Uncertainty assessment of unattended above-water radiometric data collection from research vessels with the Dynamic Above-water Radiance (L) and Irradiance (E) Collector (DALEC)

DAVID ANTOINE,[1,2,*] MATTHEW SLIVKOFF,[3] WOJCIECH KLONOWSKI,[3] CHARLES KOVACH,[4] AND MICHAEL ONDRUSEK[5]

[1]*Remote Sensing and Satellite Research Group, School of Earth and Planetary Sciences, Curtin University, Perth, WA 6845, Australia*
[2]*Sorbonne Université, CNRS, Laboratoire d'Océanographie de Villefranche, LOV, F-06230 Villefranche-sur-Mer, France*
[3]*In-situ Marine Optics Pty Ltd., Bibra Lake, WA 6163, Australia*
[4]*Global Science & Technology, Inc., Greenbelt, Maryland 20770, contractor to NOAA / NESDIS / STAR, College Park, MD 20740, USA*
[5]*NOAA / NESDIS / STAR, College Park, MD 20740, USA*
*david.antoine@curtin.edu.au

Abstract: We used above- and below-water radiometry measurements collected during a research voyage in the eastern Indian Ocean to assess uncertainties in deriving the remote sensing reflectance, R_{rs}, from unattended above-water radiometric data collection with the In-Situ Marine Optics Pty. Ltd. (IMO) Dynamic Above-water Radiance (L) and Irradiance (E) Collector (DALEC). To achieve this, the R_{rs} values derived from using the latest version of this hyperspectral radiometer were compared to values obtained from two in-water profiling radiometer systems of rather general use in the ocean optics research community, i.e., the Biospherical Instruments Inc. Compact Optical Profiling System (C-OPS) and the Seabird HyperPro II. Our results show that unattended, carefully quality-controlled, DALEC measurements provide R_{rs} for wavelengths < 600 nm that match those derived from the in-water systems with no bias and a dispersion of about 8%, provided that the appropriate technique is used to quantify the contribution of sky light reflection to the measured signal. The dispersion is larger (25-50%) for red bands, which is expected for clear oligotrophic waters as encountered during the voyage, where $\sim 2\ 10^{-5} < R_{rs} < \sim 2\ 10^{-4}\ \mathrm{sr}^{-1}$. For comparison, the two in-water systems provided R_{rs} in agreement within 4% for wavelengths < 600 nm.

1. Introduction

Producing reliable geophysical products from satellite ocean color radiometry (OCR) requires a sustained effort to collect validation data in the field [1,2]. This is needed in particular for the primary quantity that such missions deliver, i.e., the spectral remote-sensing reflectance, $R_{rs}(\lambda)$, where λ is wavelength.

Multiple efforts have led to the creation of databases collating such measurements, which are then used by space agencies to provide global validation statistics for their missions' products. Examples include the National Aeronautics Space Administration (NASA) Sea-viewing Wide Field-of-view Sensor Bio-Optical Archive and Storage System (SeaBASS) [3], the NASA bio-Optical Marine Algorithm Dataset (NOMAD) [4] and the European Space Agency (ESA)

MEdium Resolution Imaging Spectrometer (MERIS) Matchup In-Situ Database (MERMAID) [5].

Ensuring consistency of the data contributed to these archives is challenging, and significant effort is accordingly put in quality controlling the submitted data. Still, it cannot be ascertained whether all submitted data were collected following recommended data acquisition protocols, and what processing steps were followed when deriving quantities such as R_{rs} from basic radiometric quantities. Recognizing this difficulty inherent to collating data from multiple users, and aiming at improving consistency among data used for validation purposes, the international research community and space agencies defined the concept of Fiducial Reference Measurements (FRM, [6]). The FRM definition is that of a suite of independent, fully characterized, and traceable ground measurements that follow the guidelines outlined by the Group on Earth Observations (GEO) / Committee on Earth Observation Satellites (CEOS) Quality Assurance framework for Earth Observation (QA4EO).

The FRM concept emphasizes in particular the need for (data) measurements to be delivered with an uncertainty, and a description of how it was computed. The aim is to allow a critical assessment of whether or not a given data item can be used for validation purposes, in particular for OCR [7]. For instance, using measurements from a single *in situ* sensor for validation of satellite observations actually does not make much sense if the uncertainty on the *in situ* measurement is unknown. If a bias is observed, how much of it is due to a possible issue with the satellite product and how much of it might actually be due to issues with the *in situ* measurements? This situation is actually quite the rule in validation papers, with uncertainties being rarely reported.

Uncertainties in R_{rs} can stem from 1) instrument characterization and calibration, 2) the degree of adherence to recommended data acquisition protocols, 3) environment conditions during the measurements (which to some extent is related to point 2) and, 4) from data processing techniques. Radiometric uncertainty for field-deployable radiometers can generally be maintained low, at least if a number of characterization and calibration steps are performed [8]. Uncertainties due to the data collection protocols and data processing algorithms can be large, however [9].

A possible approach is to assign values to all potential sources of uncertainties, whether empirically or through laboratory experiments, and then to combine these individual values into a final uncertainty budget [10,11]. Here we follow a more pragmatic, empirical, approach by which we attempt an uncertainty assessment based on comparing two in-water and one above-water instruments. The two in-water instruments have been extensively used across the OCR community for quite a while. They are the Biospherical Instruments Inc. Compact Optical Profiling System (C-OPS) and the Seabird (formerly from Satlantic Inc.) HyperPro II. The above-water system is newer. It is the In-Situ Marine Optics Pty. Ltd. (IMO) Dynamic Above-Water Radiance (L) and Irradiance (E) Collector (DALEC).

This is not the first time this sort of comparisons were performed, however, either among above-water systems [12] or between above- and in-water system (e.g. [13,14]), yet it is the first time for the new DALEC sensor, for which we therefore provide more detailed technical information than for the two other instruments. The new features of the DALEC include 1) an improved cosine response for the irradiance collector, 2) channel-specific temperature sensors and temperature compensation, 3) upgraded tilt sensors and, 4) GYRO-stabilized heading input from the vessel (see later the instrument description and the Supplement 1). A first illustration of the DALEC capabilities was provided by [15], yet a number of instrument characteristics have changed since that 2016 paper, which also did not include a validation of the DALEC with respect to other instruments commonly used in the marine optics community.

Another objective of our study is to further document the in-water vs. above-water differences (or attended vs. unattended), which have been a number of times assessed for other instruments, and recently in the frame of the Fiducial Reference Measurements For Satellite Ocean Colour

(FRM4SOC) project [16]. Here we consider the case of an unattended instrument collecting data in continuous mode from a fixed position on the ship and adapting its viewing azimuth angle automatically from information provided by a compass (either internal to the instrument or from the ship's instrumentation). The constraints are therefore different from a situation where, for instance, the instrument would be manually oriented with respect to the sun by an operator. The three instruments we used differ in their technology, deployment techniques, and data processing. For the above-water system, we evaluate various corrections for the reflection of skylight at the air-sea interface.

2.   Theoretical background

A primary product of satellite ocean color missions is the spectral remote-sensing reflectance, $R_{rs}(\lambda)$, where λ is wavelength. All geophysical products are derived from various combinations of R_{rs} at several spectral bands, e.g., the chlorophyll concentration is derived through ratios or differences of bands in the blue and green regions of the spectrum [17,18]. $R_{rs}(\lambda)$ is defined as:

$$R_{rs}(\lambda, \theta_s, \theta_v, \phi) \;=\; L_w(\lambda, \theta_s, \theta_v, \phi) \,/\, E_s(\lambda, \theta_s), \tag{1}$$

where $E_s(\lambda, \theta_s)$ is the above-water downward (planar) irradiance for a sun zenith angle θ_s, and $L_w(\lambda, \theta_s, \theta_v, \phi)$ is the water-leaving radiance for θ_s, a nadir angle θ_v, and an azimuth difference between the sun and measurement vertical planes ϕ. The angles θ_s, θ_v and ϕ are ignored in the following (but see later, normalization to nadir and a sun at zenith).

   Validating satellite-derived $R_{rs}(\lambda)$ values requires field measurements of $E_s(\lambda)$ and of other radiometric quantities that can be used to derive $L_w(\lambda)$, which is a quantity not normally directly measured. The measurement of E_s (wavelength dependencies are now omitted in the following unless otherwise noted) is made with an above-surface (deck) irradiance sensor equipped with a cosine collector, installed in such a way that perturbations from the ship's superstructure are minimized if not completely avoided.

   Two classes of techniques have been used historically to determine L_w. They are briefly summarized here, and the reviews in [19,20] offer further reading and ample details on the subject. Another technique was recently proposed, referred to as the skylight-blocked approach (SBA) [21,22] yet is not discussed here. The way these methods have been specifically applied for the three instruments we used here is separately addressed for each instrument in the next section.

   The in-water techniques consist in measuring the vertical profile of the underwater upwelling radiance at nadir, $L_u(z)$, starting from the shallowest possible depth. The depth profile of log-transformed $L_u(z)$ from some deeper layers to this near-surface level will be used to derive a diffuse attenuation coefficient for L_u, then used to extrapolate L_u to the "0^-" depth, which is the theoretical level just beneath the air-sea interface. These measurements must be performed at sufficient distance from the ship so as to avoid any ship shading perturbation.

   The Fresnel and Snell laws are then applied to transfer the radiance across the sea-air interface, such that:

$$L_w \;=\; L_u(0^-) \,(1 - \rho(\theta)) \,/\, n^2 \tag{2}$$

where $\rho(\theta)$ is the Fresnel reflection coefficient for the water-air interface at angle θ and n is the refractive index of seawater.

   The second class of techniques consists in aiming a radiance radiometer at the sea from above the surface at an oblique angle (generally 40° from nadir), then measuring a total radiance, L_t, that is made of L_w plus reflection by the air-sea interface of the direct sun and diffuse sky radiances. Assuming that the reflection of direct sun light is avoided by selecting an appropriate measurement geometry and by further filtering the measurements, the total radiance, L_t is:

$$L_t \;=\; L_w + \rho L_s, \tag{3}$$

where L_s is the diffuse sky radiance and ρ is the reflection coefficient at the air-sea interface.

The critical processing step then consists in determining [ρ L_s] so as to finally get L_w. One possible technique to do so consists in measuring L_s at the complementary angle to the sea viewing angle (so, generally 40° from zenith in the same plane as the L_t measurement) and then to set a ρ value or estimate it through some algorithm [e.g., 23,24]. This is the technique used here for the DALEC data acquisition and processing (see later). The uncertainty in determining ρ originates from its dependence on the viewing geometry (hence the critical need to keep that geometry within predefined limits) and on the sky and sea conditions.

3. Data acquisition plan, instruments and data processing

3.1. 110E-line research voyage

The 110E-line research voyage took place 14th May – 13th June 2019 onboard the Australian Marine National Facility (R/V Investigator), starting from and returning to Fremantle, Western Australia (Fig. 1).

This voyage was organized under the umbrella of the second International Indian Ocean Expedition [25]. The overarching objective was to revisit the 110°E line that was last sampled intensively in the 1960s as part of the first International Indian Ocean Expedition [26], in search of changes and trends in physical and biological oceanographic parameters. Twenty 15-hour stations were occupied along the 110°E meridian (Fig. 1 and Table 1). Other stations and underway sampling took place on the way back from station 20 to Fremantle, but did not include in-water radiometry.

Stations south of 35°S exhibited mesotrophic conditions, with surface (<5 m) chlorophyll concentrations from about 0.2 to 0.5 mg(Chl) m^{-3} (determined through High Precision Liquid Chromatography following [27]). A rapid transition towards oligotrophic conditions occurred from station 5, reaching minimal chlorophyll concentrations of about 0.04 mg(Chl) m^{-3} at station 14.

The sky was rather cloudy for stations 1-7, then generally clear up to station 20. When feasible, the aerosol optical thickness at 870 nm ($\tau_a 870$) was determined using a handheld Microtops sun photometer. The data were processed by the AERONET project [28]. These measurements revealed extremely clear atmospheres with $\tau_a 870$ always below 0.1.

3.2. In-situ marine optics DALEC

3.2.1. DALEC design, radiometric characterization and calibration

The DALEC instrument contains three Zeiss Monolithic Miniature Spectrometers (MMS1) in a single enclosure (Fig. 2(a)), and they collect E_s, L_t and L_s spectra in a simultaneous fashion. A 2-axis gimbal hosts the DALEC so that the recommended measurement geometry is respected as far as possible, i.e., the sea-viewing channel aiming at surface with an angle of 40 degrees from nadir, the sky-viewing channel aiming at the sky at an angle of 40 degrees from zenith, and the downward irradiance collector being horizontal [29] (Fig. 2(b)).

The DALEC can be programmed to automatically rotate about the Zenith - Nadir axis in order to achieve a user-defined sun-relative azimuth viewing angle, whilst also considering manually set angular limits. This automated feature helps minimize sun-glint [e.g., 23] and avoid ship shadows and other structural perturbations throughout the transit of the sun, allowing the user to concentrate on other tasks during the field campaign. The DALEC has an internal compass, pitch and roll sensors, GPS and temperature sensors for each spectrometer. The DALEC can also record and use an ancillary RS232 serial connection or User Datagram Protocol (UDP) broadcast National Marine Electronics Association (NMEA) gyro heading to avoid potential pointing errors induced by large steel research vessels. Such a design with collocated E_s, L_t and L_s measurements avoids (minimizes) inconsistencies among them, which sometimes happen with other setups where the E_s sensor is installed in such a way that it might be recording a downward
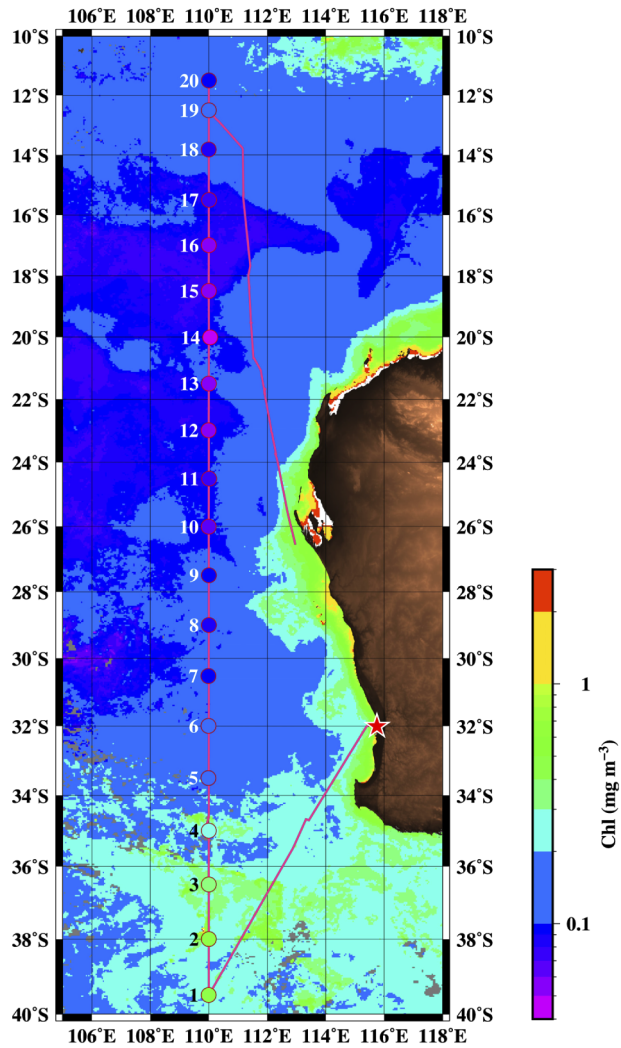
Fig. 1. The ship track (pink line) and the 20 15-hour stations (circles) occupied during the 110E-line research voyage off Western Australia, starting from and returning to Fremantle (Red star). The track is overlaid on the Aqua-MODIS monthly chlorophyll composite map for May 2019. The color inside the circles corresponds to the measured TChl-a (HPLC) at that station using the same color scale as used for the background.

**Table 1. The 20 stations positions and main characteristics**

St #	Date	hour[a]	Lat. (S)[c]	Sun zenith angle	$\tau_a 870$	Sky	Wind speed	Surface TChl-a	C-OPS[d]	HyperPro[d]
	UTC		Decimal degrees	Degrees			m s^{-1}	mg m^{-3}		
1[b]	17-May,	4:45	39.499	58.8	-	cloudy	4.7	0.42	1	10
2[b]	18-May,	4:11	38.000	57.8	-	cloudy	8.2	0.44	3	8
3[b]	19-May,	1:54	36.500	67.7	-	cloudy	12.7	0.31	0	0
4[b]	20-May,	2:28	35.030	60.4	-	cloudy	12.1	0.20	0	0
5[b]	21-May,	3:00	33.500	58.2	-	cloudy	6.9	0.16	4	14
6[b]	22-May,	4:09	32.000	52.7	-	cloudy	3.2	0.14	2	3
7[b]	23-May,	4:38	30.504	51.0	-	cloudy	5.2	0.09	3	6
8	24-May,	4:47	29.000	49.8	-	cloudy	9.9	0.09	3	7
9	25-May,	5:04	27.501	48.8	-	clear	9.2	0.09	3	9
10	26-May,	3:36	26.013	49.2	0.060	clear	9.8	0.06	5	11
11	27-May,	3:50	24.528	47.1	0.077	clear	4.8	0.07	4	13
12	28-May,	4:26	23.000	44.6	-	clear	2.0	0.05	3	6
13	29-May,	3:10	21.500	47.9	0.040	clear	5.3	0.06	8	24
14	30-May,	2:45	20.017	49.6	0.050	clear	5.1	0.05	7	36
15	31-May,	4:13	18.500	40.7	0.068	clear	7.9	0.05	3	7
16	1-Jun,	3:54	16.999	40.3	0.065	clear	12.0	0.06	3	14
17	2-Jun,	3:46	15.500	39.6	-	clear	9.	0.08	4	11
18	3-Jun,	4:15	14.009	36.6	0.087	clear	9.10	0.08	3	14
19	4-Jun,	3:42	12.501	37.4	-	clear	6.9	0.16	3	15
20	5-Jun,	4:10	11.503	34.6	0.076	clear	5.6	0.09	3	10

[a]This is the time of the first C-OPS profile.
[b]Not included in the comparison, either because no in-water profiling occurred (stations 3 and 4) or because of a cloudy sky.
[c]Longitude is not indicated here. It was 110°E for all stations at ± 0.01 decimal degree.
[d]Number of individual C-OPS profiles and number of HyperPro profiles used in the multicast technique
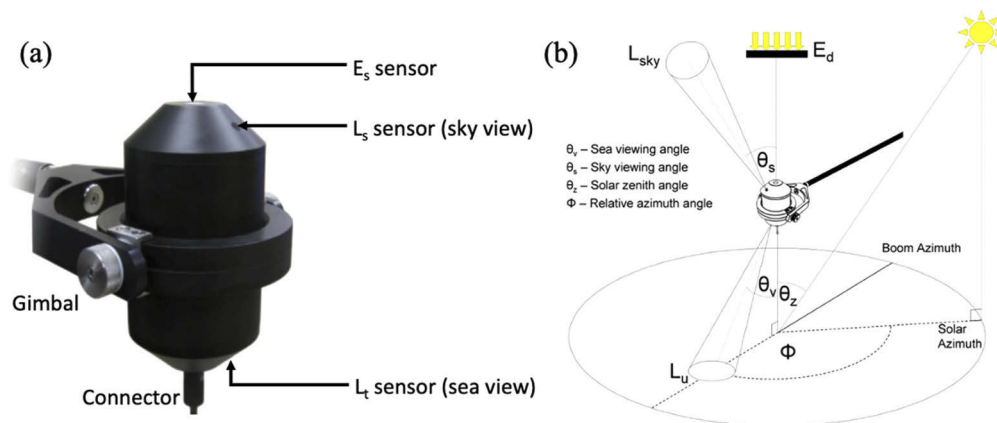
Fig. 2. (a) Picture of the DALEC instrument showing its main elements. The housing diameter is 14 cm and the height of the main enclosure is 25 cm. (b) DALEC measurement geometry.

irradiance at a different location to where the L_t measurement is performed (similarly with E_s and L_u measurements for in-water systems).

The spectrometers in the DALEC are Zeiss MMS1 Ultraviolet / Visible (UV/VIS) enhanced modules, which collect photons from ~305 to ~1200 nm across 256 pixels. Each pixel has a spectral width of approximately 10 nm, and the pixel spectral spacing is approximately 3.3 nm. The spectrometers are integrating-type, so the user defines the integration time during which to collect photons. The photon energy is digitized at the end of the integration period with a 16-bit Analog-to-Digital Converter (ADC). The DALEC can be programmed to optimize the integration times to utilize the desired span of the ADC's dynamic range, or these can be fixed to provide fixed time-period integrations.

Further details on the DALEC design and radiometric characterization and calibration are provided in the Supplement 1 to this paper, in particular Figs. S1, S2 and S3.

3.2.2. Deployment setup

Above-water radiometer systems are generally installed at the bow of the ship, either by attaching them to the handrail or to a meteorological mast when accessible or, when this proves impractical, through installation of a dedicated platform [e.g., 19]. None of these options were allowed on the R/V Investigator, and instead the DALEC was installed on a 12-m deployable boom on the port side (Fig. 3).
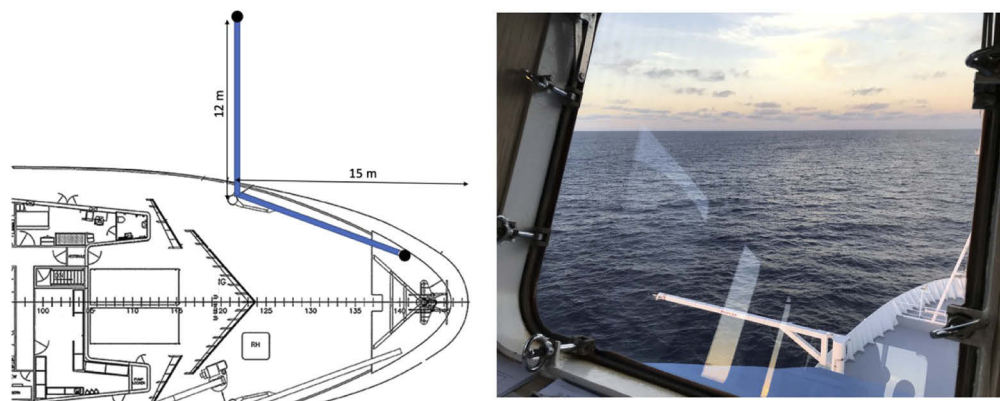
Fig. 3. DALEC set up onboard R/V Investigator. The left panel shows a map of the bow area with the boom on port side colored blue and displayed either deployed perpendicular to the ship or stored along the handrail. The DALEC is represented by the black circle. The right panel is a picture taken from the bridge and showing the boom fully deployed.

In this configuration, the instrument was about 15 m above the sea surface, 20 m from the ship's axis and 15 m behind the bow. This position would have allowed being away from the surface foam generated by the ship's bow wave, which here was not an issue because measurements were done on station. The DALEC has a field of view of 5 degrees, which, in this configuration, corresponded to an approximately 2-m diameter footprint at surface. When the sun was within the 90 degrees quadrant on port side starting from the ship stern, the viewing azimuth angle of the DALEC was basically within 10° of the ship heading, resulting in a sun-view azimuth difference from about 90 to 140 degrees, which follows the recommended protocol [19,24]).

The instrument was recording data continuously every day from dawn to dusk.

3.2.3. Data reduction

An example of calibrated DALEC E_s, L_s and L_t recorded over an entire day is shown on Fig. 4. The data are not usable in the early and late hours of the day, when the sun is low on the horizon

and multiple reflections on or shading by the ship superstructure can occur. This is responsible for the variability in the E_s and L_s time series in the morning and evening, also reinforced in the morning by clouds, which eventually passes or dissipates towards solar noon.
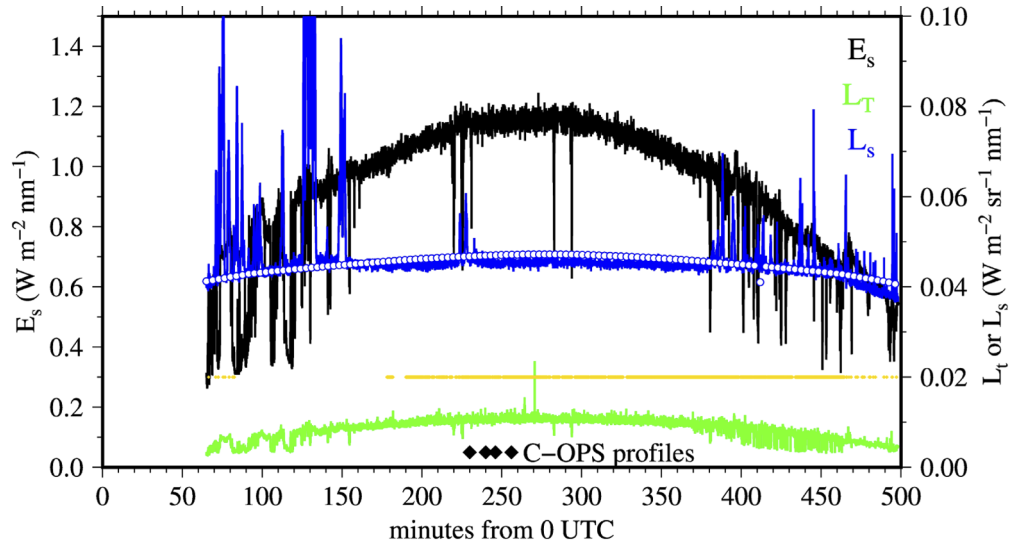
Fig. 4. Time series of E_s, L_t and L_s from the DALEC for station 11 and $\lambda = 443$ nm. The white circles are theoretical values of L_s computed for a maritime aerosol model with a relative humidity of 70% [30] and an optical thickness of 0.03 at 865 nm (radiative transfer computations as per [31]). The gold line indicates when the relative azimuth was maintained within ± 2.5 degrees around 135 degrees. The black diamonds indicate the times of the successive C-OPS profiles.

For the purpose of comparing the above-water radiometry with in-water techniques, calibrated DALEC E_s, L_t and L_s data were first extracted from the full time series based on the individual in-water radiometric profile start and stop times. Depending on the cloud conditions, between two and seven in-water profiles were performed on each station, each with approximately 3 minutes duration (Table 1).

For each profile time period, the DALEC data were screened to remove data collected at extremes in relative solar azimuth angle where increased surface reflection contamination ($80 < \Phi_{az}$) or potential ship shadow ($\Phi_{az} < 170$) can occur. Although the chosen 90 degrees boom angle mount greatly minimizes the risk of ship shadow in our setup, other deployments may need to shift this threshold closer to 150 degrees. In addition, DALEC sensor tilts greater than 5 degrees were rejected to ensure accuracy in the E_s measurements and complementary view angles in the L_t and L_s channels.

With these first stage QC filters applied, usually greater than 80% of the full day's spectra was kept, leaving just the remaining extraction to be governed by the profile start time and 3- minute duration.

The start times of E_s, L_s and L_t are all synchronized, but each spectrometer can integrate over different time scales. Therefore, the time series of each product were temporally interpolated to coincide with the middle of the sea-viewing L_t channel's integration, ready for spectral interpolation. At that stage, data are still at their full acquisition frequency.

3.2.4. R_{rs} determination

After inappropriate geometries have been filtered out of the data records, the next step is to determine [ρ L_s]. An extensive literature is available on that subject [e.g., 32,33], yet no general consensus has been reached on a ubiquitous method. The specific case of unattended measurements was addressed in particular by [34]. The simultaneously recorded L_t, L_s and E_s triplet information from the DALEC allows the calculation and assessment of *instantaneous* reflectance products, however the relatively high sampling rate also allows data to be treated statistically using percentiles at the expense of temporal resolution. This offers a number of possible processing permutations which can and should be assessed with more rigor in subsequent studies.

To minimize the scope, 5 different methods were assessed in this study, as summarized in Table 2. We do not claim that these methods are necessarily the best ones for determining [ρ L_s], although they are basically the same as previously used methods (at least in terms of the selected ρ values; the statistical treatment is seldom reported). We simply try to assess the uncertainty on the final R_{rs} values that the use of these various methods implies.

Table 2. Details of the methods used to derive R_{rs} from L_t, L_s and E_s

Method	Instantaneous Processing	Statistical treatment	Residual Glint
No correction	L_t	Lowest 5%, normalised by |Es|	
1	$(L_t - 0.022 * L_s) / E_s$	Lowest 5%	
2	$(L_t - 0.022 * L_s) / E_s$		$\|R_{rs}(720\text{-}900)\| = 0$ All methods
3	$(L_t - 0.028 * L_s) / E_s$		
4	$(L_t - \rho(\theta_s, \phi, wind) * L_s) / E_s$	Mean (methods 2 to 5)	
5	$(L_t - \text{Opti}\rho * L_s) / E_s - \text{Opti}\Delta$		

For each method, processing starts with the step referred to as instantaneous processing in Table 2, during which instantaneous R_{rs} are computed for each individual set of L_t, L_s and E_s during the considered time window. This is followed by the statistical treatment of all the instantaneous values determined previously (second column in Table 2). Using the mean of the lowest 5 percent is an attempt to eliminate most direct sun glint contamination of L_t. Finally, residual glint is subtracted the same way for all methods by setting the mean R_{rs} between 720 and 900 nm to zero, which assumes spectrally flat residual glint (and possibly foam) reflectances. This step is therefore not repeated in the subsequent descriptions of the 5 methods. It should also be noted that this residual glint correction assumes zero reflectance in the near infrared. This assumption would have to be reconsidered in coastal and eutrophic waters, and different approaches accordingly used for this residual correction.

A first processing uses no ρ skylight correction, however the mean of the lowest 5 percent of the obtained L_t distribution is calculated and normalized by the mean E_s encountered to compute R_{rs}. This approach was selected to validate the assumption that ignoring Fresnel reflection would lead to under-correction (hence overestimated R_{rs}). It basically sizes the [ρ L_s] quantity that methods 1-5 have to assess.

1- Method 1 uses a spectrally constant ρ = 0.022 as recommended by Lee et al. [35] (see also [36]), and then the mean of the lowest 5 percent of the obtained R_{rs} distribution is calculated.

2- Method 2 also uses a constant ρ = 0.022 however the full mean of the obtained R_{rs} distribution is calculated.

3- Method 3 is as Method 2 except that it uses a spectrally flat ρ = 0.028, as recommended by Mobley [23] yet only for wind speeds < 5 m s^{-1}. Although most of the wind speeds encountered were higher than 5 m s^{-1}, this approach was included for historical significance and also to see if using this value leads to under-correction (higher R_{rs}) in higher wind state conditions.

4- Therefore the alternative approach that Mobley also proposed is here our method 4, where ρ is determined as a function of sun-relative azimuth angle, solar zenith angle and wind speed determined from a lookup table [29] computed using the HydroLight radiative transfer code [37]. Another important distinction is that this ρ lookup table includes the spectral dependency which is inherent in ρ but is often overlooked [35,38].

5- Method 5, finally, utilizes a spectral optimization technique to model skylight and sun-glint artifacts by fitting a variable, spectrally independent ρ and spectrally independent near infrared offset Δ to the data between 720 and 900nm, again with the assumption that uncontaminated R_{rs}(720-900) = 0. The optimization is implemented using Levenberg-Marquardt optimization [39], with ρ being allowed to vary between 0.02 and 0.2. The near infrared offset Δ was allowed to vary between -0.01 and 0.1.

3.3. Biospherical C-OPS

3.3.1. Instrument and the deployment setup

The Biospherical Instruments Inc. (San Diego, California) Compact Optical Profiling System (C-OPS) is a multispectral radiometer built from assembling 19 microradiometers, each of which corresponding to a given spectral band. The microradiometer technology and the overall design and operation of the C-OPS were described in detail in [40,41] and examples of its performance can be found in [42].

The instrument we used here is equipped with the following 19 central wavelengths: 340, 412, 443, 465, 490, 510, 532, 555, 560, 589, 625, 665, 670, 683, 694, 710, 765, 780 and 875 nm. Each of these bands has a spectral width of about 10 nm. None of the near infrared bands are considered in the present study that uses data from clear oceanic waters where L_w is negligible for λ > 700 nm. The instrument was last calibrated by the manufacturer in February 2018 and then in November 2019, i.e., 5 months after this voyage. Responsivity changes between these two calibrations were < 3% for the E_s sensor and < 1% for the L_u sensor except for the band at 490 nm. No particular behavior of that band was observed in our data set, however, so we assumed the change occurred after the instrument was used for this Indian Ocean cruise.

A standard deployment protocol was followed where the instrument is kept at surface by keeping a moderate tension in the cable while moved away from the ship such as to avoid ship shading perturbation (in our case a minimum of 50 m), and then allowed to free fall until some depth is reached where data acquisition is stopped and the instrument pulled back at surface. The ship was positioned so that the sun was within the 90 degrees quadrant from the ship's stern to port side, ideally 45 degrees from the ship's axis. Care was also taken to avoid potential contamination by propeller-generated bubble clouds in the ship's wake. A minimum of three profiles were performed at each station, systematically targeting periods of stable above-surface irradiance.

The reference radiometer measuring E_s was installed on top of a crane deployed before each measurement session and located 5 m from the ship's stern on port side. In this configuration the E_s radiometer was 8.5 m above deck (so 12 m above water), and above any structure located in its immediate vicinity (e.g., the A-frame), and about 9 m below the top of the highest ship's superstructure located 40 m further ahead towards the bow. With the sun positioned as described above, minimal perturbations from reflection on the superstructure are to be expected yet cannot be completely ruled out. The instrument was not gimbaled.

3.3.2. Data processing

The microradiometers were set up with a 15 Hz acquisition frequency, and the C-OPS buoyancy was adjusted so the profiling speed was kept slow (typically 15 cm s^{-1} within the first meters below the surface), allowing vertical profiles of E_d and L_u to be collected with a high depth resolution, in particular close to the surface. This feature, combined with the small size of the

deployment frame, allows reducing uncertainties in deriving $L_u(0^-)$ through extrapolation of $L_u(z)$. The shallowest depth for which valid $L_u(z)$ data are obtained is generally around 50 cm, which is roughly the height of the instrument.

Our extrapolation procedure uses a local polynomial regression fitting (LOESS [43,44]), which is a non-parametric method usually employed to smooth time-series. This technique computes polynomials on a moving window. Here the window is defined in time units, and is 40 s for the $L_u(z)$ and 80 s for E_s. With a descent speed of about 15 cm s^{-1}, this time window for $L_u(z)$ corresponds to about 6 m.

This technique is used to avoid a generally subjective selection of a depth interval over which an attenuation coefficient for the upwelling radiance, K_L, is computed from the vertical profile of log-transformed $L_u(z)$ and then used to extrapolate $L_u(z)$ to the 0^- level. This log-linear extrapolation technique is still the most commonly used in field radiometry (e.g., [45]). Before the E_s values can be used to normalize the under-water $L_u(z)$, unwanted E_s oscillations due to ship's tilt are eliminated by using the Loess fitting. An example of this processing in given in Fig. 5.
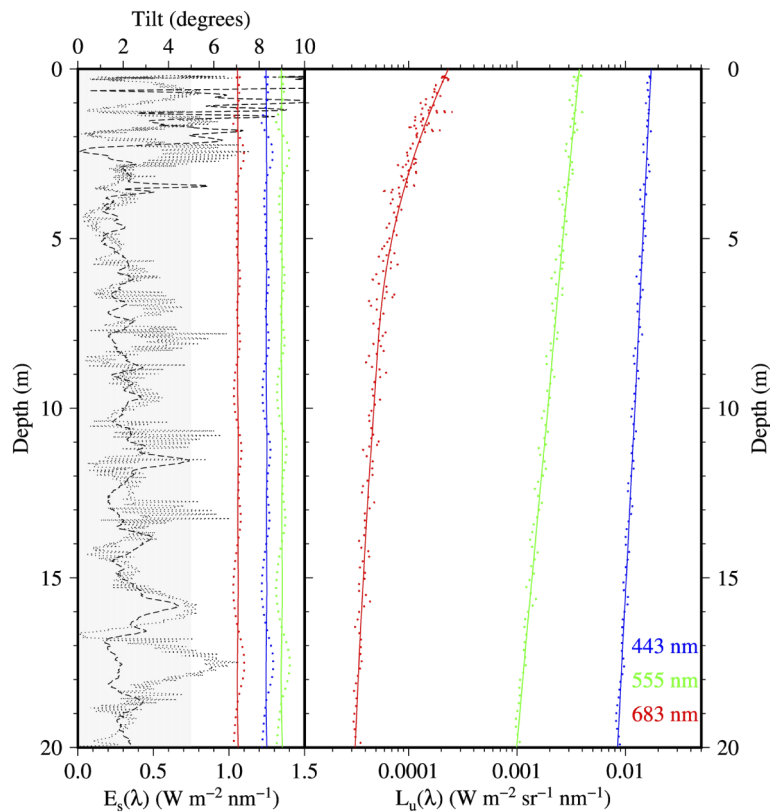
Fig. 5. Sample C-OPS data for the top 20 m of a profile collected on 2nd June 2019 and for the three wavelengths indicated (Station 17, blue waters). The left panel shows the tilt of the above-surface E_s sensor (dotted curve) and of the in-water L_u sensor (dashed curve). The greyed area is for tilt values < 5 degrees. On the same panel, the dotted colored curves are the actual E_s records and the continuous curves show the fitted data using the Loess function, both displayed as a function of matching depth of the $L_u(z)$ measurements (right panel, same line and color coding).

3.4.   *Seabird HyperPro*

3.4.1.   Instrument and the deployment setup

The Seabird HyperPro II [46] (HyperPro; formerly manufactured by Satlantic Inc., Halifax, Nova Scotia, Canada) has a downward looking HyperOCR radiometer that measures upwelling radiance, $L_u(\lambda)$, and an upward looking HyperOCI irradiance sensor to measure downward irradiance $E_d(\lambda)$ in the water column. In addition, there is an above-water upward looking HyperOCI irradiance sensor to measure downward irradiance, $E_s(\lambda)$, used as reference during data reduction.

   Each HyperOCR or HyperOCI has a 256-channel silicon-photodiode array detector with a 10-nm spectral resolution, a spectral sampling of 3.3 nm pixel^{-1}, and the radiance FOV is 8.5 degrees in water. The instruments are calibrated from 350 nm to 900 nm. They were calibrated before the cruise in June 2018 and after in August 2019. Calibrations were conducted at the NOAA/NESDIS calibration laboratory in College Park, MD, with less than a 1% difference between calibrations. The HyperOCRs have dark signal corrections performed using shutter dark measurements collected every 5th scan. The reference radiometer measuring E_s was collocated with the C-OPS E_s sensor, and the HyperPro deployment protocol was as described for the C-OPS, with the exception that for each station 2 to 10 multi-profile casts were conducted instead of successive single casts separated sometimes by several minutes as done with the C-OPS. All data in each multi-profile cast was processed to calculate R_{rs}.

3.4.2.   Data processing

The HyperOCRs utilize an adaptive gain which adjusts the integration time from 8 ms to 8192 ms depending on the light level being measured. The data acquisition frequency is 3 Hz at 128 ms integration time. The HyperPro buoyancy was adjusted so the profiling speed was kept slow and matched the C-OPS at 15 cm s^{-1}. The multicast technique logs data continuously while the HyperPro is allowed to profile up and down in the water column three to five times allowing the collection of up to hundreds of readings per meter thus meeting minimal validation depth resolution requirements outlined in [47]. At the end of the multi-cast, the logging is stopped, and the data saved into a single file. This can be repeated at each station three to 10 times depending on the desired coverage and consistency in environmental conditions.

   The individual multicast files are processed with Seabird ProSoft Software version 8.1.6. Straylight and thermal corrections are applied to the data and all profile data with tilts greater than 5 degrees is omitted from further processing. A least squared regression fit is applied to the log transformed $L_u(z)$ data and outliers are removed. The $L_u(z)$ data is normalized to the concurrent E_s data during the cast to account for variability of solar irradiance during the multicast. $L_u(0^-)$ is calculated from the intercept of the regression just below the surface. The depth of the regression is wavelength dependent and is determined by the selected optical depth (OD). The OD for each multicast is selected by processing the multicast using multiple optical depths of 0.5, 1.0, 1.5, 2.0 and 2.5 then selecting the OD that provides maximum depth while the log transformed data is still linear. This allows the maximization of $L_u(z)$ data points. L_w and R_{rs} are calculated as described in section 2. The $R_{rs}(\lambda)$ from each multicast at each station are averaged to determine the $R_{rs}(\lambda)$ for that station.

3.5.   *Reflectance normalization and band-averaged quantities*

By design, underwater profiling systems provide L_w at nadir, while the DALEC derives it at 40 degrees from nadir. The DALEC L_w's were therefore normalized to a nadir view and the DALEC, C-OPS and HyperPro L_w's were all normalized to a sun at zenith, following [48] and using the measured surface (z < 5 m) TChl-a concentration.

Radiometric quantities from the two hyperspectral instruments have been convolved with the spectral band response functions (SRFs) of either the C-OPS or the Ocean and Land Colour Imager (OLCI) instrument onboard the Copernicus Sentinel 3A satellite, in order to generate either C-OPS- or OLCI-specific L_w and E_s values, then ratioed to generate band-averaged R_{rs} [49]:

$$R_{rs,\,band-averaged} = \frac{\frac{\int_{\lambda_{i_0}}^{\lambda_{i_n}} L_w(\lambda_i) \cdot SRF(\lambda_i)\, d\lambda}{\int_{\lambda_{i_0}}^{\lambda_{i_n}} SRF(\lambda_i)\, d\lambda}}{\frac{\int_{\lambda_{i_0}}^{\lambda_{i_n}} E_s(\lambda_i) \cdot SRF(\lambda_i)\, d\lambda}{\int_{\lambda_{i_0}}^{\lambda_{i_n}} SRF(\lambda_i)\, d\lambda}}, \tag{4}$$

These quantities are used either for inter-comparison of instruments or for matchup with OLCI, in which case we used the OLCI sensor SRFs provided by EUMETSAT [50].

3.6. Statistical measures

Statistical measures have been used to quantify the comparisons of R_{rs} from instrument pairs or between an instrument and the OLCI-derived values (subscripts 1 and 2 in the following equations). For any given spectral band λ, we used the mean absolute difference (MAD) to quantify dispersion and the mean difference (MD) to quantify bias, both in R_{rs} units:

$$MAD_\lambda = \frac{1}{n} \sum_{i=1}^{n} |R_{rs}(\lambda)_{1,i} - R_{rs}(\lambda)_{2,i}| \tag{5}$$

$$MD_\lambda = \frac{1}{n} \sum_{i=1}^{n} (R_{rs}(\lambda)_{1,i} - R_{rs}(\lambda)_{2,i}) \tag{6}$$

en comparing two field instruments, we do not assume one *a priori* performing better than the other. Therefore, we also used the unbiased mean absolute percentage difference (MUAPD) and unbiased mean percentage difference (MUPD), again for each band λ:

$$MUAPD_\lambda = 200 \,\frac{1}{n} \sum_{i=1}^{n} \left| \frac{R_{rs}(\lambda)_{1,i} - R_{rs}(\lambda)_{2,i}}{R_{rs}(\lambda)_{1,i} + R_{rs}(\lambda)_{2,i}} \right| \tag{7}$$

$$MUPD_\lambda = 200 \,\frac{1}{n} \sum_{i=1}^{n} \left(\frac{R_{rs}(\lambda)_{1,i} - R_{rs}(\lambda)_{2,i}}{R_{rs}(\lambda)_{1,i} + R_{rs}(\lambda)_{2,i}} \right) \tag{8}$$

Note that the MUAPD is equal to MUPD if all differences are either positive or negative (systematic bias between the two data sets), and otherwise MUAPD > MUPD.

3.7. Satellite matchups

Our goal is not to provide a comprehensive validation of the S3A/OLCI sensor but rather to illustrate the capability of the DALEC instrument in particular. Therefore, as for the matchups between *in situ* and satellite R_{rs}, a simplified procedure was followed where average values are computed over a 5 by 5 pixel box centered on the ship position. Level2 flags were considered and a matchup rejected if any of these flags were raised. A matchup is performed only when more than half of the 25 pixels show positive R_{rs} values. Matchups are further reduced to those for which the coefficient of variation of R_{rs} within the valid pixels is lower than 0.25. This threshold is slightly larger than the recommended value (0.2 [51]). Using a CV of 0.2 would however eliminate all matchups for red bands, indicating an increased noise in the data for these bands, originating either from the instrument or the atmospheric correction process.

4.   Results and discussion

4.1.   Sample data

Examples of full E_s and L_w spectra from the DALEC and HyperPro plus the values for the discrete spectral bands of the C-OPS are displayed in Fig. 6 for station 11. The resulting R_{rs} are also displayed and, for the DALEC results, both L_w and R_{rs} are shown for the five methods used to determine ρ and also when the skylight reflection is not corrected.
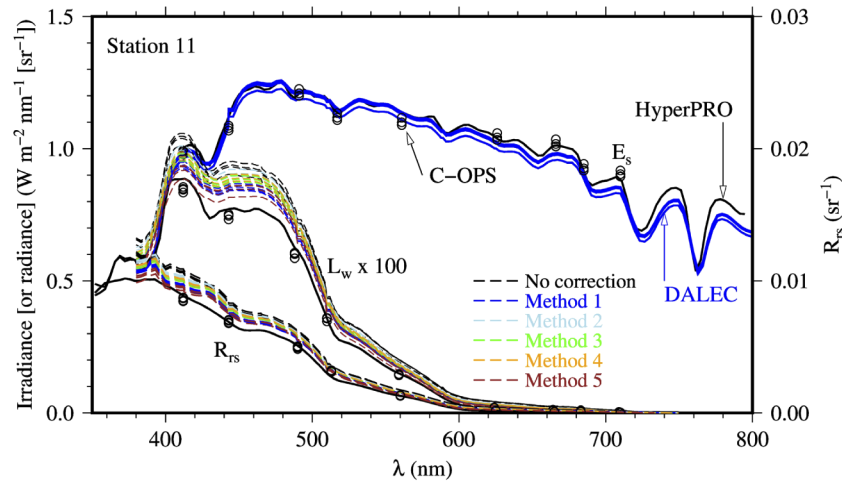
Fig. 6. E_s, L_w and R_{rs} spectra derived from the DALEC, HyperPro and C-OPS at station 11.

The sky was clear and conditions stable at station 11. Some differences are nevertheless observed for E_s, which are likely caused by both the measurements not being taken from the same place on the ship (for the DALEC and the two other instruments) and the HyperPro and C-OPS reference sensors not being gimbaled while the DALEC is. In addition, the data acquisition frequency is different for the two non-gimbaled E_s sensors, so that their sampling does not occur for identical instrument tilt (because of the ship's roll and pitch). This is simply showing that meaningful direct comparison of radiometric quantities measured at sea is generally impracticable, except if conditions are extremely stable and special care is put in making perfectly simultaneous measurements with collocated instruments. This can be achieved for E_s measurements and is virtually impossible for underwater $L_u(z)$. Collocating above-water systems on a fixed platform is another configuration where such a direct comparison can be attempted [52].

Part of the differences observed in Fig. 6 could also originate from calibration set ups and procedures, because no intercalibration to a single standard was performed before the voyage. Such differences should however normally be < 2% [e.g., [53]] and are not addressed here.

Therefore, the E_s and L_w displayed on Fig. 6 are shown as a matter of illustration, not for a quantitative comparison of their values from the different instruments. This quantitative assessment is instead made in terms of R_{rs} (also displayed on Fig. 6), which is the apparent optical property least sensitive to external factors such as the sun elevation or the diffuseness of the skylight, so that it has been referred to as a quasi-inherent optical property [54]. Provided that $L_u(z)$ or L_t are measured simultaneously to E_s, effects of small changes in illumination are also eliminated (still measurements were only taken when the E_s record was stable enough; see left panel of Fig. 5). The R_{rs} spectra on Fig. 6 confirm this by showing nearly perfect agreement between the HyperPro and the C-OPS (MUPD from 0.9% to 2.5% in the visible) and slightly larger values for the DALEC (MUPD from 5% to 10% in the visible).

### 4.2.   Comparing R_{rs} values at discrete wavelengths from the three instruments

To further compare R_{rs} obtained from the three instruments, the 3-nm resolution DALEC and HyperPro spectra have been convolved as described in the method section (Eq. (4)). The results are shown in Fig. 7 for four selected stations and show generally excellent agreement. The DALEC R_{rs} spectra obtained without skylight correction and for the five correction methods are displayed, which essentially separate for $\lambda > 600$ nm. Similarly, all (generally 3) C-OPS casts performed at each station are displayed, which essentially are superimposed one on top of each other.
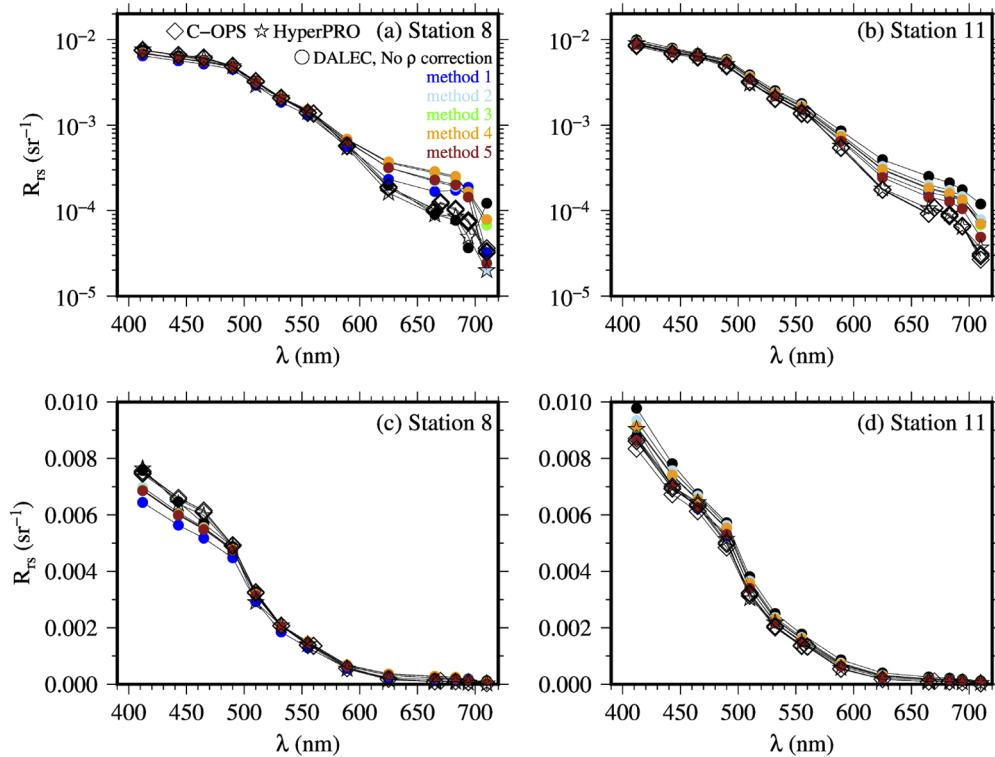

Fig. 7. Examples of R_{rs} spectra for two stations and the three instruments, as indicated. The same data are displayed with a logarithmic (top row) and linear (bottom row) scales to highlight differences in both ends of the spectrum. Data from the hyperspectral DALEC and HyperPro instruments have been convolved by 10-nm wide boxcar functions centered in the C-OPS wavelengths (Eq. (4)).

These results are generalized by comparing the R_{rs} values between each instrument pair for all clear-sky stations (8 to 20) and all wavelengths. The results are displayed in Fig. 8 for the C-OPS vs. HyperPro, Fig. 9 for the DALEC vs. C-OPS and Fig. 10 for the DALEC vs. HyperPro. Tables S1, S2 and S3 (Supplement 1) provide wavelength-specific and all-of-dataset statistical indicators. In these figures and corresponding Tables, the signed statistics are for R_{rs} of the instrument on the X-axis minus R_{rs} of the instrument on the Y-axis (e.g., C-OPS minus HyperPro for Fig. 8 and Supplement 1, Table S1).

For the comparison in Fig. 8, all individual C-OPS R_{rs} spectra at each station were averaged into a single spectrum, and compared to the single HyperPro R_{rs} spectrum derived from the multicast technique. Similarly, for Fig. 10, an average DALEC spectrum was derived over the time interval corresponding to the multicast HyperPro session. For Fig. 9, however, no such
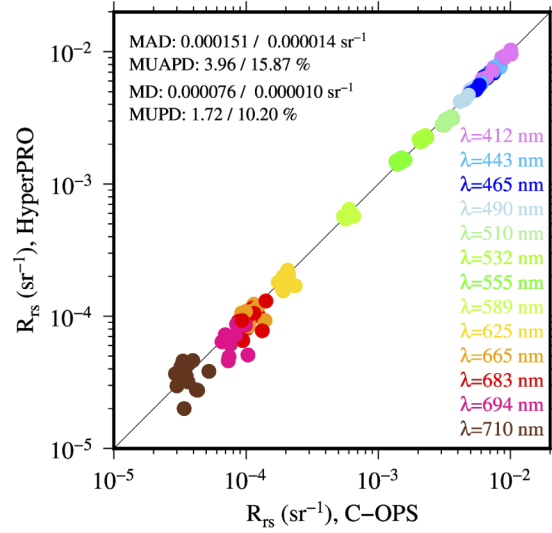
Fig. 8. Scatterplot for the HyperPro *vs.* C-OPS R_{rs} for the wavelengths indicated. The two values for each reported statistical indicator are averages for spectral bands with $\lambda < 600$ nm or $\lambda > 600$ nm (see Supplement 1, Table S1 for per-wavelength statistics).
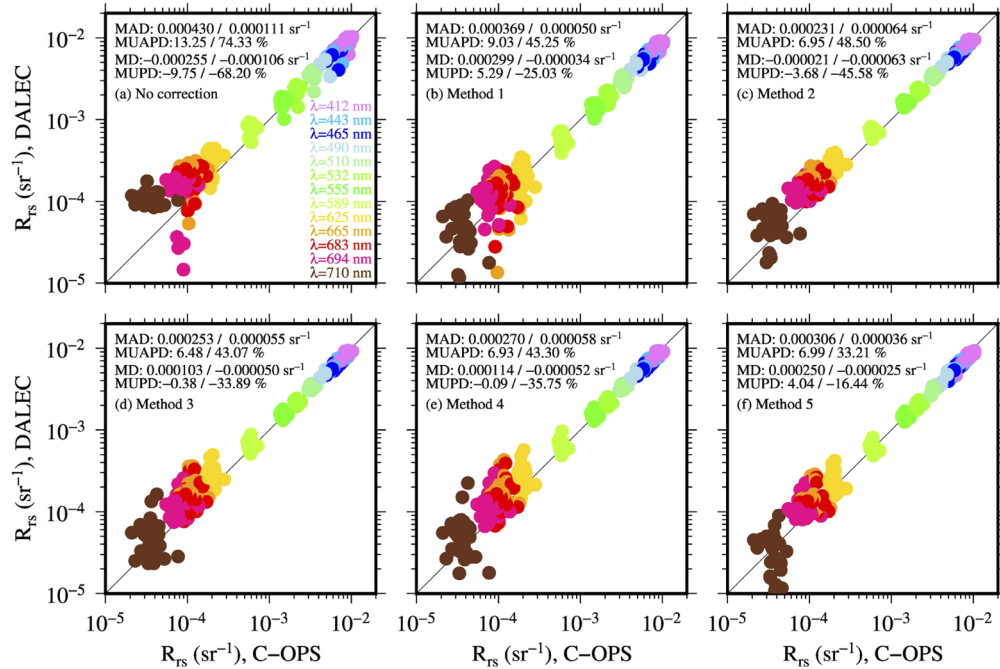
Fig. 9. As for Fig. 8 but for the DALEC vs. C-OPS R_{rs} and for the 5 processing methods. See Supplement 1, Table S2 for all statistics.
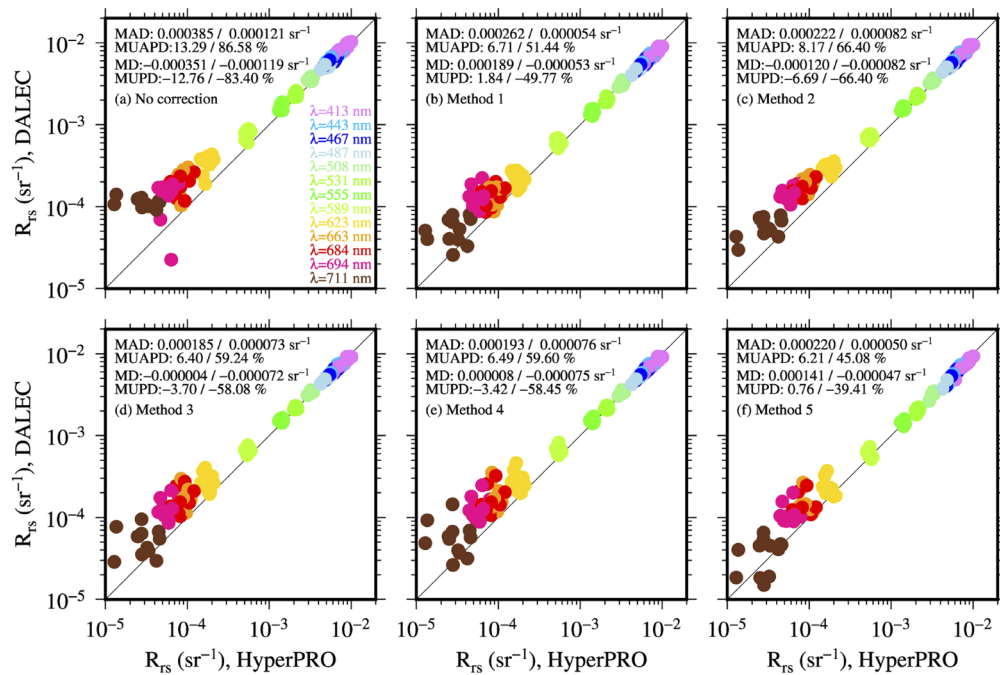
Fig. 10. As in Fig. 9 but for the DALEC vs. Hyper-Pro R_{rs}. See Supplement 1, Table S3 for all statistics.

average was necessary and each individual C-OPS cast was matched with a DALEC spectrum computed from the start time of the C-OPS cast to that time plus 3 minutes, which is about the time the profiler needs to reach the depth from which extrapolation to the surface is performed.

The agreement between the C-OPS and HyperPro is excellent (Fig. 8), with a MUD for the blue and green bands ($\lambda < 600$ nm) lower than 2% and a MUPD just below 4%. These numbers go up to 10-15% for red bands, which is still remarkable considering the quite different techniques used for data collection and processing (single cast *vs.* multi cast) and the acknowledged difficulty in determining the $L_u(0^-)$ value in that spectral range in clear waters [55].

The use of the LOESS function in the C-OPS processing was presented as a less subjective alternative to the often-used log-linear fit technique. The selection of a time interval in applying the LOESS function is admittedly also subjective. The question then is to assess which of these two imperfect techniques entails less uncertainty in determining R_{rs}.

Therefore, a sensitivity study was performed where the entire data set is reprocessed by changing the L_u and E_s fitting time interval from 40 s / 80 s to either 20/20 or 40/40 or 80/80. The average coefficients of variation within the ensemble of R_{rs} calculated for these four time-intervals and for stations 8 to 20 are provided in Table 3. Values are generally about 3% for blue bands, up to about 5% for green bands and above 10% for red bands. These larger values are expected when the $L_u(z)$ profile near the surface is strongly curved by changes in the diffuse attenuation coefficient, including because of Raman scattering [55]. The statistics reported in Supplement 1, Table S1 for wavelengths below 600 nm generally change by less than 0.5%, up to 1% in few cases. Larger changes are observed for red bands ($\lambda > 600$ nm), as expected. These uncertainties for the blue and green bands are lower than those entailed by changing the depth interval over which a single K_L is determined and used for extrapolation to the 0^- level [45]. We illustrate this here by also computing average coefficients of variation within the ensemble of R_{rs} calculated using K_L values determined over various depth intervals starting from a depth of either 1, 2 or

3 m, and extended by increments of 2 m down to 14 m deeper (second line of CV values in Table 3). These various depth intervals could all be selected as relevant, in particular to account for fluctuations of $L_u(z)$ due to wave focusing, which are particularly significant in clear-sky, clear-waters and moderate wind speeds typical of our data set [56].

Table 3. Coefficient of variation within the C-OPS R_{rs} values for stations 8 to 20 and 1- the 4 different sets of time intervals used with the Loess function and, 2- the different interpolation intervals for a K_{lu}-based extrapolation (see text).

λ (nm)	340	412	443	465	490	510	532	555	560	589
1-CV (%)	2.80	3.03	3.18	3.20	3.27	3.58	3.81	4.01	4.07	4.67
2-CV (%)	5.23	4.72	4.85	4.79	4.89	6.03	6.93	7.93	8.11	11.07
λ (nm)	625	665	670	683	694	710				
1-CV (%)	11.61	22.35	19.17	23.38	27.67	42.26				
2-CV (%)	19.68	25.53	19.48	24.35	27.56	35.57				

Another source of uncertainty is comparing multi- and hyper-spectral instruments originates from the convolution of a high-resolution spectra by a function that is supposed to accurately represent the spectral band response function (SRF) of the multispectral radiometer. These functions are generally not available from instrument manufacturers and customers rarely have the technical capability for determining these functions.

The results in Fig. 7 and 8 were actually obtained by using simplified SRFs represented by a 10-nm wide boxcar function centered onto the C-OPS wavelength and with value 1 within that interval. We did this because instrument-specific SRFs were not available for the C-OPS and the boxcar model is *a priori* a valid approximation of what the SRFs would be for the microradiometer technology (at least at full width half maximum). The manufacturer however provided class-based generic SRFs for the microradiometers, which we also used to assess the impact on the band-average R_{rs} of using them instead of the simple boxcar function. The manufacturer did not recommend the use of these SRFs for generating final results because they could not guarantee the spectral calibration. We assessed the potential spectral shift of these SRFs by searching for a nominal wavelength for each band in each of the SRFs, as being the wavelength at which the area under the curve is exactly split in two. Percent differences in R_{rs} between the C-OPS and the HyperPro obtained by convolving full HyperPro spectra with these SRFs or the 10-nm wide boxcar function were below 3% for λ = 412, 490, 510, 555, 625 and 683 nm, 4% for λ = 443 nm, and from 4.5 to 8% for other bands. When repeating the same exercise for the DALEC, numbers were below 3% for λ = 412, 490, 510, 532, 555, 625 and 665 nm, between 3% and 4% for other bands except 683 nm where it reaches 5.8%. This simple comparison shows that an average uncertainty of about 3% has to be expected from the convolution process when comparing R_{rs} values derived from multi- and hyper-spectral radiometers.

The same comparison as shown in Fig. 8, when performed with the SRF-based convolved HyperPro spectra showed a lower MUPD (0.87%) but a higher MUAPD (4.55%). Considering these small differences and the absence of instrument-specific SRFs, we choose to use the boxcar model in the rest of this study.

Another option that we also tested is to pick within the HyperPro (or DALEC) spectrum the closest spectral band to the nominal central band of the C-OPS, without any averaging, as done in [17]. They sometimes match within 1 nm but the difference can reach 3 nm for wavelengths longer than 600 nm. The rationale for doing this is that each of the spectral band in the HyperPro or DALEC spectrum (both based on the same spectrometer technology) has a spectral width of about 10 nm, which is close to the spectral width of the C-OPS bands (Full width at half maximum). Results like those displayed in Fig. 8 were slightly degraded when doing so, with both the MUAPD and MUPD being greater than 6%.

The same comparison as shown on Fig. 8 for the HyperPro vs. C-OPS R_{rs} is displayed in Fig. 9 for the DALEC vs. C-OPS R_{rs}. The figure now includes six panels, corresponding to the 5 methods for quantifying the contribution of skylight reflection to L_t, and also to the case where no correction is attempted. We thought it useful to show the latter so that the amplitude of the required correction can be appreciated.

When no correction for skylight reflection is applied, the expectation is that the DALEC will overestimate R_{rs}. This is indeed what is generally observed (Fig. 9(a)), yet some spectra also include values that are lower than those of the C-OPS, probably indicating some unidentified shading or significant difference in illumination conditions.

The bias is reduced when the various corrections are applied and the best agreement for the blue and green bands is for method 3 (Fig. 9(d)), with a MUPD at 2% and MUAPD at 6.9%. Method 5, however, performs better for the red bands, with a MUPD < 15% and MUAPD of about 30% (Fig. 9(f)). The data scatter is generally larger than for the C-OPS vs. HyperPro comparison, with the lowest MUAPD being nearly twice as large as that displayed in Fig. 8. This larger scatter is however partly due to the use of instantaneous instead of average R_{rs} spectra.

The comparison between the DALEC and HyperPro shows similar results (Fig. 10). The dispersion (MUAPD) is generally lower for the blue-green bands. Method 5 (Fig. 10(f)) seems to provide the best results for that spectral range. No method provides satisfactory results for the red bands, with negative MUPDs from about -35% to -60%, indicating possible underestimation of R_{rs} by the HyperPro for these bands (consistent with Fig. 8).

In order to summarize the information, the MUPD and MUAPD statistics are displayed across the spectrum in Fig. 11. The colored curves in panels (a) and (d) correspond to changing the fitting interval of the LOESS function (C-OPS processing), while they correspond to the different methods of correcting for the sky radiance reflection for the four other panels where the comparison involves the DALEC.

The results in Figs. 11(a) and 11(d) confirm that deriving accurate R_{rs} for $\lambda > {\sim}600$ nm from under-water measurements is very difficult in clear waters. The spread of MUPD (or MUAPD) values in that spectral domain when different fitting intervals are used for the LOESS function has a strong impact, and no independent information is available that would indicate which one is better than the others. The above-water measurement is not affected by the extrapolation issue and could be considered the reference to compare with. This one has its own significant uncertainties related to the determination of ρ, however, as shown by the spread of curves in Figs. 11(b), (c), (e), (f) (again, for $\lambda > {\sim}600$ nm). It is, quite expectedly, hardly achievable to retrieve L_w accurately in that spectral range, when the values are vanishingly small for clear waters.

For wavelengths from 412 nm to 560 nm, which is the spectral range used in most bio-optical algorithms, the uncertainty in extrapolation is essentially around 1.5% (spread of curves in Fig. 11(d)), which is consistent with, e.g., [55], and is even below the calibration uncertainty [57]. The two in-water systems agree within 3% in that blue to green spectral range (Fig. 11(d)).

The MUPD curves for the various methods applied to correct for skylight reflection come one after the other in a consistent way for the C-OPS vs. DALEC (Fig. 11(b)) and HyperPro vs. DALEC (Fig. 11(c)) comparisons. When no correction is applied, the bias is of about 10 to 15% for $\lambda < {\sim}600$ nm. The results displayed in Fig. 11 confirm what was seen in Figs. 9 and 10. They show that Methods 3 and Method 4 [23] lead to the best agreement between the C-OPS and DALEC, while method 5 (spectral optimization) seems more appropriate for the HyperPro vs. DALEC comparison, and is the best performing for the red bands in comparison with both the DALEC and HyperPro. Still, the MUAPD is at least 20% in that spectral range.

Having to use different methods to reach the best agreement between the DALEC and two in-water systems must be related to the way R_{rs} is derived from measurements by these two systems, i.e., a single- vs. a multi-cast procedures. Comparing Figs. 11(e) and 11(f) and assuming
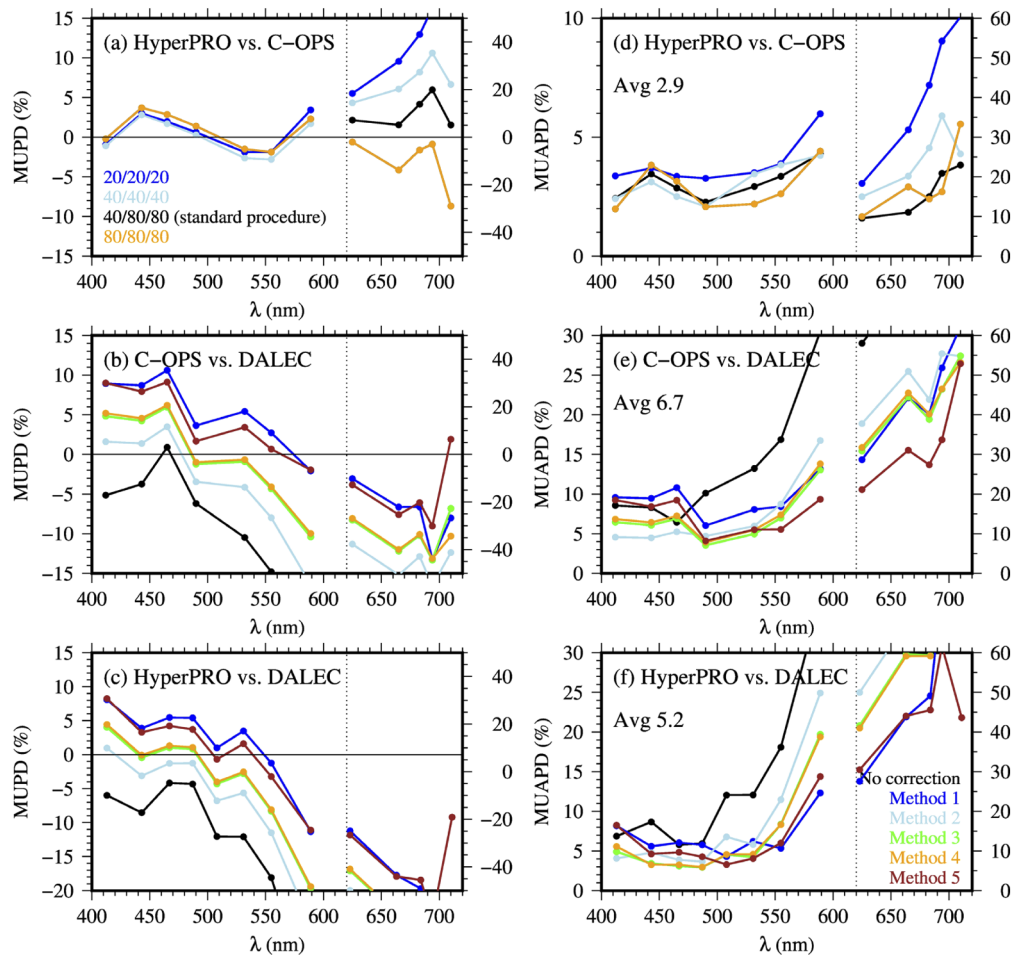
Fig. 11. MUPD (a), (b), (c) and MUAPD (d), (e), (f) for the instrument-pair comparisons of R_{rs}, as indicated. Note the different scales for the vertical axes, for $\lambda < 600$ nm (left axis) or > 600 nm (right). Colored curves when the DALEC is involved (panels b, c, e, f) correspond to the various methods for determining the sky radiance reflection. In (a) and (d) they correspond to the different intervals chosen for the LOESS function in the C-OPS processing. The average MUAPD value indicated in panels (d, e, f) is for all bands from 412 to 560 nm and all but the "no correction" curves for the DALEC and all four curves for the C-OPS.

that the above-water R_{rs} determination is closer to the truth for the red bands, the single-cast C-OPS method leads to lower MUAPDs. The opposite is true for $\lambda < 600$ nm. This better agreement between the DALEC and HyperPro can also be because these two instruments are essentially based on the same light collection technology (same spectrometers). It is also noted that the spread of the MUAPD curves for $\lambda < 600$ nm in Fig. 11(e) is reduced and their shape are more similar to those in Fig. 11(f) when no convolution is applied to the DALEC data and instead single DALEC bands are picked up for the comparison with the C-OPS. The average MUAPD is however a bit large in that case (8.2%).

These results and the shape of the curves in Figs. 11(d), (e), (f) relate to the uncertainty in extrapolating $L_u(z)$ to the 0^- level (variability shown in Fig. 11(d)), which directly stems from whether or not K_d significantly changes close to the surface [55]. If it does, as is the case for

wavelengths close to 400 nm and above 550-600 nm, the uncertainty of the extrapolation to the 0^- level increases. Because such uncertainties does not exist when using the above-water instruments, the shape of the MUAPD curves in Figs. 11(e),(f) resembles that of $K_d(\lambda)$.

4.3. Matchups with the Copernicus Sentinel3A and 3B/OLCI instruments

Here we provide results of a comparison between the C-OPS, HyperPro and DALEC (method 5) R_{rs} and the R_{rs} from the two Ocean and Land Color Imagers (OLCI) on the Copernicus Sentinel 3A and 3B satellites (S3A, S3B). We do not claim this exercise is a thorough validation of these two missions' performance, in particular because the number of matchups between in-situ and satellite observations is rather small. In addition, the processing versions for both satellite sensors data sets are not the latest so that our results do not necessarily represent the current performance of the two missions. Therefore, we do not include comparisons statistics, and rather show these results as an illustration of the capability of the DALEC and the two under-water systems to provide relevant validation data.

For matchups with both S3A (Fig. 12(a),(b),(c)) and S3B (Fig. 12(d),(e),(f)), the vertical error bar on each comparison point corresponds to plus or minus one standard deviation within the valid pixels used for the matchups. The horizontal error bar corresponds to an approximate percent error on the in-situ R_{rs} value that is set from the MUAPD values in Fig. 11. We used 8% for $\lambda \leq 443$ nm, 6% for 443 nm $< \lambda \leq 560$ nm, 15% for $\lambda = 589$ nm, and 50% for $\lambda > 600$ nm. The number of C-OPS and DALEC matchups is sometimes greater than the number of stations because we matched the satellite data with each C-OPS individual cast or each valid DALEC R_{rs} spectrum obtained at the time of that C-OPS cast. This does not happen for the HyperPRO because only one R_{rs} value is derived at each station from the multi-cast approach.
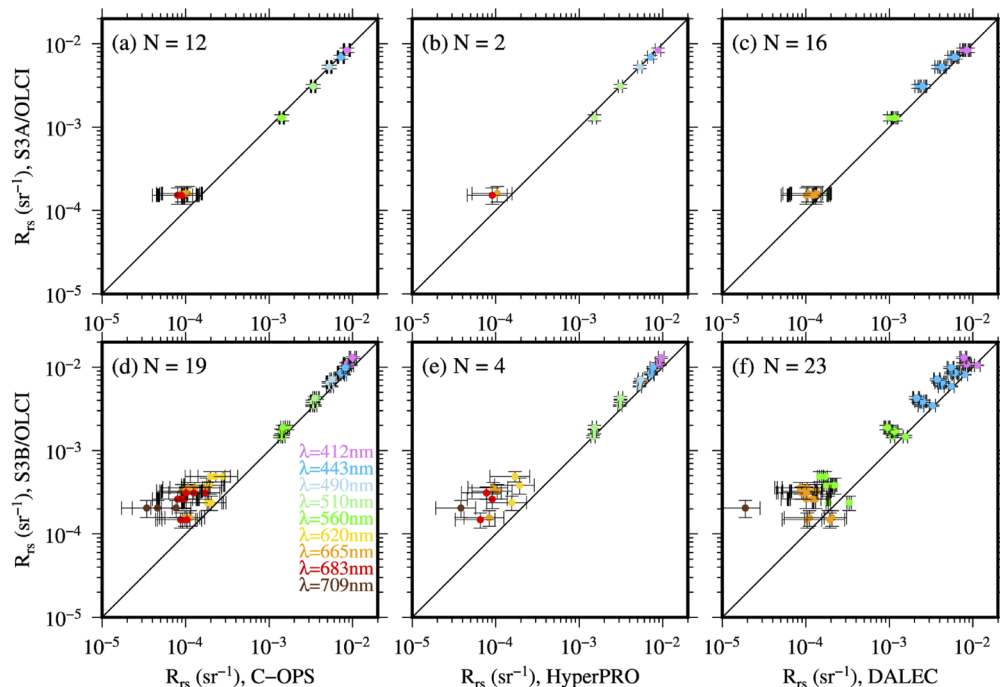
Fig. 12. Example matchups with the S3A/OLCI (bottom row) and S3B/OLCI (top row) sensors. N is the number of matchups (spectra). The results for the DALEC are using the processing method #5.

As expected from the comparisons among field instruments, the matchup results with S3A or S3B and either the C-OPS, HyperPro or DALEC are very close for the blue / green wavelengths, showing no bias for S3A and a positive bias for S3B. This difference clearly shows the importance of system vicarious calibration, which was applied to S3A but not yet to the version of S3B data we used.

Results differ for red bands, with the DALEC showing a slightly better agreement with the satellite-derived values. Although no fully unambiguous answer can be provided here, this result, combined to the comparisons summarized in Fig. 11, indicate that above-water systems are likely to be better adapted to provide relevant validation data in that spectral range ($\lambda > 600$ nm) over clear waters.

5.   Conclusion

It must be kept in mind that the conclusions we draw here derive from analyzing a data set collected in clear oceanic waters (Chl $< \sim 0.2$ mg m^{-3}) and with clear atmospheres ($\tau_a < 0.1$ at 865 nm).

Our study confirms that R_{rs} for $\lambda < 600$ nm can be derived from radiometric measurements of either above- or in-water radiometer systems with excellent agreement, here with biases (MUPD) from close to 0 to +/-6% and dispersion (MUAPD) of about 5 to 8% (Figs. 9 and 10, and Supplement 1, Tables S2 and S3). The agreement is better when comparing the two in-water systems, which here agree within about 3% in bias and dispersion (Fig. 8 and Supplement 1, Table S1).

These results are obtained with an essentially unattended above-water system, the DALEC, showing that the instrument can maintain proper viewing geometry without operator intervention. The advantage of collocated E_s, L_s and L_t measurements is also emphasized, which sometimes can be a challenge with in-water systems (for measurements of E_s and $L_u(z)$) on large vessels.

This inter-comparison exercise also shows that attempting direct comparison of radiometric quantities is hardly meaningful for data collected in the context of a research voyage into which radiometric measurements are only one part of many other operations. Specific operations at sea including dedicated and ample time are normally required for making such intercomparisons meaningful [e.g., 14]. A configuration where such a comparison can work is for two collocated gimbaled E_s sensors, still they would have to be synchronized in terms of acquisition frequency.

The significant spread of results when using five different methods for correction of the skylight reflection also emphasizes the inherent limitation of such techniques. The selection (or computation) of the reflection coefficient ρ is of course an important element of uncertainty but, similarly, the statistical treatment of the time series of individual measurements plays a significant role. This treatment, whose role is to eliminate most of the contamination of L_t by reflection of sun glint, is not systematically described in literature reporting similar comparison exercises as this one. Combining various ρ values with various data filtering techniques offers a number of possible processing permutations which can and should be assessed with more rigor in subsequent studies.

Our results do not point to a single correction method as being superior to the others for $\lambda < 600$ nm, although methods 3 and 4 (constant $\rho = 0.028$ and variable ρ following [23,29]) generally provide the best comparison results. Method 5 (spectral optimization) provides the best results for the red bands ($\lambda > 600$ nm; Fig. 11(b),(c),(e),(f)).

References

1. C. R. McClain, G. C. Feldman, and S. B. Hooker, "An overview of the SeaWiFS project and strategies for producing a climate research quality global ocean bio-optical time series," Deep Sea Res., Part II **51**(1-3), 5–42 (2004).
2. C. R. McClain, "A Decade of Satellite Ocean Color Observations," Annu. Rev. Mar. Sci. **1**(1), 19–42 (2009).
3. P.J. Werdell and S.W. Bailey, "The SeaWiFS Bio-optical Archive and Storage System (SeaBASS): Current architecture and implementation," *NASA Tech. Memo. 2002-211617*, G.S. Fargion and C.R. McClain, eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 45 pp. (2002).
4. P. J. Werdell and S. W. Bailey, "An improved in situ data set for bio-optical algorithm development and ocean color satellite validation," Rem. Sens. Environ. **98**(1), 122–140 (2005).
5. K. Barker, C. Mazeran, C. Lerebourg, M. Bouvet, D. Antoine, M. Ondrusek, G. Zibordi, and S. Lavender, "MERMAID: The MEris MAtchup In-situ Database," Proc. 2nd MERIS/(A)ATSR User Workshop : 22 - 26 September 2008, ESRIN, Frascati, Italy. ESA SP ; 666. Publisher: ESA Communication Production Office. ISBN 9789292212308
6. C. J. Donlon, P. Minnett, N. Foxn, and W. Wimmer, "Strategies for the laboratory and field deployment of ship-borne fiducial reference thermal infrared radiometers in support of satellite-derived sea surface temperature climate data records," in G. Zibordi, C. Donlon, and A. Parr, (eds.) *Optical radiometry for ocean and climate measurements, Experimental methods in the physical sciences*, Vol. 47, Elsevier, ISBN 978-0-12-417011-7 (2014).
7. A. C. Banks, R. Vendt, K. Alikas, A. Bialek, J. Kuusk, C. Lerebourg, K. Ruddick, G. Tilstone, V. Vabson, C. Donlon, and T. Casal, "Fiducial Reference Measurements for Satellite Ocean Colour (FRM4SOC)," Remote Sens. **12**(8), 1322 (2020).
8. A. Białek, S. Douglas, J. Kuusk, I. Ansko, V. Vabson, R. Vendt, and T. Casal, "Example of Monte Carlo Method Uncertainty Evaluation for Above-Water Ocean Colour Radiometry," Remote Sens. **12**(5), 780 (2020)..
9. S. B. Hooker, G. Zibordi, J-F. Berthon, D. D'Alimonte, S. Maritorena, S. McLean, and J. Sildam, "Results of the Second SeaWiFS Data Analysis Round Robin, March 2000 (DARR-00),". *NASA Tech. Memo. 2001–206892*, Vol. 15, S.B. Hooker and E.R. Firestone, eds., NASA Goddard Space Flight Center, Greenbelt, Maryland, 71 pp (2001).
10. A. Bialek, S. Douglas, J. Kuusk, I. Ansko, V. Vabson, E. Vendt, and T. Casal, "Example of Monte Carlo Method Uncertainty Evaluation for Above-Water Ocean Colour Radiometry," Remote Sens. **12**(5), 780 (2020).
11. A. Białek, V. Vellucci, B. Gentili, D. Antoine, J. Gorroño, N. Fox, and C. Underwood, "Monte Carlo–Based Quantification of Uncertainties in Determining Ocean Remote Sensing Reflectance from Underwater Fixed-Depth Radiometry Measurements," J. Atmos. Oceanic Technol. **37**(2), 177–196 (2020).
12. V. Vabson, J. Kuusk, I. Ansko, R. Vendt, K. Alikas, K. Ruddick, A. Ansper, M. Bresciani, H. Burmester, M. Costa, D. D'Alimonte, G. Dall'Olmo, B. Damiri, T. Dinter, C. Giardino, K. Kangro, M. Ligi, B. Paavel, G. Tilstone, R. Van Dommelen, S. Wiegmann, A. Bracher, C. Donlon, and T. Casal, "Field Intercomparison of Radiometers Used for Satellite Validation in the 400–900 nm Range," Remote Sens. **11**(9), 1129 (2019)..
13. S. B. Hooker and A. Morel, "Platform and Environmental Effects on Above-Water Determinations of Water-Leaving Radiances," J. Atm. Ocean Tech. **20**(1), 187–205 (2003).
14. G. Tilstone, G. Dall'Olmo, M. Hieronymi, K. Ruddick, M. Beck, M. Ligi, M. Costa, D. D'Alimonte, V. Vellucci, D. Vansteenwegen, A. Bracher, S. Wiegmann, J. Kuusk, V. Vabson, I. Ansko, R. Vendt, C. Donlon, and T. Casal, "Field Intercomparison of Radiometer Measurements for Ocean Colour Validation," Remote Sens. **12**(10), 1587 (2020)..
15. V. E. Brando, J. L. Lovell, E. A. King, D. Boadle, R. Scott, and T. Schroeder, "The Potential of Autonomous Ship-Borne Hyperspectral Radiometers for the Validation of Ocean Color Radiometry Data," Remote Sens. **8**(2), 150 (2016)..
16. K. Alikas, V. Vabson, I. Ansko, G. H. Tilstone, G. Dall'Olmo, F. Nencioli, R. Vendt, C. Donlon, and T. Casal, "Comparison of Above-Water Seabird and TriOS Radiometers along an Atlantic Meridional Transect," Remote Sens. **12**(10), 1669 (2020)..
17. J. E. O'Reilly, S. Maritorena, B. G. Mitchell, D. A. Siegel, K. L. Carder, S. A. Garver, M. Kahru, and C. R. McClain, "Ocean color chlorophyll algorithms for SeaWiFS," J. Geophys. Res. **103**(C11), 24937–24953 (1998).
18. C. Hu, Z. P. Lee, and B. Franz, "Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference," J. Geophys. Res.: Oceans **117**(C1), C01011 (2012)..
19. K. G. Ruddick, K. Voss, E. Boss, A. Castagna, R. Frouin, A. Gilerson, M. Hieronymi, B. C. Johnson, J. Kuusk, Z. Lee, M. Ondrusek, V. Vabson, and R. Vendt, "A Review of Protocols for Fiducial Reference Measurements of

Water-Leaving Radiance for Validation of Satellite Remote-Sensing Data over Water," Remote Sens. **11**(19), 2198 (2019).

20. K. G. Ruddick, K. Voss, A. C. Banks, E. Boss, A. Castagna, R. Frouin, M. Hieronymi, C. Jamet, B. C. Johnson, J. Kuusk, Z. Lee, M. Ondrusek, V. Vabson, and R. Vendt, "A Review of Protocols for Fiducial Reference Measurements of Downwelling Irradiance for the Validation of Satellite Remote Sensing Data over Water," Remote Sens. **11**(19), 2198 (2019).

21. Z. P. Lee, N. Pahlevan, Y.-H. Ahn, S. Greb, and D. O'Donnell, "Robust approach to directly measuring water-leaving radiance in the field," Appl. Opt. **52**(8), 1693–1701 (2013).

22. G. Zibordi and M. Talone, "On the equivalence of near-surface methods to determine the water-leaving radiance," Opt. Express **28**(3), 3200 (2020).

23. C. D. Mobley, "Estimation of the remote-sensing reflectance from above-surface measurements," Appl. Opt. **38**(36), 7442–7455 (1999).

24. J. L. Mueller, C. Davis, R. Arnone, R. Frouin, K. Carder, Z. P. Lee, R.G. Steward, S. B. Hooker, C. D. Mobley, and S. McLean, "Above-Water Radiance and Remote Sensing Reflectance Measurement and Analysis Protocols," Chapter 3 in Ocean Optics Protocols For Satellite Ocean Color Sensor Validation, Revision 4, Volume III, NASA/TM-2003-21621/Rev-Vol III (2003).

25. R. R. Hood, E. R. Urban, M. J. McPhaden, D. Su, and E. Raes, "The 2nd International Indian Ocean Expedition (IIOE-2): Motivating New Exploration in a Poorly Understood Basin," Limnol. Oceanogr. Bull. **25**(4), 117–124 (2016).

26. D. Behrman 1981. Assault on the largest unknown: The International Indian Ocean Expedition, 1959-1965. UNESCO Press.

27. J. Ras, J. Uitz, and H. Claustre, "Spatial variability of phytoplankton pigment distributions in the Subtropical South Pacific Ocean: comparison between in situ and modelled data," Biogeosciences **5**(2), 353–369 (2008).

28. A. Smirnov, B. N. Holben, I. Slutsker, D. M. Giles, C. R. McClain, T. F. Eck, S. M. Sakerin, A. Macke, P. Croot, G. Zibordi, P. K. Quinn, J. Sciare, S. Kinne, M. Harvey, T. J. Smyth, S. Piketh, T. Zielinski, A. Proshutinsky, J. I. Goes, N. B. Nelson, P. Larouche, V. F. Radionov, P. Goloub, K. Krishna Moorthy, R. Matarrese, E. J. Robertson, and F. Jourdin, "Maritime Aerosol Network as a component of Aerosol Robotic Network," J. Geophys. Res. **114**(D6), D06204 (2009)..

29. M. M. Slivkoff, "Ocean Colour Remote Sensing of the Great Barrier Reef Waters," PhD Thesis, Curtin University, http://hdl.handle.net/20.500.11937/798 (2014)

30. E. P. Shettle and R. W. Fenn, "Models for the aerosols of the lower atmosphere and the effects of humidity variations on their optical properties," Environmental Research Papers, AFGL-TR-79-0214, Air Force Geophysics Laboratory, 20 September 1979 (1979).

31. D. Antoine and A. Morel, "A multiple scattering algorithm for atmospheric correction of remotely-sensed ocean colour (MERIS instrument) : principle and implementation for atmospheres carrying various aerosols including absorbing ones," Int. J. Remote Sens. **20**(9), 1875–1916 (1999).

32. 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," Opt. Express **25**(16), A742–A761 (2017).

33. P. M. M. Groetsch, R. Foster, and A. Gilerson, "Exploring the limits for sky and sun glint correction of hyperspectral above-surface reflectance observations," Appl. Opt. **59**(9), 2942–2954 (2020).

34. S. G. H. Simis and J. Olson, "Unattended processing of shipborne hyperspectral reflectance measurements," Remote Sens. Environ. **135**, 202–212 (2013).

35. Z. P. Lee, Y.-H. Ahn, C. D. Mobley, and R. Arnone, "Removal of surface-reflected light for the measurement of remote-sensing reflectance from an above-surface platform," Opt. Express **18**(25), 26313–26324 (2010).

36. G. Zibordi, "Experimental evaluation of theoretical sea surface reflectance factors relevant to above-water radiometry," Opt. Express **24**(6), A446–A459 (2016).

37. C. D. Mobley, "Light and Water: Radiative Transfer in Natural Waters," Academic Press, 592 pp (1994).

38. X. Zhang, S. He, A. Shabani, P.-W. Zhai, and K. Du, "Spectral sea surface reflectance of skylight," Opt. Express **25**(4), A1–A13 (2017).

39. C. B. Markwardt, "Non-linear least-squares fitting in IDL with MPFIT," in D. A. Bohlender, D. Durand, and P. Dowler, eds., '*Astronomical Society of the Pacific Conference Series*', Vol. 411 of Astronomical Society of the Pacific Conference Series, pp. 251 (2009).

40. C. R. Booth, J. H. Morrow, and R. N. Lind, "Development of the Microradiometer, in: Advances in Measuring the Apparent Optical Properties (AOPs) of Optically Complex Waters," *NASA Tech. Memo. 2010–215856*, edited by: J. H. Morrow, S. B. Hooker, C. R. Booth, G. Bernhard, R. N. Lind, and J. W. Brown, NASA Goddard Space Flight Center, Greenbelt, Maryland, 42–50, (2010).

41. J. H. Morrow, C. R. Booth, R. N. Lind, and S. B. Hooker, "The Compact-Optical Profiling System (C-OPS)," in: *Advances in Measuring the Apparent Optical Properties (AOPs) of Optically Complex Waters, NASA Tech. Memo. 2010–215856*, edited by: J. H. Morrow, S. B. Hooker, C. R. Booth, G. Bernhard, R. N. Lind, and J. W. Brown, NASA Goddard Space Flight Center, Greenbelt, Maryland, 42–50, (2010).

42. S. B. Hooker, J. H. Morrow, and A. Matsuoka, "Apparent optical properties of the Canadian Beaufort Sea – Part 2: The 1% and 1 cm perspective in deriving and validating AOP data products," Biogeosciences **10**(7), 4511–4527 (2013).

43. W. S. Cleveland, "Robust Locally Weighted Regression and Smoothing Scatterplots," J. Am. Stat. Assoc. **74**(368), 829–836 (1979).

44. W. S. Cleveland and S. J. Devlin, "Locally Weighted Regression: An Approach to Regression Analysis by Local Fitting," J. Am. Stat. Assoc. **83**(403), 596–610 (1988).

45. G. Zibordi, K. Ruddick, I. Ansko, G. Moore, S. Kratzer, J. Icely, and A. Reinart, "In situ determination of the remote sensing reflectance: an inter-comparison," Ocean Sci. **8**(4), 567–586 (2012).

46. https://www.seabird.com/systems/free-falling-optical-profiler/family?productCategoryId=54627869942. Accessed on 5th October 2020.

47. G. Zibordi, D. D'Alimonte, and J.-F. Berthon, "An Evaluation of Depth Resolution Requirements for Optical Profiling in Coastal Waters," J. Atm. Ocean. Tech. **21**(7), 1059–1073 (2004).

48. A. Morel, D. Antoine, and B. Gentili, "Bidirectional reflectance of oceanic waters: Accounting for Raman emission and varying particle phase function," Appl. Opt. **41**(30), 6289–6306 (2002).

49. O. Burggraaff, "Biases from incorrect reflectance convolution," Opt. Express **28**(9), 13801 (2020).

50. Sentinel-3 CalVal team, 2016. Technical note: Sentinel-3 OLCI-A spectral response functions. Reports of the European Space Agency, European Space Research and Technology Center, S3-TN-ESA-OL-660, issue 2, October 2016, Noordwijk, The Netherlands.

51. S. W. Bailey and P. J. Werdell, "A multi-sensor approach for the on-orbit validation of ocean color satellite data products," Remote Sens. Environ. **102**(1-2), 12–23 (2006).

52. T. Harmel, A. Gilerson, S. Hlaing, A. Tonizzo, T. Legbandt, A. Weidemann, R. Arnone, and S. Ahmed, "Long Island Sound Coastal Observatory: Assessment of above-water radiometric measurement uncertainties using collocated multi and hyperspectral systems," Appl. Opt. **50**(30), 5842–5860 (2011).

53. C. M. Orrico, R. Van Dommelen, A. H. Barnard, R. Lamb, J. Foesenek, S. Muhammad, K. Brown, M. Dewey, A. Crisp, W. Strubhar, and C. Moore, "Uncertainty budget for Sea-Bird Scientific radiometers following cross-site calibration," Earth and Space Science Open Archive doi : 10.1002/essoar.10500048.1 (2018).

54. J. G. Allen, D. A. Siegel, N. B. Nelson, and S. Halewood, "Controls on Ocean Color Spectra Observed During the North Atlantic Aerosols and Marine Ecosystems Study (NAAMES)," Front. Mar. Sci. **7**, 567007 (2020)..

55. L. Li, D. Stramski, and R. A. Reynolds, "Effects of inelastic radiative processes on the determination of water-leaving spectral radiance from extrapolation of underwater near-surface measurements," Appl. Opt. **55**(25), 7050–7067 (2016).

56. P. Gernez and D. Antoine, "Field characterization of wave-induced underwater light field fluctuations," J. Geophys. Res. **114**(C6), C06025 (2009)..

57. G. Meister, P. Abel, R. Barnes, J. Cooper, C. Davis, M. Godin, D. Goebel, G. Fargion, R. Frouin, D. Korwan, R. Maffione, C. R. McClain, S. McLean, D. Menzies, A. Poteau, J. Robertson, and J. Sherman, "The First SIMBIOS Radiometric Intercomparison (SIMRIC-1), April-September 2001. NASA Tech. Memo. TM-2002, NASA Goddard Space Flight Center, Greenbelt, Maryland, 20771, 66 pp (2002).