Monte Carlo–Based Quantification of Uncertainties in Determining Ocean Remote Sensing Reflectance from Underwater Fixed-Depth Radiometry Measurements

Agnieszka Bialek, Vincenzo Vellucci, Bernard Gentil, David Antoine, Javier Gorroño, Nigel Fox, Craig Underwood

Monte Carlo–Based Quantification of Uncertainties in Determining Ocean Remote Sensing Reflectance from Underwater Fixed-Depth Radiometry Measurements

AGNIESZKA BIAŁEK,[a] VINCENZO VELLUCCI,[b] BERNARD GENTIL,[c] DAVID ANTOINE,[c,d] JAVIER GORROÑO,[a] NIGEL FOX,[a] AND CRAIG UNDERWOOD[e]

[a] *National Physical Laboratory, Teddington, United Kingdom*
[b] *Sorbonne Université, CNRS, Institut de la Mer Villefranche, IMEV, Villefranche-sur-Mer, France*
[c] *Sorbonne Université, UPMC Université de Paris 06, INSU-CNRS, Laboratoire d'Océanographie de Villefranche, Villefranche-sur-Mer, France*
[d] *Remote Sensing and Satellite Research Group, School of Earth and Planetary Sciences, Curtin University, Perth, Western Australia, Australia*
[e] *Surrey Space Center, University of Surrey, Guilford, United Kingdom*

(Manuscript received 29 March 2019, in final form 13 November 2019)

ABSTRACT

A new framework that enables evaluation of the in situ ocean color radiometry measurement uncertainty is presented. The study was conducted on the multispectral data from a permanent mooring deployed in clear open ocean water. The uncertainty is evaluated for each component of the measurement equation and data processing step that leads to deriving the remote sensing reflectance. The Monte Carlo method was selected to handle the data complexity such as correlation and nonlinearity in an efficient manner. The results are presented for a prescreened dataset that is suitable for system vicarious calibration applications. The framework provides uncertainty value per measurement taking into consideration environmental conditions present during acquisition. A summary value is calculated from the statistics of the individual uncertainties per each spectral channel. This summary value is below 4% ($k = 1$) for the blue and green spectral range. For the red spectral channels, the summary uncertainty value increases to approximately 5%. The presented method helps to understand the significance of various uncertainty components and to provide a way of identifying major contributors. This can be used for efficient system performance improvement in the future.

1. Introduction

Geophysical products derived from satellite remote sensing observations are now available from multiple missions, covering most compartments of the Earth system, such as the atmosphere, open oceans, coastal zones, forests and ice sheets. For a few decades, the satellite-derived geophysical variables describing these environments have been essentially used for scientific purposes, yet they are now increasingly feeding operational monitoring program (e.g., Schulz et al. 2009; Traon et al. 2015) and services to governments, industry, or the public (Zhao et al. 2005; Lafaye 2017; Harvey et al. 2015).

In both contexts, users not only require significant suites of products describing their targeted environment but also expect uncertainties to come along with these products. These uncertainties are no longer only required as generic values supposedly representative of an average uncertainty for a given sensor and associated data processing algorithms. They are often now required at the level of individual measurements (pixels). They are of paramount importance when aiming at merging datasets (Gregg and Rousseaux 2014), quantifying long-term trends (Henson et al. 2010; Saulquin et al. 2013), assimilating data into models (Mélin et al. 2016), or still when triggering warnings related to environment monitoring (Rogers et al. 2018).

Satellite ocean color radiometry (OCR) is specifically addressed here. In this case, a prerequisite to deriving geophysical product uncertainties is that uncertainties have been assigned to the quantity from which these products are derived. This quantity is the remote sensing reflectance R_{rs}, derived from the water-leaving radiance L_w, in a number of spectral bands across the near-ultraviolet to visible range; L_w is generally less than 10% of the total radiance measured by the sensor at the top of the atmosphere (TOA), and actually often

Corresponding author: Agnieszka Białek, agnieszka.bialek@npl.co.uk

DOI: 10.1175/JTECH-D-19-0049.1

around 5% only. It was demonstrated that, in such circumstances, achieving a 5% uncertainty on L_w in the blue part of the spectrum, which is a requirement set up rather early in the era of satellite OCR (Gordon and Clark 1981), was impossible without introducing what is now called a system vicarious calibration (SVC) (Gordon and Castaño 1987; Gordon 1997, 1998). The SVC process consists in adjusting the spectral calibration coefficients of the satellite sensors through minimizing differences between the observed TOA total radiances and equivalent modeled values. These modeled values are built from in situ measurements performed in the best possible conditions combined to atmospheric radiances modeled consistently with the satellite processing algorithms. This adjustment is not only needed because of possibly imperfect calibration of the sensor but also because of unavoidable errors in the atmospheric correction process, by which the dominant part of the signal is quantified in order to eventually get to L_w. This is why it is referred to as a *system* vicarious calibration.

To sustain SVC for ocean color sensors (OC-SVC in the following), the solution of deploying permanent moorings in clear open ocean waters has been adopted in the last two decades. The two currently active sites are the Marine Optical Buoy (MOBY; Clark et al. 1997) and the Bouée pour l'Acquisition de Séries Optiques á Long Terme (BOUSSOLE; Antoine et al. 2006). MOBY is deployed off Lanai (Hawaii) and is currently supported by the U.S. National Oceanic and Atmospheric Administration and used for all U.S. ocean color missions including the Moderate Resolution Imaging Spectroradiometer (Salomonson et al. 1989) and the Sea-Viewing Wide Field-of-View Sensor (SeaWiFS; Hooker et al. 1992). BOUSSOLE is deployed in the Mediterranean Sea and is currently supported by the European Space Agency (ESA) and French Centre National d'Études Spatiales (CNES), and it is used for the European missions Medium Resolution Imaging Spectrometer (*Envisat*/MERIS; Rast et al. 1999) and Ocean and Land Color Imager (*Sentinel-3A* and *Sentinel-3B*/OLCI; Donlon et al. 2012). Until recently, data from these two programs were provided to space agencies with a nominal combined standard uncertainty that was supposedly representative of average measurement conditions for the SVC dataset. It was provided for several spectral bands for MOBY (Brown et al. 2007), and as a single number for BOUSSOLE (Antoine et al. 2008b).

The need for a more thorough estimation of uncertainties has, however, become obvious for these crucial in situ measurements to be qualified for being used for OC-SVC. The goal is eventually to assign an uncertainty value to each and every measurement that the field platform provides, so that data can be selected using specific thresholds corresponding to the maximum acceptable uncertainty for various purposes. It is also to have a means of identification and correction of systematic contributions. For OC-SVC, only the data with the lowest uncertainty would be considered. Achieving this objective means that all possible sources of uncertainties are identified, individually quantified, and then combined into a single uncertainty value for each measurement. Deriving this final number is challenging, because it has to combine many uncertainty contributors of which some might be correlated. The often-used first-order approximation of the law of propagation of uncertainties (Joint Committee for Guides in Metrology 2008a), commonly known as quadratic sum of individual uncertainty values is not applicable in this situation.

This paper precisely proposes another approach (Joint Committee for Guides in Metrology 2008b), which uses a Monte Carlo method (MCM). The process involves assigning a probability distribution function (PDF) to each identified uncertainty source, either via statistical analysis of relevant data for type-A uncertainties or via other nonstatistical ways for type-B uncertainties, for example, manufacturer's specifications or a priori knowledge from literature. These uncertainty sources are explicitly introduced into the measurement equation, which is then run multiple times by randomly selecting values in the PDFs of all uncertainty sources.

In the following, we start with reminding how the remote sensing reflectance is derived from in-water radiometry measurements, and then we go through identification and quantification of uncertainty sources related to the instruments and data processing. We then present how the MCM is used to combine uncertainties, and finally we provide an example of application where a subset of BOUSSOLE data are selected following quality criteria that qualify them as a priori suitable for OC-SVC. It is important to note that we do not intend here to propose a thorough and final uncertainty evaluation for the radiometry measurements and derived quantities at BOUSSOLE. We rather use BOUSSOLE as a test bench to demonstrate the methodology. We deliberately ignore a number of uncertainty sources, for instance, when further work is actually still needed to adequately quantify them. We also do not necessarily go through a detailed and fully justified derivation of the PDFs of individual "type B" uncertainties because, again, this paper is a demonstration of the method, not a provision of a full uncertainty quantification for remote sensing reflectance derived from BOUSSOLE measurements. Finally some known sources of errors are not corrected in the current version of the processing, thus they are treated and propagated as uncertainties according to the Guide to the Expression of Uncertainty

in Measurement (GUM; Joint Committee for Guides in Metrology 2008a) recommendations for relatively small biases.

2. Deriving R_{rs} from radiometric quantities measured at fixed depths

a. Basic quantities and equations

The upper-ocean apparent optical property for which measurements uncertainties are to be quantified is the remote sensing reflectance R_{rs}. It is defined as (here omitting wavelength and angular dependencies)

$$R_{rs} = L_w/E_s \quad (\mathrm{sr}^{-1}), \tag{1}$$

where L_w is the water-leaving radiance and E_s is the downward irradiance just above the sea surface. Here, L_w is calculated according to

$$L_w = L_u(0^-)\frac{1-\rho}{n^2} \quad (\mu\mathrm{W\,cm^2\,nm^{-1}\,sr^{-1}}), \tag{2}$$

where $L_u(0^-)$ is the upwelling radiance just beneath the sea surface, ρ is the Fresnel reflection coefficient for the water–air interface, and n is the refractive index of seawater. The ratio in Eq. (2) is often given the approximate value of 0.543 (Austin 1974; Mueller et al. 2003). When underwater measurements are taken at two fixed depths, $L_u(0^-)$ is determined through

$$L_u(0^-) = L_{u,z_u}\exp\left(K_{L_u} z_u\right) \quad (\mu\mathrm{W\,cm^2\,nm^{-1}\,sr^{-1}}), \tag{3}$$

where L_{u,z_u} is the upwelling radiance at the shallowest (upper) measurement depth z_u, and K_{L_u} is the diffuse attenuation coefficient for upwelling radiance, itself computed as

$$K_{L_u} = -\frac{\ln\left(L_{u,z_l}/L_{u,z_u}\right)}{(z_l - z_u)} \quad (\mathrm{m}^{-1}), \tag{4}$$

where L_{u,z_l} is the upwelling radiance at the deepest (lower) measurement depth, that is, $z_l > z_u$. Although the above equations are valid for any zenith angle included into the Snell's cone, we here consider the specific case of upwelling radiances at nadir. Quantifying uncertainties on R_{rs} means quantifying uncertainties 1) on the measurement of the relevant radiometric quantities (here L_{u,z_u}, L_{u,z_l}, and E_s) and depth, 2) on their propagation through Eq. (4) to calculate K_{L_u}, and 3) on the transmission across the water–air interface [Eq. (2)]. Combining Eqs. (2)–(4) leads to the full measurement equation below:

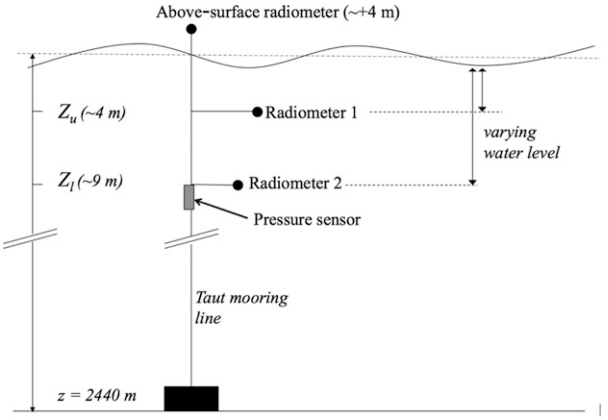

FIG. 1. Schematic layout of how the BOUSSOLE radiometers are installed on the buoy.

$$R_{rs} = \frac{L_{u,z_u}\exp\left[-\frac{\ln\left(L_{u,z_l}/L_{u,z_u}\right)}{z_l - z_u} z_u\right]\left(\frac{1-\rho}{n^2}\right)}{E_s} \quad (\mathrm{sr}^{-1}). \tag{5}$$

The practical realization in the case of the BOUSSOLE buoy is shown in a schematic way on Fig. 1, with $z_u = 4$ m and $z_l = 9$ m. Therefore, from now on L_{u4} and L_{u9} will be used to represent the upwelling radiances at nominal depths of 4 and 9 m, respectively (instead of L_{u,z_u} and L_{u,z_l}). The BOUSSOLE buoy was designed to be stable, so the depths z_u and z_l vary in a limited range (Antoine et al. 2008a).

b. Instruments used

Three multispectral radiometers of the Satlantic OCR-200 series are used to measure L_{u4}, L_{u9} and E_s. They have seven spectral channels at 412, 443, 490, 510, 555, 670, and 683 nm, each with a 10-nm bandwidth. The measurements are taken simultaneously every 15 min throughout the day and night, as 1-min measurement sequences. These radiometers have a 6-Hz frequency of acquisition, resulting in 360 measurements during each sequence. Night measurements are used as dark readings. They allow monitoring of possible drift in instrument noise levels during a deployment. BOUSSOLE functions on an approximately 6-month servicing schedule, with successive rotations on site between a calibrated and verified buoy and the buoy having been deployed for the previous 6 months or so. Therefore, two full sets of radiometers are successively used. Additional instruments on the buoy that are used as an ancillary data source for the processing include a two-axis tilt and compass sensor (Advanced Orientation Systems EZ-III) and a conductivity–temperature–depth

TABLE 1. Relative uncertainties ($k = 1$) of the absolute radiometric calibration for the Satlantic OCR-200-series radiometers used on the BOUSSOLE buoy. Uncertainties associated with the standards are shown first and then the uncertainty associated with using these standards for an absolute radiometric calibration at the NPL laboratory.

Wavelength (nm)	NPL irradiance standard	NPL irradiance calibration	Estimated Satlantic irradiance calibration	NPL reflectance standard	NPL radiance calibration	Estimated Satlantic radiance calibration
412	0.54%	0.84%	1.6%	0.50%	1.00%	2.3%
443	0.48%	0.59%	1.3%	0.50%	0.80%	2.1%
490	0.40%	0.54%	1.2%	0.50%	0.80%	2.1%
510	0.39%	0.73%	1.3%	0.50%	0.90%	2.2%
560	0.38%	0.55%	1.2%	0.50%	0.80%	2.0%
670	0.36%	0.53%	1.1%	0.50%	0.80%	2.0%
683	0.36%	0.49%	1.1%	0.50%	0.80%	2.0%

(CTD) sensor (Sea-Bird Scientific 37-SI). Both sensors are attached to the main buoy structure at 9-m depth. Meteorological data including wind direction and speed and atmospheric pressure are taken from a meteorological buoy that is deployed 2 n mi (1 n mi = 1.852 km) away from the BOUSSOLE site.

3. Instruments characterization and calibration uncertainties

In the following sections, the derivation of uncertainty values corresponding to various steps of the characterization and calibration of instruments is described. Important parameters of these uncertainties are the best estimate (e.g., average typical value) and their PDFs. Both are going to be used when combining all uncertainties via the MCM, which is described later on.

a. Absolute calibration

Since the beginning of BOUSSOLE operations, absolute radiometric calibration is undertaken about once per year for each set of radiometers, before they go at sea for a 6–9-month deployment. This calibration is performed at the manufacturer's grounds (previously Satlantic, Inc., in Halifax, Canada; since 2018 Sea-Bird Scientific in Bellevue, Washington). During that time any necessary instrument repairs are carried out. The absolute radiometric calibration coefficients are provided, yet they are not accompanied by what would be the uncertainty of the calibration process. In 2012, one set of BOUSSOLE radiometers was sent to the United Kingdom's National Measurements Institute [National Physical Laboratory (NPL)] for calibration and characterization tests, as a first step toward establishing an uncertainty budget for their absolute calibration. The difference between NPL- and Satlantic-derived calibration coefficients, combined with the NPL laboratory-calibration uncertainty values (previously evaluated including all effects that are associated with the calibration process, such as the lamp current, aging, and uniformity effects as well as alignment, instrument reading stability, and reflectance panel uniformity), form the radiometric calibration uncertainty used in this work. These derived uncertainty values are presented in Table 1 per spectral channel for irradiance and radiance.

The uncertainties in absolute radiometric calibration have a normal distribution with the mean of the multiplicative correction factor equal to 1. The standard deviation of this distribution is equal to the absolute calibration uncertainty combined with the uncertainty related to the stability of this calibration $u(\text{stab})$, a priori set here at 1% [Eq. (6) below]. The stability has a rectangular distribution and, before being combined with the calibration uncertainty, is divided by the square root of 3 to provide a value corresponding to a normal distribution:

$$u(\text{cal}) = \sqrt{u(\text{abs}_{\text{cal}})^2 + \left[\frac{u(\text{stab})}{\sqrt{3}}\right]^2} \quad (\%), \qquad (6)$$

where $u(\text{cal})$ is the radiometric calibration uncertainty for the instruments and $u(\text{abs}_{\text{cal}})$ is the uncertainty from the laboratory-based absolute radiometric calibration.

The 1% temporal stability assumption was evaluated by comparing the multispectral instruments in use here with the hyperspectral instruments also deployed on the BOUSSOLE buoy. An unbiased percent difference (UPD) was calculated for all instrument pairs (i.e., multi- vs hyperspectral L_u, E_d, and E_s sensors, where E_d is downwelling irradiance measured in water at the same depths as L_u) for a period of 3 months. The slope of the UPD-versus-time relationship gave us a daily relative sensor drift, which, when accumulated over the period under study, ended up as a 0.75% drift.

b. Cosine response

Radiometers measuring plane irradiance, such as the one measuring E_s on top of the BOUSSOLE buoy, are equipped with planar diffusers that collect radiances from all directions of a hemisphere. Ideally the response

of the radiometer is proportional to the cosine of the incident angle, having then what is referred to as a perfect cosine response. Instruments always depart from this theoretical behavior, however, and this departure has to be quantified. It is generally below 3% for incidence angles of <60° and reaches 10% or more for angles of >60° (D. Adams, Satlantic, Inc., 2012, personal communication; Multichannel Visible Detector System update guide MVDS s/n 062). Ideally this departure would be characterized for each instrument, so that a correction for that effect can be applied, and then only a residual uncertainty in that correction would be propagated (Zibordi and Bulgarelli 2007). Such instrument-specific corrections were not available. Cosine response tests were nevertheless performed occasionally to check if the instruments met these requirements, showing results within the above specifications (actually often better).

Total downward irradiance is made of the direct sun irradiance plus the diffuse sky irradiance. Both components are not affected the same way by a nonperfect cosine response. For the direct irradiance the cosine response is related to the incident angle (in this case the sun zenith angle) and is characterized for each possible value. For the diffuse part the cosine response for all angles is integrated over the whole hemisphere. If we assume an isotropic sky radiance distribution, it is calculated according to

$$k_{\cos_h} = \int_0^{\pi/2} k_{\cos}(\theta)\sin(2\theta)\,d\theta, \tag{7}$$

where $k_{\cos_h}$ is the integrated cosine response over the full hemisphere and $k_{\cos}(\theta)$ is a cosine response for a given illumination angle. Actual clear-sky radiance distributions are not isotropic and normally show larger radiances for large angles (e.g., Zibordi and Voss 1989). The $\sin(2\theta)$ factor in Eq. (7) would, however, minimize their contribution to $k_{\cos_h}$. The diffuse component of the downward irradiance for clear skies is about 30% only, further reducing the impact of $k_{\cos_h}$ on the $R_{\rm rs}$ uncertainty evaluation. A thorough evaluation of this effect would nevertheless be timely and would be considered in the succeeding version of the uncertainty evaluation.

The cosine diffuser response or its departure from a perfect cosine response is an example of a systematic error (bias). Because we do not have enough information on the actual cosine response of all our instruments, their possible deviations from an ideal response are not corrected. Therefore, we treat this as an uncertainty component. Although not normally recommended (GUM, note 6.3.1), doing so can be justified when the bias is small. Here we treat the residual bias (the signed sum of the individual biases) as an uncertainty component added to other uncertainty components. We used $k_{\cos}(\theta)$ values of 2% for angles < 20°, 3% for angles > 20° and < 60°, 5% for angles from 60° to 70°, and 10% for angles >70°. The corresponding $k_{\cos_h}$ is 3.5%. These are biases rather than uncertainties, however, which are not corrected so that they are inserted into the model as uncertainty contributors. The PDF for the $k_{\cos}(\theta)$ has a rectangular distribution with the upper limit of 1 + the bias value (relevant to a given angle cosine response) and the lower limit of 1 − the bias value. For example, for solar zenith angle (SZA) < 20° the distribution limits would be [0.98, 1.02] and for SZA = 60° the limits would be [0.95, 1.05]. A similar approach is chosen for $k_{\cos_h}$, where the lower limit is always 0.965 and the upper limit is 1.035.

c. *Temperature dependence*

The temperature dependence for the dark readings was tested at the NPL facility for a maximum temperature range that is expected at the BOUSSOLE site and varies from 5° to 30°C (air temperature outside the E_s sensor). Measurements were taken at four separate temperatures (2°, 12°, 22°, and 32°C), and each of these was repeated twice. The change in the dark signal was less than 0.05% for the entire measured temperature range. The thermal stability of the dark readings ensures that the night measurements can be used for correction of the measurements taken during the day, in particular, with maximum diurnal changes in water temperature of about 2°C at 9 m.

The thermal stability of light readings was checked on the E_s sensor data recorded on site. The check is done using the ratio of measured E_s sensor (in cloud-free conditions) to the clear-sky modeled values (Gregg and Carder 1990) and the correlation of this ratio with the temperature during the measurement. The correlations were found to be insignificant for temperature ranging from 19° to 26°C (from −0.12 for 412 nm to −0.09 for 670 nm; otherwise below −0.04), and, therefore, are not accounted for in the current uncertainty evaluation.

d. *Effects not accounted for*

A few instruments attributes and the uncertainty associated to their characterization are not included in the current version of the uncertainty budget. They include the spectral accuracy, the polarization sensitivity, and the immersion factors. This is not because we consider them negligible a priori but because we do not necessarily have the information to model them properly. The uncertainties at stake here are small anyway, yet there is definitely room for improvement in subsequent versions of the uncertainty budget. The spectral response is defined by the interference filters used in the Satlantic OCR-200-series radiometers with spectral bandwidths

of 10 nm and an out-of-band rejection value provided by the manufacturer of 10^{-6}. These filters were initially characterized but are not recharacterized on a regular basis. They are, however, replaced during manufacturer calibration if the responsivity of that channel changed significantly, minimizing then the risk of substantial spectral errors. The manufacturer does not report the polarization sensitivity for each specific instrument. Nevertheless, the same class of radiometers was used during the Seventh SeaWiFS Intercalibration Round-Robin Experiment (SIRREX-7) intercomparison exercise (Hooker et al. 2002), and showed polarization sensitivity in the range of 1.4%–2.4%, which can be considered as negligible with regard to the state of polarization of underwater nadir radiance (Austin 1974).

Class-based immersion factors for in-water radiance sensors are provided by manufacturers from a basic model based on refractive index of water and glass window (Zibordi 2006). Previous studies, for the type of instruments considered here, report a small negative bias between basic and extended sensor modeling, that is, when a more realistic optical design of the sensors is taken into account (Zibordi 2006). This bias was estimated at −0.4% with the uncertainty in the measurements at the level of 0.19%. More recent studies for the MOBY instrument confirmed that uncertainties of an improved model including transmittance and reflectance of the optical components and the instrument field of view FOV do not differ significantly from those of the basic model [0.05%, (k = 1); Feinholz et al. 2017]. Therefore, the uncertainty related to the immersion factor of radiance radiometers is not included in this budget as it is considered of negligible impact.

4. Uncertainties related to data processing

a. Raw signal statistics (data reduction)

It is reminded here that the median of the 360 values of L_{u4}, L_{u9} and E_s obtained during the 1-min acquisition sequences of instruments operating at 6 Hz are taken as representative values for subsequent calculations of derived quantities such as R_{rs}. Using the median is an efficient way to exclude possible outliers. The dispersion around this median is the result of both the instrument inherent noise plus the changing environmental conditions, in particular, focusing and defocusing by capillary waves. The instrument inherent noise, expressed as a standard deviation of the mean of the 360 individual measurements taken in laboratory conditions, was quantified below 0.1% for all spectral channels and, therefore, is ignored. The standard deviation of the mean of field measurements is used as a measure of changes in the signal caused by environmental changes during the 1-min measurement sequence. Figure 2a shows a typical 1-min series of E_s(442) measurements, showing low-frequency changes due to slight tilt variations during the measurement sequence. Figure 2b shows a typical 1-min series of L_{u4}(670) measurements, showing high-frequency changes due to wave focusing. Figure 2c shows the distribution of the same 360 values, with indication of the mean and median values. They are both represented as black vertical lines, however, nearly superimposed because the distribution is quasi Gaussian. The PDF of the instrument signal will, therefore, be represented by a normal distribution with a mean equal to the median value of 1-min readings and a standard deviation equal to the standard deviation of the mean of

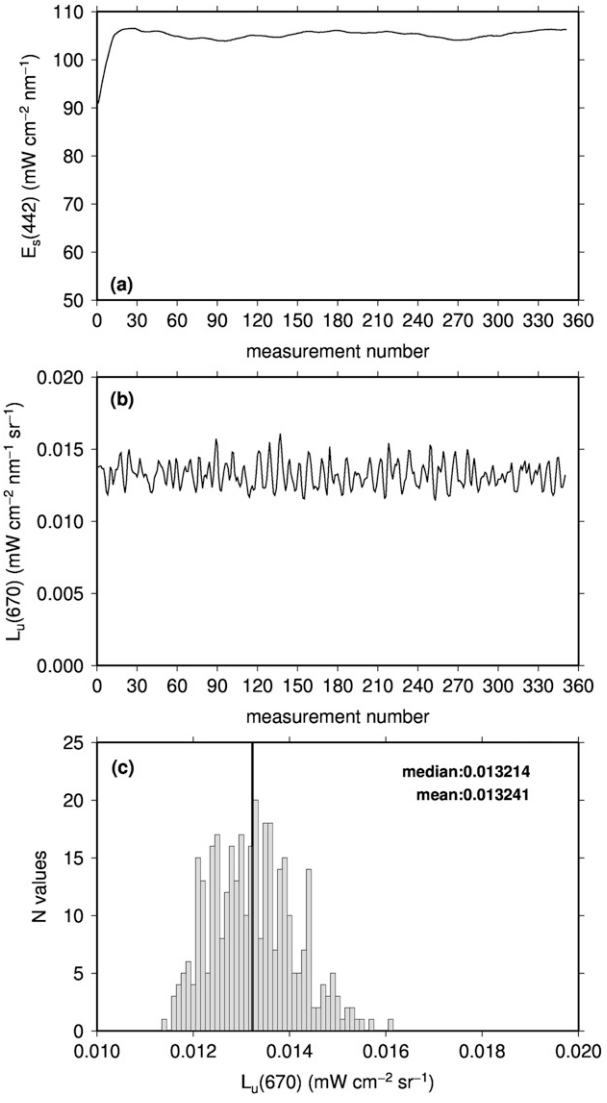

FIG. 2. Examples of 1-min time series of data acquisition for (a) E_s at 442 nm and (b) L_u at 670 nm. Also shown is (c) the distribution of the data in (b).

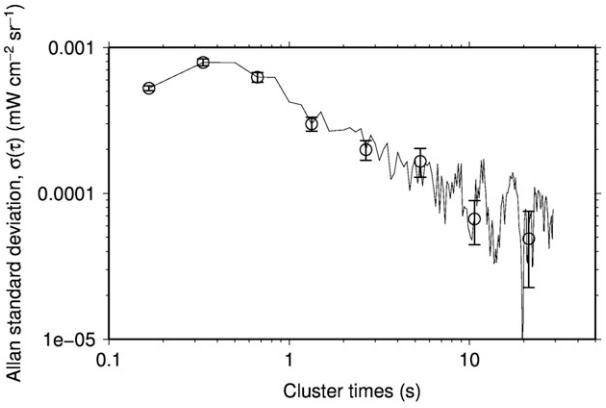

FIG. 3. Allan deviation 1-min time series of data acquisition for L_u at 670 nm.

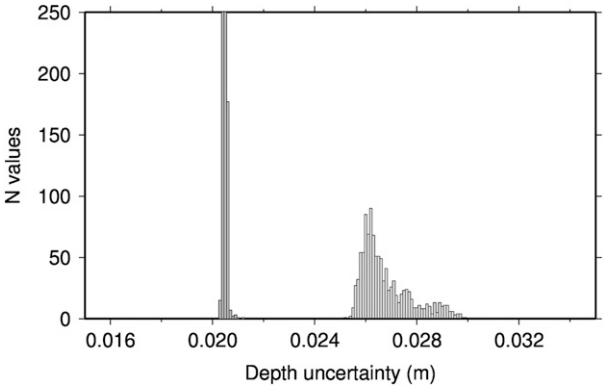

FIG. 4. Uncertainty in the measurement depth for the deepest (uncertainties around 2 cm) and shallowest (uncertainties from 2.5 to 3 cm) buoy arms.

the same signal series. Thus, the actual spread in the instrument readings, which depends mostly on the environmental conditions during the measurement sequence, is included and carried into the uncertainty estimate for each measurement. The standard deviation of the mean is a representative estimate of uncertainty for random and uncorrelated signals. To ensure that the readings in 1-min acquisitions are not correlated, the Allan deviation test (Allan 1966) was performed, which consisted in computing the standard deviation of subsamples (clusters) of various size of the 360 measurements. Figure 3 shows a standard deviation decreasing when the number of samples in each cluster increases, which confirms that measurements are not correlated. The only exception is for the second cluster for which the Allan deviation is higher than for the first sampling interval. This might be caused by matching the sampling rate with the wave-focusing period, thus generating correlation.

b. *Tilt-induced uncertainty for depth measurement*

By design the BOUSSOLE buoy stands vertical when no force are applied to it other than the balance between gravity and Archimedes buoyancy (upthrust). A pressure sensor nominally installed at a depth close to 9 m provides the buoy level with respect to the air–sea interface. The depth offset of each radiometer with respect to this pressure sensor is measured before deployment, so that the actual measurement depths of L_{u4} and L_{u9} are known. These depths change as soon as the buoy inclines under the effect of currents or oscillates when waves pass through. Geometry calculations using the measured two-axis tilt allow the actual depth of each sensor to be determined. The uncertainty in these calculations depends on the inherent accuracy of the pressure and tilt sensors, on the accuracy of the measurements of the depth offsets, and on wave height and frequency. The Sea-Bird 37-SI pressure sensor accuracy is stated by the manufacturer as 0.1%, and the stability is 0.05%. The combined uncertainty of those two gives a value of 0.11%. We assume an uncertainty of 5 mm for the depth-offset measurements, with a rectangular PDF. A specific Monte Carlo (MC) model was run to estimate uncertainty on the instruments depth calculations. Figure 4 presents histograms of the measurements uncertainty for the lower and upper arms, for a selection of data with tilt values lower than 10° and for which the depth recorded by the pressure sensor nominally at 9 m is not lower than 11 m (which means the buoy is not lowered by more than 2 m as compared with its nominal position). Most uncertainty values are around 2 cm for measurements from the deeper arm, increasing to 2.6–2.7 cm for the shallower arm. Higher uncertainties for the instrument located closer to the surface are due to the longer arm length, making the depth change for a given tilt, hence the uncertainty on this depth change, larger than the depth change of the deeper radiometer. The output of this model is integrated to the general BOUSSOLE uncertainty MC model, so that each measurement has an associated depth uncertainty PDF that accounts for actual environmental conditions.

c. *Tilt-induced uncertainty for the radiometric measurements*

Another consequence of the buoy tilt is that upwelling radiances are no longer measured exactly at nadir, and the above-surface downward irradiance no longer corresponds to the irradiance impinging on a horizontal plane. The corresponding uncertainty for upwelling radiance measurements is, therefore, related again to the uncertainty in the tilt measurement, and to changes of the upwelling radiance as a function of the viewing angle. These changes are described by the bidirectional reflectance distribution function (BRDF). This part of the uncertainty budget has not been assessed here. It could

be evaluated later on, for instance through using existing models of the BRDF as a function of the chlorophyll concentration (e.g., Morel et al. 2002). A correction is, however, applied to the direct component of the E_s measurement. This component is determined theoretically for clear-sky conditions through the Gregg and Carder (1990) model. The correction of the direct component uses the ratio of the cosine of the sun zenith angle to the cosine of the actual incident angle on the nonlevel sensor. This correction requires knowledge of the two-axis tilt values, the sun zenith angle and the sun azimuth, because a given absolute tilt value does not have the same impact whether the buoy is tilted toward the sun or away from it. The uncertainty of this tilt correction was estimated by running the calculation using normal PDFs of all input components. The mean is equal to the median value of 1-min acquisitions and the standard deviation is the standard deviation of the mean combined with an additional uncertainty related to the sensor accuracy. This accuracy was stated by the manufacturer as 0.5° for the azimuth (heading sensor) and due to the lack of such information for the tilt, the same value is used. Results are presented in Fig. 5a, for the same dataset as in Fig. 4. Most uncertainty values are lower than 1%. These uncertainties of the tilt correction increases with the sun zenith angle, as shown in Fig. 5b. The outputs from this model are integrated to the main MC model, thus each observation will have an uncertainty in the tilt correction based on the actual measurement conditions during that acquisition.

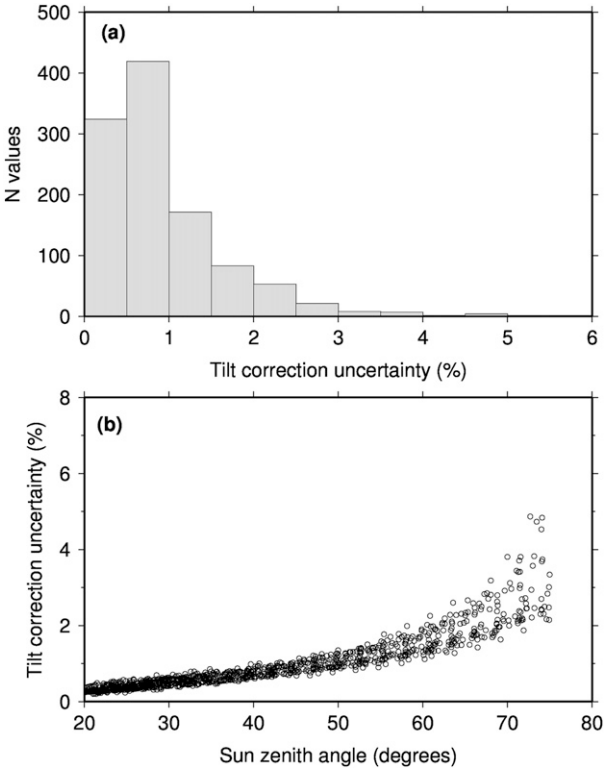

FIG. 5. (a) Distribution of the uncertainty in the tilt correction applied to the E_s measurement, and (b) its change as a function of the sun zenith angle.

d. Shading effects

Self-shading of instruments cannot be avoided and, although the BOUSSOLE buoy has been designed to minimize perturbations of the underwater light field by the buoy structure, it cannot fully eliminate such effects either. The combined self-shading of instruments and instrument shading and reflections of the BOUSSOLE buoy was modeled with the Simulation Optique (SimulO) three-dimensional (3D) backward MC code (Leymarie et al. 2010). This model takes into account the 3D structure of the buoy and instruments, and sets the sun position (zenith and azimuth angles, with 5° steps) with respect to it. Spectral optical properties are modeled as a function of the chlorophyll concentration (0.1, 0.5, 1.0, and 5.0 mg m^{-3}). Running the model with the same environmental conditions in absence of the buoy structure and instrument body allows generating a shading correction coefficients lookup table (LUT). An example of shading correction is presented in Fig. 6. The shading correction coefficient is then interpolated from LUTs and applied as a multiplicative factor to each measurement depth (z_4 and z_9) and wavelength separately. A 5% threshold for the shading correction coefficient was used as rejection criterion in the quality control procedure. The uncertainty in shading corrections is estimated from a model validation exercise that compared the outputs from the SimulO with the results of Piskozub (2004), showing differences of 2%. Although this value cannot fully represent the uncertainty in the shading correction, it is used as an indication until a better solution is found. The shading correction is derived for each spectral band at two depths for every measurement in the data series. In the MC model, that shading correction value is propagated in the form of a rectangular distribution, where the lower and upper limits are defined as the actual shading correction coefficient value ±2%. The 683-nm band is the only exception where, due to very low number of photons the SimulO cannot provide a solution. In the processing a shading correction coefficient from the closest modeled band (670 nm) is used with an increased uncertainty value of ±3%.

e. Extrapolation to the surface

The L_u measurements from nominally 4 and 9 m, L_{u4} and L_{u9}, are used to calculate K_{L_u}, and L_{u4} is then extrapolated to just below the surface to get $L_u(0^-)$

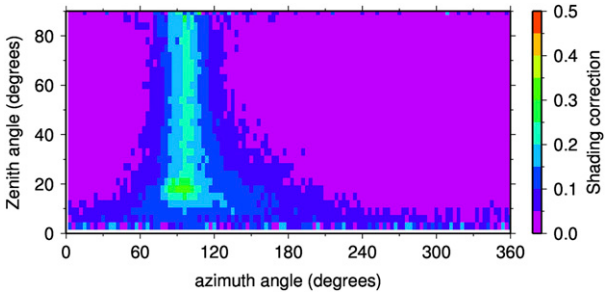

FIG. 6. Example of the shading correction, here for L_{u4} at 510 nm, for a chlorophyll concentration of 0.1 mg(Chl) m^{-3}.

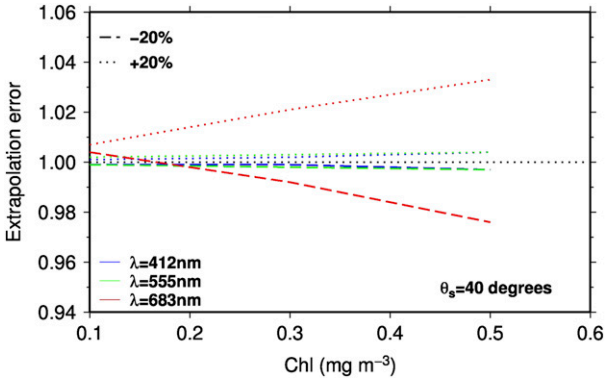

FIG. 7. Changes of the extrapolation correction as a function of the chlorophyll concentration, when the input concentration is changed by ±20% and for the wavelengths indicated.

using Eq. (3). Diffuse attenuation coefficients in the upper layers (first optical depth) are not constant with depth, however. Therefore, the K_{L_u} determined from L_{u4} and L_{u9} might not always provide accurate extrapolation from 4 m to the 0^- level. This is particularly true for wavelengths above 600 nm. A correction to Eq. (3) was, therefore, introduced in Antoine et al. (2008a, their appendix A), which is based on radiative transfer computations including multiple scattering and the Raman contribution (using Hydrolight; Mobley 1994). These computations were run for chlorophyll concentrations and solar zenith angles with the same resolution of 3D MC shading model, and results are used to build an extrapolation correction coefficient LUT. Inputs to the correction are the sun zenith angle and the chlorophyll concentration, which is used as index of the optical properties [see appendix B in Antoine et al. (2008a) for determination of this concentration]. We assume here no uncertainty due to the calculation of the sun zenith angle. The sensitivity of the correction to the chlorophyll concentration was tested for chlorophyll concentrations from 0.1 to 0.6 mg m^{-3}, to which a ±20% error was assigned. Average results over sun zenith angles in the 20°–60° range are presented in Fig. 7. They show a negligible impact in the blue–green spectral range and errors of 2%–3% for red channels. Values of 0.5% are, therefore, used for the overall uncertainty on the extrapolation correction, for wavelengths below 600 nm. For the red bands the value strongly depends on the chlorophyll concentration, the effects of SZA are minimal. Thus, this uncertainty for wavelengths above 600 nm and chlorophyll concentration below 0.25 mg m^{-3} is estimated as 2% and is set at 3% for higher concentrations but not exceeding 0.6 mg m^{-3} (consistent with the dataset considered here).

f. Transmission across the air–sea interface

Although the commonly agreed value of 0.543 for the $(1 - \rho)/n^2$ factor is used in Eq. (2), an uncertainty has to be assigned here as well. The main source of information about this number and its calculation are from Austin (1974) and Austin and Halikas (1976). More recently Wei et al. (2015) tried to confirm the theoretical value with in situ measurements and found in principal good agreements between the two, but the constant value used there was 0.54 and the level of its accuracy of 10% is far lower than the SVC needs. Voss et al. (2017b) also quantified the spectral dependency of this factor. The relative uncertainty of the transmission factor using the traditional GUM approach is given by

$$\left[\frac{u(C_{\rho n})}{C_{\rho n}}\right]^2 = \left[\frac{u(\rho)}{(1-\rho)}\right]^2 + \left[\frac{2u(n)}{n}\right]^2, \tag{8}$$

where $u(\rho)/(1 - \rho)$ is a relative uncertainty in the nominator related purely to the uncertainty in the Fresnel reflection coefficient, and $2u(n)/(n)$ is a relative uncertainty in the denominator of the same equation related to refractive index of seawater n. Because n is squared, the sensitivity coefficient assigned to its uncertainty is 2, which means that the refractive index uncertainty contributes more to the transmission factor uncertainty than that of ρ.

The index of refraction of seawater depends on water salinity, temperature, and pressure and varies spectrally. The tables from Austin and Halikas (1976) were used to verify the range of the n changes. Table 2 presents values for the first three variables recorded during one BOUSSOLE deployment (consistent with the dataset considered here). There is very little change in the salinity and atmospheric pressure for the data records in the current dataset; therefore, these two factors are not considered further. The temperature range is around 6°C; according to the data from Austin and Halikas (1976, their Table 4-2), the difference in n between 20° and 25°C is 0.04%—very small and considered to be negligible. This value increases to 0.15% for the 10°–30°C temperature range, which is representative for the entire

TABLE 2. Summary values of salinity, temperature, and atmospheric pressure during the BOUSSOLE buoy deployment under study.

Obs values	Mean	Median	Std dev σ	Actual % coverage within 1-σ limits	Min/max
Salinity (‰)	38.4	38.4	0.11	90%	37.8/39.6
Water temperature (C)	23.5	23.8	1.7	67%	19.7/26.3
Air temperature (C)	24.9	25.2	1.5	57%	21.9/28.4
Atmospheric pressure (hPa)	1014.1	1014	3.2	77%	1007/1024.6

BOUSSOLE dataset. The change in the refractive index between 412 and 683 nm was calculated using the data from Austin and Halikas (1976, their Table 4-2) for a salinity of 34.99 practical salinity units and an atmospheric pressure of 0 kg cm^{-2}. A second-order polynomial was fitted to the data from the report to estimate the refractive index values. The refractive index estimated range is between 1.3362 and 1.348 94 for wavelengths from 683 to 412 nm, and it is 1.341 99 for the 510-nm band that provides an exact 0.543 constant value. These spectral changes in n are the major contributors to the overall uncertainty, and for the two border wavelengths this effect on $u(C_{\rho n})$, calculated using only the second component of the right-hand side of Eq. (8), is 0.89%, decreasing to 0.53% for the wavelengths between. Major variables affecting the Fresnel reflection coefficient are wind speed and viewing angle. Data considered here were collected with winds speeds below 10 m s^{-2} and for a viewing angle up to 10°. This means that the ρ value could vary from 0.0211 to 0.0218 (Austin 1974). Uncertainty in $C_{\rho n}$ due to ρ, calculated using only the first component of the right-hand side of Eq. (8), is 0.042%, which when combined with the second part of the same equation does not change the final result. Thus, an uncertainty of 0.89% is assigned to 412 nm and red wavelengths, whereas a value of 0.53% is used for the remaining spectral bands. Consequently the PDF for this component has a normal distribution with a mean value 0.543 and a standard deviation varying with wavelength as stated above. This is actually a bias not an uncertainty and according to the GUM this bias should be corrected and then any residual uncertainty related to that correction can be propagated. This is not done in this case and the present uncertainty budget is calculated for the existing processing and currently there is no correction for that.

g. *Solar irradiance modeling*

When evaluating the uncertainty due to imperfect cosine response of the E_s sensor, the ratio of diffuse to direct sunlight was used. This E_s ratio is derived from the Gregg and Carder (1990) spectral solar irradiance model. The model needs wind speed, relative humidity and atmospheric pressure as inputs, which were taken from the meteorological buoy deployed in the vicinity of BOUSSOLE. It also needs the total ozone content, which is calculated daily based on the geographic position and the yearday, along with the precipitable water content, set to 2 cm. It is reported by Gregg and Carder (1990) that their model agrees within 6.2% root-mean-square (RMS) value with spectral irradiance measurements for the wavelength range from 400 to 700 nm. Again, the RMS value is not here a real uncertainty of the model, but this is a good indication of the model capabilities. Actual uncertainty in the model might be lower as the RMS value includes some of the uncertainties that are associated with the spectral irradiance measurements used for its assessment. Nevertheless, we assign a rectangular PDF to the direct to total irradiance fraction $f_{\rm dir}$, where the upper limit is defined as given E_s ratio plus 6.2% of that value and lower limit as given E_s ratio minus 6.2% of that value.

5. Combining uncertainties

a. *The final measurement equation including correction factors*

Equation (5) can now be expanded by including all factors that come into play in the data processing and for which uncertainties have been devised in the previous sections, as follows:

$$R_{\rm rs} = \frac{L_{u4}k_{\rm cal}f_{s_4}\exp\left\{-\dfrac{\ln\left[L_{u9}k_{\rm cal}f_{s_9}\Big/\left(L_{u4}k_{\rm cal}f_{s_4}\right)\right]}{z_9 - z_4}z_4\right\}f_h\left(\dfrac{1-\rho}{n^2}\right)}{E_s k_{\rm cal}k_{\cos}f_{\rm tilt}f_{\rm dir} + (1-f_{\rm dir})E_s k_{\rm cal}k_{\cos_h}}, \tag{9}$$

where L_{u4}, L_{u9}, and E_s are median values of 1-min measurement sequences (in physical units) minus dark readings, $k_{\rm cal}$ represents the uncertainty in absolute radiometric calibration, $k_{\cos}$ and $k_{\cos_h}$ are uncertainties due

TABLE 3. List of defined uncertainty sources.

Uncertainty source (contributor)	Directly affected parameters	Uncertainty evaluation type	Probability distribution shape
Radiance absolute calibration k_{cal}	L_{u4}, L_{u9}	B	Normal
Irradiance absolute calibration k_{cal}	E_s	B	Normal
Instrument stability during deployment (included in k_{cal})	L_{u4}, L_{u9}, E_s	B	Rectangular
Diffuser cosine response	E_s	B	Rectangular
Instrument random noise (statistics of 1-min raw signal)	L_{u4}, L_{u9}, E_s	A	Normal
Wave focusing (statistics of 1-min raw signal)	L_{u4}, L_{u9}	A	Normal
Depth (derived from a separated model with several uncertainty sources)	z_4, z_9, f_{tilt}	B	Normal
Solar zenith angle	k_{cos}, f_s, f_{tilt}	B	Rectangular (k_{cos}, f_s); normal (f_{tilt})
Chlorophyll concentration	f_s, f_h	B	Rectangular (f_s); normal (f_h)
Clear-sky irradiance model	f_{dir}	B	Rectangular
Air–water transmission	$C_{\rho n}$	B	Normal

to the cosine response of the irradiance sensor diffuser affecting the direct sun irradiance and the total diffuse integrated over the hemisphere, f_{s_4} and f_{s_9} are shading corrections for radiance applied at the depth of 4 and 9 m, f_h is the Hydrolight-based extrapolation correction, f_{tilt} is the tilt correction on E_s, and f_{dir} is the fraction of direct to total solar irradiance. All k factors have a mean of their PDFs equal to 1 and PDFs limits defined by a given uncertainty value.

Table 3 summarizes the main sources of uncertainty and parameters from the measurement equation [Eq. (9)] affected by them. Correlation between them can be clearly seen as often more than one parameter is related to a single uncertainty contributor.

b. The Monte Carlo method

To evaluate the uncertainty related to the derivation of R_{rs} through Eq. (9), the GUM (Joint Committee for Guides in Metrology 2008a) method is followed. A quadratic sum of the various individual uncertainties is often used to derive a relevant overall uncertainty, which is indeed defined in the GUM as simplest example to combine uncertainty of uncorrelated inputs. This is, however, clearly not adapted to Eq. (9). Here, the MCM is used, which is known as Supplement 1 to GUM (Joint Committee for Guides in Metrology 2008b). The MCM uses the PDF of each input component instead of its uncertainty value. The measurement model, here Eq. (9), is run a large number of times (here typically 10^5) for a given [L_{u4}, L_{u9}, E_s] triplet, each run including values for the various factors randomly drawn from the PDFs. The many R_{rs} values generated this way have their own PDF, from which the best estimate and its associated standard uncertainty ($k = 1$) value can be derived. When the output PDF is close to Gaussian, the mean and standard deviation are suitable to derive the representative value and its uncertainty. With this MCM, an uncertainty is produced for each R_{rs} value derived from a single [L_{u4}, L_{u9}, E_s] triplet. This uncertainty takes into account the environmental conditions that were present during its acquisition. From this, an overall uncertainty budget can be derived for any dataset that would be used for a given purpose. In the following section, such a budget is derived for a dataset that would be suitable for system vicarious calibration of an ocean color sensor, in which case uncertainties have to be kept at a minimum.

c. Selecting a demonstration dataset

To illustrate the application of the MCM to determining an uncertainty budget for a given dataset, we have selected a subset of BOUSSOLE data following criteria (Table 4) that qualify them a priori for being used for satellite ocean color SVC. These data accordingly have to be of the lowest possible uncertainty. The condition on stability of the 1-min readings is to avoid situations with rapidly changing above-water (clouds and birds) irradiances. The clear-sky test compares the measured downward irradiance with the theoretical value derived from the Gregg and Carder (1990) model and only keeps data for which the difference is within 10%. The tilt and depth conditions eliminate measurements taken with high wind speeds (typically greater than 10 m s^{-1}), which are unwanted because of whitecaps. Keeping data only for sun zenith angles less than 75° eliminates data collected in the early morning or late afternoon, or at midday in winter. Data for which the shading correction is larger than 5% are not considered. Finally, data suspected of being contaminated by biofouling are not included here, as well as data showing inter calibration issues (can be seen for some deployments where the whole dataset tends not to agree with

TABLE 4. SVC quality selection criteria used for the BOUSSOLE dataset.

Selection criterion	
1-min readings stability	<2%
Clear-sky test	$0.9 \le x \le 1.1$
Tilt	<10°
SZA	<75°
Depth	<11 m
Shading	<5%
Biofouling	*N*
Screening for intercalibration issue	Passed

that from other years). The possible presence of biofouling is determined by visual screening of the time series of the following ratios of radiometric quantities: L_{u9}/L_{u4}, L_{u4}/E_s, and L_{u9}/E_s. This is performed at individual wavelengths because biofouling might affect differently each radiometer channels (by design: one collector per wavelength). In particular, the presence of step changes in these ratios after divers' cleaning on the optical surfaces is used as an indicator of biofouling contamination.

We have applied these criteria to data from one summer deployment of BOUSSOLE, from June to August 2008, which resulted in 1090 individual observations. With the method proposed here, there is no need to claim that all of these 1090 observations are necessarily usable for SVC. The criteria we used to select them only make them a priori usable for SVC. Their use for SVC is, however, eventually and only determined by how small the individual uncertainty attached to each of them is. We could have used any equivalent BOUSSOLE data subset for this demonstration. The only difference would have been on the final proportion of measurements having an uncertainty below a predefined threshold within the set of a priori selected measurements.

d. Uncertainty budget for the SVC dataset

The results are presented at different processing steps to show the uncertainty value evolution from the measurement at a single depth through to calculation of the attenuation coefficient from two radiances at different depths and finally to the final product, R_{rs}. The uncertainties of L_{u4} are presented in Fig. 8 for four of the seven spectral channels. They include the uncertainty from instrument calibration, from the signal statistics and from the shading correction k_{cal} and f_{s_4}. These values are driven by the instrument calibration uncertainty, the signal statistics (how much the signal fluctuates during the 1-min sequences), and the uncertainty of the shading correction. The two latter are higher and dominate the uncertainties for red bands (λ = 683 nm in Fig. 8),

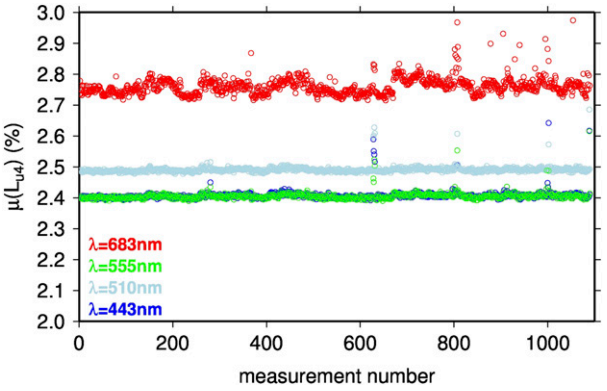

FIG. 8. Relative uncertainty in L_{u4}, (k = 1), for the four wavelengths indicated.

which also have significantly lower signal to noise ratios than the shorter-wavelength channels, for which the environment effects are not noticeable. The uncertainty is higher for λ = 510 nm than for the blue bands, because of the nonuniformity of the calibration source (the ''FEL'' lamp for irradiance and then the reflectance panel illuminated by that lamp for radiance; see Table 1 for radiometric calibration uncertainty values). Outliers with larger uncertainty values indicate imperfect data screening.

The total uncertainties of the attenuation coefficient K_{L_u} include uncertainties of radiometric measurements at 4 and 9 m, their correlation (see the appendix for more details about correlation) and uncertainties in depth measurements. This uncertainty is presented in absolute values per meter in Fig. 9. The highest uncertainties in the attenuation coefficient are observed for the 683-nm band. The remaining spectral channels have similar 0.003 m^{-1} uncertainty. Contrary to what was observed for L_{u4}, the uncertainty for 510 nm is at the same level as the other blue/green spectral bands. The calibration uncertainty that pulled up this value for L_{u4} is no longer as dominate as the calibration uncertainty for the radiometers at two depths are strongly correlated. To account for that correlation the same draw from the k_{cal} PDF is used when deriving L_{u4} and L_{u9} and combining them to get K_{L_u}.

The next processing step is the extrapolation of L_{u4} to the 0^- level, so as to get $L_u(0^-)$. This is where the uncertainty in the Hydrolight-based correction comes into play. Results are shown in Fig. 10. The uncertainty in $L_u(0^-)$ depends on the actual depth where L_{u4} was measured. Indeed, the shorter the extrapolation distance, the more certain the value of $L_u(0^-)$. If the traditional GUM approach using the law of propagation of uncertainty was used, the relative uncertainty in $L_u(0^-)$ would be expressed as

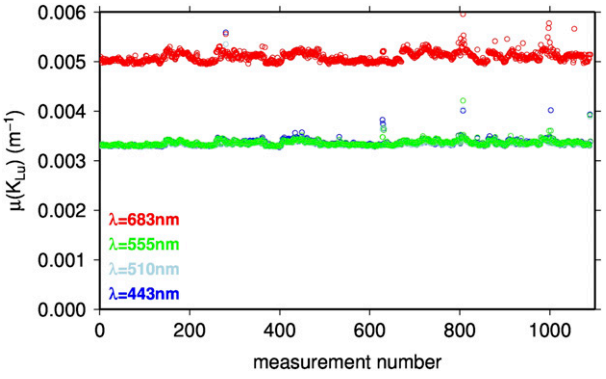

FIG. 9. Absolute uncertainty in K_{L_u}, ($k = 1$), for the four wavelengths indicated.

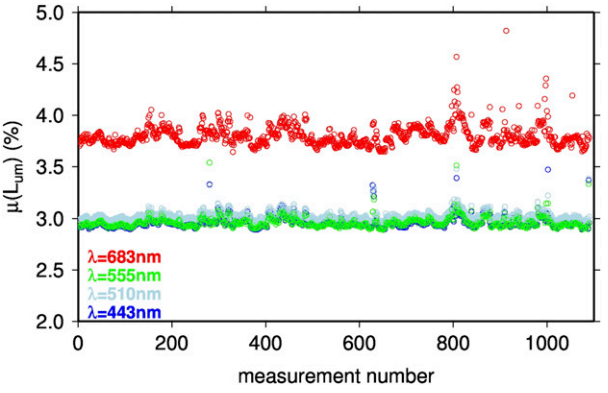

FIG. 10. Relative uncertainty in $L_u(0^-)$, ($k = 1$), for the four wavelengths indicated.

$$\frac{u^2[L_u(0^-)]}{L_u(0^-)^2} = \frac{u^2(L_{u4})}{L_{u4}^2} + z^2u^2\left(K_{L_u}\right) + K_{L_u}^2u^2(z). \quad (10)$$

Thus, the depth of the instruments has a role of sensitivity coefficient for the attenuation coefficients uncertainty. This effect is presented in Fig. 11, where the correlation between the actual depth of the radiometer and the uncertainty in $L_u(0^-)$ is clearly visible.

The uncertainty in L_w is displayed in Fig. 12, where the uncertainty on the transmission factor across the sea–air interface is now included. This uncertainty is low overall yet is slightly higher for the shortest and longest wavelengths.

The uncertainty of E_s is affected by the tilt of the buoy and the sun zenith angle, which determines proportions of diffuse and direct irradiance (Fig. 13). The decrease of uncertainty with increasing sun zenith angle up to values of 60° corresponds to the decreasing proportion of direct light. The step change at a sun zenith angle of 60° is caused by a step change in the uncertainty of the response of the cosine diffuser to direct light (as provided by the manufacturer). The time series of the uncertainty in the final product, R_{rs}, is presented in Fig. 14a. The same data are plotted as a function of SZA on Fig. 14b. The two red channels exhibit the highest uncertainty values because they are more significantly driven by the environmental conditions, as well as by the shading and the extrapolation corrections. For these wavelengths >600 nm the instrument-related uncertainties are lower and the absolute radiometric calibration sources have lower uncertainties. The blue channels have higher instrumental-related uncertainties. The environment does affect them less than the red channels, however; thus, their overall uncertainty is below 4%.

Summary values of these uncertainties are shown in Table 5, for each radiometric quantity and the final R_{rs}, and for each wavelength. The values in Table 5 are calculated as the highest density probability from all uncertainties calculated for each data point. Example histograms of uncertainties values obtained from all data products form the study dataset for selected spectral bands are shown in Fig. 15. They clearly show the most likely reachable uncertainty (mode of the histograms). They also show that lower and higher values can be observed depending on the actual conditions at the site. Key here is that these uncertainties are available for every individual measurement, which then allows selecting datasets using any threshold on uncertainties. This threshold must be low for SVC of satellite ocean color observations, might be somewhat relaxed for validation of OCR-derived R_{rs}, and will depend on the question at hand when it comes to analyzing data for science purposes. The model results were found to be stable within 0.1% by comparing the summary output values from multiple runs of the same simulation (using 10^5 individual runs each time).

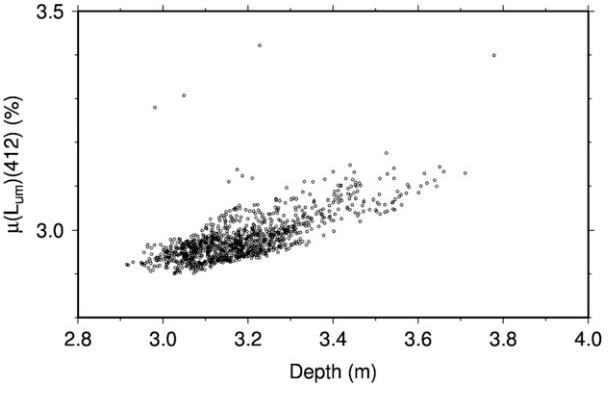

FIG. 11. Relative uncertainty in $L_u(0^-)$, ($k = 1$), at 412 nm as a function of the actual depth of the shallowest arm.

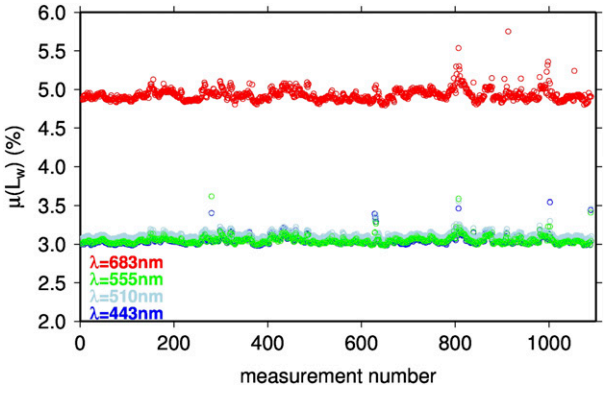

FIG. 12. Relative uncertainty in L_w, ($k = 1$), for the four wavelengths indicated.

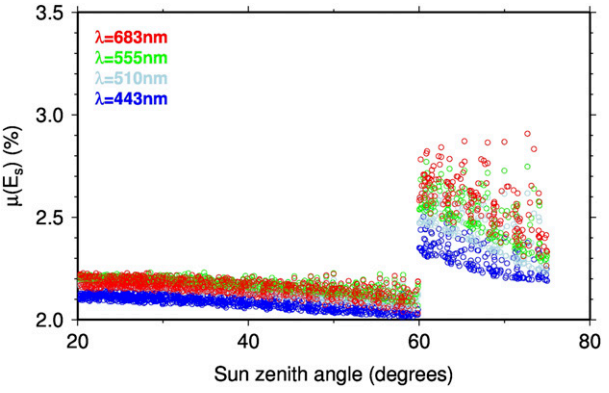

FIG. 13. Relative uncertainty in E_s, ($k = 1$), as a function of the sun zenith angle, for the four wavelengths indicated.

6. Conclusions

A method has been presented that allows providing an uncertainty estimate when deriving the remote sensing reflectance R_{rs} from measurements of underwater upwelling nadir radiances at two depths combined with the above-water plane irradiance. This uncertainty is derived from combining individual uncertainty sources (values) through a Monte Carlo method. A prerequisite is that these individual sources of uncertainty have been identified and quantified. In this particular case, their probability distribution functions have to be either determined for ''type A'' uncertainties or specified for ''type B'' uncertainties. Contrary to using a quadratic sum of individual uncertainty components, the MCM allows addressing correlations between some of the input components. This can be done either explicitly (e.g., by selective sampling of the PDFs of each input component) or implicitly through the way these input components are combined through the measurement equation [here Eq. (9)]. As a practical example, we applied the method to a subset of data from the BOUSSOLE time series. The goals are to illustrate the capability of the method and to provide a preliminary uncertainty budget for the BOUSSOLE radiometric measurements, when these measurements respect criteria making them a priori usable for satellite ocean color SVC (such as a low buoy tilt). This method has the advantage of evaluating uncertainty for each observation instead of providing one generic value for an entire dataset or project. These uncertainty values include the instrument-related uncertainties, the processing-related uncertainties, and the effect of changes in environmental conditions during the measurements. It is then up to any data user to decide whether a particular measurement with its associated uncertainty is suitable for further use or not. Any field site that delivers yearlong records of radiometry measurements taken throughout each and every day will end up delivering data whose uncertainties are going to be highly variable. When taken in ideal conditions, for example, low chlorophyll concentration, low wind speed, clear sky and high solar zenith angle, the measurement uncertainty will be low (assuming instruments are well calibrated and characterized). When those conditions are not met, higher uncertainties will occur. That is why defining a typical uncertainty for a given site does not

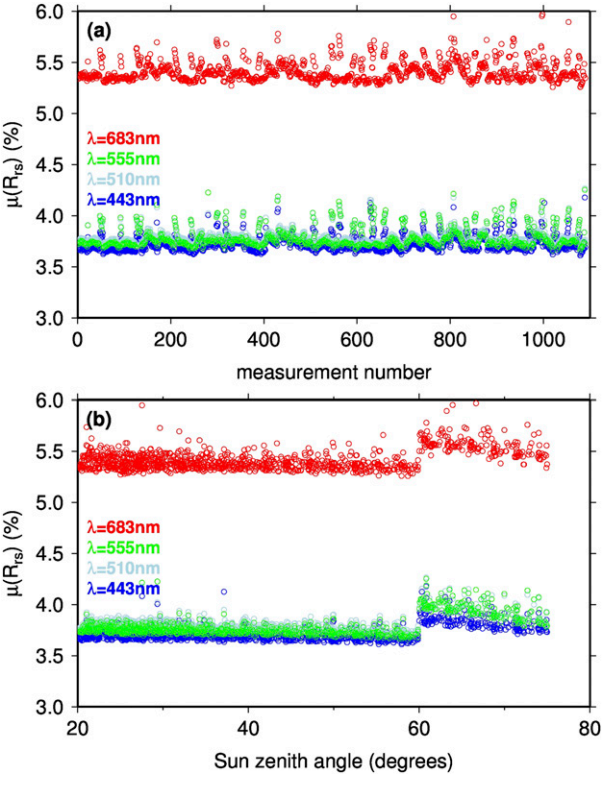

FIG. 14. (a) Relative uncertainty in R_{rs}, ($k = 1$), for the four wavelengths indicated, and (b) the same data as in (a) plotted as a function of the sun zenith angle.

TABLE 5. Uncertainty budget of Satlantic OCR-200-series radiometers deployed onto the BOUSSOLE buoy ($k = 1$).

Wavelength (nm)	E_s	L_{u4}	$L_u(0^-)$	L_w	R_{rs}
412	2.29%	2.40%	2.96%	3.14%	3.86%
443	2.11%	2.40%	2.95%	3.04%	3.68%
490	2.16%	2.40%	2.94%	3.02%	3.70%
510	2.17%	2.49%	3.01%	3.09%	3.77%
560	2.20%	2.40%	2.93%	3.02%	3.73%
670	2.23%	2.43%	3.03%	4.38%	4.88%
683	2.17%	2.78%	3.78%	4.90%	5.35%

make sense if the typical range of environment conditions is not specified. For example, an average uncertainty might be delivered for all BOUSSOLE measurements taken within an hour of solar noon and when the sky is clear. Another uncertainty budget would be obtained if all measurements from dawn to dusk are considered. And still another number if cloudier conditions would be included. That is where the individual uncertainties will reveal powerful when the full time series will have been characterized: when ocean color SVC is at stake, we can filter out any measurement whose uncertainty is larger than a given threshold, so that we provide a qualified SVC dataset. The threshold could be larger if the data were to be used for another purpose, ending up with more data being usable (could be, e.g., for geophysical products validation or for bio-optics research). This method also allows better understanding of the respective importance of various uncertainty sources in driving the final uncertainty either on a given measurement or for a given ensemble of measurements. This capability provides a way to improve either the experimental set up or the data processing chain or both. For example, the uncertainty on the air–water transmission constant has been shown to have a significant impact,

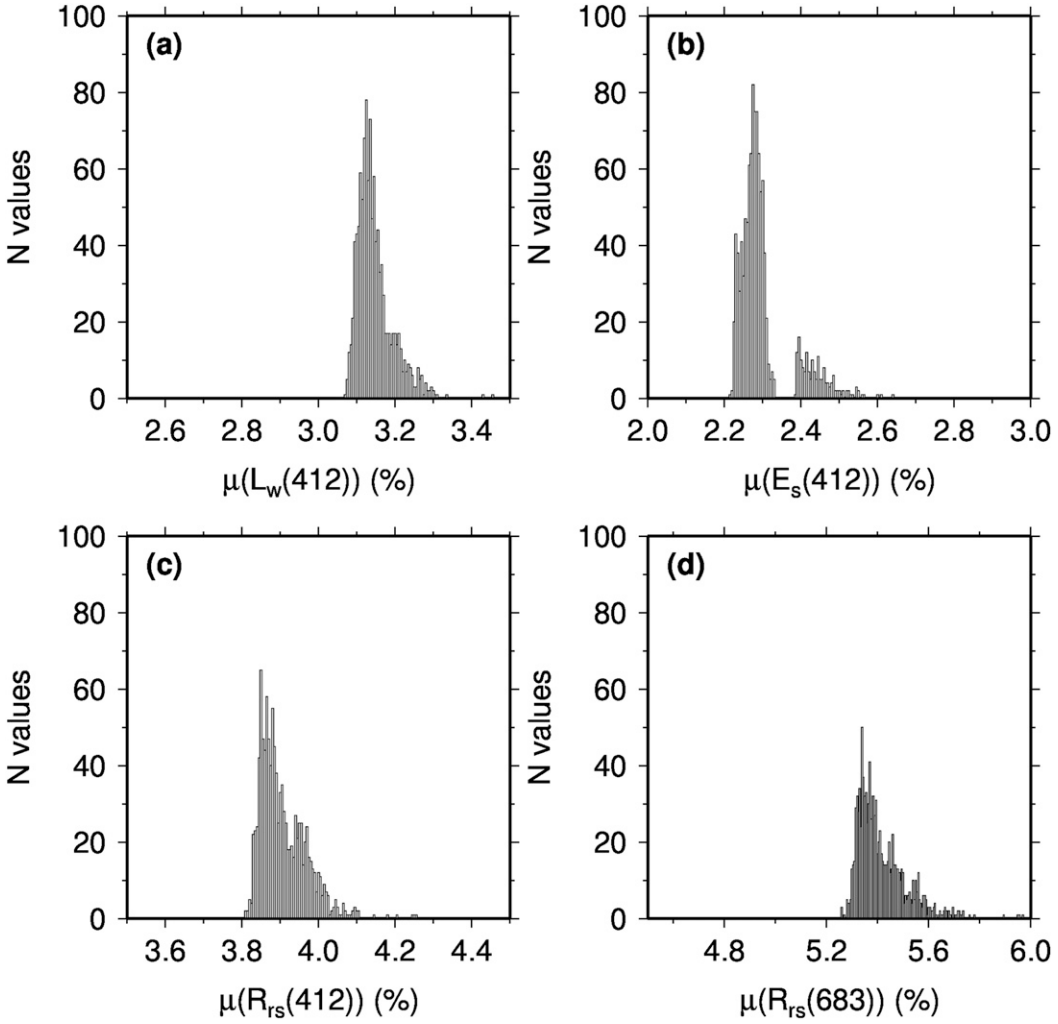

FIG. 15. Example distributions of relative uncertainties ($k = 1$) for the parameters and wavelengths indicated.

and then could be made wavelength dependent in future versions of the data processing. Similarly, the observed correlation between the depth of the shallowest instrument and $L_u(0^-)$ advocates for moving the shallowest depth of measurement closer to the surface. Stronger wave focusing would, however, affect the measurement, and then increasing uncertainty on that side. Therefore, the overall uncertainty has to be considered globally for the whole system, inasmuch as individual components separately affect different aspects of the measurements. This example shows that this framework in not only helpful to provide users with fit-for-purpose criteria but also to better understand the critical components of the system and ultimately improve it. The uncertainty in E_s is affected by the nonperfect cosine response of the diffuser, and, therefore, is related to the sun zenith angle. Recent recommendations (Mazeran et al. 2017) are to derive R_{rs} using a modeled value of E_s instead of the measured one, when R_{rs} is to be used for SVC purposes. The logic is that using the same model as used by the satellite data processing chain, rather than the in situ measurement, would improve consistency and accuracy of the SVC. If going that way, trying to reduce uncertainties on the E_s measurement would be of less importance in the context of SVC.

The R_{rs} relative uncertainty for the red channels in clear waters such those at BOUSSOLE will always be larger because of low signal levels. As the final quantity is close to zero, the relative uncertainty is high. For example, a typical value of L_w at 440 nm is $0.8458\,\mu W\,cm^2\,nm^{-1}\,sr^{-1}$ and relative uncertainty 3.11% (and 0.0226 in radiance units). At 680 nm, L_w is $0.0187\,\mu W\,cm^2\,nm^{-1}\,sr^{-1}$ with relative uncertainty 4.90% (and 0.0008 in radiance units).

Further work will apply the same methodological framework to hyperspectral radiometers that are also in operation on the BOUSSOLE buoy. These radiometers will have more uncertainty components related to their characteristics and differences in their operation mode, such as varying integration times and more dark readings interspersed among light measurements. A detailed and comprehensive uncertainty evaluation will then become reachable.

Acknowledgments. BOUSSOLE is funded by the European Space Agency, the Centre National d'Études Spatiales, the Centre National de la Recherche Scientifique (CNRS), the Institut National des Sciences de l'Univers (INSU), the Sorbonne Université, and the Institut de la Mer de Villefranche (IMEV). The authors acknowledge Météo-France for supplying data from the Azur meteorological buoy and the HyMeX database teams (ESPRI/IPSL and SEDOO/Observatoire Midi-Pyrénées) for their help in accessing the data. Edouard Leymarie is kindly acknowledged for running SimulO computations. This research was partly supported by the European Association of National Metrology Institutes (EURAMET) European Metrology Research Programme (EMRP) ENV-53 project. The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union.

APPENDIX

Error Correlation

First, we remind here the basic definitions of an error and measurement uncertainty as they are stated in the GUM and how there they are applied in the MCM approach.

An error is "the difference between a measurement value and a reference value." An uncertainty is "a parameter characterizing the dispersion of the values being attributed to a quantity, based on information used." For the MCM model the reference values are the means of PDFs, and the uncertainty is the standard deviation of the PDF, if that PDF has a Gaussian shape or is close enough to a Gaussian shape. The errors are the differences from the reference value used in the MCM processing and individual draws from the PDF assigned to this value. When we refer to correlation, it is important to note that only errors of a given quantity can be correlated; therefore, we talk about error (and not uncertainty) correlations. The error correlations between input quantities of the measurement equation have been assessed for two wavelengths (412 and 683 nm). Results are presented in Figs. A1 and A2. They show error correlations between the input quantities used to derive L_w (see the figure legend). The 1000 error values for each quantity are extracted from the MCM calculations for one measurement from the database. The diagonal of the figure shows the PDFs of each quantity. The lower triangle shows the scatterplot between two components, and the upper triangle shows the correlation value as a number and its significance level, or p value. The size of the displayed correlation coefficients depends on their value; thus, lower coefficient values are made explicitly smaller on the plot. Significance level is indicated with symbols, with three asterisks indicating $p < 0.001$, two asterisks is < 0.01, one asterisk is < 0.05, a small solid square is < 0.1, and no symbol indicates that the p value is > 0.1. As expected, significant correlation between the errors of L_{u4} and L_{u9} can be seen for both wavelengths ($r = 0.76$ for $\lambda = 412$ m; $r = 0.59$ for $\lambda = 683$ m).

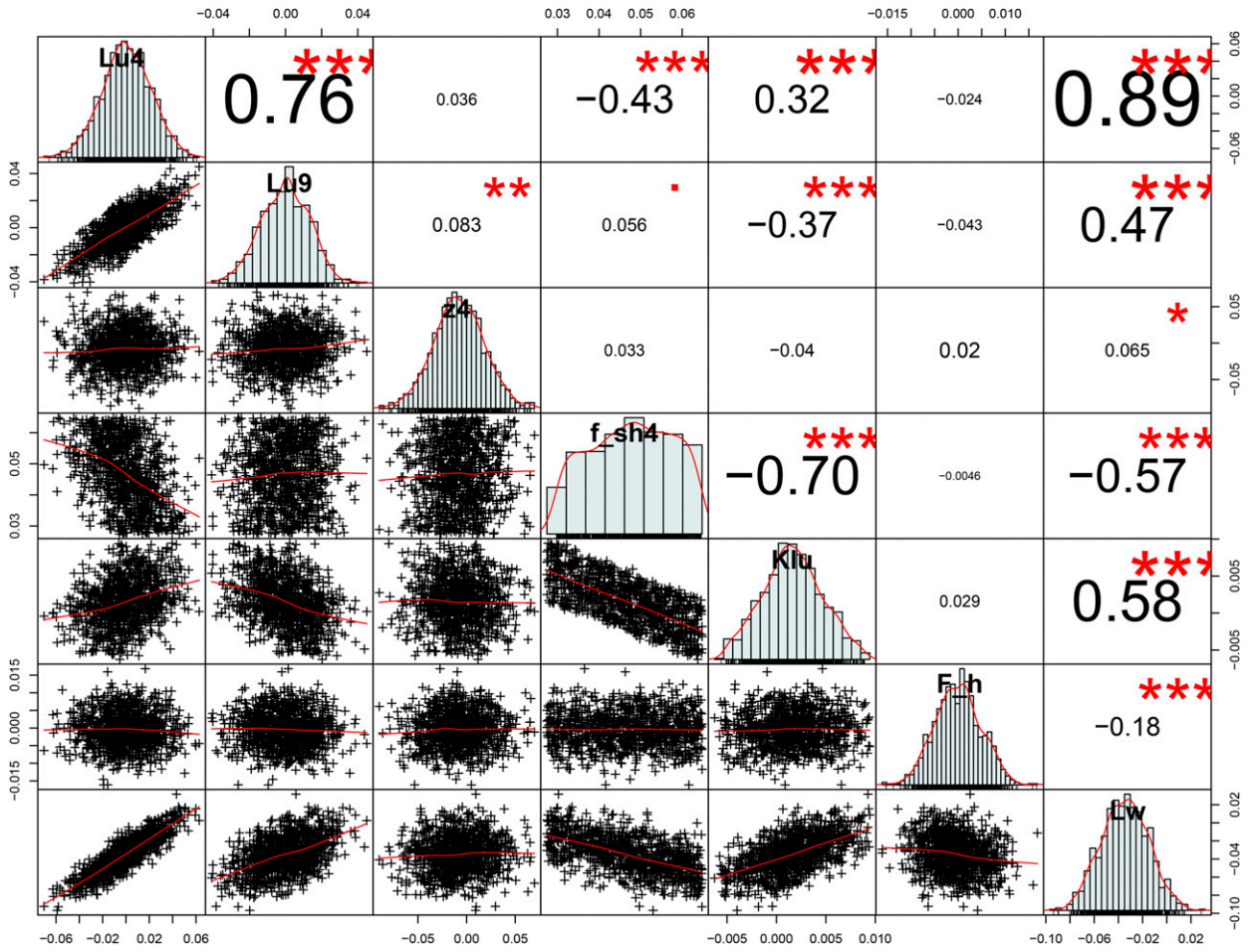

FIG. A1. Error correlation for 412-nm input components for L_w. The variables presented in columns are the upwelling radiance at 4-m depth L_{u4}, the upwelling radiance at 9-m depth L_{u9}, the depth at upper arm $z4$, the shading correction for upper arm f_{sh4}, the diffuse attenuation coefficient for upwelling radiance K_{L_u}, the Hydrolight-based extrapolation correction f_h, and the water-leaving radiance L_w. The content of this figure is explained in more detail in the appendix text.

This is caused by the absolute calibration (instruments were calibrated in the same laboratory against the same absolute radiometric standards), whose uncertainty is the main contributor to uncertainty in L_u measurements. The reason for the lower correlation at 683 nm is related to the random uncertainty components related to the signal noise, which is larger for the red wavelengths (see Fig. 8). Also, although environmental conditions are common to the measurements at 4 and 9 m, they affect them differently. For example, wave-focusing effects at 4 and 9 m are different and the random instrument noise has different levels depending on the depth.

The errors in the measurement of depth are not correlated with any of the other errors presented here. The shading correction errors of the upper arm show a negative correlation with the upwelling radiance measurement errors at this depth, and to all other components where this measurement is used, and thus to errors of K_{L_u} and L_w. The negative relationship is expected because the errors are presented in the absolute form. Thus, higher L_u values are less affected by the shading correction error. The shading correction errors are drawn from the rectangular PDF defined as 2% of the shading correction value. The error of the Hydrolight correction is uncorrelated with any of the input parameters but presents some correlation with the errors in L_w. This is relatively low for the blue channel ($r = -0.21$) and is more significant for the red one ($r = -0.63$). The above explanation for the negative correlation for shading correction applies to errors in the Hydrolight correction as well. Although, the error correlation is usually presented for the input components, in this analysis we assume the L_w as the output quantity. This is an example of additional information that can be seen on the scatterplots presented on Figs. A1 and A2. The bottom row of the scatterplots shows how the output

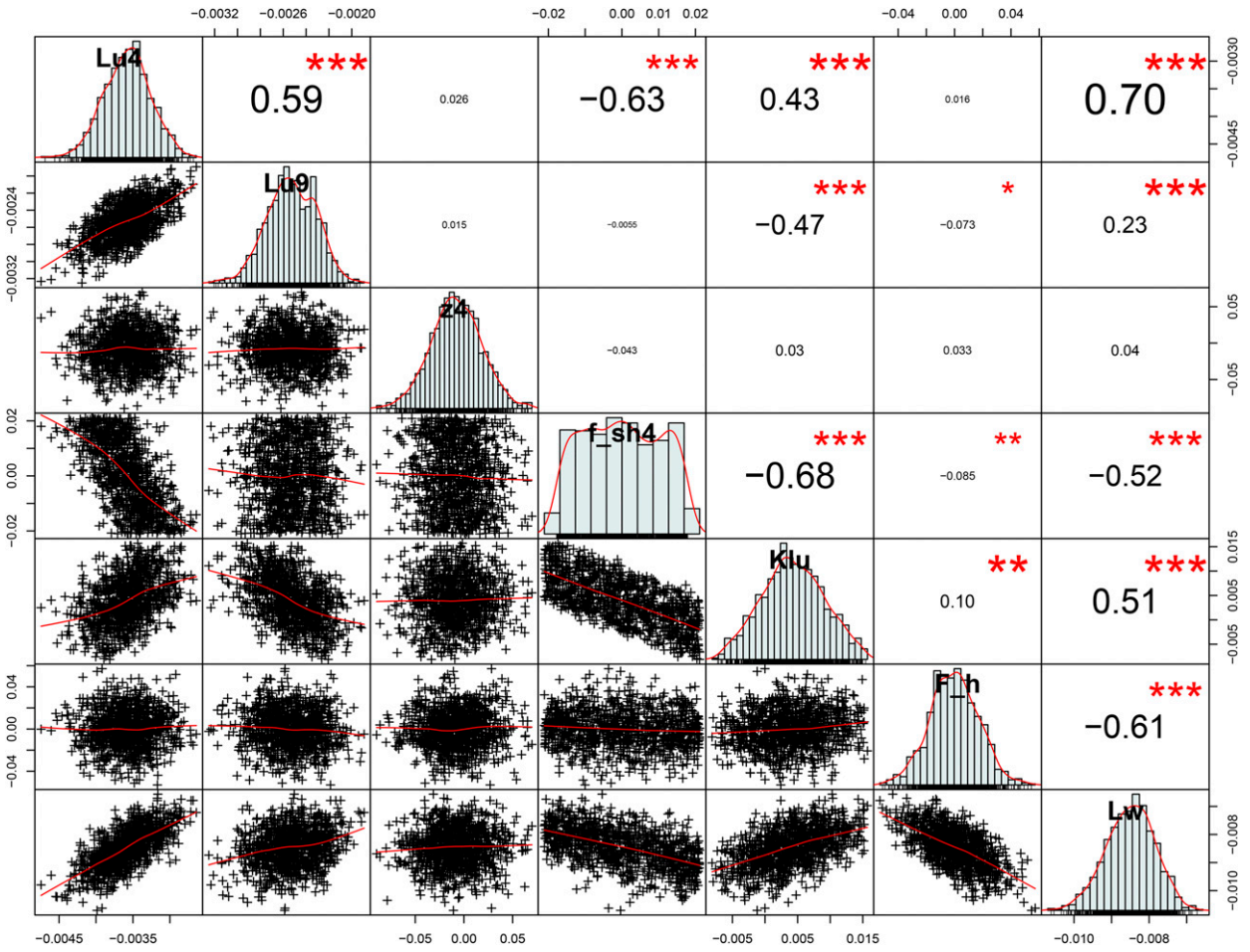

FIG. A2. As in Fig. A1, but for 683 nm.

quantity errors are correlated with the input errors. The last column holds error correlation values of the input components and L_w. This can be interpreted as how much weight a given component has on the L_w uncertainty. Thus, the uncertainty on the water-leaving radiance is mainly driven by the uncertainty in the L_u measurements at the shallowest depth, because their errors are highly correlated: the value in the top-right corner of Figs. A1 and A2 is the highest in this column. Then we start to see differences between the 412- and 683-nm spectral bands. Clearly, for the blue wavelength the measurement uncertainty from 9 m still contributes to the result with the error correlation value of 0.47 in the second row on the last column in Fig. A1, whereas for the red channel the equivalent number is only 0.18.

REFERENCES

Allan, D. W., 1966: Statistics of atomic frequency standards. *Proc. IEEE*, **54**, 221–230, https://doi.org/10.1109/PROC.1966.4634.

Antoine, D., and Coauthors, 2006: BOUSSOLE: A joint CNRS-INSU, ESA, CNES, and NASA ocean color calibration and validation activity. NASA Tech. Rep. NASA/TM-2006-214147, 68 pp., https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20070028812.pdf.

——, P. Guevel, J.-F. Deste, G. Becu, F. Louis, A. J. Scott, and P. Bardey, 2008a: The ''BOUSSOLE'' buoy—A new transparent-to-swell taut mooring dedicated to marine optics: Design, tests, and performance at sea. *J. Atmos. Oceanic Technol.*, **25**, 968–989, https://doi.org/10.1175/2007JTECHO563.1.

——, F. d'Ortenzio, S. B. Hooker, G. Becu, B. Gentili, D. Tailliez, and A. J. Scott, 2008b: Assessment of uncertainty in the ocean reflectance determined by three satellite ocean color sensors (MERIS, SeaWiFS and MODIS-A) at an offshore site in the Mediterranean sea (BOUSSOLE project). *J. Geophys. Res.*, **113**, C07013, https://doi.org/10.1029/2007JC004472.

Austin, R. W., 1974: The remote sensing of spectral radiance from below the ocean surface. *Optical Aspects of Oceanography*, N. G. Jerlov and E. Steemann-Nielsen, Eds., Elsevier, 317–344.

——, and G. Halikas, 1976: The index of refraction of seawater. Scripps Institution of Oceanography Rep.

AD-A024 800, 134 pp., https://apps.dtic.mil/dtic/tr/fulltext/u2/a024800.pdf.

Brown, S., and Coauthors, 2007: The Marine Optical Buoy (MOBY) radiometric calibration and uncertainty budget for ocean color satellite sensor vicarious calibration. *Proc. SPIE*, **6744**, 67441M, https://doi.org/10.1117/12.737400.

Clark, D. K., H. R. Gordon, K. J. Voss, Y. Ge, W. Broenkow, and C. Trees, 1997: Validation of atmospheric correction over the oceans. *J. Geophys. Res.*, **102**, 17 209–17 217, https://doi.org/10.1029/96JD03345.

Donlon, C., and Coauthors, 2012: The Global Monitoring for Environment and Security (GMES) Sentinel-3 mission. *Remote Sens. Environ.*, **120**, 37–57, https://doi.org/10.1016/j.rse.2011.07.024.

Feinholz, M., B. C. Johnson, K. Voss, M. Yarbrough, and S. Flora, 2017: Immersion coefficient for the Marine Optical Buoy (MOBY) radiance collectors. *J. Res. Natl. Inst. Stand. Technol.*, **122**, 31, https://doi.org/10.6028/jres.122.031.

Gordon, H. R., 1997: Atmospheric correction of ocean color imagery in the Earth Observing System era. *J. Geophys. Res.*, **102**, 17 081–17 106, https://doi.org/10.1029/96JD02443.

——, 1998: In-orbit calibration strategy for ocean color sensors. *Remote Sens. Environ.*, **63**, 265–278, https://doi.org/10.1016/S0034-4257(97)00163-6.

——, and D. K. Clark, 1981: Clear water radiances for atmospheric correction of coastal zone color scanner imagery. *Appl. Opt.*, **20**, 4175–4180, https://doi.org/10.1364/AO.20.004175.

——, and D. J. Castaño, 1987: Coastal zone color scanner atmospheric correction algorithm: Multiple scattering effects. *Appl. Opt.*, **26**, 2111–2122, https://doi.org/10.1364/AO.26.002111.

Gregg, W. W., and K. Carder, 1990: A simple spectral solar irradiance model for cloudless maritime atmospheres. *Limnol. Oceanogr.*, **35**, 1657–1675, https://doi.org/10.4319/lo.1990.35.8.1657.

——, and C. S. Rousseaux, 2014: Decadal trends in global pelagic ocean chlorophyll: A new assessment integrating multiple satellites, in situ data, and models. *J. Geophys. Res. Oceans*, **119**, 5921–5933, https://doi.org/10.1002/2014JC010158.

Harvey, E. T., S. Kratzer, and P. Philipson, 2015: Satellite-based water quality monitoring for improved spatial and temporal retrieval of chlorophyll-a in coastal waters. *Remote Sens. Environ.*, **158**, 417–430, https://doi.org/10.1016/J.RSE.2014.11.017.

Henson, S. A., J. L. Sarmiento, J. P. Dunne, L. Bopp, I. Lima, S. C. Doney, J. John, and C. Beaulieu, 2010: Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity. *Biogeosciences*, **7**, 621–640, https://doi.org/10.5194/bg-7-621-2010.

Hooker, S. B., W. E. Esaias, G. C. Feldman, W. W. Gregg, and C. R. Mc Clain, 1992: An overview of SeaWiFS and ocean color. NASA Tech. Memo. 104566, Vol. 1, 24 pp.

——, S. McLean, J. Sherman, M. Small, G. Lazin, G. Zibordi, and J. W. Brown, 2002: The Seventh SeaWiFS Intercalibration Round-Robin Experiment (SIRREX-7). NASA Tech. Memo. 2002-206892, Vol. 17, 69 pp.

Joint Committee for Guides in Metrology, 2008a: Evaluation of measurement data—Guide to the expression of uncertainty in measurement. Bureau International des Poids et Mesures Guidance Doc. JCGM 100, 134 pp.

——, 2008b: Evaluation of measurement data—Supplement 1 to the ''Guide to the expression of uncertainty in measurement''—Propagation of distributions using a Monte Carlo method. Bureau International des Poids et Mesures Guidance Doc. JCGM 101, 90 pp.

Lafaye, M., 2017: Benefit assessment of the application of satellite Earth observation for society and policy: Assessing the socioeconomic impacts of the development of downstream space-based Earth observation applications. *Satellite Earth Observations and Their Impact on Society and Policy*, M. Onoda and O. R. Young, Eds., Springer, 93–103.

Leymarie, E., D. Doxaran, and M. Babin, 2010: Uncertainties associated to measurements of inherent optical properties in natural waters. *Appl. Opt.*, **49**, 5415–5436, https://doi.org/10.1364/AO.49.005415.

Mazeran, C., C. Brockmann, K. Ruddick, K. Voss, and F. Zagolski, 2017: Requirements for Copernicus ocean colour vicarious calibration infrastructure. EUMETSAT Tech. Rep. SOLVO/EUM/16/VCA/D8, 92 pp.

Mélin, F., G. Sclep, T. Jackson, and S. Sathyendranath, 2016: Uncertainty estimates of remote sensing reflectance derived from comparison of ocean color satellite data sets. *Remote Sens. Environ.*, **177**, 107–124, https://doi.org/10.1016/j.rse.2016.02.014.

Mobley, C. D., 1994: *Light and Water: Radiative Transfer in Natural Waters*. Academic Press, 592 pp.

Morel, A., D. Antoine, and B. Gentili, 2002: Bidirectional reflectance of oceanic waters: Accounting for Raman emission and varying particle scattering phase function. *Appl. Opt.*, **41**, 6289–6306, https://doi.org/10.1364/AO.41.006289.

Mueller, J. L., and Coauthors, 2003: Ocean optics protocols for satellite ocean color sensor validation, revision 4, volume V: Special topics in ocean optics protocols and appendices. NASA Tech. Rep., 148 pp.

Piskozub, J., 2004: Effect of 3-D instrument casing shape on the self-shading of in-water upwelling irradiance. *Opt. Express*, **12**, 3144–3148, https://doi.org/10.1364/OPEX.12.003144.

Rast, M., J. L. Bezy, and S. Bruzzi, 1999: The ESA medium resolution imaging spectrometer MERIS a review of the instrument and its mission. *Int. J. Remote Sens.*, **20**, 1681–1702, https://doi.org/10.1080/014311699212416

Rogers, A., J. L. Blanchard, and P. J. Mumby, 2018: Fisheries productivity under progressive coral reef degradation. *J. Appl. Ecol.*, **55**, 1041–1049, https://doi.org/10.1111/1365-2664.13051.

Salomonson, V. V., W. L. Barnes, P. W. Maymon, H. E. Montgomery, and H. Ostrow, 1989: MODIS: Advanced facility instrument for studies of the Earth as a system. *IEEE Trans. Geosci. Remote Sens.*, **27**, 145–153, https://doi.org/10.1109/36.20292.

Saulquin, B., R. Fablet, A. Mangin, G. Mercier, D. Antoine, and O. Fanton d'Andon, 2013: Detection of linear trends in multisensor time series in the presence of autocorrelated noise: Application to the chlorophyll-a SeaWiFS and MERIS data sets and extrapolation to the incoming Sentinel 3-OLCI mission. *J. Geophys. Res. Oceans*, **118**, 3752–3763, https://doi.org/10.1002/jgrc.20264.

Schulz, J., and Coauthors, 2009: Operational climate monitoring from space: The EUMETSAT Satellite Application Facility on Climate Monitoring (CM-SAF). *Atmos. Chem. Phys.*, **9**, 1687–1709, https://doi.org/10.5194/acp-9-1687-2009.

Traon, P.-Y. L., and Coauthors, 2015: Use of satellite observations for operational oceanography: Recent achievements and future prospects. *J. Oper. Oceanogr.*, **8** (Suppl.), s12–s27, https://doi.org/10.1080/1755876X.2015.1022050.

Voss, K., H. Gordon, S. Flora, B. Carol Johnson, M. Yarbrough, M. Feinholz, and T. Houlihan, 2017b: A method to extrapolate the diffuse upwelling radiance attenuation coefficient to the surface as applied to the Marine Optical Buoy (MOBY). *J. Atmos. Oceanic Technol.*, **34**, 1423–1432, https://doi.org/10.1175/JTECH-D-16-0235.1.

Wei, J., Z. Lee, M. Lewis, N. Pahlevan, M. Ondrusek, and R. Armstrong, 2015: Radiance transmittance measured at the ocean surface. *Opt. Express*, **23**, 11 826–11 837, https://doi.org/10.1364/OE.23.011826.

Zhao, M., F. A. Heinsch, R. R. Nemani, and S. W. Running, 2005: Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sens. Environ.*, **95**, 164–176, https://doi.org/10.1016/j.rse.2004.12.011.

Zibordi, G., 2006: Immersion factor of in-water radiance sensors: Assessment for a class of radiometers. *J. Atmos. Oceanic Technol.*, **23**, 302–313, https://doi.org/10.1175/JTECH1847.1.

——, and K. J. Voss, 1989: Geometrical and spectral distribution of sky radiance: Comparison between simulations and field measurements. *Remote Sens. Environ.*, **27**, 343–358, https://doi.org/10.1016/0034-4257(89)90094-1.

——, and B. Bulgarelli, 2007: Effects of cosine error in irradiance measurements from field ocean color radiometers. *Appl. Opt.*, **46**, 5529–5538, https://doi.org/10.1364/AO.46.005529.