functional groups [42, 49, 214], as discussed below. Particulate backscattering increases R_{rs} across the visible range [96], and b_b, TSM, and related quantities can be estimated from R_{rs} at red wavelengths where particles are the dominant scatterers [32, 213]. More detailed information can be obtained if spectral features specific to a constituent are known [215] or with prior information on the target, for example on the typical local phytoplankton community composition [216].

A wide range of retrieval algorithms are used in practice, ranging from empirical to analytical [13, 14, 88, 195, 213, 218, 219]. Empirical algorithms are derived by comparing remote observations with field measurements or simulations (Section 1.3.4), without additional inputs. This makes them relatively quick and simple to derive, but less portable as they depend, explicitly or implicitly, on assumptions that only hold locally or under specific conditions. Many empirical algorithms are *band-ratio algorithms*, relating a single parameter to the ratio of R_{rs} at two or more wavelengths [220–223]. A major advantage of band-ratio algorithms is their robustness to correlated uncertainties and biases in L and R_{rs}, which may divide out

[3] Hence the old name *gelbstoff* or *yellow substance*.

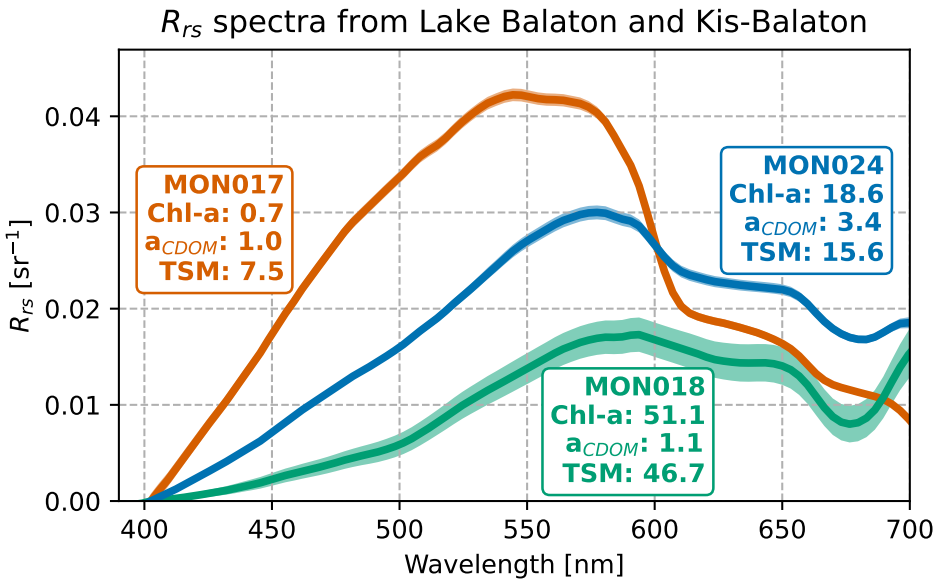


Figure 1.9: R_{rs} spectra from Lake Balaton (MON017, MON024) and Kis-Balaton (MON018) taken with TriOS RAMSES spectroradiometers during the 2019 MONOCLE field campaign [217]. Each line represents the median spectrum from a given station, the shaded areas represent the standard deviations. Biogeochemical data (chl-a concentration in mg m^{-3}, a_{CDOM}(440 nm) in m^{-1}, and TSM in g m^{-3}) from in-situ samples are also provided. As discussed in Section 1.3.3, CDOM and phytoplankton absorption induce a slope in R_{rs} at $\lambda \leq 500$ nm, with R_{rs} lower at the stations richer in CDOM and chl-a. Inflection points in the water absorption spectrum around 600 and 650 nm (Figure 1.6) are seen in R_{rs}, as is the chl-a absorption peak at 665 nm. Furthermore, MON024 is brighter than MON017 at $\lambda > 600$ nm due to increased particulate scattering. TSM at MON018 is 3× that at MON024 but R_{rs} is lower, because of differences in the suspended matter composition – Kis-Balaton contains mostly algal particles while western Lake Balaton contains many bright chalk particles [27]. Further examples of R_{rs} spectra are shown in Figures 4.2 and 5.2.

(Chapter 4). Conversely, analytical algorithms are grounded in radiative transfer models and are thus based on physics, theoretically making them more portable to different applications but also vulnerable to uncertainty in prior information. Any uncertainty in the IOPs used to derive the analytical model will propagate into every retrieved value. Semi-analytical models combine the two approaches by supplementing observations with prior information such as typical IOPs at a certain site [27, 224–226]. Generally, empirical algorithms retrieve one or more constituent concentrations, while (semi-)analytical algorithms focus on retrieving IOPs [213]. This thesis is focused primarily on empirical band-ratio algorithms, which are the most appropriate and popular type for multispectral data from consumer cameras (Chapter 4) and the type most commonly applied through spectral convolution (Chapter 5).

Most band-ratio algorithms for chl-a make use of the absorption peaks around 443 nm and 665 nm, the fluorescence peak around 680 nm, and wavelengths with minimal absorption by phytoplankton, CDOM, and particulate matter such as 550 nm and 700 nm [213,220]. For example, the OCx algorithms relate chl-a to a function of the ratio between blue- and green-band R_{rs} from any ocean colour sensor [220]. Increased chl-a leads to an increase in a(443 nm) and

thus a decrease in $R_{rs}(443\,\text{nm})$, while $R_{rs}(550\,\text{nm})$ remains relatively unchanged, leading to a decrease in the $\frac{R_{rs}(443\,\text{nm})}{R_{rs}(550\,\text{nm})}$ ratio. In Case I waters, where chl-a and CDOM covary and thus the overlap between their absorption spectra is not an issue [213], OCx algorithms can predict chl-a to within a factor of ~1.74 compared to in-situ samples [220]. In lakes with little chl-a (0–10 mg m^{-3}) and high CDOM and NAP absorption, meaning the Case I assumption does not hold, the OCx chl-a estimate is inaccurate [218]. Many alternative chl-a algorithms exist for specific sites. For example, an exponential function of the Sentinel-2/MSI red-green band ratio (B3/B4) predicts chl-a with an accuracy compared to in-situ data of 20% in Lake Ba Be in Vietnam [221]. The target accuracy for chl-a retrieval compared to in-situ measurements was 35% for SeaWiFS and 10–30% (Case I) or 10–70% (Case II) for Sentinel-3/OLCI. Most applications require an uncertainty in chl-a of 10–25% [88].

Empirical retrieval algorithms also exist for many other parameters. For example, the concentration of the cyanobacterial pigment phycocyanin can be determined from the 709/620 nm reflectance band ratio with an accuracy of 20% (typical difference) compared to in-situ samples, although the accuracy is worse when other phytoplankton species and pigments are more abundant [214]. Algorithms for particulate b_b and TSM often use the absolute R_{rs} in one or multiple red–NIR bands, where TSM is the dominant optical component [213], or a red-blue band ratio [34]. In one study in the North Sea, R_{rs} at any wavelength from 670–750 nm predicted TSM to within 30% accuracy compared to in-situ data [227]. Lastly, a_{CDOM} and S_{CDOM} at various UV–blue wavelengths can be determined from the 412 nm-670 nm and 412 nm-547 nm R_{rs} band ratios to within 24–30% (a) and 8–10% (S) accuracy [222]. As with chl-a, there are many different retrieval algorithms for the aforementioned constituents, derived for various sites and instruments. Simultaneous retrieval of multiple parameters is achieved by taking advantage of the linear summability of IOPs and the existence of wavelengths with minimal overlap between IOP spectra [228]. Finally, some algorithms convert the R_{rs} spectrum into a lower-dimensional colour measure before retrieval, such as central wavelength, hue angle, or blueness [10, 229, 230]. For a review of empirical and semi-analytical retrieval algorithms for chl-a and other algal pigments, TSM, and CDOM, the reader is referred to [213, 231].

In general, empirical algorithms reduce the vast biogeochemical parameter space, which includes concentrations of many different phytoplankton species, dissolved organic and inorganic matter, and various types of particles, into a handful of spectral bands. This simplification is possible in practice due to covariance between parameters and constraints on the parameter space in specific sites. However, this fundamentally means that empirical algorithms perform well only in the conditions they were designed and validated for, and poorly when the underlying assumptions do not hold. Analytical and semi-analytical approaches that include prior information provide a more robust alternative, although these too are limited by the accuracy of said prior information. Additionally, measuring hyperspectrally and with a wider wavelength range, for example in the UV range for CDOM, provides a wider input space, although only if measurement uncertainties are sufficiently small [87, 144, 232]. Machine learning techniques, including neural networks, are increasingly popular for deriving new algorithms and combining existing ones [219, 233]. Particularly important for future algorithm development is the generation of in-situ validation data sets with greater size and diversity [218]. For a general review of retrieval and inversion algorithms based on spectral features in R_{rs}, the reader is referred to [41, 131, 213, 231, 234].

Currently, polarisation is largely neglected in remote sensing of water, although many proof-of-concept studies have shown that including polarisation improves the accuracy of

measurements of organic and inorganic particle shape, size, and composition [163, 174, 182, 235, 236]. The use of multiangular spectropolarimetry for retrieval of particle properties is more common in the atmospheric aerosol research community, where spaceborne, airborne, and ground-based spectropolarimeters provide data products including aerosol optical depth, effective particle radius, complex refractive index[4], and aerosol layer height, separated between fine- and coarse-mode particles [125, 132, 237]. In water research, spectropolarimetry is seen as a high-potential technique for the near future [165, 238], particularly the upcoming PACE mission with its SPEXone and HARP-2 instruments [86]. As discussed in Section 1.2.3, SPEXone and HARP-2 are primarily designed for aerosol science, which will improve the retrieval of water reflectance and constituents by reducing the uncertainty due to atmospheric correction [185, 188]. However, these two instruments, especially when used in conjunction with OCI, will also enable new water research through joint atmosphere-water retrieval algorithms [186, 187, 239]. For example, coccolithophores polarise scattered light depending on their degree of calcification, which is under pressure from ocean acidification [21]; thus, polarimetric remote sensing could be used to assess coccolithophore health and the ecological impact of acidification globally [20]. At a 50°–60° viewing angle, P_L is correlated with the $\frac{c}{a}$ IOP ratio and polarisation can thus be used to distinguish between absorption and scattering [174]. Lastly, polarisation can be used to characterise and minimise the error induced by glint, since direct reflections are highly polarised compared to water-leaving radiance, although as in Figure 1.8, this is complicated by wave motion [174, 211, 240].

1.3.4 Validation

A vital aspect of remote sensing is *validation*, assessing the performance of an instrument or algorithm under realistic conditions. Data and retrieved parameters are compared with those from a different instrument or algorithm, or with in-situ sampling or simulated data. When different methods or instruments agree to within measurement uncertainties, closure is achieved [40, 88].

Validation data are obtained during dedicated campaigns, from large instrument networks, or from aggregated archival data. For example, Chapter 4 describes a three-day campaign on and around Lake Balaton, Hungary, during which we validated L, R_{rs}, and colour measurements from smartphone cameras. Instrument networks like AERONET-OC and Argo provide regular, consistent optical and biogeochemical validation data from many diverse sites [80, 117]. Lastly, archives such as LIMNADES, PANGAEA, SeaBASS [241], and Zenodo contain diverse data sets from many different cruises and instruments, which are sometimes bundled together into larger data releases [50]. These archival data can be used again for novel research, such as in Chapter 5, which discusses instrument validation through intercomparisons between multi- and hyperspectral data.

The research described in this thesis was performed within the Horizon 2020 project MONOCLE[5], which had as one of its aims to improve the validation of remote sensing data. To this end, several low-cost in-water [242] and above-water spectroradiometers, including iSPEX 2 (Chapter 6), were developed and tested in the field (Figure 1.10). The low cost of these instruments makes them attractive for autonomous use, such as mounting one on a

[4] The complex refractive index consists of a real part (Section 1.2.2) and an imaginary part describing absorption.

[5] Multiscale Observation Networks for Optical monitoring of Coastal waters, Lakes and Estuaries, https://monocle-h2020.eu/

ferry, for large-scale deployment, and for use by citizen scientists. This in turn increases the scale and diversity of the obtained validation data, while also providing the community with new instruments for novel research.

1.4 Citizen science

Citizen science or *participatory research* is the involvement of non-professionals (citizens) in the scientific process. There are several types of involvement, ranging from crowdsourcing, where citizens collect data for professional researchers, to co-creation, where the citizens are involved in various aspects of formulating and investigating a research question [95, 243]. Citizen science projects exist across this spectrum in many scientific fields, particularly in biological, ecological, and geographical research [244]. On a metalevel, research into citizen science itself is focused on determining and optimising participant motivation and retention; data quality, quantity, and value; and the interaction between professional and citizen scientists [95, 245–247].

Citizen involvement in science brings many benefits both to the citizens and to professional scientists, and provides insights that could not be obtained by either group alone. Crowdsourcing greatly expands the amount and diversity of data that can be gathered, which is desirable in most fields of environmental science (Sections 1.1.4 and 1.3.4). For example, in the first measurement campaign using the iSPEX smartphone spectropolarimeter, on 8 July 2013, 6 007 measurements of aerosol optical depth throughout the Netherlands were conducted by citizen scientists [94]. To obtain similar levels of spatial coverage with professional ground-based sensors would be prohibitively expensive, while satellites could not provide the spatial resolution and revisit time. Involvement of citizens in the processing, analysis, and interpretation of data provides professional scientists with novel insights [248, 249], such as local information about a measurement site that an external researcher otherwise would not have known (Section 1.3.3). Simultaneously, the citizen participants gain a greater understanding of the scientific process in general and their project in particular, which in turn empowers them as social and political stakeholders [66, 250, 251], as encouraged in the United Nations Sustainable Development Goals and European Green Deal [252, 253]. It is important to ensure that all parties, citizen and professional alike, fulfil their motives, and to prevent negative impacts such as interpersonal conflicts from occurring [247].

Because of its great importance to society (Section 1.1), water research has always been a popular field for citizen science. Many of these projects are conservation efforts, for example monitoring plastic pollution [77, 254]. Participants' intimate knowledge of their local study site enables them to identify changes, abnormal events, and long-term trends that outside researchers cannot [248, 249]. A major motivation for many participants is a desire to protect their local environment and community, in addition to a desire to learn more and contribute to scientific advances [255, 256]. After all, citizens are the most important societal stakeholders in and funders of water research and policy [1]. Citizen science benefits from having many, diverse participants, but recruiting and retaining them is difficult and demographics are often skewed [257]. For a broader overview of citizen science in water research, the reader is referred to [243, 248, 249, 258]; citizen motivation and retention is discussed further in Chapter 2.

Citizen science benefits from access to low-cost equipment, with the Secchi disk for water transparency measurements and Forel-Ule scale for water colour as prominent examples.

Figure 1.10: The MONOCLE team performing validation measurements at Loch Leven, United Kingdom, in August 2018. Three boats with various hand-held and mounted spectroradiometers as well as in-situ bio-optics instruments are on the lake, with three UAVs carrying hyper- and multispectral cameras in the sky above. The measurements were performed during a Sentinel-2 satellite overpass. Photograph courtesy of Stefan Simis.

The Secchi disk was first used by papal astronomer Angelo Secchi[6] in 1865 to measure water transparency [259, 260]. It is a white or black-and-white disk, normally ~30 cm in diameter, which is lowered into the water on a rope until it disappears from view. The depth at which it is no longer visible is the *Secchi disk depth* z_{SD}. The value of z_{SD} is related to the beam attenuation coefficient c (Section 1.3.1) and the diffuse attenuation coefficient K_d, an AOP describing the change in E_d with depth, although the exact relation remains a matter of debate [14, 259, 261–263]. z_{SD} is deepest (>80 m) in clear ocean waters and shallowest in turbid Case II waters with heavy particle loads [261]. For example, z_{SD} in highly turbid Lake Balaton (Chapter 4) can be as shallow as 20 cm [264]. The simplicity of the Secchi disk makes it ideal for citizen science since participants can 3D-print or craft their own disk [265, 266]. Variants such as the Secchi tube are also in use [66]. A new smartphone application for Secchi disk measurements[7] was developed within the aforementioned Horizon 2020-funded MONOCLE project.

The Forel-Ule (FU) scale is a standardised 21-colour scale to classify water colour, first devised in the 1890s by François Forel and Willi Ule [267, 268] and to some extent revived in the 2010s by Marcel Wernand [269]. Measurements are performed by comparing the apparent colour of the Secchi disk at $\frac{z_{SD}}{2}$ with a physical colour scale. The FU scale can be used without a Secchi disk by looking straight down the water column, provided the bottom is not visible, but this biases the result by up to 2 FU [270]. Century-long FU time series with global coverage enable the study of long-term changes in ocean colour and chl-a concentration [10, 271]. The FU scale is also used for validation of remote sensing products [270, 272] since multi- or hyperspectral R_{rs} can be converted to FU through spectral convolution (Chapter 5). Use of the FU scale in citizen science is discussed further in Chapters 2 and 4.

Smartphone cameras are particularly promising as low-cost spectroradiometers. Recent years have seen a flurry of smartphone-based research in environmental science [273] and many other fields (Section 3.1). Both professional and citizen science benefit from the high-end cameras in modern smartphones. As discussed in Sections 1.2 and 3.2, smartphone and other consumer cameras typically have three spectral bands, red-green-blue or RGB, covering the visible range. These capabilities can be expanded with spectroscopic add-ons [98, 143]. For example, the iSPEX add-on enabled spectropolarimetry on smartphones and, as mentioned above, was used to crowdsource AOD [94]. Citizen scientists participated out of an interest in the science and a desire to contribute to the field [256]. Unfortunately, the quality of individual observations was limited by software restrictions on the smartphone cameras, although the uncertainty was reduced by averaging over multiple devices. Furthermore, the add-on did not fit on newer smartphones. In water research, the HydroColor and EyeOnWater apps have been used by thousands of professional and citizen scientists to measure R_{rs} and water colour [121, 274]. Validation studies of both have shown agreement with reference instruments, but have also found large inaccuracies and uncertainties upwards of 30% [275, 276]. In all three examples, issues of data quality stem from the need to use JPEG data, which are heavily compressed and non-linear. These limitations are addressed in this thesis in Chapters 3, 4, and 6.

Achieving a high impact with citizen science requires trust in the robustness of methods, data, and results [277]. Trust can be achieved through standardisation of the data acquisition and processing, for example by providing participants with standardised equipment [77].

[6] Better known among astronomers for his contributions to spectroscopy and solar physics and his catalogue of double stars [259].

[7] https://monocle-h2020.eu/Sensors_and_services/Mini-secchi_disk

Keeping the methodology similar to that of professional measurements, such as HydroColor following the standard (Mobley) protocol for R_{rs} measurements [121, 209], further increases trust. At the same time, data quality is highest when the methodology is simple [254, 277]. Trust in results requires consistent reporting, using standard quantities [77] and including uncertainties [278], and rigorous validation by participants and third parties [275, 276]. Citizen scientists' trust in results is also influenced by interpersonal factors [279]. Well-designed citizen science projects can achieve a data quality similar to professional science [277]. This thesis discusses standardisation, data quality, and uncertainty in citizen science data in Chapters 2, 3, and 4.

1.5 This thesis: Accessible remote sensing of water

The aim of this thesis is to investigate and improve accessibility and uncertainty in remote sensing and citizen science, so that these techniques can better deliver the desired improvements to cost, scale, and reproducibility of water research (Section 1.1.4). Accessibility is improved by accounting for disability in citizen science (Chapter 2) and by developing consumer cameras as low-cost instruments for remote sensing by professional and citizen scientists (Chapters 3, 4, and 6). Increased accessibility is associated with a lower cost and wider scale of water research. We assess and reduce the uncertainty in measurements by citizen scientists and measurements from consumer cameras (Chapters 2, 3, and 4), as well as the uncertainty in comparing and combining data from different professional instruments (Chapter 5). Better characterised and reduced uncertainty leads to higher reproducibility. Chapter 7 provides a general synthesis, discussion, and future outlook. Short summaries of each chapter are provided below.

The research described in this thesis was performed over a 4-year period, from 2018 to 2022, within the MONOCLE project. Due to various internal and external factors, most prominently the COVID-19 pandemic, our original plans had to be adapted multiple times and we were unable to perform much of the planned field work and citizen science experiments. As a result, the research topics of Chapters 2–6 are quite diverse. However, as described above, they all deal with different aspects of two core topics, namely accessibility and uncertainty in remote sensing of water. The chapters are arranged by the dimensionality of their respective data, from multispectral (Chapters 2–4) through hyperspectral (Chapter 5) to spectropolarimetric (Chapter 6), but can be read in any order.

Chapter 2. Citizen science with colour blindness: A case study on the Forel-Ule scale. Colour vision is often used in citizen science, including Forel-Ule measurements (Section 1.4). However, up to 1 in 11 participants have deficient colour vision (Section 1.2.2). In this chapter, we simulate the effects of colour blindness on Forel-Ule measurements to quantify the resulting uncertainty. Based on the results, we suggest likely effects on participant motivation and retention, and offer general recommendations on improving accessibility and quantifying uncertainty in citizen science [280].

Chapter 3. Standardised spectral and radiometric calibration of consumer cameras. Consumer cameras, particularly onboard smartphones and UAVs, are now commonly used as scientific instruments (Section 1.4). However, their data processing pipelines are not optimised for quantitative radiometry and their calibration is more complex than that of scientific

cameras (Section 1.2). The lack of a standardised calibration methodology limits the interoperability between devices and, in the ever-changing market, ultimately the lifespan of projects using them. In this chapter, we present a standardised methodology and database (SPECTACLE) for spectral and radiometric calibrations of consumer cameras. The methodology is applied to seven popular cameras to characterise their performance and quantify uncertainties. Through calibration and the use of RAW data, we lay the groundwork for using consumer cameras to perform professional-grade spectroradiometry [281].

Chapter 4. Accuracy and reproducibility of above-water radiometry with calibrated smartphone cameras using RAW data.
In this chapter, we apply the results and recommendations from Chapter 3 to above-water radiometry. We measure R_{rs} and water colour on and around Lake Balaton, Hungary, with two smartphones and two hyperspectral reference instruments for validation (Section 1.3). We quantify the uncertainty, reproducibility, and accuracy of the resulting data and compare these to professional spectroradiometers and existing citizen science approaches (Sections 1.2 and 1.4). We find that smartphone cameras perform similarly to professional instruments. The typical uncertainty of individual smartphone measurements is ≤5% in radiance, 1.9–8.1% in R_{rs}, and 0.5–1.9% in R_{rs} band ratios. The typical difference between two smartphones is 6.9% in radiance, 5.5% in R_{rs}, and 2.9% in R_{rs} band ratios. Lastly, the typical difference between smartphone and reference data is ≤19% in radiance, ≤13% in R_{rs}, and ≤1.9% in R_{rs} band ratios, barring one outlier. Based on these results, we offer practical recommendations for using consumer cameras in professional and citizen science [282].

Chapter 5. Biases from incorrect reflectance convolution.
Reflectance measurements from hyperspectral instruments are converted to multispectral bands (Section 1.2.2) through spectral convolution. This is done to combine time series, validate instruments, and apply retrieval algorithms (Section 1.3). However, convolution is often done incorrectly, with reflectance itself convolved rather than the underlying (ir)radiances. In this chapter, the resulting error is quantified for simulated and real instruments. Significant biases, up to 5%, are found. Based on these results, we suggest that this error was partially responsible for uncertainties found in previous work and recommend that future work apply spectral convolution correctly [283].

Chapter 6. A universal smartphone add-on for portable spectroscopy and polarimetry: iSPEX 2.
In this chapter, we present a new smartphone spectropolarimeter, iSPEX 2. It succeeds the original iSPEX add-on (Section 1.4). The optical design is presented and the manufacturing process is detailed. Through universal smartphone support and a data processing pipeline based on Chapter 3, we aim to improve the accessibility and data quality compared to the original. An initial validation measurement is presented as a proof of concept. Finally, we suggest possible applications of iSPEX 2 for professional and citizen science [284].

Chapter 7. General discussion and future outlook.
In this chapter, we synthesise the overall findings of Chapters 2–6 and discuss them relative to the current state and future direction of the field. Initial findings from a follow-up project on spectropolarimetry of floating debris are also presented. Finally, the major conclusions and recommendations of this thesis are summarised.

2 | Citizen science with colour blindness: A case study on the Forel-Ule scale

Adapted from

Olivier Burggraaff, Sanjana Panchagnula, Frans Snik

PLOS ONE, 16(4), e0249755 (2021)

doi:10.1371/journal.pone.0249755

Many citizen science projects depend on colour vision. Examples include classification of soil or water types and biological monitoring. However, up to 1 in 11 participants are colour blind. We simulate the impact of various forms of colour blindness on measurements with the Forel-Ule scale, which is used to measure water colour by eye with a 21-colour scale. Colour blindness decreases the median discriminability between Forel-Ule colours by up to 33% and makes several colour pairs essentially indistinguishable. This reduces the precision and accuracy of citizen science data and the motivation of participants. These issues can be addressed by including uncertainty estimates in data entry forms and discussing colour blindness in training materials. These conclusions and recommendations apply to colour-based citizen science in general, including other classification and monitoring activities. Being inclusive of the colour blind increases both the social and scientific impact of citizen science.

2.1 Introduction

Colour measurements are common in citizen science. They are often done using red-green-blue (RGB) consumer cameras such as smartphones [121, 281, 285], but also with the human eye. Human colour measurements are used in such diverse fields as coral reef monitoring [286], snail evolution [287], soil surveying [288], climate adaptation [289], and water colour [265, 290, 291]. The data are expressed through a qualitative label [287] or by comparison with a colour chart [265, 286, 288, 290–292]. Colour is a useful proxy for underlying properties such as chemical composition [292, 293] and the simplicity of measuring with the eye enables low-cost measurements over large areas and long time series [271, 293].

Accessibility and inclusivity are key to successful citizen science [294, 295]. A large and diverse group of participants increases the social and scientific impact of citizen science [255, 256, 287, 294]. However, recruiting and retaining participants is challenging [95, 294–296]. Important motivations to participate are a feeling of contributing to science and environmental protection [95, 255, 256, 288, 296], learning [256, 289, 296], and simply having fun [95, 255, 256, 290, 294]. Common reasons to stop participating include mis- or not understanding the project [256], perceiving the data as not valuable [255, 286, 296], and difficulty in performing the measurements [288, 290, 296, 297].

While colour vision is often assumed to be universal, many differences exist between individuals. Colour blindness, or colour deficiency, affects up to 9% of men and 2% of women, depending on ethnicity and other genetic factors [155]. It reduces or even eliminates one's ability to distinguish certain colours, most commonly red and green [155]. Colour blindness is typically congenital [154, 155, 298–300], but can also be acquired through age or disease [300, 301].

Three forms of colour blindness exist, namely *anomalous trichromacy*, *dichromacy*, and *monochromacy*. Each affects the eye's three pigments in a different way. These pigments are labelled LMS for long-, medium-, and short-wave, respectively, with peak sensitivity wavelengths of 560, 530, and 420 nm [154]. In anomalous trichromacy, a single pigment has an atypical spectral response, reducing one's colour discrimination abilities [154, 298, 302]. This is called *protanomaly*, *deuteranomaly*, or *tritanomaly*, for the respective LMS pigments. Dichromacy is a complete lack of one pigment, similarly called *protanopia*, *deuteranopia*, or *tritanopia* [154]. Finally, monochromacy is a complete lack of multiple cones, causing a full lack of colour vision. Monochromacy is exceedingly rare [154, 299] and is not discussed further in this work.

Colour blindness is often treated as a continuous spectrum from regular colour vision (all pigments present and typical) through degrees of anomalous trichromacy (one pigment atypical) to dichromacy (one pigment wholly missing) [154, 155, 301]. For simplicity, the three LMS deficiencies are referred to as *protan*, *deutan*, and *tritan*, respectively [155]. Protan and deutan are the most common, affecting for example up to 9% of men and 0.6% of women in Europe, as well as 7% of men and 2% of women in China [155]. The prevalence of tritan in the West is on the order of 1:10 000 [300], though higher prevalences have been reported in other locations [303].

Colour blindness limits the accessibility of citizen science that involves colour measurements for up to 1 in 11 participants. However, to our knowledge, little research has gone into its potentially far-reaching consequences. Such work has been done for science communication, for example in designing inclusive colour maps [302, 304].

As a case study, we investigate the impact of colour blindness on water colour measure-

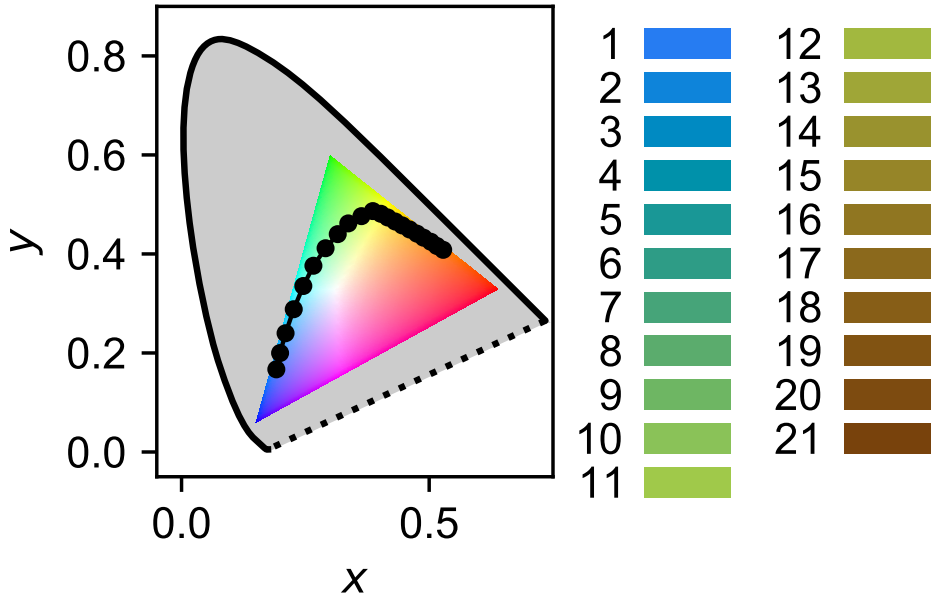


Figure 2.1: The Forel-Ule scale. The individual FU colours are shown on the right, a comparison to the human gamut on the left. The gamut is plotted in (x, y) chromaticity, normalized from CIE XYZ and shown with a constant brightness, and converted to sRGB. The FU scale increases from 1 (bottom left) to 21 (far right). The shaded area represents the full gamut of regular colour vision, while the coloured triangle represents the sRGB colour space, which most computer monitors are limited to. The perceived colours may vary depending on monitor or printer settings and the reader's own colour vision.

ments with the Forel-Ule (FU) scale. This scale quantifies human water colour measurements [305] by assigning a numerical value from 1–21 to a predetermined set of colours, shown in Figure 2.1. These range from indigo blue (FU 1) through green (FU 11) to cola brown (FU 21). First used in the 1890s by Forel and Ule [267, 268], it provides the longest continuous record of ocean colour [271]. For instance, Wernand and Van der Woerd used 17 171 archival FU measurements from 1930 to 1999 to derive long-term biogeochemical trends in the Pacific Ocean [293]. Properties of a water body that can be derived from its colour include suspended particles, dissolved organic matter, and algal pigments such as chlorophyll-a [10, 271, 293].

The FU scale is commonly used by professionals [272, 293] and by citizen scientists [265, 291]. Measurements are done by comparing a physical standard colour scale to a water body. For citizen science, the original scale made from 21 vials of pigment mixes [306] may be replaced with plastic filters [291] or a printed version [265], making it easier to use. Having this physical reference reduces the effects of variations in illumination, though in all cases it is difficult to guarantee colour consistency.

We use simulations to determine the effects of colour blindness on FU measurements. Such digital simulations accurately reproduce colour blind vision [302, 307]. The discriminability of the resulting shifted colours is assessed using the CIE ΔE_{00} colour difference measure [308]. This way, the impact of colour blindness on FU measurements is quantified.

Based on these results, we make general recommendations for dealing with colour blindness in citizen science. These include guidelines for data entry protocols and training materials, benefiting citizen motivation and data quality. Moreover, the methods applied in this work are easily generalized to other colour-based tools. This enables authors to account for colour blindness in the design stage of new citizen science projects. While some projects have opted for simplified colour scales [292], this significantly reduces the information content [10] of all data, including those from colour blind participants. Simplified colour scales are thus generally not an ideal solution to this problem.

Section 2.2 describes the methods used to simulate colour deficiency and assess colour discriminability. Results are presented in Section 2.3 and discussed in Section 2.4. Finally, conclusions and recommendations are drawn up in Section 2.5.

2.2 Methods

The colour blindness simulations and analysis were implemented in custom Python scripts available from GitHub[8].

2.2.1 Forel-Ule scale

Tristimulus (CIE XYZ) values for the FU scale were derived by Novoa et al. from transmission spectroscopy [306]. The corresponding (x, y) chromaticities are shown in Figure 2.1.

Four illuminants were considered, namely E (equal-energy) and D55, D65, and D75 (daylight). These illuminants quantify differences in lighting conditions and are used to express colour appearance in a standardised manner [101]. The FU scale is defined with an E illuminant [306] but measurements take place in daylight, making D-type illuminants more representative [291]. Conversion between illuminants was done in XYZ space using the Bradford chromatic adaptation matrices provided on Bruce Lindbloom's website [309].

The tristimulus values were first converted to the LMS colour space, representing the relative excitations of the LMS cones [302, 310]. This was done through the Hunt-Pointer-Estevez matrix [310], as shown in Equation (2.1). Here $[\,L\;M\;S\,]^T$ and $[\,X\;Y\;Z\,]^T$ are the vector representations of a single colour in LMS and XYZ, respectively.

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.38971 & 0.68898 & -0.07868 \\ -0.22981 & 1.18340 & 0.04641 \\ 0.00000 & 0.00000 & 1.00000 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{2.1}$$

2.2.2 Simulation of colour blindness

Colour blindness was simulated by mapping colours from the LMS colour space representing regular vision to a reduced colour space representing colour deficiency [302,307,311]. This is a mathematical representation of how colour appearances shift due to colour blindness, based on the observed colour perceptions of dichromats [307]. Since for dichromats and anomalous trichromats, two out of three cones are unaffected, the responses of those cones to a given colour are unchanged. The simulation determines the response of the third, deficient cone that imitates for a regular observer the colour perceived by a colour blind person [307, 311].

[8] https://github.com/burggraaff/cbfu

This in turn allows us to apply discriminability metrics developed for regular colour vision to the simulated perceived colours.

The LMS-space vectors $\vec{c}_L$ were modified using a cone-deficiency transfer matrix $\mathbf{T_k}$. $\mathbf{T_k}$ is the identity matrix $\mathbf{I_3}$ with one diagonal element (T_k^{00}, T_k^{11}, T_k^{22} for protan, deutan, tritan, respectively) reduced to a relative cone contribution k. This is shown in Equations (2.2) and (2.3) for protan with its respective matrix $\mathbf{T_k^p}$ [302,311]. k ranges continuously from 1 (regular vision) to 0 (dichromacy). It represents the relative contribution of a specific cone to colour vision but does not correspond directly to a physical property of the eye. The elements q_1 and q_2 of $\mathbf{T_k}$ shift the response from the deficient cone (L in the example) to the others.

$$\begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} = \begin{bmatrix} k & q_1^p & q_2^p \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} L \\ M \\ S \end{bmatrix} \tag{2.2}$$

$$\vec{c}_L' = \mathbf{T_k^p} \vec{c}_L \tag{2.3}$$

The cone transfer matrices for protan $\mathbf{T_k^p}$, deutan $\mathbf{T_k^d}$, and tritan $\mathbf{T_k^t}$ are as follows:

$$\mathbf{T_k^p} = \begin{bmatrix} k & q_1^p & q_2^p \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{T_k^d} = \begin{bmatrix} 1 & 0 & 0 \\ q_1^d & k & q_2^d \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{T_k^t} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ q_1^t & q_2^t & k \end{bmatrix} \tag{2.4}$$

The elements q_1, q_2 were determined by noting that colour blind people retain regular vision for white and a complementary colour (blue for protan and deutan, red for tritan) [302, 307, 311]. In other words, $\mathbf{T_k}$ has eigenvectors $\vec{w}_L = [1\ 1\ 1]^T$ (white) and either $\vec{b}_L$ (blue) or $\vec{r}_L$ (red) with eigenvalues 1. This is shown in Equations (2.5) and (2.6).

$$\mathbf{T_k^p} \vec{b}_L = \vec{b}_L \qquad \mathbf{T_k^d} \vec{b}_L = \vec{b}_L \qquad \mathbf{T_k^t} \vec{r}_L = \vec{r}_L \tag{2.5}$$

$$\mathbf{T_k^p} \vec{w}_L = \vec{w}_L \qquad \mathbf{T_k^d} \vec{w}_L = \vec{w}_L \qquad \mathbf{T_k^t} \vec{w}_L = \vec{w}_L \tag{2.6}$$

For each case, a system of two equations with two unknowns q_1, q_2 and one variable k was derived, with L_b, M_b, S_b the LMS coordinates of the blue reference vector $\vec{b}_L$ and L_r, M_r, S_r those of $\vec{r}_L$:

$$kL_b + q_1^p M_b + q_2^p S_b = L_b \qquad kM_b + q_1^d L_b + q_2^d S_b = M_b \qquad kS_r + q_1^t L_r + q_2^t M_r = S_r \tag{2.7}$$

$$k + q_1^p + q_2^p = 1 \qquad k + q_1^d + q_2^d = 1 \qquad k + q_1^t + q_2^t = 1 \tag{2.8}$$

Solving for q_1, q_2 gave the following expressions:

$$q_1^p = 1 - k - q_2^p \qquad q_1^d = 1 - k - q_2^d \qquad q_1^t = 1 - k - q_2^t \tag{2.9}$$

$$q_2^p = (1-k)\frac{M_b - L_b}{M_b - S_b} \qquad q_2^d = (1-k)\frac{L_b - M_b}{L_b - S_b} \qquad q_2^t = (1-k)\frac{L_r - S_r}{L_r - M_r} \tag{2.10}$$

The sRGB blue and red primaries are typically used for $\vec{b}_L$ and $\vec{r}_L$, respectively, as this technique is used in the field of computer graphics [302, 311]. While other primaries could

be used, such as monochromatic wavelengths [307], this makes little difference [302] so we followed the convention.

We calculated $\mathbf{T_k}$ for protan, deutan, and tritan with $1 \geq k \geq 0$ in intervals of 0.01, and transformed the 21 FU colours with each $\mathbf{T_k}$. The modified vectors were then transformed back to XYZ and analyzed. This was implemented in Python through NumPy's `einsum` method [312].

2.2.3 Colour discrimination

Discriminability of the transformed FU colours was assessed in the CIE Lab (1976) colour space. CIE Lab is approximately perceptually uniform, its components representing lightness (L^*), green-red (a^*), and blue-yellow (b^*) [101]. While FU colour assignment is typically done in (x, y) chromaticity (normalized XYZ) through the hue angle [10, 285], this approach does not work for dichromacy, which reduces the chromaticity plane to a line [307]. The Euclidean distance in XYZ coordinates also could not be used, as XYZ is not perceptually uniform [313].

Discriminability was quantified through the ΔE_{00} metric [308], which expresses the difference between colour pairs. The full formula for ΔE_{00} is given in [308] and not reprinted here due to its length; our Python implementation passed all the example cases in said paper. A value of $\Delta E_{00} = 2.3$ corresponds to a just-noticeable difference (JND), the smallest difference an average observer can distinguish [101, 314].

For each deficiency simulation, the ΔE_{00} difference between each of the 21 transformed FU colours was calculated, giving a 21×21 confusion matrix. In this, any colour pairs where $\Delta E_{00} < 1$ JND cannot be discriminated at all, while pairs with $1 \leq \Delta E_{00} \leq 3$ JND are discriminable with difficulty. Pairs with $\Delta E_{00} > 3$ JND were considered discriminable.

2.3 Results

2.3.1 Colour blindness simulation

The appearance of the FU scale with varying degrees of colour blindness, simulated as in Section 2.2.2, is shown in Figure 2.2. The observed changes qualitatively match those seen in previous work [302, 311] and were anecdotally confirmed by one of the authors (deuteranomalous) and a colleague (protanopic). The largest colour shifts are seen for tritan, as expected since it affects the perception of blue light and many FU colours are shades of blue.

Colour blindness narrows the gamut of the FU scale, as shown in Figure 2.3. It has little effect on the lightness (L^*) of the FU scale but affects its colour components. Protan and deutan (red-green blindness) reduce the range of a^* (red-green) while tritan reduces the range of b^* (blue-yellow). These shifts imply that colour blindness reduces the ability to discriminate FU colours based on hue, meaning the user will have to rely more on lightness.

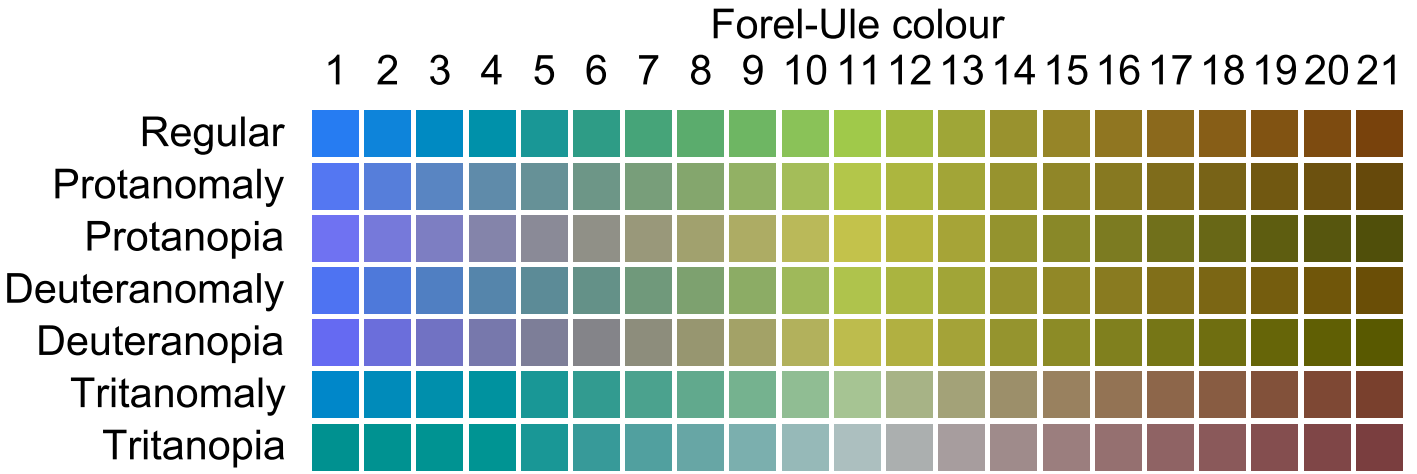


Figure 2.2: Apparent Forel-Ule colours with regular and deficient colour vision. The (modified) XYZ coordinates were adapted to a D65 illuminant, then converted to the sRGB colour space and gamma expanded [281] for visualization purposes. The perceived colours may vary depending on monitor or printer settings and the reader's own colour vision. Readers who cannot distinguish between the colours shown here may benefit from taking a colour vision test; many variants are freely available online. The anomalous examples correspond to $k = 0.50$.

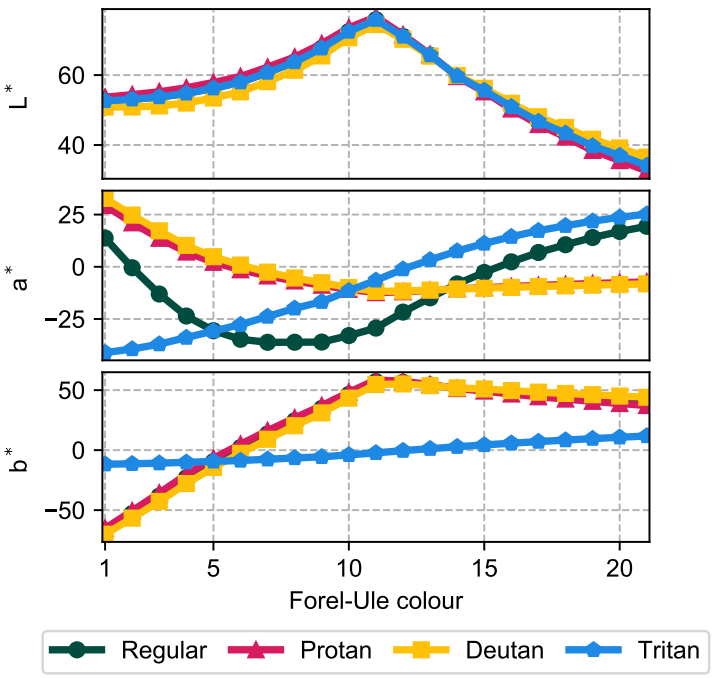


Figure 2.3: Forel-Ule colours in CIE Lab space. Both regular and deficient vision are included. Regular vision is hidden in the top and bottom panels behind protan and deutan. These affect the a* (green-red) coordinate the most while tritan affects b* (blue-yellow) the most. None of the deficiencies significantly affect L* (lightness).

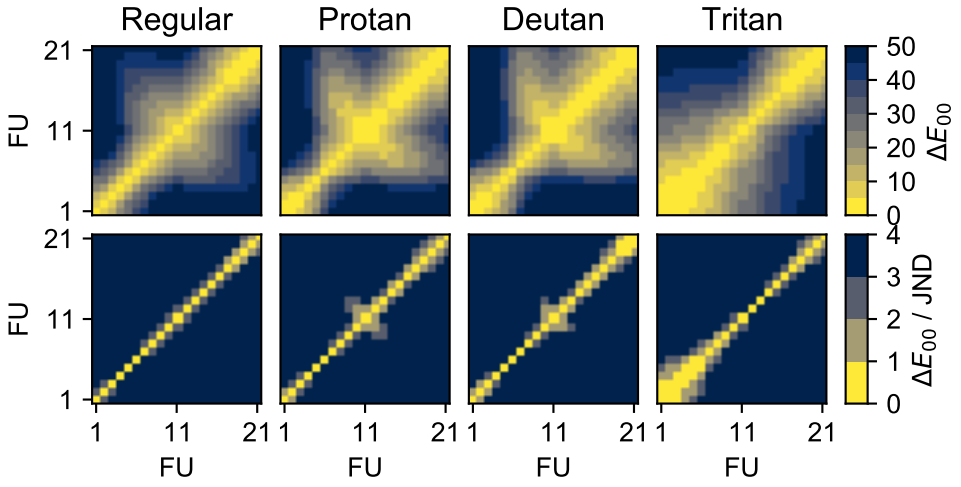


Figure 2.4: Confusion matrices for regular and deficient colour vision. The top panels show the full range of ΔE_{00}, while the bottom panels have a narrower colour bar, in units of just-noticeable difference (JND, $\Delta E_{00} = 2.3$). Even with regular vision, some pairs of FU colours are difficult to distinguish ($\Delta E_{00} \leq 3$ JND) Protan and deutan primarily decrease the discriminability of the middle (green) and high (brown) colours, while tritan primarily affects the low (blue) colours, as expected.

2.3.2 Colour discrimination

The discriminability of FU colours is reduced by colour blindness. The confusion matrices for regular and deficient vision, calculated as in Section 2.2.3, are shown in Figure 2.4. They show that the reduced range in a^* (red-green) for protan and deutan and in b^* (blue-yellow) for tritan, observed in Section 2.3.1, reduce the discriminability at opposite ends of the FU scale. The former primarily affect FU 10–21 (green–brown) while tritan affects FU 1–9 (blue–green).

Several pairs of FU colours become fully indistinguishable. Deuteranopia causes two colour pairs (FU 19-20 and 20-21) to fall within 1 JND and thus become indistinguishable. For tritanopia, six pairs become indistinguishable, namely 1-2, 1-3, 2-3, 3-4, 4-5, and 5-6. Protanopia does not cause indistinguishable pairs.

Additionally, many more pairs exhibit reduced discriminability. While most adjacent pairs are <3 JND apart even with regular colour vision, deficiency extends this further off the diagonal. In particular, protan and deutan cause confusion between the central colours (FU 9–13), which is also apparent from Figure 2.3 as they have similar L^*, a^*, and b^*. On the other hand, tritan significantly reduces the discriminability of FU 1–9. As seen in Figure 2.5, the number of pairs within 3 JND increases from 17 (regular) to 24 (protanopia), 21/24 (deuteranopia/deuteranomaly), or 30 (tritanopia).

These trends also apply to partial colour blindness (anomalous trichromacy). Figure 2.5 shows the relation between k and median/minimum ΔE_{00} as well as the number of indistinguishable pairs. The median decreases smoothly for protan, deutan, and tritan (from 33 to 27, 26, and 22, respectively) from $k = 1$ to 0. The minimum ΔE_{00} decreases smoothly for protan and deutan (from 3.3 to 2.5 and 2.2, respectively) while the tritan curve is piecewise smooth. Fully indistinguishable pairs ($\Delta E_{00} < 2.3$) appear at $k \leq 0.20$ for deutan and tritan.

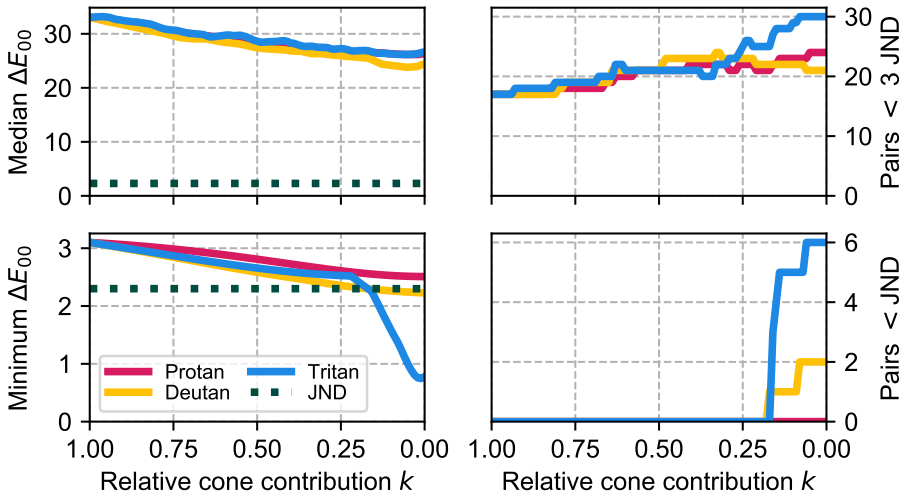


Figure 2.5: Discriminability of Forel-Ule colours. The median and minimum (left) ΔE_{00} difference between FU colour pairs, and the number of pairs within 3 and 1 JND (right), are shown as a function of the relative cone contribution k. k ranges from 1 (full colour vision) to 0 (dichromacy), with intermediate values representing partial colour blindness (anomalous trichromacy). Pairs with $\Delta E_{00} < 1$ JND are fully indistinguishable, pairs with <3 JND are difficult to distinguish (Section 2.3.2).

Chromatic adaptation with a daylight illuminant (Section 2.2.1) did not affect these results. While the ΔE_{00} between some pairs changed by up to 1 JND, the patterns seen in Figures 2.4 and 2.5 remained, as did the previously discussed pairs of non-discriminable colours.

2.3.3 Practical consequences

In practice, FU measurements always have an uncertainty of ≥1 FU units. This is due to viewing conditions at the time of measurement including waves, specular reflections, and uneven illumination. As seen in Section 2.3.2, adjacent pairs of FU colours are difficult to distinguish ($\Delta E_{00} < 3$ JND) even with regular vision.

Colour blindness increases the uncertainty on FU measurements. Observers with protan or deutan experience increased difficulty in distinguishing adjacent pairs. Moreover, protans have difficulty distinguishing FU 9–13 while for deutans, FU 19-20 and 20-21 are completely indistinguishable. For a FU 20-type water body, a deutan cannot specify their observation more precisely than 19–21. Furthermore, $\Delta E_{00} = 2.33$ for FU 18 and 20, further reducing this precision to 18–21 given imperfect viewing conditions. Similarly, since tritans cannot distinguish six pairs of colours in the FU 1–6 range, they can provide little precision on 99% of global surface waters [10].

This increased uncertainty affects data quality and user motivation. This is further discussed in Section 2.4 and recommended guidelines for considering these issues are given in Section 2.5.

2.4 Discussion

Simulating the effects of colour blindness on Forel-Ule (FU) measurements, we have found significant reductions in colour discriminability and hence precision (Sections 2.3.2 and 2.3.3). This matches the authors' and colleagues' experiences in the field, and the simulation methods are well-attested in other contexts [302, 304, 307]. However, wider validation specific to the FU scale, with participants representing different types of colour blindness, is desirable.

The reduction in precision due to colour blindness reduces the quality and value of citizen science data. The magnitude of this effect depends on the type and severity of colour blindness, as described in Section 2.3.3. Protans and deutans, the vast majority of colour blind people in the West [155, 300], experience a reduction in median discriminability (ΔE_{00}) between FU colours of up to 21%; for tritans this is 33%. The uncertainty in FU data increases correspondingly, though not evenly. For example, tritans' ability to identify green-brown waters (FU 10–21) changes little, but they cannot distinguish the blue water types (FU 1–6) that represent most global surface waters [10].

This reduction in data value can be addressed by modifying data entry protocols to include uncertainties. Currently, many citizen science projects require users to provide a single value, for example FU 9 or 10. An entry field for uncertainty, or allowing the user to enter multiple values, accounts for the decrease in selectivity. Participants can estimate this uncertainty themselves. Even FU measurements by participants with regular colour vision have a typical uncertainty of ±1 FU (Section 2.3.2), which should be accounted for when using them to validate remote sensing data [272]. Colour blindness, particularly dichromacy, increases this uncertainty to up to ±3 FU.

We propose three methods to include uncertainties in data entry forms. The first is simply to include two fields, one for the best estimate (for example FU 9) and one for the estimated uncertainty (e.g. ±2 FU). This method is commonly used in scientific publications but it may be difficult for citizens to understand and apply [278], especially for asymmetric uncertainties. The second method is to have participants estimate a sequential range of possible values (e.g. FU 8–11), optionally including a single best estimate (e.g. FU 9). This is intuitive, simple to apply, and easily translated into traditional uncertainty intervals. It is most applicable for sequential scales like FU where confusion occurs primarily between adjacent numbers (Figure 2.4). The third method is to have participants select any number of possible values (e.g. FU 8, 9, 11). This is the most general method for discrete colour scales but makes the uncertainties more difficult to process. It is best suited to colour scales with many non-adjacent indistinguishable pairs. Our Python code (Section 2.2) can be adapted to other colour scales to determine which method is most suitable. A more detailed discussion on handling uncertainty in citizen science data is provided in [315].

Colour blindness can also affect the motivation of citizen scientists. As discussed in Section 2.1, participants need to feel they are contributing to science with valuable data. A participant presented with a colour scale where multiple colours appear indistinguishable may dismiss the method as either too difficult or nonsensical, and stop participating [255, 256, 286, 288, 290, 296, 297]. This is especially true for one unaware of their colour blindness. Since citizen science benefits from a large and diverse group of participants [255, 256, 287, 294], participant retention is important.

Demotivation can be prevented by modifying training materials. Explaining the choice of colour scale and how colour blindness affects its appearance helps participants understand the method. Particular care should be taken in emphasising the value of citizen data, even with colour blindness. For example, while tritans cannot distinguish the FU colours covering the open sea, their ability to distinguish FU 10–21 differs little from regular vision. These cover many inland waters [285, 316], which are commonly studied with the FU scale [305], so training materials should emphasise the value of tritans' observations there.

Training participants to estimate and provide uncertainties would further help them understand the value of their data [278]. Moreover, since uncertainty estimation is an integral part of professional science, citizen scientists may even gain motivation from learning about it [256, 289, 296]. For existing applications, if modifying data entry forms is impossible, explaining why colours may appear similar and how to pick a single colour would reduce the perceived difficulty.

The severity of these motivational effects and the efficacy of these preventative measures should be tested in practice. Comparing the retention of participants with regular and deficient colour vision, with and without modified training materials and data entry forms, would serve this purpose. This is ideally done in the design stage, as part of a co-creation process [256, 289].

Additional future work includes investigating the effects of other variations in colour perception. Even among those with regular colour vision, variations in colour perception exist [154], including demographical trends [155, 303]. Moreover, monochromacy was not discussed in this work because of its rarity [299] but likely has an even more pronounced effect on colour discriminability than the deficiencies investigated here.

Finally, unrelated to human observations, Figure 2.3 highlights the importance of lightness in distinguishing FU colours. Many FU index algorithms, which apply the FU scale to remote sensing data, only account for chromaticity [10, 272, 285]. Introducing lightness to

these algorithms may improve their precision and accuracy.

2.5 Conclusions & recommendations

Citizen science projects that depend on colour vision should account for colour blindness, which affects up to 1 in 11 participants. For Forel-Ule water colour measurements, colour blindness reduces the median discriminability between colours by up to 33% and makes multiple pairs of colours fully indistinguishable. This affects data quality and citizen motivation.

Modifying data entry forms to include uncertainty estimates would reduce the impact on data quality. This can be done by letting participants estimate the uncertainty in their measurement or choose multiple colours on the scale. Our provided Python code can be adapted to determine the best suited method for different colour scales. Learning how to estimate uncertainties may also increase participants' motivation and understanding of science.

The impact on motivation is reduced by including colour blindness in training materials. This includes explaining the colour scale and the difficulties colour blind participants may face, but also emphasising the continued value of their data. Through improved retention, this increases the number and diversity of the participants, which in turn increases both the social and scientific impact of citizen science.

Funding

European Commission Horizon 2020 programme (grant nr. 776480, MONOCLE).

Acknowledgements

The authors wish to thank Mortimer Werther, Steele Farnsworth, Emmanuel Boss, and Akupara Panchagnula for valuable discussions relating to this work. Data analysis and visualization were done using the Matplotlib, NumPy, and colorio libraries for Python. Figure 2.4 uses the *cividis* colour map [304]; the line colours in Figures 2.3 and 2.5 were obtained from https://davidmathlogic.com/colorblind/.

3 | Standardised spectral and radiometric calibration of consumer cameras

Adapted from

Olivier Burggraaff, Norbert Schmidt, Jaime Zamorano, Klaas Pauly, Sergio Pascual, Carlos Tapia, Evangelos Spyrakos, Frans Snik

Optics Express, 27(14), 19075–19101 (2019)

doi:10.1364/OE.27.019075

Consumer cameras, particularly onboard smartphones and UAVs, are now commonly used as scientific instruments. However, their data processing pipelines are not optimised for quantitative radiometry and their calibration is more complex than that of scientific cameras. The lack of a standardised calibration methodology limits the interoperability between devices and, in the ever-changing market, ultimately the lifespan of projects using them. We present a standardised methodology and database (SPECTACLE) for spectral and radiometric calibrations of consumer cameras, including linearity, bias variations, read-out noise, dark current, ISO speed and gain, flat-field, and RGB spectral response. This includes golden standard ground-truth methods and do-it-yourself methods suitable for non-experts. Applying this methodology to seven popular cameras, we found high linearity in RAW but not JPEG data, inter-pixel gain variations >400% correlated with large-scale bias and read-out noise patterns, non-trivial ISO speed normalisation functions, flat-field correction factors varying by up to 2.79 over the field of view, and both similarities and differences in spectral response. Moreover, these results differed wildly between camera models, highlighting the importance of standardisation and a centralised database.

3.1 Introduction

Consumer cameras have seen increasing scientific use in recent years. Their low cost makes them ideal for projects involving large scale deployment, autonomous monitoring, or citizen science. Successful scientific applications include environmental monitoring [56, 94, 121, 162, 276, 285, 317–323], cosmic ray detection [324], vegetation mapping [325–329], colour science [285, 330–335], and biomedical applications [336–343]. However, the use of consumer cameras is made difficult by limited software controls and camera specifications. Inter-calibration of multiple camera models is complex and laborious and the market constantly shifting, and for these reasons many projects are limited to only a few devices. These constraints severely affect both the data quality and the sustainability of projects using consumer cameras.

Smartphones, in particular, have become a common tool for research, thanks to their wide availability and features such as wireless connectivity. Many scientific applications (apps) using smartphone cameras have been developed, across a variety of fields. A recent example is HydroColor, a citizen science tool for remote sensing of water, specifically turbidity and remote sensing reflectance R_{rs}. These are derived from RGB colour photographs using standard inversion algorithms. Results from this app agree well with professional standard equipment, with mean errors in R_{rs} and turbidity $\leq$26% compared to reference sensors. However, due to software constraints, the app uses compressed JPEG data rather than raw sensor data and assumes identical spectral responses for all cameras. These factors severely limit the possible data quality. Nevertheless, HydroColor has already seen significant adoption by the community, and future developments may reduce the aforementioned limitations [121, 276, 318]. Another recent application of smartphone cameras is bioluminescent-based analyte quantitation by smartphone (BAQS), a technique for the detection of bioluminescent bacteria. Using BAQS, flux intensities down to the pW scale can be detected on some smartphone models; however, on others, software constraints and dark noise severely limit its sensitivity [344]. As a final example, Skandarajah et al. used smartphone cameras with conventional microscopes for micron-scale imaging, for example of stained blood samples. Resolutions comparable to scientific cameras were achieved, but intensity and colour measurements were limited by a lack of camera control and factors including non-linearity and white balance [342]. A full review of smartphone science is outside the scope of this work, and we instead refer the reader to a number of extensive reviews by other authors [143, 345–351].

Smartphone spectroscopy is an active field of development [98, 143]. Many spectroscopic add-ons have been developed, including do-it-yourself models costing less than $10 at Public Lab[9]. One early smartphone spectrometer was iSPEX, a spectropolarimetric add-on for iPhone devices used by >3000 citizen scientists to measure aerosol optical thickness (AOT) in the Netherlands in 2013. iSPEX data were found to agree well with reference sensors, with a correlation coefficient of $r = 0.81$ between AOT values observed with iSPEX and with the Moderate Resolution Imaging Spectroradiometer (MODIS) Aqua and Terra satellites [94]. However, the iSPEX data were limited in their polarimetric accuracy (absolute uncertainties in the degree of linear polarisation (DoLP) $\approx$ 0.03), preventing quantitative measurements of aerosol compositions and sizes [94]. This relatively large error stemmed from a lack of camera controls, such as the inability to fix the focus of the camera to a controlled and reproducible position. Furthermore, the sustainability of iSPEX in the fast-moving smartphone market was limited by its need for device-specific calibrations.

[9] https://publiclab.org/wiki/spectrometry

Consumer uncrewed aerial vehicles (UAVs) with RGB cameras have similarly become common scientific instruments. They provide a low-cost, high-resolution replacement for, or complement to, satellite and airplane imagery, especially for environmental monitoring [325–327, 352–354]. UAV data are increasingly being integrated with data from other platforms, such as satellites [355].

However, few scientific consumer camera projects progress past a proof-of-concept on a handful of camera models, which often become obsolete within two years, particularly in the constantly shifting smartphone market. This severely limits the sustainability of projects that require calibrations specific to each camera model. Difficulties in upscaling and future-proofing such calibrations are an oft cited constraint on the combination of multiple camera models [94, 319, 323, 324, 331, 349, 356]. Further complications are introduced by the lack of control over camera hardware and software parameters such as focus and white balance [94, 142, 320, 325, 342, 344]. For example, the dominant smartphone operating systems, Android and iOS, only introduced support for unprocessed (RAW) imagery as recently as 2014 (Android 5.0 'Lollipop') and 2016 (iOS 10). Previously, third-party developers could only use JPEG data, which introduce a number of systematic errors due to their lossy compression and bit-rate reduction [94, 121, 142, 317, 319, 334, 342, 349, 357]. Other common problems in consumer camera data include non-linearity and the gamma correction [94, 121, 162, 285, 317, 322, 329, 332, 334, 338, 342, 357–362], electronic and thermal noise [118, 319, 324, 344, 354, 363, 364], and highly variable (between camera models) spectral response functions which are not provided by manufacturers [121, 317, 323, 331, 333, 334, 349, 351, 354, 358, 365]. These factors limit the accuracy of radiometric measurements done with consumer cameras by introducing systematic errors. Furthermore, the accuracy of colour measurements and their conversion to standard measures, such as the CIE 1931 XYZ and CIELAB colour spaces, is limited by distortions in the observed colours [330] and differences in spectral response functions [331, 333–335].

Extensive (spectro-)radiometric calibrations of consumer cameras are laborious and require specialised equipment and are thus not commonly performed [333, 361, 366]. A notable exception is the spectral and absolute radiometric calibration of a Raspberry Pi 3 V2 webcam by Pagnutti et al. [118], including calibrations of linearity, exposure stability, thermal and electronic noise, flat-field, and spectral response. Using this absolute radiometric calibration, digital values can be converted into SI units of radiance. However, the authors noted the need to characterise a large number of these cameras before the results could be applied in general. Moreover, certain calibrations are device-dependent and would need to be done separately on each device. Spectral and radiometric calibrations of seven cameras, including the Raspberry Pi, are given in [358]. These calibrations include dark current, flat-fielding, linearity, and spectral characterisation. However, for the five digicams included in this work, JPEG data were used, severely limiting the quality and usefulness of these calibrations, as described above.

Spectral characterisations are more commonly published since these are vital for quantitative colour analysis. Using various methods, the spectral responses of several Canon [317, 329, 333, 334, 354, 358, 361, 367], Nikon [317, 323, 333, 334, 361, 365–368], Olympus [333, 334, 358], Panasonic [354], SeaLife [323], Sigma [366], and Sony [317, 333, 354, 358, 367] digital cameras (digicams), as well as a number of smartphones [121, 333], have been measured. Direct comparisons between >2 different camera models are given in [121, 317, 333, 354, 358]. Common features include the peak response wavelengths for the RGB colour filters, approximately 600, 520, and 470 nm, respectively, as well as a roughly Gaussian profile around the

peak. Differences are found especially in the wings, notably the locations of secondary peaks and near-infrared (NIR) and ultraviolet (UV) cut-off wavelengths. These may cause significant differences in observed colours between cameras [331,333], especially for narrow-band sources.

Camera calibrations in the literature are often limited to a small number of cameras or properties, either to narrow down the scope or because of limitations in time and equipment. Furthermore, calibration data are published in varying formats, locations, and quality, complicating their use by others. Standardised formats exist, such as those for vignetting, bias, and colour corrections described in Adobe's digital negative (DNG) standard [369], but have seen limited adoption. The European Machine Vision Association (EMVA) standard 1288 [370] for characterisation of cameras is extremely thorough, but has also seen limited adoption due to the high-end equipment required [361] and its scope simply being too broad for many practical purposes. Similarly, standardised data sets or databases, for example containing spectral response curves [333,366], have been created but these are limited in scope and, again, adoption. To our knowledge, no widely adopted standardised methodology or centralised database containing spectral and radiometric calibration data for consumer cameras has been created thus far.

In this work, we present a standardised methodology for the calibration of consumer cameras and a database, SPECTACLE (Standardised Photographic Equipment Calibration Technique And CataLoguE), containing calibration data for the most popular devices. The calibration methodology is focused on simplicity and facilitating measurements by non-experts and those lacking expensive equipment, similarly to [361] but with a broader scope including software, optics, and sensor characteristics. The database is designed with openness and sustainability in mind, focusing on community contributions. Furthermore, we strive to follow previously existing standards, such as DNG [369] and EMVA 1288 [370], where practical. Our focus is on radiometric and photometric measurements but these calibration data can equally be used for colour science purposes, in particular to convert between colour spaces using the measured spectral response curves. We stress that we have no financial nor commercial interests in consumer cameras, and any comparison between devices is purely scientific. The aim of our standardised methodology and the SPECTACLE database is merely to simplify the use of data from consumer cameras, not to cast judgment on their quality.

Section 3.2 contains an overview of hardware and software trends in consumer cameras. We present the standardised calibration methodology in Section 3.3. Section 3.4 contains results from its application to several popular cameras and a description of the SPECTACLE database. Finally, in Section 3.5 we compare our findings with previous work and discuss future perspectives.

3.2 Trends in consumer cameras

Consumer cameras can be divided into four categories, namely smartphones, UAVs, digicams (DSLR and mirrorless), and webcams. Despite serving very diverse purposes, these cameras share common characteristics and can be calibrated with the same methods.

CMOS-based sensors now dominate the consumer camera market [98]. These are often not produced in-house by camera manufacturers, but acquired from external parties, such as Sony and Samsung. Different cameras often use the same sensor, such as the Sony IMX298 which is used in 12 smartphone models from 10 different manufacturers.

Most colour cameras use Bayer filters, on-chip RGB filters arranged in a checkerboard pattern, with two green pixels (G and G_2) for every red or blue one [138]. The spectral responses of these filters differ strongly between cameras and are further modified by fore-optics [331]. Alternate pixelated filter arrangements exist, intended for example to reduce aliasing, but with little adoption so far [371]. Data from the separate $RGBG_2$ pixels can be recombined through a process known as demosaicking to retrieve an image with interpolated RGB values for each pixel. Many different schemes exist for this [371], ranging from simple bilinear interpolation [118, 330, 354] to complex computational methods [140]. Consumer camera software often includes proprietary demosaicking algorithms [329, 342] which may introduce complex, unpredictable effects. Depending on their implementation, demosaicking schemes typically mix data from different filters and remove their mutual independence, leading to undesirable cross-feed effects [118, 121]. In any case, the added data are fully synthetic and thus do not offer any new radiometric information. It is thus preferable for radiometric applications to treat the $RGBG_2$ images completely independently [329] and demosaic data for visualisation purposes [118] only.

As discussed previously, the most commonly used digital file formats are JPEG (or JPG) and RAW. In both formats, data are saved on a pixel-by-pixel basis in analogue-digital units (ADU). ADU are alternately referred to as digital numbers (DN) in the literature, but in this work we will use the ADU nomenclature. JPEG (ISO 10918) is based on lossy spatial compression and downsampling to 8-bit values, optimal for small file sizes while maintaining aesthetic qualities. Due to camera-specific processing and compression artefacts, JPEG images lose information and are not recommended for quantitative analysis [94, 121, 142, 317, 319, 329, 334, 342, 349, 357]. While standardisations exist, such as the standard Red Green Blue (sRGB) colour space and gamma curve [369], these are not strictly adhered to and cannot be assumed in data processing [372]. Conversely, RAW files contain relatively unprocessed sensor output, intended for manual post-processing. One factor complicating the reduction of RAW data is their mosaicked nature, due to which they must be demosaicked or treated as multiple independent images, as discussed above. Despite these complications, their unprocessed nature makes RAW data highly preferable for scientific purposes [118, 121, 329, 331, 334, 349].

Available camera controls generally include focus, exposure time, ISO speed (sensitivity), and aperture. Focus and aperture are changed by physical movement of camera optics, though most webcams and smartphones only allow a single, fixed aperture. ISO speed is set by changing the camera gain, through analogue amplification or digital processing. Analogue amplification involves varying the gain of the CMOS amplifiers, which can be done on the level of individual pixels. Conversely, digital gain is implemented in post-processing by simply re-scaling and interpolating measured digital values. Since ISO speed is a measure of the overall sensitivity of the camera, including fore-optics, each camera (and possibly each pixel) has a unique relation between ISO speed and gain. Finally, exposure time may be set by a physical shutter (common in digicams) or an electronic one (common in smartphones). Other parameters like white balance only affect processed imagery and are not relevant to RAW photography.

Many cameras include some built-in calibrations, most notably for non-linearity, dark current, and flat-fielding effects. Non-linearity corrections are typically based on previously measured correction curves [104]. Dark current corrections (autodarking) are commonly done using unilluminated exposures or permanently dark pixels around the sensor. Finally, flat-fielding (specifically vignetting) is typically corrected using a pre-made correction map.

A variety of methods for generating such maps exists, based for example on computational methods using regular photographs [330, 373–376], simply averaging many exposures [329], and simply imaging white paper [377]. These maps are typically parametrised, for which various methods also exist [118, 329, 330, 369, 374–377], the simplest being the $\cos^4$ model, a combination of inverse square falloff, Lambert's law, and foreshortening [374]. Alternately, a pixel-by-pixel map of vignetting correction coefficients may be used. Such maps may be device-specific or generalised for a camera model. Notably, iOS-based smartphones use the seven-parameter parametrisation described in the DNG standard [369] (see Section 3.3.8) while Android-based smartphones use pixel-by-pixel maps.

3.2.1 Smartphones

The smartphone market has become remarkably homogeneous in recent years, with virtually all models using the slate form factor, featuring a large touch screen, few buttons, and a camera on either side of the device. The most popular smartphones are all iOS- or Android-based. Both these operating systems now support RAW photography using Adobe's DNG standard [369], though not on all devices. Hardware properties are rarely released by manufacturers, and are instead often provided by reviewers through disassembly of the smartphone.

Smartphone cameras aim to reproduce the human eye and thus have similar optical properties [121]. Sensors, most commonly from the Sony Exmor series, are compact with 12–16 megapixels and a diagonal of 5–8 mm. Some devices focus on high-resolution imagery with many pixels, while others are optimised for low-light conditions, with fewer but larger pixels.

Smartphones now increasingly have multiple rear cameras. These secondary cameras offer features such as different fixed focal lengths and higher sensitivity, for example with a different lens or a monochromatic sensor. All rear cameras are typically placed in a cluster at the top right or top center of the smartphone.

3.3 Methods

In this section we describe the standardised methods for the calibration of consumer cameras. We developed a custom data processing pipeline, implemented in Python scripts available on GitHub[10] and iOS[11] and Android apps[12].

Section 3.3.1 describes the experimental setups and data processing used in calibration measurements. The methods used to characterise and calibrate the camera responses are given in Sections 3.3.2–3.3.9. Finally Section 3.3.10 describes how consumer camera data are converted into relative radiometric units using the previously described calibration measurements. These units provide a constant scale, independent of exposure parameters and individual device characteristics, for each camera model, a constant factor K per model away from absolute radiometric units (W m^{-2} sr^{-1}). Absolute radiometric calibration is outside the scope of this work.

[10] https://github.com/monocle-h2020/camera_calibration
[11] https://github.com/monocle-h2020/spectacle_ios
[12] https://github.com/monocle-h2020/spectacle_android

3.3.1 Experimental setup

This section describes the setups used in our golden standard ground-truth experiments. Descriptions of do-it-yourself (DIY) calibration methods are given in the relevant sections. All images from all cameras were taken in RAW format; for the linearity measurements, simultaneous RAW and JPEG images were taken for comparison. As discussed in Section 3.2, demosaicking schemes introduce synthetic data and undesirable cross-feed effects. For this reason, in our data reduction the RAW images were split into separate $RGBG_2$ images which were analysed individually [329]. Multiple images were taken and stacked for each measurement to improve the signal-to-noise ratio (SNR). On smartphones, the aforementioned iOS and Android apps were used to control the camera and automatically take multiple exposures. Exposure settings, including ISO speeds and exposure times were obtained from camera controls where possible, since EXIF metadata values for these were found (Section 3.4.1) to be unreliable.

The setup for measuring linearity, ISO-gain relations, and inter-pixel gain variations on smartphones is shown in Figure 3.1. A halogen light source (OceanOptics HL-2000-LL) was used, specified by the manufacturer to be stable to 0.15% peak-to-peak and drift <0.3% per hour after a warm-up of 10 minutes. Its light was fed into an optical fiber (Thorlabs M25L02) and collimated using two lenses (Thorlabs AC254-030-A with $f = 30$ mm and AC508-200-A with $f = 200$ mm). Two linear polarisers (both Thorlabs LPVISE100-A, with an extinction ratio ≥495 from 400–700 nm), the first rotatable and the second fixed, were used to attenuate the light beam entering an integrating sphere (Thorlabs IS200). Using Malus's law ($I = I_0 \cos^2 \theta$), the rotation angle between the polarisers could be used to calculate the attenuation. A calibration detector was not necessary since all experiments done with this setup involve relative measurements only. Malus's law was first fitted to a series of exposures over the entire rotation range to determine the reference angle. The rotation angle of the first polariser could be determined visually up to 2° precision, giving a typical uncertainty on the attenuated intensity of 2.5%. Finally, smartphones were placed on top of the integrating sphere, flush against the view-port. The farthest possible focus was used (infinity on Android devices, an arbitrary number on iOS). All experiments done with this setup involved analysis on the individual pixel and (broad-band) filter level, without any spatial averaging. Because of this, differences in illumination due to spectral dependencies in the polariser throughput or the integrating sphere output did not affect any of the experiments.

The linear polarisers can be replaced with alternate methods for attenuation, such as neutral density filters. Attenuation can also be replaced completely by varying exposure times instead, though physical attenuation may be more precise [118]. The integrating sphere may

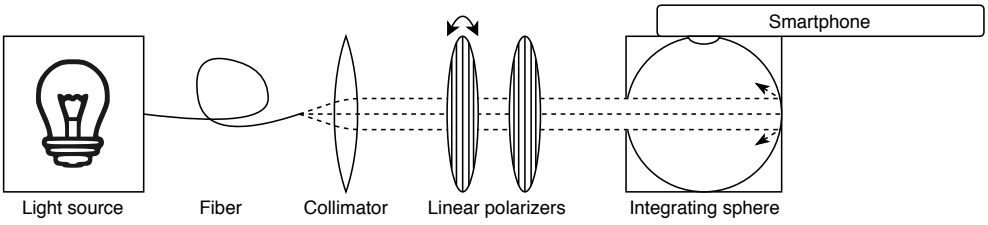


Figure 3.1: Setup used to measure linearity, ISO-gain relations, and inter-pixel gain variations on smartphones. The first linear polariser was rotatable, the second fixed. Smartphones were placed with their camera flush against the view-port at the top of the integrating sphere.

be replaced by another diffuse surface, such as a Spectralon target. If sufficiently wide, the light beam may also be shone directly onto the sensor; such a setup was used for digicams, with the digicam in place of the collimator in Figure 3.1 at a sufficient distance to completely illuminate the sensor. This was done to simplify the alignment process since our digicams had large physical CMOS sizes. Since all measurements were done on the individual pixel level, they were not affected by the added differences in illumination.

Bias, read-out noise, and dark current were measured on all devices by placing the camera flush against a flat surface (such as a table), pointing down, in a dark room. The setups for flat-fielding and spectral characterisation are described in Sections 3.3.8 and 3.3.9, respectively.

3.3.2 General properties

General hardware and software properties were retrieved from official specifications and community reviews. A survey across these provided an overview of basic physical and optical parameters of cameras. On Android smartphones, the Camera2 API provides ample information on such parameters, facilitating automatic data collection using an app.

The retrieved device properties included the camera type, manufacturer, product code and internal identifiers, release year, the number of cameras (for smartphones), camera module identifiers, and CMOS sensor models. Sensor properties included physical size, pixel pitch, resolution, orientation with respect to the device, colour filter pattern, and bit depth. Camera optic properties included focal length, f-number, neutral density filters (for high-end smartphones), and a vignetting model if available. Finally, software and firmware properties included supported software versions, RAW and JPEG support, estimated bias value (see Section 3.3.4), ISO speed range, exposure time range, and the active part of the sensor (accounting for dark pixels, see Section 3.3.5).

3.3.3 Linearity

Sensor linearity was quantified by measuring the camera response to varying exposures, either by attenuating a light source or by varying the exposure time, as discussed in Section 3.3.1. We used the setup shown in Figure 3.1 with two linear polarisers to attenuate the light for smartphones, since exposure times on those are not completely trustworthy (Section 3.4.1). Conversely, for digicams, exposure times are reliable [361, 366] and thus were used instead of physical attenuation to simplify the setup. A third method, varying the physical aperture, changes the distribution of light on the sensor [374] and thus cannot be used to measure linearity.

Two common types of non-linearity exist, either across the entire intensity range or only at high intensities. The former is common in JPEG imagery due to the gamma correction [329, 342] while the latter is expected in both JPEG and RAW data. We only investigated the former since it has the largest impact on data quality, as described in Section 3.1. Non-linearity at high intensities is easily negated by discarding data above a threshold value; we use a threshold of ≥95% of the maximum digital value.

The linearity of each pixel was expressed through the Pearson correlation coefficient r, a measure of the linear correlation between intensity and camera response. Pixels were analysed individually to negate differences in illumination and vignetting effects (Section 3.3.8). Simulated responses of a perfectly linear camera with a mean error of 5% in the incoming intensity simulating, for example, errors in exposure parameters or polariser alignment in the

setup described in Section 3.3.1, as well as Poisson noise ($\sigma_N = \sqrt{N}$) and 10 ADU read noise in the response were analysed. This included simulated measurements at 15 different exposures, averaged over 10 images per exposure. These simulated data resulted in a mean value of $r = 0.996 \pm 0.002$ and a lower 0.1 percentile $P_{0.1}(r) = 0.985$. To account for unforeseen measurement errors, we set the cut-off for linearity at $r \geq 0.980$.

Additionally, the JPEG data were compared to sRGB-like profiles to determine whether gamma inversion [285] is possible. The sRGB-like profiles are described by Equation (3.1), with J_C the JPEG response (0–255) in band C, n a normalisation factor, γ the gamma correction factor and I the incoming intensity in arbitrary units. The JPEG response of each pixel was individually fit to Equation (3.1) with n and γ as free parameters. Additionally, profiles with standard γ values (2.2 and 2.4 [285]) were fit to the JPEG data (with n free) to determine the accuracy of these standards.

$$J_C = 255 \times \begin{cases} 12.92nI & \text{if } nI < 0.0031308 \\ 1.055(nI)^{1/\gamma} - 0.055 & \text{otherwise} \end{cases} \tag{3.1}$$

3.3.4 Bias & read-out noise

Bias (or 'black level') and read-out noise (RON) were measured by stacking short dark exposures. The bias and RON in individual pixels are given by the mean and variance, respectively, of their values in each stack. Many (>50) images per stack are required to distinguish bias variations from RON. Temporal variations were probed by repeating this process several times. While EXIF metadata often contain a bias value, this is only an estimate and should be validated by measurement.

3.3.5 Dark current

Dark current (thermal noise) was measured by taking dark exposures with different lengths and fitting a linear relation between exposure time and camera response to determine the dark current in ADU s^{-1}. For cameras that have autodarking (see Section 3.2), the residual dark current was characterised instead. Depending on the autodark precision, the exposure-response relation may be non-linear in this case.

3.3.6 ISO speed

The relation between camera sensitivity and ISO speed was measured by taking identically exposed images at different ISO speeds. These were bias-corrected and pixel values were divided by those at the lowest ISO speed. A relation between ISO speed and normalisation factor was then fitted. Like the linearity measurements (Section 3.3.3), this was done individually per pixel to negate illumination differences and vignetting effects.

This relation may be any combination of linear and constant functions, depending on the implementation of ISO speed ratings. Linear relations correspond to analogue gain, while digital gain may result in linear or constant relations, as described in Section 3.2.

3.3.7 Gain variations

Inter-pixel and inter-filter gain variations were characterised using the mean-variance method [378], which exploits the Poissonian nature of photo-electrons in a sensor. We applied this

method to individual pixels rather than averaging over the sensor, to measure inter-pixel variations and remove the need for flat-fielding prior to this calibration. The response of a digital camera to incoming light is given by Equation (3.2), with M the mean response in ADU, I the exposure in photo-electrons, D the dark current in e^-, B the bias in ADU, and G the gain in ADU/e^-. Both I and D are integrated over the exposure time.

$$M = IG + DG + B \tag{3.2}$$

The variance in the response of a pixel is a combination of shot noise on the photo-electrons and dark current, and read noise. The shot noise follows a Poissonian distribution with a standard deviation $\sigma_I = \sqrt{I}$ and thus a variance $V_I = I$. The total variance in the response is expressed in Equation (3.3), with V the variance in ADU^2 and RON the read noise in ADU.

$$V = IG^2 + DG^2 + RON^2 \tag{3.3}$$

After correcting for bias and dark current, and assuming DG^2 is negligible, a linear relation between mean and variance is found, shown in Equation (3.4).

$$V = GM_{\text{cor}} + RON^2 \tag{3.4}$$

Equation (3.4) was fitted to mean and variance values from several image stacks taken under different illumination conditions. Within each stack, all images were exposed identically, while the illumination varied between stacks. A large amount of data (>10 stacks of >50 images each) was necessary to constrain the fitted gain values sufficiently (typical relative errors in individual pixels <15%). ISO normalisation functions derived in Section 3.3.6 may be used to extrapolate measured values to different ISO speeds.

3.3.8 Flat-field correction

Flat-fielding was performed by imaging a uniform light source. Unlike telescopes, most consumer cameras have fields-of-view (FoV) too large to use the twilight sky for this. Instead, a large integrating sphere was used to create an isotropic light field, as described in [118]. We used a LabSphere HELIOS USLR-D12L-NMNN lit by three halogen lamps with a specified luminance uniformity of ±1.0%, sequentially placing each camera before its aperture.

Any significant chromatic differences in the flat-field response were measured automatically, since all filters were exposed simultaneously. The $RGBG_2$ images were split out and each normalised to their maximum value, then recombined and smoothed with a Gaussian filter ($\sigma = 10$ pixels); both the individual $RGBG_2$ images and the recombined image were analysed. Since vignetting, often the dominant flat-field component, is caused by the camera aperture, the flat-field response changes and must be measured again when varying the aperture [374].

Vignetting can be parametrised in a number of different ways, as discussed in Section 3.2. For consistency, we used the DNG seven-parameter $(k_0 \dots k_4, \hat{c}_x, \hat{c}_y)$ model, also used internally in iOS smartphones, for the flat-field correction factor $g(x, y)$, expressed in Equation (3.5), with r the normalised Euclidean distance from pixel (x, y) to the optical center $(\hat{c}_x, \hat{c}_y)$.

$$g(x, y) = 1 + k_0 r^2 + k_1 r^4 + k_2 r^6 + k_3 r^8 + k_4 r^{10} \tag{3.5}$$

Three simpler, alternate methods were also tested. The first involved imaging an overcast sky, the second imaging the sun with a piece of paper taped onto the camera as a diffuser similarly to the Hukseflux Pyranometer app[13]. For the final method, the camera, again with a paper diffuser, was held flush against a computer monitor displaying a white screen, somewhat similarly to [361]. In all three cases, the camera was dithered and rotated 360° during measurements to average out anisotropies. Data from all three methods were processed in the same way as the integrating sphere data, to compare their efficacy.

3.3.9 Spectral response

The spectral response of a camera, which is a product of the individual spectral responses of its fore-optics, filters, and sensor, was measured in two ways. The first method, using a monochromator, is simple processing-wise as the data are simply a series of images at different wavelengths with known intensities [118, 354, 368]. It also allows for the measurement of inter-pixel variations in spectral response. The second, a spectrometer add-on such as iSPEX [94], is more accessible than monochromators but its spectral data are more difficult to calibrate and process.

We used a double monochromator (OL 750-M-D) at the NERC Field Spectroscopy Facility to scan a wavelength range of 390–700 nm. This wavelength range was chosen because no significant response was found outside it on any of the test cameras. The effective spectral resolution (half bandwidth) of the monochromator was 4 nm, calculated from the grating (1200 grooves/mm) and slits (2.5 mm entrance/exit and 5.0 mm central slit) used. The wavelength range was critically sampled at 2 nm intervals. A laser-driven light source (Energetiq EQ-99X) was used, and its spectral output calibrated using a silicon photodiode (Gooch & Housego OL DH-300C with a Hamamatsu S1337-1010BQ sensor). The system was NIST-traceably calibrated in 2012 and is described in more detail in [354].

Spectral characterisation was also done using a modified (removing polarisers and retarders) iSPEX add-on [94]. iSPEX has a slit consisting of two parts, one 0.4 mm ('broad') and the other 0.17 mm ('narrow') wide and a 1000 grooves/mm transmission grating foil (Edmund Optics #52-116). Using this foil, a similar spectrometer can be built for any other camera.

The reflection of sunlight on a piece of white paper was measured using the iSPEX on an iPhone SE. iSPEX projects a spectrum onto the sensor, so the pixel responses must be corrected for bias, dark current, and flat-field to obtain a quantitative spectrum. The 436.6, 544.5, and 611.6 nm spectral lines of a commercial fluorescent lamp were used for the wavelength calibration, fitting a quadratic relation between pixel position and wavelength. A stray light correction was done by subtracting the mean pixel value per column above and below the spectrum from the narrow and broad slit spectra, respectively. Two theoretical reference spectra were used to normalise the observed spectra, namely a 5777 K black body (approximating the Sun) and a diffuse solar irradiance spectrum generated using the Simple Model for the Atmospheric Radiative Transfer of Sunshine (SMARTS2) [379, 380] and smoothed to the 5 nm resolution of narrow-slit iSPEX spectra. For the latter, the location and time of the iSPEX measurements as well as the built-in urban aerosol and ground albedo models were used instead of default parameters. The models differed significantly (RMS 34%) due to the diffuse sky irradiance factored into the SMARTS2 model. Finally, the observed spectra

[13] http://www.hukseflux.com/product/pyranometer-app

were corrected for the transmission of the iSPEX optics, determined by measuring the zero-order transmission using a halogen lamp and spectrometer (OceanOptics HL-2000-LL and USB2000+, respectively).

Instead of the sun, a previously calibrated commercial lamp may be used. For example, the LICA-UCM database[14] contains spectra of common commercial lamps which can be used as standard light sources for spectral response measurements [381]. This method has the advantage of independence from weather conditions and higher reproducibility compared to solar measurements. Combined with the new version of iSPEX we are currently developing, featuring a universal smartphone hardware interface, this enables volunteer measurements of smartphone camera spectral responses.

The spectral curves $R_C(\lambda)$ thus derived were normalised to the global maximum transmission in all bands and used for calibration of spectral measurements and in the radiometric correction of imaging data (Section 3.3.10) to calculate effective spectral bandwidths Λ_C. These are defined as $\Lambda_C = \int_C R'_C(\lambda)\mathrm{d}\lambda$, with $R'_C(\lambda)$ the spectral response $R_C(\lambda)$ normalised to the maximum in band C [99,118]. This integral was calculated using the composite trapezoid method, implemented in the NumPy function `numpy.trapz` [312].

3.3.10 Relative radiometric calibration

The calibrations described in the previous section are used to convert digital values to radiance. Following the methods described in [99, 118, 382, 383], a digital value M (in ADU) in band C ($\mathrm{RGBG_2}$ for Bayer filters) can be converted to effective radiance L_C, in units of W m^{-2} sr^{-1}.

Since absolute radiometric calibration is outside the scope of this work, we instead determined the relative effective radiance $L'_C = L_C/K$, in relative radiometric units (RRU) m^{-2} sr^{-1}, with K an extra factor accounting for the absolute quantum efficiency and transmission of the lens. Measuring these requires a previously calibrated light source with a known radiance.

The expression for converting M to L'_C is given in Equation (3.6). The advantage of the piece-wise calibration given in Equation (3.6) over a black-box approach containing all calibration components is its adaptability when a small subset of parameters are changed, such as due to firmware updates or manufacturing changes. This way, calibration data can be re-used rather than requiring a full re-calibration with every change.

$$L'_C = hc\frac{1}{A_d\Lambda_C}g\left[\frac{4(f\#)^2}{\pi\tau N}\right](M - B - D\tau) \tag{3.6}$$

First, the bias B (ADU; Section 3.3.4) and dark current $D\tau$ (with D in ADU s^{-1} and τ the exposure time in seconds; Section 3.3.5) are subtracted. Linearisation of digital values [383] is not necessary since we only used sufficiently linear pixels ($r \geq 0.980$; Section 3.3.3).

Next, the image is corrected for the exposure parameters, dividing by the exposure time τ, ISO speed normalisation factor N (Section 3.3.6), and aperture, approximated as $\pi/4(f\#)^2$, with $f\#$ the f-number of the camera [118]. This approximation causes a systematic error of 4% at $f/2.0$ [118]; for fixed-aperture systems like smartphones, this error is not relevant. For systems with adjustable apertures, an exact solution may be preferable if operating at very low f-numbers. These corrections yield a response in normalised ADU s^{-1} sr^{-1}.

The third step is the flat-field correction. The response is multiplied by the flat-field correction g (unitless; Section 3.3.8). The flat-fielding methods used here account for both

[14] https://guaix.fis.ucm.es/lamps_spectra

optical and electronic variations in sensitivity, so a separate correction for inter-pixel gain variations (Section 3.3.7) is not necessary. Since absolute transmission and quantum efficiency were not measured, this step yields a response in relative counts s^{-1} sr^{-1}, proportional to the number of photo-electrons s^{-1} sr^{-1}.

Next, sensor properties are corrected for. The response is divided by the pixel size A_d (m^2; Section 3.3.2) to give a response in relative counts s^{-1} m^{-2} sr^{-1}. It is then divided by the effective spectral bandwidth of band C, $\Lambda_C = \int_C R'_C(\lambda)\mathrm{d}\lambda$ (Section 3.3.9).

Finally, the result is converted to a relative radiance by multiplying by a factor hc, with h Planck's constant and c the speed of light. This yields L'_C in RRU m^{-2} sr^{-1}.

For specific applications, Equation (3.6) may be simplified or adjusted. For example, inter-pixel bias and dark current variations are typically negligible in bright conditions. In those cases, B and D may be approximated by constants, and inter-pixel variations incorporated in the error budget. For spectroscopic applications, a relative spectral radiance $L'_{C,\lambda}$ in RRU m^{-2} sr^{-1} nm^{-1} is measured, which is not averaged over band C. In this case, the energy per photon is simply hc/λ and only the transmission at wavelength λ, $R_C(\lambda)$ is relevant; furthermore, the result must be divided by the wavelength coverage of each pixel $\Delta\lambda$. This is expressed in Equation (3.7).

$$L'_{C,\lambda} = \frac{hc}{\lambda}\frac{1}{A_d R_C(\lambda)\Delta\lambda} g\left[\frac{4(f\#)^2}{\pi\tau N}\right](M_\lambda - B - D\tau) \tag{3.7}$$

3.4 Results

The methodology described in Section 3.3 was applied to three iOS smartphones (Apple iPhone SE, 6S, and 7 Plus), two Android devices (Samsung Galaxy S6 and S8), one digicam (Nikon D5300), and one UAV camera (DJI Phantom Pro 4). This section contains an overview of results from these various calibration steps. Results for all devices are included in the SPECTACLE database further described in Section 3.4.9.

3.4.1 General properties

General hardware and software properties were retrieved from the survey described in Section 3.3.2, with a specific focus on smartphones using the previously described Android app. Little variation was found in these general properties, especially for smartphones. For example, virtually all main cameras on smartphones have apertures of $f/2.4$–$f/1.5$, focal lengths of 3.8–4.5 mm, and sensors of 3.4–6.7 × 2.7–4.7 mm, giving fields-of-view (FoVs) of 60°–75° × 45°–55°.

It was found from test images that EXIF metadata from some cameras are inaccurate. For example, the iPhone SE can use unrounded exposure times of 1/3.0 s and 1/3.9 s but records both as simply 1/3 s in metadata. Assuming the recorded exposure time of 1/3 s for a real exposure of 1/3.9 s would lead to photometric errors up to 30%. To counteract this, exposure parameters such as ISO speed and exposure time should be recorded separately from default EXIF metadata, for example with custom EXIF tags or extra files.

3.4.2 Linearity

The linearity of two smartphones (iPhone SE and Galaxy S8) and one digicam (Nikon D5300) was measured using the methods described in Section 3.3.3 and the setup described in Section 3.3.1 and shown in Figure 3.1. The smartphones were analysed using rotating linear polarisers while the D5300 was analysed by varying exposure times. Simultaneous RAW and JPEG images were taken on each device (using the Fine JPEG setting on the D5300) to compare their responses. JPEG images were taken with a fixed white balance.

The Pearson r coefficients of the RAW and JPEG responses of all pixels were calculated and their histograms are shown in Figure 3.2. The JPEG responses of all pixels in all cameras were well below the linearity threshold ($r \geq 0.980$), showing again that JPEG data are highly non-linear. Conversely, nearly all RAW responses were well within the bounds for linearity, with 99.9% of r values ≥0.997 (iPhone SE), ≥0.996 (Galaxy S8), and ≥0.999 (D5300). The Galaxy S8 was the only camera with RAW responses having $r < 0.980$, though only in 56 pixels.

The JPEG and RAW responses of individual pixels in the iPhone SE and Galaxy S8 cameras are shown in Figure 3.3. The JPEG responses are visibly non-linear ($r = 0.956, 0.918$) while the RAW responses are linear within measurement errors ($r = 0.999, 0.998$). Furthermore, the dynamic range of the JPEG data is much smaller than that of the RAW data. These differences highlight the advantages of RAW data.

Finally, Figure 3.4 shows the best-fitting γ for the JPEG response per pixel as well as the accuracy of two standard values ($\gamma = 2.2$ and 2.4, expressed in RMS relative difference (1 − data/fit). Large inter-pixel, inter-filter, and inter-device differences in best-fitting γ exist, indicating an sRGB gamma inversion with a single γ value is not possible. Furthermore, the $\gamma = 2.2$ and 2.4 models are both clearly very inaccurate for all cameras. For the $\gamma = 2.2$ and 2.4 cases respectively, 99.9% of pixels had RMS relative differences between observations and the sRGB model of >7% and >10% (iPhone SE), >13% and >15% (Galaxy S8), and >19% and >21% (Nikon D5300).

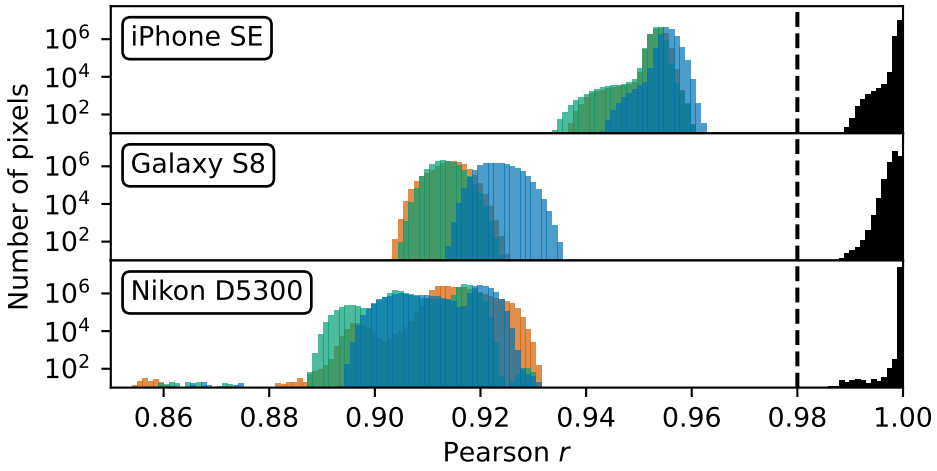


Figure 3.2: Histogram of Pearson r coefficients for RAW (black, all filters combined) and JPEG (red/green/blue) responses. The $r \geq 0.980$ cut-off is shown with a dashed black line. Note the logarithmic vertical scale.

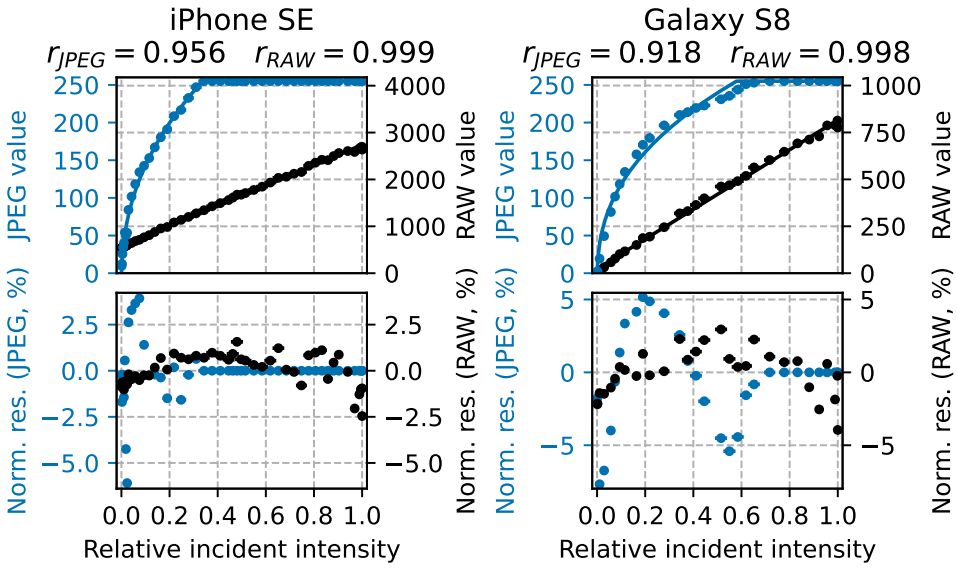


Figure 3.3: JPEG (blue, left vertical axis) and RAW (black, right axis) response of a single B pixel in the iPhone SE (left) and Galaxy S8 (right) rear cameras, under varying incident intensities. Each point represents the mean of a stack of 10 images at the same exposure. Vertical error bars are smaller than the dot size. The black and blue lines represent the best-fitting linear (RAW) and sRGB-like (JPEG) profiles, respectively. The lower row shows the residuals, normalised to the dynamic range.

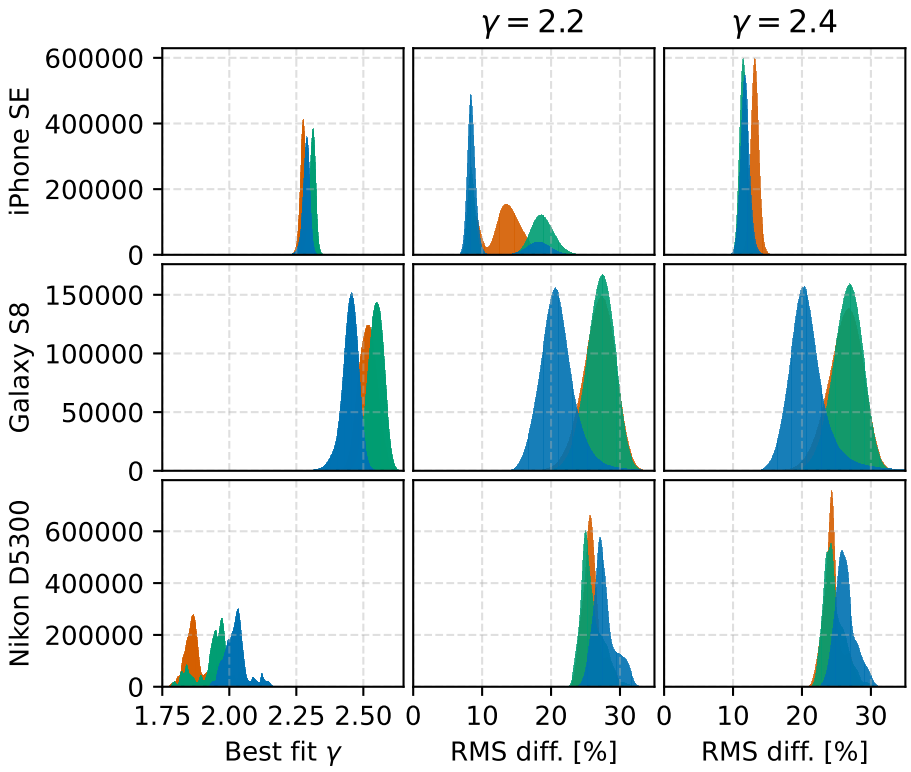


Figure 3.4: Histogram of best-fitting γ and RMS relative difference between JPEG data and fit (for models with γ = 2.2 and 2.4) in the RGB bands.

3.4.3 Bias & read noise

Bias and read noise variations in four smartphone cameras (iPhone SE and 7 Plus, Galaxy S6 and S8), one digicam (Nikon D5300), and one UAV camera (Phantom Pro 4) were analysed using the methods from Section 3.3.4.

Bias values in all cameras deviated systematically from the EXIF values by <1 ADU on average, with standard deviations also <1 ADU. However, large outliers were found, such as some pixels in our Galaxy S6 which even saturated in bias frames. Phantom Pro 4 data are scaled up from 12-bit (its native bit depth) to 16-bit, increasing the observed bias variations. Scaled down to 12 bits, its bias variations are similar to those in the other cameras.

Typical observed RON values were distributed similarly to inter-pixel bias variations. The smartphones and D5300 show RON distributions consisting of one or two main components <3 ADU, which correlate with inter-pixel gain variations (Section 3.4.6), and a long but shallow tail towards RON values >20 ADU. As with the bias variations above, the Phantom Pro 4 showed a comparatively high mean RON (14 ADU at ISO speed 100) in 16-bit (scaled-up) data but a comparable value (1.8 ADU) when scaled down to its native bit depth of 12 bits.

Large-scale patterns in inter-pixel and inter-filter bias and RON variations were observed

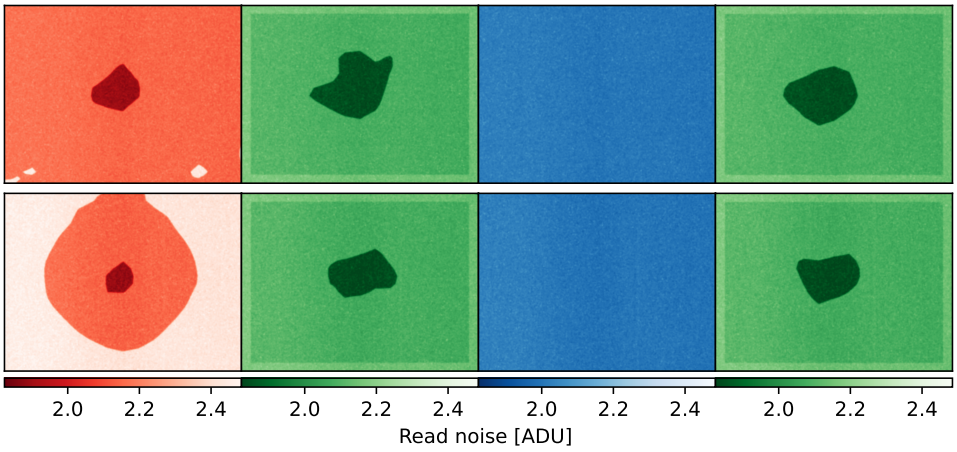


Figure 3.5: Read-out noise per pixel of two iPhone SE devices (top and bottom) at ISO speed 23, in the $RGBG_2$ filters from left to right. Darker colours correspond to lower read-out noise. A two-dimensional Gaussian filter ($\sigma = 5$ pixels) has been applied to better visualise large-scale variations. The G image shows similar patterns to Figure 3.7.

in several cameras, most prominently in the smartphones. Figure 3.5 shows the RON per pixel in the sensors of two iPhone SE devices. The RON and bias patterns on each device are strongly correlated, suggesting a common origin. The RMS difference in bias between these two devices was 0.31 ADU, larger than the standard deviation on either device (0.24 and 0.21 ADU). The large-scale patterns persisted over time scales of months, indicating that they are systematic.

Both bias variations and RON decreased with ISO speed when normalised (Section 3.3.6). This may be a result of better amplifier or ADC performance at a higher gain. Similarly, large-scale patterns such as those in Figure 3.5 become less distinct at high ISO speeds.

Either a map of mean bias per pixel at a given ISO speed $B(x, y, ISO)$ or a mean value B is used in Equation (3.6). For low-light applications such as dark-sky measurements [320] or spectroscopy, a detailed map is necessary since a single 'bad' pixel with an abnormally high output may cause a significant systematic error. Being manufacturing defects, bad pixels are in different locations even on two cameras of the same model, and thus a map is required for each device. Conversely, for bright conditions, the bias variations are not significant and thus a mean value can be used. Similarly, RON values can be incorporated in the error budget separately for individual pixels or using the RMS value as an ensemble estimate.

3.4.4 Dark current

The methods described in Section 3.3.5 were applied to two smartphones (iPhone SE and Galaxy S8) to measure their dark current properties. Both cameras have built-in dark current calibrations (autodark; see Section 3.2). Measurements were done at room temperature, with short breaks between differently exposed stacks to prevent overheating the sensor. However, sensor temperatures were not obtained from the camera software.

A separate data set consisting of 96 images taken with 4 seconds between each on the iPhone SE, during which the entire device palpably warmed up, was analysed to identify

thermal effects. Pearson r correlations between response and time stamps (as a proxy for temperature) were calculated for the individual pixels. These r values were well-described by a normal distribution with $\mu = 0.00$ and $\sigma = 0.10$, indicating that no strong relation exists between temperature and residual dark current. However, we note that again no direct sensor temperatures could be obtained.

In both cameras, a small residual (positive or negative) dark current signal was observed. Most pixels in both cameras had little dark current (RMS <2 ADU s^{-1}, 99.9th percentile of absolute values <6 ADU s^{-1}), though notable outliers were found, such as >300 pixels in our Galaxy S8 with dark current >50 ADU s^{-1}. The residual dark current decreased at higher ISO speeds, similar to RON and bias variations (Section 3.4.3), but showed no large-scale patterns.

These results show that autodarking accurately corrects most pixels, but is inadequate for outliers. Since autodarking is built into camera chips, it cannot be disabled. For outliers and in low-light conditions, it should be augmented with a manual dark current correction. As with bias variations, the dark current map $D(x, y, ISO)$ is used in Equation (3.6) for low-light conditions, but an approximation is adequate for bright conditions. For autodarked cameras like the ones tested here, a mean value of $D = 0$ ADU s^{-1} is assumed, and the RMS variation incorporated into the error budget. Outliers may be masked in either case.

3.4.5 ISO speed

The normalisation of data at different ISO speeds was measured using the methods from Section 3.3.6 on two smartphones (iPhone SE and Galaxy S8) and one digicam (Nikon D5300).

The measured and best-fit normalisation curves are depicted in Figure 3.6. The Nikon D5300 and Galaxy S8 were best fit with a single linear relation, while the iPhone SE curve is clipped at ISO 184. This clipping is not due to image saturation, as none of the pixels in any image reached saturation. The linear part of the iPhone SE relation passes through the origin, while the Nikon D5300 and Galaxy S8 curves do not, instead showing significant (>5%) systematic errors when using the simplest mathematical model (zero offset and slope 1/minimum ISO speed).

The clipping of the iPhone SE curve likely corresponds to a transition from purely analogue to purely digital gain. However, data from the Camera2 API on the Galaxy S8 indicated that it too uses digital gain, at ISO speeds >640. This suggests that there are different implementations of gain for RAW photography.

The large observed differences in ISO speed normalisation can lead to significant systematic errors when combining data taken at different ISO speeds, if not adequately calibrated. Data are normalised by dividing by N, as expressed in Equation (3.6).

3.4.6 Gain

The methods from Section 3.3.7 were used to characterise inter-pixel gain variations in two smartphones (iPhone SE and Galaxy S8).

Significant inter-pixel variations were observed, as shown in Figure 3.7 for the G pixels in both cameras. Since the measurement protocol is applied on the individual pixel level, the observed variations are only due to differences in gain, rather than external factors such as vignetting effects. The iPhone SE showed small variations, with higher gain values at the edges and lower values in the center. This pattern is similar to that seen in Figure 3.5, suggesting

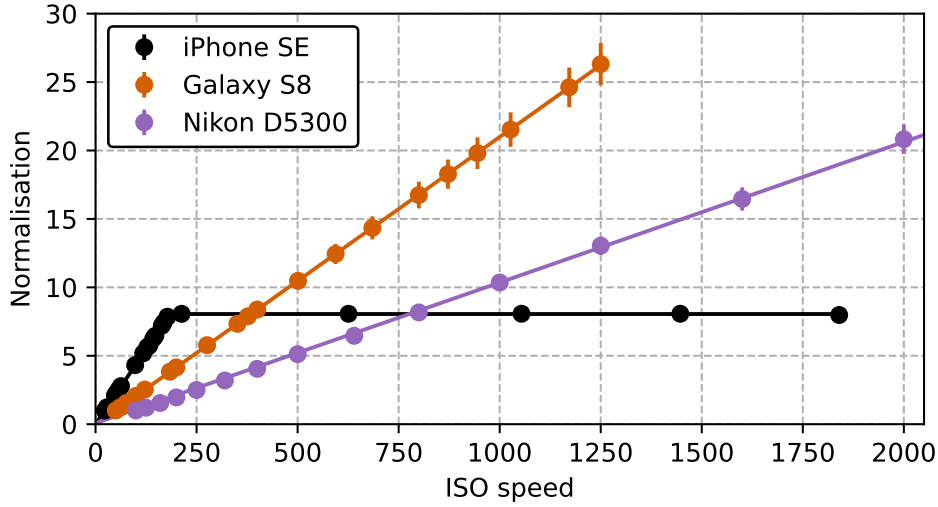


Figure 3.6: ISO speed normalisation for the iPhone SE, Samsung Galaxy S8, and Nikon D5300. Dots indicate means of image stacks divided by the mean value per pixel at the lowest ISO speed. Lines indicate the best-fitting relationships.

a common origin. Conversely, on the Galaxy S8 a concentric pattern with a very wide range is clearly visible, likely intended as a first-order vignetting correction. Both showed similar ranges in gain (0.58–2.74 and 0.59–2.45 ADU/e$^-$, respectively), though on the iPhone SE most variations were on small scales and thus are not visible in the smoothed image.

Histograms of gain values for both cameras are shown in Figure 3.8. Inter-filter differences are small in the Galaxy S8 but obvious in the iPhone SE. In particular, the R, G, and B distributions in the latter clearly have different mean values and widths (means and standard deviations of 1.97 ± 0.24, 1.78 ± 0.29, and 1.73 ± 0.30 ADU/e$^-$, respectively). Furthermore, the G distribution is bimodal while both others are unimodal; no significant differences be-

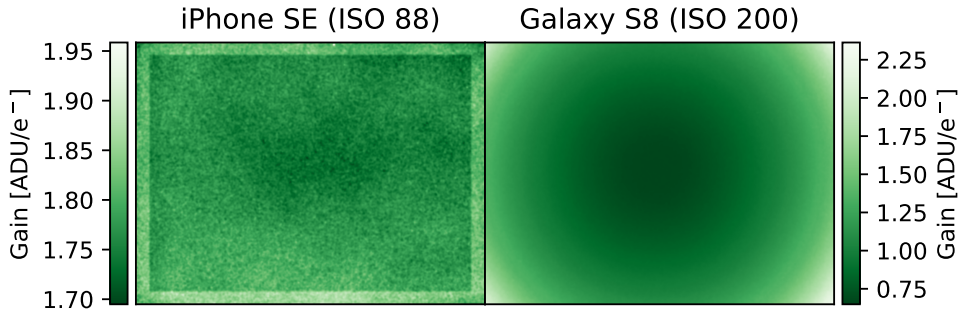


Figure 3.7: Gain values of G pixels in the iPhone SE (left; ISO speed 88) and Galaxy S8 (right; ISO speed 200) sensors. Darker colours indicate lower gain values. A two-dimensional Gaussian filter (σ = 5) has been applied to better visualise large-scale fluctuations. The iPhone SE patterns are similar to the read noise shown in Figure 3.5.

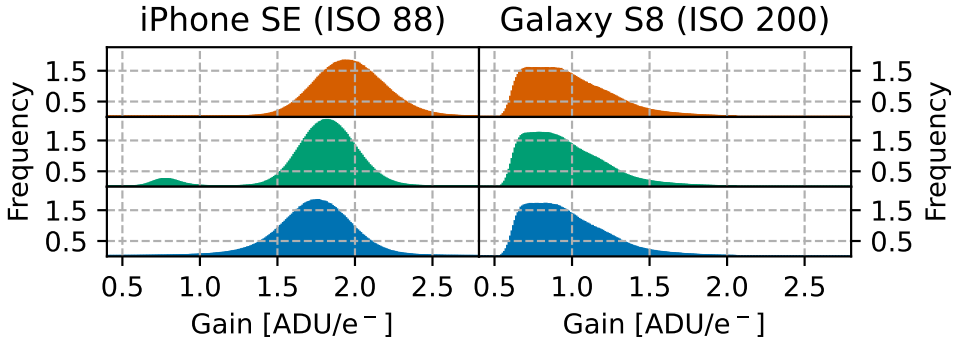


Figure 3.8: Histogram of gain values in the R (top), G and G_2 (middle), and B pixels (bottom) in the iPhone SE (left; ISO speed 88) and Galaxy S8 (right; ISO speed 200) sensors. The vertical axes were normalised to account for the different amounts of pixels.

tween the G and G_2 gain distributions were found, so this is not the cause for the observed bimodality.

The observed gain variations are significant and provide insights into camera design and fabrication, specifically the origins of some of the phenomena seen in flat-field data (Section 3.4.7). However, they are not necessary in the radiometric calibration of data, since our flat-field correction method (Sections 3.3.8 and 3.4.7) accounts for these inter-pixel gain variations as well as all other differences in sensitivity, such as vignetting, as discussed in Section 3.3.10.

3.4.7 Flat-field correction

Three smartphones (iPhone SE and 6S, and Galaxy S8) and one UAV (DJI Phantom Pro 4) were flat-fielded using an integrating sphere as described in Section 3.3.8. An aperture of $f/2.8$ was used for the Phantom Pro 4, and on each device the maximum focus was used. 300 images were taken with the iPhone SE and Galaxy S8, 224 with the Phantom Pro 4, and 30 with the iPhone 6S. The latter was flat-fielded using a different set-up, with a Newport 819D-SF-4 integrating sphere and taking only 30 images as this was sufficient for an SNR >3 in >99% of its pixels.

Significant vignetting was found in all cameras. The observed correction factors of the iPhone SE, the best-fitting model, and residuals between the two are shown in Figure 3.9. The smooth pattern suggests optical vignetting is the main flat-field component; the same is true in the iPhone 6S and Galaxy S8. The Phantom Pro 4 data showed an additional steep cut-off near the corners, suggesting mechanical vignetting. To counteract the latter, the outermost 250 pixels on all sides of the images from all cameras were removed prior to further analysis. Correction factors up to 2.42 (iPhone SE), 2.03 (iPhone 6S), 1.43 (Galaxy S8), and 2.79 (Phantom Pro 4) were observed. No significant chromatic differences were found, so the recombined data were used instead of separate $RGBG_2$ data.

As seen in Figure 3.9, the DNG model fitted the data well with only small residuals remaining. The RMS of the residuals, normalised to the unsmoothed observed values, was 1.5% (iPhone SE), 1.4% (Galaxy S8), 3.1% (iPhone 6S), and 2.0% (Phantom Pro 4). These differences drop to ≤0.7% on all cameras when using the spatially smoothed data, implying

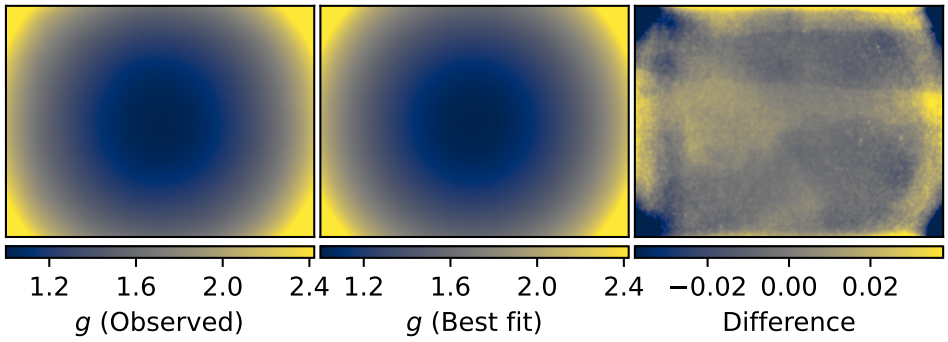


Figure 3.9: Flat-field correction factor g for the iPhone SE camera. From left to right: observed values (inverse of observed relative sensitivity), best-fitting DNG model, and residuals.

that they are mostly due to pixel-by-pixel variations and noise in the observations. These small residuals show that the DNG model is an adequate approximation for most applications; a pixel-by-pixel map per device is necessary only if sub-percent precision is required. Estimated errors in the model were <0.01 on the polynomial coefficients and $<10^{-5}$ on the optical center (in relative coordinates) for all cameras. Anomalous dots can be seen throughout the difference image in Figure 3.9, possibly due to dust particles or inter-pixel gain variations (Section 3.4.6).

Since iOS also uses the DNG model for its internal vignetting correction, a direct comparison between correction models for the iPhone SE was made. The RMS relative residual between our smoothed data and the internal model was 5.9%, more than 10 times that of our model (0.5%). While the iOS model is symmetric ($\hat{c}_x = \hat{c}_y = 0.5$), ours had a slight offset ($\hat{c}_x = 0.494226(1)$ and $\hat{c}_y = 0.503718(2)$). The polynomial coefficients all differed by $>400\sigma$, with σ the standard error on our model derived by the fitting routine. Finally, the RMS difference between the models per pixel was 5.7%.

The three alternate methods described in Section 3.3.8 were tested on the Galaxy S8. 40 images of the overcast sky were taken, as well as 40 of the sun and 50 of a monitor with a paper diffuser. The Galaxy S8 was used because its integrating sphere data show a large asymmetry ($\hat{c}_x = 0.449391(5), \hat{c}_y = 0.426436(9)$), providing a simple comparison metric. The RMS difference between the smoothed data from the integrating sphere and alternative methods relative to the sphere data were 4%, 4%, and 5%, respectively. The best-fitting optical centers of all three data sets differed significantly both from the sphere data and from each other ($\hat{c}_x = 0.53447(1), 0.501989(4), 0.490794(4)$ and $\hat{c}_y = 0.38837(2), 0.449426(7), 0.477590(7)$, for the sky, sun, and monitor methods, respectively). This causes a typical systematic error on the order of 5% in all three cases. Finally, six replicate measurement sets (50 images each) were taken using the monitor method to assess the effects of nonuniformities in the paper diffusers, generating a correction model for each set. The typical variation, expressed as the RMS of the standard deviation per pixel relative to the mean value per pixel, was 3%, smaller than the typical deviations between the do-it-yourself methods and ground truth data. The effect of paper nonuniformities thus does not significantly impact the quality of do-it-yourself data.

The flat-field correction is incorporated in the radiometric correction expressed in Equation (3.6) as the factor $g = g(x, y)$. For cameras with a fixed aperture, such as smartphones,

one measurement is sufficient; otherwise, g varies with aperture. This corrects for the systematic error induced by flat-fielding effects but pixels at the edges still receive fewer photons than those in the center. The former correspondingly have a smaller SNR due to shot noise, scaling as $SNR \propto g^{-1/2}$. Therefore, objects of interest are preferably imaged near the optical center of the camera.

3.4.8 Spectral response

Two smartphones (iPhone SE and Galaxy S8) and one UAV (DJI Phantom Pro 4) were spectrally calibrated using a monochromator, and the iPhone SE using iSPEX, as described in Section 3.3.9.

Figure 3.10 shows the normalised spectral response curves derived from the monochromator data, calibrated to the spectral throughput of the monochromator and spectral irradiance of the light source. This calibration was done by measuring its output under the same conditions as during the measurements, using a pre-calibrated silicon photodiode. Parts of the spectra were measured with different exposure settings and monochromator filters; these were first calibrated and then normalised and averaged on overlapping sections. The peak response wavelengths and effective bandwidths of the $RGBG_2$ filters in the different cameras are given in Table 3.1.

Some similarities and differences between the cameras are immediately obvious from Figure 3.10 and Table 3.1. Notably, no significant differences between G and G_2 were found in any camera (RMS differences ≤0.003); the different peak wavelength for the Phantom Pro 4 is likely due to noise. The peak response wavelengths are very similar or even identical between cameras, as are the effective bandwidths, with two notable exceptions. The Galaxy S8 B filter is significantly broader than the others, with a comparatively high response at $\lambda > 500$ nm. Conversely, the Phantom Pro 4 has a relatively narrow R filters due to its NIR cut-off around 670 nm rather than 680 nm. Moreover, the R filters in all three cameras show a secondary peak around 535 nm and nearly identical responses between 570–650 nm.

The spectral response curves measured with iSPEX, shown in Figure 3.11, were simi-

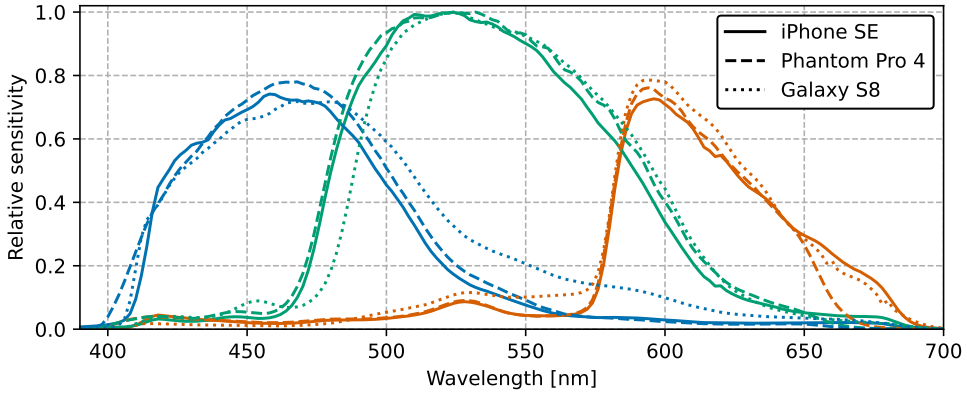


Figure 3.10: Spectral response curves of the iPhone SE, Galaxy S8, and Phantom Pro 4, derived from monochromator data. The responses are normalised to the global maximum per camera, giving relative sensitivities. G is the average of the G and G_2 responses over the wavelength axis, since no significant differences were found. RMS errors are ≤0.005.

Camera	$\lambda_{P,R}$	Λ_R	$\lambda_{P,G}$	Λ_G	$\lambda_{P,G2}$	Λ_{G2}	$\lambda_{P,B}$	Λ_B
iPhone SE	596	72	524	110	524	109	458	93
Galaxy S8	594	73	524	109	524	109	468	117
Phantom Pro 4	594	65	524	115	532	116	468	94

Table 3.1: Peak response wavelength $\lambda_{P,C}$ and effective spectral bandwidth Λ_C of each filter in the three cameras, derived from monochromator measurements. All values are in nm.

lar to those derived from the monochromator data but showed small though significant systematic differences. No significant differences were found between narrow- and broad-slit spectra, so these were averaged. RMS differences between iSPEX- and monochromator-derived responses were 0.04, 0.02, and 0.02 (SMARTS2 normalisation) and 0.12, 0.10, and 0.10 (black-body normalisation), in RGB respectively. The black-body under-estimated the irradiance <500 nm and over-estimated it >500 nm compared to the SMARTS2 model, resulting in large deviations in the retrieved spectral response. The RMS difference between the monochromator-derived and black-body-normalised iSPEX-derived spectral responses could be reduced to 0.05, 0.10, and 0.04 by multiplying each filter with an empirical constant. However, systematic differences >0.2 remained in the G filter at wavelengths of 500–600 nm. Conversely, the SMARTS2-normalised iSPEX-derived spectral responses only showed a significant systematic difference compared to monochromator data at wavelengths >650 nm, the origins of which are unclear.

The observed differences between devices have important implications for RGB colour measurements and spectroscopy, for example for colour measurements as discussed in Section 3.1. The effective spectral bandwidths are incorporated into the radiometric calibration of imaging data as described in Section 3.3.10. Furthermore, smartphone spectrometers naturally require calibration for the spectral response of the camera, as expressed in Equation (3.7).

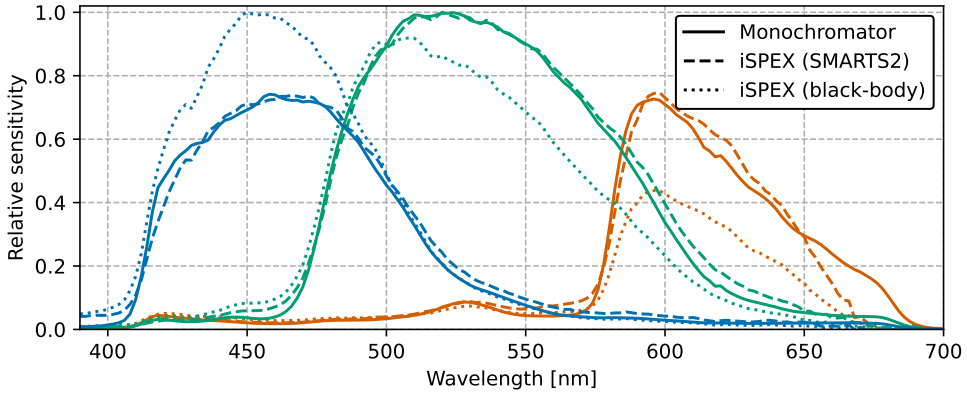


Figure 3.11: Comparison of the iPhone SE spectral response curves measured with the monochromator and iSPEX. iSPEX data are normalised using a 5777 K black-body and a SMARTS2 model, as described in Section 3.3.9.

3.4.9 SPECTACLE database

To facilitate the use of consumer cameras in scientific projects and improve future compatibility, we have created the SPECTACLE (Standardised Photographic Equipment Calibration Technique And CataLoguE) database. It includes the calibration data required for radiometric corrections (Section 3.3.10), for the most popular cameras. The data are given in standardised formats, split into three categories (device, camera, and software) to minimise the amount of data required. For example, two devices using the same camera module have the same spectral response curves and flat-field response, while software parameters such as bias and ISO speed settings vary. The former can thus be combined while keeping the latter separate. Since the properties of a camera may change with firmware updates or changes in manufacturing, database entries may be split according to device version, rather than assuming devices of the same model are clones. Finally, given calibration data for multiple identical devices, statistics on variations within a camera model may be included. The open design of the SPECTACLE database, based on the Parse platform, allows anyone to use or contribute data, particularly using the calibration apps we have developed. Submitted data are currently curated by the authors to ensure their quality. As the database grows, community curation or automated curation based on outlier analysis may become preferable. SPECTACLE can be accessed at http://spectacle.ddq.nl/.

3.5 Discussion & conclusions

In this work, we have presented a standardised calibration methodology for the most important factors limiting the quality of consumer camera data, the first to our knowledge. Furthermore, we have developed the SPECTACLE database, containing calibration data for the most popular devices. The standardised methodology and the SPECTACLE database have the potential to improve the sustainability of projects using these cameras, by simplifying their calibration and the use of multiple camera models.

The main difference between our approach and those in much of the literature is the use of RAW data. Software constraints previously forced the use of JPEG data, which are compressed and heavily processed, introducing systematic effects that negatively affect the data quality and are difficult to calibrate [94,121,142,285,319,334,342,349,357]. The desire to use RAW data has been expressed widely in the literature [118,121,285,329,331,334,349,351], and their superiority is clearly demonstrated by the highly linear response and larger dynamic range found in Section 3.4.2. The former is especially notable since non-linearity and the associated gamma correction are among the most cited problems of JPEG data [94,121,162,285,317,322,329,332,334,338,342,357–360,362]. While JPEG non-linearity corrections exist, either fully empirical or based on the sRGB standard [285,342,358], the wide (1.7–2.6) variations in gamma and large (>30%) deviations from sRGB profiles shown in Section 3.4.2 and Figure 3.4 indicate that these are inaccurate and difficult to generalise. The highly linear nature of RAW data was previously demonstrated in [118,361,366] and may be a result of internal linearity corrections in the CMOS chip [104]. Furthermore, RAW data are not affected by white balance, a colour correction in JPEG processing which severely affects colourimetric measurements, is difficult to calibrate, and differs strongly between measurements and cameras [56,94,285,317,318,323,342,349,351,376]. This variable gamma correction and white balance make it impossible to invert the JPEG algorithm and recover RAW data. However, RAW data are no panacea, since they still require further calibrations.

Furthermore, not all consumer cameras support RAW imagery, especially low-end smartphones; hence the low adoption rate in literature until now. Still, we consider the linearity, larger dynamic range, and lack of unknown post-processing affecting the data worth relying on RAW data, especially in a market trending towards broader support.

Inter-pixel and inter-device bias variations and read noise were found to be small in general ($\sigma < 1$ for bias variations, mean RON < 3 ADU), though with large outliers (Section 3.4.3). These distributions are similar to those found in several smartphones [324] and a Raspberry Pi camera [118], though neither work distinguishes between bias variations, read noise, and dark current. The large-scale patterns seen in Figure 3.5 were not found in the literature. Their cause is unclear, though correlations with inter-pixel gain variations (Section 3.4.6) suggest a common origin. Ultimately, since both phenomena are small, for most applications these patterns are merely a curiosity and an estimate in the error budget and masking of outliers is sufficient for further radiometric calibrations (Section 3.3.10).

While dark current has been implicated in the literature as a major noise source [118, 319, 324, 344, 354, 363, 364], the results presented in Section 3.4.4 indicate that it is actually typically quite minor. The RMS dark current in the iPhone SE and Galaxy S8 (<2 ADU s^{-1}) is similar to values found in [319, 358, 363, 364], while we found larger outliers, such as >300 pixels with >50 ADU s^{-1} in our Galaxy S8. Similarly to [364], no significant relationship was found between temperature and residual dark current, though this experiment should be repeated under more controlled conditions and using internal sensor temperatures to draw strong conclusions. In general, a quantitative comparison with the literature is difficult, since those studies used JPEG data, not RAW. While our sample of two cameras is insufficient to draw broad conclusions, these results suggest that dark current is less important than previously thought. As discussed in Section 3.4.4 and similarly to the aforementioned bias and RON variations, extensive characterisation of the dark current in individual pixels is necessary for low-light applications and spectroscopy as these are significantly affected by a few 'bad' pixels. Conversely, for bright-light conditions the dark response is typically negligible and an ensemble estimate in the error budget and masking of outliers are sufficient.

ISO speed normalisation is typically done by simply dividing digital values by the ISO speed [121, 367], but the results presented in Section 3.4.5 and Figure 3.6 contradict the validity of this method. This discrepancy was also identified in [368]. Observed relations differ significantly from the naïve linear model in shape, offset and slope. For example, differences between the two models of >5% were found in the Galaxy S8. More extremely, the expected and observed normalisation factor at ISO speed 1840 on the iPhone SE differ by a factor of 10. Moreover, Android documentation suggests that more complex curves with mixed analogue and digital gain may also be in use. Thus, to prevent similar systematic errors, either a single ISO speed per device must be used or these relations must be calibrated.

Significant inter-pixel gain variations were found in Section 3.4.6, as shown in Figures 3.7 and 3.8. The Galaxy S8 showed a strong radial pattern, likely intended as a first-order vignetting correction; this was not seen in the iPhone SE. Conversely, gain values in the latter differed significantly between colour filters. This may be a colour correction called analogue white balance, which is described in the DNG standard [369]; however, in this case it is not clear why significant inter-pixel variations exist. No previous discussion of such variations in gain in a consumer camera was found in the literature. Typically, an equal gain in all pixels is assumed in absolute radiometric calibrations [118, 368] but the variations found here cast doubt on the generality of this assumption.

Strong flat-field effects were found in Section 3.4.7, with correction factors up to 2.79.

Similarly large correction factors have been found for other cameras, for instance approximately 2.8 in a Canon EOS 400D [329] and 650D [330], 4 in a Raspberry Pi camera [118], 1.8 in a Canon EOS 10D [374], and 1.5 in a Nikon E775 [375]. It should be noted that vignetting is highly aperture-dependent and thus these correction factors will change with varying apertures [374]. Interestingly, we did not find the large chromatic differences described in [118, 329]. Notably, the Galaxy S8 showed a much weaker vignetting effect ($g_{max} = 1.43$) than the other cameras ($g_{max} > 2$), likely due to the aforementioned inter-pixel gain variations. These may also explain the strong asymmetry ($\hat{c}_x = 0.449391(5), \hat{c}_y = 0.426436(9)$) seen in the Galaxy S8, due to the main symmetrical component having been corrected already.

The 7-parameter vignetting model described in the DNG standard [369] fits our data very well (RMSE ≤3.1% for raw data, ≤0.7% for smoothed data), without significant systematic differences. Since the typical difference between observed and modeled corrections is small, pixel-by-pixel flat-fielding is necessary only for applications requiring sub-percent precision. For those, a flat-field map would be made for each individual device, rather than using the same map for multiple devices of the same model. Flat-field measurements of the latter could be used to quantify typical variations in flat-field response among identical devices and further determine when pixel-by-pixel or modeled flat-field corrections are preferable. The DNG model is also used for internal vignetting correction in iOS. While this correction is sometimes considered a major advantage of JPEG data over RAW data, the internal model of the iPhone SE was shown to be significantly less accurate (RMSE = 5.9%) than one based on our data (RMSE = 0.5%), contradicting this notion. Moreover, residual vignetting effects up to 15% have been observed in JPEG data [358]. A comparison to the internal correction data in Android smartphones, consisting of pixel-by-pixel look-up tables, has not yet been done since these data are relatively difficult to access.

Finally, three simpler alternative flat-fielding methods were tested, namely imaging the sky, the sun, and a computer monitor, as described in Section 3.3.8. Applied on the Galaxy S8, data from these methods differed from the integrating sphere data by ≤5% RMS. These errors mainly result from a difference in the location of the optical center. The cause of these discrepancies is unclear, though insufficiently isotropic light sources are an obvious explanation. Nevertheless, the RMS difference of ≤5% is small compared to the overall flat-field correction of up to 179% and better than the internal correction of the iPhone (RMS 5.9%). These methods thus serve as a useful first estimate for the flat-field correction in the absence of integrating sphere data. As discussed in Section 3.2, many further alternative flat-fielding methods exist [329, 330, 373–377]. Our data may be useful as a ground truth for a thorough comparison of such methods akin to [330, 361].

The spectral responses found in Section 3.4.8 and shown in Figure 3.10 agree well with those found in the literature [118, 121, 317, 323, 329, 333, 334, 354, 358, 361, 365–368], with the RGB curves centered around 600, 520, and 470 nm, respectively. Notably, the strong secondary peaks seen in [121, 358] were not found in our data and may be JPEG artefacts. Differences are mainly found in the wings, such as the NIR cut-offs [329, 354] and harmonics. The comparatively high response of the Galaxy S8 B filter at wavelengths >500 nm is also seen in the Nokia N900 [333] and Sony A7SII [317], and to a lesser extent the Galaxy S5 [121], but is otherwise uncommon. The early NIR cut-off of the Phantom Pro 4 appears to be similarly uncommon but not unique [121, 317, 333, 354]. These differences again show the importance of spectral characterisation for normalising smartphone spectrometer data. Furthermore, the significant variations show that the common assumption of sRGB responses [285, 332] does not hold, as has been suggested previously [331], and characteri-

sation of the spectral response is necessary to convert observed colours to colour spaces such as CIE 1931 XYZ or CIELAB [333,335]. However, colour measurements still depend on the incident light spectrum [335]; hyperspectral measurements, for example with iSPEX [94], and characterisation of common light sources [317,381] may provide valuable additional information. Finally, while no significant response was found at wavelengths <390 or >700 nm on our test cameras, it may be worthwhile in the future and the SPECTACLE database to use a spectral range of 380–780 nm to follow colourimetric standards [335,359,363].

Spectral response measurements done with the iSPEX smartphone spectrometer [94] agreed well (RMS differences ≤0.04) with the monochromator measurements (Section 3.4.8 and Figure 3.11). The only systematic difference was an under-estimation at wavelengths >650 nm, though it is unclear what causes this. The good agreement shows that iSPEX measurements are an adequate replacement for monochromator data if the latter are not available. This will be especially useful with the new iSPEX we are developing, which will also feature universal smartphone hardware interface. One downside of this method is that it requires an accurate solar reference spectrum. We used one generated with SMARTS2 [379,380]; this model matches observed solar spectra very well but it is not very portable or user-friendly for non-expert users. A 5777 K black-body approximation was also used but reproduced the SMARTS2 spectrum poorly (RMSE of 34%) and accurate spectral response curves could not be retrieved this way. A more portable model or set of standard spectra could improve the user-friendliness of this calibration method.

Further alternative methods for spectral response characterisation include those based on multispectral measurements using computational methods to enhance their resolution [335, 372, 384, 385] or those using a linear variable edge filter [334]. However, the former are not sufficiently accurate [366] while the latter is not necessarily more accessible than a monochromator. Our data may be used as a ground-truth for testing other methods akin to [366] but with the advantage of smartphones being more accessible than the cameras used therein.

Finally, we have created the SPECTACLE database containing the calibration data described above. The aim of this database is to facilitate the use of consumer cameras in scientific projects by reducing the labor required for calibration. Data sets containing spectral responses [333,366] and extensive calibrations of single cameras [118] have been published before but to our knowledge SPECTACLE is the first comprehensive, centralised spectral and radiometric calibration database. It is designed with community participation in mind, relying on volunteer contributions to become and remain complete in the rapidly evolving camera market. This will require a critical mass of users to maintain it, which is easier if more accessible calibration methods, like those discussed previously, can be used. We have kick-started this process with the calibrations done in this paper and will continue this while developing iSPEX.

Though extensive, our calibration methodology is not complete. The two most prominent missing components are geometric distortions and absolute radiometric calibration. The former are a well-known phenomenon with a large impact on image quality but relatively simple to measure and correct [322, 326–328, 353, 377]. A parametric model for distortion is given in the DNG standard [369] and a comparison between measured distortions and the internal correction models of different cameras, similar to that done in Section 3.4.7 for vignetting corrections, may be used to determine the accuracy of the latter. Absolute radiometric calibration is extremely valuable for quantitative measurements, as described in Section 3.3.10. In principle, our methods and calibration data contain most of the information

required for this, bar a constant K. Absolute radiometric calibration of consumer cameras has been demonstrated before, notably in the Raspberry Pi camera [118], and Nikon D300 and Canon 40D [368], though only for a small number of devices. Another notable example is the Hukseflux Pyranometer app (Section 3.3.8) for measurements of solar irradiance, though it is intended for education and entertainment rather than scientific measurements. Finally, most of our calibrations were done on a single device, and differences between devices may exist, as shown in Figure 3.5. Calibration of multiple devices per camera model would allow the characterisation of these differences and the associated errors when using multiple devices. Additionally, differences may be introduced by changes in manufacturing or camera software. Characterisation of different generations of the same model camera will be necessary to characterise these, which may result in separate entries in the SPECTACLE database for each camera version being necessary. However, the modular design of the SPECTACLE database makes it simple to extend. The simple, standardised calibration methods described in this work and the SPECTACLE database have the potential to greatly improve the data quality and sustainability of future scientific projects using consumer cameras.

Funding

European Commission Horizon 2020 programme (grant nr. 776480, MONOCLE and grant nr. 824603, ACTION).

Acknowledgements

The authors wish to thank Molly and Chris MacLellan of the NERC Field Spectroscopy Facility for experimental help and invaluable insights in the flat-field and spectral response measurements. Figure 3.1 was drawn using *draw.io*. Data analysis and visualisation were done using the AstroPy, ExifRead, Matplotlib, NumPy, RawPy, and SciPy libraries for Python. Finally, the authors wish to thank the two anonymous reviewers for their thorough and constructive reviews.

4 | Accuracy and reproducibility of above-water radiometry with calibrated smartphone cameras using RAW data

Adapted from

Olivier Burggraaff, Mortimer Werther, Emmanuel Boss, Stefan Simis, Frans Snik

Frontiers in Remote Sensing, 3, 940 096 (2022)

doi:10.3389/frsen.2022.940096

Consumer cameras, especially on smartphones, are popular and effective instruments for above-water radiometry. The remote sensing reflectance R_{rs} is measured above the water surface and used to estimate inherent optical properties and constituent concentrations. Two smartphone apps, HydroColor and EyeOnWater, are used worldwide by professional and citizen scientists alike. However, consumer camera data have problems with accuracy and reproducibility between cameras, with systematic differences of up to 40% in intercomparisons. These problems stem from the need, until recently, to use JPEG data. Lossless data, in the RAW format, and calibrations of the spectral and radiometric response of consumer cameras can now be used to significantly improve the data quality. Here, we apply these methods to above-water radiometry. The resulting accuracy in R_{rs} is around 10% in the red, green, and blue (RGB) bands and 2% in the RGB band ratios, similar to professional instruments and up to 9 times better than existing smartphone-based methods. Data from different smartphones are reproducible to within measurement uncertainties, which are on the percent level. The primary sources of uncertainty are environmental factors and sensor noise. We conclude that using RAW data, smartphones and other consumer cameras are complementary to professional instruments in terms of data quality. We offer practical recommendations for using consumer cameras in professional and citizen science.

4.1 Introduction

The remote sensing reflectance $R_{rs}(\lambda)$ is an apparent optical property that contains a wealth of information about the substances within the water column [2]. In above-water radiometry, R_{rs} is measured using one or more (spectro)radiometers deployed above the water surface [120]. The absorption and scattering coefficients and concentrations of coloured dissolved organic matter (CDOM), suspended particulate matter, and prominent phytoplankton pigments such as chlorophyll-a (chl-a) can be determined from R_{rs} [234]. Due to spectral range and long-term stability requirements, the equipment necessary for accurate measurements of R_{rs} is often expensive. High costs limit the uptake and, therefore, impact of these instruments.

Consumer cameras have long been seen as a low-cost alternative or complement to professional instruments [102]. Work in this direction has mostly focused on hand-held digital cameras, which measure the incoming radiance in red-green-blue (RGB) spectral bands typically spanning the visible range from 390–700 nm [323]. Uncrewed aerial vehicles (UAVs or drones) and webcams have similar optical properties, often contain the same sensors, and are also increasingly used in remote sensing [281]. Consumer cameras have been used to retrieve CDOM, chl-a, and suspended mineral concentrations through above-water radiometry [34, 323]. They are particularly useful for measuring at small spatial scales, short cadence, and over long time periods [386, 387].

Smartphones are especially effective as low-cost sensing platforms thanks to their wide availability, cameras, and functionalities including accelerometers, GPS, and wireless communications. They are already commonly used in place of professional sensors in laboratories [56, 388]. However, what smartphones truly excel at is providing a platform for citizen science in the field [94, 389]. There is a vibrant ecosystem of applications (apps) using the smartphone camera for environmental citizen science purposes [273]. Some use additional fore-optics to measure hyperspectrally [284, 390], while most use the camera as it is [121,274,391]. Smartphone science apps are also commonly used for educational purposes and in professional research [318, 392, 393].

Two apps are currently widely used for above-water radiometry, namely HydroColor [121] and EyeOnWater [274]. HydroColor measures R_{rs} in the RGB bands using the Mobley protocol [209], guiding the user to the correct pointing angles with on-screen prompts. Through an empirical algorithm based on the red band of R_{rs}, the app estimates the turbidity, suspended matter concentration, and backscattering coefficient of the target body of water. EyeOnWater uses the WACODI algorithm [285] to determine the hue angle α of the water, representing its intrinsic colour. From α it also estimates the Forel-Ule (FU) index, a discrete water colour scale with a century-long history [306]. α and the FU index are reasonable first-order indicators of the surface chl-a concentration and optical depth [10].

While these apps and other consumer camera-based methods provide useful data, improvements to the accuracy and reproducibility are necessary to derive high-quality end products. Validation campaigns have consistently found the radiance, R_{rs} in the RGB bands, and hue angle from consumer cameras to be well-correlated with reference instruments, but often with a wide dispersion and a significant bias. For R_{rs}, the mean difference between smartphone and reference match-up data is typically ≥ 0.003 sr^{-1} or $\geq 30\%$, but varies wildly between studies [121, 276, 391, 394]. As an extreme example, Malthus et al. found no correlation at all between HydroColor and reference R_{rs} data [275], albeit under challenging observing conditions. The typical accuracy in α is around 10° or 1–2 FU [275,285,391,395]. Differences in R_{rs} between smartphones can be as large as 40% [276]. The uncertainties,

as well as the differences between smartphones and reference instruments, in observed optical properties and derived water constituent concentrations are often even greater than 40% [275, 396, 397], although this may be explained in part by differences in inherent optical properties and observing conditions between study sites.

A major source of uncertainty in existing methods is the use of the JPEG data format. Until recently JPEG was the only format available to third-party developers on most smartphones and other consumer cameras. JPEG data are irreversibly compressed and post-processed for visual appeal, at the cost of radiometric accuracy and dynamic range. Most importantly, they are very non-linear, meaning a 2× increase in radiance does not cause a 2× increase in response [281]. Instead, in a process termed gamma correction or gamma compression, the radiance is scaled by a power law. The non-linearity of JPEG data is a significant contributor to the uncertainty in R_{rs} obtained from consumer cameras and apps such as HydroColor [275, 281, 394]. Some approaches, including WACODI, attempt to correct for non-linearity through an inverse gamma correction [285, 394]. This inverse correction cannot be performed consistently because the smartphone JPEG processing differs between smartphone brands, models, and firmware versions [281].

A secondary source of uncertainty are the spectral response functions (SRFs) of the cameras. Because exact SRF profiles are laborious to measure and are rarely provided by manufacturers, it is often necessary to use simplified SRFs and assume them to be device-independent [121, 285]. However, the SRFs of different cameras actually vary significantly [281].

The quality of consumer camera radiometry can be improved significantly by using lossless data, in the RAW format, and camera calibrations. RAW data are almost entirely unprocessed and thus are not affected by the uncertainties introduced by the JPEG format. Furthermore, through calibration and characterisation of the radiometric and spectral response, consumer cameras can be used as professional-grade (spectro)radiometers [281].

In this work, we assess the uncertainty, reproducibility, and accuracy of calibrated smartphone cameras, using RAW data, for above-water radiometry. By comparing in situ observations from two smartphone cameras and two hyperspectral instruments, we test the hypothesis that the new methods decrease the uncertainty and increase the reproducibility and accuracy of data from consumer cameras. To our knowledge, this is the first time that the new methods have been applied or assessed in a field setting.

Section 4.2 describes the data acquisition and processing as well as the performed experiments. The results are presented in Section 4.3. In Section 4.4, we discuss the results, compare them to the literature, and present some recommendations for projects using smartphones. Finally, the conclusions of the analysis are presented in Section 4.5.

4.2 Methods

Smartphone and reference data were gathered on and around Lake Balaton, Hungary, from 3–5 July 2019. Lake Balaton is the largest (597 km^2) lake in central Europe, with a mean depth of only 3.3 m, and is well-studied. It has a high concentration of suspended mineral particles and appears very bright and turquoise (bluish-green) to the eye (Figure 4.1, further discussed in Section 4.2.1). Due to inflow from the Zala river, the western side of the lake is richer in nutrients than the eastern side. The adjacent Kis-Balaton reservoir is hypereutrophic with chl-a concentrations up to 160 mg m^{-3}. More detailed descriptions of this site are given

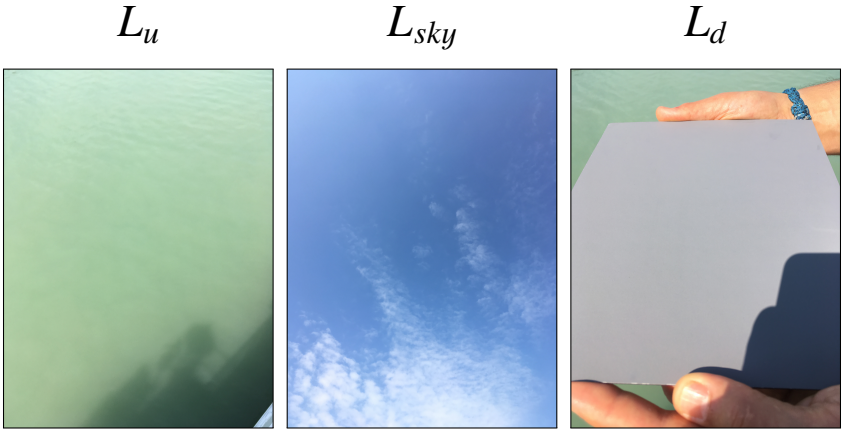


Figure 4.1: Example iPhone SE images of L_u, L_{sky}, and L_d, taken at Lake Balaton on 3 July 2019 at 07:47 UTC. Little wave motion is visible on the water surface in L_u, while L_{sky} shows patchy cloud coverage. The conditions seen here were representative for the entire campaign.

in [27, 264].

Two smartphones were used, an Apple iPhone SE and a Samsung Galaxy S8, and two hyperspectral spectroradiometer instruments were used as references. The reference instruments were a set of three TriOS RAMSES instruments mounted on a prototype Solar-tracking Radiometry (So-Rad) platform [123] to maintain a favorable viewing geometry throughout the day, and a hand-held Water Insight WISP-3 spectroradiometer [126]. The spectral and radiometric calibration of the smartphones is described in [281]; manufacturer calibrations were used for the So-Rad and WISP-3.

Data processing and analysis were done using custom Python scripts based on the NumPy [312], SciPy [398], and SPECTACLE [281] libraries, available from GitHub[15]. The smartphone data processing pipeline supports RAW data from most consumer cameras. The processing of the reference and smartphone data is further discussed in Sections 4.2.2–4.2.4, the analysis in Sections 4.2.5 and 4.2.6.

4.2.1 Data acquisition

Observations were performed on 3 July 2019 from the Tihany-Szántód ferry on eastern Lake Balaton, performing continuous transects around 46°53'00"N 17°53'43"E, facing southwest before 10:00 UTC (12:00 local time) and northeast afterwards. Data were also acquired on 4 July in the Kis-Balaton reservoir at 46°39'41"N 17°07'45"E and on 5 July on western Lake Balaton at 46°45'15"N 17°15'09"E, 46°42'25"N 17°15'53"E, 46°43'59"N 17°16'34"E, and 46°45'04"N 17°24'46"E. The So-Rad, which was mounted on the ferry, was only used in the morning on 3 July; the two smartphones and WISP-3 were used at all stations. All data, including a detailed station log, are available from Zenodo[16].

The upwelling radiance L_u, sky radiance L_{sky}, and either downwelling radiance L_d (smartphones) or downwelling irradiance E_d (references) were measured. The So-Rad and WISP-3 data were hyperspectral, the smartphones multispectral in different RGB bands [281]. A

[15] https://github.com/burggraaff/smartphone-water-colour
[16] https://dx.doi.org/10.5281/zenodo.4549621

Brandess Delta 1 18% gray card was used to measure L_d, which is discussed in Section 4.2.3. The observations on 3 and 5 July were done under a partially clouded sky (Figure 4.1), which introduced uncertainties in L_{sky} and R_{rs} by increasing the variability of the sky brightness and causing cloud glitter effects on the water surface [209]. Simultaneous measurements from different instruments were affected in the same way, meaning an intercomparison was still possible. However, for measurements taken farther apart in time and space, the match-up error may be significant. On 4 July, the sky was overcast.

Following standard procedure [120, 209], the smartphone observations were performed pointing 135° away from the solar azimuth in the direction furthest from the observing platform and 40° from nadir (L_u, L_d) or zenith (L_{sky}). The smartphones were taped together and aligned in azimuth by eye and in elevation using the tilt sensors in the iPhone SE, to approximately 5° precision. Example smartphone images are shown in Figure 4.1. The same viewing geometry is used in HydroColor, but not EyeOnWater [275]. The reference observations were performed in the same way, following standard procedure for the respective sensors [90, 126].

The So-Rad and WISP-3 each recorded L_u, L_{sky}, and E_d simultaneously while the smartphones took sequential L_u, L_d, and L_{sky} images within one minute. Using the SPECTACLE apps for iOS and Android smartphones [281], the iPhone SE took one RAW image and one JPEG image simultaneously, and the Galaxy S8 took 10 sequential RAW images per exposure. The exposure settings on both smartphones were chosen manually to prevent saturation and were not recorded, but were kept constant throughout the campaign.

In total, 304 and 453 sets of WISP-3 and So-Rad spectra, respectively, and 28 sets each of iPhone SE and Galaxy S8 images were obtained. For the WISP-3, one set of spectra (5 July at 10:35:51 UTC) was manually removed because it appeared excessively noisy. Six sets of smartphone data were discarded due to saturation.

4.2.2 Reference data processing

R_{rs} spectra were calculated from the WISP-3 and So-Rad data (Figure 4.2). For the WISP-3, the Mobley method [209] shown in Equation (4.1), with a sea surface reflectance factor of $\rho = 0.028$, was used. Wavelength dependencies are dropped for brevity. The value of $\rho = 0.028$ was chosen for the WISP-3 and smartphone data processing (Section 4.2.3) to enable a direct comparison to HydroColor, which uses the same value [121]. Given the brightness of Lake Balaton, the relative magnitude of ρL_{sky} compared to L_u was small (typically <5% in the WISP-3 data) for any value of ρ around 0.03, and thus the effect of a small difference in ρ on R_{rs} was negligible. The So-Rad data, having a wider spectral range, were processed using the three-component (3C) method, which subtracts an additional glint term Δ and determines ρ empirically from a spectral optimisation [210, 399].

$$R_{rs} = \frac{L_u - \rho L_{sky}}{E_d} \tag{4.1}$$

The general appearance of the reflectance spectra (Figure 4.2) is that of a broad peak around 560 nm. On the short wavelength side of this peak, absorption by phytoplankton and CDOM suppresses R_{rs} to approximately 25% of the peak amplitude. Towards longer wavelengths, the effects of increasing absorption by water are clearly seen around 600 nm and beyond 700 nm, and R_{rs} reaches near-zero amplitude at the edge of the visible spectrum. The reflectance is ultimately skewed towards blue-green wavelengths, giving the water a turquoise appearance. A minor absorption feature of chl-a and associated accessory pigments is visible

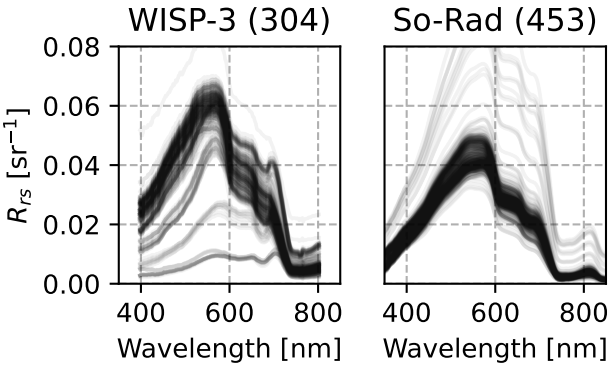


Figure 4.2: Reference R_{rs} spectra derived from measurements on and around Lake Balaton. There is a difference in normalisation between the two data sets, which is discussed in Section 4.4.3.

around 675 nm. Sun-induced chl-a fluorescence is visible at 680–690 nm in the WISP-3 spectra taken on 4 and 5 July, but not the WISP-3 or So-Rad spectra taken on 3 July.

4.2.3 Smartphone data processing

The RAW smartphone images were processed using a SPECTACLE-based [281] pipeline (Figure 4.3). The images were first corrected for bias or black level, which shifts the pixel values in each image by a constant amount. On the Galaxy S8, the nominal black level was 0 analogue-digital units (ADU), while on the iPhone SE it was 528 ADU or 13% of the dynamic range, as determined from the RAW image metadata and validated experimentally [281]. Next, a flat-field correction was applied, correcting for pixel-to-pixel sensitivity variations. The sensitivity varies by up to 142% across the iPhone SE sensor [281], although in the central 100×100 pixels, the variations are only 0.2% on the iPhone SE and 1.6% on the Galaxy S8. A central slice of 100×100 pixels was taken to decrease the uncertainties introduced by spatial variations across the image [121]. The central pixels were then demosaicked into separate images for the $RGBG_2$ channels, where G_2 is the duplicate green channel present in most consumer cameras [281]. The $RGBG_2$ images were flattened into lists of 10 000 samples per channel and normalised by the effective spectral bandwidths of the channels, determined from the SRFs [281]. The mean radiance was calculated per channel, after which the G and G_2 channels, which have identical SRFs, were averaged together. Finally, R_{rs} was calculated from L_u, L_{sky}, and L_d using Equation (4.2) [209]. Like for the WISP-3 (Section 4.2.2) and in HydroColor, a constant $\rho = 0.028$ was used. R_{ref} is the gray card reference reflectance, nominally 0.18.

$$R_{rs} = \frac{L_u - \rho L_{sky}}{\frac{\pi}{R_{ref}} L_d} \tag{4.2}$$

For R_{ref}, a Brandess Delta 1 18% gray card was used by manually holding it horizontal in front of the camera. The nominal reflectance of $R_{ref} = 18\%$ was verified to within 0.5 percent point in the smartphone RGB bands by comparing spectroradiometer measurements of L_d on a similar gray card to cosine collector measurements of E_d. Angular variations in R_{ref} were found to be $\lesssim$1 percent point for nadir angles of 35°–45° in a laboratory experiment

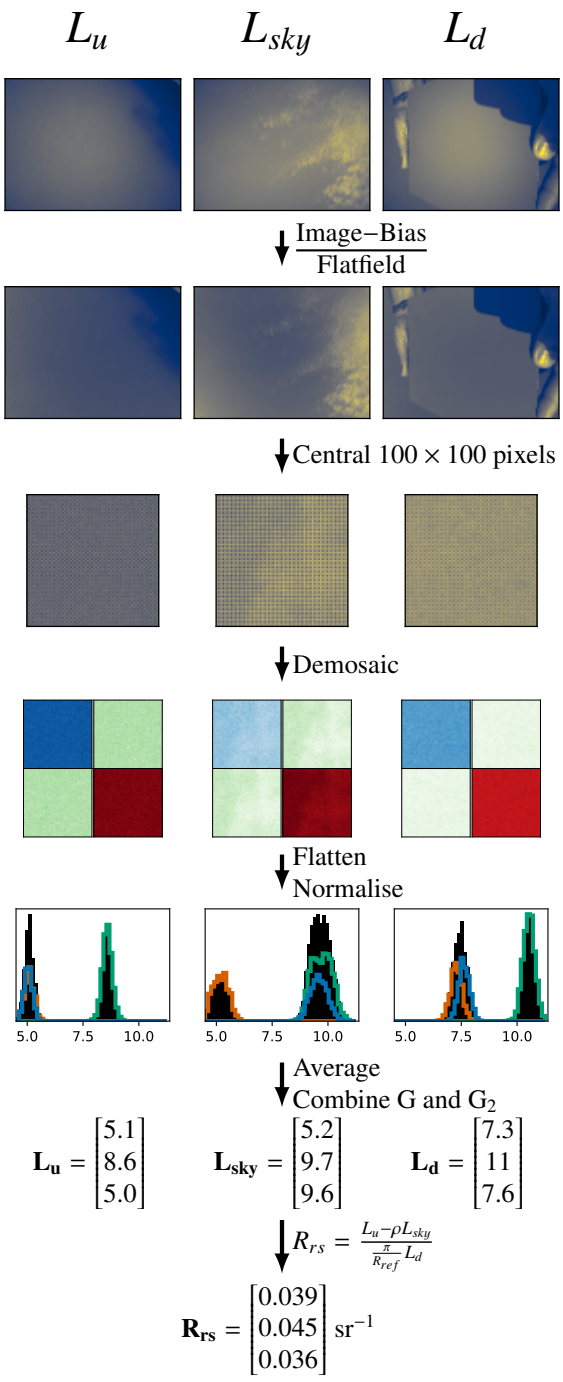


Figure 4.3: Smartphone data processing pipeline, from RAW images to multispectral R_{rs}. The example input images are those from Figure 4.1. Some processing steps have been combined for brevity. The histograms show the distribution of normalised pixel values in the central 100 × 100 pixels for the RGBG_2 channels separately (coloured lines, G and G_2 combined) and together (black bars). The order of elements in **L** and $\mathbf{R_{rs}}$ is RGB.

with the iPhone SE. This value is similar to previous characterisations of different consumer-grade gray cards [400]. To account for these factors as well as fouling, an uncertainty of $\sigma_{R_{ref}} = 0.01$, or 1 percent point, was used in our data processing. This does not account for systematic errors (Section 4.4.3).

Unlike EyeOnWater, which selects multiple sub-images from different parts of each image, our pipeline only used a central slice of 100×100 pixels. The use of sub-images was not necessary since all images were manually curated and sub-imaging has been shown to have little impact on the data quality [275]. The 100×100 size was chosen to minimise spatial variations, but a comparison of box sizes from 50–200 pixels showed that the exact size made little difference. For example, the mean radiance typically varied by <0.4%, less than the typical uncertainty on the radiance estimated from each image (Section 4.3.1). Furthermore, the signal-to-noise ratio (SNR) varied by <3% for L_u and L_d but up to 19% in L_{sky} due to the patchy cloud coverage.

The iPhone SE JPEG data were processed using a simplified version of the RAW pipeline, lacking the bias and flat-field corrections and G-G_2 averaging. Smartphone cameras perform these three tasks internally for JPEG data [281]. The processing was repeated with an additional linearisation step, like in WACODI and EyeOnWater, to determine whether linearisation improves the data quality. Following WACODI, the default sRGB inverse gamma curve was used, although this curve has already been shown to be poorly representative of real smartphones [281].

The uncertainties in the image data, determined from the sample covariance matrix of the 10 000 pixels per channel per image, were propagated analytically as described in Section 4.A. The pixel values were approximately normally distributed (Figure 4.3). Significant correlations between the RGBG_2 channels were found. For example, the iPhone SE L_{sky} image from 3 July 2019 at 07:47 UTC had a correlation of $r_{RG} = 0.68$ between R and G, while in the 08:01 image this was only $r_{RG} = 0.09$. The observed correlations were likely due to spatial structures in the images [140], such as patchy clouds for L_{sky} and waves for L_u. In larger data sets, the presence of strong correlations could be used as a means to filter out images that are not sufficiently homogeneous. The propagated uncertainties in R_{rs} were typically 5–10% of the mean R_{rs} and similarly correlated between channels. For example, the 07:47 data had correlations in R_{rs} of $r_{RG} = 0.67$, $r_{RB} = 0.57$, and $r_{GB} = 0.72$.

4.2.4 Colour

In addition to absolute R_{rs} in the RGB bands, several relative colour measurements were investigated, namely RGB band ratios, hue angle, and FU index.

The band ratios were calculated as specific combinations of R_{rs} bands. For simplicity in notation, the ratios are expressed as, for example, G/R instead of $R_{rs}(G)/R_{rs}(R)$. Following the literature, the numerators and denominators were chosen as G/R, B/G, and R/B. The G/R ratio is sensitive to water clarity and optical depth [391]. B/G is sensitive to the chl-a concentration [323], at least in water types where phytoplankton covaries with other absorbing substances. Finally, the R/B ratio is particularly sensitive to broad features such as CDOM absorption, as well as the concentration of scatterers (turbidity, suspended matter concentrations), as described in [34, 323].

To calculate the hue angle, the data were first transformed to the CIE XYZ colour space. CIE XYZ is a standard colour space representing the colours that a person with average colour vision can experience [101]. The reference data were spectrally convolved with the XYZ

colour matching functions [401]. The spectral convolution was applied directly to R_{rs}, since R_{rs} represents the true colour of the water [283]. For the smartphone data, transformation matrices calculated from the smartphone camera SRFs (Section 4.B) were used [402, 403]. These matrices are given in Equations (4.3) and (4.4). The uncertainties on the matrix elements were not included since this would require a full re-analysis of the raw SRF data [404], which is outside the scope of this work. The resulting colours were relative to an E-type (flat spectrum, $x = y = 1/3$) illuminant.

$$\mathbf{M}_{RGB\to XYZ}^{iPhone\ SE} = \begin{bmatrix} 0.5709 & 0.2452 & 0.1839 \\ 0.3760 & 0.4346 & 0.1894 \\ 0.0439 & 0.0913 & 0.8648 \end{bmatrix} \tag{4.3}$$

$$\mathbf{M}_{RGB\to XYZ}^{Galaxy\ S8} = \begin{bmatrix} 0.5611 & 0.1451 & 0.2938 \\ 0.3944 & 0.2391 & 0.3666 \\ 0.0231 & 0.0416 & 0.9353 \end{bmatrix} \tag{4.4}$$

From XYZ, the chromaticity (x, y) and hue angle α were calculated as shown in Equations (4.5) and (4.6). Chromaticity is a normalisation of the XYZ colour space that removes information on brightness [101]. The FU index was determined from α using a look-up table [10, 306]. The uncertainties in R_{rs} were propagated analytically into XYZ and (x, y), as described in Section 4.A. However, further propagation into α was not feasible, since the linear approximation of Equation (4.6) breaks down near the white point $(x, y) = (1/3, 1/3)$, especially with highly correlated x and y [405].

$$x = \frac{X}{X + Y + Z} \qquad y = \frac{Y}{X + Y + Z} \tag{4.5}$$

$$\alpha = \arctan2\,(y - 1/3, x - 1/3) \mod 2\pi \tag{4.6}$$

4.2.5 Replicate analysis

The Galaxy S8 data were taken in sets of 10 sequential replicates per image (Section 4.2.1). The variability between these replicates was analysed to assess the uncertainty in smartphone data.

The processing chain described in Section 4.2.3 was applied to each image from each set, resulting in 10 measurements per channel of L_u, L_{sky}, and L_d. R_{rs} was calculated from each combination of images, resulting in 1 000 values. From these, the band ratios, α, and FU were calculated.

The relative uncertainty in L_u, L_{sky}, L_d, R_{rs}, and the band ratios was estimated through the coefficient of variation $\frac{\sigma}{\mu}$, σ being the standard deviation and μ the mean value. Because α and FU have arbitrary zero-points, relative uncertainties are not applicable to them, and σ was instead used to estimate the absolute uncertainty.

4.2.6 Match-up analysis

Simultaneous data taken with the various sensors were matched up and compared. There were 27 pairs of iPhone SE and Galaxy S8 images, taken on average 50 s apart. On the ferry, which

had an average speed of 8 km/h, a 50 s delay corresponded to a distance along the transect of approximately 120 m. The smartphone images were also matched to reference spectra taken within a 10-minute time frame, resulting in 1–41 reference spectra per match-up. The reference R_{rs} spectra were convolved to the smartphone RGB bands by first convolving the reference radiances [283]. For match-ups with multiple reference spectra per smartphone image, the median value of each variable in the reference spectra was used, with the standard deviation as an estimate for the uncertainty. For match-ups with a single reference spectrum per smartphone image, the uncertainty was instead estimated as the median uncertainty on the multiple-spectrum match-ups, for each variable. Match-up reference spectra with large uncertainties, for example relative uncertainties of >10% in R_{rs}, were not discarded because these represent common measurement scenarios.

The match-up data were compared using the metrics shown in Equations (4.7)–(4.10). Here P, Q are any two data sets with elements p_i, q_i; $\mathrm{cov}(P, Q)$ is their covariance; σ_P, σ_Q are the standard deviations in P and Q, respectively; Med_i is the median evaluated over the indices i; and sgn is the sign function. The RGB channels were treated as separate samples, as were the three band ratios.

$$r = \frac{\mathrm{cov}(P, Q)}{\sigma_P \sigma_Q} \tag{4.7}$$

$$\mathcal{M} = \mathrm{Med}_i \left(|q_i - p_i| \right) \tag{4.8}$$

$$\zeta = \exp \left[\mathrm{Med}_i \left(\left| \ln \frac{q_i}{p_i} \right| \right) - 1 \right] \tag{4.9}$$

$$\mathcal{B} = \mathrm{sgn} \left[\mathrm{Med}_i \left(\ln \frac{q_i}{p_i} \right) \right] \times \left[\exp \left(\left| \mathrm{Med}_i \left(\ln \frac{q_i}{p_i} \right) \right| \right) - 1 \right] \tag{4.10}$$

The Pearson correlation r and median absolute deviation $\mathcal{M}$ are well-known [406, 407]. The median symmetric accuracy ζ and signed symmetric percentage bias $\mathcal{B}$, both expressed as a percentage, are recent introductions, which we chose to use for their robustness, symmetry, and ability to span multiple orders of magnitude in the data [406]. r expresses the degree of linear correlation between variables, from −1 to 1, but is sensitive to outliers and the data range. $\mathcal{M}$ and ζ measure the typical random error or dispersion between variables in absolute and relative terms, respectively. Both are robust to outliers. $\mathcal{B}$ is similar to ζ but measures the bias towards over- or underestimation. The covariance, standard deviations, and median calculated in r and $\mathcal{M}$ were weighted by $w_i = \frac{1}{\sigma_{p_i}^2 + \sigma_{q_i}^2}$. ζ and $\mathcal{B}$ are unweighted.

The FU indices were also compared by the number of matches [395, 407], considering both full ($\Delta\mathrm{FU} = 0$) and near-matches ($\Delta\mathrm{FU} \leq 1$). The typical uncertainty on human observations is 1 FU [280].

5–95% confidence intervals (CIs) on the metrics were estimated by bootstrapping over pairs of (p_i, q_i), and w_i if applicable. Bootstrapping involves randomly resampling the data with replacement, mimicking the original sampling process [408]. This was necessary to account for the relatively small size of our data set, which increases the effects of outliers, even on robust metrics like $\mathcal{M}$ or ζ. The bootstraps were evaluated with 9 999 resamples, sufficient to obtain consistent results matching the analytical formula for CIs on unweighted r to 4 decimals [408].

Some data were also compared through a linear regression ($y = ax + b$ with parameters a, b), to convert data to the same units or account for normalisation differences. The regres-

sion was done through the `scipy.odr` function for orthogonal distance regression, which minimises differences and accounts for weights on both axes. The same process was used to fit a power law ($y = ax^b$) in the JPEG data comparison (Section 4.3.4).

4.3 Results

4.3.1 Replicate analysis

The Galaxy S8 replicate analysis showed that among the radiances, L_u had the largest relative variability with a quartile range (QR, the 25–75% percentile range of variability among the sets of replicate observations) of 1.8–5.8%, followed by L_{sky} with 1.1–3.4%, and L_d with 0.4–1.2% (Figure 4.4). L_u and L_{sky} were affected primarily by cloud and wave movement, shaking of the camera, and movement of the ferry on 3 July. Therefore, the variability in L_u and L_{sky} was largely methodological in nature, as discussed further in Section 4.4.1. Since L_d was measured on a bright, stable gray card, it was not affected by the above factors, and its variability best represented the radiometric stability of the smartphone camera.

The RGB R_{rs} varied by 1.9–8.1%, while the R_{rs} band ratios only varied by 0.5–1.9%. The difference can be explained by correlations between channels. For example, wave movements between successive images affected all three RGB channels of L_u equally, changing the individual R_{rs} values, but having little effect on their ratios. The same held true for other environmental variations and camera stability issues.

Finally, there was a variability in hue angle α of 2.1°–6.8° and in FU index of 0.19–0.62 FU. The variability distributions of α and FU index did not have the same shape because the hue angle difference between successive FU indices varies greatly.

The variability between replicates represents the typical uncertainty associated with random effects on our data. However, there are some caveats. First, systematic effects such as an error in R_{ref} would affect successive measurements equally, and not cause random variations. Second, the uncertainty in individual images may be larger due to spatial structures, which the uncertainty propagation described in Section 4.2.3 does account for. Both of these issues explain differences between the replicate and propagated uncertainties in our data. For

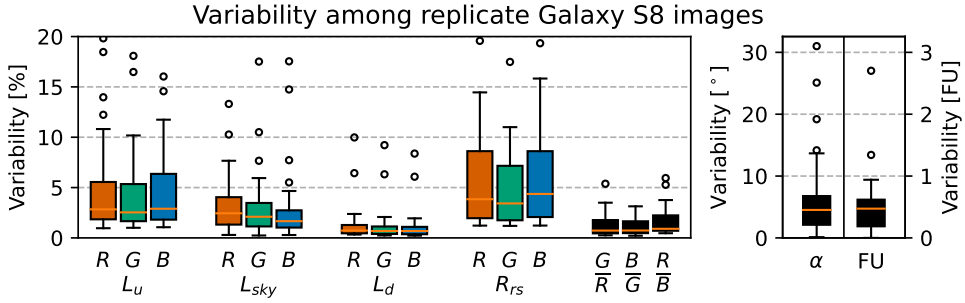


Figure 4.4: Variability in radiance, R_{rs}, and colour between replicate Galaxy S8 images. The boxes show the distribution, among 27 individually processed sets of 10 replicates, of the variability between replicate images. The orange lines indicate the medians, the boxes span the quartile range (QR), the whiskers extend to 1.5 times the QR, and circles indicate outliers. Up to two outliers per column fell outside the y-axis range.

example, the propagated uncertainty in individual images was 6.6–9.0% for RGB R_{rs} and 4.5–7.0% for the band ratios. While the exact uncertainties will differ between campaigns, sites, and even smartphones, the trends seen here can be generalised.

As a point of comparison, the uncertainty QRs for the spectrally convolved WISP-3 data in the Galaxy S8 match-up (Section 4.3.3), were 4.2–38% in L_u, 4.8–14% in L_{sky}, 2.5–30% in E_d, 2.6–7.2% in RGB R_{rs}, 0.7–2.9% in R_{rs} band ratios, 0.4°–2.8° in α, and 0–0.46 in FU. While the Galaxy S8 and WISP-3 variability cannot be compared 1:1 due to differences in data acquisition and processing and in the uncertainty estimation, the order of magnitude of the uncertainties in the Galaxy S8 and WISP-3 reference data was the same.

4.3.2 Smartphone comparison

There was a strong correlation, $r = 0.94$ (CI 0.90, 0.96), between the iPhone SE and Galaxy S8 radiances (Figure 4.5). Due to differences in exposure settings, both cameras measured radiance in different, arbitrary units (a.u.). After re-scaling the Galaxy S8 data through a linear regression (Section 4.2.6), the median absolute deviation was $\mathcal{M} = 0.39$ (CI 0.29, 0.52) in iPhone SE units and the median symmetric accuracy was $\zeta = 6.9\%$ (CI 5.1%, 8.7%). The value of ζ was comparable to the variability between replicate images (Section 4.3.1).

The R_{rs} match-ups between the two smartphones, in both RGB (Figure 4.6) and band ratios (Figure 4.7), showed excellent agreement. The data were strongly correlated, with $r = 0.98$ (CI 0.95, 0.99) for RGB and $r = 0.99$ (CI 0.99, 1.00) for band ratio R_{rs}. The typical difference in RGB R_{rs} was $\mathcal{M} = 0.0010$ (CI 0.0005, 0.0013) sr^{-1} or $\zeta = 5.5\%$ (CI 3.8%, 8.2%). For band ratios, the typical difference was $\mathcal{M} = 0.032$ (CI 0.026, 0.035), unitless, and $\zeta = 2.9\%$ (CI 2.3%, 3.7%). Both values of ζ are consistent with Section 4.3.1, as is the observation that band ratios are more reproducible than RGB R_{rs}. Finally, the signed symmetric percentage bias in RGB R_{rs}, $\mathcal{B} = -2.7\%$ (CI −7.0%, −1.8%), was smaller than the typical uncertainty. There was no significant offset in the band ratios, with $\mathcal{B} = -1.1\%$ (CI −1.8%, +0.7%).

The agreement in α and FU was poorer but still similar to the expected uncertainties (Figure 4.8). The typical difference was $\mathcal{M} = 8.3°$ (CI 5.0°, 11°) in α and $\mathcal{M} = 1$ (CI 0, 2) in FU index. 33% (CI 15%, 48%) of the match-up pairs had the same FU index, 59% (CI 37%, 74%) had a difference $\Delta\text{FU} \leq 1$. The wide CIs are due to the relatively small number ($N = 27$) of match-ups. The data did not span the full range of α, but were mostly concentrated into two clusters, around 50° (FU 14–16, greenish brown) and 90° (FU 8–9, bluish green). Interestingly, while the 90° cluster was centred roughly on the 1:1 line, the 50° cluster fell entirely underneath it. However, due to the small N and the uncertainties on the data, it is difficult to say whether this was significant.

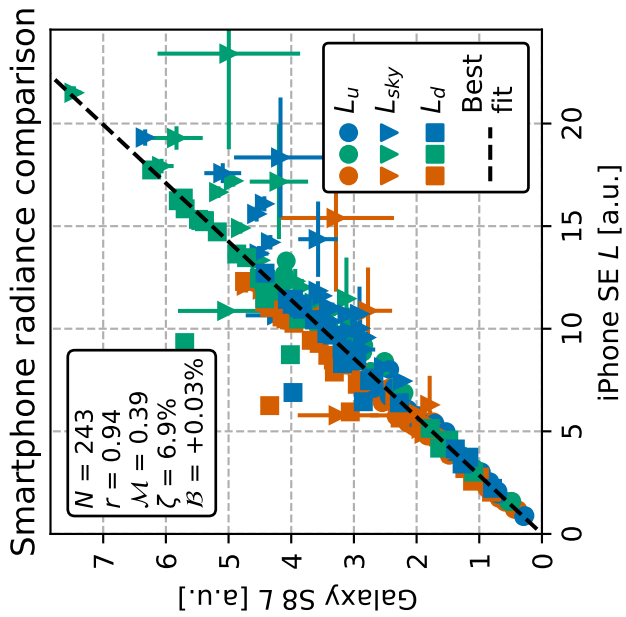


Figure 4.5: Comparison between iPhone SE and Galaxy S8 radiance measurements. The axes are in different units due to differences in exposure settings. The RGB channels are shown in their respective colours, with different symbols for L_u, L_{sky}, and L_d. The statistics in the text box are relative to the regression line.

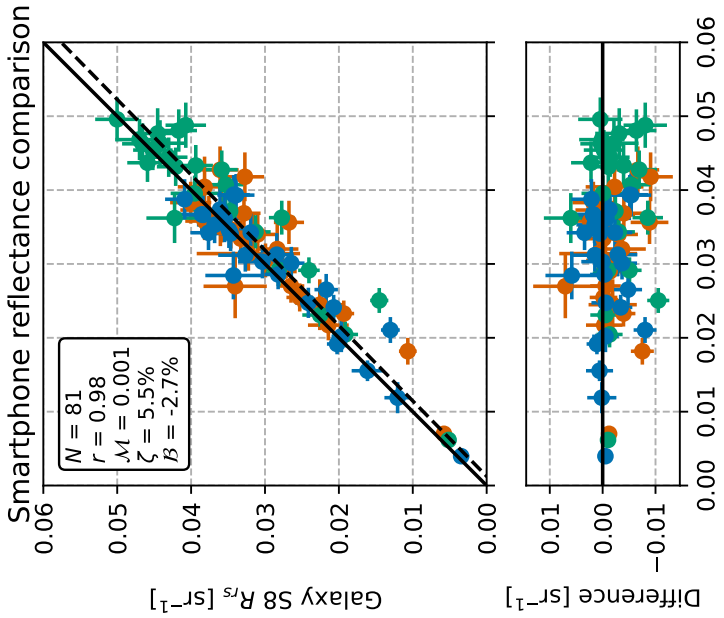


Figure 4.6: Comparison between iPhone SE and Galaxy S8 R_{rs} measurements in the RGB bands. The solid line corresponds to a 1:1 relation, the dashed line is the best-fitting linear regression line. The statistics in the text box are based on a 1:1 comparison, as are the differences in the lower panel.

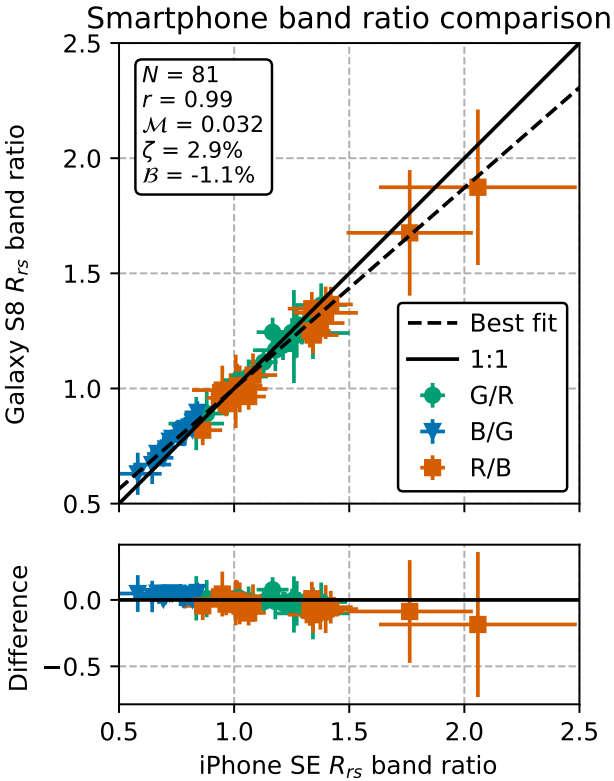


Figure 4.7: Comparison between iPhone SE and Galaxy S8 R_{rs} band ratios. The solid line corresponds to a 1:1 relation, the dashed line is the best-fitting linear regression line. The statistics in the text box are based on a 1:1 comparison, as are the differences in the lower panel.

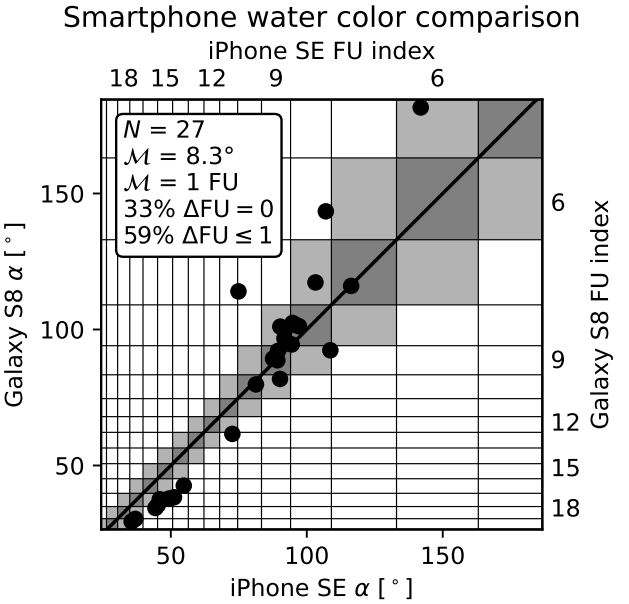


Figure 4.8: Comparison between iPhone SE and Galaxy S8 measurements of hue angle and FU index. The solid line corresponds to a 1:1 relation. The dark gray squares indicate a full FU match, the light gray ones a near-match. Accurate uncertainties on individual points could not be determined (Section 4.2.4). The statistics in the text box are based on a 1:1 comparison.

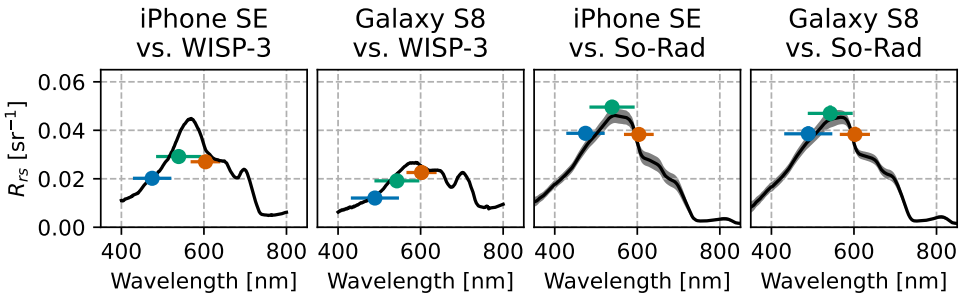


Figure 4.9: Examples of smartphone vs. reference R_{rs} match-ups at different stations. The solid lines show the reference spectrum, with uncertainties in gray. The RGB dots show the smartphone data, with error bars indicating the effective bandwidth (horizontal) and R_{rs} uncertainty (vertical). In some panels, the vertical error bars are smaller than the data point size.

4.3.3 Smartphone vs. reference comparison

A total of 72 pairs of smartphone vs. reference match-up spectra were analysed, four of which are shown in Figure 4.9. There were 27 match-ups between the WISP-3 and each smartphone and 9 between the So-Rad and each smartphone. Except for the normalisation difference that was also present between the So-Rad and WISP-3 (Figure 4.2, discussed in Section 4.4.3), there was good agreement between the instruments (Figure 4.9).

The full statistics of the match-up analysis are given in Table 4.1. The correlation between smartphone and reference radiance was $r \geq 0.71$ in all pairs of instruments (Figure 4.10). The median symmetric accuracy ζ ranged between 12–19%, larger than the typical uncertainties and the value from the smartphone vs. smartphone comparison. This larger difference in observed radiance is not surprising, since the smartphone vs. reference match-ups typically differed more in time and location than the smartphone vs. smartphone match-ups. No significant differences in the match-up statistics between the individual RGB bands were found.

The RGB R_{rs} data were strongly correlated between smartphone and reference sensors ($r \geq 0.94$ for the WISP-3) and showed a relatively small dispersion, although with a normalisation difference in the WISP-3 comparisons (Figure 4.11), similar to that between the WISP-3 and So-Rad data (Figure 4.2). To negate the normalisation issue, the smartphone data were re-scaled based on a linear regression (Section 4.2.6) for the smartphone vs. WISP-3 RGB R_{rs} comparison. The So-Rad and smartphone data were compared 1:1. The typical differences in R_{rs}, then, were on the order of 10^{-3} sr^{-1} for the So-Rad and 10^{-4} sr^{-1} for the WISP-3, differing mostly due to their different ranges. The difference in range of R_{rs} also decreased the correlation coefficient r for the So-Rad comparisons. In the four smartphone vs. reference R_{rs} comparisons, ζ was between 9–13%, twice the value seen in the smartphone vs. smartphone comparison but similar to the differences between smartphone and reference radiances.

The agreement between smartphone and reference R_{rs} band ratios was better than the agreement in RGB R_{rs} (Figure 4.12). In all four band ratio comparisons, the correlation was near-perfect ($r \geq 0.97$), and the typical differences ($1.1\% \leq \zeta \leq 3.8\%$) were consistent with the uncertainties in the data. The WISP-3 normalisation difference did not affect this comparison since it divided out.

The agreement in α and FU was not as good as that in L and R_{rs}, like in the smartphone

		N	r, $\mathcal{M}(\alpha)$	$\mathcal{M}$	ζ, ΔFU = 0	$\mathcal{B}$, ΔFU $\leq$ 1
		WISP-3				
iPhone SE	L	162	0.71	0.009	19%	−7.8%
			(0.57, 0.80)	(0.007, 0.010)	(15%, 24%)	(−11%, +0.21%)
	R_{rs}	81	0.97	0.0014 sr^{-1}	22%	−22%
			(0.91, 0.99)	(0.0001, 0.0063)	(18%, 27%)	(−28%, −20%)
	R_{rs}	81	0.97	0.0004 sr^{-1}	9.7%	+1.9%
	(regr.)		(0.90, 0.99)	(0.0001, 0.0025)	(6.7%, 13%)	(−1.1%, +4.8%)
	Band	81	0.98	0.013	1.9%	+0.25%
	ratios		(0.97, 0.99)	(0.009, 0.019)	(1.2%, 2.7%)	(−0.61%, +0.87%)
	α, FU	27	9.4°	1 FU	26%	59%
			(6.3°, 12°)	(1, 2)	(7.4%, 41%)	(37%, 74%)
Galaxy S8	L	162	0.75	0.009	19%	−3.2%
			(0.66, 0.83)	(0.007, 0.011)	(15%, 24%)	(−7.5%, +2.5%)
	R_{rs}	81	0.94	0.0025 sr^{-1}	31%	−31%
			(0.75, 0.98)	(0.0013, 0.0072)	(24%, 35%)	(−36%, −25%)
	R_{rs}	81	0.93	0.0007 sr^{-1}	13%	+5.9%
	(regr.)		(0.70, 0.97)	(0.0005, 0.0041)	(9.2%, 14%)	(+3.2%, +11%)
	Band	81	0.98	0.010	1.7%	+0.04%
	ratios		(0.96, 0.99)	(0.007, 0.012)	(1.1%, 2.0%)	(−0.59%, +0.68%)
	α, FU	27	16°	2 FU	19%	48%
			(11°, 21°)	(1, 4)	(3.7%, 33%)	(26%, 63%)
		So-Rad				
iPhone SE	L	54	0.87	0.006	13%	−4.7%
			(0.75, 0.94)	(0.004, 0.007)	(8.4%, 16%)	(−12%, −0.01%)
	R_{rs}	27	0.70	0.004 sr^{-1}	13%	+12%
			(0.49, 0.86)	(0.003, 0.005)	(9.9%, 16%)	(+6.5%, +14%)
	Band	27	0.97	0.013	3.8%	+0.82%
	ratios		(0.95, 0.98)	(0.006, 0.042)	(0.98%, 5.6%)	(−1.3%, +1.6%)
	α, FU	9	11°	1 FU	11%	89%
			(6.1°, 13°)	(1, 1)	(0%, 33%)	(34%, 100%)
Galaxy S8	L	54	0.83	0.005	12%	−5.4%
			(0.69, 0.93)	(0.003, 0.007)	(9.2%, 16%)	(−12%, +1.4%)
	R_{rs}	27	0.75	0.003 sr^{-1}	8.5%	+6.4%
			(0.49, 0.87)	(0.002, 0.005)	(5.5%, 13%)	(−0.08%, +9.1%)
	Band	27	0.99	0.004	1.1%	+0.36%
	ratios		(0.98, 0.99)	(0.003, 0.009)	(0.42%, 2.4%)	(−0.35%, +0.45%)
	α, FU	9	16°	1 FU	11%	56%
			(12°, 23°)	(1, 2)	(0%, 33%)	(11%, 78%)

Table 4.1: Summary of the smartphone vs. reference match-up analysis. The values between parentheses indicate the 5–95% CI determined from bootstrapping. N is the number of matching observations; the other metrics are described in Section 4.2.6. $\mathcal{M}(L)$ is in units of W m^{-2} nm^{-1} sr^{-1}. For the WISP-3, R_{rs} was compared 1:1 and with a linear regression (regr.).

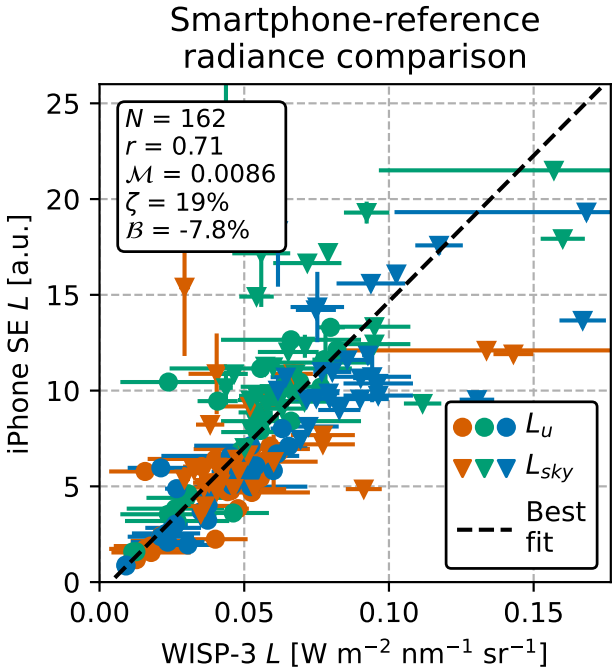


Figure 4.10: Comparison between iPhone SE and spectrally convolved WISP-3 radiance measurements. The RGB channels are shown in their respective colours, with different symbols for L_u and L_{sky}. The statistics in the text box are relative to the regression line. We note that this regression line cannot be used as a general absolute radiometric calibration for the iPhone SE due to the arbitrary choice of exposure settings.

intercomparison (Section 4.3.2). For each smartphone, there were only $N = 27$ WISP-3 match-ups and even fewer So-Rad ones, making the CIs wide and the interpretation difficult. The difference between the WISP-3 and iPhone SE was slightly larger than in the smartphone comparison, at $\mathcal{M}(\alpha) = 9.4°$ (CI 6.3°, 12°) and $\mathcal{M}$(FU) = 1 (CI 1, 2). The Galaxy S8 and WISP-3 differed more, at $\mathcal{M}(\alpha) = 16°$ (CI 11°, 21°) and $\mathcal{M}$(FU) = 2 (CI 1, 4). The cause for this difference is unclear but may simply be an artifact of the small number of match-ups; the Galaxy S8 also differed more in RGB R_{rs} but not in the band ratios. Both smartphones performed similarly in the FU match-ups, with 19–26% of the match-ups agreeing fully and 48–59% to within 1 FU, although these figures had particularly wide CIs.

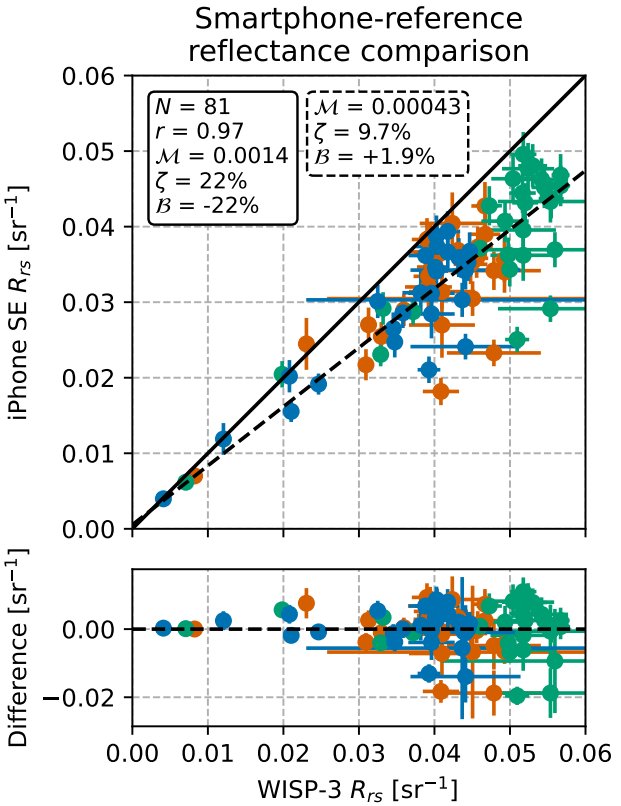


Figure 4.11: Comparison between iPhone SE and spectrally convolved WISP-3 R_{rs} measurements in the RGB bands. The solid line corresponds to a 1:1 relation, the dashed line is the best-fitting linear regression line. The statistics in the solid-outline text box are based on a 1:1 comparison, those in the dashed-outline text box are based on the regression line. The differences in the lower panel are based on the regression line.

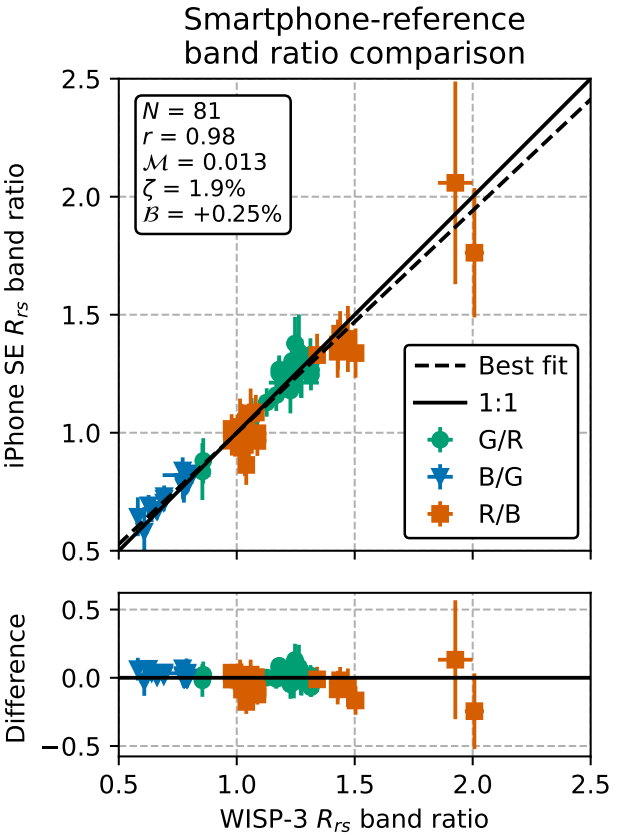


Figure 4.12: Comparison between iPhone SE and spectrally convolved WISP-3 R_{rs} band ratios. The solid line corresponds to a 1:1 relation, the dashed line is the best-fitting linear regression line. The statistics in the text box are based on a 1:1 comparison, as are the differences in the lower panel.

4.3.4 JPEG data

28 sets of JPEG images from the iPhone SE, taken simultaneously with the RAW images, were analysed and compared to the RAW and reference data.

The relationship between JPEG and RAW radiances was highly non-linear (Figure 4.13). Each RGB channel had a different best-fitting power law, with exponents ranging from 0.477±0.005 for B to 0.949±0.013 for R. Due to differences between the RAW and JPEG data processing, the power law exponents are not equivalent to sRGB gamma exponents [281]. Figure 4.13 also shows the significant dispersion of the data around the power law curves. Comparing the RAW and re-scaled JPEG data yielded ζ ranging from 8.9% (CI 7.5%, 11%) for B to 38% (CI 29%, 43%) for R.

The JPEG vs. RAW R_{rs} match-ups agreed better, particularly in the band ratios. The RGB R_{rs} were strongly correlated, with $r = 0.92$ (CI 0.84, 0.97), but the JPEG data showed a large, consistent overestimation of $\mathcal{B} = +52\%$ (CI +39%, +59%). Comparing R_{rs} through a linear regression removed this offset, although a significant dispersion of $\zeta = 15\%$ (CI 12%, 21%) remained. Conversely, the R_{rs} band ratios were more similar with $r = 0.97$ (CI 0.95, 0.98), $\mathcal{M} = 0.033$ (CI 0.023, 0.042), and $\zeta = 4.9\%$ (CI 3.6%, 6.8%).

Finally, the agreement in α and FU was similar to the smartphone vs. smartphone and smartphone vs. reference comparisons. $\mathcal{M}$ was 11° (CI 3.6°, 14°) in α and 1 (CI 0, 2) in FU. 39% (CI 18%, 54%) of match-up pairs had the same FU index, while 61% (CI 39%, 75%) agreed to within 1 FU.

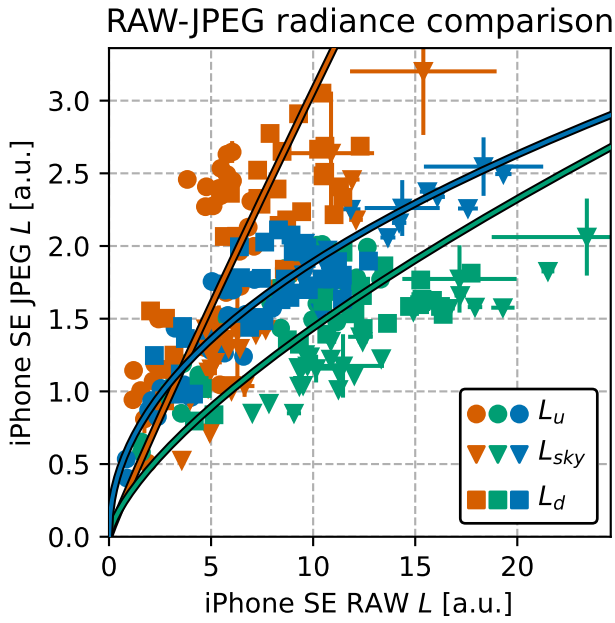


Figure 4.13: Comparison between RAW- and JPEG-based iPhone SE radiance measurements. The axes are in different units due to differences in exposure settings and normalisation. The RGB channels are shown in their respective colours, with different symbols for L_u, L_{sky}, and L_d. The coloured lines show the best-fitting power law for each channel.

The agreement between JPEG and reference data was notably worse than between RAW and reference data. While the JPEG vs. reference radiance match-ups appeared to follow a single linear relationship, rather than the multiple power laws seen in the JPEG vs. RAW comparison, they were only weakly correlated, with $r = 0.39$ (CI 0.22, 0.52) in the JPEG vs. WISP-3 comparison. The dispersion around the regression line was $\zeta = 31\%$ (CI 26%, 41%), 1.6× larger than for the RAW data.

The JPEG data consistently overestimated R_{rs} compared to the references, and were widely dispersed. In the JPEG vs. WISP-3 comparison, $\mathcal{B} = +17\%$ (CI +10%, +19%), although this was reduced to $\mathcal{B} = +1.1\%$ (CI −7.3%, +5.8%) when comparing to a regression line instead of the 1:1 line, as in Section 4.3.3. However, the dispersion remained significant at $\mathcal{M} = 0.0039$ (CI 0.0018, 0.0047) sr^{-1} or $\zeta = 21\%$ (CI 12%, 24%), with $\mathcal{M}$ 9× as much as for the RAW data, and ζ 2.1×.

The JPEG band ratios deviated from the WISP-3 by >2.5× as much as the RAW data, with $\mathcal{M} = 0.032$ (CI 0.023, 0.041) and $\zeta = 5.5\%$ (CI 3.7%, 6.4%). The So-Rad comparison showed a similarly stark difference. However, while this represents a serious reduction in performance, a typical difference of 5.5% is still relatively small.

It was only in α and FU that the JPEG vs. reference and RAW vs. reference agreements were similar. $\mathcal{M}(\alpha)$ in the JPEG vs. WISP-3 comparison was even marginally better at 7.1° (CI 5.0°, 11°); in the JPEG vs. So-Rad comparison it was 13° (CI 3.8°, 16°), almost identical to Table 4.1. $\mathcal{M}$(FU) and the fraction of FU matches were also similar, at $\mathcal{M}(\text{FU}) = 1$ (CI 1, 2), with 26% (CI 7.4%, 41%) full and 59% (CI 37%, 74%) partial FU matches between the JPEG and WISP-3 data. The agreement between JPEG and reference α and FU is discussed further in Section 4.4.3.

The effectiveness of an sRGB linearisation applied to the JPEG data, like in WACODI, was also investigated (Section 4.2.3). In α and FU, the main outputs from WACODI, the linearisation had very little effect. In the JPEG vs. WISP-3 comparison, $\mathcal{M}(\alpha)$ changed from 7.1° (CI 5.0°, 11°) originally to 7.0° (CI 5.4°, 9.4°) with linearisation. In radiance and R_{rs}, the linearisation made all comparison metrics significantly worse.

4.4 Discussion

4.4.1 Uncertainty

The uncertainty of the smartphone data as derived from replicate measurements (Section 4.3.1) is comparable to that of professional spectroradiometers. This was shown by the comparison with WISP-3 replicate measurements, which had a variability similar to, and in some cases larger than, the Galaxy S8. In general, the uncertainty from instrumental effects, excluding environmental factors and photon noise, in professional spectroradiometer data is around 1% [92]. In field data, the typical uncertainty is 1–7% [110]. The Galaxy S8 replicate variability, which was 0.4–1.2% (L_d), 1.1–3.4% (L_{sky}), and 1.8–5.8% (L_u), falls within this range.

The same is true for the smartphone R_{rs} uncertainty, both in RGB (1.9–8.1%) and in band ratios (0.5–1.9%). R_{rs} is typically measured with an uncertainty of 5% at blue and green wavelengths [88] and this is the target for satellites like PACE [86]. The 5% target also applies to narrower bands than the smartphone SRFs and to waters considerably darker than Lake Balaton, which increases the influence of sensor noise. The reduced uncertainty in band ratios is well-known and can be attributed to correlated uncertainties dividing out [146].

Propagated into the mineral suspended sediment (MSS) algorithm described in [34], for R/B ranging from 1.0–1.4, a 2% uncertainty in R/B results in a relative MSS uncertainty of only 1%. In the chl-a algorithm from [323], a 2% uncertainty in B/G induces a relative chl-a uncertainty of 9%. This level of uncertainty is well within the desired limits for many end users [88].

Finally, the uncertainty of the Galaxy S8 α (2.1°–6.8°) and FU index (0.19–0.62 FU) estimates is similar to the uncertainty of satellite and human measurements as well as the existing EyeOnWater app. Through propagation from R_{rs}, Pitarch et al. found uncertainties on SeaWiFS-derived α of 6°–18° [10], although it is difficult to compare these values due to the vastly different water types examined. Furthermore, propagated and replicate-based uncertainty estimates may vary significantly due to differences in sensitivity to various factors (Section 4.3.1). A more representative comparison point is the standard deviation of 3.15° among replicate EyeOnWater observations by Malthus et al. [275], which falls squarely within the range found in this work. The similarity in uncertainty is interesting because EyeOnWater is based on JPEG data, not RAW. However, since we did not take replicate JPEG images, a direct comparison in uncertainty between JPEG and RAW could not be made. The accuracy of JPEG and RAW data, including α and FU index, is compared in Section 4.4.3. The uncertainty of 0.19–0.62 FU is 5.3–1.6× better than human measurements, which have a typical uncertainty of 1 FU with perfect colour vision [280].

Since the use of RAW data eliminates virtually all smartphone-specific sources of uncertainty [281], the primary remaining sources are those that apply to all (spectro)radiometers as well as environmental factors. For a thorough overview of the former, we refer the reader to [106, 110]; for the latter, to [88]. Read-out noise, thermal dark current, and digitisation noise are negligible for well-lit smartphone images [281]. Since L_d was measured on a stable target, its variability of 0.4–1.2% between replicates can be ascribed mostly to sensor noise (Section 4.3.1). Sensor noise scales with the square root of the number of photons, so the induced uncertainty will be larger in darker conditions such as overcast days, highly absorbing waters, and low solar elevation angles. In practice, smartphone observations under dark conditions will require longer exposure times or multiple images to attain similar levels of uncertainty. The impact of sun glint, which is estimated from L_{sky}, on the uncertainty in R_{rs} is also larger for darker waters. The sensitivity of smartphone cameras to temperature variations and polarisation is unknown, although the latter is expected to be negligible unless special fore-optics are used [284]. Because our data were gathered in a single 3-day campaign, long-term sensor drift is unlikely to have had any effect; in general, sensor drift does not affect relative measurements like R_{rs} and α. Environmental factors, such as the patchy clouds that were present during our campaign (Figure 4.1), likely contributed the bulk of the uncertainty in L_{sky} and L_u. These environmental factors also affected the reference measurements and are inherent to above-water radiometry.

4.4.2 Reproducibility

As there are hundreds of different smartphone models, reproducibility between devices is key. This is a major problem with HydroColor, as reported to us directly by users and as reported in the literature. For example, HydroColor measurements of R_{rs} with different smartphones regularly differ by as much as 50% or 0.005 sr^{-1} [121, 276]. This is largely due to the use of JPEG data, which are processed differently on every smartphone model, leading to a wide variety of errors and uncertainties that cannot be reliably corrected [281]. On the other

hand, Goddijn et al. reported smaller differences (4% ± 4%) between JPEG data from two high-quality digital cameras [323], suggesting that some of the problems may be specific to smartphones.

In Section 4.3.2, we showed that with RAW data and camera calibrations, excellent agreement and thus reproducibility between smartphones can be achieved. Near-simultaneous iPhone SE and Galaxy S8 measurements of radiance and R_{rs} were nearly perfectly correlated ($r \geq 0.94$), and their dispersion could be explained by the uncertainties in the individual measurements. The typical difference in R_{rs} was 0.0010 (CI 0.0005, 0.0013) sr^{-1} or 5.5% (CI 3.8%, 8.2%), both major improvements over HydroColor. In fact, the dispersion in radiance between the two smartphones, $\zeta = 6.9\%$ (CI 5.1%, 8.7%), is only slightly larger than that between professional instruments in a similar experiment [92].

On the contrary, the smartphone JPEG processing algorithm was found to be poorly constrained and highly inconsistent between the RGB channels (Section 4.3.4). Moreover, the internal JPEG processing in the smartphone is re-tuned every time a camera session is started [281]. Combined, the differences between channels and between sessions highly limit the reproducibility of JPEG-based measurements of radiance and R_{rs}. As discussed below, white-balancing further reduces the reproducibility of JPEG-based R_{rs} band ratios and hue angles. Finally, the JPEG processing algorithms differ between manufacturers, further reducing the reproducibility of JPEG data between devices [281]. Due to limitations in the SPECTACLE app in 2019, we did not collect Galaxy S8 JPEG data in this study, meaning a direct comparison between the RAW vs. RAW and JPEG vs. JPEG reproducibility could not be performed. Reproducing JPEG data from the RAW data was not possible, due to the aforementioned proprietary smartphone algorithms.

Differences in smartphone SRFs set some minor fundamental limits on the reproducibility between different cameras [331]. However, since most natural waters have broad and smooth spectra, this should only lead to minor differences. In theory, JPEG data do not have this problem because they are always in the sRGB colour space [285], but in practice the various proprietary colour algorithms cause larger differences in JPEG data than in RAW [281]. Furthermore, to account for illumination differences, JPEG data are white-balanced, changing the relative intensity of each channel. The re-normalisation directly reduces the accuracy of band ratio and hue angle measurements and is difficult to correct post-hoc [281, 391]. The white-balance setting may be locked between exposures [121, 323], but this does not guarantee consistency between different devices. Finally, due to differences in field-of-view between cameras, the central slice of 100 × 100 pixels does not always subtend the same solid angle. In future work, it may be advisable to use a constant solid angle rather than a constant pixel slice [121].

4.4.3 Accuracy

In Section 4.3.3, we compared smartphone and reference data to determine the accuracy of the smartphone data, but this comes with important caveats. While each instrument measured L_u and L_{sky}, they did not do so in exactly the same way, having differences in field of view, spectral response, spectral resolution, and time and location. While the smartphones measured L_d on a gray card, the references measured E_d with a cosine collector. Due to these differences, the true 'ground truth' value of each measurand is not known [106, 409]. The reference data can be used to approximate the true values and achieve closure [234], but one must be aware of the uncertainties and systematic errors that may be present. Additionally,

one must exercise caution when comparing different metrics, such as the median symmetric accuracy ζ and the mean percentage deviation, which measure the same quantity but are calculated differently and on different data.

The WISP-3 and So-Rad R_{rs} spectra were similarly shaped, but differently normalised (Section 4.2.2). Both were similar to spectra from previous work in shape, with the So-Rad more similar in magnitude [27, 264]. Normalisation differences and offsets have been seen in previous comparisons between the WISP-3 and other instruments [92, 126], so we felt confident in using a linear regression to re-scale R_{rs} in the smartphone vs. WISP-3 comparisons. In fact, since each smartphone R_{rs} measurement was based on three images from the same camera, rather than from three separate sensors like the WISP-3, and the gray card reference was independently verified, we can be more confident in the normalisation of the smartphone R_{rs} than that of the WISP-3, at least for the particular unit and calibration settings used during our campaign. These results suggest that smartphones and other low-cost cameras could be used to provide closure when there is tension between data from professional instruments (Section 4.4.5).

Considering the above, the level of closure between smartphone and reference data was comparable to intercomparisons between professional radiometers to within a factor of 2–3. The dispersion ζ in radiance was relatively large at 12–19%, 2–3× that reported in a comparison of hyperspectral instruments on a single, stable platform [92], but as discussed previously, our radiance measurements were particularly affected by environmental factors and were taken at slightly different times and positions between instruments. Patchy clouds can increase the dispersion in radiance match-ups by a factor of 10 or more [126]. In R_{rs}, the typical difference was on the order of 10^{-4}–10^{-3} sr^{-1} or 9–13%. Comparing hyperspectral radiometers, Tilstone et al. found mean differences between sensors on the order of 10^{-3} sr^{-1} or 1–8%, with outliers up to 13% [114]. A comparison between WISP-3 and RAMSES sensors under cloudy conditions, similar to ours, found differences in R_{rs} of 20–30% [126].

Most importantly, the smartphone and reference measurements of R_{rs} band ratios agreed to within 2% in three out of four comparisons. The difference was only larger in the iPhone SE vs. So-Rad comparison, at 3.8%. Since band ratios are what most inversion algorithms for inherent optical properties and constituent concentrations are based on, it is the band ratio accuracy that determines the usefulness of smartphones as spectroradiometers. An accuracy and uncertainty of around 2% is well within most user requirements (Section 4.4.1).

The accuracy of the JPEG data was considerably worse (Section 4.3.4). In R_{rs}, the dispersion in the JPEG vs. WISP-3 comparison was 0.0039 (CI 0.0018, 0.0047) sr^{-1} or 21% (CI 12%, 24%), which is in line with previous validation efforts for HydroColor [121, 276] and other JPEG-based methods [391, 394]. At $\mathcal{M} = 0.032$ (CI 0.023, 0.041) and $\zeta = 5.5\%$ (CI 3.7%, 6.4%), the same is true for the R_{rs} band ratios [276]. The RAW data performed better on each of these metrics, most notably by 9× for the RGB R_{rs} and 2.5× for the band ratios. These results do not completely invalidate previous JPEG-based methods nor HydroColor specifically [275], but demonstrate the significant increase in accuracy and decrease in uncertainty obtained by using RAW data.

The results for the hue angle α and FU index were less conclusive. While at first glance the dispersion of approximately 10° or 1 FU appears to be in line with previous studies [275, 285, 395], our measurement protocol (Section 4.2.1) did not follow the EyeOnWater protocol exactly, so the results cannot be compared directly to the aforementioned validation efforts. Additionally, our data only contained 27 smartphone vs. WISP-3 match-ups and even fewer for the So-Rad, with little diversity. Lastly, hue angles derived from narrow-band

multispectral satellite data have been shown to differ systematically by several degrees, up to 20° in extreme cases, compared to hue angles derived from hyperspectral data [10, 410]. This effect may also be present in the smartphone data and a correction term in the hue angle algorithm may be necessary [411]. This work used the original hue angle algorithm, which is based only on the SRFs [402], to enable a comparison between RAW and JPEG data and between the current study and previous works, particularly the WACODI algorithm [285]. We recommend that future work be done to investigate the magnitude of the hue angle bias in consumer camera data. Interestingly, there was little difference in accuracy between the RAW- and JPEG-derived hue angles and FU indices. It is unclear whether this is because the method is inherently robust to JPEG-induced errors [285], although Gao et al. have suggested that it is not [391]. More data, from more diverse waters, will be necessary to compare the accuracy of RAW- and JPEG-based hue angles and FU indices.

A potentially important source of systematic error is the 18% gray card. While the gray card used here did not deviate significantly from $R_{ref} = 18\%$ (Section 4.2.3), this may not be true in general. Since many smartphone radiometry projects are aimed at citizen scientists, who may purchase a wide variety of gray cards and may not always use them correctly, this presents an important possible source for error. Even a small difference in R_{ref} can significantly bias R_{rs}. One possible solution to this problem is to issue or recommend standardised gray cards [391]. Characterising the most popular gray cards is another possibility [400], which may itself be done through citizen science. The use of relative quantities like band ratios negates this problem.

4.4.4 Recommendations

Based on previous work and the results discussed above, several recommendations can be made. Some are specific to smartphones, but most apply in general to above-water radiometry with consumer cameras since the cameras in most smartphones, digital cameras, UAVs, and webcams are extremely similar [281].

RAW data provide professional-grade radiometric performance and should be used whenever possible. Most consumer cameras now support this natively and many smartphone apps provide this capacity. Within the MONOCLE[17] project, a universal smartphone library for RAW acquisition and processing is in development. In the future, apps like HydroColor may simply import this library and use RAW data without further work from the user. The SPECTACLE Python library (Section 4.2.3) provides this functionality on PCs.

Few calibration data are necessary for above-water radiometry. Our processing pipeline contains bias and flatfield corrections, demosaics the data to the $RGBG_2$ channels, and normalises by the SRF spectral bandwidths (Figure 4.3). RAW files from virtually all cameras contain metadata describing the bias correction and demosaicking pattern. The flatfield correction requires additional data, which can be obtained through do-it-yourself methods [281], but may also be neglected at little cost in accuracy because its effect is typically small (0.2% for the iPhone SE and 1.6% for the Galaxy S8) in the central 100×100 pixels. The flat-field correction is more important in approaches that require a wider field-of-view like the multiple gray card approach [391]. The bandwidth normalisation divides out in the calculation of R_{rs} and thus is only necessary to obtain accurate radiances. The SRFs are also required to accurately calculate α and convolve hyperspectral data in validation efforts, but may be approximated by standard profiles [121]. Low-cost smartphone spectrometers and other novel

[17] https://monocle-h2020.eu/

methods will soon enable on-the-fly SRF calibrations [284, 412].

As discussed in [281], it is important to accurately record exposure settings. In the current study, the exposure settings were not recorded, so it is not possible to combine our data with data from other studies, taken with different settings. The most important exposure settings are ISO speed and exposure time, which strongly affect the observed signal, but are not recorded accurately in the image metadata (EXIF). The settings must therefore be recorded by the user or the app. Since ISO speed does not affect the signal-to-noise ratio (SNR), a constant value maybe used. Longer exposure times increase the SNR but run the risk of saturation. Ideally, an automatic exposure time is determined and recorded for each image; if this is not possible, a single value may be used.

Algorithms to retrieve inherent optical properties from smartphone-based R_{rs} measurements are best based on band ratios since they are the most precise, reproducible, and accurate. Algorithms based on absolute R_{rs} in RGB [121, 391] are more susceptible to uncertainty and systematic errors. Because the RGB SRFs are broad and overlapping, some narrow spectral features like pigment absorption peaks cannot be distinguished, and retrieval algorithms require tuning to specific sites [34]. In edge cases where spectral features fall on wavelengths where SRFs vary significantly between devices, the reproducibility of retrieval algorithms between devices may also vary. For example, the iPhone SE and Galaxy S8 B-band SRFs differ greatly between 550–600 nm [281]. Algorithms that use spectrally distinct peaks, for example to retrieve chl-a concentrations, should be unaffected. Distinguishing between chl-a and CDOM, which both absorb in the B and G bands, may require a three-band algorithm that also estimates the backscattering coefficient b_b from the R-band [228]. Alternative colour spaces like relative RGB [34, 386], hue-saturation-intensity [388], and CIE L*a*b* [230] are also worth exploring. Potential algorithms may be identified through spectral convolution of archival R_{rs} spectra [283].

4.4.5 Outlook

The findings presented in this work extend to other methods for smartphone (spectro)radiometry and to most consumer cameras. This study was performed as a precursor to the field validation for the iSPEX 2 smartphone spectropolarimeter [284]. The uncertainty, accuracy, and reproducibility of iSPEX 2 data will be comparable to what was found in this study, although longer exposure times will be necessary to attain similar photon counts. The low uncertainty and high accuracy of the R_{rs} band ratios is particularly promising since iSPEX 2 will measure hyperspectrally across the visible range, enabling many such algorithms. Also applicable to iSPEX 2 are some of the limitations found in this work, primarily the dependence on a gray card and the question of sensitivity in low-light conditions.

There is also potential for low-cost cameras, like webcams and UAV cameras, to augment professional spectroradiometers. Removal of the direct sun glint remains challenging, requiring assumptions about the spectrum and wave statistics [120, 210]. Low-cost camera images, taken simultaneously with the spectra, could be used to determine the wave statistics akin to [413] but for individual exposures. A similar system, which flags spectra if the associated image has saturated pixels, was already demonstrated in [414], and there are further opportunities for image-based anomaly detection. Finally, low-cost cameras can serve as simple validation checks for other sensors, for example to identify normalisation problems.

4.5 Conclusions

In this work, we have assessed the performance of smartphones as multispectral above-water radiometers. We have extended the existing smartphone-based approaches by using RAW data, processed through the SPECTACLE method for calibration of consumer cameras [281]. Using field data gathered under realistic observing conditions on and around Lake Balaton, we have analysed the uncertainty, reproducibility, and accuracy of above-water radiometry data taken with smartphone cameras. Furthermore, by comparing RAW and JPEG data, we have determined to what extent our new method improves upon existing work.

The uncertainty of the smartphone data, determined from replicate observations, was on the percent level and was comparable to professional radiometers. The typical uncertainty on R_{rs} band ratios was 0.5–1.9%, leading to percent-level uncertainties in retrieved inherent optical properties and constituent concentrations. This level of uncertainty falls within the desired limits for many end users.

The reproducibility between smartphones was excellent, representing a significant improvement over existing methods, in some cases nearly tenfold. Any differences in the data between smartphones could be explained by measurement uncertainties.

The accuracy of smartphone data, as determined from match-ups with reference instruments, was comparable to professional instruments. The typical difference between smartphone and reference instruments was 10^{-4}–10^{-3} sr^{-1} or 9–13% in RGB R_{rs}, and 0.004–0.013 or 1.1–3.8% in R_{rs} band ratios. These differences were an improvement of 9× and 2.5×, respectively, over JPEG data.

Based on the findings of this study, we recommend the use of RAW data for above-water radiometry with smartphones by professional and citizen scientists alike. We further recommend that retrieval algorithms be based on R_{rs} band ratios rather than absolute RGB R_{rs}. Potential algorithms may be identified through spectral convolution of archival hyperspectral data. The conclusions and recommendations described above extend to other consumer cameras and to hyperspectral approaches like iSPEX 2. Future work should focus on determining the limitations of consumer cameras, primarily in terms of sensitivity, and exploring opportunities for complementary use of consumer cameras and professional spectroradiometers.

4.A Uncertainty propagation

As discussed above, significant correlations were found between channels in the smartphone data, as well as between images within one data set. To account for this, the inter-channel and inter-image covariances were incorporated into the uncertainty propagation.

4.A.1 Covariance and correlation

First, the mean radiances were combined into a single vector $\mathbf{L}$, containing 12 elements, corresponding to each channel in each image. The radiance vector $\mathbf{L}$ for the 07:47 data set is given in Equation (4.11). The elements of $\mathbf{L}$ are in analogue-digital units (ADU) nm^{-1}. To save space, $\mathbf{L}$ is shown in its row vector form $\mathbf{L}^T$.

$$\begin{aligned}\mathbf{L}^T &= \begin{bmatrix} L_u(R) & L_u(G) & L_u(B) & L_u(G_2) & L_{sky}(R) & L_{sky}(G) & L_{sky}(B) & L_{sky}(G_2) & L_d(R) & L_d(G) & L_d(B) & L_d(G_2) \end{bmatrix} \\ &= \begin{bmatrix} 5.10 & 8.58 & 5.05 & 8.59 & 5.19 & 9.71 & 9.60 & 9.73 & 7.28 & 10.53 & 7.63 & 10.54 \end{bmatrix}\end{aligned} \tag{4.11}$$

The radiance vector $\mathbf{L}$ has a corresponding covariance matrix $\mathbf{\Sigma_L}$. The diagonal elements of $\mathbf{\Sigma_L}$ contain the variances of individual elements of $\mathbf{L}$, namely $\sigma^2_{L_u(R)}, \sigma^2_{L_u(G)}, \ldots$, while the off-diagonal elements contain the covariances, namely $\sigma_{L_u(R)L_u(G)}, \sigma_{L_u(R)L_u(B)}, \ldots$, all in units of ADU^2 nm^{-2}. The covariance matrix corresponding to Equation (4.11) is given in Equation (4.12). The rows and columns are in the same order as the elements of $\mathbf{L}$, with $\sigma^2_{L_u(R)}$ and $\sigma^2_{L_d(G_2)}$ in the top left and bottom right corners, respectively.

$$\mathbf{\Sigma_L} = \begin{bmatrix} 0.043 & 0.003 & 0.002 & 0.003 & 0.006 & 0.007 & 0.007 & 0.007 & 0.003 & 0.002 & 0.002 & 0.004 \\ 0.003 & 0.050 & 0.002 & 0.004 & 0.002 & 0.006 & 0.004 & 0.001 & 0.004 & 0.009 & 0.003 & 0.003 \\ 0.002 & 0.002 & 0.033 & 0.002 & -0.002 & -0.004 & 0.004 & -0.004 & 0.003 & 0.003 & 0.007 & 0.003 \\ 0.003 & 0.004 & 0.002 & 0.047 & 0.004 & 0.007 & 0.005 & 0.009 & 0.003 & 0.003 & 0.003 & 0.010 \\ 0.006 & 0.002 & -0.002 & 0.004 & 0.114 & 0.101 & 0.075 & 0.108 & -0.003 & -0.007 & -0.006 & -0.007 \\ 0.007 & 0.006 & -0.004 & 0.007 & 0.101 & 0.195 & 0.103 & 0.151 & -0.002 & -0.003 & -0.007 & -0.013 \\ 0.007 & 0.004 & 0.004 & 0.005 & 0.075 & 0.103 & 0.142 & 0.107 & -0.004 & -0.007 & 0.001 & -0.005 \\ 0.007 & 0.001 & -0.004 & 0.009 & 0.108 & 0.151 & 0.107 & 0.206 & -0.004 & -0.010 & -0.008 & -0.008 \\ 0.003 & 0.004 & 0.003 & 0.003 & -0.003 & -0.002 & -0.004 & -0.004 & 0.066 & 0.007 & 0.007 & 0.008 \\ 0.002 & 0.009 & 0.003 & 0.003 & -0.007 & -0.003 & -0.007 & -0.010 & 0.007 & 0.068 & 0.007 & 0.013 \\ 0.002 & 0.003 & 0.007 & 0.003 & -0.006 & -0.007 & 0.001 & -0.008 & 0.007 & 0.007 & 0.056 & 0.010 \\ 0.004 & 0.003 & 0.003 & 0.010 & -0.007 & -0.013 & -0.005 & -0.008 & 0.008 & 0.013 & 0.010 & 0.073 \end{bmatrix} \tag{4.12}$$

It is often easier to think in terms of correlation r, which is dimensionless and ranges from -1 to 1. The correlation between two quantities is simply their covariance normalised by their individual uncertainties: $r_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$. $r = 0$ implies no correlation is present, while $r = 1$ and $r = -1$ imply a perfect positive and negative correlation, respectively. When applied to a covariance matrix, this yields a correlation matrix $\mathbf{R}$. The matrix $\mathbf{R_L}$ corresponding to Equation (4.12), is shown in Equation (4.13). The diagonal elements of $\mathbf{R_L}$ are always 1 and the matrix is always symmetric.

$$\mathbf{R_L} = \begin{bmatrix} 1.00 & 0.05 & 0.06 & 0.08 & 0.09 & 0.07 & 0.09 & 0.08 & 0.06 & 0.04 & 0.04 & 0.07 \\ 0.05 & 1.00 & 0.06 & 0.09 & 0.02 & 0.06 & 0.05 & 0.01 & 0.07 & 0.16 & 0.06 & 0.05 \\ 0.06 & 0.06 & 1.00 & 0.05 & -0.03 & -0.05 & 0.06 & -0.05 & 0.06 & 0.06 & 0.16 & 0.07 \\ 0.08 & 0.09 & 0.05 & 1.00 & 0.06 & 0.07 & 0.06 & 0.09 & 0.05 & 0.05 & 0.07 & 0.16 \\ 0.09 & 0.02 & -0.03 & 0.06 & 1.00 & 0.68 & 0.59 & 0.71 & -0.03 & -0.08 & -0.07 & -0.08 \\ 0.07 & 0.06 & -0.05 & 0.07 & 0.68 & 1.00 & 0.62 & 0.75 & -0.02 & -0.03 & -0.07 & -0.11 \\ 0.09 & 0.05 & 0.06 & 0.06 & 0.59 & 0.62 & 1.00 & 0.63 & -0.04 & -0.07 & 0.01 & -0.05 \\ 0.08 & 0.01 & -0.05 & 0.09 & 0.71 & 0.75 & 0.63 & 1.00 & -0.04 & -0.09 & -0.07 & -0.07 \\ 0.06 & 0.07 & 0.06 & 0.05 & -0.03 & -0.02 & -0.04 & -0.04 & 1.00 & 0.11 & 0.11 & 0.12 \\ 0.04 & 0.16 & 0.06 & 0.05 & -0.08 & -0.03 & -0.07 & -0.09 & 0.11 & 1.00 & 0.11 & 0.19 \\ 0.04 & 0.06 & 0.16 & 0.07 & -0.07 & -0.07 & 0.01 & -0.07 & 0.11 & 0.11 & 1.00 & 0.15 \\ 0.07 & 0.05 & 0.07 & 0.16 & -0.08 & -0.11 & -0.05 & -0.07 & 0.12 & 0.19 & 0.15 & 1.00 \end{bmatrix} \tag{4.13}$$

4.A.2 Averaging the G and G_2 channels

The G and G_2 channels in each image were averaged because these bands have the same spectral response function. This was done by multiplying $\mathbf{L}$ with a simple transfer matrix as shown in Equation (4.14); $\mathbf{M}$ is shown in Equation (4.15).

$$\mathbf{L}_{RGB} = \mathbf{M}\mathbf{L} \tag{4.14}$$

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{4.15}$$

Since Equation (4.14) is a simple linear transformation, the propagation of $\mathbf{\Sigma}_{\mathbf{L}}$ to $\mathbf{\Sigma}_{\mathbf{L}_{RGB}}$ is a simple matrix multiplication with $\mathbf{M}$ and its transpose, as shown in Equation (4.16). In the special case that all covariances are 0, this reduces to the well-known sum-of-squares equation.

$$\mathbf{\Sigma}_{\mathbf{L}_{RGB}} = \mathbf{M}\mathbf{\Sigma}_{\mathbf{L}}\mathbf{M}^T \tag{4.16}$$

For illustration, the resulting radiance vector $\mathbf{L}_{RGB}$ and correlation matrix $\mathbf{R}_{\mathbf{L}_{RGB}}$ are shown in Equations (4.17) and (4.18), respectively.

$$\begin{aligned} \mathbf{L}_{RGB}^T &= \begin{bmatrix} L_u(R) & L_u(G) & L_u(B) & L_{sky}(R) & L_{sky}(G) & L_{sky}(B) & L_d(R) & L_d(G) & L_d(B) \end{bmatrix} \\ &= \begin{bmatrix} 5.10 & 8.59 & 5.05 & 5.19 & 9.72 & 9.60 & 7.28 & 10.54 & 7.63 \end{bmatrix} \end{aligned} \tag{4.17}$$

$$\mathbf{R}_{\mathbf{L}_{RGB}} = \begin{bmatrix} 1.00 & 0.09 & 0.06 & 0.09 & 0.08 & 0.09 & 0.06 & 0.07 & 0.04 \\ 0.09 & 1.00 & 0.07 & 0.05 & 0.09 & 0.07 & 0.08 & 0.19 & 0.09 \\ 0.06 & 0.07 & 1.00 & -0.03 & -0.05 & 0.06 & 0.06 & 0.09 & 0.16 \\ 0.09 & 0.05 & -0.03 & 1.00 & 0.74 & 0.59 & -0.03 & -0.10 & -0.07 \\ 0.08 & 0.09 & -0.05 & 0.74 & 1.00 & 0.67 & -0.03 & -0.10 & -0.07 \\ 0.09 & 0.07 & 0.06 & 0.59 & 0.67 & 1.00 & -0.04 & -0.08 & 0.01 \\ 0.06 & 0.08 & 0.06 & -0.03 & -0.03 & -0.04 & 1.00 & 0.15 & 0.11 \\ 0.07 & 0.19 & 0.09 & -0.10 & -0.10 & -0.08 & 0.15 & 1.00 & 0.17 \\ 0.04 & 0.09 & 0.16 & -0.07 & -0.07 & 0.01 & 0.11 & 0.17 & 1.00 \end{bmatrix} \tag{4.18}$$

4.A.3 Remote sensing reflectance

The remote sensing reflectance R_{rs}, in units of sr^{-1}, was calculated from L_u, L_{sky}, and L_d using Equation (4.2). This equation applies to each band individually, meaning that for example $R_{rs}(R)$ only depends on $L_u(R)$, $L_{sky}(R)$, and $L_d(R)$. However, since there were non-zero covariances between bands and between images, these were propagated through to R_{rs}.

$$R_{rs} = \frac{L_u - \rho L_{sky}}{\frac{\pi}{R_{ref}} L_d} \qquad \text{(4.2 rev.)}$$

While the reference reflectance R_{ref} was assumed to be a constant $R_{ref} = 0.18$ in each band, it had its own uncertainty $\sigma_{R_{ref}} = 0.01$, which was also propagated through to R_{rs}. This was done by appending its uncertainty to $\mathbf{\Sigma}_{\mathbf{L}_{RGB}}$, as shown in block matrix form in Equation (4.19).

$$\mathbf{\Sigma}_{\mathbf{L}_{RGB}}^{R_{ref}} = \begin{bmatrix} \mathbf{\Sigma}_{\mathbf{L}_{RGB}} & \mathbf{0} \\ \mathbf{0} & \sigma_{R_{ref}}^2 \end{bmatrix} \tag{4.19}$$

Since Equation (4.2) is not a simple linear transformation, a linear approximation was used in the uncertainty propagation, as shown in Equation (4.20). This is similar to Equation (4.16), but using the Jacobian matrix of Equation (4.2). This approximation is valid for functions that are locally well-approximated by a linear function.

$$\mathbf{\Sigma}_{\mathbf{R_{rs}}} = \mathbf{J}_{\mathbf{R_{rs}}} \mathbf{\Sigma}_{\mathbf{L}_{RGB}}^{R_{ref}} \mathbf{J}_{\mathbf{R_{rs}}}{}^{T} \tag{4.20}$$

The Jacobian matrix $\mathbf{J}_{\mathbf{R_{rs}}}$ contains all first-order derivatives of R_{rs} in each band, as shown in Equation (4.21).

$$\mathbf{J}_{\mathbf{R_{rs}}} = \begin{bmatrix} \frac{\partial R_{rs}(R)}{\partial L_u(R)} & \frac{\partial R_{rs}(R)}{\partial L_u(G)} & \frac{\partial R_{rs}(R)}{\partial L_u(B)} & \frac{\partial R_{rs}(R)}{\partial L_{sky}(R)} & \frac{\partial R_{rs}(R)}{\partial L_{sky}(G)} & \frac{\partial R_{rs}(R)}{\partial L_{sky}(B)} & \frac{\partial R_{rs}(R)}{\partial L_d(R)} & \frac{\partial R_{rs}(R)}{\partial L_d(G)} & \frac{\partial R_{rs}(R)}{\partial L_d(B)} & \frac{\partial R_{rs}(R)}{\partial R_{ref}} \\ \frac{\partial R_{rs}(G)}{\partial L_u(R)} & \frac{\partial R_{rs}(G)}{\partial L_u(G)} & \frac{\partial R_{rs}(G)}{\partial L_u(B)} & \frac{\partial R_{rs}(G)}{\partial L_{sky}(R)} & \frac{\partial R_{rs}(G)}{\partial L_{sky}(G)} & \frac{\partial R_{rs}(G)}{\partial L_{sky}(B)} & \frac{\partial R_{rs}(G)}{\partial L_d(R)} & \frac{\partial R_{rs}(G)}{\partial L_d(G)} & \frac{\partial R_{rs}(G)}{\partial L_d(B)} & \frac{\partial R_{rs}(G)}{\partial R_{ref}} \\ \frac{\partial R_{rs}(B)}{\partial L_u(R)} & \frac{\partial R_{rs}(B)}{\partial L_u(G)} & \frac{\partial R_{rs}(B)}{\partial L_u(B)} & \frac{\partial R_{rs}(B)}{\partial L_{sky}(R)} & \frac{\partial R_{rs}(B)}{\partial L_{sky}(G)} & \frac{\partial R_{rs}(B)}{\partial L_{sky}(B)} & \frac{\partial R_{rs}(B)}{\partial L_d(R)} & \frac{\partial R_{rs}(B)}{\partial L_d(G)} & \frac{\partial R_{rs}(B)}{\partial L_d(B)} & \frac{\partial R_{rs}(B)}{\partial R_{ref}} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{R_{ref}}{\pi L_d(R)} & 0 & 0 & \frac{-\rho R_{ref}}{\pi L_d(R)} & 0 & 0 & \frac{-R_{rs}(R)}{L_d(R)} & 0 & 0 & \frac{R_{rs}(R)}{R_{ref}} \\ 0 & \frac{R_{ref}}{\pi L_d(G)} & 0 & 0 & \frac{-\rho R_{ref}}{\pi L_d(G)} & 0 & 0 & \frac{-R_{rs}(G)}{L_d(G)} & 0 & \frac{R_{rs}(G)}{R_{ref}} \\ 0 & 0 & \frac{R_{ref}}{\pi L_d(B)} & 0 & 0 & \frac{-\rho R_{ref}}{\pi L_d(B)} & 0 & 0 & \frac{-R_{rs}(B)}{L_d(B)} & \frac{R_{rs}(B)}{R_{ref}} \end{bmatrix} \tag{4.21}$$

For illustration, the resulting $\mathbf{R_{rs}}$, its covariance matrix $\mathbf{\Sigma}_{\mathbf{R_{rs}}}$, and its correlation matrix $\mathbf{R}_{\mathbf{R_{rs}}}$ are shown in Equations (4.22) and (4.23), respectively. Evidently, there are strong correlations between the RGB channels in $\mathbf{R_{rs}}$. The uncertainty in R_{ref} is particularly important; halving its value to $\sigma_{R_{ref}} = 0.005$ approximately halves the off-diagonal elements of $\mathbf{R}_{\mathbf{R_{rs}}}$.

$$\mathbf{R_{rs}}^{T} = \begin{bmatrix} R_{rs}(R) & R_{rs}(G) & R_{rs}(B) \end{bmatrix} = \begin{bmatrix} 0.039 & 0.045 & 0.036 \end{bmatrix} \tag{4.22}$$

$$\mathbf{\Sigma}_{\mathbf{R_{rs}}} = \begin{bmatrix} 9.0 & 5.5 & 4.4 \\ 5.5 & 7.5 & 5.1 \\ 4.4 & 5.1 & 6.6 \end{bmatrix} \times 10^{-6} \qquad \mathbf{R}_{\mathbf{R_{rs}}} = \begin{bmatrix} 1.00 & 0.67 & 0.57 \\ 0.67 & 1.00 & 0.72 \\ 0.57 & 0.72 & 1.00 \end{bmatrix} \tag{4.23}$$

4.A.4 Band ratios

The calculation of the G/R, B/G, and R/B R_{rs} band ratios was straightforward and the Jacobian matrix approximation was used again. This is shown in Equation (4.24). We assumed

this approximation to be valid here due to the small range of our data and relatively small uncertainties in $\mathbf{R_{rs}}$.

$$\mathbf{J}_{\mathbf{R_{rs,br}}} = \begin{bmatrix} \frac{\partial(G/R)}{\partial R} & \frac{\partial(G/R)}{\partial G} & \frac{\partial(G/R)}{\partial B} \\ \frac{\partial(B/G)}{\partial R} & \frac{\partial(B/G)}{\partial G} & \frac{\partial(B/G)}{\partial B} \\ \frac{\partial(R/B)}{\partial R} & \frac{\partial(R/B)}{\partial G} & \frac{\partial(R/B)}{\partial B} \end{bmatrix} = \begin{bmatrix} -G/R^2 & 1/R & 0 \\ 0 & -B/G^2 & 1/G \\ 1/B & 0 & -R/B^2 \end{bmatrix} \tag{4.24}$$

The resulting band ratio vector $\mathbf{R_{rs,br}}$ for the 07:47 UTC data and its covariance and correlation matrices $\mathbf{\Sigma}_{\mathbf{R_{rs,br}}}$ and $\mathbf{R}_{\mathbf{R_{rs,br}}}$ are shown in Equations (4.25) and (4.26).

$$\mathbf{R_{rs,br}}^T = \begin{bmatrix} G/R & B/G & R/B \end{bmatrix} = \begin{bmatrix} 1.2 & 0.79 & 1.1 \end{bmatrix} \tag{4.25}$$

$$\mathbf{\Sigma}_{\mathbf{R_{rs,br}}} = \begin{bmatrix} 4.5 & -0.54 & -3.5 \\ -0.54 & 1.6 & -1.7 \\ -3.5 & -1.7 & 5.6 \end{bmatrix} \times 10^{-3} \qquad \mathbf{R}_{\mathbf{R_{rs,br}}} = \begin{bmatrix} 1.00 & -0.20 & -0.69 \\ -0.20 & 1.00 & -0.57 \\ -0.69 & -0.57 & 1.00 \end{bmatrix} \tag{4.26}$$

4.A.5 Chromaticity and hue angle

Since the colour space transformation from RGB to XYZ was a simple linear transformation, so was the uncertainty propagation, analogous to Equation (4.16). As discussed in the main paper, the uncertainties on the elements of the transformation matrices were ignored here. The results for the 07:47 data are shown in Equations (4.27) and (4.28). The resulting correlations were very strong, particularly for $R_{rs}(X)$ and $R_{rs}(Y)$, due to the spectral overlap between the XYZ colour-matching functions amplifying the existing correlations in $\mathbf{R_{rs}}(RGB)$.

$$\mathbf{R_{rs,XYZ}}^T = \begin{bmatrix} R_{rs}(X) & R_{rs}(Y) & R_{rs}(Z) \end{bmatrix} = \begin{bmatrix} 0.040 & 0.041 & 0.037 \end{bmatrix} \tag{4.27}$$

$$\mathbf{\Sigma}_{\mathbf{R_{rs,XYZ}}} = \begin{bmatrix} 6.5 & 6.3 & 5.2 \\ 6.3 & 6.2 & 5.3 \\ 5.2 & 5.3 & 6.2 \end{bmatrix} \times 10^{-6} \qquad \mathbf{R}_{\mathbf{R_{rs,XYZ}}} = \begin{bmatrix} 1.00 & 0.98 & 0.81 \\ 0.98 & 1.00 & 0.85 \\ 0.81 & 0.85 & 1.00 \end{bmatrix} \tag{4.28}$$

The resulting uncertainties were further propagated analytically into (x, y) chromaticity, again using the Jacobian matrix approximation as shown in Equations (4.29) and (4.30).

$$\mathbf{R_{rs,xy}} = \begin{bmatrix} x & y \end{bmatrix} = \begin{bmatrix} \frac{X}{X+Y+Z} & \frac{Y}{X+Y+Z} \end{bmatrix} \tag{4.29}$$

$$\mathbf{J}_{\mathbf{R_{rs,xy}}} = \begin{bmatrix} \frac{Y+Z}{(X+Y+Z)^2} & \frac{-X}{(X+Y+Z)^2} & \frac{-X}{(X+Y+Z)^2} \\ \frac{-Y}{(X+Y+Z)^2} & \frac{X+Z}{(X+Y+Z)^2} & \frac{-Y}{(X+Y+Z)^2} \end{bmatrix} \tag{4.30}$$

The results for the 07:47 data are shown in Equation (4.31) and (4.32)

$$\mathbf{R_{rs,xy}}^T = \begin{bmatrix} x & y \end{bmatrix} = \begin{bmatrix} 0.34 & 0.35 \end{bmatrix} \tag{4.31}$$

$$\mathbf{\Sigma}_{\mathbf{R_{rs,xy}}} = \begin{bmatrix} 2.7 & 1.2 \\ 1.2 & 1.3 \end{bmatrix} \times 10^{-5} \qquad \mathbf{R}_{\mathbf{R_{rs,xy}}} = \begin{bmatrix} 1.00 & 0.65 \\ 0.65 & 1.00 \end{bmatrix} \tag{4.32}$$

The uncertainty propagation from (x, y) to hue angle α is problematic. α itself is calculated using Equation (4.6), giving $\alpha = 71°$ for the 07:47 data. The associated Jacobian matrix is given in Equation (4.33) and can be applied, giving an uncertainty of $\sigma_\alpha = 4°$ in the example.

$$\alpha = \arctan2\left(y - 1/3, x - 1/3\right) \mod 2\pi \qquad \text{(4.6 rev.)}$$

$$\mathbf{J}_\alpha = \begin{bmatrix} \frac{-(y-1/3)}{(x-1/3)^2+(y-1/3)^2} & \frac{x-1/3}{(x-1/3)^2+(y-1/3)^2} \end{bmatrix} \tag{4.33}$$

However, for $(x, y) \to (1/3, 1/3)$, the assumption underpinning this method, namely that the transformation can be locally approximated by a linear one, breaks down. This can result in extremely large uncertainty estimates, such as $\alpha = (89 \pm 59)°$ for the 2019-07-03 12:32 UTC Galaxy S8 data, where $(x, y) = (0.3334, 0.3409)$. For this reason, the uncertainty in α is better estimated from replicate observations rather than this analytical propagation.

4.B Smartphone RGB to XYZ transformation

The RGB to XYZ transformation matrices for the iPhone SE and Galaxy S8 were calculated by determining the locations of the RGB primary vectors in XYZ space, following https://www.ryanjuckett.com/rgb-color-space-conversion/. This process was implemented in Python and integrated into the SPECTACLE module, available from https://github.com/monocle-h2020/camera_calibration.

The transformation matrix $\mathbf{M}_{RGB\to XYZ}$ had the RGB primary vectors, $\mathbf{r}_{XYZ}, \mathbf{g}_{XYZ}, \mathbf{b}_{XYZ}$, as its columns, as shown in block matrix form in Equation (4.34).

$$\mathbf{M}_{RGB\to XYZ} = \begin{bmatrix} \mathbf{r}_{XYZ} & \mathbf{g}_{XYZ} & \mathbf{b}_{XYZ} \end{bmatrix} \tag{4.34}$$

First, the RGB SRFs were convolved with the CIE XYZ colour matching functions (CMFs) to give $\mathbf{r}'_{XYZ}, \mathbf{g}'_{XYZ}, \mathbf{b}'_{XYZ}$. This is shown in Equation (4.35), where $S_{R,G,B}$ is the SRF for the RGB bands, and $\bar{x}, \bar{y}, \bar{z}$ are the CMFs. λ dependencies are dropped for clarity. The integrals were evaluated numerically from 390–700 nm.

$$\mathbf{r}'_{XYZ} = \begin{bmatrix} \int S_R \bar{x} \mathrm{d}\lambda \\ \int S_R \bar{y} \mathrm{d}\lambda \\ \int S_R \bar{z} \mathrm{d}\lambda \end{bmatrix} \qquad \mathbf{g}'_{XYZ} = \begin{bmatrix} \int S_G \bar{x} \mathrm{d}\lambda \\ \int S_G \bar{y} \mathrm{d}\lambda \\ \int S_G \bar{z} \mathrm{d}\lambda \end{bmatrix} \qquad \mathbf{b}'_{XYZ} = \begin{bmatrix} \int S_B \bar{x} \mathrm{d}\lambda \\ \int S_B \bar{y} \mathrm{d}\lambda \\ \int S_B \bar{z} \mathrm{d}\lambda \end{bmatrix} \tag{4.35}$$

The normalisation of $\mathbf{r}'_{XYZ}, \mathbf{g}'_{XYZ}, \mathbf{b}'_{XYZ}$ was arbitrary, and they needed to be re-normalised to the desired E-type illuminant white point $\mathbf{w}_{XYZ} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$. First, the corresponding vectors in xyz chromaticity were calculated as shown in Equation (4.36).

$$\mathbf{r}_{xyz} = \frac{\mathbf{r}'_{XYZ}}{\sum_{XYZ} \mathbf{r}'_{XYZ}} \qquad \mathbf{g}_{xyz} = \frac{\mathbf{g}'_{XYZ}}{\sum_{XYZ} \mathbf{g}'_{XYZ}} \qquad \mathbf{b}_{xyz} = \frac{\mathbf{b}'_{XYZ}}{\sum_{XYZ} \mathbf{b}'_{XYZ}} \tag{4.36}$$

Similarly, the following held for the column vectors of $\mathbf{M}_{RGB\to XYZ}$:

$$\mathbf{r}_{XYZ} = \mathbf{r}_{xyz} \sum_{XYZ} \mathbf{r}_{XYZ} \qquad \mathbf{g}_{XYZ} = \mathbf{g}_{xyz} \sum_{XYZ} \mathbf{g}_{XYZ} \qquad \mathbf{b}_{XYZ} = \mathbf{b}_{xyz} \sum_{XYZ} \mathbf{b}_{XYZ} \tag{4.37}$$

Combining Equations (4.34) and (4.37) led to the following:

$$\mathbf{M}_{RGB\to XYZ} = \begin{bmatrix} \mathbf{r}_{xyz} \sum_{XYZ} \mathbf{r}_{XYZ} & \mathbf{g}_{xyz} \sum_{XYZ} \mathbf{g}_{XYZ} & \mathbf{b}_{xyz} \sum_{XYZ} \mathbf{b}_{XYZ} \end{bmatrix} \tag{4.38}$$

$$= \begin{bmatrix} r_x & g_x & b_x \\ r_y & g_y & b_y \\ r_z & g_z & b_z \end{bmatrix} \begin{bmatrix} \sum_{XYZ} \mathbf{r}_{XYZ} & 0 & 0 \\ 0 & \sum_{XYZ} \mathbf{g}_{XYZ} & 0 \\ 0 & 0 & \sum_{XYZ} \mathbf{b}_{XYZ} \end{bmatrix} \tag{4.39}$$

The unknowns in the diagonal matrix on the right were determined by defining the white point $\mathbf{w}$ to be an eigenvector of $\mathbf{M}_{RGB\to XYZ}$ with eigenvalue 1, as shown in Equation (4.40).

$$\mathbf{M}_{RGB\to XYZ}\mathbf{w} = \mathbf{w} \tag{4.40}$$

$$\begin{bmatrix} r_x & g_x & b_x \\ r_y & g_y & b_y \\ r_z & g_z & b_z \end{bmatrix} \begin{bmatrix} \sum_{XYZ} \mathbf{r}_{XYZ} & 0 & 0 \\ 0 & \sum_{XYZ} \mathbf{g}_{XYZ} & 0 \\ 0 & 0 & \sum_{XYZ} \mathbf{b}_{XYZ} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{4.41}$$

$$\begin{bmatrix} r_x & g_x & b_x \\ r_y & g_y & b_y \\ r_z & g_z & b_z \end{bmatrix} \begin{bmatrix} \sum_{XYZ} \mathbf{r}_{XYZ} \\ \sum_{XYZ} \mathbf{g}_{XYZ} \\ \sum_{XYZ} \mathbf{b}_{XYZ} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{4.42}$$

$$\begin{bmatrix} \sum_{XYZ} \mathbf{r}_{XYZ} \\ \sum_{XYZ} \mathbf{g}_{XYZ} \\ \sum_{XYZ} \mathbf{b}_{XYZ} \end{bmatrix} = \begin{bmatrix} r_x & g_x & b_x \\ r_y & g_y & b_y \\ r_z & g_z & b_z \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{4.43}$$

Since all elements on the right-hand side of Equation (4.43) were known, the elements on the left-hand side could be calculated. Finally, plugging these back into Equation (4.39) resulted in the RGB to XYZ transformation matrix. The resulting matrices for the iPhone SE and Galaxy S8 are provided in the main paper.

Funding

European Commission Horizon 2020 programme (grant nr. 776480, MONOCLE).

Acknowledgements

The authors would like to thank Tom Jordan, Victor Martínez-Vicente, Aser Mata, Caitlin Riddick, Norbert Schmidt, and Anna Windle for their help in the data acquisition and processing, and Thomas Leeuw, Sanjana Panchagnula, and Hans van der Woerd for valuable discussions relating to this work.

The datasets generated for this study are available from Zenodo (https://dx.doi.org/10.5281/zenodo.4549621).

5 | Biases from incorrect reflectance convolution

Adapted from

Olivier Burggraaff

Optics Express, 28(9), 13801–13816 (2020)

doi:10.1364/OE.391470

Reflectance, a crucial earth observation variable, is converted from hyperspectral to multispectral through convolution. This is done to combine time series, validate instruments, and apply retrieval algorithms. However, convolution is often done incorrectly, with reflectance itself convolved rather than the underlying (ir)radiances. Here, the resulting error is quantified for simulated and real multispectral instruments, using 18 radiometric data sets (N = 1799 spectra). Biases up to 5% are found, the exact value depending on the spectrum and band response. This significantly affects extended time series and instrument validation, and is similar in magnitude to errors seen in previous validation studies. Post-hoc correction is impossible, but correctly convolving (ir)radiances prevents this error entirely. This requires publication of original data alongside reflectance.

5.1 Introduction

Reflectance, the spectral fraction of light reflected by a surface, is an essential earth observation (EO) variable. It forms the basis for data products such as chlorophyll and suspended matter in water [62, 227, 232], and canopy cover and biomass on land [415, 416]. As such, it is a routine data product for EO satellites, including NASA's Landsat and ESA's Sentinel programmes, and in situ radiometers.

Spectral data are divided into two categories, namely multispectral and hyperspectral. Multispectral instruments observe in several broad, discrete wavelength bands. Examples include the Moderate Resolution Imaging Spectrometer (MODIS) and the Visible Infrared Imaging Radiometer Suite (VIIRS), but also in situ instruments including unmanned aerial vehicles (UAVs) and even smartphones [281]. Conversely, hyperspectral instruments provide continuous wavelength coverage with a fine spectral resolution. Examples include the TriOS RAMSES, Sea-Bird HyperOCR, and ASD FieldSpec field-going spectroradiometers, as well as the Ocean Color Instrument (OCI) due to fly on the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission. Hyperspectral data have a finer spectral sampling and, typically, resolution and thus contain more information than multispectral ones, but depending on the instrument design, often collect less light in each band, giving a worse signal-to-noise ratio.

Since the current EO landscape is a mixture of both types, it is often desirable to convert data between the two, typically from hyper- to multispectral. Three common use cases for this process exist, namely combining time series, instrument validation, and retrieval algorithms.

The first use case is merging and extending time series using different sensors. Long-term, high temporal resolution time series are necessary to study fundamental biogeochemical processes and long-term effects [195, 220] such as climate change [417]. Current efforts focus on merging multispectral time series, on the radiance or reflectance level [91, 418, 419], achieving relative errors on reflectance <5% [91, 419], or on the end product level [220, 420]. Future efforts will focus on extending multispectral time series with new hyperspectral sensors, for example extending MODIS/VIIRS aerosol optical depth (AOD) series with OCI (PACE) data [421]. This is done by converting hyperspectral data to the multispectral sensor's bands, to simulate what the latter would have measured. However, calibration differences and sensor characterisation imperfections can introduce significant biases, for example up to 0.10 AOD for OCI-MODIS/VIIRS [421].

The second use case is the validation of multispectral (often satellite) data using in situ hyperspectral sensors. This is done by comparing simultaneous match-up measurements from both instruments [152]. Validation is done on all products, including normalised radiance [422], reflectance [415, 418, 423, 424], and derived products such as chlorophyll [220, 223] and inherent optical properties (IOPs) [424]. Similar validation is done for in situ multispectral sensors, such as UAVs [425] and smartphones [121, 276]. Vicarious calibration similarly involves comparing match-up data, but aimed at determining satellite gain factors [113]. Since vicarious calibration is performed on (normalised) radiance rather than reflectance, it is outside the scope of this work, though a brief discussion is given in Section 5.2.4.

The third use case is the application of multispectral retrieval algorithms to hyperspectral data. Such algorithms are commonly based on the ratio between spectral bands and are thus called band-ratio algorithms. For example, band-ratio algorithms relating chlorophyll to Sentinel-2A (S2A) Multispectral Instrument (MSI) bands have been developed for Vietnamese [221] and Estonian [62] lakes, the latter with a mean standard error in chlorophyll-a of 5%. While derived on multispectral data, such algorithms are also applied to hyperspectral

data, both to derive products and for validation, requiring a spectral conversion. Differences between these converted data and in situ data have been found [426], which may be due in part to incorrect treatment of radiometry. It should be noted that instead of a spectral conversion, often only the central band wavelength reflectance is used [223, 227].

Converting hyperspectral data to multispectral bands is commonly, though not exclusively (see Section 5.2.2), termed spectral convolution. An in-depth description is provided in Section 5.2, but in short, the hyperspectral data are multiplied by the spectral response function (SRF) of the multispectral band and the product is integrated. This is done for quantities including radiance [427, 428], optical thickness [428, 429], IOPs [424, 430], vegetation indices [416], and reflectance [415, 430].

However, reflectance is often convolved incorrectly. As shown in Section 5.2, hyperspectral reflectance cannot simply be convolved to simulate what a multispectral instrument would observe. Instead, the numerator and denominator, (ir)radiances, should be convolved separately and then divided. This error occurs frequently in the literature, for example in [62, 121, 221, 227, 276, 415, 416, 418, 420, 423–425, 431, 432], with few works convolving radiances before division [223, 426, 433].

This work quantifies the error induced by incorrect spectral convolution of reflectance in each of the three use cases, for a variety of synthetic and real instruments using 18 archival data sets totaling $N = 1799$ spectra. To narrow the scope, this work focuses on remote sensing of ocean colour. However, the principles and methods apply broadly to any fractional quantity, including other reflectances (soil, vegetation), attenuation coefficients, and degree of polarisation, as well as spatial convolution [434]. While the existence of this error has been pointed out previously [430, 433, 435] and quantified at $\lesssim 1\%$ for a single data set and sensor [430], a large-scale quantitative assessment has not yet been published.

This work fits into a wider field of EO error analysis. Recent efforts include investigations into the out-of-band response of EO sensors [115], the impact of differing spectral [433] and spatial [434] resolutions on satellite match-up analyses, and the impact of hyperspectral SRFs having a non-zero bandwidth [436]. On the experimental side, significant efforts have gone into glint removal in above-water radiometry [120, 206, 209] and rigorous characterisation of instrumental [281, 425] and methodological [152, 435] uncertainties. A broad, in-depth overview of uncertainties in ocean colour data is provided in the recently published International Ocean Color Coordination Group (IOCCG) report number 18 [88].

Section 5.2 describes the theoretical background of reflectance and spectral convolution. Section 5.3 describes the data used in this work and the method for quantifying the convolution error. Results are presented in Section 5.4. Finally, Section 5.5 contains a discussion of the results and conclusions.

5.2 Theoretical background

5.2.1 Reflectance

Reflectance R is the ratio of upwelling over downwelling (ir)radiance. Radiance $L(\lambda, \theta, \phi)$ is the radiant energy per wavelength λ propagating in a direction (θ, ϕ), in W m^{-2} nm^{-1} sr^{-1}, while irradiance $E(\lambda)$ is L integrated over a solid angle, in W m^{-2} nm^{-1}. The units of R depend on which ratio is taken. Since this work deals only with wavelength dependence, (θ, ϕ) terms are dropped henceforth for clarity.

Different reflectances can be defined by dividing different (ir)radiances. Examples include the bi-directional radiance reflectance [415], the non-directional irradiance reflectance [432], and the uni-directional remote sensing reflectance R_{rs} used in ocean colour [120,209]. As defined in Equation (5.1), R_{rs} is the ratio of water-leaving radiance L_w [120] over downwelling irradiance E_d [97], in units of sr^{-1}. This work focuses on R_{rs}, but the same mathematics apply to any reflectance.

$$R_{rs}(\lambda) = \frac{L_w(\lambda)}{E_d(\lambda)} \tag{5.1}$$

5.2.2 Spectral convolution

Multispectral data are simulated from hyperspectral data through spectral convolution. As shown in Equation (5.2), this involves multiplying the hyperspectral data $L(\lambda)$ by the multispectral band SRF $S_B(\lambda)$, integrating the result over all wavelengths in the band ($\int_{\lambda\in B} \mathrm{d}\lambda$), and normalising by the effective bandwidth. In this work, convolved quantities are denoted by a bar, such as $\bar{L}(B)$ in Equation (5.2). In practice, spectral convolution is often a sum over discrete L and S_B data. The convolution process is shown graphically in Figure 5.1.

$$\bar{L}(B) = \frac{\int_{\lambda\in B} L(\lambda) S_B(\lambda) \mathrm{d}\lambda}{\int_{\lambda\in B} S_B(\lambda) \mathrm{d}\lambda} \tag{5.2}$$

Convolving hyperspectral data is really an approximation, due to the finite spectral resolution of hyperspectral sensors. As derived in Appendix 5.A, this method is valid if the full width at half maximum (FWHM) of the multispectral band is at least double that of the hyperspectral sensor.

Nomenclature

Various names for this process are used in the literature, including convolution or convolving [97, 152, 276, 415, 416, 418–420, 423, 425, 427, 432, 433, 435, 436], SRF-weighting [429], simulation [426], and band-averaging [121, 281, 424, 427, 428, 430]. Since it is the most common term, 'spectral convolution' is used in this work. However, it should be noted that this term may instead refer to smoothing the spectrum with a kernel [99]. Finally, since neither process involves transforming the SRF, both are actually cross-correlations rather than convolutions.

5.2.3 Reflectance convolution

Just as the hyperspectral remote sensing reflectance $R_{rs}(\lambda)$ is $L_w(\lambda)$ over $E_d(\lambda)$, the convolved $\bar{R}_{rs}(B)$ is $\bar{L}_w(B)$ over $\bar{E}_d(B)$. Both are calculated as in Equation (5.2) and then divided, as shown in Equation (5.3). Convolving (ir)radiances to calculate a band-average reflectance will be referred to in this work as working in *radiance space* or *L-space*, and the result as $\bar{R}^L_{rs}(B)$. Mathematically, this is the correct method for convolving R_{rs} to simulate multispectral data.

$$\bar{R}^L_{rs}(B) = \frac{\bar{L}_w(B)}{\bar{E}_d(B)} = \frac{\int_{\lambda\in B} L_w(\lambda) S_B(\lambda) \mathrm{d}\lambda}{\int_{\lambda\in B} E_d(\lambda) S_B(\lambda) \mathrm{d}\lambda} \tag{5.3}$$

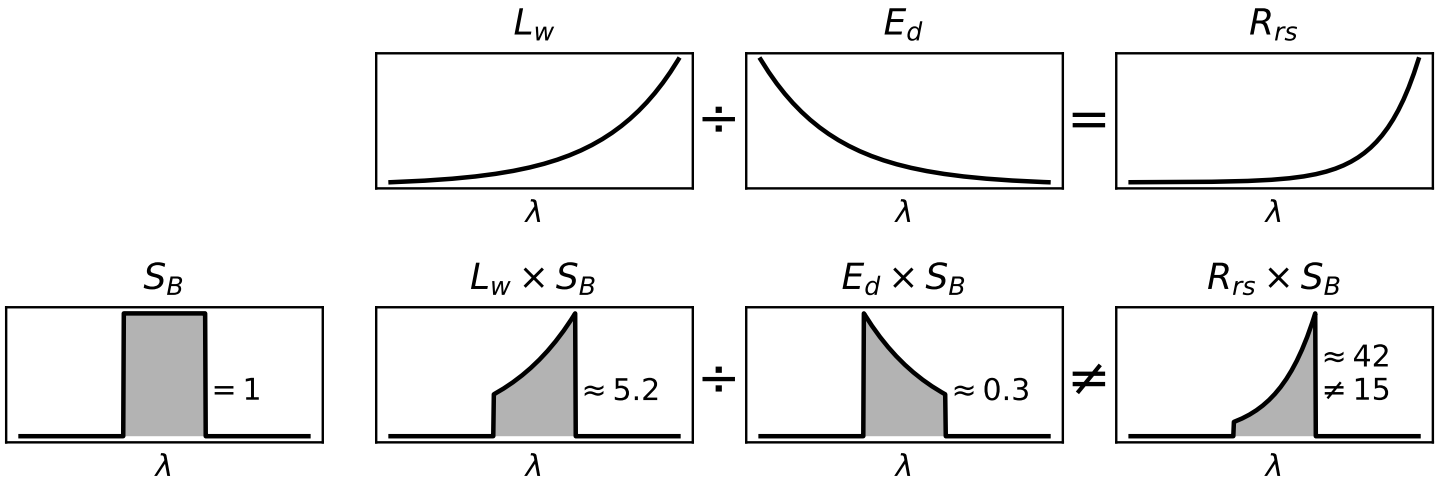


Figure 5.1: Demonstration of spectral convolution with a boxcar SRF S_B (bottom left). Each column shows the convolution of a spectrum (L_w, E_d, R_{rs}) with S_B, the shaded area showing the integral as calculated with Equation (5.2). The $\times$ symbol indicates a spectral convolution. The band-average value of R_{rs} calculated in R-space is 42, significantly different from the correct (L-space) value of $5.2/0.3 \approx 15$. Axis labels and units are dropped for clarity.

Instead, one might simply convolve R_{rs} itself. This will be referred to as working in *reflectance space* or *R-space* and the result as $\bar{R}^R_{rs}(B)$. The expression for $\bar{R}^R_{rs}(B)$ is given in Equation (5.4).

$$\bar{R}^R_{rs}(B) = \frac{\int_{\lambda\in B} R_{rs}(\lambda) S_B(\lambda)\mathrm{d}\lambda}{\int_{\lambda\in B} S_B(\lambda)\mathrm{d}\lambda} = \frac{\int_{\lambda\in B} \frac{L_w(\lambda)}{E_d(\lambda)} S_B(\lambda)\mathrm{d}\lambda}{\int_{\lambda\in B} S_B(\lambda)\mathrm{d}\lambda} \tag{5.4}$$

Working in R-space is incorrect, as shown in Figure 5.1 and the following example. First, let the SRF $S_B(\lambda)$ be a boxcar response of 1 for $0 \leq \lambda \leq 1$ and 0 elsewhere. Then all integrals need only be evaluated for those wavelengths and the SRF bandwidth is 1. Second, let $L_w(\lambda) = e^{e\lambda}$ and $E_d(\lambda) = e^{-e\lambda}$. Such spectra are not physical but demonstrate the mathematical principles well. As shown in Equations (5.5) and (5.6), $\bar{R}^L_{rs}(B) \approx 15$ and $\bar{R}^R_{rs}(B) \approx 42$ differ significantly.

$$\bar{R}^L_{rs}(B) = \frac{\int_0^1 e^{e\lambda}\mathrm{d}\lambda}{\int_0^1 e^{-e\lambda}\mathrm{d}\lambda} \approx \frac{5.2}{0.3} \approx 15 \tag{5.5}$$

$$\bar{R}^R_{rs}(B) = \frac{\int_0^1 \frac{e^{e\lambda}}{e^{-e\lambda}}\mathrm{d}\lambda}{1} = \int_0^1 e^{2e\lambda}\mathrm{d}\lambda \approx 42 \tag{5.6}$$

5.2.4 General rule

Convolution is a useful tool, but the order of operations is not always intuitive. A general rule of thumb can be used, which applies to any kind of convolution (spectral or spatial) when converting high- to low-resolution (spectral or spatial) data. For other purposes, such as smoothing, reflectance itself can be transformed.

As a rule of thumb, only quantities the lower-resolution sensor would observe can be convolved. This includes the at-sensor (ir)radiance (in physical units [435]) but not reflectance and derived products. Propagation of in situ radiances, through surfaces when measured underwater [120] or through the atmosphere for vicarious calibration [113], must occur prior to convolution to accurately simulate the radiance at a multispectral sensor. Simplifications may be necessary [115, 428, 437] but should be mathematically justified. Finally, hyperspectral upwelling radiance L_u, measured in- or above-water, should be converted to L_w [120] before convolution when comparing it to multispectral L_w.

5.3 Methods

Archival data sets containing (ir)radiance and reflectance data were used to test the principles described in Section 5.2 and quantify the errors resulting from working in R-space rather than L-space. All analysis was done using custom Python scripts, available from GitHub[18].

5.3.1 Radiometric data

18 archival radiometric data sets were used [37, 438–451], totaling $N = 1799$ spectra. Data were sourced from the SeaWiFS Bio-optical Archive and Storage System (SeaBASS) [241]

[18] https://github.com/burggraaff/reflectance_convolution

and PANGAEA[19]. Only data sets including either the original radiometric data or R_{rs} and L_w or E_d were used. In most cases, the given R_{rs} and E_d were used and $L_w = R_{rs}E_d$ was reconstructed. This reduced the amount of post-processing, such as glint removal, that was necessary. All spectra used in the further analysis are shown in Figure 5.2. An overview of the data and post-processing is provided in Appendix 5.B.

Small imperfections in the resulting data, such as residual atmospheric bands in R_{rs} (Figure 5.2), are no problem. For this work, it is only necessary to obtain a set of realistic spectra, not to determine IOPs. Negative R_{rs} were removed since they are not physical but instead the result of measurement error or over-correction of glint; this is no problem for the same reason.

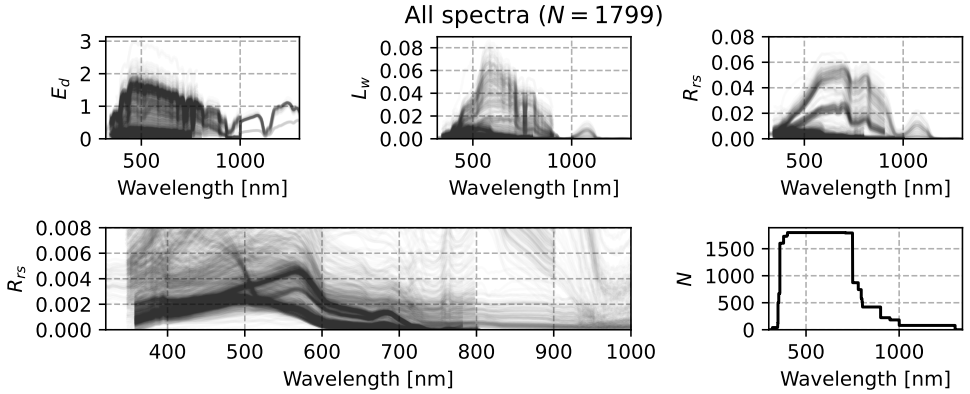


Figure 5.2: All radiometric data used in this work, in SI units. $N(\lambda)$ is the number of spectra that include wavelength λ. The bottom left panel is a zoom on the top right one. Individual spectra are plotted with high transparency.

5.3.2 Spectral convolution

Spectral convolution was implemented in the custom Python library described above. The radiometric data were interpolated to the SRF wavelengths. If the radiometric data and SRF wavelengths did not overlap fully, the convolution was only done if the integral of the SRF over the non-overlapping wavelengths was ≤5% of its total integral. The integration was done using the SciPy implementation of Simpson's rule in the `integrate.simps` function [398].

In each experiment, data were convolved in both L- and R-space, and the resulting reflectances were compared in absolute and relative terms. The absolute difference is $\Delta\bar{R}_{rs} = \bar{R}^R_{rs}(B) - \bar{R}^L_{rs}(B)$, meaning a positive $\Delta\bar{R}_{rs}$ corresponds to an overestimation in R-space. The relative difference was normalised to $\bar{R}^L_{rs}(B)$, and set to 0% if $\bar{R}^L_{rs}(B) = 0$ sr^{-1}. All spectra were treated separately, enabling a statistical analysis of the difference on varying input spectra. Due to the finite spectral resolution of the hyperspectral data, some data sets could not be convolved with some multispectral SRFs (see Appendix 5.A).

[19] https://pangaea.de/

5.3.3 Spectral response functions

Synthetic

The dependence of $\Delta\bar{R}_{rs}$ on band location and width was investigated by generating various synthetic boxcar and Gaussian filters. Both are common approximations of real SRFs [416,430]. Boxcars were evaluated on wavelengths with an non-zero response, Gaussians on wavelengths from 320–800 nm. For both, a 0.1 nm step size was used to properly sample narrow bands. Central wavelengths between 330–809 nm (1 nm steps) and FWHMs of 6–65 nm (1 nm steps) were used, representative of real multispectral instruments (Section 5.3.3).

Real instruments

The behavior of $\Delta\bar{R}_{rs}$ for real multispectral instruments, namely eleven satellite instruments and three low-cost sensors, was also investigated. A selection of these is shown in Figure 5.3. Panchromatic bands were not used as they are intended for spatial sharpening, not reflectance measurements. Only bands fitting the radiometric data (within 320–1300 nm) were used. The satellite instruments were the Enhanced Thematic Mapper Plus (ETM+) aboard Landsat 7 [452], Operational Land Imager (OLI) aboard Landsat 8 [453], Coastal Zone Color Scanner (CZCS), Sea-Viewing Wide Field-of-View Sensor (SeaWiFS), MODIS aboard Aqua and Terra, Medium Resolution Imaging Spectrometer (MERIS), VIIRS aboard Suomi NPP [454], MSI aboard S2A/B [455], and the Ocean and Land Colour Instrument (OLCI) aboard Sentinel-3A/B (S3A/B) [456]. These are all commonly used to measure R_{rs}. The low-cost sensors were one UAV, the DJI Phantom Pro 4, and two smartphones, the iPhone SE and Samsung Galaxy S8 [281]. Such sensors have become popular in their own right as they can provide radiance data, if radiometrically calibrated [121,281,425], but also serve as proxies for new cubesat sensors such as the Planet Labs RapidEye and Dove series.

The radiometric response and SRF may be affected by mechanical and electronic effects, including satellite launch and sensor drift, as well as by viewing angle and electronic crosstalk. Using up-to-date calibration data from the instrument developer negates these problems. Here, the SRFs recommended by instrument developers or in literature were used, representative of what is done in the wider literature.

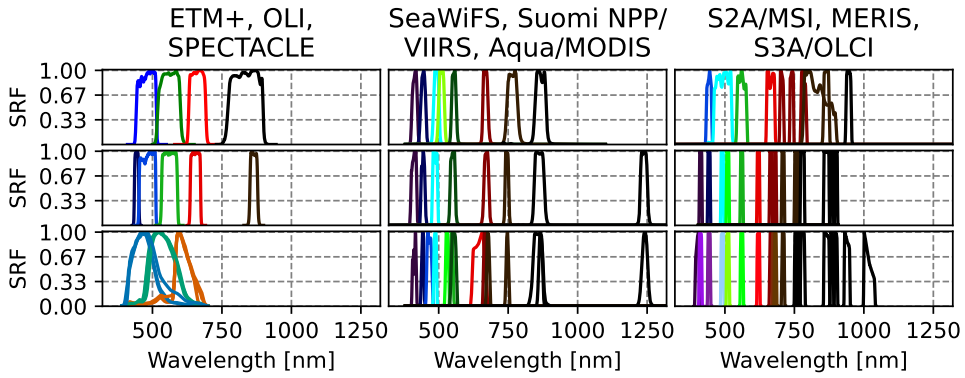


Figure 5.3: Selected spectral response functions (SRFs) of real sensors used in this work, labelled from top to bottom. For this plot, each SRF was normalised to a maximum of 1.

5.3.4 Retrieval algorithm propagation

Finally, the error induced by R-space convolution was propagated through several retrieval algorithms. These were the polynomial OCx chlorophyll-a (Chl-a) algorithms [220] for MODIS (OC6, OC3), SeaWiFS (OC4), MERIS (OC4), VIIRS (OC3), and CZCS (OC3), the exponential Ha+17 S2A/MSI Chl-a algorithm [221], and the polynomial Lymburner+16 (LL+16) OLI total suspended matter (TSM) algorithm [420]. These are representative of most multispectral retrieval algorithms in the literature, which differ only in bands used or coefficient values.

Equation (5.7) describes OCx, with [Chl-a] in mg m^{-3}, a_i instrument-specific empirical coefficients, and λ_B, λ_G the instrument's blue and green bands. The Ha+17 algorithm is given in Equation (5.8), with B3, B4 the R_{rs} in the respective S2A/MSI bands. The LL+16 algorithm is given in Equation (5.9), with G, R the R_{rs} in the OLI Green and Red bands, and TSM in mg L^{-1}.

$$\log_{10}([\text{Chl-a}]) = a_0 + \sum_{i=1}^{4} a_i \left[\log_{10}\left(\frac{R_{rs}(\lambda_B)}{R_{rs}(\lambda_G)}\right)\right]^i \tag{5.7}$$

$$[\text{Chl-a}] = 0.80 \exp\left(0.35 \frac{\text{B3}}{\text{B4}}\right) \tag{5.8}$$

$$\text{TSM} = 3957 \left(\frac{G+R}{2}\right)^{1.6436} \tag{5.9}$$

For each input spectrum, both $\bar{R}_{rs}^{L}$ and $\bar{R}_{rs}^{R}$ were propagated through each algorithm and the results were compared, analogous to the $\bar{R}_{rs}$ comparison described in Section 5.3.2.

5.4 Results

5.4.1 Simulated instruments

The reflectance convolution error $\Delta\bar{R}_{rs}$ was calculated for the synthetic SRFs described in Section 5.3.3. As an example, Figure 5.4 shows $\Delta\bar{R}_{rs}$ as a function of central wavelength λ_c and FWHM for the *seaswir-a* (see Table 5.1) data. The sign and magnitude of the error depend on the input spectrum. For example, the local minima around 400 and 520 nm correspond to local maxima in the derivative E_d spectrum $\mathrm{d}E_d/\mathrm{d}\lambda$. Similarly, the local maxima at 480 nm correspond to a local minimum in $\mathrm{d}E_d/\mathrm{d}\lambda$. Furthermore, the magnitude of $\Delta\bar{R}_{rs}$ increases with wider FWHMs. This is expected since E_d, L_w, and R_{rs} are less spectrally flat over a wider spectral range [430].

Rather than a random error around a median of 0, the difference is a systematic bias in either direction. This is especially clear in Figure 5.4 at $\lambda_c \leq 460$ nm. Being a bias, it needs to be corrected rather than simply incorporated into an error budget. This will be discussed in Section 5.5.

Similar trends were found in the other data sets and with the Gaussian SRFs. For the latter, the λ_c–$\Delta\bar{R}_{rs}$ relation was similar to boxcars with the same FWHM, but larger in magnitude and smoother. This is due to the Gaussian wings covering more of the spectrum than the boxcar's sharp edges. For example, for $\lambda_c = 420$ nm, $\Delta\bar{R}_{rs} = (-1.5 \pm 0.2)\%$ for a 30 nm boxcar and $(-3.8 \pm 0.4)\%$ for a 30 nm Gaussian, error bars indicating the 5%–95% range, for the

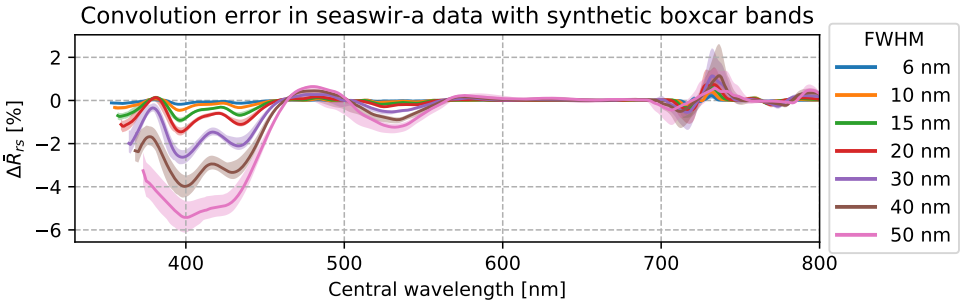


Figure 5.4: Relative reflectance convolution error (Section 5.3.2) in the *seaswir-a* data with boxcar filters of varying FWHM, as a function of central wavelength λ_c. Lines indicate the median error for each filter, shaded areas the 5%–95% range.

seaswir-a data. Finally, the same boxcar filter applied to the *tarao* data gave $\Delta\bar{R}_{rs}(420\text{ nm}) = (+0.01 \pm 0.02)\%$. This value is much smaller since the *tarao* spectra are smoother than the *seaswir-a* ones; a similar trend was seen across all data sets. These differences highlight the importance of determining this error for each filter and data set, as an ensemble correction is impossible.

5.4.2 Real instruments

$\Delta\bar{R}_{rs}$ was also calculated using the real SRFs described in Section 5.3.3. For example, Figure 5.5 shows the distribution of $\Delta\bar{R}_{rs}$ across all data for the five OLI bands. As with the synthetic sensors, $\Delta\bar{R}_{rs}$ is typically a bias in one direction rather than a random error and its magnitude and sign depend on the input spectrum. For example, in the OLI Blue band $\Delta\bar{R}_{rs} > 0$ for 77% (1380/1799) of spectra while in Green $\Delta\bar{R}_{rs} < 0$ for 80% (1444/1799). Furthermore, a similar trend for larger errors with wider bands was seen, for example in the OLI Green band (λ_c = 562 nm, FWHM = 57 nm) $\Delta\bar{R}_{rs} = (-0.2^{+0.4}_{-0.9})\%$ while in the similar S3A/OLCI Oa6 band (λ_c = 560 nm, FWHM = 10 nm) $\Delta\bar{R}_{rs} = (-0.00^{+0.03}_{-0.05})\%$. No significant differences were found between paired instruments such as S3A/OLCI and S3B/OLCI. Some multispectral band-data set combinations are technically invalid (Appendix 5.A); however, these need not be excluded from these overall statistics, as they do not affect the observed trends.

Comparing the convolution error between data sets, as in Figure 5.6 for the OLI Green band, again revealed significant differences. Depending on the data, $\Delta\bar{R}_{rs}$ was a systematic underestimation (*tarao* $\Delta\bar{R}_{rs} = (-0.7 \pm 0.2)\%$), overestimation (*seaswir-r* $\Delta\bar{R}_{rs} = (+0.2 \pm 0.1)\%$), or a random error around 0 (*orinoco* $\Delta\bar{R}_{rs} = (+0.1^{+0.3}_{-0.5})\%$). This is similar to what was observed in Section 5.4.1 and again shows that the error must be quantified separately for each filter and data set.

Low-cost sensors

Finally, the SPECTACLE low-cost sensors [281] are particularly interesting due to their broad bands. The convolution error in their RGB bands, using all data, is shown in Figure 5.7. Interestingly, $\Delta\bar{R}_{rs}$ was largest in the relatively narrow R bands, possibly due to the shapes of the input spectra or the multi-peaked SRFs [281]. Overall, the large magnitude of $\Delta\bar{R}_{rs}$

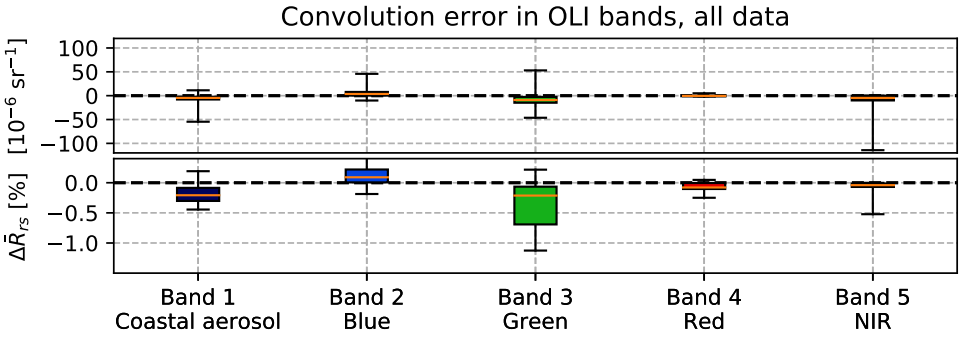


Figure 5.5: Reflectance convolution error in the OLI bands using all data. The boxplots show the error distribution across the data, with an orange median line, first–third quartile boxes, and 5%–95% percentile whiskers. Outliers are not shown for clarity.

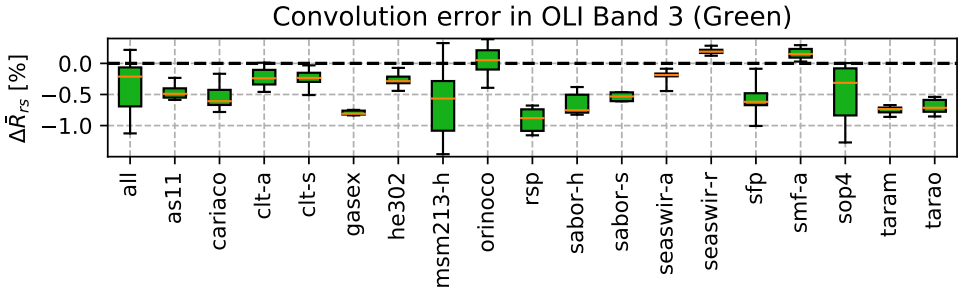


Figure 5.6: Reflectance convolution error in the OLI green band for each data set. The boxplots represent the error distribution within each data set as in Figure 5.5.

(down to −5% in the R bands) highlights the importance of correct spectral convolution for these sensors.

5.4.3 Retrieval algorithms

Finally, the reflectance convolution error was propagated through the retrieval algorithms described in Section 5.3.4. The results for the *smf-a* data set are shown in Figure 5.8. As in the previous sections, the propagated error in Chl-a and TSM was a bias of a few percent. Its sign varied by data set and by algorithm; for example, for the *seaswir-a* data, VIIRS OC3 underestimated Chl-a (ΔChl-a = $-1.4^{+0.7}_{-0.3}$%) while CZCS OC3 overestimated it (ΔChl-a = $+0.8^{+0.3}_{-0.1}$%). The magnitude of the error was consistently on the percent level for all data sets and algorithms. These results are representative for most band-ratio algorithms, as discussed in Section 5.3.4.

5.5 Discussion & conclusions

In this work, the effects of incorrectly convolving reflectance when simulating multispectral data (Section 5.2) were investigated. While this error has been pointed out previously [430,

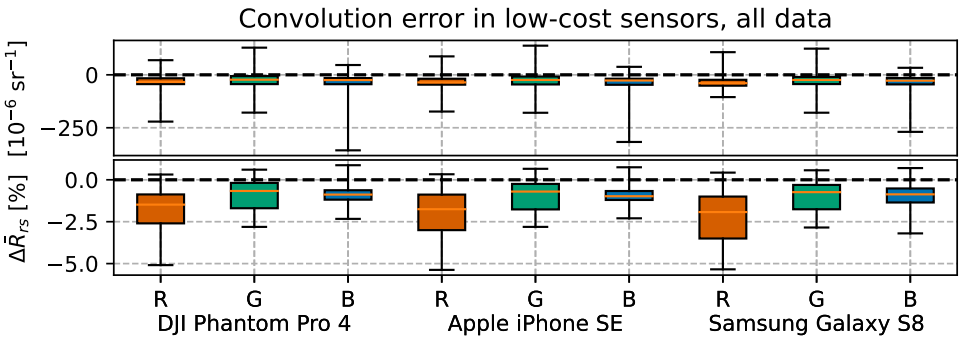


Figure 5.7: Reflectance convolution error in the SPECTACLE sensor RGB bands using all data. The boxplots represent the error distribution within each data set as in Figure 5.5.

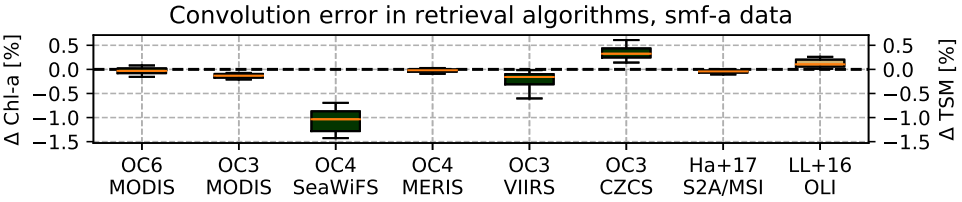


Figure 5.8: Propagated convolution error in the Chl-a (OCx, Ha+17) and TSM (LL+16) retrieval algorithms (Section 5.3.4), using the *smf-a* data set. The boxplots represent the error distribution within each data set as in Figure 5.5.

5

433, 435], it still commonly occurs in the literature (see Section 5.1). Only one quantitative analysis was found, in which for one data set and one sensor the difference was found to be $\lesssim$1% and neglected [430]. However, this result cannot be generalised to all data sets and sensors, as shown in this work.

Significant errors, up to several percent, in the remote sensing reflectance ($\Delta\bar{R}_{rs}$) were found for all data sets and sensor bands (Sections 5.4.1 and 5.4.2). The error was largest near features in the input spectra, particularly peaks in the derivative of E_d ($\mathrm{d}E_d/\mathrm{d}\lambda$), and for sensors with wide FWHMs, especially low-cost sensors (Section 5.4.2). For example, in the narrow (FWHM $\approx$ 10 nm) OLCI bands, $|\Delta\bar{R}_{rs}| \lesssim 0.1\%$, while in the wide (FWHM > 50 nm) R bands of low-cost sensors, $|\Delta\bar{R}_{rs}| > 5\%$ for >5% of the spectra. Furthermore, the magnitude and sign of $\Delta\bar{R}_{rs}$ differed significantly between data sets due to varying spectral shapes.

Since uncertainty requirements are typically ±5% for satellite-derived R_{rs}, and even stricter for validation data [88], errors on this scale are significant. Moreover, the error was typically a bias, causing a systematic over- or underestimation of $\bar{R}_{rs}$ and derived products. Preventing such biases is crucial to obtain representative data [88]. Finally, the convolution error is important simply due to its prevalence in the literature [62, 121, 221, 227, 276, 415, 416, 418, 420, 423–425, 431, 432].

If not prevented, the convolution error will create dubious patterns in combined time series. Depending on the data set and sensor, the convolution error is similar to or larger than errors found in existing band-shifting algorithms for combining multispectral time series [91, 419]. With the launch of PACE, for which time series extension is a primary goal [421],

this effect must be accounted for to achieve desired uncertainty requirements [88, 433].

Similarly, incorrect reflectance convolution in instrument validation leads to systematic over- or undercorrections. For example, reflectances from the HydroColor smartphone app have been validated using WISP [121] and HyperSAS [276] data, convolved in R-space, finding significant errors and biases. Biases of -9.5×10^{-4} to $+1.3 \times 10^{-4}$ sr^{-1} were found in the WISP comparison; in Section 5.4.2, the convolution error caused biases on the order of 10^{-4} sr^{-1} for 5%–14% of spectra, varying per band. Errors in the HyperSAS comparison were on the percent level, similar to the errors up to 5% found in Section 5.4.2. Interestingly, in both studies the convolved data underestimated the multispectral data, as would be expected from the negative biases found in this work. This suggests that the convolution error may have contributed a significant part of the error in both studies. However, a direct comparison is difficult due to differing input spectra, as shown in Figure 5.6, and band responses. Thus, the error in these cases cannot definitively be attributed to incorrect convolution. Additionally, many other factors causing significant errors in low-cost sensor data are known [281].

This importance for validation also applies to satellites. For example, in [424], systematic underestimations up to 1% were found in band-average $\bar{R}_{rs}$ (compared to hyperspectral R_{rs}) convolved in R-space with the OLI, MSI, and ETM+ SRFs. This is similar to, and may be explained by, the reflectance convolution error found for these sensors in Section 5.4.2 and shown in Figures 5.5 and 5.6. The same study found no significant errors in convolved VIIRS and OLCI reflectance, agreeing with the correlation between FWHM and error demonstrated in Section 5.4.1.

Conversely, the effects on retrieval algorithms are minor. The convolution error in Chl-a and TSM algorithms (Section 5.4.3) was on the percent level. Since errors in satellite-retrieved Chl-a can be up to 500% [232], a bias of a few percent can safely be neglected. Typical TSM errors are less extreme but still significantly larger than the ≤1% found here [227, 420]. While only a few algorithms were tested, as discussed in Section 5.3.4, these results are representative for most band-ratio algorithms. While many studies opt to use only the central band wavelength, not the full SRF [223, 227], in which case the convolution error does not occur, comparing narrow- and wide-band data that way introduces similar problems, described in [430].

Prevention of the convolution error is straight-forward while post-hoc correction is not. As explained in Section 5.2, simply convolving (ir)radiances instead of reflectance prevents the error from occurring, and is the only mathematically correct procedure. Of course this requires the original data to be available, which is not always true. Post-hoc correction is impossible since the error is highly variable across different sensors and data sets. When lacking original data, the reported uncertainty may simply be increased by a few percentage points [430] but this fails to account for systematic biases. An estimate may be made, for example by reconstructing L_w from a reported R_{rs} and simulated E_d, but this introduces further assumptions.

To this end, it is recommendable that published data sets, intended for satellite validation, contain not only products such as reflectance but also the raw data, at-sensor (ir)radiance data, and calibration data. This way, the convolution error can be avoided. Furthermore, it would greatly increase the amount of data available for other studies requiring radiometric data, such as those into glint removal [206]. Finally, publication of original data, as well as sensor characteristics, allows for traceability, which is crucial for quality control [88].

While this work focused on the remote sensing reflectance R_{rs} using ocean colour data, the principles and conclusions are broadly applicable. A general rule of thumb on convolution

practice is given in Section 5.2.4. In short, the principles outlined in this work are relevant to the simulation of low-resolution data from high-resolution data. This includes all types of reflectance, as well as other divisional quantities such as attenuation coefficients and degree of polarisation. Furthermore, it includes all types of convolution, including spectral and spatial. In all cases, a correct order of operations is crucial to prevent systematic errors. Any simplifications should be justified mathematically, not made at whim.

5.A Validity of spectral convolution

Consider a source radiance spectrum $L_{src}(\lambda)$, observed by a hyperspectral sensor with N bands. The spectral radiance arriving at the sensor $L(\lambda)$ is the product of the source spectrum and any atmospheric effects. However, the data used in this work were recorded in or only a few meters above the source, so atmospheric effects can safely be ignored:

$$L(\lambda) \approx L_{src}(\lambda) \tag{5.10}$$

The hyperspectral sensor records the radiance in bands $h = 1, 2, ..., N$, with central wavelengths λ_h. Each band h has its own SRF $S_h(\lambda)$, and the radiance recorded in band h, L_h, is the spectral convolution of the at-sensor spectrum $L(\lambda)$ with $S_h(\lambda)$. The integral is evaluated over all wavelengths; for clarity, this is not written explicitly in this section. The denominator in Equation (5.11) corrects for the spectral (or 'quantum') efficiency of the sensor in band h.

$$L_h = \frac{\int L(\lambda) S_h(\lambda) \mathrm{d}\lambda}{\int S_h(\lambda) \mathrm{d}\lambda} \tag{5.11}$$

The resulting spectrum measured by the hyperspectral sensor, $L_H(\lambda)$, consists of the individual band spectra:

$$L_H(\lambda) = L_1, L_2, ..., L_N = \frac{\int L(\lambda) S_1(\lambda) \mathrm{d}\lambda}{\int S_1(\lambda) \mathrm{d}\lambda}, \frac{\int L(\lambda) S_2(\lambda) \mathrm{d}\lambda}{\int S_2(\lambda) \mathrm{d}\lambda}, ..., \frac{\int L(\lambda) S_N(\lambda) \mathrm{d}\lambda}{\int S_N(\lambda) \mathrm{d}\lambda} \tag{5.12}$$

S_h affects $L(\lambda)$ in two ways. The first is to lower it due to the spectral efficiency of the sensor. This is described by an overall SRF $S_H(\lambda)$; dividing the data by $S_H(\lambda)$ corrects for this. The second effect is to smoothen the data: since in practice S_h is never a delta function, band h records not only the radiance at its central wavelength λ_h but also at other wavelengths where $S_h(\lambda) > 0$.

The smoothening can be described as a cross-correlation ($\star$) between the observed radiance $S_H(\lambda)L(\lambda)$ and a bandwidth function G. In reality, each band will have a slightly different G_h, for example due to stray light; however, for simplicity, here G is assumed to be the same for all bands. Then $L_H(\lambda)$ can be described as in Equation (5.13).

$$L_H(\lambda) = \frac{(S_H L) \star G}{S_H}(\lambda) \tag{5.13}$$

Now consider a multispectral band M with SRF $S_M(\lambda)$. Following the same logic, Equation (5.11) gives the radiance recorded in band M, L_M, as in Equation (5.14).

$$L_M = \frac{\int L(\lambda) S_M(\lambda) \mathrm{d}\lambda}{\int S_M(\lambda) \mathrm{d}\lambda} \tag{5.14}$$

However, when simulating multispectral data from hyperspectral data, the original radiance $L(\lambda)$ is not available. Instead, the recorded hyperspectral radiance $L_H(\lambda)$ is used. This means that in practice, one does not calculate L_M as in Equation (5.14) but an approximation L_M^H, as in Equation (5.15).

$$L_M^H = \frac{\int L_H(\lambda) S_M(\lambda) \mathrm{d}\lambda}{\int S_M(\lambda) \mathrm{d}\lambda} = \frac{\int \frac{(S_H L) \star G}{S_H}(\lambda) S_M(\lambda) \mathrm{d}\lambda}{\int S_M(\lambda) \mathrm{d}\lambda} \tag{5.15}$$

The approximation $L_M^H \approx L_M$ holds in two cases. The first is if $L_H(\lambda) \approx L(\lambda)$, that is if Equation (5.16) holds. From information theory it follows that this is true if G is significantly narrower than typical features in $L(\lambda)$.

$$\frac{(S_H L) \star G}{S_H}(\lambda) \approx L(\lambda) \tag{5.16}$$

The second case where $L_M^H \approx L_M$ holds is when the multispectral band M is significantly wider than G and typical features in $L(\lambda)$. Then, any radiance redistributed from λ_h to surrounding wavelengths in $L_H(\lambda)$ is still captured in the integral $\int L_H(\lambda) S_M(\lambda) \mathrm{d}\lambda$, and the value of the integral is the same.

E_d has narrow line features, as does L_w by extension. Hyperspectral (ir)radiance sensors typically undersample these features [97, 120], so Equation (5.16) does not hold in practice.

However, the second case does hold, if M is significantly wider than the hyperspectral band (or G). The Nyquist-Shannon theorem provides, to first order, a requirement: $L_M^H \approx L_M$ if the FWHM of M is at least twice that of the hyperspectral data. Then, hyperspectral data $L_H(\lambda)$ adequately approximate the original radiance $L(\lambda)$ for spectral convolution purposes.

5.B Radiometric data

Table 5.1 lists the radiometric data sets used in this work. Some of these contain unphysical data due to measurement errors, environmental effects, and instrumental problems [206, 431, 437]. This appendix describes how the data were filtered and homogenised before processing.

First, all spectra were converted to SI units. Second, negative R_{rs} values were clipped to 0 if $-10^{-4} < R_{rs}(\lambda) < 0$ as this is within typical measurement errors; spectra with any $R_{rs}(\lambda) \leq -10^{-4}$ were removed wholly. For *as11*, 5 spectra with negative R_{rs} were removed. For *cariaco*, 230 spectra missing R_{rs} and 11 missing E_d values were removed, as were 64 spectra with negative and 1 spectrum with unphysically high (>0.8) R_{rs}. For *clt-a*, 6 spectra with negative R_{rs} were removed. For *clt-s*, the spectra within 3-minute windows suggested in the accompanying documentation were averaged and L_w was calculated from L_u and L_t following the Mobley protocol [209]; 19 spectra with missing and 36 with negative R_{rs} were removed. For *gasex*, wavelengths $\lambda > 710$ nm were removed due to incomplete data. For *he302*, 3 spectra with $R_{rs}(800\,\mathrm{nm}) \geq 0.003$ were removed as outliers; the original authors noted the difficulty in normalising these data [206]. For *msm213-h*, L_w was used to reconstruct E_d; 179 spectra with missing data were removed, as were spectra with unphysically large jumps in E_d, namely 21 with $|E_d(\lambda_1) - E_d(\lambda_2)| \geq 0.2$ and $|E_d(\lambda_2) - E_d(\lambda_3)| \geq 0.2$ and 1

Label	Reference	Location	Instrument	Spectra			
				N	Range	FWHM	Step
as11	[438]	Arabian Sea	HyperTSRB	8	398 – 798	10	3.3
cariaco	[439]	Caribbean Sea	SpectraScan PR650	58	380 – 780	8	4
clt-a	[440]	Chesapeake Bay	ASD	56	350 – 1000	3	2
clt-s	[440]	Chesapeake Bay	HyperSAS	161	354 – 796	10	2
gasex	[441]	South Atlantic Ocean	HyperTSRB	7	352 – 710	10	2
he302	[442]	North Sea	RAMSES	41	320 – 950	10	5
msm213-h	[443]	North Atlantic Ocean	HyperPro	56	360 – 800	10	5
orinoco	[444]	Caribbean Sea, Orinoco River	SpectraScan PR650	70	380 – 780	8	4
rsp	[445]	Atlantic Ocean, Pacific Ocean	HyperPro	27	349 – 802	10	3.3
sabor-h	[446]	Atlantic Ocean	HyperPro	12	349 – 802	10	3.3
sabor-s	[446]	Atlantic Ocean	HyperSAS	6	358 – 749	10	1
seaswir-a	[37]	Gironde, La Plata	ASD	80	350 – 1300	3	1
seaswir-r	[37]	Gironde, La Plata, Scheldt	RAMSES	198	350 – 900	10	2.5
sfp	[447]	Florida Strait	SVC GER 1500	16	399 – 801	3	1.5
smf-a	[448]	Gulf of Mexico	ASD	46	350 – 1000	3	2
sop4	[449]	North Sea, Skagerrak	RAMSES	870	360 – 750	10	5
taram	[450]	Mediterranean Sea	HyperPro	42	348 – 803	10	3.3
tarao	[451]	Global	HyperPro	45	401 – 750	10	3.3

Table 5.1: Summary of radiometric data sets used in this work. Spectral range, FWHM, and step are in nanometers, the former rounded to the nearest integer. The number of spectra *N* is the number after post-processing and removal of bad data.

with $|E_d(\lambda_1) - E_d(\lambda_2)| \geq 0.35$, with λ_1, λ_2, λ_3 subsequent wavelengths. For *orinoco*, 1 spectrum with negative R_{rs} was removed. For *sabor-s*, which contains polarised and unpolarised spectra, only the latter were used and the wavelength range was clipped to 358–749 nm because of incomplete and noisy data elsewhere. For *seaswir-a*, the provided plaque radiance L_d was used to calculate $E_d = \pi L_d$ [209], assuming a plaque reflectance of $R_g \approx 1$ [431]; no units were given for these data, so the resulting E_d spectra were divided by 10^5 to be in line with the others. 'Water reflectance' R_w was provided instead of R_{rs}; comparing [431] and [209] showed that $R_{rs} = R_w/\pi$. Finally, 35 spectra with missing E_d and 39 with negative R_{rs} were removed. For *seaswir-r*, R_w was similarly converted to R_{rs} and 2 spectra with negative R_{rs} were removed. For *sop4*, the provided L_u and L_s were used to calculate R_{rs} following the Mobley protocol [209] for simplicity [206]. The R_{rs} spectra were then normalised by subtracting $R_{rs}(750\text{ nm})$ [209] and the results used to re-calculate L_w. Next, 885 spectra with unphysical $\max(E_d(\lambda)) < 0.01$ were removed and the wavelength range cropped to 360–750 nm to remove noisy data. Spectra with unphysical features were then removed, namely 23 with $|R_{rs}(\lambda_1) - R_{rs}(\lambda_2)| \geq 0.005$ and $|R_{rs}(\lambda_2) - R_{rs}(\lambda_3)| \geq 0.005$ and 3 with $E_d(400\text{ nm}) - E_d(405\text{ nm}) > 0.01$; finally, 289 spectra with negative R_{rs} were removed. The remaining data sets (*sabor-h*, *sfp*, *rsp*, *taram*, and *tarao*) required no post-processing. For *sfp*, only the mean R_{rs} spectra were used.

Funding

European Commission Horizon 2020 programme (grant nr. 776480, MONOCLE).

Acknowledgements

The author wishes to thank Emmanuel Boss, Stefan Simis, Kiwi and Frans Snik, and Mortimer Werther for their suggestions and support in writing this work. Additionally, the author wishes to thank the two anonymous referees for their thorough and constructive reviews. Data analysis and visualisation were done using the AstroPy, Matplotlib, NumPy, and SciPy libraries for Python.

6 | A universal smartphone add-on for portable spectroscopy and polarimetry: iSPEX 2

Adapted from

Olivier Burggraaff, Armand Perduijn, Robert van Hek, Norbert Schmidt, Christoph Keller, Frans Snik

Micro- and Nanotechnology Sensors, Systems, and Applications XII, Proceedings of the SPIE 11389, 113892K (2020)

doi:10.1117/12.2558562

Spectropolarimetry is a powerful technique for remote sensing of the environment. It enables the retrieval of particle shape and size distributions in air and water to an extent that traditional spectroscopy cannot. SPEX is an instrument concept for spectropolarimetry through spectral modulation, providing snapshot, and hence accurate, hyperspectral intensity and degree and angle of linear polarisation. Successful SPEX instruments have included groundSPEX and SPEX airborne, which both measure aerosol optical thickness with high precision, and soon SPEXone, which will fly on PACE. Here, we present a low-cost variant for consumer cameras, iSPEX 2, with universal smartphone support. Smartphones enable citizen science measurements which are significantly more scaleable, in space and time, than professional instruments. Universal smartphone support is achieved through a modular hardware design and SPECTACLE data processing. iSPEX 2 is manufactured through injection moulding and 3D printing. A smartphone app for data acquisition and processing is in active development. Production, calibration, and validation are ongoing, with promising initial results. Scientific applications will include citizen science measurements of aerosol optical thickness and surface water reflectance, as well as low-cost laboratory and portable spectroscopy.

6.1 Introduction

Spectropolarimetry, the characterisation of reflected or emitted light at different wavelengths and polarisation states, is a powerful technique for remote sensing of the environment [132, 238, 457]. Most prominently, the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) satellite due for launch in 2024 will fly two spectropolarimetric instruments, namely HARP-2 and SPEXone [86, 111]. HARP-2 will observe linear polarisation (LP) in four spectral bands (440, 550, 670, 870 nm) at 10–60 angles with a polarimetric accuracy of <0.005 in Degree of Linear polarisation (DoLP) [86, 185]. Meanwhile, SPEXone will observe at five discrete angles (0°, ±20°, ±57°) with continuous spectral coverage from 385–770 nm and a DoLP accuracy of 0.0025 [184, 185]. Instruments observing circular polarisation are also under active development, such as the Life Signature Detection polarimeter (LSDpol) [458], but current efforts typically focus on linear polarimetry, as does this work.

Science cases for linear spectropolarimetry include the retrieval of aerosol and hydrosol particle properties, the beam attenuation and absorption coefficients (c, a) in water, and the study of vegetation covers. For aerosols, there is already a long history of multi-angle spectropolarimetric observations, from which parameters including particle size and shape distributions, spatial distributions, and chemical composition can be derived [132, 185]. More recently, this has been extended to oceanic hydrosols, where the bulk refractive index, particle size distribution, and c can be derived from DoLP [163]. This has been demonstrated for example by Gilerson et al. with a retrieval algorithm for c and a from multi-angular DoLP data [174]. Finally, spectropolarimetry of vegetation probes its physical characteristics, such as leaf orientation, and provides reflectance distribution functions, which are crucial for improving the accuracy of air- or space-based aerosol retrieval algorithms [459].

Combining spectral and polarimetric measurements can be done in multiple ways [170]. First, regular spectroradiometers can be fitted with rotating polarising filters, as was done in the aforementioned studies of water and vegetation [174, 459]. A second method is *channelled* spectropolarimetry, where polarisation information is encoded into the spectrum itself. One method for channelled linear spectropolarimetry is SPEX [175], the basis for SPEXone [184]. In SPEX, incoming light is modulated with a sine wave with an amplitude and phase depending on the DoLP and the Angle of Linear polarisation (AoLP), respectively [175]. This is further explained in Section 6.2.2.

The SPEX technique has been applied successfully in two high-end field-going instruments measuring aerosol optical thickness (AOT, sometimes termed aerosol optical depth, AOD), namely groundSPEX [125] and SPEX airborne [177]. GroundSPEX is a ground-based instrument based on a dual-channel fiber-optic spectrometer with SPEX optics on a moving mount, allowing sequential measurements at multiple angles. Its AOT measurements are well-correlated (Pearson r = 0.932) [125] with data from AERONET, the global network of photometers observing the solar almucantar and principal plane [134]. SPEX airborne, as the name implies, is an airborne instrument, simultaneously observing at nine fixed viewing angles. A 2017 campaign on a NASA ER-2 high-altitude aircraft demonstrated excellent agreement (RMS DoLP differences of 0.004–0.02) with coflying instruments [177].

A third successful SPEX variant was iSPEX, a smartphone-based version [94, 256]. Developed as a low-cost citizen science (CS) tool for AOT measurements, in 2013 iSPEX was used in CS campaigns yielding ~10 000 observations in the Netherlands. iSPEX data agreed well with AERONET reference data, showing typical standard errors and offsets in AOT of <0.1, while the typical absolute DoLP uncertainties were ~0.03 [94]. However, the original

iSPEX add-on, app, and data had several limitations. First and foremost, the add-on was tailored to the iPhone 4 and 5 and did not work on later iPhone models (bar the iPhone SE) or any Android devices, limiting its reach and future compatibility. Second, at the time iOS only offered very limited camera controls for third-party applications, meaning iSPEX spectra were gathered at very coarse resolution, in the highly non-linear JPEG format, and with varying and uncontrollable exposure settings [281]. Thus, iSPEX data were only reliable when averaged over at least 50 individual measurements. Finally, iSPEX had a single-beam SPEX implementation, meaning the polarisation modulation could not be distinguished from inherent spectral features [94].

We present iSPEX 2, an upgraded version of iSPEX, solving the problems faced by its predecessor. First, the iSPEX 2 hardware is designed to universally support all smartphones. Second, using our SPECTACLE method and database for camera calibration, smartphone cameras offer data similar in quality to professional radiometers [281]. Third, a dual-beam SPEX implementation facilitates distinguishing between spectral and polarimetric signals. This improves the polarimetric accuracy and enables pure spectroscopy for uniform targets.

iSPEX 2 was developed specifically for aerosol and ocean colour measurements. Like its predecessor, it can be used for large CS campaigns to measure AOT [94], with higher quality data, but also for individual AOT measurements like other SPEX variants [125, 177], though with a coarser spectral resolution. These can be used to fill in temporal and spatial gaps in AOT coverage for satellite atmospheric correction algorithms. The AOT data may be further improved through aureole, almucantar, and near-horizon measurements [134, 460, 461]. Ocean colour measurements will include unpolarised remote sensing reflectance (R_{rs}), similar to the HydroColor app [121] but hyperspectral, and polarised reflectance as discussed above [174, 238]. These too can be used to provide coverage in scenarios without coverage by high-end sensors, but also to validate satellite measurements. Finally, iSPEX 2 can be used as a low-cost instrument for portable or laboratory spectroscopy [98, 143].

The working principle of SPEX and its implementation in iSPEX 2 are described in Section 6.2. The physical design of the add-on is described in Section 6.3. Section 6.4 describes the production process. The current and planned data acquisition and processing pipeline are given in Section 6.5. Finally, Section 6.6 contains current progress on and future plans for calibration, validation, and scientific applications of iSPEX 2.

6.2 Working principle

6.2.1 Definitions

Spectral polarisation states are most easily described using a wavelength-dependent Stokes vector $\vec{S}(\lambda)$, defined in Equation (6.1). Here $I(\lambda)$ is the total spectral radiance, $Q(\lambda)$ and $U(\lambda)$ the linear polarisation state, and $V(\lambda)$ the circular polarisation state. Here, we define $+Q$ as horizontal and $-Q$ as vertical polarisation, $+U$ and $-U$ as $+45°$ and $-45°$ from the horizontal, and $+V$ and $-V$ as right- and left-handed circular polarisation, respectively. Lowercase q, u, v are the fractional polarisation, normalised by $I(\lambda)$. I, Q, U, V are sometimes referred to as S_0, S_1, S_2, S_3 respectively [457]. In this work, circular polarisation in incoming light will be neglected as typically $v \lesssim 10^{-3}$ in nature [458].

$$\vec{S}(\lambda) = \begin{bmatrix} I(\lambda) \\ Q(\lambda) \\ U(\lambda) \\ V(\lambda) \end{bmatrix} = I(\lambda) \begin{bmatrix} 1 \\ q(\lambda) \\ u(\lambda) \\ v(\lambda) \end{bmatrix} = \begin{bmatrix} I_0(\lambda) + I_{90}(\lambda) \\ I_0(\lambda) - I_{90}(\lambda) \\ I_{45}(\lambda) - I_{-45}(\lambda) \\ I_R(\lambda) - I_L(\lambda) \end{bmatrix} \tag{6.1}$$

The state of linear polarisation is also described by the degree and angle of linear polarisation, DoLP or $P_L(\lambda)$ and AoLP or $\phi_L(\lambda)$ respectively. These are defined in Equations (6.2) and (6.3) [175]. In practice, the arctan2 operator is used in Equation (6.3).

$$P_L(\lambda) = \frac{\sqrt{Q(\lambda)^2 + U(\lambda)^2}}{I(\lambda)} = \sqrt{q(\lambda)^2 + u(\lambda)^2} \tag{6.2}$$

$$\phi_L(\lambda) = \frac{1}{2} \arctan\left(\frac{U(\lambda)}{Q(\lambda)}\right) = \frac{1}{2} \arctan\left(\frac{u(\lambda)}{q(\lambda)}\right) \tag{6.3}$$

Finally, optical elements are described through their 4×4 Mueller matrix M, describing how the element modifies the incident $\vec{S}(\lambda)$. Each element of M can have its own wavelength dependence. Passing through an element X modifies $\vec{S}(\lambda)$ to be $M_X\vec{S}(\lambda)$. The Mueller matrix of a chain of elements X, Y, Z is simply the product of their individual Mueller matrices $M_Z M_Y M_X$.

6.2.2 SPEX polarisation modulation optics

The SPEX polarisation modulation optics (PMO) consist of three elements, namely a quarter-wave plate (QWP), multi-order retarder (MOR), and analysing linear polariser (ALP) [175]. Their orientations and function are as follows:

- *Quarter-wave plate*: The QWP has its fast axis at 0° (+Q, horizontal). It should be highly achromatic and interchanges the Stokes U and V components, making the instrument insensitive to circular polarisation. Residual chromaticity from misalignment, deviations in retardance $\delta_{QWP}(\lambda)$, and other effects must be calibrated [125]. In iSPEX 2, an Edmund Optics WP140HE (#88-253) $\lambda/4$ polymer retarder foil is used.

- *Multi-order retarder*: The MOR has its fast axis at +45° from horizontal (+U). Its retardance $\delta_{MOR}(\lambda)$ is highly chromatic, exchanging the incoming Q and V components by a fraction depending on the wavelength. As with the QWP, the performance of the MOR requires extensive calibration [125]. The first iSPEX 2 units contain a stack of two Meadowlark B4 polymer retarder foils [462], with a nominal retardance of 4λ each at 560 nm. For the future, alternatives are being investigated, as described in Section 6.4.1.

- *Analysing linear polariser*: The ALP imprints the modulation onto the exiting spectrum in Stokes I, and can be implemented in several ways. The single-beam approach uses a single linear polariser, parallel or orthogonal to the slit, as in the original iSPEX [94]. This approach does not allow for full linear spectropolarimetry, as the modulation and inherent spectral properties cannot be fully distinguished. This is possible in the dual-beam approach, where both directions are measured. A polarising beamsplitter is used in groundSPEX [125], SPEX airborne [177], and SPEXone [184]. In iSPEX 2, a pair of Polarization.com PFSC NA foils is used, oriented parallel or orthogonal to the two slits (Section 6.3.1).

Together, the PMO induce a modulation in the outgoing Stokes I radiance $I_{\pm}(\lambda)$ (where the sign ± corresponds to the two ALP orientations) that depends only on the incoming radiance $I_{in}(\lambda)$, $P_L(\lambda)$, $\phi_L(\lambda)$, λ, and $\delta_{MOR}(\lambda)$. This is described in Equation (6.4). The modulation is a sine wave on the radiance spectrum, quasi-periodic in $1/\lambda$, its amplitude and phase corresponding to the DoLP (P_L) and AoLP (ϕ_L), respectively [175]. These parameters are retrieved by fitting Equation (6.4).

$$I_{\pm}(\lambda) = \frac{I_{in}(\lambda)}{2}\left[1 \pm P_L(\lambda)\cos\left(\frac{2\pi\delta_{MOR}(\lambda)}{\lambda} + 2\phi_L(\lambda)\right)\right] \tag{6.4}$$

In dual-beam mode, the modulations in $I_+(\lambda)$ and $I_-(\lambda)$ are exactly opposite for a uniform target, so the total radiance and modulation can be disentangled as shown in Equations (6.5) and (6.6). However, this is complicated in practice due to imperfections in the optical elements, misalignments, differences in transmission between the two beams, and nonnormal incidence [125, 177].

$$I_+(\lambda) + I_-(\lambda) = I_{in}(\lambda) \tag{6.5}$$

$$\frac{I_+(\lambda) - I_-(\lambda)}{I_+(\lambda) + I_-(\lambda)} = P_L(\lambda)\cos\left(\frac{2\pi\delta_{MOR}(\lambda)}{\lambda} + 2\phi_L(\lambda)\right) \tag{6.6}$$

6.3 Add-on design

The iSPEX 2 add-on is a whole divided into three parts, as shown in Figure 6.1. These are a tube containing the PMO and other optics (Section 6.3.1), a clip to clamp onto a smartphone (Section 6.3.2) and a backplate for aligning the tube with the smartphone camera (Section 6.3.3). A cross-section is shown in Figure 6.2. The tube can be used on any camera, including smartphones but also UAVs and webcams. The clip can be used with nearly all smartphones, as most models have a similar form factor [281]. Finally, the backplate is unique to each smartphone model.

6.3.1 Optical tube

The iSPEX 2 optics consist of a double slit (side by side), the SPEX PMO (Section 6.2.2), a collimator lens, and a transmission grating, as shown in Figure 6.3. There are two slits, each 0.25 mm wide and 9 mm long, located side by side to measure in quasi-dual-beam mode. The PMO are placed directly behind the slits to minimise instrumental polarisation through stray light. A small plastic cradle holds the PMO foils in place. Dual-beam mode is achieved by having a horizontal ALP in the PMO behind the left slit and a vertical one behind the right slit. Dual-beam mode requires a uniform target between the two slits; this assumption holds in the center for smooth surfaces like sky polarisation [125], but not toward the edges. From the PMO, the modulated light propagates to a custom-made collimator lens (f = 35 mm) and a 1000 line/mm holographic transmission grating foil (Edmund Optics #52-116), dispersing the light onto the smartphone camera; the camera optics then register the spectra, as shown in Figures 6.2 and 6.7. A rubber seal blocks stray light, as shown in Figure 6.2.

The optics are located in a plastic tube, as shown in Figure 6.3. The tube itself is 35 mm long along its optical axis, which is angled +17.3° upwards to project the entire zeroth and

(a) iSPEX 2, whole.

(b) iSPEX 2, exploded view.

Figure 6.1: Render of the iSPEX 2 add-on, as a whole attached to a smartphone (left) and exploded into its three components (right). These are, from left to right, the smartphone clip, optical tube, and smartphone backplate. The smartphone is seen from the back.

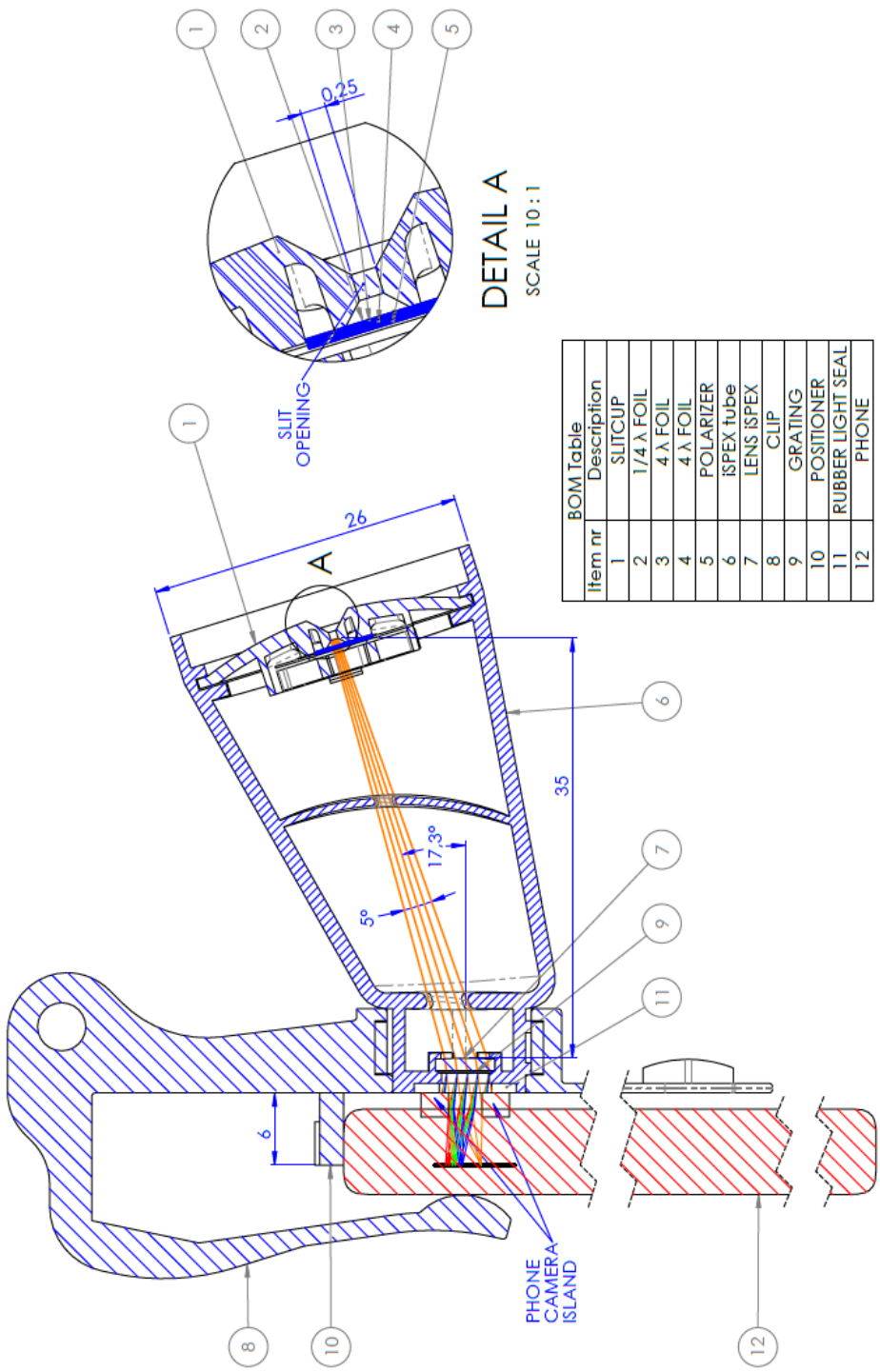


Figure 6.2: Cross section of iSPEX 2 (blue) attached to a smartphone (red) with a zoom on the slit (detail A). Five rays are traced from the slit to the grating in orange, then dispersed onto the camera chip. Sizes are in mm.

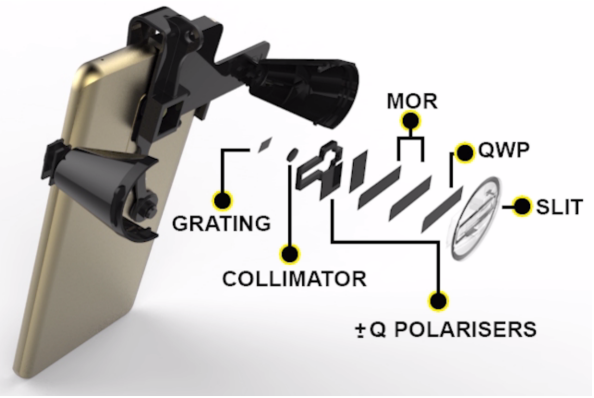
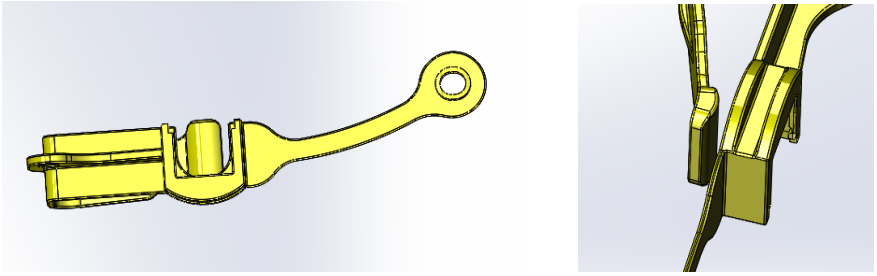


Figure 6.3: Exploded view of the iSPEX 2 optical tube with the slit, quarter-wave plate (QWP), multi-order retarder (MOR), 0° and 90° polarisers, collimator lens, and grating foil indicated.

first orders of the spectrum on the smartphone camera. The tube consists of two halves (left and right) which are produced separately and click together along its length (Section 6.4.4). A baffle is located halfway along the tube, consisting of overlapping protrusions from either tube half. This ensures overlapping coverage and thus reduces light leakage. The slit end of the tube has two 'ears' to which additional add-ons can be attached, such as a cuvette holder for transmission spectroscopy. The camera end has two ridges to which the clip (Section 6.3.2) attaches; these can also be used for custom attachments for different cameras.

6.3.2 Smartphone clip

iSPEX 2 attaches to smartphones using a clip, as seen in Figure 6.1. The clip design is shown in detail in Figure 6.4. It attaches to the smartphone with a clamp on the front (screen) side, directly behind the camera. This clamp is 14 mm wide and made from soft plastic to prevent scratching. It is attached to the clip over the top of the smartphone. Additionally, the clip has a 'clapper' extending 49 mm below the camera along the back side, with a suction cup at the end. This attaches to the flat back surface of the smartphone. This double attachment



(a) Smartphone clip.

(b) Zoom on clamp.

Figure 6.4: CAD model of the iSPEX 2 smartphone clip. The optical tube and backplate slot into the central opening, opposite the clamp. The long, curved extension is the 'clapper' with a suction cup at the end. Ridges along the clip provide stiffness and strength. The clip is fabricated in black.

prevents rotation of the add-on and ensures the slit is always projected horizontally onto the camera. The clapper is curved 11 mm to the right of the camera (seen from the back, as shown in Figure 6.5) so the suction cup does not fall off the edge on devices with cameras near the edge, such as iPhones.

Several iterations of the clip design were necessary. Originally, the clip had no clapper and attached to the right side of the smartphone (seen from the front), but this acted as a lever and caused the add-on to rotate under its own weight. This was solved by having the clip attach over the top, so its weight rests on the smartphone, and adding a clapper. The clapper originally had a small magnet at its end instead of a suction cup, but smartphone backsides were found to be only weakly magnetic in only a few places.

6.3.3 Smartphone backplate

iSPEX 2 is placed in front of the smartphone camera using a backplate unique to each smartphone model, as shown in Figure 6.5. Smartphones with multiple backside cameras typically feature one wide-view camera with generic optical properties and a focal length of 3.8–4.5 mm (corresponding to a field-of-view of 60°–75° × 45°–55° [281]), which is used for iSPEX 2. The backplate consists of a universal plate with two positioners with different positions for each smartphone model. One rests on top of the smartphone, the other on the right (seen from the back). Since most smartphone cameras are in the center or on the left, this design ensures that most of the weight of iSPEX 2 rests on the smartphone, rather than create a lever.

The positioner locations depend on the dimensions of the smartphone, the curvature of its top corners, and the locations of buttons along the side. We are compiling a database of popular smartphone models including these parameters. The top positioner has a small notch to accommodate the iPhone SE, one of the few popular models with buttons along the top. The positioners are 21.5 mm (side) and 30 mm (top) long and 6 mm wide, providing sufficient coverage even on smartphones with curved edges.

Figure 6.5: Render of iSPEX 2 attached to several smartphones with different dimensions and camera locations, seen from the back. Each has a backplate with unique positioner locations to place the optics directly in front of the camera.

6.4 Production

6.4.1 Multi-order retarder foils

As described in Section 6.2.2, iSPEX 2 uses polymer retarder foils to produce the SPEX polarisation modulation. The first units use two Meadowlark B4 foils [462] with 4λ retardance (at 560 nm) each, as did the original iSPEX [94], for a nominal total retardance of 8λ at 560 nm. This induces a modulation with 7 full periods across the typical spectral range of smartphone cameras (as seen in Figure 6.10b), which is 390–700 nm [281].

These foils are produced by stretching transparent sheets of polymer, such as polyvinyl alcohol (PVA), polycarbonate (PC) or poly(methyl methacrylate) (PMMA) [462]. While the pre-fabricated foils provide the desired retardance with great consistency, they are prohibitively expensive for low-cost CS purposes.

To enable high-volume throughput, integrated in our production line, we are exploring internal production of MOR foils, based on a setup previously used to stretch sheet metal [463]. Initial experiments are focused on finding the optimal material from PVA, PC, PMMA, polyethylene terephthalate (PET), and Zeonor cyclo olefin polymer, at various thicknesses from 50–200 µm. Further experiments will determine the efficacy of softening the foils through heat (up to 80 °C) and the maximum achievable retardance. Visual inspection and spectrally resolved measurements through a crossed polariser setup will be used to measure the retardance of sections of the foil during the stretching process, as spatial variations in retardance and fast axis orientation are expected [462]. The end goal is to mass produce low-cost MOR foils with sufficient quality for iSPEX 2, not necessarily for high-end commercial purposes.

6.4.2 Injection moulding

Like the original [94], iSPEX 2 is produced through injection moulding. This is inexpensive yet precise. Components that can be injection moulded include the collimator lens, optical tube, smartphone clip, and backplate. The suction cup (Section 6.3.2) and optical foils (Sections 6.2.2 and 6.3.1) are purchased or self-produced by other means (Section 6.4.1).

Various plastics are used for iSPEX 2. Thin parts, including the tube and backplates, are manufactured from polycarbonate-acrylonitrile butadiene styrene (PC-ABS) coloured black with masterbatch (MB). PC-ABS and polypropylene (PP) are being evaluated as materials for the clip. The collimators are produced separately from Zeonor 330 plastic.

The optical tube and clip are manufactured from a single mould (Figure 6.6), while a separate mould is used for the backplate. The backplate mould is complex, requiring two sliders to account for the device-dependent positioners. By parameterising the slider positions based on the smartphone dimensions, the mould can instantly be adjusted to a different device.

6.4.3 3D printing

Except for the optical components, iSPEX 2 units can also be 3D printed, for which we will provide model files. This was used extensively in the development phase for quick testing and will be useful for future compatibility. For example, this allows users to self-produce backplates tailored to new smartphone models not included in our database. Local production through 3D printing also reduces the unit cost, especially valuable in resource-poor areas, one of the prime target audiences for smartphone spectroscopy [143].

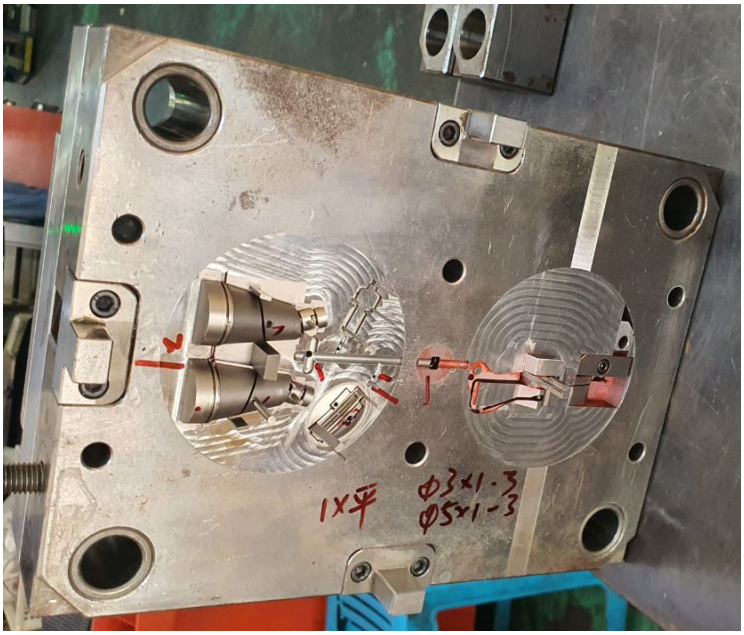


Figure 6.6: Injection mould used to produce the iSPEX 2 optical tube and smartphone clip.

However, 3D printing introduces several difficulties. First of all, the PMO and grating foils cannot be 3D printed and thus must still be purchased and cut to size. Second, low-cost 3D printing techniques inherently have wider production tolerances than injection moulding. With 3D printed prototypes we often found it necessary to manually file or cut components to make them fit tightly. Finally, some 3D printing materials such as PA nylon are translucent; we found this easiest to counteract by covering the entire unit twice over with a felt-tip pen. Because of these complications, we use 3D printing only for prototyping, not production.

6.4.4 Assembly

Assembly of iSPEX 2 units is straight-forward. First, the PMO foils are placed in their cradle in the correct orientations. Next, the cradle, collimator lens, and grating foil are placed into the corresponding slots in the tube halves, which are then clicked together. When used with a smartphone, the optical tube is then slotted into the smartphone clip, followed by the backplate corresponding to the smartphone model. The backplate is easily removed for use with a different device. For other uses, a custom attachment between tube and camera can easily be manufactured by the user. The tube can even be used with the naked eye for a qualitative measurement or a demonstration.

For the injection moulded units, the PMO and grating foils are punched into an asymmetric shape to prevent confusion of their optical axes; the cradle has corresponding protuberances. Like with the original iSPEX, a calibration setup in the factory is used to verify that all optics are oriented correctly [94]. For 3D printed units, this must be done carefully by hand.

6.5 Data acquisition and processing

This section describes the acquisition, calibration, and processing of iSPEX 2 data. These are acquired as RAW images on a smartphone using an app based on SPECTACLE [281], described in Section 6.5.1. Currently, these RAW data are manually uploaded to a PC for calibration and processing. Our goal is to move as much of this as possible to the smartphone, possibly with additional cloud computing in a back-end server for devices with insufficient computational power.

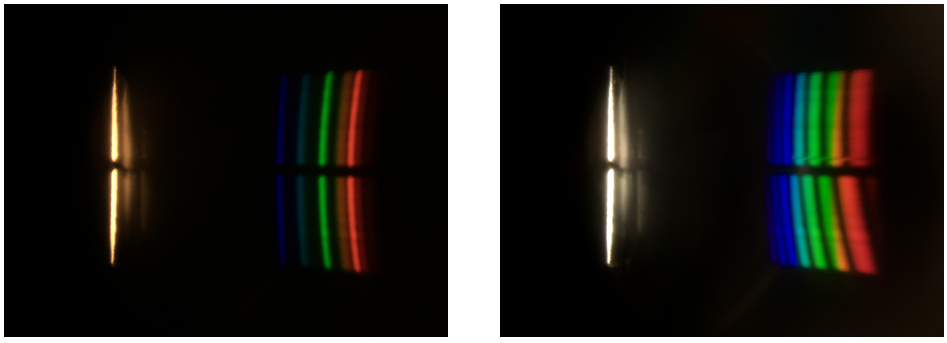
6.5.1 Smartphone app

Data acquisition on smartphones is done using a custom-designed app, based on SPECTACLE [281], currently in development for iOS and Android. Significant changes to iOS mean few elements from the 2013 iSPEX app [94] can be reused for iSPEX 2. However, user feedback on the original app is taken into account. For example, some users misunderstood the scientific aims and methods of iSPEX because these were not explained clearly [256]. Difficulties in installing the add-on and interpreting feedback from the app were also noted [297].

Data are obtained in RAW image format because of its high linearity and dynamic range [281]. This is in contrast to the JPEG images taken with the original, where non-linearity and white balance introduced significant problems [94, 281]. Two examples, taken with a 3D printed prototype, are shown in Figure 6.7. Aside from problems due to faults in the prototype, such as stray light, which will be reduced in the final product, these images are representative examples of iSPEX 2 data. Processing of these images (Section 6.5.2) is currently done on PC, but will be done in-app in the future.

The following data acquisition protocols will be included in the initial release of the app:

- Wavelength calibration: Single observation of a fluorescent light, as described in Section 6.5.3.



(a) Fluorescent light.

(b) Reflected sunlight through a 100% polariser.

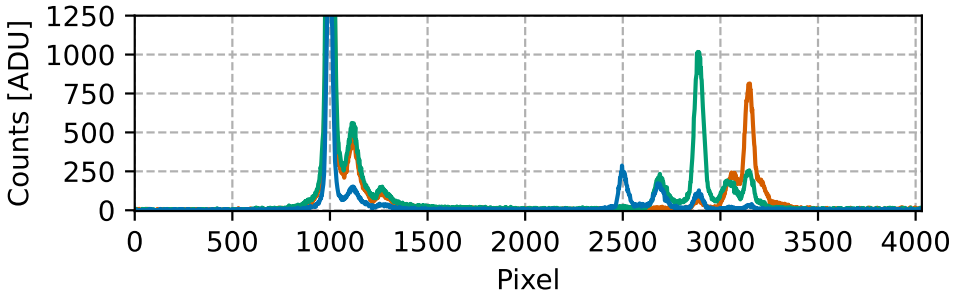
Figure 6.7: Spectra taken with an iSPEX 2 prototype on an iPhone SE. The fluorescent light spectrum (left) is used in the wavelength calibration, and clearly shows the smile and keystone effects. The 100% polarised image (right) shows the SPEX modulation in the two $\pm Q$ spectra, with an additional curvature due to variations in retardance. The small glitch in the top spectrum and the stray light are due to manufacturing faults in this prototype. These images were taken in RAW format and converted to JPEG for visualisation; all data processing is done on RAW images only.

- Aerosols (AOT): Series of observations from horizon to zenith along the principal (observer-Sun-zenith) plane, as with other SPEX variants [94, 125, 177]. The optimal number of observations is to be determined based on data quality and computational considerations, specifically the speed at which RAW images can be saved on smartphones.

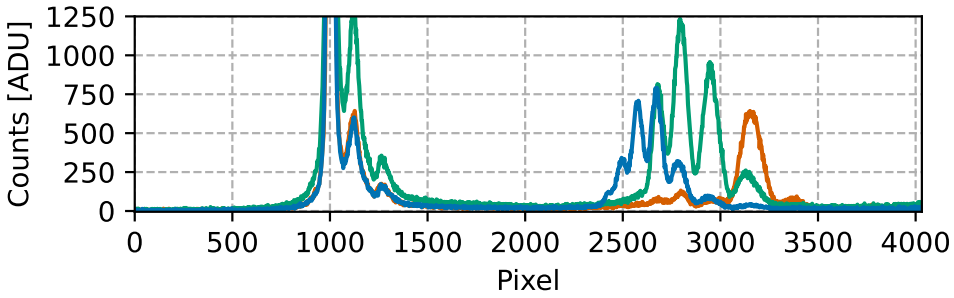
- Water (R_{rs}): Series of observations according to the Mobley protocol [209], measuring sky radiance at 40° from zenith, upwelling radiance at 40° from nadir, and downwelling irradiance with a grey card at 40° from nadir. The same protocol is used in the HydroColor app, which does multispectral (RGB) measurements [121].

Users are guided through these protocols with text explaining what to do and, for example, arrows to guide them in the right direction, using the smartphone compass and accelerometers. Citizen scientists have been involved in the development of these protocols from the start, to ensure user-friendliness. Optimal exposure settings for each protocol are currently hard-coded but in the future will be determined automatically.

6.5.2 Data pre-processing



(a) Fluorescent light.



(b) Reflected sunlight through a 100% polariser.

Figure 6.8: Two adjacent rows of pixels (one GR, one BG_2) in the spectra shown in Figure 6.7, split into the RGB channels. The G_2 channel is not shown here. The slit and several ghosts are visible from 800–1500 pixels, the first order spectrum from 2300–3500 pixels. The data have been corrected for bias and flat-field.

Data are processed according to the SPECTACLE method [281], which was originally developed for iSPEX 2. A Python library for data processing specific to iSPEX 2 is currently in development[20]. iSPEX 2 data are corrected for bias and flat-field using SPECTACLE; in spectra taken with prototypes, dark current is negligible compared to stray light, but this may change with injection moulded iSPEX 2 units or on certain devices.

The corrected image is split into the $\pm Q$ component spectra, currently based on hard-coded windows for specific cameras but in the future automatically, and demosaicked. The $RGBG_2$ channels of the Bayer-filter camera are treated separately rather than combined through interpolation, since interpolated data add no extra information [281]. After demosaicking, there are eight separate spectra, namely the combinations of $\pm Q$ and $RGBG_2$. Figure 6.8 shows two examples of demosaicked RGB spectra. Finally, each row is convolved with a Gaussian kernel ($\sigma = 6$ pixels) corresponding to an FWHM of ~3.8 nm (see Section 6.5.3). This reduces the noise on the spectrum without reducing the spectral resolution, being narrower than the image of the slit.

6.5.3 Wavelength calibration

The pre-processed data are then wavelength-calibrated. As seen in Figure 6.7, iSPEX 2 data exhibit significant smile (variations in dispersion along the slit) and keystone (deformation of the spectrum into a trapezoid). Both effects are common in long-slit spectrometers [177,464]. Smile is corrected by doing the wavelength calibration per pixel row; a keystone correction is still in development.

The wavelength calibration is done using a reference spectrum of a fluorescent light, like that shown in Figure 6.7a. These have three sharp spectral lines corresponding to the $RGBG_2$ channels, at 611.6 (R), 544.5 (G/G_2), and 436.6 (B) nm. Because of mosaicking, the raw data have R and B values only in every alternate row and column, while combining G and G_2 gives full row coverage but in alternating columns. The column gaps are filled in by the Gaussian kernel convolution (Section 6.5.2), which also reduces noise. The maximum value per channel in each pixel row is determined to find the line centers. A quadratic fit is made to these, and the resulting fitted line centers are used. This fills in the gaps in R and B and reduces the effect of noisy rows. Figure 6.9 shows the line centers thus derived from the spectrum in Figure 6.7a.

A wavelength solution map, with the central wavelength for each pixel, is generated by fitting a quadratic relation between the spectral line wavelengths and the line centers from Figure 6.9. The resulting dispersion is typically ~0.27 nm/px, depending on the camera optics, exposure settings, and pixel position (as seen in Figure 6.9). For the images shown in Figure 6.7, where the slit is ~35 pixels wide (FWHM), this gives a spectral resolution (FWHM) of 9 nm. The FWHM varies slightly based on camera optics and exposure settings, mainly focus. 9 nm resolution is comparable to common ocean colour sensors such as the TriOS RAMSES and HyperOCR [283] and satellite instruments like Sentinel-3/OLCI [283] and HARP-2 [86], and only 2–3 times wider than PACE/OCI (5 nm) [86] and SPEXone (2–3 nm) [86, 184].

The overall wavelength map is converted into individual wavelength maps for the $RGBG_2$ channels by demosaicking it, as if it were an image itself. Finally, all rows are interpolated to the same 390–700 nm range in 1 nm steps, giving wavelength-calibrated spectra as shown in Figure 6.10.

[20] https://github.com/burggraaff/ispex2

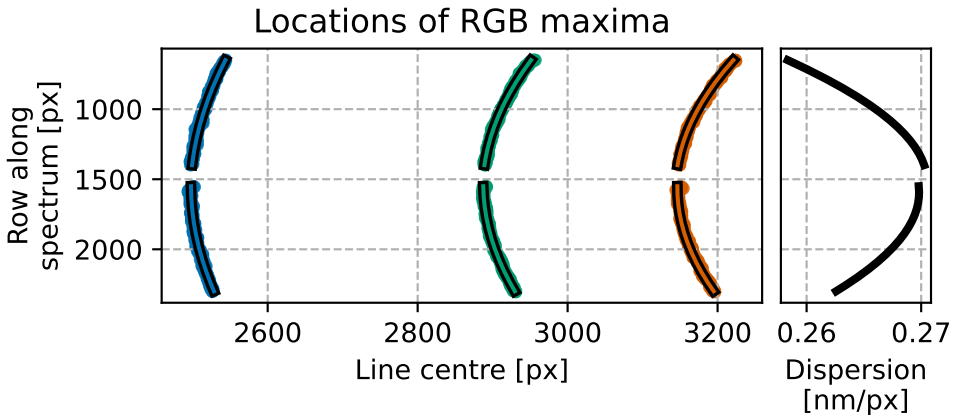


Figure 6.9: Pixel positions of the spectral lines in Figure 6.7a (left) and the derived dispersion (right). The coloured dots indicate the maxima in the B, G/G_2, R (from left to right) channels, while the coloured lines with a black outline indicate the fitted positions. The fits are done separately for the $\pm Q$ spectra. The smile effect is clearly visible. The mean dispersion between the B (436.6 nm) and R (611.6 nm) lines is shown; a smaller value in nm/px corresponds to a wider dispersion.

Since fluorescent lights are less ubiquitous than in the past, and will likely be fully replaced by LEDs in the foreseeable future, an alternative method may become necessary. For example, common features of SRFs such as the ~580 nm edge in R bands [281] may be used instead of spectral lines.

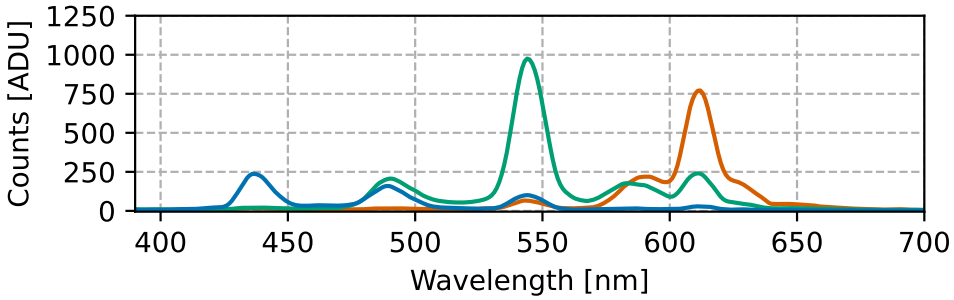
6.5.4 Spectral response calibration

The wavelength-calibrated spectra are corrected for the RGBG_2 spectral response functions (SRFs), again using SPECTACLE [281]. The SRFs are interpolated to the same wavelengths as the data, after which the data are divided by the SRF. This gives radiances, a constant factor away from absolute radiometric units [281], as shown in Figure 6.11. To prevent amplifying noise and stray light, currently only wavelengths where the SRFs are >0.15 (in relative units) are used. This restriction will be relaxed with better stray light reduction and correction. Even so, both spectra in Figure 6.11 show excellent agreement between the RGB radiances. For fully unpolarised light measurements, this is the final calibration step.

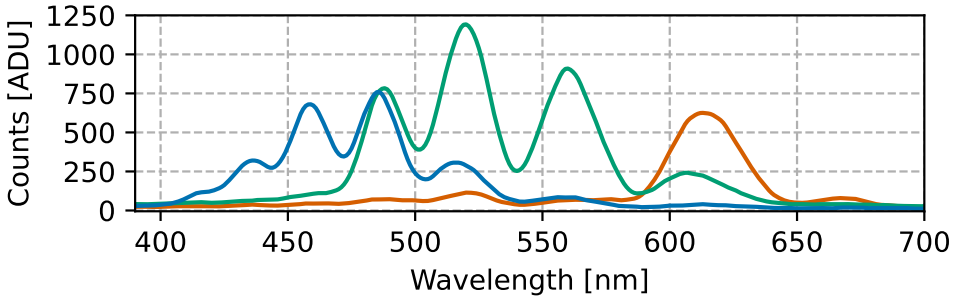
For cameras without SPECTACLE SRFs, iSPEX 2 itself can be used to measure these using a known light source. This has been attempted with the original iSPEX using skylight reflected off white paper, with an RMS error of 5% compared to reference data, increasing towards longer wavelengths [281]. The skylight spectrum was model-based, introducing assumptions that cannot be tested in the same measurement. For this reason, it may be preferable to instead use consumer lamps, such as those characterised in the LICA database [381].

6.5.5 Polarimetric demodulation

Finally, the DoLP and AoLP are retrieved by inverting Equation (6.4) to demodulate the calibrated spectra. First, the retardance and polarimetric efficiency of the instrument must be



(a) Fluorescent light.



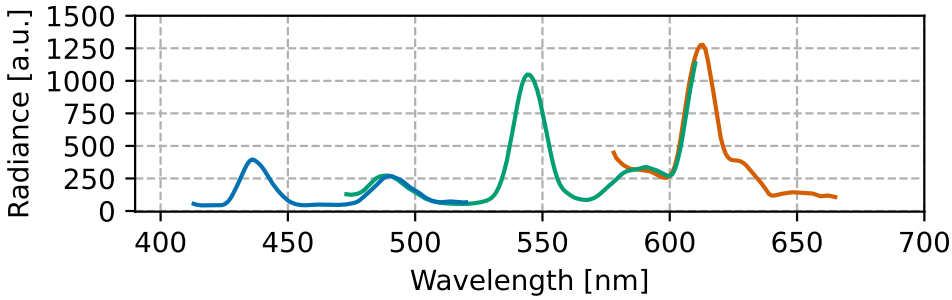
(b) Reflected sunlight through a 100% polariser.

Figure 6.10: Spectra from Figure 6.8 after wavelength calibration.

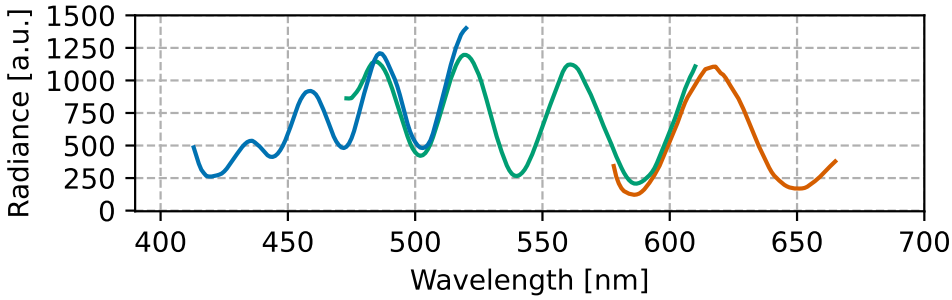
calibrated using a known 100% polarised light source [177]. For iSPEX 2, we plan to do this upon assembly using a rotating wiregrid polariser.

As can be seen in Figure 6.12, variations in retardance and efficiency exist along the $\pm Q$ slits of 3D-printed prototypes. These are partially due to fabrication issues, such as MOR foils flexing in their bracket due to wide production tolerances (Section 6.4.3), which will be resolved in the injection moulded product. However, they are also partially due to issues including nonnormal incidence, since the optical path length and refractive index vary with the direction of propagation [465]; this will persist in the final product, necessitating a spatially dependent calibration of retardance and efficiency. We are currently characterising these effects. A final complication is the fact that different sections of the slits see different targets, meaning inherent variations in DoLP and AoLP exist. The demodulation pipeline will have to account for this too.

The demodulation algorithm is in development. An iterative approach is likely necessary, fitting not only DoLP, AoLP, and the unpolarised spectrum $I_{in}(\lambda)$, but also instrumental effects including the QWP and MOR retardance, alignment and orientation of foils, the relative transmission between the two slits, and variations in the source spectrum along the slits [125, 177]. The derived DoLP, AoLP, and $I_{in}(\lambda)$ are used to determine AOT for aerosol and R_{rs} for ocean colour measurements. The typical spectral resolution in DoLP and AoLP is approximately the modulation period [175] of 25–60 nm, though this can be lowered to the native spectral resolution of ~9 nm (Section 6.5.3), albeit with a lower accuracy [176]. The



(a) Fluorescent light.



(b) Reflected sunlight through a 100% polariser.

Figure 6.11: Spectra from Figure 6.10 after spectral response calibration.

original iSPEX had a polarimetric accuracy (typical uncertainty) of ~0.03 in DoLP, mostly limited by defocus due to lacking camera controls [94]. Having solved these problems [281], we hope to increase the accuracy to sub-percent levels which enable retrieval of parameters such as effective radii and refractive indices (real and imaginary) [466].

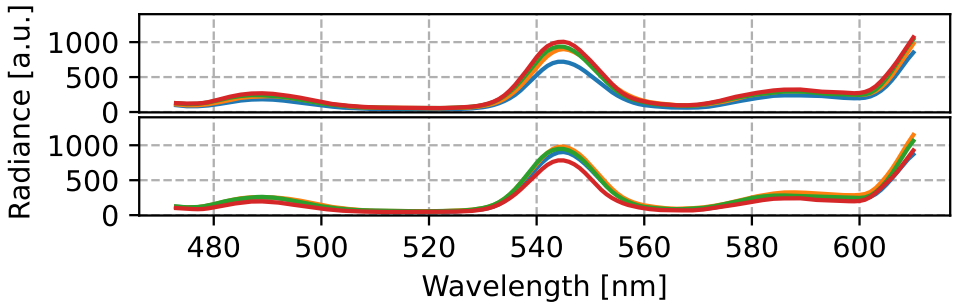
6.6 Future perspective

This section describes the ongoing calibration and validation efforts for iSPEX 2 as well as current and future scientific applications and opportunities.

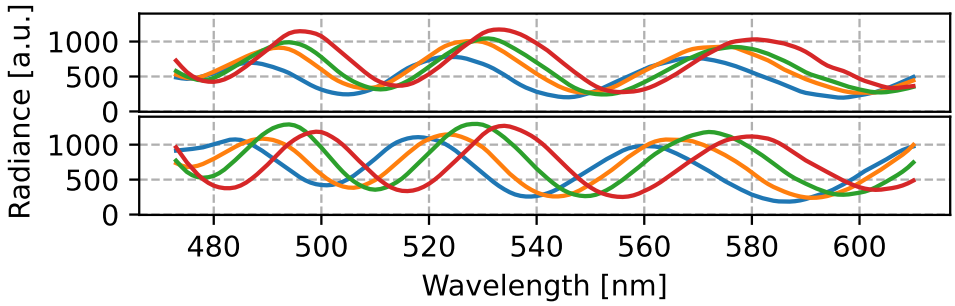
6.6.1 Calibration & validation

Each iSPEX 2 unit will be factory-calibrated for retardance and polarimetric efficiency with a 100% polariser (Section 6.5.5) and a small number will be validated in the lab at various DoLP and AoLP using a glass plate setup [457]. This will allow for a thorough comparison in performance between iSPEX 2 and other sensors, as well as between iSPEX 2 units and between smartphones. The calibration data for each iSPEX 2 unit will be linked to its serial number in a database, from which the app will retrieve them.

iSPEX 2 AOT and R_{rs} measurements are being validated through simultaneous observations with other instruments. For both, this will include groundSPEX, which is based



(a) Fluorescent light.



(b) Reflected sunlight through a 100% polariser.

Figure 6.12: Four rows in the G-band $+Q$ (top) and $-Q$ (bottom) spectra from Figure 6.7. Each coloured line corresponds to a row in the spectrum, though these are not the same between $+Q$ and $-Q$. Large variations are seen in the 100% polarised spectrum but not the unpolarised fluorescent light spectrum.

on the same principle but with a much higher spectral resolution and a polarimetric accuracy of ~1% [125]. AOT match-ups will also be done with MicroTOPS II, a handheld Sun photometer [122], and AERONET [134]. AERONET has previously been used to validate groundSPEX [125]. R_{rs} match-ups are being performed with WISP-3 handheld and TriOS RAMSES shipborne spectroradiometers, similar to the HydroColor app [121]. Validation has thus far taken place largely within field campaigns organised through the MONOCLE consortium[21]. Opportunities for further validation are under investigation.

An initial experiment comparing iSPEX 2 and WISP-3 R_{rs} measurements showed good agreement overall but also highlighted several areas of improvement (Figure 6.13). The spectra were similarly shaped, particularly the CDOM-phytoplankton absorption slope at 400–570 nm. The B- and G-band spectra showed residual modulation at $\lambda \leq 500$ nm, which will be reduced in the future through an improved demodulation algorithm. Future work on the SRF calibration (Section 6.5.4) will reduce edge effects such as that seen in the G-band at 470 nm.

[21] https://monocle-h2020.eu/

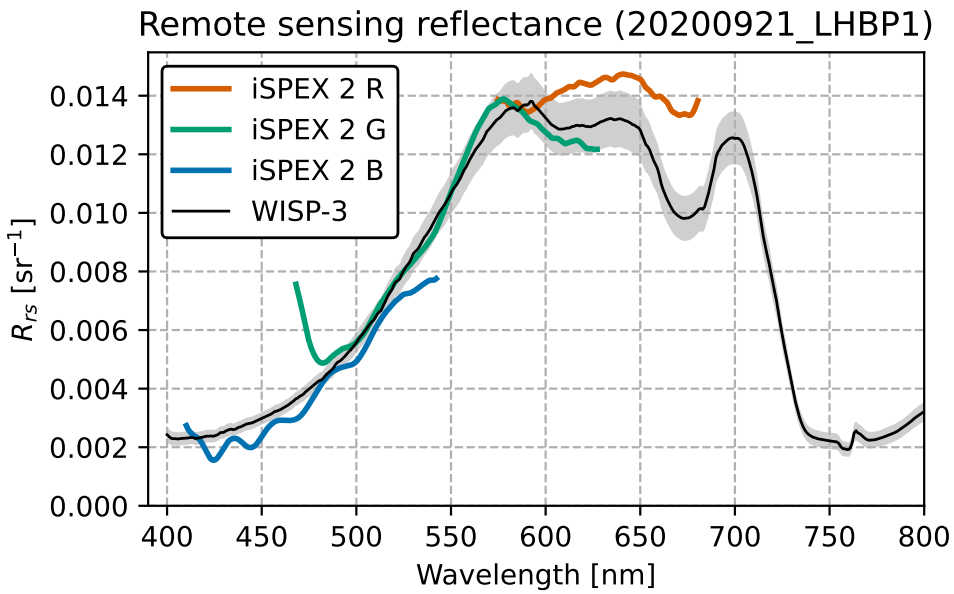


Figure 6.13: Comparison between near-simultaneous iSPEX 2 (red, green, blue) and WISP-3 (black) R_{rs} spectra. Data were gathered on 21 September 2020 at a pond in Leiden, the Netherlands (52°10'2.5"N 04°28'23.7"E). The iSPEX 2 spectra were normalised to the WISP-3 spectrum following [282]. The overall agreement between the two instruments is good, but there are some areas of improvement (Section 6.6.1).

6.6.2 Scientific applications

The main application of iSPEX 2 is as a low-cost instrument for citizen science remote sensing of air and water, specifically measuring AOT and R_{rs}. Both top-down and bottom-up approaches will be used for this. In the top-down approach, citizen scientists will be prompted by researchers to observe at a certain place or time, similar to the original iSPEX [94]. Conversely, in the bottom-up approach, citizen scientists can use iSPEX 2 independently, with researchers only providing support such as data processing and interpretation.

Planned top-down scientific applications of iSPEX 2 include high spatial resolution measurements of AOT and R_{rs}, extension of existing time series, and validation of satellite or airborne instruments. iSPEX 2 provides point measurements in arbitrary locations, facilitating extremely high spatial resolution. For example, a small group of citizen scientists standing along a lake shore can simultaneously map its reflectance (and thus its inherent properties) on meter scales. iSPEX 2 can also be used to fill in gaps in existing time series, for example if clouds prevented measurements during a satellite overpass. Finally, push notifications can be used to prompt citizen scientists to take validation measurements during a satellite overpass.

6.6.3 Future opportunities

In addition to the currently planned applications, more experimental work with iSPEX 2 is also possible. For example, while we are currently focused on unpolarised R_{rs}, polarised R_{rs} may provide additional information on water composition [174, 238]. However, an optimal

protocol for measuring polarised R_{rs} will need to be found. Additionally, extending the principal plane measurements to the solar aureole, almucantar, and horizon may improve the AOT data and particle size distributions [134,460,461]. Outside remote sensing, iSPEX 2 will also be useful as a low-cost device for lab or field-going spectroscopy, for biological assaying and point-of-care diagnostics, among other purposes [98, 143].

Non-smartphone platforms also provide interesting opportunities. Unmanned aerial vehicles (UAVs) and webcams like the Raspberry Pi have cameras capable of professional-grade radiometry [281]. Raspberry Pi-based systems could be used as low-cost autonomous field-going spectroradiometers. Already, UAVs with pushbroom spectrometers are delivering data products like R_{rs} with high spatial and spectral resolution in a single fly-over [136]. Using iSPEX 2, any camera can become a hyperspectral and polarimetric sensor.

Funding

European Commission Horizon 2020 programme (grant nr. 776480, MONOCLE).

Acknowledgements

The authors wish to thank Akupara and Sanjana Panchagnula for providing pandemic-proof laboratory space. Data analysis and visualisation were done using the AstroPy, Matplotlib, NumPy, and SciPy libraries for Python.

7 | General discussion and future outlook

The aim of this thesis has been to investigate and improve accessibility and uncertainty in remote sensing and citizen science, so that these techniques can better deliver the desired improvements to cost, scale, and reproducibility of water research (Section 1.1.4). Accessibility was improved by accounting for disability in citizen science (Chapter 2) and by developing consumer cameras as low-cost instruments for remote sensing by professional and citizen scientists (Chapters 3, 4, and 6). The uncertainty in measurements by citizen scientists and measurements from consumer cameras was assessed and reduced (Chapters 2, 3, and 4), as was the uncertainty in comparing and combining data from different professional instruments (Chapter 5). Here, a general synthesis, discussion, and future outlook are provided.

The current direction of research in remote sensing of water is towards higher-dimensional data by including a wider wavelength range (UV–NIR), finer spectral sampling, and polarisation; automation; and more in-depth analysis of uncertainty and information content.

Satellite remote sensing of water has come a long way since its origins in the 1970s, when Landsat-1 data were used to map ocean currents based on a dynamic range of only 4 ADU [96]. Successive generations of satellite instruments developed by different organisations, such as the NASA Satellite Ocean Colour programme and Landsat missions and the EU/ESA Copernicus programme [85, 467], continue to provide high-resolution, high-quality data. Planned to launch in 2024, the PACE mission will further expand satellite remote sensing of water with its hyperspectral and spectropolarimetric instruments OCI, SPEXone, and HARP-2. Combined observations from these instruments will enable joint aerosol-water reflectance retrieval algorithms, improving the accuracy of both aerosol and water constituent products. Furthermore, PACE will provide insights into aerosol-water interactions, which play a key role in global biogeochemical cycles and climate change [86]. Ongoing pre-launch research includes characterisation of its instruments and capabilities [111, 112, 184], generation of synthetic data [186], and development of algorithms for data processing and analysis [185, 188, 421].

Current in-situ instrumentation development is focused on automation of sensors and data processing to enable autonomous data collection on larger scales and with greater consistency [89, 109, 117, 124, 399]. This improves the cross-validation between satellite and in-situ radiometry and makes it possible to measure high-frequency local time series. Optical measurements are being further integrated with existing networks such as Argo [80], providing additional information and improving the accuracy of retrieval algorithms [216].

These new developments will be valuable tools for studying the changes occurring in global waters. Climate change is affecting phytoplankton community compositions [42, 232, 417] and the availability of habitats for aquatic animals [468]. Similar trends are caused by pollution with nutrients, causing eutrophication and algal blooms [47], and toxins, which kill organisms and damage ecosystems [69]. Equally important to study are the resulting feedback mechanisms, such as the role of phytoplankton in carbon capture and export [469, 470]. Increased remote sensing capabilities, in particular wide spatial coverage with high-frequency observations, combined with a general increase in adoption of remote sensing by water managers and policymakers, are necessary to meet data requirements for modelling and understanding the changes occurring in waters worldwide [89, 471].

This chapter places the research described in the previous chapters into the context discussed above and in Chapter 1. Comparisons between individual results and the existing

literature are presented in the discussion sections of the respective chapters, and will not be repeated here. Section 7.1 discusses our research into uncertainty characterisation and mitigation. Section 7.2 discusses accessibility and inclusion in remote sensing of water, including our studies of the Forel-Ule scale and smartphone cameras. In Section 7.3, we discuss the potential of spectropolarimetry in remote sensing of water and present initial results from a follow-up project on spectropolarimetric sensing of floating debris. Finally, Section 7.4 presents the general conclusions and recommendations to be drawn from this thesis.

7.1 Uncertainty in remote sensing of water

Formal analysis of uncertainty has historically been underutilised in the remote sensing community, to the extent that data and products were often distributed without any uncertainty estimate [88, 106]. Instead, field intercomparisons between different instruments, different algorithms, or between data and simulations have been the primary method for characterisation of uncertainty [88]. While validation with different instruments is an important step towards achieving closure (Sections 1.3.4 and 4.4.3), it is fundamentally limited. For example, data taken at slightly different times or locations and with different instruments or setups are never truly identical [409]. Uncertainty and error come from many sources, including instability and change in sensor responses, variability in targets due to waves and similar factors, shot noise, and uncertainty in sensor characterisation and calibration materials (Section 1.2.1). Advances in instrumentation, such as hyperspectral measurements or polarimetry, are pointless if measurement error and uncertainty are not appropriately characterised and minimised [144]. Improving our understanding of uncertainty is crucial to making optimal use of new platforms and technology [87, 238].

Currently, there is a strong push within the community to improve the understanding and analysis of uncertainty. For example, 2019 saw the release of IOCCG report 18, which reviewed the current state of affairs and provided numerous recommendations for future research [88]. The FRM4SOC project[22] improved the state of the art by standardising protocols, intercomparing commonly used spectroradiometers, and improving methods for uncertainty estimation and propagation [92, 97, 108, 110, 114, 120]. Additional recent research into uncertainty has included standardisation of robust comparison metrics [406, 407] and terminology [472], identification and quantification of individual contributors to the overall uncertainty budget [208, 218], and improvements to the visualisation and communication of uncertainty [278, 473]. The most important recommendation for the future is to treat uncertainty as an integral part of the measurement process, meaning uncertainty and error should be characterised as comprehensively as possible, reported as consistently as possible, and propagated as accurately as possible.

In Chapters 3 and 4, we investigated and improved the uncertainty associated with using consumer cameras as spectroradiometers. By characterising and calibrating the optical properties of several devices in the same way as professional sensors, we showed that consumer cameras can achieve an accuracy and uncertainty similar to those professional sensors, when RAW data are used. This was a major improvement on the state of the art, which until recently had been based on data in the inferior JPEG format. Our calibration methodology and data have already been adopted by other groups [412, 474, 475]. While the SPECTACLE database (Section 3.4.9) unfortunately has not materialised as intended, the associated Python

[22] Fiducial Reference Measurements for Satellite Ocean Colour, https://frm4soc.org/

package is in active development and has been published on the PyPI repository[23].

Since the original publication of Chapter 3 in 2019, the smartphone market has moved towards increasing the number of cameras per device, which further increases the amount of calibration data necessary. To facilitate collecting these data, future work will be needed to improve the accessibility of the SPECTACLE package itself, for example through a web interface. Future smartphone cameras will likely include more advanced features such as additional spectral bands and polarisation filters. The SPECTACLE methodology will need to be expanded to include calibration of the associated optical properties. Some aspects of calibration that were neglected in this thesis, such as temperature sensitivity, also require further investigation. Overall, however, we have clearly demonstrated that consumer cameras are valuable scientific instruments.

There is no single optimal method for uncertainty estimation or propagation, but any reasonable estimate is better than none. In Chapter 4, two methods for estimating radiometric uncertainty were compared, namely analytical propagation of the inter-pixel variability within one image vs. the variability in products derived from different images. The resulting uncertainty estimates were relatively close, but not identical. Notably, the analytical propagation accounted for uncertainties in calibration materials like the 18% reference grey card, which affect each measurement in the same way and thus cannot be determined from replicates. On the other hand, analytical propagation to R_{rs} is not exact (Section 4.A.3), resulting in an overestimation of the uncertainty in relative quantities like band ratios. Both methods are ultimately limited in their capacity to handle systematic errors, which are often difficult or impossible to determine and propagate statistically. These trade-offs apply to any data, and similar considerations apply to other methods for uncertainty propagation, such as Monte Carlo simulation and neural networks [88, 110]. Other considerations include the choice between absolute and relative uncertainty as well as the choice of uncertainty metric, such as coefficient of variation or interquartile range.

The importance of characterising the uncertainty in calibration materials like the 18% reference grey card extends beyond our measurements. For example, diffuse reflectance standards are often used in professional spectroradiometry to estimate E_d in the same way. The associated uncertainty has been shown to be 1%–6.5%, comparable to the overall uncertainty from other sources [435]. As recommended in Section 4.4.3, the impact of calibration material uncertainty could be reduced by characterising the materials on a large scale [400] or by issuing standard ones [394]. The same applies to the calibration methods proposed in Chapters 3 and 6.

Chapter 5 addressed a specific source of systematic error, namely incorrect spectral convolution of reflectance. The resulting error in R_{rs} was up to 5% for consumer cameras, ~1% for broad-band satellite sensors, and <1% for narrow-band satellite sensors. As the relative uncertainty in remote sensing measurements decreases below the current standards of ~5% in R_{rs} [86, 88], these systematic errors become more significant. This is especially true when they skew validation results. Spectral convolution of reflectance is now more commonly performed correctly [32, 109, 476], but other potential sources of systematic error remain to be investigated. For instance, hyperspectral data are often convolved to multispectral bands with a similar bandwidth, effectively convolving the input signal twice, while real narrow-band filters only convolve once. This discrepancy may affect the results of instrument validation studies. Future work should investigate this and other potential errors that are based on mathematical simplifications.

[23] https://pypi.org/project/pyspectacle/

The impact and communication of uncertainty in citizen science were explored in Chapter 2. Specifically, we investigated the effects of colour blindness on Forel-Ule (FU) measurements, combining current research on colour blindness in science [304] and on the FU scale in general [269,270]. Two forms of colour blindness, deuteranopia and tritanopia, were found to significantly increase the uncertainty in simulated FU measurements by decreasing the distinguishability of colour pairs. This previously neglected factor likely affects the data quality of FU measurements and their value in validation studies. The inferred implications for inclusivity and participant motivation are discussed in Section 7.2.

Citizen science data are rarely reported with uncertainties. Based on our results, we recommend that researchers incorporate uncertainty into citizen science by having participants estimate it themselves, which requires detailed instructions [278, 315], or through post-hoc analysis. Involving citizen scientists in this way may improve public understanding of scientific uncertainty in general. This, in turn, may improve decision-making and trust in science, which the COVID-19 pandemic has shown can be somewhat lacking [477,478].

Finally, our investigations into consumer cameras and spectropolarimetry present opportunities to decrease uncertainties in satellite and above-water remote sensing. The main source of uncertainty in satellite remote sensing of water is the atmospheric correction [208]. Spectropolarimetry provides greater information on atmospheric properties than spectroradiometry does, including aerosol particle properties, which can be used to improve the atmospheric correction. Joint retrieval algorithms for aerosol optical depth (AOD) and water-leaving radiance are in development, particularly focusing on the PACE mission [187]. The original iSPEX demonstrated the possibility for citizen scientists to measure AOD [94], and iSPEX 2 will improve the accuracy of AOD measurements through its dual-beam design and SPECTACLE-based data processing (Chapter 6). Citizen scientists could be asked through a push notification to measure AOD during a satellite overpass. Polarisation can also be used to characterise and reduce sun and sky glint [211, 240], which are major sources of uncertainty and error in above-water radiometry [120]. These possibilities can be explored with iSPEX 2 after its calibration and validation are complete (Chapter 6) or with similar instruments like groundSPEX [125]. Glint removal for spot spectroradiometers may also be improved through combined measurements with low-cost cameras, which can provide real-time wave statistics and thus improved estimates of the surface reflectivity (Chapter 4). Additionally, low-cost cameras can be deployed in the field to autonomously obtain long time series with a short cadence [479,480].

7.2 Accessibility of water research

Science benefits from being accessible to a wide audience and inclusive of a diverse group of researchers [481, 482]. Diversity of people provides diversity of ideas, interpretations, and applications. Improving equity, diversity, and inclusion (EDI) in science has an inherent social value and increases the quality and quantity of science and its impact on society [482]. While recent years have shown significant improvements to EDI in science, including in remote sensing, there is still a long road ahead [481]. The increased focus on EDI in science is part of a wider trend towards equity in society.

Accessibility is affected not only by social factors, but also by economics. As discussed in Chapter 1, research often requires expensive equipment and specialised training. Economic disparity means that those who are most affected by environmental changes and pollution are

often those with the least access to research.

Citizen science, the involvement of non-professionals in the scientific process, has experienced a boom in the last twenty years [243, 389]. This boom can largely be attributed to technological innovations such as increased internet and smartphone usage [389, 483]. Compared to traditional research conducted by professionals, citizen science provides greater data collection capabilities through crowdsourcing and increased social relevance through stakeholder participation and co-creation [95]. Citizen science is often touted as an example of inclusive science and as a method of empowerment for socially or economically disadvantaged people and nations.

However, like professional science, achieving real inclusion and equity in citizen science remains a challenge. The demographic imbalance seen in professional science is mirrored among citizen scientists, with the majority of participants belonging to socially privileged groups [484]. This imbalance is an unintentional result of the way citizen science is coordinated and used by professional researchers from a top-down perspective. In fact, the term *citizen* itself and the distinction from *professional science* influence its perception among the general public [485]. Translating the increased awareness of EDI into tangible improvements will require significant efforts in science communication, community engagement, research planning, and funding allocation [484–486]. However, the results will be worthwhile. Obtaining diverse data, for example spanning many different water bodies, requires diverse participants. Valorisation of scientific results is also improved by diversity among citizen scientists, since the participants become well-informed stakeholders, who can translate scientific knowledge into social action [294, 487]. In addition to EDI issues, professionals working with citizen scientists also need to be more aware and considerate of the citizens' desires and well-being [247, 488].

In Chapter 2, we investigated the impact of disability, specifically colour blindness, on inclusion and motivation in citizen science. Colour blindness was found to reduce the data quality in Forel-Ule (FU) measurements. Based on previous work and personal experience, we inferred that participants would be demotivated by the increased difficulty of measuring and decreased quality of results, leading to a decrease in engagement and thus inclusion. Since the FU scale represents the true colours of natural waters, it cannot be changed, and we instead made recommendations regarding data entry forms, manuals, and communication. For example, allowing participants to enter a range instead of a single value could largely mitigate the problems associated with colour blindness. Our research has led to an increased awareness of colour blindness in citizen science and the development of more inclusive techniques [489]. Future work should investigate different forms of disability and strive towards a general understanding and inclusion of disabled people in citizen science. This would likely be achieved by involving disabled people in the design of measurement protocols and manuals in a form of co-creation. In general, the quality of training materials is improved by involving participants in their development. It is often impossible for professional scientists to envision all possible problems, questions, and even unintended use cases that arise when citizens use their equipment [256, 297, 490].

In Chapters 3, 4, and 6, we investigated the use of smartphone cameras as low-cost remote sensing instruments. As discussed in Section 7.1, the quality of smartphone radiometric data was improved to a level comparable to professional sensors. This research improved upon the existing iSPEX, HydroColor, and EyeOnWater apps [94, 121, 274]. These apps have been used by professional scientists in lieu of more expensive equipment and by thousands of citizen scientists [318, 392, 393, 491]. This way, they have made remote sensing of water

accessible to new audiences.

One of our aims in improving the data quality was to enable more independent research by citizens. For example, the original iSPEX could only obtain reliable data when multiple people nearby measured at the same time, during a top-down campaign [94]. The results of Chapter 4 suggest that with our improved methodology, individual citizen scientists will now be able to observe when and where they want. Additional validation for iSPEX 2 (Section 6.6.1) is ongoing, having been delayed by the COVID-19 pandemic, but we expect comparable results.

A side effect of using RAW data is the exclusion of some users whose smartphone cameras do not support RAW photography. Several potential users of iSPEX 2 have indicated that this requirement prohibits them from using the app. Unfortunately, it is up to smartphone manufacturers to enable this functionality and until that happens, these users are excluded. We have decided to fully exclude these devices rather than offer a lower-quality JPEG-based version of the app, to avoid confusion and lower-quality data. Fortunately, increased consumer demand means RAW photography is now available on all new iOS and most new Android devices, in all price ranges, so this limitation is quickly disappearing.

Smartphone science and smartphone spectroscopy are being used by thousands of people, but this is still only a minute fraction of the global population. The literature is rife with examples, proofs-of-concept, and potential use cases, and even with reviews thereof [98,143,273,345–351,492–494]. This thesis itself provides several examples. However, none have made the step towards adoption by millions of users in their daily lives. This discrepancy can be attributed to a simple lack of demand. Demonstrating a scientifically interesting use case is not enough to generate commercial interest and investment. For smartphone science and spectroscopy to progress beyond the proof-of-concept stage, a *killer app* is necessary, an application so lucrative that investment and social interest follow naturally [490]. Scientific use cases for iSPEX 2 are discussed in Section 6.6.2, focused on remote sensing of air and water. While low-cost in-situ electrochemical sensors are the norm for citizen science of aerosols [495, 496], iSPEX 2 has the advantage of being more directly comparable and complementary to professional measurements from satellites and AERONET. Combining both types of low-cost sensor delivers the best of both worlds. In principle, iSPEX 2 can be applied to any field where visible-light spectroradiometry or spectropolarimetry is used. Potential commercial use cases include characterisation of paint colour, electric lights, and food freshness, and detection of contamination and health issues like skin cancer [98, 346, 349, 490]. Future work is necessary to develop these use cases and, by demonstrating the quality and value of smartphone measurements, identify the killer app.

7.3 Spectropolarimetry of floating debris

Spectropolarimetry plays a prominent role in the future outlook for remote sensing of water. In an interesting parallel with smartphone science (Section 7.2), despite many scientifically interesting proof-of-concept studies, polarimetry has not yet been embraced by the wider community [165]. The upcoming PACE mission, with its SPEXone and HARP-2 instruments, will provide multiangular hyperspectral polarimetry with global coverage [86], offering many new opportunities for research [238]. Initially, these will be focused on aerosol and climate science and on reducing the uncertainty in atmospheric correction algorithms [185,188,421]. Specific areas of interest for water research include the detection and characterisation of coc-

colithophores [21], oil spills [497], suspended particles [163, 164], and wind and wave conditions [240]. Spectropolarimetry also provides new opportunities for reducing uncertainties from glint and atmospheric correction (Section 7.1). These use cases also apply to other satellite instruments, airborne sensors, and terrestrial sensors including iSPEX 2. Scientific use cases for spectropolarimetry are discussed further in Chapter 6 and in Sections 1.2.3 and 1.3.3.

This section contains the initial results from a study on spectropolarimetry of floating debris. This work was done as part of the OP^3 project[24] funded by ESA and led by Shungu Garaba, Tristan Harmel, and Paolo Corradi. Within OP^3, we are investigating the spectropolarimetric properties of light scattered by micro- and macroplastics and other floating debris through radiative transfer simulations [498] and laboratory experiments [476]. The aim of the project is to determine the value of visible-light spectropolarimetry in a marine debris observing system [75].

Measurements were conducted during an ESA campaign in the Deltares Atlantic Basin[25] from 24 January–4 February 2022. The basin is 75 m long, 8.7 m wide, and 1.3 m deep, with a maximum water depth of 1.0 m. Its bottom is made of grey concrete and partially covered in sand. During the campaign, the basin was filled with clear, unaltered tap water. It features a wave cradle capable of producing waves with a realistic spectrum, similar to natural waves in the North Sea. Various types of macroplastics and other debris, such as plywood and rope, were manually thrown into the basin and studied as they floated on the water surface.

Two instruments were used, namely the groundSPEX spectropolarimeter[26] and a FLIR BlackFly BFS-U3-51S5PC-C RGB polarisation camera. GroundSPEX performs snapshot hyperspectral measurements of radiance and polarisation within its 0.9°-diameter field of view through the SPEX technique [175], further described in Section 6.2. Its original calibration in 2011–2014 is described in [125] and laboratory tests indicated that these calibration data were still valid. A Python implementation[27] of the original data processing and demodulation pipeline was used. The BlackFly camera is based on a Sony IMX 250 MYR sensor with a double Bayer pattern (Section 1.2.2), consisting of a layer of RGB filters and a layer of polarisation filters (0°, 45°, 90°, and 135°). The calibration of a different camera based on the monochromatic version of this sensor is described in [499]. The BlackFly data were processed using the Polanalyser package[28]. The BlackFly camera was mounted on top of groundSPEX so that they were roughly aligned. Observations were done at a nadir angle of 40°, following the standard protocol for above-water radiometry [120, 209]. The basin was illuminated by dozens of fluorescent lights spread across the ceiling, which for safety reasons could only be turned off during one measurement session. For additional light, a halogen lamp was positioned next to the instruments and angled toward their field of view, with an effective azimuth angle of ~90° and zenith angle of ~35°.

Initial results from the groundSPEX measurements (Figure 7.1) showed a moderate degree of linear polarisation ($0.05 \leq P_L \leq 0.15$) for several types of plastic debris. In Figures 7.1b–7.1d, P_L appears to increase as the total radiance decreases at $\lambda > 700$ nm, which is similar to the Umov effect [500, 501]. The high values of P_L seen at $\lambda > 710$ nm in Fig-

[24] Ocean Plastics Polarization Properties, funded through the Discovery Element of the European Space Agency's Basic Activities contract no. 4000132037/20/NL/GLC, https://uol.de/en/icbm/marine-sensor-systems/current-projects/ocean-plastics-polarization-properties-op3

[25] https://www.deltares.nl/en/facilities/atlantic-basin-3/

[26] Courtesy of RIVM.

[27] https://github.com/burggraaff/SPEX

[28] https://github.com/elerac/polanalyser

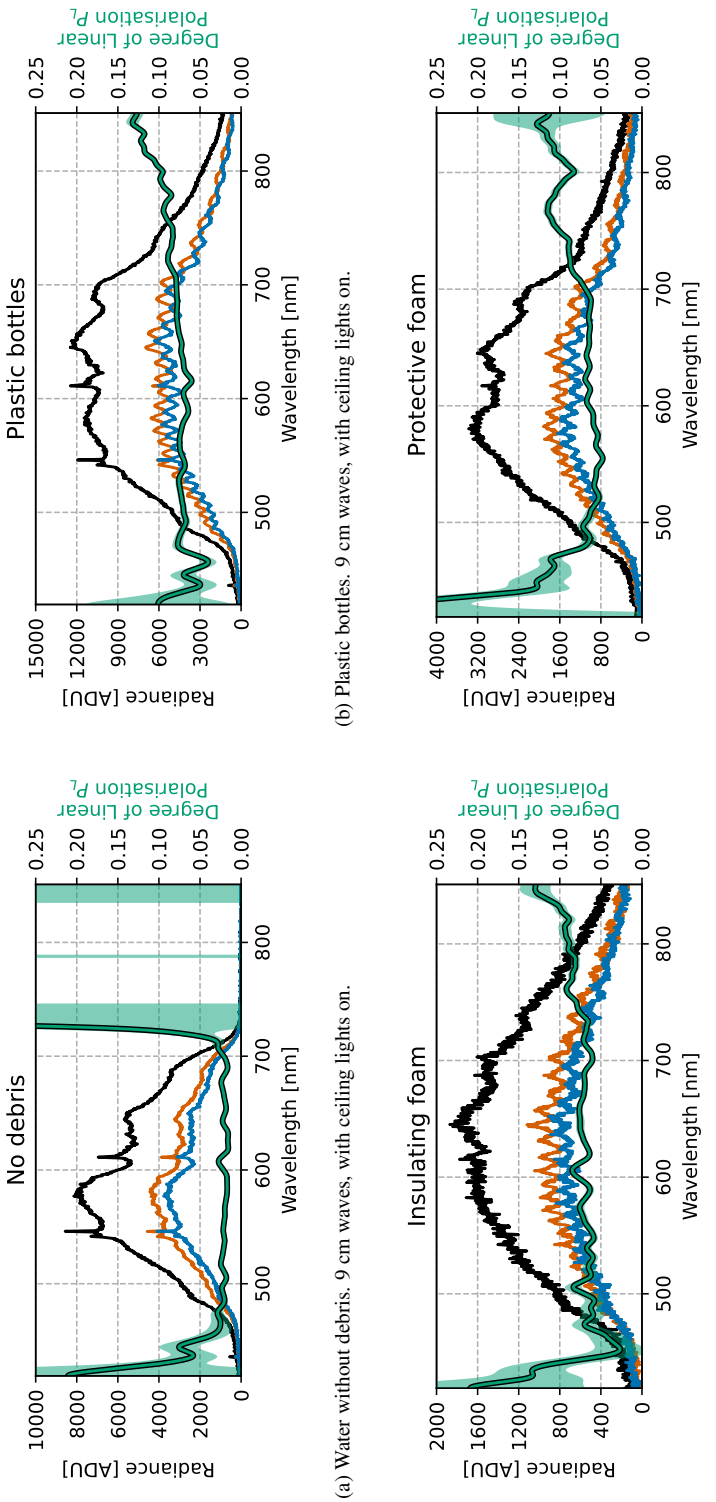
ure 7.1a are the result of measurement noise. P_L was otherwise approximately spectrally flat for virtually all types of debris investigated. Because groundSPEX is a spot radiometer without imaging capabilities, it is difficult to determine the mechanism causing the observed polarisation. The line features seen in Figures 7.1a and 7.1b were caused by the fluorescent ceiling lights and are present in most data sets. The offset between the two groundSPEX channels seen in each panel, most clearly in Figure 7.1a, is due to a difference in transmission between the two channels and is corrected in the demodulation algorithm, although future work is necessary to improve the accuracy of this correction.

Similar trends were found with the BlackFly camera. RGB polarimetric images of the protective foam sample taken roughly simultaneously with the groundSPEX observations (Figure 7.1d) showed a similar degree of polarisation, namely $0.0 \leq P_L \leq 0.2$ (Figure 7.2). Comparing the RGB photograph and G-band P_L image suggests two primary mechanisms for polarisation. First, as the foam floated near the water surface, small puddles formed on top of it, which can be identified in both images. Second, the foam was slightly crinkled, resulting in diagonal lines that can again be identified in both images. Thus, both the plastic itself and its interactions with the water surface appear to be sources of polarised light. Both mechanisms were also seen in images of plastics fully submerged just under the water surface or floating entirely on top of it. Based on calibration data, it will be possible to match the field of view of groundSPEX with specific pixels in the BlackFly images to further investigate the sources of polarisation.

The preliminary conclusion from this measurement campaign is that many types of floating debris impart a measurable polarisation ($0.05 \leq P_L \leq 0.20$) on reflected light, making polarimetry a useful addition to a marine observing system. Both debris itself and its interactions with the water surface seem to induce polarisation. The observed P_L represents a situation where the sensor field of view is entirely filled with debris, which is not realistic for satellite sensors. Even in garbage patches, floating debris covers a small fraction of the 10 m–10 km satellite pixel footprints. It is unlikely that spectropolarimetric sensors on satellites will be able to distinguish the polarised reflectance of debris from other constituents, bubbles and whitecaps, specular reflections, and atmospheric signals. Simulations suggest that small microplastics are more suitable for satellite detection than macroplastics [498]. This limitation does not apply to air- and shipborne sensors, which have much smaller pixel footprints that can realistically be fully covered by a piece of debris. Thus, we suggest that polarimetry may be used to aid in the detection of floating debris from airborne platforms like UAVs and from ships, in particular to better distinguish between debris and water when the two are similarly bright, and to distinguish between types of debris. Since P_L appears to be largely spectrally flat, hyperspectral measurements are not necessary and an RGB camera can be used.

Future work on this experiment will include investigating all collected data in more detail, improving the data processing pipeline, and more precisely determining the uncertainties on the data and results. Future experiments should focus on adding other constituents such as phytoplankton, CDOM, suspended minerals, and microplastics to determine the contribution of debris to the overall reflectance in realistic settings.

Lastly, we have developed a goniometer setup for measuring the bidirectional polarised reflectance distribution functions (BPDFs) of micro- and macroplastics in a laboratory setting (Figure 7.3). Using groundSPEX, samples are observed at four instrument elevation angles, corresponding to the viewing angles of SPEXone and the Mobley protocol [86, 209]. A laser-driven light source provides broad-spectrum light at arbitrary azimuth and elevation



(a) Water without debris. 9 cm waves, with ceiling lights on.

(b) Plastic bottles. 9 cm waves, with ceiling lights on.

(c) Dark grey insulating foam. 9 cm waves, with ceiling lights off.

(d) Transparent protective foam. 9 cm waves, with ceiling lights off.

Figure 7.1: Upwelling radiance spectra of four targets observed with groundSPEX. The calibrated data from the $\pm Q$ channels are shown in blue and red, the total radiance in black, and P_L in green. The shaded green area shows the uncertainty in P_L, estimated by the demodulation fitting routine.

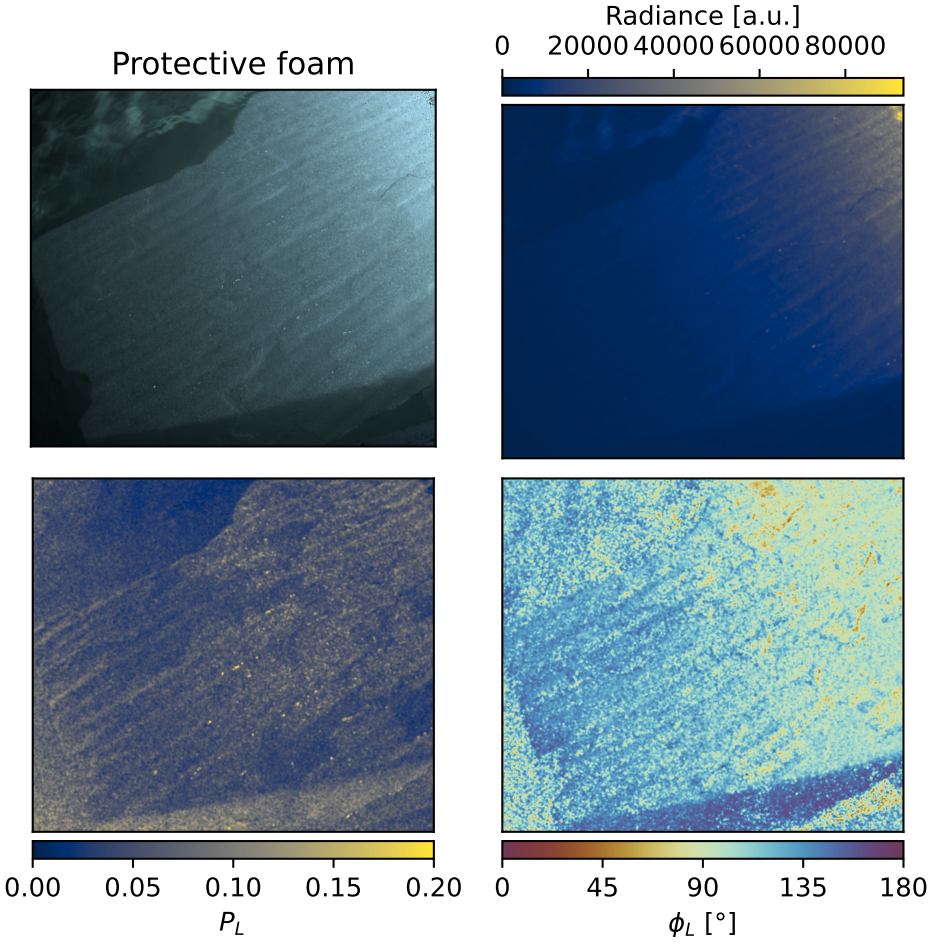


Figure 7.2: BlackFly observations of floating protective foam. The data were taken during the same session as Figure 7.1d, but not simultaneously. The RGB photograph (top left) is not white balanced and is slightly bluer than true colour. The G-band radiance (top right), P_L (bottom left), and ϕ_L (bottom right) images were convolved with a two-dimensional Gaussian with $\sigma = 3$ pixels to reduce noise. The zero-point on ϕ_L is arbitrary.

angles. In the future, polarised input light may be used to measure the full Mueller matrix (Section 1.2.3). The primary aims of this experiment are to validate BPDF simulations [498] and to provide the community with additional BPDF data for a wide variety of samples. Similar work has recently been done on characterising the effects of turbidity and salinity on the BPDF of water [502, 503] and on determining the BPDFs of various types of land cover and vegetation [166, 504, 505]. These data are valuable inputs for atmospheric correction algorithms, (exo)planetary atmosphere models, and vegetation reflectance models [166, 188, 506]. The goniometer setup has been built and is currently being commissioned.

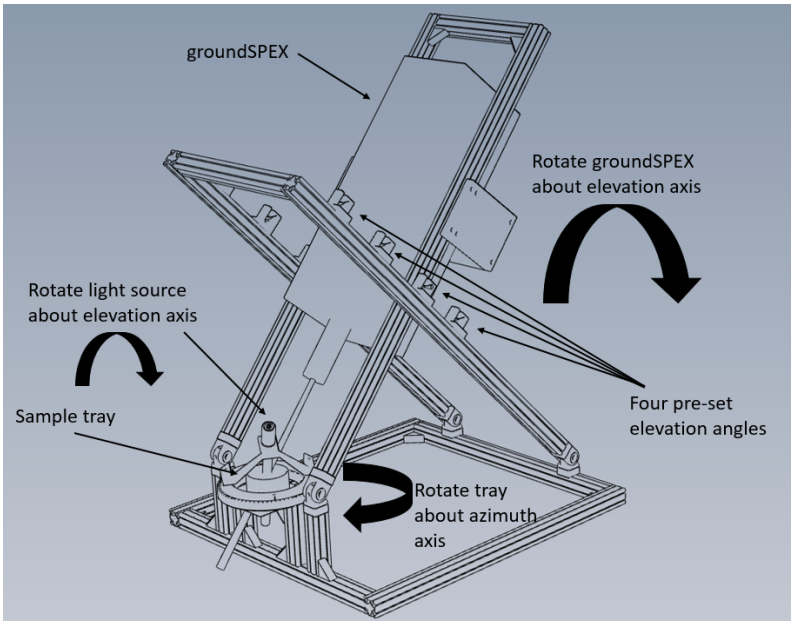


Figure 7.3: Render of the laboratory setup for measuring BPDFs with groundSPEX. The most important elements and degrees of freedom are indicated. Image courtesy of Remko Stuik.

Figure 7.4: Sunset over the Damariscotta River. Photograph taken at the Darling Marine Center, Walpole, Maine, USA, on 9 June 2019.

7.4 Conclusions and recommendations

This section provides a summary of the primary conclusions and recommendations of this thesis. A full summary of the thesis, in Dutch and in English, is provided after the bibliography.

Conclusions

1. Colour blindness significantly increases the uncertainty on Forel-Ule colour measurements for a significant fraction of users (Chapter 2).
2. Consumer cameras, including smartphone cameras, can perform professional-grade spectroradiometry when using RAW data (Chapters 3 and 4).
3. Spectral convolution of hyperspectral reflectance is often performed incorrectly, causing significant systematic errors (Chapter 5).
4. The iSPEX 2 add-on enables accurate spectropolarimetry using smartphone cameras (Chapter 6).

Recommendations

1. Vague terms like *water quality* should be replaced with specific quantities like constituent concentrations and inherent optical properties (Chapter 1).
2. Results should always be reported with an uncertainty estimate (Chapters 2, 3, 4, 5, and 7).
3. To ensure reproducibility and facilitate novel research, data should be published in full, including raw data and calibration materials (Chapters 2, 4, and 5).
4. To improve accessibility and data quality, citizen science protocols should be co-created with a diverse group of participants, including people with disabilities (Chapters 2 and 7).
5. To ensure consistency and reproducibility, instruments should be calibrated and characterised using standardised methods, and calibration data should be published (Chapters 3 and 4).
6. To reduce measurement uncertainties, calibration materials should themselves be calibrated thoroughly and regularly (Chapter 4).
7. To achieve a realistic uncertainty estimate, multiple methods should be compared, such as replicate observations and analytical propagation (Chapters 4 and 7).
8. To reduce systematic errors, assumptions about the accuracy of approximations and mathematical methods should always be challenged or justified (Chapter 5).
9. To ensure future compatibility and optimal accessibility, smartphone science add-ons should be designed as universally as possible (Chapter 6).
10. To maximise adoption and impact, citizen science tools should be designed with both top-down and bottom-up research in mind (Chapters 6 and 7).

Bibliography

[1] United Nations. *The United Nations world water development report 2021: Valuing water*. UNESCO, Paris, France, 2021. ISBN 978-92-3-100434-6.

[2] IOCCG. *Why Ocean Colour? The Societal Benefits of Ocean-Colour Technology*, volume 7 of *Reports of the International Ocean Colour Coordinating Group*. International Ocean Colour Coordinating Group (IOCCG), Dartmouth, Canada, 2008. ISBN 978-1-896246-56-7. doi:10.25607/OBP-97.

[3] E. W. Schwieterman, N. Y. Kiang, M. N. Parenteau, C. E. Harman, S. DasSarma, et al. *Exoplanet Biosignatures: A Review of Remotely Detectable Signs of Life*. Astrobiology 18(6), 663–708 (2018). doi:10.1089/ast.2017.1729.

[4] Environmental Audit Committee. *Water quality in rivers*. Technical report, House of Commons, London, United Kingdom (2022).

[5] F. van Gaalen, L. Osté, and E. van Boekel. *Nationale Analyse Waterkwaliteit*. Technical report, Planbureau voor de Leefomgeving, Den Haag, Netherlands (2020). doi:10.13140/RG.2.2.22104.11528.

[6] S. Wuijts, P. P. J. Driessen, and H. F. M. W. van Rijswick. *Towards more effective water quality governance: A review of social-economic, legal and ecological perspectives and their interactions*. Sustainability 10(4), 914 (2018). doi:10.3390/su10040914.

[7] STOWA. *Inventarisatie kennisbehoefte waterkwaliteit*. Technical report, STOWA, Amersfoort, Netherlands (2017).

[8] European Parliament and Council of the European Union. *Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy*. Official Journal of the European Communities L(269), 1–73 (2000).

[9] D. G. Smith and R. J. Davies-Colley. *Perception of Water Clarity and Colour in Terms of Suitability for Recreational Use*. Journal of Environmental Management 36(3), 225–235 (1992). doi:10.1016/S0301-4797(05)80136-7.

[10] J. Pitarch, H. J. van der Woerd, R. J. W. Brewin, and O. Zielinski. *Optical properties of Forel-Ule water types deduced from 15 years of global satellite ocean color observations*. Remote Sensing of Environment 231, 111 249 (2019). doi:10.1016/j.rse.2019.111249.

[11] J. M. Bogert, J. Ellers, S. Lewandowsky, M. M. Balgopal, and J. A. Harvey. *Reviewing the relationship between neoliberal societies and nature: implications of the industrialized dominant social paradigm for a sustainable future*. Ecology and Society 27(2), 7 (2022). doi:10.5751/ES-13134-270207.

[12] United Nations. *A/RES/70/1 – Transforming our world: the 2030 Agenda for Sustainable Development* (2015). https://sdgs.un.org/2030agenda.

[13] D. Blondeau-Patissier, J. F. R. Gower, A. G. Dekker, S. R. Phinn, and V. E. Brando. *A review of ocean color remote sensing methods and statistical techniques for the detection, mapping and analysis of phytoplankton blooms in coastal and open oceans*. Progress in Oceanography 123, 123–144 (2014). doi:10.1016/j.pocean.2013.12.008.

[14] C. D. Mobley. *The Oceanic Optics Book*. International Ocean Colour Coordinating Group (IOCCG), Dartmouth, Canada, 2022. doi:10.25607/OBP-1710.

[15] P. J. Durack. *Ocean Salinity and the Global Water Cycle*. Oceanography 28(1), 20–31 (2015). doi:10.5670/oceanog.2015.03.

[16] F. J. Millero, R. Feistel, D. G. Wright, and T. J. McDougall. *The composition of Standard Seawater and the definition of the Reference-Composition Salinity Scale*. Deep-Sea Research Part I: Oceanographic Research Papers 55(1), 50–72 (2008). doi:10.1016/j.dsr.2007.10.001.

[17] P. L. Munday, D. L. Dixson, J. M. Donelson, G. P. Jones, M. S. Pratchett, et al. *Ocean acidification impairs olfactory discrimination and homing ability of a marine fish.* Proceedings of the National Academy of Sciences 106(6), 1848–1852 (2009). doi:10.1073/pnas.0809996106.

[18] S. B. Tavakoly Sany, R. Hashim, M. Rezayi, A. Salleh, and O. Safari. *A review of strategies to monitor water and sediment quality for a sustainability assessment of marine environment*. Environmental Science and Pollution Research 21(2), 813–833 (2014). doi:10.1007/s11356-013-2217-5.

[19] R. W. Dunford, D. N. M. Donoghue, and T. P. Burt. *Forest land cover continues to exacerbate freshwater acidification despite decline in sulphate emissions*. Environmental Pollution 167, 58–69 (2012). doi:10.1016/j.envpol.2012.03.022.

[20] J. Salisbury, D. Vandemark, B. Jönsson, W. Balch, S. Chakraborty, et al. *How Can Present and Future Satellite Missions Support Scientific Studies that Address Ocean Acidification?* Oceanography 28(2), 108–121 (2015). doi:10.5670/oceanog.2015.35.

[21] W. M. Balch. *The ecology, biogeochemistry, and optical properties of coccolithophores*. Annual Review of Marine Science 10, 71–98 (2018). doi:10.1146/annurev-marine-121916-063319.

[22] M. J. Behrenfeld, T. K. Westberry, E. S. Boss, R. T. O'Malley, D. A. Siegel, et al. *Satellite-detected fluorescence reveals global physiology of ocean phytoplankton*. Biogeosciences 6(5), 779–794 (2009). doi:10.5194/bg-6-779-2009.

[23] Q. Wei, R. Nagi, K. Sadeghi, S. Feng, E. Yan, et al. *Detection and Spatial Mapping of Mercury Contamination in Water Samples Using a Smart-Phone*. ACS Nano 8(2), 1121–1129 (2014). doi:10.1021/nn406571t.

[24] C. T. Reinhard and N. J. Planavsky. *The History of Ocean Oxygenation*. Annual Review of Marine Science 14, 331–353 (2022). doi:10.1146/annurev-marine-031721-104005.

[25] S. Li, J. Zhang, G. Mu, H. Ju, R. Wang, et al. *Spatiotemporal Characterization of Chromophoric Dissolved Organic Matter (CDOM) and CDOM-DOC Relationships for Highly Polluted Rivers*. Water 8(9), 399 (2016). doi:10.3390/w8090399.

[26] K. R. Matthews and N. H. Berg. *Rainbow trout responses to water temperature and dissolved oxygen stress in two southern California stream pools*. Journal of Fish Biology 50(1), 50–67 (1997). doi:10.1006/jfbi.1996.0274.

[27] C. A. L. Riddick. *Remote sensing and bio-geo-optical properties of turbid, productive inland waters: a case study of Lake Balaton*. Ph.D. thesis, University of Stirling (2016).

[28] J. S. Turner, J. L. Pretty, and A. M. P. McDonnell. *Marine particles in the Gulf of Alaska shelf system: Spatial patterns and size distributions from in situ optics*. Continental Shelf Research 145, 13–20 (2017). doi:10.1016/j.csr.2017.07.002.

[29] A. Hommersom, M. R. Wernand, S. Peters, and J. de Boer. *A review on substances and processes relevant for optical remote sensing of extremely turbid marine areas, with a focus on the Wadden Sea*. Helgoland Marine Research 64(2), 75–92 (2010). doi:10.1007/s10152-010-0191-6.

[30] F. P. Maciel, P. E. Santoro, and F. Pedocchi. *Spatio-temporal dynamics of the Río de la Plata turbidity front; combining remote sensing with in-situ measurements and numerical modeling*. Continental Shelf Research 213, 104 301 (2021). doi:10.1016/j.csr.2020.104301.

[31] I. Ogashawara. *Bibliometric Analysis of Remote Sensing of Inland Waters Publications from 1985 to 2020*. Geographies 1(3), 346–361 (2021). doi:10.3390/geographies1030019.

[32] J. Tavora, E. Boss, D. Doxaran, and P. Hill. *An Algorithm to Estimate Suspended Particulate Matter Concentrations and Associated Uncertainties from Remote Sensing Reflectance in Coastal Environments*. Remote Sensing 12(13), 2172 (2020). doi:10.3390/rs12132172.

[33] H. Burchard, H. M. Schuttelaars, and D. K. Ralston. *Sediment trapping in estuaries*. Annual Review of Marine Science 10, 371–395 (2018). doi:10.1146/annurev-marine-010816-060535.

[34] A. M. Hoguane, C. L. Green, D. G. Bowers, and S. Nordez. *A note on using a digital camera to measure suspended sediment load in Maputo Bay, Mozambique*. Remote Sensing Letters 3(3), 259–266 (2012). doi:10.1080/01431161.2011.566287.

[35] J. S. Turner, P. St-Laurent, M. A. Friedrichs, and C. T. Friedrichs. *Effects of reduced shoreline erosion on Chesapeake Bay water clarity*. Science of the Total Environment 769, 145 157 (2021). doi:10.1016/j.scitotenv.2021.145157.

[36] E. Knaeps, K. G. Ruddick, D. Doxaran, A. I. Dogliotti, B. Nechad, et al. *A SWIR based algorithm to retrieve total suspended matter in extremely turbid waters*. Remote Sensing of Environment 168, 66–79 (2015). doi:10.1016/j.rse.2015.06.022.

[37] E. Knaeps, D. Doxaran, A. I. Dogliotti, B. Nechad, K. Ruddick, et al. *The SeaSWIR dataset*. PANGAEA (2018). doi:10.1594/PANGAEA.886287.

[38] M. Jafar-Sidik, F. Gohin, D. Bowers, J. Howarth, and T. Hull. *The relationship between Suspended Particulate Matter and Turbidity at a mooring station in a coastal environment: consequences for satellite-derived products*. Oceanologia 59(3), 365–378 (2017). doi:10.1016/j.oceano.2017.04.003.

[39] E. Kari, S. Kratzer, J. M. Beltrán-Abaunza, E. T. Harvey, and D. Vaičiūtė. *Retrieval of suspended particulate matter from turbidity – model development, validation, and application to MERIS data over the Baltic Sea*. International Journal of Remote Sensing 38(7), 1983–2003 (2017). doi:10.1080/01431161.2016.1230289.

[40] I. Kostakis, M. Twardowski, C. Roesler, R. Röttgers, D. Stramski, et al. *Hyperspectral optical absorption closure experiment in complex coastal waters*. Limnology and Oceanography: Methods 19(9), 589–625 (2021). doi:10.1002/lom3.10447.

[41] IOCCG. *Earth Observations in Support of Global Water Quality Monitoring*, volume 17 of *Reports of the International Ocean Colour Coordinating Group*. International Ocean Colour Coordinating Group (IOCCG), Dartmouth, Canada, 2018. doi:10.25607/OBP-113.

[42] IOCCG. *Phytoplankton Functional Types from Space*, volume 15 of *Reports of the International Ocean Colour Coordinating Group*. International Ocean Colour Coordinating Group (IOCCG), Dartmouth, Canada, 2014. doi:10.25607/OBP-106.

[43] J.-O. Irisson, S.-D. Ayata, D. J. Lindsay, L. Karp-Boss, and L. Stemmann. *Machine Learning for the Study of Plankton and Marine Snow from Images*. Annual Review of Marine Science 14, 277–301 (2022). doi:10.1146/annurev-marine-041921-013023.

[44] C. B. Field, M. J. Behrenfeld, J. T. Randerson, and P. Falkowski. *Primary production of the biosphere: Integrating terrestrial and oceanic components*. Science 281(5374), 237–240 (1998). doi:10.1126/science.281.5374.237.

[45] F. E. Muller-Karger, E. Hestir, C. Ade, K. Turpie, D. A. Roberts, et al. *Satellite sensor requirements for monitoring essential biodiversity variables of coastal ecosystems*. Ecological Applications 28(3), 749–760 (2018). doi:10.1002/eap.1682.

[46] R. Riegman and A. Rowe. *Nutritional status and pigment composition of phytoplankton during spring and summer Phaeocystis blooms in Dutch coastal waters (Marsdiep area)*. Netherlands Journal of Sea Research 32(1), 13–21 (1994). doi:10.1016/0077-7579(94)90024-8.

[47] M. Lürling and M. Mucci. *Mitigating eutrophication nuisance: in-lake measures are becoming inevitable in eutrophic waters in the Netherlands*. Hydrobiologia 847(21), 4447–4467 (2020). doi:10.1007/s10750-020-04297-9.

[48] L. Carvalho, C. McDonald, C. de Hoyos, U. Mischke, G. Phillips, et al. *Sustaining recreational quality of European lakes: Minimizing the health risks from algal blooms through phosphorus control*. Journal of Applied Ecology 50(2), 315–323 (2013). doi:10.1111/1365-2664.12059.

[49] T. Holtrop, J. Huisman, M. Stomp, L. Biersteker, J. Aerts, et al. *Vibrational modes of water predict spectral niches for photosynthesis in lakes and oceans*. Nature Ecology & Evolution 5, 55–66 (2021). doi:10.1038/s41559-020-01330-x.

[50] A. Valente, S. Sathyendranath, V. Brotas, S. Groom, M. Grant, et al. *A compilation of global bio-optical in situ data for ocean-colour satellite applications – version two*. Earth System Science Data 11(3), 1037–1068 (2019). doi:10.5194/essd-11-1037-2019.

[51] E. Spyrakos, R. O'Donnell, P. D. Hunter, C. Miller, M. Scott, et al. *Optical types of inland and coastal waters*. Limnology and Oceanography 63(2), 846–870 (2018). doi:10.1002/lno.10674.

[52] A. Morel and L. Prieur. *Analysis of variations in ocean color*. Limnology and Oceanography 22(4), 709–722 (1977). doi:10.4319/lo.1977.22.4.0709.

[53] E. Courtecuisse, K. Oxborough, G. H. Tilstone, E. Spyrakos, P. D. Hunter, et al. *Determination of optical markers of cyanobacterial physiology from fluorescence kinetics*. Journal of Plankton Research 44(3), 365–385 (2022). doi:10.1093/plankt/fbac025.

[54] S. B. Groom, S. Sathyendranath, Y. Ban, S. Bernard, B. Brewin, et al. *Satellite ocean colour: Current status and future perspective*. Frontiers in Marine Science 6, 485 (2019). doi:10.3389/fmars.2019.00485.

[55] Z. Zhao, L. Wei, M. Cao, and M. Lu. *A smartphone-based system for fluorescence polarization assays*. Biosensors and Bioelectronics 128, 91–96 (2019). doi:10.1016/j.bios.2018.12.031.

[56] A. Friedrichs, J. A. Busch, H. J. van der Woerd, and O. Zielinski. *SmartFluo: A method and affordable adapter to measure chlorophyll a fluorescence with smartphones*. Sensors 17(4), 678 (2017). doi:10.3390/s17040678.

[57] B. Nechad, K. Ruddick, T. Schroeder, K. Oubelkheir, D. Blondeau-Patissier, et al. *CoastColour Round Robin data sets: A database to evaluate the performance of algorithms for the retrieval of water quality parameters in coastal waters*. Earth System Science Data 7(2), 319–348 (2015). doi:10.5194/essd-7-319-2015.

[58] M. Royo-Llonch, P. Sánchez, C. Ruiz-González, G. Salazar, C. Pedrós-Alió, et al. *Compendium of 530 metagenome-assembled bacterial and archaeal genomes from the polar Arctic Ocean*. Nature Microbiology 6, 1561–1574 (2021). doi:10.1038/s41564-021-00979-9.

[59] R. Astoreca, V. Rousseau, and C. Lancelot. *Coloured dissolved organic matter (CDOM) in Southern North Sea waters: Optical characterization and possible origin*. Estuarine, Coastal and Shelf Science 85(4), 633–640 (2009). doi:10.1016/j.ecss.2009.10.010.

[60] A. Bricaud, A. Morel, and L. Prieur. *Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domains*. Limnology and Oceanography 26(1), 43–53 (1981). doi:10.4319/lo.1981.26.1.0043.

[61] E. K. Duursma. *The production of dissolved organic matter in the sea, as related to the primary gross production of organic matter*. Netherlands Journal of Sea Research 2(1), 85–94 (1963). doi:10.1016/0077-7579(63)90007-3.

[62] K. Toming, T. Kutser, A. Laas, M. Sepp, B. Paavel, et al. *First Experiences in Mapping Lake Water Quality Parameters with Sentinel-2 MSI Imagery*. Remote Sensing 8(8), 640 (2016). doi:10.3390/rs8080640.

[63] R. A. Shuchman, G. Leshkevich, M. J. Sayers, T. H. Johengen, C. N. Brooks, et al. *An algorithm to retrieve chlorophyll, dissolved organic carbon, and suspended minerals from Great Lakes satellite data*. Journal of Great Lakes Research 39, 14–33 (2013). doi:10.1016/j.jglr.2013.06.017.

[64] S. Vignudelli, C. Santinelli, E. Murru, L. Nannicini, and A. Seritti. *Distributions of dissolved organic carbon (DOC) and chromophoric dissolved organic matter (CDOM) in coastal waters of the northern Tyrrhenian Sea (Italy)*. Estuarine, Coastal and Shelf Science 60(1), 133–149 (2004). doi:10.1016/j.ecss.2003.11.023.

[65] Y. Sugimura and Y. Suzuki. *A high-temperature catalytic oxidation method for the determination of non-volatile dissolved organic carbon in seawater by direct injection of a liquid sample*. Marine Chemistry 24(2), 105–131 (1988). doi:10.1016/0304-4203(88)90043-6.

[66] H. A. Moshi, D. A. Shilla, I. A. Kimirei, C. O'Reilly, W. Clymans, et al. *Community monitoring of coliform pollution in Lake Tanganyika*. PLOS ONE 17(1), e0262 881 (2022). doi:10.1371/journal.pone.0262881.

[67] R. van Oosten. *The Dutch Great Stink: The End of the Cesspit Era in the Pre-Industrial Towns of Leiden and Haarlem*. European Journal of Archaeology 19(4), 704–727 (2016). doi:10.1080/14619571.2016.1147677.

[68] E. Fabbri and S. Franzellitti. *Human pharmaceuticals in the marine environment: Focus on exposure and biological effects in animal species*. Environmental Toxicology and Chemistry 35(4), 799–812 (2016). doi:10.1002/etc.3131.

[69] H. S. Barmentlo. *Neonicotinoids in nature: The effects on aquatic invertebrates and their role in ecosystems*. Ph.D. thesis, Universiteit Leiden (2020). ISBN 978-90-5191-194-7.

[70] A. Estim and R. Sudirman. *Types and abundance of macro- and micro-marine debris at Sebatik Island, Tawau, Sabah*. Borneo Journal of Marine Science and Aquaculture 1, 57–64 (2017). doi:10.51200/bjomsa.v1i0.992.

[71] E. Blondel and F. A. Buschman. *Vertical and Horizontal Plastic Litter Distribution in a Bend of a Tidal River*. Frontiers in Environmental Science 10, 861 457 (2022). doi:10.3389/fenvs.2022.861457.

[72] M. E. L. Felismino, P. A. Helm, and C. M. Rochman. *Microplastic and other anthropogenic microparticles in water and sediments of Lake Simcoe*. Journal of Great Lakes Research 47(1), 180–189 (2021). doi:10.1016/j.jglr.2020.10.007.

[73] M. MacLeod, H. P. H. Arp, M. B. Tekman, and A. Jahnke. *The global threat from plastic pollution*. Science 373(6550), 61–65 (2021). doi:10.1126/science.abg5433.

[74] G. Zhong and X. Peng. *Transport and accumulation of plastic litter in submarine canyons—The role of gravity flows*. Geology 49(5), 581–586 (2021). doi:10.1130/G48536.1.

[75] N. Maximenko, P. Corradi, K. L. Law, E. van Sebille, S. P. Garaba, et al. *Toward the Integrated Marine Debris Observing System*. Frontiers in Marine Science 6, 447 (2019). doi:10.3389/fmars.2019.00447.

[76] R. de Vries, M. Egger, T. Mani, and L. Lebreton. *Quantifying Floating Plastic Debris At Sea Using Vessel-Based Optical Data and Artificial Intelligence*. Remote Sensing 13(17), 3401 (2021). doi:10.3390/rs13173401.

[77] T. Bosker, P. Behrens, and M. G. Vijver. *Determining global distribution of microplastics by combining citizen science and in-depth case studies*. Integrated Environmental Assessment and Management 13(3), 536–541 (2017). doi:10.1002/ieam.1908.

[78] C. Lee. *Passing the baton to the next generation: A few problems that need solving*. Annual Review of Marine Science 11, 1–13 (2019). doi:10.1146/annurev-marine-010318-095342.

[79] G. C. Johnson, S. Hosoda, S. R. Jayne, P. R. Oke, S. C. Riser, et al. *Argo—Two Decades: Global Oceanography, Revolutionized*. Annual Review of Marine Science 14, 379–403 (2022). doi:10.1146/annurev-marine-022521-102008.

[80] IOCCG. *Bio-Optical Sensors on Argo Floats*, volume 11 of *Reports of the International Ocean Colour Coordinating Group*. International Ocean Colour Coordinating Group (IOCCG), Dartmouth, Canada, 2011. doi:10.25607/OBP-102.

[81] D. L. Rudnick. *Ocean Research Enabled by Underwater Gliders*. Annual Review of Marine Science 8, 519–541 (2016). doi:10.1146/annurev-marine-122414-033913.

[82] A. Olita, A. Capet, M. Claret, A. Mahadevan, P. M. Poulain, et al. *Frontal dynamics boost primary production in the summer stratified Mediterranean sea*. Ocean Dynamics 67(6), 767–782 (2017). doi:10.1007/s10236-017-1058-z.

[83] N. Haëntjens and E. Boss. *Inlinino: A Modular Software Data Logger for Oceanography*. Oceanography 33(1), 80–84 (2020). doi:10.5670/oceanog.2020.112.

[84] IOCCG. *Inherent Optical Property Measurements and Protocols: Best Practices for the Collection and Processing of Ship-Based Underway Flow-Through Optical Data*. Technical report, International Ocean Colour Coordinating Group (IOCCG), Dartmouth, Canada (2019). doi:10.25607/OBP-691.

[85] C. R. McClain, B. A. Franz, and P. J. Werdell. *Genesis and Evolution of NASA's Satellite Ocean Color Program*. Frontiers in Remote Sensing 3, 938 006 (2022). doi:10.3389/frsen.2022.938006.

[86] P. J. Werdell, M. J. Behrenfeld, P. S. Bontempi, E. Boss, B. Cairns, et al. *The plankton, aerosol, cloud, ocean ecosystem mission status, science, advances*. Bulletin of the American Meteorological Society 100(9), 1775–1794 (2019). doi:10.1175/BAMS-D-18-0056.1.

[87] H. M. Dierssen, S. G. Ackleson, K. E. Joyce, E. L. Hestir, A. Castagna, et al. *Living up to the Hype of Hyperspectral Aquatic Remote Sensing: Science, Resources and Outlook*. Frontiers in Environmental Science 9, 649 528 (2021). doi:10.3389/fenvs.2021.649528.

[88] IOCCG. *Uncertainties in Ocean Colour Remote Sensing*, volume 18 of *Reports of the International Ocean Colour Coordinating Group*. International Ocean Colour Coordinating Group (IOCCG), Dartmouth, Canada, 2019. ISBN 978-1-896246-68-0. doi:10.25607/OBP-696.

[89] G. Free, M. Bresciani, M. Pinardi, C. Giardino, K. Alikas, et al. *Detecting climate driven changes in chlorophyll-a using high frequency monitoring: The impact of the 2019 european heatwave in three contrasting aquatic systems*. Sensors 21(18), 6242 (2021). doi:10.3390/s21186242.

[90] S. G. H. Simis and J. Olsson. *Unattended processing of shipborne hyperspectral reflectance measurements*. Remote Sensing of Environment 135, 202–212 (2013). doi:10.1016/j.rse.2013.04.001.

[91] M. Hieronymi. *Spectral band adaptation of ocean color sensors for applicability of the multi-water biogeo-optical algorithm ONNS*. Optics Express 27(12), A707–A724 (2019). doi:10.1364/OE.27.00A707.

[92] V. Vabson, J. Kuusk, I. Ansko, R. Vendt, K. Alikas, et al. *Field Intercomparison of Radiometers Used for Satellite Validation in the 400–900 nm Range*. Remote Sensing 11(9), 1129 (2019). doi:10.3390/rs11091129.

[93] L. N. H. Verbrugge, W. Ganzevoort, J. M. Fliervoet, K. Panten, and R. J. G. van den Born. *Implementing participatory monitoring in river management: The role of stakeholders' perspectives and incentives*. Journal of Environmental Management 195, 62–69 (2017). doi:10.1016/j.jenvman.2016.11.035.

[94] F. Snik, J. H. H. Rietjens, A. Apituley, H. Volten, B. Mijling, et al. *Mapping atmospheric aerosols with a citizen science network of smartphone spectropolarimeters*. Geophysical Research Letters 41(20), 7351–7358 (2014). doi:10.1002/2014GL061462.

[95] K. Vohland, A. Land-Zandstra, L. Ceccaroni, R. Lemmens, J. Perelló, et al. *The Science of Citizen Science*. Springer International Publishing, 2017, 1st edition. ISBN 978-3-030-58277-7. doi:10.1007/978-3-030-58278-4.

[96] H. R. Gordon. *Physical Principles of Ocean Color Remote Sensing*. University of Miami, Miami, Florida, USA, 2019. doi:10.33596/ppocrs-19.

[97] K. G. Ruddick, K. Voss, A. C. Banks, E. Boss, A. Castagna, et al. *A Review of Protocols for Fiducial Reference Measurements of Downwelling Irradiance for the Validation of Satellite Remote Sensing Data over Water*. Remote Sensing 11(15), 1742 (2019). doi:10.3390/rs11151742.

[98] R. A. Crocombe. *Portable Spectroscopy*. Applied Spectroscopy 72(12), 1701–1751 (2018). doi:10.1177/0003702818809719.

[99] J. R. Schott. *Remote sensing: the image chain approach*. Oxford University Press, 2007, 2nd edition. ISBN 978-0-19-972439-0.

[100] A. Einstein. *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt*. Annalen der Physik 322(6), 132–148 (1905). doi:10.1002/andp.19053220607.

[101] G. Sharma. *Digital Color Imaging Handbook*. CRC Press LLC, Boca Raton, Florida, USA, 2003, 1st edition. ISBN 0-8493-0900-X.

[102] T. Sunamura and K. Horikawa. *Visible-region photographic remote sensing of nearshore waters*. Coastal Engineering Proceedings 1(16), 1439–1453 (1978). doi:10.9753/icce.v16.85.

[103] A. B. Klugh. *Ecological Photometry and a New Instrument for Measuring Light*. Ecology 6(3), 203–237 (1925). doi:10.2307/1929198.

[104] H. Zhang and E. Sanchez-Sinencio. *Linearization techniques for CMOS low noise amplifiers: A tutorial*. IEEE Transactions on Circuits and Systems I: Regular Papers 58(1), 22–36 (2011). doi:10.1109/TCSI.2010.2055353.

[105] M. Talone and G. Zibordi. *Non-linear response of a class of hyper-spectral radiometers*. Metrologia 55(5), 747–758 (2018). doi:10.1088/1681-7575/aadd7f.

[106] J. Mittaz, C. J. Merchant, and E. R. Woolliams. *Applying principles of metrology to historical Earth observations from satellites*. Metrologia 56(3), 032 002 (2019). doi:10.1088/1681-7575/ab1705.

[107] C. Huang, G. Xiang, Y. Chang, L. Han, M. Zhang, et al. *Pre-flight calibration of a multi-angle polarimetric satellite sensor directional polarimetric camera*. Optics Express 28(9), 13 187–13 215 (2020). doi:10.1364/OE.391078.

[108] V. Vabson, J. Kuusk, I. Ansko, R. Vendt, K. Alikas, et al. *Laboratory intercomparison of radiometers used for satellite validation in the 400-900 nm range*. Remote Sensing 11(9), 1101 (2019). doi:10.3390/rs11091101.

[109] D. Antoine, M. Slivkoff, W. Klonowski, C. Kovach, and M. Ondrusek. *Uncertainty assessment of unattended above-water radiometric data collection from research vessels with the Dynamic Above-water Radiance (L) and Irradiance (E) Collector (DALEC)*. Optics Express 29(3), 4607–4631 (2021). doi:10.1364/OE.412022.

[110] A. Białek, S. Douglas, J. Kuusk, I. Ansko, V. Vabson, et al. *Example of Monte Carlo Method Uncertainty Evaluation for Above-Water Ocean Colour Radiometry*. Remote Sensing 12(5), 780 (2020). doi:10.3390/rs12050780.

[111] G. Meister, J. J. Knuble, L. H. Chemerys, H. Choi, N. R. Collins, et al. *Test Results From the Prelaunch Characterization Campaign of the Engineering Test Unit of the Ocean Color Instrument of NASA's Plankton, Aerosol, Cloud and Ocean Ecosystem (PACE) Mission*. Frontiers in Remote Sensing 3, 875 863 (2022). doi:10.3389/frsen.2022.875863.

[112] J. Rietjens, J. Campo, M. Smit, R. Winkelman, R. Nalla, et al. *Optical and system performance of SPEXone, a multi-angle channeled spectropolarimeter for the NASA PACE mission*. In *International Conference on Space Optics — ICSO 2020*, volume 11582, p. 1185234. International Society for Optics and Photonics, 2021. doi:10.1117/12.2599531.

[113] IOCCG. *In-flight Calibration of Satellite Ocean-Colour Sensors*, volume 14 of *Reports of the International Ocean Colour Coordinating Group*. International Ocean Colour Coordinating Group (IOCCG), Dartmouth, Canada, 2013. doi:10.25607/OBP-105.

[114] G. Tilstone, G. Dall'Olmo, M. Hieronymi, K. Ruddick, M. Beck, et al. *Field Intercomparison of Radiometer Measurements for Ocean Colour Validation*. Remote Sensing 12(10), 1587 (2020). doi:10.3390/rs12101587.

[115] M. Wang, P. Naik, and S. Son. *Out-of-band effects of satellite ocean color sensors*. Applied Optics 55(9), 2312–2323 (2016). doi:10.1364/AO.55.002312.

[116] G. van Harten. *Spectropolarimetry for planetary exploration*. Ph.D. thesis, Universiteit Leiden (2014). ISBN 978-94-6259-458-6.

[117] G. Zibordi, B. Holben, F. Mélin, D. D'Alimonte, J. F. Berthon, et al. *AERONET-OC: An overview*. Canadian Journal of Remote Sensing 36(5), 488–497 (2010). doi:10.5589/m10-073.

[118] M. Pagnutti, R. E. Ryan, G. Cazenavette, M. Gold, R. Harlan, et al. *Laying the foundation to use Raspberry Pi 3 V2 camera module imagery for scientific and engineering purposes*. Journal of Electronic Imaging 26(1), 013 014 (2017). doi:10.1117/1.JEI.26.1.013014.

[119] K. J. Voss and N. Souaidia. *POLRADS: polarization radiance distribution measurement system*. Optics Express 18(19), 19 672–19 680 (2010). doi:10.1364/OE.18.019672.

[120] K. G. Ruddick, K. Voss, E. Boss, A. Castagna, R. Frouin, et al. *A Review of Protocols for Fiducial Reference Measurements of Water-Leaving Radiance for Validation of Satellite Remote-Sensing Data over Water*. Remote Sensing 11(19), 2198 (2019). doi:10.3390/rs11192198.

[121] T. Leeuw and E. Boss. *The HydroColor app: Above water measurements of remote sensing reflectance and turbidity using a smartphone camera*. Sensors 18(1), 256 (2018). doi:10.3390/s18010256.

[122] M. Morys, F. M. I. Mims, S. Hagerup, S. E. Anderson, A. Baker, et al. *Design, calibration, and performance of MICROTOPS II handheld ozone monitor and Sun photometer*. Journal of Geophysical Research: Atmospheres 106(D13), 14 573–14 582 (2001). doi:10.1029/2001JD900103.

[123] A. Wright and S. Simis. *Solar-tracking Radiometry platform (So-Rad)*. Technical report, MONOCLE (2021). doi:10.5281/zenodo.4485805.

[124] D. Vansteenwegen, K. Ruddick, A. Cattrijsse, Q. Vanhellemont, and M. Beck. *The Pan-and-Tilt Hyperspectral Radiometer System (PANTHYR) for Autonomous Satellite Validation Measurements—Prototype Design and Testing*. Remote Sensing 11(11), 1360 (2019). doi:10.3390/rs11111360.

[125] G. van Harten, J. de Boer, J. H. H. Rietjens, A. Di Noia, F. Snik, et al. *Atmospheric aerosol characterization with a ground-based SPEX spectropolarimetric instrument*. Atmospheric Measurement Techniques 7(6), 5741–5768 (2014). doi:10.5194/amt-7-4341-2014.

[126] A. Hommersom, S. Kratzer, M. Laanen, I. Ansko, M. Ligi, et al. *Intercomparison in the field between the new WISP-3 and other radiometers (TriOS Ramses, ASD FieldSpec, and TACCS)*. Journal of Applied Remote Sensing 6(1), 063 615 (2012). doi:10.1117/1.JRS.6.063615.

[127] M. Hieronymi. *Polarized reflectance and transmittance distribution functions of the ocean surface*. Optics Express 24(14), A1045–A1068 (2016). doi:10.1364/OE.24.0A1045.

[128] Z. Yang, T. Albrow-Owen, W. Cai, and T. Hasan. *Miniaturization of optical spectrometers*. Science 371(6528), eabe0722 (2021). doi:10.1126/science.abe0722.

[129] N. Hagen and M. W. Kudenov. *Review of snapshot spectral imaging technologies*. Optical Engineering 52(9), 090 901 (2013). doi:10.1117/1.OE.52.9.090901.

[130] M. Davies, M. B. Stuart, M. J. Hobbs, A. J. S. McGonigle, and J. R. Willmott. *Image Correction and In Situ Spectral Calibration for Low-Cost, Smartphone Hyperspectral Imaging*. Remote Sensing 14(5), 1152 (2022). doi:10.3390/rs14051152.

[131] IOCCG. *Ocean-Colour Observations from a Geostationary Orbit*, volume 12 of *Reports of the International Ocean Colour Coordinating Group*. International Ocean Colour Coordinating Group (IOCCG), Dartmouth, Canada, 2012. doi:10.25607/OBP-103.

[132] O. Dubovik, Z. Li, M. I. Mishchenko, D. Tanré, Y. Karol, et al. *Polarimetric remote sensing of atmospheric aerosols: Instruments, methodologies, results, and perspectives*. Journal of Quantitative Spectroscopy and Radiative Transfer 224, 474–511 (2019). doi:10.1016/j.jqsrt.2018.11.024.

[133] M. Chami and D. McKee. *Determination of biogeochemical properties of marine particles using above water measurements of the degree of polarization at the Brewster angle.* Optics Express 15(15), 9494–9509 (2007). doi:10.1364/OE.15.009494.

[134] B. N. Holben, T. F. Eck, I. Slutsker, D. Tanré, J.-P. Buis, et al. *AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization*. Remote Sensing of Environment 66(1), 1–16 (1998). doi:10.1016/S0034-4257(98)00031-5.

[135] E. J. Knight and G. Kvaran. *Landsat-8 Operational Land Imager Design, Characterization and Performance*. Remote Sensing 6(11), 10 286–10 305 (2014). doi:10.3390/rs61110286.

[136] H. Aasen, E. Honkavaara, A. Lucieer, and P. Zarco-Tejada. *Quantitative Remote Sensing at Ultra-High Resolution with UAV Spectroscopy: A Review of Sensor Technology, Measurement Procedures, and Data Correction Workflows*. Remote Sensing 10(7), 1091 (2018). doi:10.3390/rs10071091.

[137] Y. August and A. Stern. *Compressive sensing spectrometry based on liquid crystal devices*. Optics Letters 38(23), 4996–4999 (2013). doi:10.1364/OL.38.004996.

[138] B. E. Bayer. *Color imaging array* (1975). https://patents.google.com/patent/US3971065A/en.

[139] C. Zhang, Y. Li, J. Wang, and P. Hao. *Universal demosaicking of color filter arrays*. IEEE Transactions on Image Processing 25(11), 5173–5186 (2016). doi:10.1109/TIP.2016.2601266.

[140] D. Menon, S. Andriani, and G. Calvagno. *Demosaicing with directional filtering and a posteriori decision*. IEEE Transactions on Image Processing 16(1), 132–141 (2007). doi:10.1109/TIP.2006.884928.

[141] S. M. A. Sharif and Y. J. Jung. *Deep color reconstruction for a sparse color sensor*. Optics Express 27(17), 23 661–23 681 (2019). doi:10.1364/OE.27.023661.

[142] C. Zhang, G. Cheng, P. Edwards, M.-D. Zhou, S. Zheng, et al. *G-Fresnel smartphone spectrometer*. Lab on a Chip 16(2), 246–250 (2016). doi:10.1039/C5LC01226K.

[143] A. J. S. McGonigle, T. C. Wilkes, T. D. Pering, J. R. Willmott, J. M. Cook, et al. *Smartphone spectrometers*. Sensors 18(1), 223 (2018). doi:10.3390/s18010223.

[144] B. B. Cael, A. Chase, and E. Boss. *Information content of absorption spectra and implications for ocean color inversion*. Applied Optics 59(13), 3971–3984 (2020). doi:10.1364/AO.389189.

[145] Y. Wang, Z. Lee, J. Wei, S. Shang, M. Wang, et al. *Extending satellite ocean color remote sensing to the near-blue ultraviolet bands*. Remote Sensing of Environment 253, 112 228 (2021). doi:10.1016/j.rse.2020.112228.

[146] Z. Lee, S. Shang, C. Hu, and G. Zibordi. *Spectral interdependence of remote-sensing reflectance and its implications on the design of ocean color satellite sensors*. Applied Optics 53(15), 3301–3310 (2014). doi:10.1364/AO.53.003301.

[147] Sea-Bird Scientific. *HyperOCR Data Sheet*. Technical report, Sea-Bird Scientific, Bellevue, Washington, USA (2022).

[148] TriOS. *RAMSES Data Sheet*. Technical report, TriOS, Rastede, Germany (2018).

[149] Malvern Panalytical. *ASD FieldSpec 4 Brochure*. Technical report, Malvern Panalytical, Westborough, Massachusetts, USA (2022).

[150] J. Nieke, F. Borde, C. Mavrocordatos, B. Berruti, Y. Delclaud, et al. *The Ocean and Land Colour Imager (OLCI) for the Sentinel 3 GMES Mission: status and first test results*. In *Proc. SPIE*, volume 8528, p. 85280C. International Society for Optics and Photonics, Kyoto, Japan, 2012. doi:10.1117/12.977247.

[151] M. Drusch, U. Del Bello, S. Carlier, O. Colin, V. Fernandez, et al. *Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services*. Remote Sensing of Environment 120, 25–36 (2012). doi:10.1016/j.rse.2011.11.026.

[152] IOCCG. *Protocols for Satellite Ocean Color Data Validation: In situ Optical Radiometry*. Technical report, International Ocean Colour Coordinating Group (IOCCG), Dartmouth, Canada (2019). doi:10.25607/OBP-691.

[153] IOCCG. *Ocean-Colour Data Merging*, volume 6 of *Reports of the International Ocean Colour Coordinating Group*. International Ocean Colour Coordinating Group (IOCCG), Dartmouth, Canada, 2007. doi:10.25607/OBP-100.

[154] J. Neitz and M. Neitz. *The genetics of normal and defective color vision*. Vision Research 51(7), 633–651 (2011). doi:10.1016/j.visres.2010.12.002.

[155] J. Birch. *Worldwide prevalence of red-green color deficiency*. Journal of the Optical Society of America A 29(3), 313–320 (2012). doi:10.1364/JOSAA.29.000313.

[156] F. Snik, A. López Ariste, and S. Bagnulo. *Definitions and Conventions for Polarimetry* (2015).

[157] J. J. Foster, S. E. Temple, M. J. How, I. M. Daly, C. R. Sharkey, et al. *Polarisation vision: overcoming challenges of working with a property of light we barely see*. The Science of Nature 105, 27 (2018). doi:10.1007/s00114-018-1551-3.

[158] C. H. L. Patty. *Circular spectropolarimetric sensing of life*. Ph.D. thesis, Vrije Universiteit Amsterdam (2019). ISBN 978-94-028-1361-6.

[159] J. C. Kemp, G. D. Henson, C. T. Steiner, I. S. Beardsley, and E. R. Powell. *The optical polarization of the Sun measured at a sensitivity of parts in ten million*. Nature 326, 270–273 (1987). doi:10.1038/326270a0.

[160] F. Snik, S. P. Bos, S. A. Brackenhoff, D. S. Doelman, E. H. Por, et al. *Detection of polarization neutral points in observations of the combined corona and sky during the 21 August 2017 total solar eclipse*. Applied Optics 59(21), F71–F77 (2020). doi:10.1364/AO.391814.

[161] L. M. Eshelman and J. A. Shaw. *The VIS–SWIR spectrum of skylight polarization*. Applied Optics 57(27), 7974–7986 (2018). doi:10.1364/AO.57.007974.

[162] A. Kreuter, C. Emde, and M. Blumthaler. *Measuring the influence of aerosols and albedo on sky polarization*. Atmospheric Research 98(2-4), 363–367 (2010). doi:10.1016/j.atmosres.2010.07.010.

[163] A. Ibrahim, A. Gilerson, J. Chowdhary, and S. Ahmed. *Retrieval of macro- and micro-physical properties of oceanic hydrosols from polarimetric observations*. Remote Sensing of Environment 186, 548–566 (2016). doi:10.1016/j.rse.2016.09.004.

[164] H. Volten, J. F. de Haan, J. W. Hovenier, R. Schreurs, W. Vassen, et al. *Laboratory measurements of angular distributions of light scattered by phytoplankton and silt*. Limnology and Oceanography 43(6), 1180–1197 (1998). doi:10.4319/lo.1998.43.6.1180.

[165] T. Harmel. *Recent developments in the use of light polarization for marine environment monitoring from space*. In *Light Scattering Reviews 10: Light Scattering and Radiative Transfer*, pp. 41–84. Springer Berlin Heidelberg, 2016. ISBN 978-3-662-46762-6. doi:10.1007/978-3-662-46762-6_2.

[166] B. Yang, L. Yan, and S. Liu. *Polarization of Light Reflected by Grass: Modeling Using Visible-Sunlit Areas*. Photogrammetric Engineering & Remote Sensing 86(12), 745–752 (2020). doi:10.14358/PERS.86.12.745.

[167] K. J. Voss and E. S. Fry. *Measurement of the Mueller matrix for ocean water*. Applied Optics 23(23), 4427–4439 (1984). doi:10.1364/AO.23.004427.

[168] J. E. Hansen and J. W. Hovenier. *Interpretation of the Polarization of Venus*. Journal of the Atmospheric Sciences 31(4), 1137–1160 (1974). doi:10.1175/1520-0469(1974)031<1137:IOTPOV>2.0.CO;2.

[169] F. Snik and C. U. Keller. *Astronomical polarimetry: Polarized views of stars and planets*. In T. D. Oswalt and H. E. Bond (eds.), *Planets, Stars and Stellar Systems*, chapter 4, pp. 175–221. Springer, Dordrecht, Netherlands, 2013. ISBN 978-94-007-5618-2. doi:10.1007/978-94-007-5618-2_4.

[170] J. S. Tyo, D. L. Goldstein, D. B. Chenault, and J. A. Shaw. *Review of passive imaging polarimetry for remote sensing applications*. Applied Optics 45(22), 5453–5469 (2006). doi:10.1364/AO.45.005453.

[171] R. Foster, D. Gray, J. Bowles, D. Korwan, I. Slutsker, et al. *Mantis: an all-sky visible-to-near-infrared hyper-angular spectropolarimeter*. Applied Optics 59(20), 5896–5909 (2020). doi:10.1364/AO.393822.

[172] J. Hao, Y. Wang, K. Zhou, X. Yu, and Y. Yu. *New diagonal micropolarizer arrays designed by an improved model in fourier domain*. Scientific Reports 11(1), 5778 (2021). doi:10.1038/s41598-021-85103-x.

[173] D. Klindžić, D. M. Stam, F. Snik, C. U. Keller, H. J. Hoeijmakers, et al. *LOUPE: observing Earth from the Moon to prepare for detecting life on Earth-like exoplanets*. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 379(2188), 20190577 (2021). doi:10.1098/rsta.2019.0577.

[174] A. Gilerson, C. Carrizo, A. Ibrahim, R. Foster, T. Harmel, et al. *Hyperspectral polarimetric imaging of the water surface and retrieval of water optical parameters from multi-angular polarimetric data*. Applied Optics 59(10), C8–C20 (2020). doi:10.1364/AO.59.0000C8.

[175] F. Snik, T. Karalidi, and C. U. Keller. *Spectral modulation for full linear polarimetry*. Applied Optics 48(7), 1337–1346 (2009). doi:10.1364/AO.48.001337.

[176] G. van Harten, F. Snik, J. H. H. Rietjens, J. M. Smit, and C. U. Keller. *Spectral line polarimetry with a channeled polarimeter*. Applied Optics 53(19), 4187–4194 (2014). doi:10.1364/AO.53.004187.

[177] J. M. Smit, J. H. H. Rietjens, G. van Harten, A. Di Noia, W. Laauwen, et al. *SPEX airborne spectropolarimeter calibration and performance*. Applied Optics 58(21), 5695–5719 (2019). doi:10.1364/AO.58.005695.

[178] A. Ivanoff and T. H. Waterman. *Factors, mainly depth and wavelength, affecting the degree of underwater light polarization*. Journal of Marine Research 16(3), 283–307 (1958).

[179] H. Volten, J.-P. Jalava, K. Lumme, J. F. de Haan, W. Vassen, et al. *Laboratory measurements and T-matrix calculations of the scattering matrix of rutile particles in water*. Applied Optics 38(24), 5232–5240 (1999). doi:10.1364/AO.38.005232.

[180] M. Ottaviani, R. Foster, A. Gilerson, A. Ibrahim, C. Carrizo, et al. *Airborne and shipborne polarimetric measurements over open ocean and coastal waters: Intercomparisons and implications for spaceborne observations*. Remote Sensing of Environment 206, 375–390 (2018). doi:10.1016/j.rse.2017.12.015.

[181] T. Harmel, A. Tonizzo, A. Ibrahim, A. Gilerson, J. Chowdhary, et al. *Measuring underwater polarization field from above-water hyperspectral instrumentation for water composition retrieval*. In *Proc. SPIE*, volume 8175, p. 817509. International Society for Optics and Photonics, 2011. doi:10.1117/12.898261.

[182] R. Foster, D. J. Gray, D. Koestner, A. El-Habashi, and J. Bowles. *Hydrosol Scattering Matrix Inversion Across a Fresnel Boundary*. Frontiers in Remote Sensing 2, 791 048 (2022). doi:10.3389/frsen.2021.791048.

[183] P.-Y. Deschamps, F.-M. Bréon, M. Leroy, A. Podaire, A. Bricaud, et al. *The POLDER Mission: Instrument Characteristics and Scientific Objectives*. IEEE Transactions on Geoscience and Remote Sensing 32(3), 598–615 (1994). doi:10.1109/36.297978.

[184] O. P. Hasekamp, G. Fu, S. P. Rusli, L. Wu, A. Di Noia, et al. *Aerosol measurements by SPEXone on the NASA PACE mission: expected retrieval capabilities*. Journal of Quantitative Spectroscopy and Radiative Transfer 227, 170–184 (2019). doi:10.1016/j.jqsrt.2019.02.006.

[185] L. A. Remer, K. Knobelspiesse, P.-W. Zhai, F. Xu, O. V. Kalashnikova, et al. *Retrieving Aerosol Characteristics From the PACE Mission, Part 2: Multi-Angle and Polarimetry*. Frontiers in Environmental Science 7, 94 (2019). doi:10.3389/fenvs.2019.00094.

[186] P.-W. Zhai, M. Gao, B. A. Franz, P. J. Werdell, A. Ibrahim, et al. *A Radiative Transfer Simulator for PACE: Theory and Applications*. Frontiers in Remote Sensing 3, 840 188 (2022). doi:10.3389/frsen.2022.840188.

[187] M. Gao, K. Knobelspiesse, B. A. Franz, P.-W. Zhai, V. Martins, et al. *Adaptive Data Screening for Multi-Angle Polarimetric Aerosol and Ocean Color Remote Sensing Accelerated by Deep Learning*. Frontiers in Remote Sensing 2, 757 832 (2021). doi:10.3389/frsen.2021.757832.

[188] R. J. Frouin, B. A. Franz, A. Ibrahim, K. Knobelspiesse, Z. Ahmad, et al. *Atmospheric Correction of Satellite Ocean-Color Imagery During the PACE Era*. Frontiers in Earth Science 7, 145 (2019). doi:10.3389/feart.2019.00145.

[189] L. Qie, Z. Li, S. Zhu, H. Xu, Y. Xie, et al. *In-flight radiometric and polarimetric calibration of the Directional Polarimetric Camera onboard the GaoFen-5 satellite over the ocean*. Applied Optics 60(24), 7186–7199 (2021). doi:10.1364/AO.422980.

[190] S. Stamnes, R. Baize, P. Bontempi, B. Cairns, E. Chemyakin, et al. *Simultaneous Aerosol and Ocean Properties From the PolCube CubeSat Polarimeter*. Frontiers in Remote Sensing 2, 709 040 (2021). doi:10.3389/frsen.2021.709040.

[191] U. Homberg, S. Heinze, K. Pfeiffer, M. Kinoshita, and B. el Jundi. *Central neural coding of sky polarization in insects*. Philosophical Transactions of the Royal Society B: Biological Sciences 366(1565), 680–687 (2011). doi:10.1098/rstb.2010.0199.

[192] J. Marshall, T. W. Cronin, N. Shashar, and M. Land. *Behavioural evidence for polarisation vision in stomatopods reveals a potential channel for communication*. Current Biology 9(14), 755–758 (1999). doi:10.1016/S0960-9822(99)80336-4.

[193] R. P. O'Shea, G. P. Misson, and S. E. Temple. *Seeing polarization of light with the naked eye*. Current Biology 31(4), R178–R179 (2021). doi:10.1016/j.cub.2020.12.037.

[194] W. Haidinger. *Ueber das directe Erkennen des polarisirten Lichts und der Lage der Polarisationsebene*. Annalen der Physik 139(9), 29–39 (1844). doi:10.1002/andp.18441390903.

[195] IOCCG. *Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications*, volume 5 of *Reports of the International Ocean Colour Coordinating Group*. International Ocean Colour Coordinating Group (IOCCG), Dartmouth, Canada, 2006. doi:10.25607/OBP-96.

[196] R. M. Pope and E. S. Fry. *Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements*. Applied Optics 36(33), 8710–8723 (1997). doi:10.1364/AO.36.008710.

[197] A. Bricaud, M. Babin, A. Morel, and H. Claustre. *Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: Analysis and parameterization*. Journal of Geophysical Research 100(C7), 13 321–13 332 (1995). doi:10.1029/95JC00463.

[198] C. S. Roesler, M. J. Perry, and K. L. Carder. *Modeling in situ phytoplankton absorption from total absorption spectra in productive inland marine waters*. Limnology and Oceanography 34(8), 1510–1523 (1989). doi:10.4319/lo.1989.34.8.1510.

[199] A. Bricaud, H. Claustre, J. Ras, and K. Oubelkheir. *Natural variability of phytoplanktonic absorption in oceanic waters: Influence of the size structure of algal populations*. Journal of Geophysical Research: Oceans 109, C11 010 (2004). doi:10.1029/2004JC002419.

[200] S. W. Jeffrey, S. W. Wright, and M. Zapata. *Microalgal classes and their signature pigments*. In S. Roy, C. Llewellyn, E. Egeland, and G. Johnsen (eds.), *Phytoplankton Pigments*, chapter 1, pp. 3–77. Cambridge University Press, Cambridge, UK, 2012. ISBN 978-0-511-73226-3. doi:10.1017/CBO9780511732263.004.

[201] X. Zhang, L. Hu, D. Gray, and Y. Xiong. *Shape of particle backscattering in the North Pacific Ocean: the χ factor*. Applied Optics 60(5), 1260–1266 (2021). doi:10.1364/AO.414695.

[202] W. H. Slade and E. Boss. *Spectral attenuation and backscattering as indicators of average particle size*. Applied Optics 54(24), 7264–7277 (2015). doi:10.1364/AO.54.007264.

[203] X. Zhang and L. Hu. *Light Scattering by Pure Water and Seawater: Recent Development*. Journal of Remote Sensing 2021, 9753 625 (2021). doi:10.34133/2021/9753625.

[204] T. J. Petzold. *Volume Scattering Functions for Selected Ocean Waters*. Technical report, Scripps Institution of Oceanography, La Jolla, California, USA (1972).

[205] T. Harmel, J. Agagliate, M. Hieronymi, and P. Gernez. *Two-term Reynolds–McCormick phase function parameterization better describes light scattering by microalgae and mineral hydrosols*. Optics Letters 46(8), 1860–1863 (2021). doi:10.1364/OL.420344.

[206] S. P. Garaba, D. Voß, J. Wollschläger, and O. Zielinski. *Modern approaches to shipborne ocean color remote sensing*. Applied Optics 54(12), 3602–3612 (2015). doi:10.1364/AO.54.003602.

[207] H. R. Gordon. *Evolution of ocean color atmospheric correction: 1970–2005*. Remote Sensing 13(24), 5051 (2021). doi:10.3390/rs13245051.

[208] A. Gilerson, E. Herrera-Estrella, R. Foster, J. Agagliate, C. Hu, et al. *Determining the Primary Sources of Uncertainty in Retrieval of Marine Remote Sensing Reflectance From Satellite Ocean Color Sensors*. Frontiers in Remote Sensing 3, 857 530 (2022). doi:10.3389/frsen.2022.857530.

[209] C. D. Mobley. *Estimation of the remote-sensing reflectance from above-surface measurements*. Applied Optics 38(36), 7442–7455 (1999). doi:10.1364/AO.38.007442.

[210] P. M. M. Groetsch, P. Gege, S. G. H. Simis, M. A. Eleveld, and S. W. M. Peters. *Validation of a spectral correction procedure for sun and sky reflections in above-water reflectance measurements*. Optics Express 25(16), A742–A761 (2017). doi:10.1364/OE.25.00A742.

[211] B. Fougnie, R. Frouin, P. Lecomte, and P.-Y. Deschamps. *Reduction of skylight reflection effects in the above-water measurement of diffuse marine reflectance*. Applied Optics 38(18), 3844–3856 (1999). doi:10.1364/AO.38.003844.

[212] M. Lo Prejato, D. McKee, and C. Mitchell. *Inherent Optical Properties-Reflectance Relationships Revisited*. Journal of Geophysical Research: Oceans 125(11), e2020JC016 661 (2020). doi:10.1029/2020JC016661.

[213] D. Odermatt, A. Gitelson, V. E. Brando, and M. Schaepman. *Review of constituent retrieval in optically deep and complex waters from satellite imagery*. Remote Sensing of Environment 118, 116–126 (2012). doi:10.1016/j.rse.2011.11.013.

[214] S. G. H. Simis, S. W. M. Peters, and H. J. Gons. *Remote sensing of the cyanobacterial pigment phycocyanin in turbid inland water*. Limnology and Oceanography 50(1), 237–245 (2005). doi:10.4319/lo.2005.50.1.0237.

[215] S. P. Garaba, M. Arias, P. Corradi, T. Harmel, R. de Vries, et al. *Concentration, anisotropic and apparent colour effects on optical reflectance properties of virgin and ocean-harvested plastics*. Journal of Hazardous Materials 406, 124 290 (2021). doi:10.1016/j.jhazmat.2020.124290.

[216] A. P. Chase, E. S. Boss, N. Haëntjens, E. Culhane, C. Roesler, et al. *Plankton Imagery Data Inform Satellite-Based Estimates of Diatom Carbon*. Geophysical Research Letters 49(13), e2022GL098 076 (2022). doi:10.1029/2022GL098076.

[217] P. Hunter, C. Riddick, M. Werther, V. Martinez-Vicente, O. Burggraaff, et al. *In situ bio-optical data in Lake Balaton, MONOCLE H2020 project*. Zenodo (2021). doi:10.5281/zenodo.4559501.

[218] M. Werther, D. Odermatt, S. G. H. Simis, D. Gurlin, D. S. F. Jorge, et al. *Characterising retrieval uncertainty of chlorophyll-a algorithms in oligotrophic and mesotrophic lakes and reservoirs*. ISPRS Journal of Photogrammetry and Remote Sensing 190, 279–300 (2022). doi:10.1016/j.isprsjprs.2022.06.015.

[219] M. Werther, E. Spyrakos, S. G. H. Simis, D. Odermatt, K. Stelzer, et al. *Meta-classification of remote sensing reflectance to estimate trophic status of inland and nearshore waters*. ISPRS Journal of Photogrammetry and Remote Sensing 176, 109–126 (2021). doi:10.1016/j.isprsjprs.2021.04.003.

[220] J. E. O'Reilly and P. J. Werdell. *Chlorophyll algorithms for ocean color sensors - OC4, OC5 & OC6*. Remote Sensing of Environment 229, 32–47 (2019). doi:10.1016/j.rse.2019.04.021.

[221] N. T. T. Ha, N. T. P. Thao, K. Koike, and M. T. Nhuan. *Selecting the Best Band Ratio to Estimate Chlorophyll-a Concentration in a Tropical Freshwater Lake Using Sentinel 2A Images from a Case Study of Lake Ba Be (Northern Vietnam)*. ISPRS International Journal of Geo-Information 6(9), 290 (2017). doi:10.3390/ijgi6090290.

[222] A. Mannino, M. G. Novak, S. B. Hooker, K. Hyde, and D. Aurin. *Algorithm development and validation of CDOM properties for estuarine and continental shelf waters along the northeastern U.S. coast*. Remote Sensing of Environment 152, 576–602 (2014). doi:10.1016/j.rse.2014.06.027.

[223] A. A. Gilerson, A. A. Gitelson, J. Zhou, D. Gurlin, W. Moses, et al. *Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands*. Optics Express 18(23), 24 109–24 125 (2010). doi:10.1364/OE.18.024109.

[224] A. I. Dogliotti, K. G. Ruddick, B. Nechad, D. Doxaran, and E. Knaeps. *A single algorithm to retrieve turbidity from remotely-sensed data in all coastal and estuarine waters*. Remote Sensing of Environment 156, 157–168 (2015). doi:10.1016/j.rse.2014.09.020.

[225] C. B. Mouw, H. Chen, G. A. McKinley, S. Effler, D. O'Donnell, et al. *Evaluation and optimization of bio-optical inversion algorithms for remote sensing of Lake Superior's optical properties*. Journal of Geophysical Research: Oceans 118(4), 1696–1714 (2013). doi:10.1002/jgrc.20139.

[226] Z. Lee, K. L. Carder, and R. A. Arnone. *Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters*. Applied Optics 41(27), 5755 (2002). doi:10.1364/AO.41.005755.

[227] B. Nechad, K. G. Ruddick, and Y. Park. *Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters*. Remote Sensing of Environment 114(4), 854–866 (2010). doi:10.1016/j.rse.2009.11.022.

[228] F. E. Hoge and P. E. Lyon. *Satellite retrieval of inherent optical properties by linear matrix inversion of oceanic radiance models: An analysis of model and radiance measurement errors*. Journal of Geophysical Research 101(C7), 16 631–16 648 (1996). doi:10.1029/96JC01414.

[229] R. A. Vandermeulen, A. Mannino, S. E. Craig, and P. J. Werdell. *150 shades of green: Using the full spectrum of remote sensing reflectance to elucidate color shifts in the ocean*. Remote Sensing of Environment 247, 111 900 (2020). doi:10.1016/j.rse.2020.111900.

[230] S. Watanabe, W. F. Vincent, J. Reuter, S. J. Hook, and S. G. Schladow. *A quantitative blueness index for oligotrophic waters: Application to Lake Tahoe, California-Nevada*. Limnology and Oceanography: Methods 14(2), 100–109 (2016). doi:10.1002/lom3.10074.

[231] M. H. Gholizadeh, A. M. Melesse, and L. Reddi. *A comprehensive review on water quality parameters estimation using remote sensing techniques*. Sensors 16(8), 1298 (2016). doi:10.3390/s16081298.

[232] H. M. Dierssen. *Perspectives on empirical approaches for ocean color remote sensing of chlorophyll in a changing climate*. Proceedings of the National Academy of Sciences 107(40), 17 073–17 078 (2010). doi:10.1073/pnas.0913800107.

[233] A. El-habashi, I. Ioannou, M. C. Tomlinson, R. P. Stumpf, and S. Ahmed. *Satellite Retrievals of Karenia brevis Harmful Algal Blooms in the West Florida Shelf Using Neural Networks and Comparisons with Other Techniques*. Remote Sensing 8(5), 377 (2016). doi:10.3390/rs8050377.

[234] P. J. Werdell, L. I. W. McKinna, E. Boss, S. G. Ackleson, S. E. Craig, et al. *An overview of approaches and challenges for retrieving marine inherent optical properties from ocean color remote sensing*. Progress in Oceanography 160, 186–212 (2018). doi:10.1016/j.pocean.2018.01.001.

[235] D. Koestner, D. Stramski, and R. A. Reynolds. *Characterization of suspended particulate matter in contrasting coastal marine environments with angle-resolved polarized light scattering measurements*. Applied Optics 60(36), 11 161–11 179 (2021). doi:10.1364/AO.441226.

[236] H. Loisel, L. Duforet, D. Dessailly, M. Chami, and P. Dubuisson. *Investigation of the variations in the water leaving polarized reflectance from the POLDER satellite data over two biogeochemical contrasted oceanic areas*. Optics Express 16(17), 12 905–12 918 (2008). doi:10.1364/OE.16.012905.

[237] F. Xu, G. van Harten, D. J. Diner, O. V. Kalashnikova, F. C. Seidel, et al. *Coupled retrieval of aerosol properties and land surface reflection using the Airborne Multiangle SpectroPolarimetric Imager*. Journal of Geophysical Research: Atmospheres 122(13), 7004–7026 (2017). doi:10.1002/2017JD026776.

[238] G. Neukermans, T. Harmel, M. Galí, N. Rudorff, J. Chowdhary, et al. *Harnessing remote sensing to address critical science questions on ocean-atmosphere interactions*. Elementa: Science of the Anthropocene 6(1), 71 (2018). doi:10.1525/elementa.331.

[239] M. Gao, P.-W. Zhai, B. A. Franz, Y. Hu, K. Knobelspiesse, et al. *Inversion of multiangular polarimetric measurements over open and coastal ocean waters: a joint retrieval algorithm for aerosol and water-leaving radiance properties*. Atmospheric Measurement Techniques 12(7), 3921–3941 (2019). doi:10.5194/amt-12-3921-2019.

[240] T. Harmel and M. Chami. *Estimation of the sunglint radiance field from optical satellite imagery over open ocean: Multidirectional approach and polarization aspects*. Journal of Geophysical Research: Oceans 118(1), 76–90 (2013). doi:10.1029/2012JC008221.

[241] P. J. Werdell, S. Bailey, G. Fargion, C. Pietras, K. Knobelspiesse, et al. *Unique data repository facilitates ocean color satellite validation*. Eos, Transactions American Geophysical Union 84(38), 377–387 (2003). doi:10.1029/2003EO380001.

[242] C. Rodero, R. Bardaji, J. Salvador, E. Olmedo, and J. Piera. *Water quality monitoring program through the KduSTICK, a low-cost and Do-It-Yourself instrument connected by the Internet of Things*. In *9th International Workshop on Marine Technology (MARTECH 2021)*, 21, pp. 64–65. SARTI, 2021.

[243] N. Njue, J. Stenfert Kroese, J. Gräf, S. R. Jacobs, B. Weeser, et al. *Citizen science in hydrological monitoring and ecosystem services management: State of the art and future prospects*. Science of the Total Environment 693, 133 531 (2019). doi:10.1016/j.scitotenv.2019.07.337.

[244] C. Kullenberg and D. Kasperowski. *What Is Citizen Science? – A Scientometric Meta-Analysis*. PLOS ONE 11(1), e0147 152 (2016). doi:10.1371/journal.pone.0147152.

[245] L. Alfonso, M. Gharesifard, and U. Wehn. *Analysing the value of environmental citizen-generated data: Complementarity and cost per observation*. Journal of Environmental Management 303, 114 157 (2022). doi:10.1016/j.jenvman.2021.114157.

[246] A. Johnston, E. Matechou, and E. B. Dennis. *Outstanding challenges and future directions for biodiversity monitoring using citizen science data*. Methods in Ecology and Evolution pp. 1–14 (2022). doi:10.1111/2041-210X.13834.

[247] D. W. Walker, M. Smigaj, and M. Tani. *The benefits and negative impacts of citizen science applications to water as experienced by participants and communities*. Wiley Interdisciplinary Reviews: Water 8(1), e1488 (2021). doi:10.1002/wat2.1488.

[248] G. Gawarkiewicz and A. Malek Mercer. *Partnering with Fishing Fleets to Monitor Ocean Conditions*. Annual Review of Marine Science 11, 391–411 (2019). doi:10.1146/annurev-marine-010318-095201.

[249] W. Buytaert, Z. Zulkafli, S. Grainger, L. Acosta, T. C. Alemie, et al. *Citizen science in hydrology and water resources: opportunities for knowledge generation, ecosystem service management, and sustainable development*. Frontiers in Earth Science 2, 26 (2014). doi:10.3389/feart.2014.00026.

[250] D. C. McKinley, A. J. Miller-Rushing, H. L. Ballard, R. Bonney, H. Brown, et al. *Citizen science can improve conservation science, natural resource management, and environmental protection*. Biological Conservation 208, 15–28 (2017). doi:10.1016/j.biocon.2016.05.015.

[251] R. Bonney, T. B. Phillips, H. L. Ballard, and J. W. Enck. *Can citizen science enhance public understanding of science?* Public Understanding of Science 25(1), 2–16 (2016). doi:10.1177/0963662515607406.

[252] I. M. Alsamara, S. F. Beinert, J. C. de Jong, M. J. B. Klappe, V. S. Marewski, et al. *IMPACT: Ideal Measures for Participation and Awareness of Climate Change: Stronger Together Citizen participation in achieving the European Green Deal in the Meuse-Rhine Euroregion*. South Eastern European Journal of Public Health (SEEJPH) 3 (2022). doi:10.11576/seejph-5603.

[253] S. M. Woods, M. Daskolia, A. Joly, P. Bonnet, K. Soacha, et al. *How Networks of Citizen Observatories Can Increase the Quality and Quantity of Citizen-Science-Generated Data Used to Monitor SDG Indicators*. Sustainability 14(7), 4078 (2022). doi:10.3390/su14074078.

[254] S. E. Nelms, E. Easman, N. Anderson, M. Berg, S. Coates, et al. *The role of citizen science in addressing plastic pollution: Challenges and opportunities*. Environmental Science and Policy 128, 14–23 (2022). doi:10.1016/j.envsci.2021.11.002.

[255] B. Alender. *Understanding volunteer motivations to participate in citizen science projects: a deeper look at water quality monitoring*. Journal of Science Communication 15(03), A04 (2016). doi:10.22323/2.15030204.

[256] A. M. Land-Zandstra, J. L. A. Devilee, F. Snik, F. Buurmeijer, and J. M. van den Broek. *Citizen science on a smartphone: Participants' motivations and learning*. Public Understanding of Science 25(1), 45–60 (2016). doi:10.1177/0963662515602406.

[257] B. Weeser, J. Gräf, N. K. Njue, P. Cerutti, M. C. Rufino, et al. *Crowdsourced Water Level Monitoring in Kenya's Sondu-Miriu Basin—Who Is "The Crowd"?* Frontiers in Earth Science 8, 602 422 (2021). doi:10.3389/feart.2020.602422.

[258] S. Lucrezi. *Remote Public Engagement in Coastal Citizen Science: A Systematic Scoping Review*. Journal of Coastal Research 37(6), 1271–1287 (2021). doi:10.2112/JCOASTRES-D-21-00034.1.

[259] J. Pitarch. *A Review of Secchi's Contribution to Marine Optics and the Foundation of Secchi Disk Science*. Oceanography 33(3), 26–37 (2020). doi:10.5670/oceanog.2020.301.

[260] P. A. Secchi. *Relazione delle esperienze fatte a bordo della pontificia pirocorvetta l'Immacolata concezione per determinare la trasparenza del mare*. Il Nuovo Cimento 20, 205–238 (1865). doi:10.1007/BF02726911.

[261] A. Castillo-Ramírez, E. Santamaría-del Ángel, A. González-Silvera, R. Frouin, M.-T. Sebastiá-Frasquet, et al. *A New Algorithm to Estimate Diffuse Attenuation Coefficient from Secchi Disk Depth*. Journal of Marine Science and Engineering 8(8), 558 (2020). doi:10.3390/jmse8080558.

[262] Z. Lee, S. Shang, C. Hu, K. Du, A. Weidemann, et al. *Secchi disk depth: A new theory and mechanistic model for underwater visibility*. Remote Sensing of Environment 169, 139–149 (2015). doi:10.1016/j.rse.2015.08.002.

[263] R. W. Preisendorfer. *Secchi disk science: Visual optics of natural waters*. Limnology and Oceanography 31(5), 909–926 (1986). doi:10.4319/lo.1986.31.5.0909.

[264] S. C. J. Palmer. *Remote sensing of spatiotemporal phytoplankton dynamics of the optically complex Lake Balaton*. Ph.D. thesis, University of Leicester (2015).

[265] R. J. W. Brewin, T. G. Brewin, J. Phillips, S. Rose, A. Abdulaziz, et al. *A Printable Device for Measuring Clarity and Colour in Lake and Nearshore Waters*. Sensors 19(4), 936 (2019). doi:10.3390/s19040936.

[266] Secchi Disk Seafarers, S. Lavender, G. Beaugrand, N. Outram, N. Barlow, et al. *Seafarer citizen scientist ocean transparency data as a resource for phytoplankton and climate research.* PLOS ONE 12(12), e0186 092 (2017). doi:10.1371/journal.pone.0186092.

[267] W. Ule. *Die bestimmung der Wasserfarbe in den Seen*. Kleinere Mittheilungen. Dr. A. Petermanns Mittheilungen aus Justus Perthes geographischer Anstalt pp. 70–71 (1892).

[268] F. A. Forel. *Une nouvelle forme de la gamme de couleur pour l'étude de l'eau des lacs*. Archives des Sciences Physiques et Naturelles/Société de Physique et d'Histoire Naturelle de Genève 6(25) (1890).

[269] M. R. Wernand. *Poseidon's paintbox: Historical archives of ocean colour in global-change perspective*. Ph.D. thesis, Universiteit Utrecht (2011). ISBN 978-90-6464-509-9.

[270] J. Pitarch. *Biases in ocean color over a Secchi disk*. Optics Express 25(24), A1124–A1131 (2017). doi:10.1364/OE.25.0A1124.

[271] M. R. Wernand, H. J. van der Woerd, and W. W. C. Gieskes. *Trends in Ocean Colour and Chlorophyll Concentration from 1889 to 2000, Worldwide*. PLOS ONE 8(6), e63 766 (2013). doi:10.1371/journal.pone.0063766.

[272] Y. Nie, J. Guo, B. Sun, and X. Lv. *An evaluation of apparent color of seawater based on the in-situ and satellite-derived Forel-Ule color scale*. Estuarine, Coastal and Shelf Science 246, 107 032 (2020). doi:10.1016/j.ecss.2020.107032.

[273] M. Andrachuk, M. Marschke, C. Hings, and D. Armitage. *Smartphone technologies supporting community-based environmental monitoring and implementation: a systematic scoping review*. Biological Conservation 237, 430–442 (2019). doi:10.1016/j.biocon.2019.07.026.

[274] J. A. Busch, R. Bardaji, L. Ceccaroni, A. Friedrichs, J. Piera, et al. *Citizen Bio-Optical Observations from Coast- and Ocean and Their Compatibility with Ocean Colour Satellite Measurements*. Remote Sensing 8(12), 879 (2016). doi:10.3390/rs8110879.

[275] T. J. Malthus, R. Ohmsen, and H. J. van der Woerd. *An Evaluation of Citizen Science Smartphone Apps for Inland Water Quality Assessment*. Remote Sensing 12(10), 1578 (2020). doi:10.3390/rs12101578.

[276] Y. Yang, L. L. E. Cowen, and M. Costa. *Is ocean reflectance acquired by citizen scientists robust for science applications?* Remote Sensing 10(6), 835 (2018). doi:10.3390/rs10060835.

[277] M. Kosmala, A. Wiggins, A. Swanson, and B. Simmons. *Assessing data quality in citizen science*. Frontiers in Ecology and the Environment 14(10), 551–560 (2016). doi:10.1002/fee.1436.

[278] A. Gustafson and R. E. Rice. *A review of the effects of uncertainty in public science communication*. Public Understanding of Science 29(6), 614–633 (2020). doi:10.1177/0963662520942122.

[279] T. Thornton and J. Leahy. *Trust in Citizen Science Research: A Case Study of the Groundwater Education Through Water Evaluation & Testing Program*. Journal of the American Water Resources Association 48(5), 1032–1040 (2012). doi:10.1111/j.1752-1688.2012.00670.x.

[280] O. Burggraaff, S. Panchagnula, and F. Snik. *Citizen science with colour blindness: A case study on the Forel-Ule scale*. PLOS ONE 16(4), e0249 755 (2021). doi:10.1371/journal.pone.0249755.

[281] O. Burggraaff, N. Schmidt, J. Zamorano, K. Pauly, S. Pascual, et al. *Standardized spectral and radiometric calibration of consumer cameras*. Optics Express 27(14), 19 075–19 101 (2019). doi:10.1364/OE.27.019075.

[282] O. Burggraaff, M. Werther, E. S. Boss, S. G. H. Simis, and F. Snik. *Accuracy and Reproducibility of Above-Water Radiometry With Calibrated Smartphone Cameras Using RAW Data*. Frontiers in Remote Sensing 3, 940 096 (2022). doi:10.3389/frsen.2022.940096.

[283] O. Burggraaff. *Biases from incorrect reflectance convolution*. Optics Express 28(9), 13 801–13 816 (2020). doi:10.1364/OE.391470.

[284] O. Burggraaff, A. B. Perduijn, R. F. van Hek, N. Schmidt, C. U. Keller, et al. *A universal smartphone add-on for portable spectroscopy and polarimetry: iSPEX 2*. In *Proc. SPIE*, volume 11389, p. 113892K. International Society for Optics and Photonics, 2020. doi:10.1117/12.2558562.

[285] S. Novoa, M. R. Wernand, and H. J. van der Woerd. *WACODI: A generic algorithm to derive the intrinsic color of natural waters from digital images*. Limnology and Oceanography: Methods 13(12), 697–711 (2015). doi:10.1002/lom3.10059.

[286] N. J. Marshall, D. A. Kleine, and A. J. Dean. *CoralWatch: Education, monitoring, and sustainability through citizen science*. Frontiers in Ecology and the Environment 10(6), 332–334 (2012). doi:10.1890/110266.

[287] J. P. Worthington, J. Silvertown, L. Cook, R. Cameron, M. Dodd, et al. *Evolution MegaLab: A case study in citizen science methods*. Methods in Ecology and Evolution 3(2), 303–309 (2012). doi:10.1111/j.2041-210X.2011.00164.x.

[288] J. Bone, M. Archer, D. Barraclough, P. Eggleton, D. Flight, et al. *Public participation in soil surveys: Lessons from a pilot study in England*. Environmental Science and Technology 46(7), 3687–3696 (2012). doi:10.1021/es203880p.

[289] S. Bremer, M. M. Haque, S. B. Aziz, and S. Kvamme. *'My new routine': Assessing the impact of citizen science on climate adaptation in Bangladesh*. Environmental Science and Policy 94, 245–257 (2019). doi:10.1016/j.envsci.2018.12.029.

[290] A. B. Scott and P. C. Frost. *Monitoring water quality in Toronto's urban stormwater ponds: Assessing participation rates and data quality of water sampling by citizen scientists in the FreshWater Watch*. Science of the Total Environment 592, 738–744 (2017). doi:10.1016/j.scitotenv.2017.01.201.

[291] S. Novoa, M. R. Wernand, and H. J. van der Woerd. *The modern Forel-Ule scale: A 'do-it-yourself' colour comparator for water monitoring*. Journal of the European Optical Society: Rapid Publications 9, 14 025 (2014). doi:10.2971/jeos.2014.14025.

[292] E. P. Castilla, D. G. F. Cunha, F. W. F. Lee, S. Loiselle, K. C. Ho, et al. *Quantification of phytoplankton bloom dynamics by citizen scientists in urban and peri-urban environments*. Environmental Monitoring and Assessment 187(11), 690 (2015). doi:10.1007/s10661-015-4912-9.

[293] M. R. Wernand and H. J. van der Woerd. *Ocean colour changes in the North Pacific since 1930*. Journal of the European Optical Society: Rapid Publications 5, 10 015s (2010). doi:10.2971/jeos.2010.10015s.

[294] S. Brouwer and L. K. Hessels. *Increasing research impact with citizen science: The influence of recruitment strategies on sample diversity*. Public Understanding of Science 28(5), 1–16 (2019). doi:10.1177/0963662519840934.

[295] L. Rambonnet, S. C. Vink, A. M. Land-Zandstra, and T. Bosker. *Making citizen science count: Best practices and challenges of citizen science projects on plastics in aquatic environments*. Marine Pollution Bulletin 145, 271–277 (2019). doi:10.1016/j.marpolbul.2019.05.056.

[296] D. Asingizwe, P. M. Poortvliet, C. J. M. Koenraadt, A. J. H. van Vliet, C. M. Ingabire, et al. *Why (not) participate in citizen science? Motivational factors and barriers to participate in a citizen science program for malaria control in Rwanda*. PLOS ONE 15(8), e0237 396 (2020). doi:10.1371/journal.pone.0237396.

[297] M. Budde, A. Schankin, J. Hoffmann, M. Danz, T. Riedel, et al. *Participatory Sensing or Participatory Nonsense? - Mitigating the Effect of Human Error on Data Quality in Citizen Science*. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 1(3), 1–23 (2017). doi:10.1145/3131900.

[298] G. Jordan, S. S. Deeb, J. M. Bosten, and J. D. Mollon. *The dimensionality of color vision in carriers of anomalous trichromacy*. Journal of Vision 10(8), 1–19 (2010). doi:10.1167/10.8.12.

[299] L. Pentao, R. A. Lewis, D. H. Ledbetter, P. I. Patel, and J. R. Lupski. *Maternal uniparental isodisomy of chromosome 14: Association with autosomal recessive rod monochromacy*. American Journal of Human Genetics 50(4), 690–699 (1992).

[300] M. Alpern, K. Kitahara, and D. H. Krantz. *Classical tritanopia*. The Journal of Physiology 335(1), 655–681 (1983). doi:10.1113/jphysiol.1983.sp014557.

[301] G. Verriest. *Further studies on acquired deficiency of color discrimination*. Journal of the Optical Society of America 53(1), 185–195 (1963). doi:10.1364/JOSA.53.000185.

[302] F. Viénot, H. Brettel, and J. D. Mollon. *Digital Video Colourmaps for Checking the Legibility of Displays by Dichromats*. Color Research and Application 24(4), 243–252 (1999). doi:10.1002/(SICI)1520-6378(199908)24:4<243::AID-COL5>3.3.CO;2-V.

[303] H. Hashemi, M. Khabazkhoob, R. Pakzad, A. Yekta, J. Heravian, et al. *The prevalence of color vision deficiency in the northeast of Iran*. Journal of Current Ophthalmology 31(1), 80–85 (2019). doi:10.1016/j.joco.2017.05.005.

[304] F. Crameri, G. E. Shephard, and P. J. Heron. *The misuse of colour in science communication*. Nature Communications 11, 5444 (2020). doi:10.1038/s41467-020-19160-7.

[305] M. R. Wernand and H. J. van der Woerd. *Spectral analysis of the Forel-Ule ocean colour comparator scale*. Journal of the European Optical Society: Rapid Publications 5, 10 014s (2010). doi:10.2971/jeos.2010.10014s.

[306] S. Novoa, M. R. Wernand, and H. J. van der Woerd. *The Forel-Ule scale revisited spectrally: preparation protocol, transmission measurements and chromaticity*. Journal of the European Optical Society: Rapid Publications 8, 13 057 (2013). doi:10.2971/jeos.2013.13057.

[307] H. Brettel, F. Viénot, and J. D. Mollon. *Computerized simulation of color appearance for dichromats*. Journal of the Optical Society of America A 14(10), 2647–2655 (1997). doi:10.1364/JOSAA.14.002647.

[308] G. Sharma, W. Wu, and E. N. Dalal. *The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations*. Color Research and Application 30(1), 21–30 (2005). doi:10.1002/col.20070.

[309] B. Lindbloom. *Chromatic Adaptation*. Website (2017). http://www.brucelindbloom.com/index.html?Eqn_ChromAdapt.html.

[310] R. W. G. Hunt. *Revised colour-appearance model for related and unrelated colours*. Color Research and Application 16(3), 146–165 (1991). doi:10.1002/col.5080160306.

[311] J. Schmitz. *Color Blindness Simulation Research*. Website (2020). https://ixora.io/projects/colorblindness/color-blindness-simulation-research/.

[312] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, et al. *Array programming with NumPy*. Nature 585, 357–362 (2020). doi:10.1038/s41586-020-2649-2.

[313] K. D. Chickering. *Perceptual Significance of the Differences between CIE Tristimulus Values*. Journal of the Optical Society of America 59(8), 986–990 (1969). doi:10.1364/JOSA.59.000986.

[314] M. Mahy, L. Van Eycken, and A. Oosterlinck. *Evaluation of Uniform Color Spaces Developed after the Adoption of CIELAB and CIELUV*. Color Research and Application 19(2), 105–121 (1994). doi:10.1111/j.1520-6378.1994.tb00070.x.

[315] M. Jiménez, I. Triguero, and R. John. *Handling uncertainty in citizen science data: Towards an improved amateur-based large-scale classification*. Information Sciences 479, 301–320 (2019). doi:10.1016/j.ins.2018.12.011.

[316] S. Wang, J. Li, B. Zhang, E. Spyrakos, A. N. Tyler, et al. *Trophic state assessment of global inland waters using a MODIS-derived Forel-Ule index*. Remote Sensing of Environment 217, 444–460 (2018). doi:10.1016/j.rse.2018.08.026.

[317] A. Sánchez de Miguel, C. C. M. Kyba, M. Aubé, J. Zamorano, N. Cardiel, et al. *Colour remote sensing of the impact of artificial light at night (I): The potential of the International Space Station and other DSLR-based platforms*. Remote Sensing of Environment 224, 92–103 (2019). doi:10.1016/j.rse.2019.01.035.

[318] J. B. Gallagher and C. H. Chuan. *Chlorophyll a and turbidity distributions: Applicability of using a smartphone app across two contrasting bays*. Journal of Coastal Research 34(5), 1236–1243 (2018). doi:10.2112/JCOASTRES-D-16-00221.1.

[319] D. P. Igoe, A. V. Parisi, A. Amar, N. J. Downs, and J. Turner. *Atmospheric total ozone column evaluation with a smartphone image sensor*. International Journal of Remote Sensing 39(9), 2766–2783 (2018). doi:10.1080/01431161.2018.1433895.

[320] A. Hänel, T. Posch, S. J. Ribas, M. Aubé, D. Duriscoe, et al. *Measuring night sky brightness: methods and challenges*. Journal of Quantitative Spectroscopy and Radiative Transfer 205, 278–290 (2018). doi:10.1016/j.jqsrt.2017.09.008.

[321] F. Cai, W. Lu, W. Shi, and S. He. *A mobile device-based imaging spectrometer for environmental monitoring by attaching a lightweight small module to a commercial digital camera*. Scientific Reports 7, 15 602 (2017). doi:10.1038/s41598-017-15848-x.

[322] S. Sumriddetchkajorn, K. Chaitavon, and Y. Intaravanne. *Mobile-platform based colorimeter for monitoring chlorine concentration in water*. Sensors and Actuators B: Chemical 191, 561–566 (2014). doi:https://doi.org/10.1016/j.snb.2013.10.024.

[323] L. Goddijn-Murphy, D. Dailloux, M. White, and D. G. Bowers. *Fundamentals of in situ digital camera methodology for water quality monitoring of coast and ocean*. Sensors 9(12), 5825–5843 (2009). doi:10.3390/s90705825.

[324] D. Whiteson, M. Mulhearn, C. Shimmin, K. Cranmer, K. Brodie, et al. *Searching for ultra-high energy cosmic rays with smartphones*. Astroparticle Physics 79, 1–9 (2016). doi:10.1016/j.astropartphys.2016.02.002.

[325] M. Ruwaimana, B. Satyanarayana, V. Otero, A. M. Muslim, M. Syafiq A., et al. *The advantages of using drones over space-borne imagery in the mapping of mangrove forests*. PLOS ONE 13(7), e0200 288 (2018). doi:10.1371/journal.pone.0200288.

[326] J. Rasmussen, G. Ntakos, J. Nielsen, J. Svensgaard, R. N. Poulsen, et al. *Are vegetation indices derived from consumer-grade cameras mounted on UAVs sufficiently reliable for assessing experimental plots?* European Journal of Agronomy 74, 75–92 (2016). doi:10.1016/j.eja.2015.11.026.

[327] K. Flynn and S. Chapra. *Remote sensing of submerged aquatic vegetation in a shallow non-turbid river using an unmanned aerial vehicle*. Remote Sensing 6(12), 12 815–12 836 (2014). doi:10.3390/rs61212815.

[328] X. Liang, A. Jaakkola, Y. Wang, J. Hyyppä, E. Honkavaara, et al. *The use of a hand-held camera for individual tree 3D mapping in forest sample plots*. Remote Sensing 6(7), 6587–6603 (2014). doi:10.3390/rs6076587.

[329] V. Lebourgeois, A. Bégué, S. Labbé, B. Mallavan, L. Prévot, et al. *Can commercial digital cameras be used as multispectral sensors? A crop monitoring test*. Sensors 8(11), 7300–7322 (2008). doi:10.3390/s8117300.

[330] A. Kordecki, A. Bal, and H. Palus. *A study of vignetting correction methods in camera colorimetric calibration*. In *Proc. SPIE*, volume 10341, p. 103410X. International Society for Optics and Photonics, 2017. doi:10.1117/12.2268532.

[331] R. Nguyen, D. K. Prasad, and M. S. Brown. *Raw-to-raw: Mapping between image sensor color responses*. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3398–3405. IEEE, Columbus, Ohio, USA, 2014. doi:10.1109/CVPR.2014.434.

[332] R. Charrière, M. Hébert, A. Trémeau, and N. Destouches. *Color calibration of an RGB camera mounted in front of a microscope with strong color distortion*. Applied Optics 52(21), 5262–5271 (2013). doi:10.1364/AO.52.005262.

[333] J. Jiang, D. Liu, J. Gu, and S. Süsstrunk. *What is the space of spectral sensitivity functions for digital color cameras?* In *2013 IEEE Workshop on Applications of Computer Vision (WACV)*, pp. 168–179. IEEE, Clearwater Beach, Florida, USA, 2013. doi:10.1109/WACV.2013.6475015.

[334] D. L. Bongiorno, M. Bryson, D. G. Dansereau, and S. B. Williams. *Spectral characterization of COTS RGB cameras using a linear variable edge filter*. In *Proc. SPIE*, volume 8660, p. 86600N. International Society for Optics and Photonics, 2013. doi:10.1117/12.2001460.

[335] V. Cheung, C. Li, J. Hardeberg, D. Connah, and S. Westland. *Characterization of trichromatic color cameras by using a new multispectral imaging technique*. Journal of the Optical Society of America A 22(7), 1231–1240 (2005). doi:10.1364/JOSAA.22.001231.

[336] A. Khedekar, B. Devarajan, K. Ramasamy, V. Muthukkaruppan, and U. Kim. *Smartphone-based application improves the detection of retinoblastoma*. Eye 33, 896–901 (2019). doi:10.1038/s41433-018-0333-7.

[337] R. G. Mannino, D. R. Myers, E. A. Tyburski, C. Caruso, J. Boudreaux, et al. *Smartphone app for non-invasive detection of anemia using only patient-sourced photos*. Nature Communications 9(1), 4924 (2018). doi:10.1038/s41467-018-07262-2.

[338] H. Ding, C. Chen, S. Qi, C. Han, and C. Yue. *Smartphone-based spectrometer with high spectral accuracy for mHealth application*. Sensors and Actuators A: Physical 274, 94–100 (2018). doi:10.1016/j.sna.2018.03.008.

[339] L.-J. Wang, Y.-C. Chang, R. Sun, and L. Li. *A multichannel smartphone optical biosensor for high-throughput point-of-care diagnostics*. Biosensors and Bioelectronics 87, 686–692 (2017). doi:10.1016/j.bios.2016.09.021.

[340] K. D. Long, E. V. Woodburn, H. M. Le, U. K. Shah, S. S. Lumetta, et al. *Multimode smartphone biosensing: the transmission, reflection, and intensity spectral (TRI)-analyzer*. Lab on a Chip 17(19), 3246–3257 (2017). doi:10.1039/C7LC00633K.

[341] A. Lam and Y. Kuno. *Robust heart rate measurement from video using select random patches*. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 3640–3648. IEEE, Santiago, Chile, 2015. doi:10.1109/ICCV.2015.415.

[342] A. Skandarajah, C. D. Reber, N. A. Switz, and D. A. Fletcher. *Quantitative imaging with a mobile phone microscope*. PLOS ONE 9(5), e96 906 (2014). doi:10.1371/journal.pone.0096906.

[343] Z. J. Smith, K. Chu, A. R. Espenson, M. Rahimzadeh, A. Gryshuk, et al. *Cell-phone-based platform for biomedical device development and education applications*. PLOS ONE 6(3), e17 150 (2011). doi:10.1371/journal.pone.0017150.

[344] H. Kim, Y. Jung, I.-J. Doh, R. A. Lozano-Mahecha, B. Applegate, et al. *Smartphone-based low light detection for bioluminescence application*. Scientific Reports 7, 40 203 (2017). doi:10.1038/srep40203.

[345] M. Grossi. *A sensor-centric survey on the development of smartphone measurement and sensing systems*. Measurement 135, 572–592 (2019). doi:10.1016/j.measurement.2018.12.014.

[346] A. A. Zaidan, B. B. Zaidan, O. S. Albahri, M. A. Alsalem, A. S. Albahri, et al. *A review on smartphone skin cancer diagnosis apps in evaluation and benchmarking: coherent taxonomy, open issues and recommendation pathway solution*. Health and Technology 8(4), 223–238 (2018). doi:10.1007/s12553-018-0223-9.

[347] S. Kanchi, M. I. Sabela, P. S. Mdluli, Inamuddin, and K. Bisetty. *Smartphone based bioanalytical and diagnosis applications: A review*. Biosensors and Bioelectronics 102, 136–149 (2018). doi:10.1016/j.bios.2017.11.021.

[348] X. Huang, D. Xu, J. Chen, J. Liu, Y. Li, et al. *Smartphone-based analytical biosensors*. Analyst 143(22), 5339–5351 (2018). doi:10.1039/C8AN01269E.

[349] G. Rateni, P. Dario, and F. Cavallo. *Smartphone-based food diagnostic technologies: A review*. Sensors 17(6), 1453 (2017). doi:10.3390/s17061453.

[350] F. Li, Y. Bao, D. Wang, W. Wang, and L. Niu. *Smartphones for sensing*. Science Bulletin 61(3), 190–201 (2016). doi:10.1007/S11434-015-0954-1.

[351] K. E. McCracken and J.-Y. Yoon. *Recent approaches for optical smartphone sensing in resource-limited settings: a brief review*. Analytical Methods 8(36), 6591–6601 (2016). doi:10.1039/C6AY01575A.

[352] S. Manfreda, M. F. McCabe, P. E. Miller, R. Lucas, V. Pajuelo Madrigal, et al. *On the use of unmanned aerial systems for environmental monitoring*. Remote Sensing 10(4), 641 (2018). doi:10.3390/rs10040641.

[353] F. Tauro, M. Porfiri, and S. Grimaldi. *Surface flow measurements from drones*. Journal of Hydrology 540, 240–245 (2016). doi:10.1016/j.jhydrol.2016.06.012.

[354] E. Berra, S. Gibson-Poole, A. MacArthur, R. Gaulton, and A. Hamilton. *Estimation of the spectral sensitivity functions of un-modified and modified commercial off-the-shelf digital cameras to enable their use as a multispectral imaging system for UAVs*. In *International Conference on Unmanned Aerial Vehicles in Geomatics*, pp. 207–214. 2015. doi:10.5194/isprsarchives-XL-1-W4-207-2015.

[355] P. C. Gray, J. T. Ridge, S. K. Poulin, A. C. Seymour, A. M. Schwantes, et al. *Integrating drone imagery into high resolution satellite remote sensing assessments of estuarine environments*. Remote Sensing 10(8), 1257 (2018). doi:10.3390/rs10081257.

[356] S. Levin, S. Krishnan, S. Rajkumar, N. Halery, and P. Balkunde. *Monitoring of fluoride in water samples using a smartphone*. Science of the Total Environment 551-552, 101–107 (2016). doi:10.1016/j.scitotenv.2016.01.156.

[357] A. Kreuter, M. Zangerl, M. Schwarzmann, and M. Blumthaler. *All-sky imaging: a simple, versatile system for atmospheric research*. Applied Optics 48(6), 1091–1097 (2009). doi:10.1364/AO.48.001091.

[358] C. A. Coburn, A. M. Smith, G. S. Logie, and P. Kennedy. *Radiometric and spectral comparison of inexpensive camera systems used for remote sensing*. International Journal of Remote Sensing 39(15-16), 4869–4890 (2018). doi:10.1080/01431161.2018.1466085.

[359] Y. Q. Xu, J. Hua, Z. Gong, W. Zhao, Z. Q. Zhang, et al. *Visible light communication using dual camera on one smartphone*. Optics Express 26(26), 34 609–34 621 (2018). doi:10.1364/OE.26.034609.

[360] Y. Wang, M. M. A. Zeinhom, M. Yang, R. Sun, S. Wang, et al. *A 3D-printed, portable, optical-sensing platform for smartphones capable of detecting the herbicide 2,4-dichlorophenoxyacetic acid*. Analytical Chemistry 89(17), 9339–9346 (2017). doi:10.1021/acs.analchem.7b02139.

[361] A. Manakov. *Evaluation of computational radiometric and spectral sensor calibration techniques*. In *Proc. SPIE*, volume 9896, p. 98960O. International Society for Optics and Photonics, 2016. doi:10.1117/12.2228071.

[362] A. Cescatti. *Indirect estimates of canopy gap fraction based on the linear conversion of hemispherical photographs: Methodology and comparison with standard thresholding techniques*. Agricultural and Forest Meteorology 143(1-2), 1–12 (2007). doi:10.1016/j.agrformet.2006.04.009.

[363] J. Turner, A. V. Parisi, D. P. Igoe, and A. Amar. *Detection of ultraviolet B radiation with internal smartphone sensors*. Instrumentation Science & Technology 45(6), 618–638 (2017). doi:10.1080/10739149.2017.1298042.

[364] D. P. Igoe, A. V. Parisi, and B. Carter. *A method for determining the dark response for scientific imaging with smartphones*. Instrumentation Science & Technology 42(5), 586–592 (2014). doi:10.1080/10739149.2014.915557.

[365] D. Javoršek, T. Jerman, and A. Javoršek. *Comparison of two digital cameras based on spectral data estimation obtained with two methods*. Acta Polytechnica Hungarica 12(7), 183–197 (2015). doi:10.12700/APH.12.7.2015.7.11.

[366] M. M. Darrodi, G. Finlayson, T. Goodman, and M. Mackiewicz. *Reference data set for camera spectral sensitivity estimation*. Journal of the Optical Society of America A 32(3), 381–391 (2015). doi:10.1364/JOSAA.32.000381.

[367] H. Zhao, R. Kawakami, R. T. Tan, and K. Ikeuchi. *Estimating basis functions for spectral sensitivity of digital cameras*. In *Meeting on Image Recognition and Understanding (MIRU)*. 2009.

[368] F. Sigernes, M. Dyrland, N. Peters, D. A. Lorentzen, T. Svenøe, et al. *The absolute sensitivity of digital colour cameras*. Optics Express 17(22), 20 211–20 220 (2009). doi:10.1364/OE.17.020211.

[369] Adobe Systems Incorporated. *Digital Negative (DNG) specification, version 1.4.0.0*. Technical report, Adobe Systems Incorporated (2012).

[370] European Machine Vision Association. *EMVA standard 1288: Standard for characterization of image sensors and cameras*. Technical report, European Machine Vision Association (2016).

[371] C.-H. Lin, K.-L. Chung, and C.-W. Yu. *Novel chroma subsampling strategy based on mathematical optimization for compressing mosaic videos with arbitrary RGB color filter arrays in H.264/AVC and HEVC*. IEEE Transactions on Circuits and Systems for Video Technology 26(9), 1722–1733 (2016). doi:10.1109/TCSVT.2015.2472118.

[372] G. Finlayson, M. M. Darrodi, and M. Mackiewicz. *Rank-based camera spectral sensitivity estimation*. Journal of the Optical Society of America A 33(4), 589–599 (2016). doi:10.1364/JOSAA.33.000589.

[373] V. De Silva, V. Chesnokov, and D. Larkin. *A novel adaptive shading correction algorithm for camera systems*. In *Electronic Imaging, Digital Photography and Mobile Imaging XII*, pp. 1–5. 2016. doi:10.2352/ISSN.2470-1173.2016.18.DPMI-249.

[374] D. B. Goldman and J.-H. Chen. *Vignette and exposure calibration and compensation*. IEEE Transactions on Pattern Analysis and Machine Intelligence 32(12), 2276–2288 (2010). doi:10.1109/TPAMI.2010.55.

[375] Y. Zheng, S. Lin, C. Kambhamettu, J. Yu, and S. Bing Kang. *Single-image vignetting correction*. IEEE Transactions on Pattern Analysis and Machine Intelligence 31(12), 2243–2256 (2009). doi:10.1109/TPAMI.2008.263.

[376] S. J. Kim and M. Pollefeys. *Robust radiometric calibration and vignetting correction*. IEEE Transactions on Pattern Analysis and Machine Intelligence 30(4), 562–576 (2008). doi:10.1109/TPAMI.2007.70732.

[377] W. Yu, Y. Chung, and J. Soh. *Vignetting distortion correction method for high quality digital imaging*. In *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, pp. 666–669. IEEE, 2004. doi:10.1109/ICPR.2004.1334617.

[378] I. S. McLean. *Electronic imaging in astronomy: detectors and instrumentation*. Springer-Verlag Berlin Heidelberg, 2008, 2nd edition. ISBN 978-3-642-09532-0. doi:10.1007/978-3-540-76583-7.

[379] C. A. Gueymard. *Parameterized transmittance model for direct beam and circumsolar spectral irradiance*. Solar Energy 71(5), 325–346 (2001). doi:10.1016/S0038-092X(01)00054-8.

[380] C. A. Gueymard. *SMARTS2, A Simple Model of the Atmospheric Radiative Transfer of Sunshine: Algorithms and performance assessment*. Technical report, Florida Solar Energy Center/University of Central Florida (1995).

[381] C. E. Tapia Ayuga and J. Zamorano. *LICA AstroCalc, a software to analyze the impact of artificial light: Extracting parameters from the spectra of street and indoor lamps*. Journal of Quantitative Spectroscopy and Radiative Transfer 214, 33–38 (2018). doi:10.1016/j.jqsrt.2018.04.022.

[382] U. Beisl. *Absolute spectroradiometric calibration of the ADS40 sensor*. In *Proc. of the ISPRS Commission I Symposium "From Sensors to Imagery", Paris – Marne-la-Vallée*. 2006.

[383] S. W. Brown, T. C. Larason, C. Habauzit, G. P. Eppeldauer, Y. Ohno, et al. *Absolute radiometric calibration of digital imaging systems*. In *Proc. SPIE*, volume 4306, pp. 13–21. International Society for Optics and Photonics, 2001. doi:10.1117/12.426952.

[384] S. Chaji, A. Pourreza, H. Pourreza, and M. Rouhani. *Estimation of the camera spectral sensitivity function using neural learning and architecture*. Journal of the Optical Society of America A 35(6), 850–858 (2018). doi:10.1364/JOSAA.35.000850.

[385] P.-K. Yang. *Determining the spectral responsivity from relative measurements by using multicolor light-emitting diodes as probing light sources*. Optik 126(21), 3088–3092 (2015). doi:10.1016/j.ijleo.2015.07.097.

[386] M. Iwaki, N. Takamura, S. Nakada, and H. Oguma. *Monitoring of Lake Environment Using a Fixed Point and Time-lapse Camera — Case Study of South Basin of Lake Biwa*. Journal of The Remote Sensing Society of Japan 41(5), 563–574 (2021). doi:10.11440/rssj.41.563.

[387] H. S. Lim, M. Z. Mat Jafri, K. Abdullah, and M. N. Abu Bakar. *Water quality mapping using digital camera images*. International Journal of Remote Sensing 31(19), 5275–5295 (2010). doi:10.1080/01431160903283843.

[388] D. Hatiboruah, T. Das, N. Chamuah, D. Rabha, B. Talukdar, et al. *Estimation of trace-mercury concentration in water using a smartphone*. Measurement 154, 107 507 (2020). doi:10.1016/j.measurement.2020.107507.

[389] C. Garcia-Soto, J. J. C. Seys, O. Zielinski, J. A. Busch, S. I. Luna, et al. *Marine Citizen Science: Current State in Europe and New Technological Developments*. Frontiers in Marine Science 8, 621 472 (2021). doi:10.3389/fmars.2021.621472.

[390] M. B. Stuart, A. J. McGonigle, M. Davies, M. J. Hobbs, N. A. Boone, et al. *Low-cost hyperspectral imaging with a smartphone*. Journal of Imaging 7(8), 136 (2021). doi:10.3390/jimaging7080136.

[391] M. Gao, J. Li, S. Wang, F. Zhang, K. Yan, et al. *Smartphone–Camera–Based Water Reflectance Measurement and Typical Water Quality Parameter Inversion*. Remote Sensing 14(6), 1371 (2022). doi:10.3390/rs14061371.

[392] K. H. D. Al-Ghifari, S. Nurdjaman, S. Nur, B. D. P. P. Cahya, D. A. Widiawan, et al. *Low Cost Method Of Turbidity Estimation Using A Smartphone Application In Cirebon Waters, Indonesia*. Borneo Journal of Marine Science and Aquaculture 5(1), 32–36 (2021). doi:10.51200/bjomsa.v5i1.2713.

[393] A. O. Ayeni and J. I. Odume. *Analysis of Algae Concentration in the Lagos Lagoon Using Eye on Water and Algae Estimator Mobile App*. FUTY Journal of the Environment 14(2), 105–115 (2020).

[394] M. Gao, J. Li, F. Zhang, S. Wang, Y. Xie, et al. *Measurement of water leaving reflectance using a digital camera based on multiple reflectance reference cards*. Sensors 20(22), 6580 (2020). doi:10.3390/s20226580.

[395] J. A. Busch, I. Price, E. Jeansou, O. Zielinski, and H. J. van der Woerd. *Citizens and satellites: Assessment of phytoplankton dynamics in a NW Mediterranean aquaculture zone*. International Journal of Applied Earth Observation and Geoinformation 47, 40–49 (2016). doi:10.1016/j.jag.2015.11.017.

[396] I. A. Pratama, H. Hariyadi, A. Wirasatriya, L. Maslukah, and M. Yusuf. *Validasi Pengukuran Turbiditas dan Material Padatan Tersuspensi di Banjir Kanal Barat, Semarang dengan Menggunakan Smartphone*. Indonesian Journal of Oceanography 3(2), 22–29 (2021). doi:10.14710/ijoce.v3i2.11158.

[397] Y. O. Ouma, J. Waga, M. Okech, O. Lavisa, and D. Mbuthia. *Estimation of Reservoir Bio-Optical Water Quality Parameters Using Smartphone Sensor Apps and Landsat ETM+: Review and Comparative Experimental Results*. Journal of Sensors 2018, 3490 757 (2018). doi:10.1155/2018/3490757.

[398] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, et al. *SciPy 1.0: fundamental algorithms for scientific computing in Python*. Nature Methods 17(3), 261–272 (2020). doi:10.1038/s41592-019-0686-2.

[399] T. M. Jordan, S. G. H. Simis, P. M. M. Grötsch, and J. Wood. *Incorporating a Hyperspectral Direct-Diffuse Pyranometer in an Above-Water Reflectance Algorithm*. Remote Sensing 14(10), 2491 (2022). doi:10.3390/rs14102491.

[400] R. J. Soffer, J. W. Harron, and J. R. Miller. *Characterization of Kodak Grey Cards as Reflectance Reference Panels in Support of BOREAS Field Activities*. In *Proceedings of the 17th Canadian Symposium on Remote Sensing*, pp. 357–362. Saskatoon, Saskatchewan, Canada, 1995.

[401] I. Nimeroff. *Propagation of errors in tristimulus colorimetry*. Journal of the Optical Society of America 47(8), 697–702 (1957). doi:10.1364/JOSA.47.000697.

[402] M. R. Wernand, A. Hommersom, and H. J. van der Woerd. *MERIS-based ocean colour classification with the discrete Forel-Ule scale*. Ocean Science 9, 477–487 (2013). doi:10.5194/os-9-477-2013.

[403] R. Juckett. *RGB Color Space Conversion*. Website (2010). https://www.ryanjuckett.com/rgb-color-space-conversion/.

[404] G. Wyszecki. *Propagation of errors in colorimetric transformations*. Journal of the Optical Society of America 49(4), 389–393 (1959). doi:10.1364/JOSA.49.000389.

[405] H. Onusic and D. Mandic. *Propagation of errors in chromaticity coefficients (x, y) obtained from spectro-radiometric curves: tristimulus covariances included*. International Journal of Vehicle Design 10(1), 79–88 (1989). doi:10.1504/IJVD.1989.061564.

[406] S. K. Morley, T. V. Brito, and D. T. Welling. *Measures of Model Performance Based On the Log Accuracy Ratio*. Space Weather 16(1), 69–88 (2018). doi:10.1002/2017SW001669.

[407] B. N. Seegers, R. P. Stumpf, B. A. Schaeffer, K. A. Loftin, and P. J. Werdell. *Performance metrics for the assessment of satellite data products: an ocean color case study*. Optics Express 26(6), 7404–7422 (2018). doi:10.1364/OE.26.007404.

[408] L. A. Wasserman. *All of Statistics: A Concise Course in Statistical Inference*. Springer, 2004. ISBN 978-0-387-40272-7.

[409] I. H. Woodhouse. *On 'ground' truth and why we should abandon the term*. Journal of Applied Remote Sensing 15(4), 041 501 (2021). doi:10.1117/1.JRS.15.041501.

[410] H. J. van der Woerd and M. R. Wernand. *Hue-Angle Product for Low to Medium Spatial Resolution Optical Satellite Sensors*. Remote Sensing 10(2), 180 (2018). doi:10.3390/rs10020180.

[411] H. J. van der Woerd and M. R. Wernand. *True Colour Classification of Natural Waters with Medium-Spectral Resolution Satellites: SeaWiFS, MODIS, MERIS and OLCI*. Sensors 15(10), 25 663–25 680 (2015). doi:10.3390/s151025663.

[412] S. Tominaga, S. Nishi, R. Ohtera, and H. Sakai. *Improved method for spectral reflectance estimation and application to mobile phone cameras*. Journal of the Optical Society of America A 39(3), 494–508 (2022). doi:10.1364/JOSAA.449347.

[413] C. Cox and W. Munk. *Measurement of the Roughness of the Sea Surface from Photographs of the Sun's Glitter*. Journal of the Optical Society of America 44(11), 838–850 (1954). doi:10.1364/JOSA.44.000838.

[414] S. P. Garaba, J. Schulz, M. R. Wernand, and O. Zielinski. *Sunglint detection for unmanned and automated platforms*. Sensors 12(9), 12 545–12 561 (2012). doi:10.3390/s120912545.

[415] M. Badawi, D. Helder, L. Leigh, and X. Jing. *Methods for earth-observing satellite surface reflectance validation*. Remote Sensing 11(13), 1543 (2019). doi:10.3390/rs11131543.

[416] S. L. Cundill, H. M. A. van der Werff, and M. van der Meijde. *Adjusting Spectral Indices for Spectral Response Function Differences of Very High Spatial Resolution Sensors Simulated from Field Spectra*. Sensors 15(3), 6221–6240 (2015). doi:10.3390/s150306221.

[417] S. Dutkiewicz, A. E. Hickman, O. Jahn, S. Henson, C. Beaulieu, et al. *Ocean colour signature of climate change*. Nature Communications 10, 578 (2019). doi:10.1038/s41467-019-08457-x.

[418] N. Pahlevan, B. Smith, C. Binding, and D. M. O'Donnell. *Spectral band adjustments for remote sensing reflectance spectra in coastal/inland waters*. Optics Express 25(23), 28 650–28 667 (2017). doi:10.1364/OE.25.028650.

[419] F. Mélin and G. Sclep. *Band shifting for ocean color multi-spectral reflectance data*. Optics Express 23(3), 2262–2279 (2015). doi:10.1364/OE.23.002262.

[420] L. Lymburner, E. Botha, E. Hestir, J. Anstee, S. Sagar, et al. *Landsat 8: Providing continuity and increased precision for measuring multi-decadal time series of total suspended matter*. Remote Sensing of Environment 185, 108–118 (2016). doi:10.1016/j.rse.2016.04.011.

[421] L. A. Remer, A. B. Davis, S. Mattoo, R. C. Levy, O. V. Kalashnikova, et al. *Retrieving Aerosol Characteristics From the PACE Mission, Part 1: Ocean Color Instrument*. Frontiers in Earth Science 7, 152 (2019). doi:10.3389/feart.2019.00152.

[422] S. W. Bailey and P. J. Werdell. *A multi-sensor approach for the on-orbit validation of ocean color satellite data products*. Remote Sensing of Environment 102(1-2), 12–23 (2006). doi:10.1016/j.rse.2006.01.015.

[423] B. B. Barnes, J. P. Cannizzaro, D. C. English, and C. Hu. *Validation of VIIRS and MODIS reflectance data in coastal and oceanic waters: An assessment of methods*. Remote Sensing of Environment 220, 110–123 (2019). doi:10.1016/j.rse.2018.10.034.

[424] Z. Cao, R. Ma, H. Duan, and K. Xue. *Effects of broad bandwidth on the remote sensing of inland waters: Implications for high spatial resolution satellite data applications*. ISPRS Journal of Photogrammetry and Remote Sensing 153, 110–122 (2019). doi:10.1016/j.isprsjprs.2019.05.001.

[425] F. H. Holman, A. B. Riche, M. Castle, M. J. Wooster, and M. J. Hawkesford. *Radiometric Calibration of 'Commercial off the Shelf' Cameras for UAV-Based High-Resolution Temporal Crop Phenotyping of Reflectance and NDVI*. Remote Sensing 11(14), 1657 (2019). doi:10.3390/rs11141657.

[426] M. Kampel, J. A. Lorenzzetti, C. M. Bentz, R. A. Nunes, R. Paranhos, et al. *Simultaneous Measurements of Chlorophyll Concentration by Lidar, Fluorometry, above-Water Radiometry, and Ocean Color MODIS Images in the Southwestern Atlantic*. Sensors 9(1), 528–541 (2009). doi:10.3390/s90100528.

[427] D. K. Clark, H. R. Gordon, K. J. Voss, Y. Ge, W. Broenkow, et al. *Validation of atmospheric correction over the oceans*. Journal of Geophysical Research: Atmospheres 102(14), 17 209–17 217 (1997). doi:10.1029/96JD03345.

[428] H. R. Gordon. *Remote sensing of ocean color: a methodology for dealing with broad spectral bands and significant out-of-band response*. Applied Optics 34(36), 8363–8374 (1995). doi:10.1364/AO.34.008363.

[429] M. Wang. *Rayleigh radiance computations for satellite remote sensing: accounting for the effect of sensor spectral response function*. Optics Express 24(11), 12 414–12 429 (2016). doi:10.1364/OE.24.012414.

[430] Z. Lee. *Applying narrowband remote-sensing reflectance models to wideband data*. Applied Optics 48(17), 3177–3183 (2009). doi:10.1364/AO.48.003177.

[431] E. Knaeps, A. I. Dogliotti, D. Raymaekers, K. Ruddick, and S. Sterckx. *In situ evidence of non-zero reflectance in the OLCI 1020nm band for a turbid estuary*. Remote Sensing of Environment 120, 133–144 (2012). doi:10.1016/j.rse.2011.07.025.

[432] A. G. Dekker, R. J. Vos, and S. W. M. Peters. *Comparison of remote sensing data, model results and in situ data for total suspended matter (TSM) in the southern Frisian lakes*. Science of the Total Environment 268(1-3), 197–214 (2001). doi:10.1016/S0048-9697(00)00679-3.

[433] G. Zibordi, M. Talone, K. J. Voss, and B. C. Johnson. *Impact of spectral resolution of in situ ocean color radiometric data in satellite matchups analyses*. Optics Express 25(16), A798–A812 (2017). doi:10.1364/OE.25.00A798.

[434] N. Pahlevan, S. Sarkar, and B. A. Franz. *Uncertainties in coastal ocean color products: Impacts of spatial sampling*. Remote Sensing of Environment 181, 14–26 (2016). doi:10.1016/j.rse.2016.03.022.

[435] A. Castagna, B. C. Johnson, K. Voss, H. M. Dierssen, H. Patrick, et al. *Uncertainty in global downwelling plane irradiance estimates from sintered polytetrafluoroethylene plaque radiance measurements*. Applied Optics 58(16), 4497–4511 (2019). doi:10.1364/AO.58.004497.

[436] W. Wu, X. Liu, X. Xiong, Y. Li, Q. Yang, et al. *An Accurate Method for Correcting Spectral Convolution Errors in Intercalibration of Broadband and Hyperspectral Sensors*. Journal of Geophysical Research: Atmospheres 123(17), 9238–9255 (2018). doi:10.1029/2018JD028585.

[437] S. P. Garaba and O. Zielinski. *Comparison of remote sensing reflectance from above-water and in-water measurements west of Greenland, Labrador Sea, Denmark Strait, and west of Iceland*. Optics Express 21(13), 15 938–15 950 (2013). doi:10.1364/OE.21.015938.

[438] C. Roesler. *ArabianSea_2011*. SeaBASS (2011). doi:10.5067/SeaBASS/ARABIANSEA_2011/DATA001.

[439] C. Hu and F. Muller-Karger. *CARIACO*. SeaBASS (2017). doi:10.5067/SeaBASS/CARIACO/DATA001.

[440] G. Cota and R. Zimmerman. *Chesapeake_Light_Tower*. SeaBASS (2007). doi:10.5067/SeaBASS/CHESAPEAKE_LIGHT_TOWER/DATA001.

[441] W. Balch, M. Behrenfeld, F. Chavez, C. Del Castillo, H. Dierssen, et al. *GasEx*. SeaBASS (2008). doi:10.5067/SeaBASS/GASEX/DATA001.

[442] S. P. Garaba, M. R. Wernand, and O. Zielinski. *Radiance, irradiance, and remote sensing reflectance during the North Sea Coast Harmful Algal Bloom (NORCOHAB II) HEINCKE cruise HE302*. PANGAEA (2011). doi:10.1594/PANGAEA.759690.

[443] S. P. Garaba and O. Zielinski. *Downwelling solar irradiance, upwelling solar radiance, sky leaving radiance, and cloud cover observed during ARCHEMHAB study (on Maria S. Merian Leg MSM21/3) from 2012-07-26 to 2012-08-10*. PANGAEA (2013). doi:10.1594/PANGAEA.819717.

[444] F. Muller-Karger. *ORINOCO*. SeaBASS (2000). doi:10.5067/SeaBASS/ORINOCO/DATA001.

[445] I. Cetinic. *RemSensPOC*. SeaBASS (2014). doi:10.5067/SeaBASS/REMSENSPOC/DATA001.

[446] M. Behrenfeld, I. Cetinic, A. Gilerson, and M. Twardowski. *sabor*. SeaBASS (2014). doi:10.5067/SeaBASS/SABOR/DATA001.

[447] F. Muller-Karger. *SFP*. SeaBASS (2018). doi:10.5067/SeaBASS/SFP/DATA001.

[448] R. Zimmerman. *Seagrass_Mapping_Florida*. SeaBASS (2010). doi:10.5067/SeaBASS/SEAGRASS_MAPPING_FLORIDA/DATA001.

[449] S. P. Garaba, D. Voß, J. Wollschläger, and O. Zielinski. *Shipborne ocean color remote sensing*. PANGAEA (2015). doi:10.1594/PANGAEA.845444.

[450] E. Boss and H. Claustre. *Tara Mediterranean 2014*. SeaBASS (2014). doi:10.5067/SeaBASS/TARA_MEDITERRANEAN/DATA001.

[451] E. Boss and H. Claustre. *Tara Oceans expedition*. SeaBASS (2012). doi:10.5067/SeaBASS/TARA_OCEANS_EXPEDITION/DATA001.

[452] NASA. *Spectral Characterization*. NASA website. https://landsat.gsfc.nasa.gov/about/spectral-characterization/.

[453] NASA. *Spectral Response of the Operational Land Imager In-Band, Band-Average Relative Spectral Response*. NASA website. https://landsat.gsfc.nasa.gov/preliminary-spectral-response-of-the-operational-land-imager-in-band-band-average-relative-spectral-response/.

[454] NASA. *Spectral Bandpass Integration*. NASA website. https://oceancolor.gsfc.nasa.gov/docs/rsr/.

[455] ESA. *Sentinel-2 Spectral Response Functions (S2-SRF)*. ESA website. https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/document-library/-/asset_publisher/Wk0TKajiISaR/content/sentinel-2a-spectral-responses.

[456] ESA. *Spectral Characterisation Data*. ESA website. https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-3-olci/olci-instrument/spectral-characterisation-data.

[457] F. Snik, J. Craven-Jones, M. Escuti, S. Fineschi, D. Harrington, et al. *An overview of polarimetric sensing techniques and technology with applications to different research fields*. In *Proc. SPIE*, volume 9099, p. 90990B. International Society for Optics and Photonics, 2014. doi:10.1117/12.2053245.

[458] F. Snik, C. U. Keller, D. S. Doelman, J. Kühn, C. H. L. Patty, et al. *A snapshot full-Stokes spectropolarimeter for detecting life on Earth*. In *Proc. SPIE*, volume 11132, p. 111320A. International Society for Optics and Photonics, 2019. doi:10.1117/12.2529548.

[459] Z. Sun, D. Wu, Y. Lv, and Y. Zhao. *Polarized reflectance factors of vegetation covers from laboratory and field: A comparison with modeled results*. Journal of Geophysical Research: Atmospheres 122(2), 1042–1065 (2017). doi:10.1002/2016JD025892.

[460] T. Vlemmix, A. J. M. Piters, P. Stammes, P. Wang, and P. F. Levelt. *Retrieval of tropospheric NO_2 using the MAX-DOAS method combined with relative intensity measurements for aerosol correction*. Atmospheric Measurement Techniques 3(5), 1287–1305 (2010). doi:10.5194/amt-3-1287-2010.

[461] A. Deepak, G. P. Box, and M. A. Box. *Experimental validation of the solar aureole technique for determining aerosol size distributions*. Applied Optics 21(12), 2236–2243 (1982). doi:10.1364/AO.21.002236.

[462] T. G. Baur and M. Kraemer. *The power of polymers and liquid crystals for polarization control*. In *Proc. SPIE*, volume 11132, p. 111320L. International Society for Optics and Photonics, 2019. doi:10.1117/12.2530705.

[463] A. B. Perduijn and S. M. Hoogenboom. *The pure bending of sheet*. Journal of Materials Processing Technology 51(1-4), 274–295 (1995). doi:10.1016/0924-0136(94)01596-S.

[464] L. Yuan, J. Xie, Z. He, Y. Wang, and J. Wang. *Optical design and evaluation of airborne prism-grating imaging spectrometer*. Optics Express 27(13), 17 686–17 700 (2019). doi:10.1364/OE.27.017686.

[465] P. D. Hale and G. W. Day. *Stability of birefringent linear retarders (waveplates)*. Applied Optics 27(24), 5146–5153 (1988). doi:10.1364/AO.27.005146.

[466] A. Di Noia, O. P. Hasekamp, G. van Harten, J. H. H. Rietjens, J. M. Smit, et al. *Use of neural networks in ground-based aerosol retrievals from multi-angle spectropolarimetric observations*. Atmospheric Measurement Techniques 8, 281–299 (2015). doi:10.5194/amt-8-281-2015.

[467] S. Jutz and M. P. Milagro-Pérez. *Copernicus: the European Earth Observation programme*. Revista de Teledetección 56, V–XI (2020). doi:10.4995/raet.2020.14346.

[468] T. A. C. Gordon, H. R. Harding, F. K. Clever, I. K. Davidson, W. Davison, et al. *Fishes in a changing world: learning from the past to promote sustainability of fish populations*. Journal of Fish Biology 92(3), 804–827 (2018). doi:10.1111/jfb.13546.

[469] A. S. Rigual Hernández, T. W. Trull, S. D. Nodder, J. A. Flores, H. Bostock, et al. *Coccolithophore biodiversity controls carbonate export in the Southern Ocean*. Biogeosciences 17, 245–263 (2020). doi:10.5194/bg-17-245-2020.

[470] M. L. Estapa, M. L. Feen, and E. Breves. *Direct Observations of Biological Carbon Export From Profiling Floats in the Subtropical North Atlantic*. Global Biogeochemical Cycles 33(3), 282–300 (2019). doi:10.1029/2018GB006098.

[471] A. R. Boon and J. C. Kromkamp. *Climate change and intensifying human use call for a monitoring upgrade of the Dutch North Sea*. Journal of Sea Research 182, 102 185 (2022). doi:10.1016/j.seares.2022.102185.

[472] C. E. Bulgin, C. M. Thomas, J. A. Waller, and E. R. Woolliams. *Representation Uncertainty in the Earth Sciences*. Earth and Space Science 9(6), e2021EA002 129 (2022). doi:10.1029/2021EA002129.

[473] L. Padilla, S. C. Castro, and H. Hosseinpour. *A review of uncertainty visualization errors: Working memory as an explanatory theory*. In K. D. Federmeier (ed.), *Psychology of Learning and Motivation*, volume 74, chapter 7, pp. 275–315. Academic Press, 2021. ISBN 978-0-12-824586-6. doi:10.1016/bs.plm.2021.03.001.

[474] R. Zeng, C. M. Mannaerts, and Z. Shang. *A Low-Cost Digital Colorimetry Setup to Investigate the Relationship between Water Color and Its Chemical Composition*. Sensors 21(20), 6699 (2021). doi:10.3390/s21206699.

[475] J. Zhang, Y. Meuret, X. Wang, and K. A. G. Smet. *Improved and Robust Spectral Reflectance Estimation*. LEUKOS 17(4), 359–379 (2021). doi:10.1080/15502724.2020.1798246.

[476] S. P. Garaba and T. Harmel. *Top-of-atmosphere hyper and multispectral signatures of submerged plastic litter with changing water clarity and depth*. Optics Express 30(10), 16 553–16 571 (2022). doi:10.1364/OE.451415.

[477] D. Leffrang and O. Müller. *Should I Follow this Model? The Effect of Uncertainty Visualization on the Acceptance of Time Series Forecasts*. In *2021 IEEE Workshop on TRust and EXpertise in Visual Analytics (TREX)*, pp. 20–26. IEEE, New Orleans, Louisiana, USA, 2021. doi:10.1109/TREX53765.2021.00009.

[478] C. F. Manski. *Forming COVID-19 Policy under Uncertainty*. Journal of Benefit-Cost Analysis 11(3), 341–356 (2020). doi:10.1017/bca.2020.20.

[479] T. C. Wilkes, T. D. Pering, A. J. S. McGonigle, J. R. Willmott, R. Bryant, et al. *The PiSpec: A Low-Cost, 3D-Printed Spectrometer for Measuring Volcanic SO_2 Emission Rates*. Frontiers in Earth Science 7, 65 (2019). doi:10.3389/feart.2019.00065.

[480] O. Burggraaff, G. J. J. Talens, J. Spronck, A.-L. Lesage, R. Stuik, et al. *Studying bright variable stars with the Multi-site All-Sky CAmeRA (MASCARA)*. Astronomy & Astrophysics 617, A32 (2018). doi:10.1051/0004-6361/201833142.

[481] K. E. Joyce, C. L. Nakalembe, C. Gómez, G. Suresh, K. Fickas, et al. *Discovering Inclusivity in Remote Sensing: Leaving No One Behind*. Frontiers in Remote Sensing 3, 869 291 (2022). doi:10.3389/frsen.2022.869291.

[482] T. H. Swartz, A.-G. S. Palermo, S. K. Masur, and J. A. Aberg. *The Science and Value of Diversity: Closing the Gaps in Our Understanding of Inclusion and Diversity*. Journal of Infectious Diseases 220(S2), S33–S41 (2019). doi:10.1093/infdis/jiz174.

[483] A. Skarlatidou, A. Hamilton, M. Vitos, and M. Haklay. *What do volunteers want from citizen science technologies? A systematic literature review and best practice guidelines*. Journal of Science Communication 18(01), A02 (2019). doi:10.22323/2.18010202.

[484] R. E. Pandya. *A framework for engaging diverse communities in citizen science in the US*. Frontiers in Ecology and the Environment 10(6), 314–317 (2012). doi:10.1890/120007.

[485] C. B. Cooper, C. L. Hawn, L. R. Larson, J. K. Parrish, G. Bowser, et al. *Inclusion in citizen science: The conundrum of rebranding*. Science 372(6549), 1386–1388 (2021). doi:10.1126/science.abi6487.

[486] C. S. Lowry and K. F. Stepenuck. *Is Citizen Science Dead?* Environmental Science & Technology 55(8), 4194–4196 (2021). doi:10.1021/acs.est.0c07873.

[487] J. D. Paul, W. Buytaert, S. Allen, J. A. Ballesteros-Cánovas, J. Bhusal, et al. *Citizen science for hydrological risk reduction and resilience building*. Wiley Interdisciplinary Reviews: Water 5(1), e1262 (2018). doi:10.1002/wat2.1262.

[488] J. Hall, M. Gaved, and J. Sargent. *Participatory Research Approaches in Times of Covid-19: A Narrative Literature Review*. International Journal of Qualitative Methods 20, 1–15 (2021). doi:10.1177/16094069211010087.

[489] S. Hornby, J. Benn, R. Vinkenoog, S. Goldberg, and M. J. Pound. *Methods in melissopalynology: colour determination of pollen pellets for colour vision deficient individuals*. Palynology (2022). doi:10.1080/01916122.2022.2062476.

[490] A. Scheeline. *Selfie spectrometry: Why tablets, laptops, and cell phones have not taken over visible spectrometry*. TrAC Trends in Analytical Chemistry 149, 116 564 (2022). doi:10.1016/j.trac.2022.116564.

[491] L. Ceccaroni, J. Piera, M. R. Wernand, O. Zielinski, J. A. Busch, et al. *Citclops: A next-generation sensor system for the monitoring of natural waters and a citizens' observatory for the assessment of ecosystems' status*. PLOS ONE 15(3), e0230 084 (2020). doi:10.1371/journal.pone.0230084.

[492] S. Di Nonno and R. Ulber. *Smartphone-based optical analysis systems*. Analyst 146(9), 2749–2768 (2021). doi:10.1039/d1an00025j.

[493] B. Hunt, A. J. Ruiz, and B. W. Pogue. *Smartphone-based imaging systems for medical applications: a critical review*. Journal of Biomedical Optics 26(4), 040 902 (2021). doi:10.1117/1.jbo.26.4.040902.

[494] A. Scheeline. *Smartphone Technology – Instrumentation and Applications*. In R. Crocombe, P. Leary, and B. Kammrath (eds.), *Portable Spectroscopy and Spectrometry 1: Technologies and Instrumentation*, chapter 9, pp. 209–235. John Wiley & Sons, Ltd., 2021, first edition. ISBN 978-1-119-63648-9. doi:10.1002/9781119636489.ch9.

[495] J. Venkatraman Jagatha, A. Klausnitzer, M. Chacón-Mateos, B. Laquai, E. Nieuwkoop, et al. *Calibration Method for Particulate Matter Low-Cost Sensors Used in Ambient Air Quality Monitoring and Research*. Sensors 21(12), 3960 (2021). doi:10.3390/s21123960.

[496] J. Wesseling, H. de Ruiter, C. Blokhuis, D. Drukker, E. Weijers, et al. *Development and Implementation of a Platform for Public Information on Air Quality, Sensor Measurements, and Citizen Science*. Atmosphere 10(8), 445 (2019). doi:10.3390/atmos10080445.

[497] M. Ottaviani, B. Cairns, J. Chowdhary, B. Van Diedenhoven, K. Knobelspiesse, et al. *Polarimetric retrievals of surface and cirrus clouds properties in the region affected by the Deepwater Horizon oil spill*. Remote Sensing of Environment 121, 389–403 (2012). doi:10.1016/j.rse.2012.02.016.

[498] T. Harmel, O. Burggraaff, S. Garaba, R. Foster, and P. Corradi. *Polarization signatures of nano- and microplastics suspended in the water column simulated at the water surface and top-of-atmosphere levels*. In *ESA Living Planet Symposium (LPS2022)*. Bonn, Germany, 2022. doi:10.13140/RG.2.2.22794.75208.

[499] C. Lane, D. Rode, and T. Rösgen. *Calibration of a polarization image sensor and investigation of influencing factors*. Applied Optics 61(6), C37–C45 (2022). doi:10.1364/AO.437391.

[500] M. K. Kupinski, C. L. Bradley, D. J. Diner, F. Xu, and R. A. Chipman. *Angle of linear polarization images of outdoor scenes*. Optical Engineering 58(8), 082 419 (2019). doi:10.1117/1.OE.58.8.082419.

[501] N. Umov. *Chromatische Depolarisation durch Lichtzerstreuung*. Physikalische Zeitschrift 6(20), 674–676 (1905).

[502] Z. Han, X. Gu, X. Zuo, K. Bi, and S. Shi. *Semi-Empirical Models for the Bidirectional Water-Leaving Radiance: An Analysis of a Turbid Inland Lake*. Frontiers in Environmental Science 9, 818 557 (2022). doi:10.3389/fenvs.2021.818557.

[503] Z. Sun, D. Wu, and Y. Lv. *Effects of water salinity on the multi-angular polarimetric properties of light reflected from smooth water surfaces*. Applied Optics 61(15), 4527–4534 (2022). doi:10.1364/AO.458737.

[504] J. I. Peltoniemi, J. Järvinen, N. Zubko, and M. Gritsevich. *Spectropolarimetric characterization of pure and polluted land surfaces*. International Journal of Remote Sensing 41(13), 4865–4878 (2019). doi:10.1080/01431161.2019.1672219.

[505] M. Ottaviani, B. van Diedenhoven, and B. Cairns. *Photopolarimetric retrievals of snow properties*. The Cryosphere 9(5), 1933–1942 (2015). doi:10.5194/tc-9-1933-2015.

[506] V. J. H. Trees and D. M. Stam. *Blue, white, and red ocean planets: Simulations of orbital variations in flux and polarization colors*. Astronomy & Astrophysics 626, A129 (2019). doi:10.1051/0004-6361/201935399.

English summary

Water is all around us. We use it to drink, wash, play, fish, sail, and much more. Natural waters, like streams, rivers, lakes, seas, and oceans, are full of life and interesting chemistry. Because of its importance, we need to study water intensively, measuring the various compounds and life forms that inhabit it. This way, we can better understand the world around us, our impact on it, and its impact on us.

Oceans and inland waters contain many different *constituents*. These include various chemical compounds, particles, phytoplankton, and pollution. Each constituent in a water body plays a role in its chemistry and biology. For example, phytoplankton produces half of the organic carbon and oxygen in the world, which are vital for all other life forms. By studying the concentrations and properties of constituents in a water body, we can understand its ecosystem. This knowledge has inherent value and also allows us to intervene against pollution, climate change, and other harmful events.

Concentrations and properties of constituents are traditionally measured by taking samples and analysing them in a laboratory. However, this method faces three important challenges. First, it requires expensive equipment and trained personnel, which are not available everywhere nor to everyone. Second, it is difficult to scale these measurements up in time and space. Studying global processes like climate change requires global coverage, but the funds and personnel required to sample every water body regularly are not available. Studying local processes requires speed and local access, which are also not readily available. Third, there is limited standardisation in sampling and analysis protocols between researchers. This causes a large uncertainty on results and makes it difficult to compare results from different groups or locations.

Remote sensing, measuring the light reflected by water, enables global measurements with rapid response times and high consistency. Using different wavelengths and polarisation states, we can determine the concentrations and properties of various constituents. For example, the phytoplankton concentration can be estimated by comparing the reflectance at blue vs. green wavelengths. There is a wide variety of remote sensing instruments on satellites, aeroplanes, ships, and stationary platforms. Recently, consumer cameras like those on drones and smartphones have become popular for low-cost remote sensing.

Citizen science involves non-professionals in the scientific process by taking measurements, interpreting results, and thinking of new research. This provides opportunities to reduce cost and improve scalability by increasing the accessibility of water research. Citizens provide many data and local knowledge that, in collaboration with professional researchers, can lead to new possibilities, shared insights, and tailored local interpretation. At the same time, the citizen scientists learn more about their environment by conducting their own research and they are socially and politically empowered as stakeholders.

The aim of this thesis is to investigate and improve accessibility and uncertainty in remote sensing and citizen science, so that these techniques can better deliver the desired improvements to cost, scale, and reproducibility of water research. *Accessibility* here means the degree to which people can create, use, and interpret data, without being limited by physical ability or financial status. *Uncertainty* refers to the spread in measured values caused by random effects and to errors caused by known or unknown systematic effects.

Chapter 2. Citizen science with colour blindness: A case study on the Forel-Ule scale.
Many citizen science projects depend on colour vision, but up to 1 in 11 participants are colour blind. We simulate the impact of various forms of colour blindness on measurements with the Forel-Ule scale, which is used to measure water colour by eye. We find that colour blindness decreases the ability to distinguish between Forel-Ule colours. This reduces the precision and accuracy of citizen science data and the motivation of participants. These issues can be addressed by including uncertainty estimates in data entry forms and discussing colour blindness in training materials. These conclusions and recommendations apply to colour-based citizen science in general, including other classification and monitoring activities. Being inclusive of the colour blind increases both the social and scientific impact of citizen science.

Chapter 3. Standardised spectral and radiometric calibration of consumer cameras.
Consumer cameras, particularly onboard smartphones and UAVs, are now commonly used as scientific instruments. However, their data processing pipelines are not optimised for quantitative radiometry and their calibration is more complex than that of scientific cameras. The lack of a standardised calibration methodology limits the reproducibility between devices and, in the ever-changing market, ultimately the lifespan of projects using them. We present a standardised methodology and database (SPECTACLE) for spectral and radiometric calibrations of consumer cameras. The methodology is applied to seven popular cameras to characterise their performance and quantify uncertainties. Through calibration and the use of RAW data, we lay the groundwork for using consumer cameras to perform professional-grade spectroradiometry.

Chapter 4. Accuracy and reproducibility of above-water radiometry with calibrated smartphone cameras using RAW data.
In this chapter, we apply the results and recommendations from Chapter 3 to above-water radiometry. We measure R_{rs} and water colour on and around Lake Balaton, Hungary, with two smartphones and two hyperspectral reference instruments for validation. We quantify the uncertainty, reproducibility, and accuracy of the resulting data and compare these to professional spectroradiometers and existing citizen science approaches. We find that smartphone cameras perform similarly to professional instruments in terms of uncertainty, accuracy, and reproducibility. Based on these results, we offer practical recommendations for using consumer cameras in professional and citizen science.

Chapter 5. Biases from incorrect reflectance convolution.
Reflectance measurements from different instruments are converted and compared through spectral convolution. This is done to combine time series, validate instruments, and apply retrieval algorithms. However, convolution is often done incorrectly, with reflectance itself convolved rather than the underlying (ir)radiances. We quantify the resulting error for simulated and real instruments and find biases up to 5%. Based on these results, we suggest that this error was partially responsible for uncertainties found in previous work and recommend that future work apply spectral convolution correctly.

Chapter 6. A universal smartphone add-on for portable spectroscopy and polarimetry: iSPEX 2.
We present a new smartphone spectropolarimeter, iSPEX 2. It succeeds the original iSPEX add-on for citizen science measurements of atmospheric aerosols. The optical design is presented and the manufacturing process is detailed. Through universal smartphone support and a data processing pipeline based on Chapter 3, we aim to improve the accessibility and data quality compared to the original. An initial validation measurement is presented as a proof of concept. Finally, we suggest possible applications of iSPEX 2 for professional and citizen science.

Chapter 7. General discussion and future outlook.
In this chapter, we synthesise the overall findings of Chapters 2–6 and discuss them relative to the current state and future direction of the field. Research in remote sensing of water is focused on obtaining higher-dimensional data by including more wavelengths and polarisation, on automation, and on more in-depth analysis of uncertainty and information content. Citizen science has experienced a boom in the last twenty years, but much work remains to be done. Improvements to equity, diversity, and inclusion will increase the social and scientific impact of citizen science. Further social and technological research is necessary to attain these improvements. Additionally, this chapter contains the initial findings from a follow-up project on spectropolarimetry of floating debris. We observe various types of debris, mostly plastics, in a wave basin under realistic conditions. We find a significant polarisation signal for several types of debris, suggesting that polarisation can be used to detect debris in nature.

Nederlandstalige samenvatting

Water is overal om ons heen. We gebruiken het om te drinken, wassen, spelen, vissen, varen en nog veel meer. Natuurlijke wateren, zoals beken, rivieren, meren, zeeën en oceanen, zitten vol leven en interessante chemie. Omdat water zo belangrijk is, moeten we het intensief bestuderen en de verschillende verbindingen en levensvormen die erin voorkomen, meten. Zo kunnen we de wereld om ons heen, onze invloed daarop en de invloed daarvan op ons beter begrijpen.

Oceanen en binnenwateren bevatten veel verschillende *bestanddelen*. Hieronder vallen diverse chemische verbindingen, deeltjes, fytoplankton en vervuiling. Elk bestanddeel in een waterlichaam speelt een rol in de chemie en biologie ervan. Fytoplankton produceert bijvoorbeeld de helft van de organische koolstof en zuurstof in de wereld; beide stoffen zijn van vitaal belang voor alle andere levensvormen. Door de concentraties en eigenschappen van bestanddelen in een waterlichaam te bestuderen, kunnen we het ecosysteem daarvan begrijpen. Deze kennis heeft een inherente waarde en stelt ons ook in staat om in te grijpen tegen vervuiling, klimaatverandering en andere schadelijke gebeurtenissen.

Concentraties en eigenschappen van bestanddelen worden traditioneel gemeten door monsters te nemen en deze in een laboratorium te analyseren. Deze methode kent echter drie belangrijke uitdagingen. Ten eerste zijn er dure apparatuur en opgeleid personeel voor nodig, die niet overal of voor iedereen beschikbaar zijn. Ten tweede is het moeilijk om deze metingen op te schalen in tijd en ruimte. Het bestuderen van wereldwijde processen zoals klimaatverandering vereist een wereldwijde dekking, maar de middelen en het personeel die nodig zijn om elk waterlichaam regelmatig te bemonsteren, zijn niet beschikbaar. Het bestuderen van lokale processen vereist snelheid en lokale toegang, beide ook niet gemakkelijk beschikbaar. Ten derde zijn de bemonsterings- en analyseprotocollen van onderzoekers slechts in beperkte mate gestandaardiseerd. Dit veroorzaakt een grote onzekerheid in de resultaten en maakt het moeilijk om resultaten van verschillende groepen of locaties met elkaar te vergelijken.

Met *remote sensing* (letterlijk "meten op afstand"), waarbij door water gereflecteerd licht wordt gemeten, kan er wereldwijd met snelle reactietijden en hoge consistentie worden gemeten. Door verschillende golflengtes en polarisatiestaten te gebruiken, kunnen we de concentraties en eigenschappen van verschillende bestanddelen bepalen. Zo kan de fytoplanktonconcentratie bijvoorbeeld worden geschat door de reflectantie bij blauwe en groene golflengtes te vergelijken. Er bestaat een grote verscheidenheid aan instrumenten voor remote sensing, op satellieten, vliegtuigen, schepen en vaste platforms. Recentelijk zijn consumentencamera's zoals die op drones en smartphones populair geworden voor goedkope remote sensing.

Citizen science (letterlijk "burgerwetenschap") betrekt mensen die geen beroepswetenschappers zijn bij het wetenschappelijke proces via het verrichten van metingen, interpreteren van resultaten en het bedenken van nieuw onderzoek. Door de toegankelijkheid van wateronderzoek te vergroten, biedt dit kansen om de kosten daarvan te verlagen en de schaalbaarheid te verbeteren. Burgers verstrekken veel gegevens en lokale kennis die, in samenwerking met professionele onderzoekers, kunnen leiden tot nieuwe mogelijkheden, gedeelde inzichten en lokale interpretatie op maat. Tegelijkertijd leren de *citizen scientists* meer over hun omgeving door zelf onderzoek te doen en worden zij als belanghebbenden sociaal en politiek mondiger.

Het doel van dit proefschrift is om de toegankelijkheid en onzekerheid van remote sens-

ing en citizen science te onderzoeken en te verbeteren, zodat deze technieken beter in staat zijn om de gewenste verbeteringen in kosten, schaal en reproduceerbaarheid van wateronderzoek op te leveren. *Toegankelijkheid* betekent hier de mate waarin mensen gegevens kunnen creëren, gebruiken en interpreteren, zonder beperkt te worden door fysieke mogelijkheden of financiële status. *Onzekerheid* verwijst naar de spreiding in gemeten waarden veroorzaakt door willekeurige effecten en naar fouten veroorzaakt door bekende of onbekende systematische effecten.

Hoofdstuk 2. Citizen science met kleurenblindheid: Een casestudy over de Forel-Ule-schaal.
Veel citizen-science-projecten zijn afhankelijk van kleurwaarneming, maar tot 1 op de 11 deelnemers is kleurenblind. We bootsen het effect na van verschillende vormen van kleurenblindheid op metingen met de Forel-Ule-schaal, die wordt gebruikt om met het oog de kleur van water te meten. We constateren dat kleurenblindheid het vermogen om onderscheid te maken tussen kleuren op de Forel-Ule-schaal vermindert. Dit vermindert de precisie en nauwkeurigheid van citizen-science-data en de motivatie van deelnemers. Deze problemen kunnen worden aangepakt door onzekerheidsschattingen op te nemen in formulieren voor data-invoer en door kleurenblindheid te bespreken in handleidingen. Deze conclusies en aanbevelingen gelden algemeen voor citizen science op basis van kleurmetingen, inclusief andere classificatie- en monitoringsactiviteiten. Het betrekken van kleurenblinde mensen vergroot zowel de sociale als de wetenschappelijke impact van citizen science.

Hoofdstuk 3. Gestandaardiseerde spectrale en radiometrische kalibratie van consumentencamera's.
Consumentencamera's, met name die op smartphones en UAV's, worden tegenwoordig veel gebruikt als wetenschappelijke instrumenten. De dataverwerkingsstromen van deze camera's zijn echter niet geoptimaliseerd voor kwantitatieve radiometrie en hun kalibratie is complexer dan die van wetenschappelijke camera's. Het ontbreken van een gestandaardiseerde kalibratiemethode beperkt de reproduceerbaarheid tussen apparaten en, in de steeds veranderende markt, uiteindelijk ook de levensduur van projecten die ze gebruiken. Wij presenteren een gestandaardiseerde methodologie en database (SPECTACLE) voor spectrale en radiometrische kalibraties van consumentencamera's. De methodologie wordt toegepast op zeven populaire camera's om hun prestaties te karakteriseren en onzekerheden te kwantificeren. Door middel van kalibratie en het gebruik van RAW-data leggen we de basis voor het gebruik van consumentencamera's voor professionele spectroradiometrie.

Hoofdstuk 4. Nauwkeurigheid en reproduceerbaarheid van bovenwaterradiometrie met gekalibreerde smartphonecamera's met behulp van RAW-data.
In dit hoofdstuk passen we de resultaten en aanbevelingen uit Hoofdstuk 3 toe op bovenwaterradiometrie. We meten R_{rs} en waterkleur op en rond het Balatonmeer, Hongarije, met twee smartphones en twee hyperspectrale referentie-instrumenten ter validatie. We kwantificeren de onzekerheid, reproduceerbaarheid en nauwkeurigheid van de resulterende gegevens en vergelijken deze met professionele spectroradiometers en bestaande citizen-science-benaderingen. We constateren dat smartphonecamera's vergelijkbaar presteren met professionele instrumenten wat betreft onzekerheid, nauwkeurigheid en reproduceerbaarheid. Op basis van deze resultaten bieden we praktische aanbevelingen voor het gebruik van consumentencamera's in professionele wetenschap en citizen science.

Hoofdstuk 5. Systematische fouten door onjuiste convolutie van reflectantie.
Reflectantiemetingen van verschillende instrumenten worden omgerekend en vergeleken door middel van spectrale convolutie. Dit wordt gedaan om tijdreeksen te combineren, instrumenten te valideren en *retrieval*-algoritmes toe te passen. Convolutie wordt echter vaak onjuist uitgevoerd, waarbij de reflectantie zelf wordt geconvolueerd in plaats van de onderliggende (ir)radianties. We kwantificeren de resulterende fout voor gesimuleerde en echte instrumenten en vinden systematische fouten tot 5%. Op basis van deze resultaten stellen we voor dat deze fout gedeeltelijk verantwoordelijk is voor de onzekerheden die in eerder werk zijn gevonden en raden we aan dat spectrale convolutie in toekomstig werk op de juiste manier wordt toegepast.

Hoofdstuk 6. Een universeel smartphone-opzetstuk voor draagbare spectroscopie en polarimetrie: iSPEX 2.
We presenteren een nieuwe spectropolarimeter voor smartphones, iSPEX 2. Dit is de opvolger van de oorspronkelijke iSPEX, een opzetstuk voor citizen-science-metingen van atmosferische aerosolen. Het optische ontwerp wordt gepresenteerd en het productieproces wordt gedetailleerd beschreven. Door universele smartphone-ondersteuning en een gegevensverwerkingslijn op basis van Hoofdstuk 3 willen we de toegankelijkheid en gegevenskwaliteit verbeteren ten opzichte van het origineel. Een eerste validatiemeting wordt gepresenteerd als proof of concept. Ten slotte stellen we mogelijke toepassingen van iSPEX 2 voor professionele en burgerwetenschap voor.

Hoofdstuk 7. Algemene discussie en toekomstperspectief.
In dit hoofdstuk synthetiseren we de algemene bevindingen van Hoofdstukken 2–6 en bespreken we ze met betrekking tot de huidige stand van zaken en de toekomstige richting van het vakgebied. Onderzoek in remote sensing van water is gericht op het verkrijgen van data met meer dimensies door meer golflengtes en polarisatie mee te nemen, op automatisering, en op een meer diepgaande analyse van onzekerheid en informatie-inhoud. Citizen science heeft de afgelopen twintig jaar een hoge vlucht genomen, maar er is nog veel werk aan de (wetenschaps)winkel. Verbeteringen op het gebied van gelijkwaardigheid, diversiteit en inclusiviteit zullen de sociale en wetenschappelijke impact van citizen science vergroten. Verder sociaal en technologisch onderzoek is nodig om deze verbeteringen te realiseren. Daarnaast bevat dit hoofdstuk de eerste bevindingen van een vervolgproject over spectropolarimetrie van drijvend afval. We observeren verschillende soorten afval, voornamelijk plastic, in een golfbassin onder realistische omstandigheden. We zien een significant polarisatiesignaal voor verschillende soorten afval, wat suggereert dat polarisatie kan worden gebruikt om afval in de natuur te detecteren.

List of publications

Refereed publications

1. *Studying bright variable stars with the Multi-site All-Sky CAmeRA (MASCARA)*
   Olivier Burggraaff; Geert Jan J. Talens; Julien F. P. Spronck; Anna-Léa Lesage; Remko Stuik; Gilles P. P. L. Otten; Vincent Van Eylen; Don Pollacco; Ignas A. G. Snellen
   Astronomy and Astrophysics, 617, A32 (2018)
   doi:10.1051/0004-6361/201833142

2. *Standardized spectral and radiometric calibration of consumer cameras*
   Olivier Burggraaff; Norbert Schmidt; Jaime Zamorano; Klaas Pauly; Sergio Pascual; Carlos Tapia; Evangelos Spyrakos; Frans Snik
   Optics Express, 27(14), 19075–19101 (2019)
   doi:10.1364/OE.27.019075

3. *Biases from incorrect reflectance convolution*
   Olivier Burggraaff
   Optics Express, 28(9), 13801–13816 (2020)
   doi:10.1364/OE.391470

4. *Citizen science with colour blindness: A case study on the Forel-Ule scale*
   Olivier Burggraaff; Sanjana Panchagnula; Frans Snik
   PLOS ONE, 16(4), e0249755 (2021)
   doi:10.1371/journal.pone.0249755

5. *Accuracy and Reproducibility of Above-Water Radiometry With Calibrated Smartphone Cameras Using RAW Data*
   Olivier Burggraaff; Mortimer Werther; Emmanuel S. Boss; Stefan G. H. Simis; Frans Snik
   Frontiers in Remote Sensing, 3, 940 096 (2022)
   doi:10.3389/frsen.2022.940096

Non-refereed publications

1. *A universal smartphone add-on for portable spectroscopy and polarimetry: iSPEX 2*
   Olivier Burggraaff; Armand B. Perduijn; Robert F. van Hek; Norbert Schmidt; Christoph U. Keller; Frans Snik
   Proceedings of the SPIE, 11389, 113892K (2020)
   doi:10.1117/12.2558562

Co-authored publications

1. *Data calibration for the MASCARA and bRing instruments*
   Geert Jan J. Talens; Erik R. Deul; Remko Stuik; Olivier Burggraaff; Anna-Léa Lesage; Julien F. P. Spronck; Samuel N. Mellon; John I. Bailey; Eric E. Mamajek; Matthew A. Kenworthy; Ignas A. G. Snellen
   Astronomy and Astrophysics, 619, A154 (2018)
   doi:10.1051/0004-6361/201834070

2. *Discovery of δ Scuti Pulsations in the Young Hybrid Debris Disk Star HD 156623*
   Samuel N. Mellon; Eric E. Mamajek; Konstanze Zwintz; Trevor J. David; Remko Stuik; Geert Jan J. Talens; Patrick Dorval; Olivier Burggraaff; Matthew A. Kenworthy; John I. Bailey; Blaine B. D. Lomberg; Rudi B. Kuhn; Michael J. Ireland; Steven M. Crawford
   The Astrophysical Journal, 870(1), 36 (2019)
   doi:10.3847/1538-4357/aaf008

3. *Bright Southern Variable Stars in the bRing Survey*
   Samuel N. Mellon; Eric E. Mamajek; Remko Stuik; Konstanze Zwintz; Matthew A. Kenworthy; Geert Jan J. Talens; Olivier Burggraaff; John I. Bailey; Patrick Dorval; Blaine B. D. Lomberg; Rudi B. Kuhn; Michael J. Ireland
   The Astrophysical Journal Supplement Series, 244(1), 15 (2019)
   doi:10.3847/1538-4365/ab3662

Curriculum vitae

I was born in Zoetermeer, the Netherlands, in 1996, and grew up there and in neighbouring Bergschenhoek. I attended Passe Partout primary school from 2000–2006 and from 2006–2012 I attended Erasmus College Zoetermeer, from which I obtained my vwo/gymnasium high school diploma.

I started my Bachelor's degrees in physics and astronomy at Leiden University in 2012. In my second year, I participated in a research project involving observations of the Sagittarius stellar stream using the Isaac Newton Telescope on La Palma, Spain. I graduated cum laude from both degrees in 2015 with a thesis titled *iSPEX: Quantitative Spectroscopy on a Smartphone*, supervised by Frans Snik and Christoph Keller. In that thesis, we attempted to empirically correct the errors caused by JPEG compression, achieving limited success and laying the groundwork for Chapter 3 five years down the line.

I continued my studies in Leiden by pursuing a Master's degree in Research in Astronomy from 2015–2017. This programme included elective computer science courses, studying data mining techniques and neural networks. For my first research project, I worked with Ignas Snellen on the detection and characterisation of bright variable stars using the Multi-site All-Sky CAmeRA (MASCARA). This project eventually led to a publication in Astronomy & Astrophysics, as listed in the previous section. I worked with Anthony Brown for my second research project, investigating the velocities of stars in the solar neighbourhood using the first Gaia data release (DR1). We used the extreme deconvolution (XD) statistical technique to infer the six-dimensional position-velocity distribution of stars from the five-dimensional DR1 data, which did not include radial velocity measurements. Our results hinted towards previously unknown structures in the stellar velocity distribution, which were later confirmed by other groups using DR2. I graduated cum laude in 2017.

I worked as a teaching assistant (TA) on the *Praktische Sterrenkunde* (Practical Astronomy) course from 2015–2020. This course exposed first-year Bachelor's students to all aspects of astronomical work, including data acquisition with a telescope, data analysis in Python, and academic writing. As a TA, I was involved in every aspect of the course, including teaching practicals, grading reports, and building instrumentation. In 2017, I worked with Frans Snik on a LEGO® replica of the Paranal Very Large Telescope (VLT) complex, commissioned by the European Southern Observatory (ESO).

With the astronomy student association L.A.D. 'F. Kaiser', I organised a series of public lectures on astronomy, the Kaiser Spring Lectures, in 2016. We welcomed >300 guests over the course of four lectures. From 2016–2017 I also served as the chair-treasurer for Kaiser.

In 2018, I started my PhD at Leiden University, working in the astronomy and environmental sciences departments under supervision of Frans Snik and Martina Vijver, with Christoph Keller as an additional promotor. My research (Chapters 1–7) was part of the Horizon 2020 consortium MONOCLE. In 2019, I attended the Ocean Optics course at the University of Maine, sponsored by NASA and EUMETSAT. I have attended and presented at several conferences, including SPIE Defense + Commercial Sensing, PittCon, and RSPSoc's Wavelength conference, and co-organised the 2018 Remote Sensing in Ecology workshop in Leiden. I have served as a peer reviewer for the journals Applied Optics, Biomedical Optics Express, Optics Express, and Photogrammetric Engineering & Remote Sensing. Finally, I have given guest lectures on remote sensing and polarimetry in several courses at Leiden University and the University of Siena.

Acknowledgements

At the time of writing, I have spent over a third of my life at Leiden University. I would like to use this section to thank those who worked with or supported me during this period.

First of all, I would like to thank my supervisors/promotors, Christoph Keller, Martina Vijver, and Frans Snik, for their professional and personal support. I appreciate their great enthusiasm and the many opportunities they have provided me.

I am very grateful to everyone at CML and the Observatory who makes the institutes pleasant to be in and work in, including the housekeeping, scientific, and secretarial staff. I would like to thank all members of the Environmental Biology, Remote Sensing, and Astronomical Instrumentation groups for creating a friendly atmosphere. I am particularly grateful to David Doelman for his invaluable advice during the PhD process and to Juriën Huisman and Rafaël Mostert for making my Bachelor's, Master's, and PhD much more enjoyable.

My research was part of the Horizon 2020 consortium MONOCLE and I would like to thank all participants for the pleasant collaborations, field campaigns, and even committee meetings. I am especially thankful to Jess Heard and Stefan Simis for leading the consortium and for always offering a helping hand.

My primary collaborator for smartphone science has been Norbert Schmidt, whom I would like to thank for his enthusiastic work on the SPECTACLE and iSPEX 2 apps, even as I kept suggesting new features. Equally, I would like to thank Armand Perduijn for all the energy poured into designing and producing iSPEX 2 and into teaching me how to think like an engineer – I'll get it one day! I am also grateful to everyone else at DDQ and BLS for their important contributions, in particular Rob van Hek and Joep van der Heiden.

There have been many other pleasant collaborations in the past four years. I would like to thank Hester Volten and Daan Swart of RIVM, Jeroen Rietjens and Martijn Smit of SRON, and Gerard van Harten of JPL for advising and collaborating on various SPEX instruments. I am particularly indebted to my former student Aaron Seymour who single-handedly revived the groundSPEX instrument after many years in an attic (the instrument, not the student). I am grateful to Shungu Garaba, Tristan Harmel, and Paolo Corradi for inviting me into OP^3, to Remko Stuik for designing the groundSPEX goniometer, and to the participants of the ESA measuring campaigns, particularly Wout Bakker and Anton de Fockert of Deltares and Suhyb Salama of ITC. Lastly, I would like to thank Arie van der Meijden of Naturalis for collaborating on the beetle scanner project.

I am very grateful to have been given the opportunity to attend the 2019 Ocean Optics course at the University of Maine, funded by NASA and EUMETSAT. I would like to thank the professors, teaching assistants, classmates and research interns, and the Darling Marine Center staff for providing a highly educational and enjoyable, and utterly gruelling, experience. Most of all, I would like to thank Emmanuel Boss and Mortimer Werther for advising and collaborating on much of my work since then.

I would like to thank the teachers/professors who led me here. At school, these were Dirk Kuilman, Jaap Focke, Meindert Boskma, and Richard van Rooijen. From Leiden, I am particularly grateful to Ignas Snellen and Anthony Brown for supervising my Master's theses. Lastly, I would like to thank Michiel Hogerheijde, Huib Jan van Langevelde, Maarten van 't Zelfde, and Steven Loiselle for providing me with valuable teaching experience.

Additionally, I would like to thank my Leiden friends Coen, Corjan, Daan, Eveline, Teun, and Tim, as well as my online friends Colin, Kyrylo, Steele, Tom, and Vitus for their great

personal and occasionally professional support over the years. I would also like to thank my pal the tortoise, Aku, who lives the life of a madman. Most of all I would like to thank Sanj for all her support, all the new perspectives she has given me, and for the music she has added to my life over these past years.

Finally, I would like to thank my family. My siblings Sebastiaan and Carlijn have always been supportive and have taught me much, especially the art of wordplay. I am a proud uncle to Eléonora and Willem, who I hope will one day read this. I would like to thank my late grandfather Thom, who enthusiastically read every essay I wrote for my Bachelor's courses. My father, Jan Karel, has been instrumental in feeding my insatiable curiosity since I was little, which eventually led to this thesis and for which I am very grateful. Lastly, I would like to thank my mother, Karin, who unfortunately did not get to see me finish my PhD, for all her loving support over the years.

I conclude this thesis with a paraphrase from *Cynegeticus*[29], written almost 2000 years ago by Arrian of Nicomedia:

> For I myself, you know, have raised a hound with eyes as grey as the greyest, and he is both fast and diligent and of good spirit and most gentle. [...] And so I think I should not hesitate to record the name of the hound, because truly I have a most swift, most wise, and most wonderful hound, whose name is Dorus.

[29] Following the translation in *Lives in Interaction: Animal 'Biographies' in Graeco-Roman Literature?*, Thorsten Fögen (2017), doi:10.1515/9783110545623-005.