Estimating pixel-level uncertainty in ocean color retrievals from MODIS

Minwei Zhang[1, 2*], Amir Ibrahim[2], Bryan A. Franz[2], Ziauddin Ahmad[1, 2], and Andrew M. Sayer[3, 2]

[1]*Science Application International Corp., McLean, VA, United States*

[2]*Ocean Ecology Laboratory, Goddard Space Flight Center, National Aeronautics and Space Administration, Greenbelt, MD, United States*

[3]*GESTAR II, University of Maryland Baltimore County, Baltimore, MD, United States*

[*]*minwei.zhang@nasa.gov*

Abstract: The spectral distribution of marine remote sensing reflectance, R_{rs}, is the fundamental measurement of ocean color science, from which a host of bio-optical and biogeochemical properties of the water column can be derived. Estimation of uncertainty in these derived properties is thus dependent on knowledge of the uncertainty in satellite-retrieved R_{rs} ($u_c(R_{rs})$) at each pixel. Uncertainty in R_{rs}, in turn, is dependent on the propagation of various uncertainty sources through the R_{rs} retrieval process, namely the atmospheric correction (AC). A derivative-based method for uncertainty propagation is established here to calculate the pixel-level uncertainty in R_{rs}, as retrieved using NASA's multiple-scattering epsilon (MSEPS) AC algorithm and verified using Monte Carlo (MC) analysis. The approach is then applied to measurements from the Moderate Resolution Imaging Spectroradiometer (MODIS) on the Aqua satellite, with uncertainty sources including instrument random noise, instrument systematic uncertainty, and forward model uncertainty. The $u_c(R_{rs})$ is verified by comparison with statistical analysis of coincident retrievals from MODIS and in situ R_{rs} measurements, and our approach performs well in most cases. Based on analysis of an example 8-day global products, we also show that relative uncertainty in R_{rs} at blue bands has a similar spatial pattern to the derived concentration of the phytoplankton pigment chlorophyll-a (chl-a), and around 7.3%, 17.0%, and 35.2% of all clear water pixels (chl-a ≤ 0.1 mg/m^3) with valid $u_c(R_{rs})$ have a relative uncertainty ≤ 5% at bands 412 nm, 443 nm, and 488 nm respectively, which is a common goal of ocean color retrievals for clear waters. While the analysis shows that $u_c(R_{rs})$ calculated from our derivative-based method is reasonable, some issues need further investigation, including improved knowledge of forward model uncertainty and systematic uncertainty in instrument calibration.

1. Introduction

Ocean color products contain a certain degree of uncertainty resulting from imperfect calibration, sensor noise, uncertainty in ancillary data, and retrieval algorithms [1-5]. Providing retrieval-level uncertainty estimates within ocean color products has been recommended by Group on Earth Observations (GEO) and International Ocean-Colour Coordinating Group (IOCCG) within the quality assurance framework for earth observation and should be a general requirement of any satellite missions [3, 6]. Traditionally, uncertainty in remote sensing reflectance (R_{rs}) retrievals is derived through statistical comparison of satellite retrievals with collocated in situ measurements. Due to the limited availability of such matchups and their sparse distribution in space and time, statistical measures over all matching pairs or large groupings are typically used to gauge the uncertainty in the retrieved R_{rs}. Such derived uncertainties have at least three issues. First, the statistical measures represent an averaged value for the measurement conditions when and where the in situ data are collected. In practice, however, every pixel within the satellite image represents a different set of observing conditions, including radiant path geometry between the sensor, surface, and Sun, aerosol type and concentration, and surface conditions (e.g., Sun glint), and these different observational conditions result in different retrieval uncertainties. The averaged uncertainty doesn't represent the value at a specific pixel. Second, in situ data have uncertainties [7, 8] that contribute to the perceived mismatch and are typically included in the statistical measures. Although it is possible to account for the in situ data uncertainties in evaluating the mismatch [9], accurate estimates of in situ data uncertainty are not always available. Last, the spatial and temporal differences between satellite retrievals and in situ data could result in additional uncertainty in the statistical measures. While in situ data are measured at one point, satellite values used in the matchup represent the measurement average over the satellite pixel footprint (e.g., ~1 km^2), and are typically derived by averaging over even larger areas (e.g., 5×5 satellite pixels centered on the location of the in situ measurement, as recommended by [10]). Furthermore, the satellite and in situ measurements are rarely collected at exactly the same time, and within the matchup time window (e.g., 3 hours recommended by [10]) the optical properties could change, especially in coastal waters [11]. Together, these factors mean it is not an apples-to-apples comparison, and especially the effect from spatial and temporal differences is hard to quantify. The European Space Agency's Ocean Color Climate Change Initiative program (OC_CCI) does provide pixel-level uncertainty in R_{rs} merged from various missions [12], as computed using the weighted average of the uncertainty within optical water classes associated with that pixel. As the uncertainty of each class is derived from the validation against in situ data, issues with the validation described above still exist. Furthermore, optical water classes do not capture spatiotemporal variations in some key drivers of variability in atmospheric correction (AC) uncertainty, such as geometry and atmospheric turbidity.

Given the issues with the validation against in situ data as a measure of uncertainty, several image-based approaches were developed to estimate uncertainty in satellite retrieved R_{rs}. For example, [13] used a bias-resistant algorithm for concentration of chlorophyll-a (chl-a) to determine R_{rs} with the highest quality, which was then used as a surrogate for "ground truth" to estimate R_{rs} uncertainty. Using coincident daily R_{rs} from two sensors or matching satellite retrieved and in situ R_{rs}, [14] established an approach based on collocation analysis to generate R_{rs} uncertainty associated with random effects. Using geostationary measurement from Geostationary Ocean Color Imager (GOCI) collected over the course of a day, and with the assumption that no detectable changes occur in the optical properties over waters with low productivity during the daytime period, uncertainty in R_{rs} is calculated as twice the standard deviation of multiple observations in one day [15]. Although some issues with validation using in situ data could be resolved by the image-based approaches, the uncertainty derived is either valid for a specific dataset or only includes the random uncertainty.

Although there have been some studies on the pixel-level uncertainty in inherent optical properties (IOPs) [16, 17], few studies focus on R_{rs}, which is critical for calculating physical, biological and biogeochemical products, including IOPs. Uncertainty in R_{rs} is either neglected [16] or assumed constant [17], which is unrealistic [13]. These studies need a relatively realistic pixel-level uncertainty in R_{rs}. A derivative approach was developed to propagate sensor noise into uncertainty in R_{rs} [18], as retrieved from Ocean and Land Colour Instrument (OLCI) onboard Sentinel-3 using an AC algorithm for clear water [19]. A Monte Carlo (MC) approach was used to propagate sensor noise into uncertainty in R_{rs} [20], as retrieved from the standard NASA AC algorithm [21], which is based on the algorithm of Gordon &Wang (1994) [22] (GW94) but includes an iterative method to improve performance in highly productive or turbid waters. With the assumption that R_{rs} can be expressed as first-order approximation of top-of-atmosphere (TOA) radiance (L_t), Gillis et al. propagated sensor noise into R_{rs} [23], as retrieved using the Tafkaa AC algorithm [24]. Only sensor noise is included in those studies, while in reality there are other significant uncertainty sources that should be considered including instrument systematic uncertainties, ancillary data uncertainties, model uncertainties and assumptions in the AC algorithms [3], some of which have been indicated to play a more significant role than sensor noise in R_{rs} uncertainty [25]. The AC algorithms used by [18, 23] are different from the algorithm used operationally by the Ocean Biology Processing Group (OBPG) at NASA for processing ocean color data. Since uncertainty propagation depends on the AC algorithm, the propagation method developed by [18, 23] cannot be directly applied to the NASA algorithm. MC approaches such as [20] provide a generalized mechanism for calculating pixel-level R_{rs} uncertainties for any algorithms, but they are computationally intensive and therefore impracticable for routine production.

OBPG has been distributing global ocean color products for more than two decades. While these products have been used widely, pixel-level uncertainties in R_{rs} have not yet been provided. OBPG is planning the next ocean color reprocessing using a Multiple-Scattering EPSilon AC algorithm (MSEPS) [25, 26], which has been shown to perform better than GW94 [27]. In this study, we will establish a derivative method to propagate instrument random noise, instrument systematic uncertainty, and forward model uncertainty through MSEPS, with the goal of generating and verifying pixel-level uncertainty in R_{rs} retrieved from MODIS and establishing a framework for computationally efficient generation of pixel-level R_{rs} uncertainties that can be applied for all ocean color missions processed and distributed by NASA.

2. Data and methodology

2.1. MODIS data

Uncalibrated (Level-1A) data from MODIS aboard the Aqua satellite were downloaded from NASA's Ocean Biology Distributed Active Archive Center (OB.DAAC), and processed into calibrated and geolocated (Level-1B) data using the SeaDAS software package (seadas.gsfc.nasa.gov) and latest instrument calibration coefficients, as also distributed by the OB.DAAC.

2.2. In situ data

Using the OB.DAAC's in situ data archive and validation search utility tool (SeaBASS, https://seabass.gsfc.nasa.gov/search#val), coincident matchups spanning the years 2002-2019 were collected between MODIS-Aqua R_{rs} retrievals and in situ data from three sources: Marine Optical Buoy (MOBY) [28], Acqua Alta Oceanographic Tower (AAOT) [29], and BOUée pour l'acquiSition d'une Série Optique à Long termE (BOUSSOLE) [30].

2.3. MSEPS atmospheric correction

The purpose of AC is to retrieve spectral water-leaving radiance, L_w, from observed radiance, L_t, at the top of the atmosphere (for a complete list of symbols describing the AC process, see Table A1 in Appendix A). The NASA standard atmospheric correction algorithm is detailed in [21]. Briefly, L_t can be expressed as:

$$L_t = (L_r + L_a + t_{vr}L_f + TL_g + t_vL_w)t_g \quad (1)$$

where L_r is the radiance resulting from multiple scattering by air molecules in the absence of aerosol, L_a is the radiance resulting from multiple scattering by aerosols including the interaction between air molecular and aerosol scattering, L_f is the radiance resulting from scattering by surface whitecaps, L_g is sun glint, T is the direct transmittance from surface to sensor, t_{vr} and t_v are the diffuse transmittance from surface to sensor with the former only including molecular scattering and the latter including molecular and aerosol scattering, and t_g is two-way gas transmittance. Due to the known composition of air molecules, L_r can be accurately calculated using vector radiative transfer simulations that account for polarization, multiple scattering, and sea state to an uncertainty within 0.1% [31]. L_f is calculated using wind speed based on an empirical model [32, 33]. A sun glint coefficient is calculated using a statistical model [34] and then applied to a model developed by [35] to calculate L_g. Due to the high spatial and temporal variation of aerosol, it is challenging to calculate L_a, which is described briefly here for completeness.

MSEPS is based on the relationship between aerosol reflectance, ρ_a, and aerosol optical thickness, τ_a:

$$\ln(\rho_a) = \sum_{i=0}^{n} c_i \left(\ln(\tau_a) \right)^i \tag{2}$$

c_i are calculated through least-square fitting of $ln(\tau_a)$ to $ln(\rho_a)$ and stored in lookup tables (LUTs), which were generated for 80 aerosol models with eight relative humidity (rh) values and ten fine-mode fractions [36]. n is 2 for the MODIS band at 869 nm and 4 for other visible (VIS) and near infrared (NIR) bands. $\rho_a(748)$ and $\rho_a(869)$ are derived with the assumption that L_w is zero or can be accurately estimated [37], and thus epsilon (ε) can be calculated from Eq.(3).

$$\varepsilon = \frac{\rho_a(748)}{\rho_a(869)} \tag{3}$$

Two rh values are selected from the LUTs that closely bracket the rh of a MODIS pixel, assuming $rh1 < rh < rh_2$. For each aerosol model *i* in the ten models that correspond to rh1 (or rh2), $\rho_a(869)$ is converted to $\tau_a(869)$ using the model coefficients *c* through Eq.(2). Through the extinction coefficients, $\tau_a(869)$ is extrapolated to $\tau_a(748)$, which is then used to calculate $\rho_a(748)$. Dividing $\rho_a(748)$ by $\rho_a(869)$, we can derive ε_i for model *i*. Two aerosol models with the corresponding ε_x, ε_y (x, y indicate aerosol model number, assuming $\varepsilon_x < \varepsilon < \varepsilon_y$) that closely bracket ε are selected. Using the coefficients *c* for models x and y, $\rho_a(869)$ can be converted to $\tau_a(869)$. Given $\tau_a(869)$, τ_a at other bands can be calculated using extinction coefficients and then converted to ρ_a. ρ_a derived for models x and y are linearly interpolated using a ratio of $\frac{\varepsilon - \varepsilon_x}{\varepsilon_y - \varepsilon_x}$. Such interpolated ρ_a can be derived for rh_1 and rh_2, denoted by ρ_{a1} and ρ_{a2}. The actual ρ_a over the MODIS pixel is linearly interpolated from ρ_{a1} and ρ_{a2} using a ratio of $\frac{rh - rh_1}{rh_2 - rh_1}$. After removing L_r, L_a, L_f, and TL_g from L_t in Eq. (1), L_w can be derived and hereafter R_{rs}:

$$R_{rs} = \left(L_t/t_g - L_r - TL_g - t_{vr}L_f - L_a \right) f_b \big/ (t_v F_0 t_s cos\theta_s) \tag{4}$$

where f_b is bidirectional reflectance correction [38], t_s is diffuse transmittance from Sun to surface, θ_s is solar zenith angle, and F_0 is extraterrestrial solar irradiance corrected for earth-Sun distance. In effect, R_{rs} is water-leaving radiance normalized to the downwelling irradiance.

2.4. Uncertainty propagation through AC

The major categories contributing to R_{rs} uncertainties as shown by Eq. (4) are:

1. uncertainty in L_t due to instrument random noise (i.e., sensor noise);
2. instrument systematic uncertainty (e.g., absolute calibration uncertainty); and
3. uncertainties in the forward model to calculate L_a, L_r, L_f, L_g, f_b, t_g, T, t_{vr}, t_s and t_v.

A derivative approach is used to propagate all those uncertainties into R_{rs}.

In general, a variable y that is a function of variables x_i can be expressed as:

$$y = f(x_1,\ x_2\ , \ldots, x_n) \tag{5}$$

The uncertainty in y (u_c(y)) can be calculated from the uncertainty in x_i ($u(x_i)$) through

$$u_c^2(\mathrm{y}) \ = \sum_{\mathrm{i}=1}^{n} \left(\frac{\partial f}{\partial x_i}\right)^2 u^2(x_i) + 2\sum_{\mathrm{i}=1}^{n-1}\sum_{\mathrm{j}=\mathrm{i}+1}^{n} \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j} u(x_i, x_j) \tag{6}$$

n is the number of variables. Following the definitions outlined in [39], u represents standard uncertainty and u_c represents combined standard uncertainty. $u(x_i, x_j)$ is the covariance of error in variables x_i and x_j. $\frac{\partial f}{\partial x_i}$ is the partial derivative of y with respect to x_i.

To calculate $u_c(R_{rs})$, partial derivatives of R_{rs} with respect to each term with known uncertainty on the right side of Eq.(4) (L_t, t_g, L_r, TL_g, t_{vr}, L_f, L_a, t_v, t_s, and f_b) are needed. For those terms, uncertainty in L_r results from the uncertainty in wind speed (*ws*) and surface pressure (*pr*). Uncertainty in t_{vr} results from uncertainty in *pr*. Uncertainty in L_f results from the uncertainty in *ws* [40]. Uncertainty in t_g results from uncertainty in gas concentration (including ozone (*oz*), water vapor (*wv*), nitrogen dioxide (no_2)). Uncertainty in f_b results from chl-a uncertainty. The uncertainty in the model of L_r, L_f, t_g, L_g, f_b, and t_{vr} are included in the forward model uncertainty calculated through the system vicarious calibration (SVC) that is described in Appendix B, which is applied to L_t. Uncertainty in L_t, L_r, t_{vr}, L_f, and t_g can be calculated relatively straightforwardly without aerosol information. Grouping these terms together, Eq. (4) at band λ_i can be rewritten as:

$$R_{rs}(\lambda_i) = \Big(L_{rfc}(\lambda_i) - T(\lambda_i)L_g(\lambda_i) - L_a(\lambda_i)\Big)f_b(\lambda_i)\Big/\big(t_v(\lambda_i)F_0(\lambda_i)t_s(\lambda_i)cos\theta_s\big) \tag{7}$$

where L_{rfc} is defined as:

$$L_{rfc} = L_t/t_g - L_r - t_{vr}L_f \tag{8}$$

Uncertainty in $T(\lambda_i)$ results from $\tau_a(\lambda_i)$ [35], denoted by $\tau'_a(\lambda_i)$. Using L_a(NIR) (NIR refers to MODIS bands at 748 and 869 nm), $L_a(\lambda_i)$, $t_v(\lambda_i)$, and $t_s(\lambda_i)$ are calculated through MSEPS, which deploys an iterative approach to account for the non-zero L_w(NIR).

$$L_a(NIR) = L_{rfc}(NIR) - T(NIR)L_g(NIR) - t'_v(NIR)L_w(NIR) \tag{9}$$

where t'_v and the corresponding uncertainty are from last iteration. For the first iteration, t'_v is equal to t_{vr}. Uncertainty in T(NIR) is from τ_a(NIR) [35], denoted by $\tau'_a(NIR)$, which is from last iteration and equals to a predefined value for first iteration. L_w(NIR) is extrapolated from L_w at red band using a spectral model that is a function of chl-a. So, uncertainty in L_w(NIR) results from chl-a uncertainty [37]. Since *rh* is used to select aerosol models and interpolate ρ_a, uncertainty in *rh* should be another uncertainty source. In summary from Eqs.(7) and (9), uncertainty in $R_{rs}(\lambda_i)$ mainly results from L_{rfc}(NIR), $L_{rfc}(\lambda_i)$, chl-a, τ'_a(NIR), $\tau'_a(\lambda_i)$, and rh. Since τ'_a(NIR) and $\tau'_a(\lambda_i)$ are perfectly correlated with $\tau'_a(869)$ through the extinction coefficients, those two uncertainty variables can be represented by $\tau'_a(869)$. A vector for uncertainty variables is defined as:

$$X_i = [L_{rfc}(\mathrm{NIR}), L_{rfc}(\lambda_i), \mathrm{chl-a}, \tau'_a(869),\ \mathrm{rh}\ ] \tag{10}$$

To calculate $u_c(R_{rs}(\lambda_i))$ using Eq.(6), the partial derivative of $R_{rs}(\lambda_i)$ with respect to X_i ($\frac{\partial R_{rs}(\lambda_i)}{\partial X_i}$) as well as $u_c(X_i)$ are needed. The calculation of $\frac{\partial R_{rs}(\lambda_i)}{\partial X_i}$ is detailed in Appendix A and $u_c(X_i)$ are calculated in Section 2.5.

While Fig. 1 shows a general flow chart, the calculation of $u_c(R_{rs})$ is detailed step by step as follows:

(1) With the input of MODIS L1B, GEO, and ancillary data, L_a(NIR) can be calculated after removing L_r, L_f, L_g, and L_w from L_t. Note L_w(NIR) is assumed zero for the first iteration that is used to account for non-zero L_w(NIR);

(2) Based on L_a(NIR), MSEPS is applied to calculate L_a, t_s, t_v, τ_a, $\frac{\partial L_a}{\partial X}$, $\frac{\partial t_s}{\partial X}$, $\frac{\partial t_v}{\partial X}$, and $\frac{\partial \tau_a}{\partial X}$;

(3) Using $\frac{\partial L_a}{\partial X}$, $\frac{\partial t_s}{\partial X}$, and $\frac{\partial t_v}{\partial X}$, $\frac{\partial R_{rs}}{\partial X}$ can be derived. Using L_a, t_s, and t_v, R_{rs} can be derived.

(4) $\frac{\partial R_{rs}}{\partial X}$ and u_c(X) are used to calculate $u_c(R_{rs})$. Note for the first iteration, u_c(chl-a) and $u_c(\tau'_a(869))$ are assumed 0;

(5) R_{rs} at red band is used as one criterion to determine if the iteration converges. Readers are referred to [37] for detailed convergence criteria;

(6) If the iteration converges, R_{rs} and $u_c(R_{rs})$ are output and the iteration stops;

(7) If the iteration doesn't converge, $u_c(R_{rs})$ is used to calculate u_c(chl-a) [17]. $\frac{\partial \tau_a}{\partial X}$ and u_c(X) are used to calculate $u_c(\tau'_a(869))$. u_c(chl-a) and $u_c(\tau'_a(869))$ are used for the next iteration. chl-a can be calculated from R_{rs} and then applied to calculate L_w(NIR). From L_w(NIR), L_a(NIR) is derived and another iteration starts.

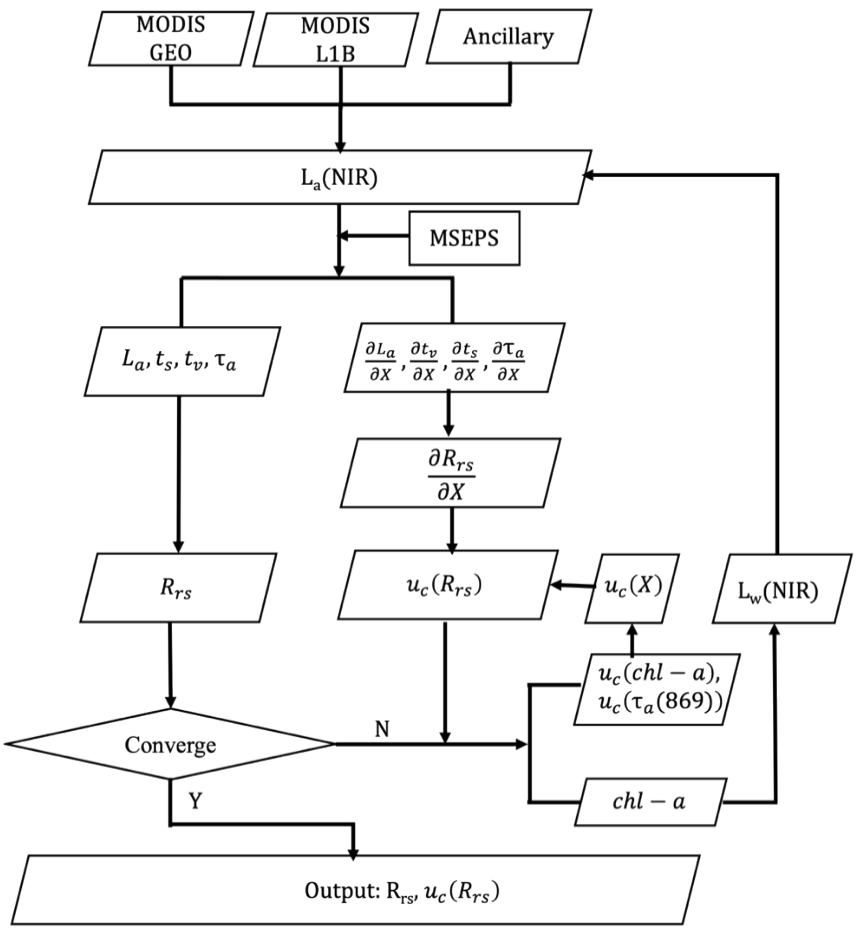

Fig.1. The flow chart for the calculation of $u_c(R_{rs})$. For detailed convergence criteria of the iteration to account for non-zero L_w(NIR), readers are referred to [37].

2.5. Estimation of uncertainty sources for MODIS

Uncertainty in L_t comes from sensor noise and systematic error in instrument calibration. For MODIS, the sensor noise (χ) is modeled as:

$$\chi(\lambda) = [A_0(\lambda) + A_1(\lambda)L_t(\lambda)]S(\lambda) \tag{11}$$

where A_0 and A_1 are derived from fitting the lab measured L_t and χ. The values are listed in Table 1. It should be noted that L_t is in the unit of $W.m^{-2}.\mu m^{-1}.sr^{-1}$. As MODIS' land bands with a spatial resolution of 500 m and 250 m are aggregated to 1000-m resolution to match the ocean bands, S is applied to correct for the spatial resolution difference [41], which is equal to 2 and 4 for bands with spatial resolution of 500 m and 250 m respectively.

The instrument systematic uncertainty is more difficult to quantify, but the largest source is the absolute instrument calibration that relates measured counts to radiance in geophysical units. For MODIS ocean color processing, NASA updates the prelaunch counts to radiance conversion using SVC approach [42]. In the SVC process, in situ measurements of L_w from MOBY are matched-up in time and space with satellite observations, and the coincident MOBY measurements are propagated to the TOA using forward model of the AC algorithm to derive an expected TOA radiance at each visible band. The ratio of expected TOA radiance to MODIS-observed L_t is a measure of the absolute calibration gain, and many such samples are collected and averaged over the mission lifetime to derive the mean SVC gain that effectively replaces the pre-launch calibration. We thus adopt here the uncertainty in SVC gain as a first-order estimate of systematic uncertainty on L_t. It should be noted from Table 2 the low systematic uncertainty at bands 412-748 nm with respect to 869-nm band, which results from the low percentage of aerosol radiance in the TOA radiance. As the systematic uncertainty is mainly due to the uncertainty in aerosol radiance that is extrapolated from 869-nm band, the low percentage of aerosol radiance will result in a low systematic uncertainty.

While sensor noise is spectrally independent, uncertainty from SVC, as with systematic instrument calibration errors in general, will exhibit some level of spectral covariance that should be considered in the uncertainty propagation. Specifically, in the SVC process, MODIS 748-nm band is calibrated relative to the 869-nm band, and the VIS bands are calibrated relative to those two NIR bands. Due to the spectral dependence of the $u(L_t)$, the covariance of error in L_t at VIS and at NIR bands should be included, i.e., $u(L_t(\lambda_i), L_t(748))$, $u(L_t(\lambda_i), L_t(869))$, and $u(L_t(748), L_t(869))$ should be included in calculating $u_c(R_{rs}(\lambda_i))$. Using the correlation coefficient (r) between ρ_t at bands λ_i and 748 nm derived in Appendix C, $u(L_t(\lambda_i), L_t(748))$ can be calculated from:

$$u(L_t(\lambda_i), L_t(748)) = r(\rho_t(\lambda_i), \rho_t(748))\, u(L_t(\lambda_i))\, u(L_t(748)) \tag{12}$$

$u(L_t(\lambda_i), L_t(869))$ and $u(L_t(748), L_t(869))$ are calculated using the same approach as Eq. (12). $u(L_t)$ only includes the uncertainty from SVC. Xiong et al. indicate that the calibration uncertainty of MODIS reflective solar bands could meet their specified requirement of 2% [43], which is taken as the instrument systematic uncertainty on MODIS 869-nm band in this study.

Forward model uncertainty is also difficult to estimate. It derives from algorithm assumptions and modeling errors in determining L_a, L_r, L_f, f_b, and L_g, as well as ancillary data uncertainties and other unknown sources. Here we again take advantage of the SVC process and assume that the variance in the individual SVC gain samples provides an estimate of the total uncertainty on L_t, combined with the uncertainty in the in situ measurements after propagation to TOA. Thus, forward model uncertainties are derived by removing other terms from the standard deviation of the SVC gains, including (1) standard error of the mean SVC gain, (2) sensor noise, (3) uncertainty in ancillary data, and (4) uncertainty in MOBY L_w. For the uncertainty in ancillary data, we adopt the local temporal variability (i.e., the difference between the two temporal samples that bound the time of satellite observation) as a first-order estimate. The estimation of the forward model uncertainty is further detailed in Appendix B. Table 2 lists the instrument systematic uncertainty and the forward model uncertainty, expressed as a percentage of L_t. The high forward model uncertainty at band 748 nm is due to the fixed aerosol model used in the SVC [42], as the true aerosol type in the SVC region of the SPG (see Appendix B) may vary with time.

Using the estimation of uncertainty sources described above, $u_c(X)$ can be calculated:

(1) $u_c(L_{rfc})$. The partial derivative of L_{rfc} with respect to L_t, t_g, L_r, t_{vr}, and L_f are calculated based on Eq.(8). $u_c(L_t)$ is derived from multiplying L_t by the systematic uncertainty plus model uncertainty in Table 2 as well as adding sensor noise from Eq.(11). $u_c(t_g)$ are calculated based on Eq.(6) using the partial derivative of t_g with respect to gas concentration and the uncertainty in gas concentration. Similar approach is used to calculate $u_c(L_r)$, $u_c(t_{vr})$, and $u_c(L_f)$. Note that $u_c(L_r)$ results from uncertainty in *ws* and *pr*. $u_c(t_{vr})$ results from uncertainty in *pr*. $u_c(L_f)$ results from uncertainty in *ws*.

(2) $u_c(\tau'_a(869))$ and u_c(chl-a). For the first iteration that is used to account for the non-zero L_w(NIR), $\tau'_a(869)$ and chl-a are assumed constant with uncertainty of zero. $u_c(R_{rs})$ and $u_c(\tau_a)$ derived from the i^{th} iteration are then applied to calculate u_c(chl-a) and $u_c(\tau_a(869))$ used for the $(i+1)^{th}$ iteration.

(3) $u_c(rh)$. Calculated as the local temporal variability as described above.

Table 1. Coefficients in Eq.(11) for calculating sensor noise of MODIS.

	412	443	469	488	531	547	555	645	667	678	748	869
A_0 (10^{-3})	55.0	29.4	119.3	19.3	14.0	11.4	87.7	104.1	5.0	4.3	4.2	3.1
A_1(10^{-5})	8.3	9.4	8.2	9.5	10.0	16.5	7.0	8.5	14.1	13.2	21.3	18.6

Table 2. Instrument systematic uncertainty (Sys) and forward model uncertainty (Mod).

	412	443	469	488	531	547	555	645	667	678	748	869
Sys (%)	0.14	0.13	0.13	0.13	0.10	0.095	0.095	0.089	0.065	0.068	0.085	2.0
Mod (%)	1.0	0.94	0.91	0.86	0.68	0.62	0.60	0.49	0.37	0.38	1.27	0.0

2.6. Verification of uncertainty propagation using Monte Carlo analysis

Monte Carlo analysis is used to verify $u_c(R_{rs})$ derived from the derivative method when only instrument random noise is included. A Gaussian random noise is generated as:

$$L_{noise} = N(0, \frac{\chi}{L_t})L_t \tag{13}$$

where χ is sensor noise calculated from Eq. (11). A random noise L_{noise} is added to L_t, providing L'_t, which is defined as:

$$L'_t = L_t + L_{noise} \tag{14}$$

MSEPS is applied to L'_t with the resulting R_{rs} denoted by R'_{rs}. If a total of N samples of R'_{rs} are generated, the root mean square error (RMSE) can be calculated as

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(R'_{rsi} - R_{rsi}\right)^2}{N}} \tag{15}$$

where R_{rs} is derived from applying MSEPS to L_t without sensor noise. This RMSE represents the uncertainty in R_{rs} resulting from sensor noise and is used to verify the corresponding

u_c(R_{rs}) derived from the derivative method.

2.7. Evaluation of $u_c(R_{rs})$ using validation results

Following the approach presented by [44, 45], u_c(R_{rs}) is evaluated using the matchups between MODIS retrieved and in situ R_{rs} at MOBY, AAOT, and BOUSSOLE. By adding in quadrature u_c(R_{rs}) from the derivative method, which represents uncertainty in MODIS retrieved R_{rs}, uncertainty in in situ R_{rs}, and the spatial and temporal difference between these two measurements, we calculate an expected discrepancy (Δ_D) between MODIS-retrieved and in situ R_{rs}. The uncertainty-normalized difference, Δ_N, is defined as the ratio of actual retrieval difference to Δ_D, i.e.,

$$\Delta_N = \frac{R_{rs}^m - R_{rs}^f}{\Delta_D} \tag{16}$$

where R_{rs}^m and R_{rs}^f represent MODIS-retrieved and in situ R_{rs} respectively. If the uncertainties in MODIS retrieved R_{rs} and in situ R_{rs} and the spatiotemporal mismatch effects are calculated appropriately, and the sample size is sufficient, the ensemble of Δ_N should be close to a Gaussian distribution with mean 0 if there is no bias and variance 1. So we can first qualitatively evaluate u_c(R_{rs}) by checking the probability density function (PDF) of Δ_N against that of a Gaussian distribution. Taking this a step further, a total of N matchups is divided into *n* equally populated bins based on Δ_D indexed from low to high. For each bin, the 68th percentile of absolute difference between retrieved and in situ R_{rs}, which is close to the standard deviation (1σ) for a Gaussian distribution, is plotted against the average Δ_D. If Δ_D is reasonable, the points should lie along the 1:1 line. Dividing all the matchups into n bins allows examination of the skill of u_c(R_{rs}) to distinguish between low- and high-uncertainty conditions, as opposed to just population-average behavior.

3. Results

3.1. Evaluation of the derivative method against MC analysis for instrument random noise

Fig. 2 shows one example of u_c(R_{rs}(443)) calculated from the derivative method for MODIS data over the South Pacific ocean with very clear waters. Only instrument random noise is included to compare with uncertainty derived from MC. We can see from Fig. 2 that u_c(R_{rs}(443)) from these two methods shows a similar spatial pattern, higher at the edge than at the center of the swath. The higher u_c(R_{rs}) at edges is due to the longer path length and higher relative contribution of path radiance to L_t, increasing the uncertainty in L_a, which is then propagated to R_{rs}. Higher L_t also means larger random noise as A_1 in Eq. (11) is positive. Fig. 3 shows the quantitative comparison of spectral u_c(R_{rs}) derived from these two methods over 4 pixels, ranging from high to low values. The spectral u_c(R_{rs}) agree very well. The higher jump in uncertainty at 469 nm and 555 nm is due to the lower signal-to-noise ratio (SNR) at ocean signal levels for those bands, which were designed with a much higher dynamic range to support land applications. Fig. 4 shows the mean ratio between u_c(R_{rs}) from the derivative method against that from MC over all valid pixels in Fig. 2. This is generally between 0.9 and 1.1 across the VIS bands, with the derivative method tending to underestimate u_c(R_{rs}) at blue bands compared with MC. Note that 2000 random samples were used for the MC calculations, as this seems sufficient to get stable results for this type of scene (Fig. 5). Figs.2-4 show that u_c(R_{rs}) derived from the derivative method compares reasonably well with that from MC method, indicating the reliability of the derivative method when only instrument random noise is included in the uncertainty budget.

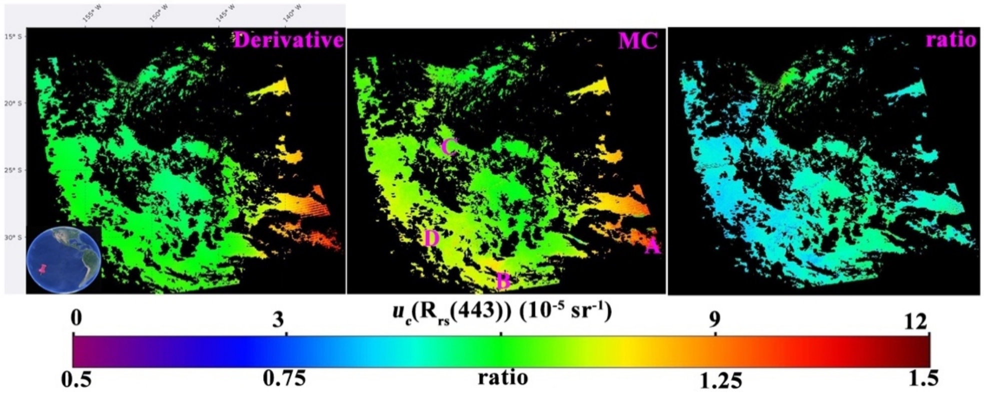

Fig. 2. $u_c(R_{rs}(443))$ derived by applying the derivative method and MC to MODIS data over South Pacific ocean on Apr. 19, 2017. Only instrument random noise is included. The ratio is calculated from dividing $u_c(R_{rs}(443))$ from derivative method by that from MC.

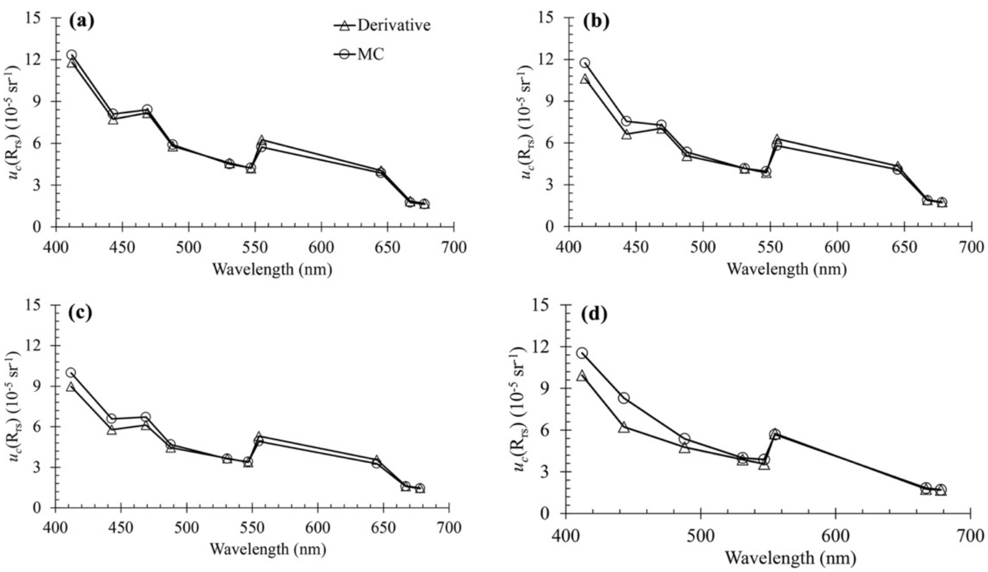

Fig. 3. Spectral $u_c(R_{rs})$ from the derivative method compared with that from MC over pixels denoted by (a) 'A', (b) 'B', (c) 'C', and (d) 'D' in Fig. 2. Only instrument random noise is included in the calculation.

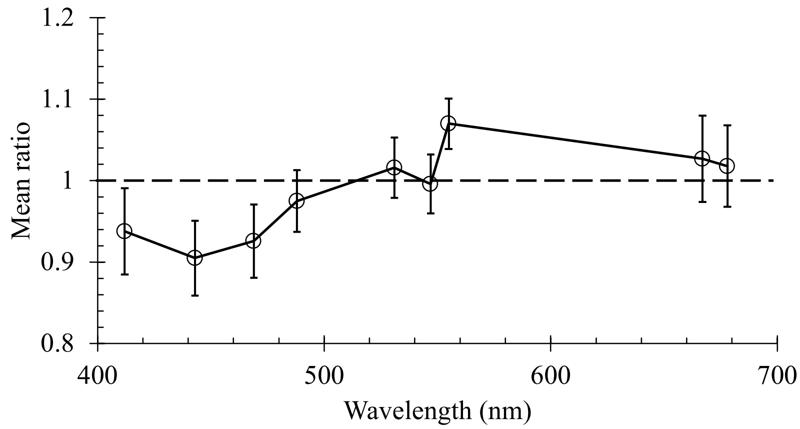

Fig. 4. Mean ratio of derivative to MC $u_c(R_{rs})$ over all the valid pixels in Fig. 2. Error bars indicate the standard deviations.

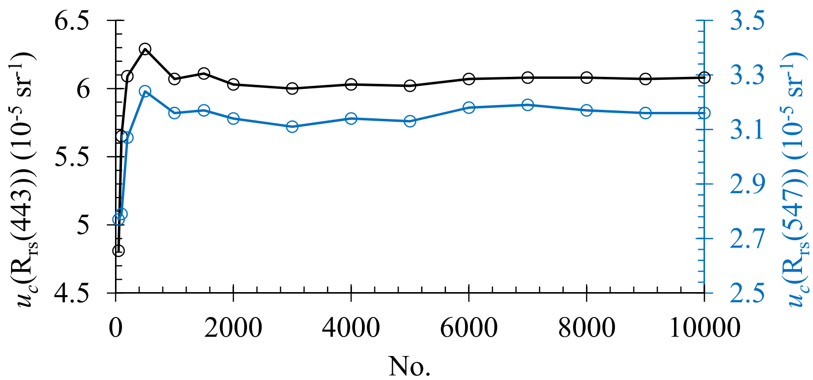

Fig. 5. Variation of $u_c(R_{rs}(443))$ and $u_c(R_{rs}(547))$ from instrument random noise, as a function of the number of random samples used in the MC calculation. The calculation is based on pixel 'D' in Fig. 2.

3.2. Evaluation of $u_c(R_{rs})$ including all modeled uncertainty sources

3.2.1. Evaluation of spatial patterns

We first assess $u_c(R_{rs})$ estimates, including all modeled uncertainty sources, through the expected spatial patterns. Fig. 6 shows one example of $u_c(R_{rs})$ in absolute terms and expressed as relative uncertainty, δ ($u_c(R_{rs})$×100/ R_{rs}). $u_c(R_{rs})$ is high at the edge of sun glint and over regions near thin clouds. The higher $u_c(R_{rs}(443))$ than $u_c(R_{rs}(547))$ results from the large sensor noise as well as the large systematic and forward model uncertainty (see Table 2). Note from Fig. 6b that some pixels in the circled area lack valid values due to the absence of an upper bounding aerosol model. As described in Section 2.3, a lower bounding aerosol model x (with the corresponding epsilon ε_x) and upper bounding aerosol model y (with the corresponding epsilon ε_y) are selected to interpolate ρ_a using a ratio of $\frac{\varepsilon - \varepsilon_x}{\varepsilon_y - \varepsilon_x}$, where ε is calculated from $L_{rfc}(748)$ and $L_{rfc}(869)$, and ε_x and ε_y are calculated from $L_{rfc}(869)$ using the coefficients in the LUTs. Model x (or y) is considered as the aerosol for this pixel when there is no upper bounding aerosol (or no lower bounding aerosol). Then, the ratio is assumed to be 1 and L_a in other wavelengths is calculated from $L_{rfc}(869)$ without using $L_{rfc}(748)$, which means that the uncertainty in $L_{rfc}(748)$ cannot be propagated to L_a, resulting in the underestimation of $u_c(R_{rs})$. Pixels without lower or upper bounding aerosol models are currently masked out until an improved implementation can be realized.

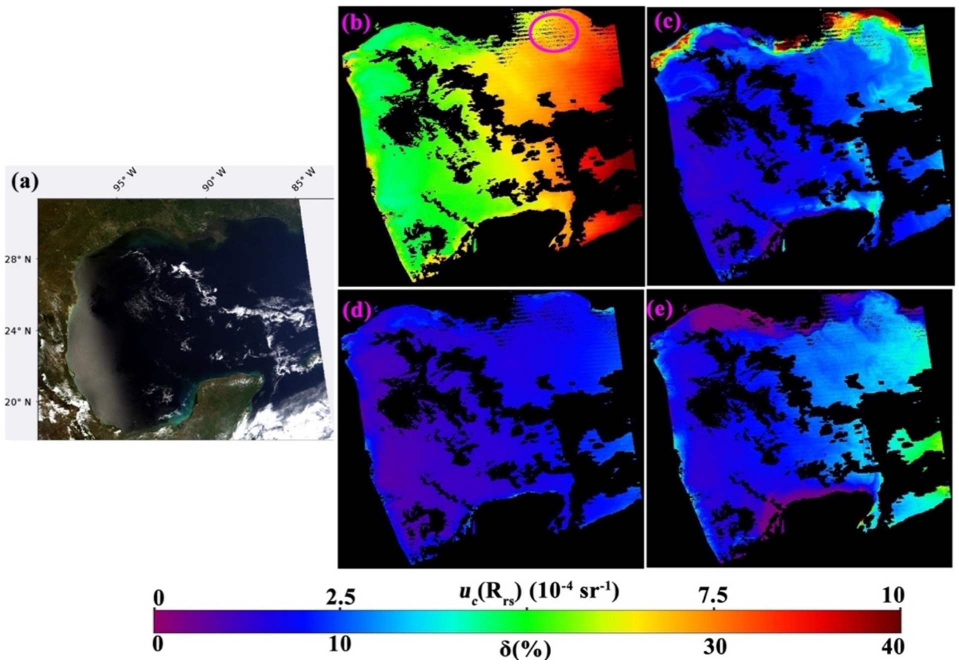

Fig. 6. Spatial analysis of a MODIS scene over the Gulf of Mexico on May 6, 2017 showing (a) true-color image, (b) $u_c(R_{rs}(443))$, (c) $\delta(R_{rs}(443))$, (d) $u_c(R_{rs}(547))$, and (e) $\delta(R_{rs}(547))$. Uncertainty sources include instrument random noise, instrument systematic uncertainty, and forward model uncertainty.

3.2.2. Closure analysis with results from validation against in situ data

$u_c(R_{rs})$ calculated using all uncertainty sources is further evaluated using the validation results derived from matchup comparison between MODIS retrieved R_{rs} and in situ measurements at MOBY, AAOT, and BOUSSOLE. Following the approach presented by [10], a spatial window of 5×5 pixels centered on the location of in situ data and time window of ± 3h were used to search matching pairs. A valid matching pair also requires spatial homogeneity with a coefficient of variation (i.e., ratio of standard deviation to mean over a 5×5 pixels region) smaller than 15%. For each matching pair, we have MODIS retrieved R_{rs}, in situ R_{rs}, and $u_c(R_{rs})$. As described in Section 2.7, our approach requires knowledge of uncertainty in in situ R_{rs} and the spatial and temporal difference between MODIS retrieved and in situ R_{rs} to evaluate $u_c(R_{rs})$.

MOBY includes three L_u sensors deployed at depths of 1 m (top), 5 m (middle), and 9 m (bottom), with the uncertainty in L_u measured by the top sensor increasing from 2.1% in blue wavelengths to 3.3% in red wavelengths, for good scans on good days [46]. Combining this L_u uncertainty with the environmental uncertainty (personal communication with Kenneth J Voss) as well as the uncertainty in downward irradiance just above the sea surface (E_d), we used a constant 5% uncertainty in MOBY R_{rs} at VIS bands in this study. For BOUSSOLE, Białek et al. show R_{rs} uncertainty of less than 4% in blue and green wavelengths and less than 5% in red wavelengths [47], and we adopt those values here. The SeaPRISM system used to collect data at AAOT has an uncertainty of 5.3%, 4.8%, 4.6%, 4.9%, and 7.3% in wavelengths of 412 nm, 443 nm, 488 nm, 551 nm, and 667 nm respectively[48], and we adopt those values here for AAOT data. Note 4.9% is used for bands at 531 nm, 547 nm and 555 nm. 7.3% is used for bands at 667 nm and 678 nm. A temporal variation of 2%, 3%, 4% is indicated for normalized water-leaving radiance at 551nm, $L_{wn}(551)$, at time difference of 0.5, 1.0, 1.5 h at AAOT [49]. As a result, we add a 3% per hour spectrally independent uncertainty to account for the time difference between satellite and in situ data at this site. Temporal variation is neglected at MOBY and BOUSSOLE due to the stability of the optical properties [50]. The standard deviation over the box of 5×5 pixels centered on the location of in situ data is used to represent the spatial variation between retrieved and in situ R_{rs}. Then, the Δ_D for a given matchup is calculated by adding in quadrature $u_c(R_{rs})$, uncertainty in in situ R_{rs}, and the spatial and temporal variability estimates. Because uncertainty is a measure of the statistical dispersion of retrievals relative to truth, the evaluation needs to be done on a statistical rather than pairwise basis. Fig. 7 shows two methods for this with the number of matchups listed in Table 3. The left column shows the PDF of normalized difference (Eq. 16) and the theoretical Gaussian distribution. These two distributions should ideally match if the uncertainty estimates are reliable [44]. Results are reasonable at band 443 nm at BOUSSOLE and at bands 412 nm-531 nm at MOBY. R_{rs} at bands 547 nm and 555 nm at MOBY tend to be biased (PDFs not centered around zero) and Δ_D estimates are overconfident (PDFs wider than expected). Δ_D estimates are underconfident in red wavelengths at MOBY and BOUSSOLE (PDFs narrower than expected). R_{rs} in all wavelengths tends to be biased at AAOT, probably due to the different conditions (including water and aerosol) from that at the SVC site (i.e., MOBY). The different conditions complicate the calculation of L_a in the atmospheric correction, either because L_w(NIR) is not well represented or because the aerosol models are not able to properly model the actual aerosol condition. The right column shows binned Δ_D vs. 1σ of the absolute difference between retrieved and in situ R_{rs} within each bin. At least 100 matchups are needed for each bin, for better statistical robustness. Please note the exception of 412 nm and 667 nm at BOUSSOLE, with the former having one bin with 88 matchups and the latter having two bins with 81 matchups for each. There is only one bin with 65 matchups for 678-nm band at AAOT. Overall, Δ_D agrees reasonably well with 1σ points of absolute difference, especially at MOBY. This shows that the derivative method has skill in distinguishing relatively low-uncertainty cases from high-uncertainty cases and capturing the spectral dependence of uncertainty. The underestimate of Δ_D at AAOT could be partly due to the approximation of temporal and spatial variation, which are challenging to quantify considering the complicated water environment in a transitional zone from coastal to open ocean. The underestimate of Δ_D could also result from the bias shown in Fig. 7c. This hypothesis is supported by Fig. 8, which is the same as Fig. 7d, except for subtracting the mean R_{rs} bias (i.e., bias-correction) before calculating the absolute difference. In this case the points are much closer to the 1:1 line, suggesting some systematic error in the retrieval at this site but a reasonable estimate of dispersion.

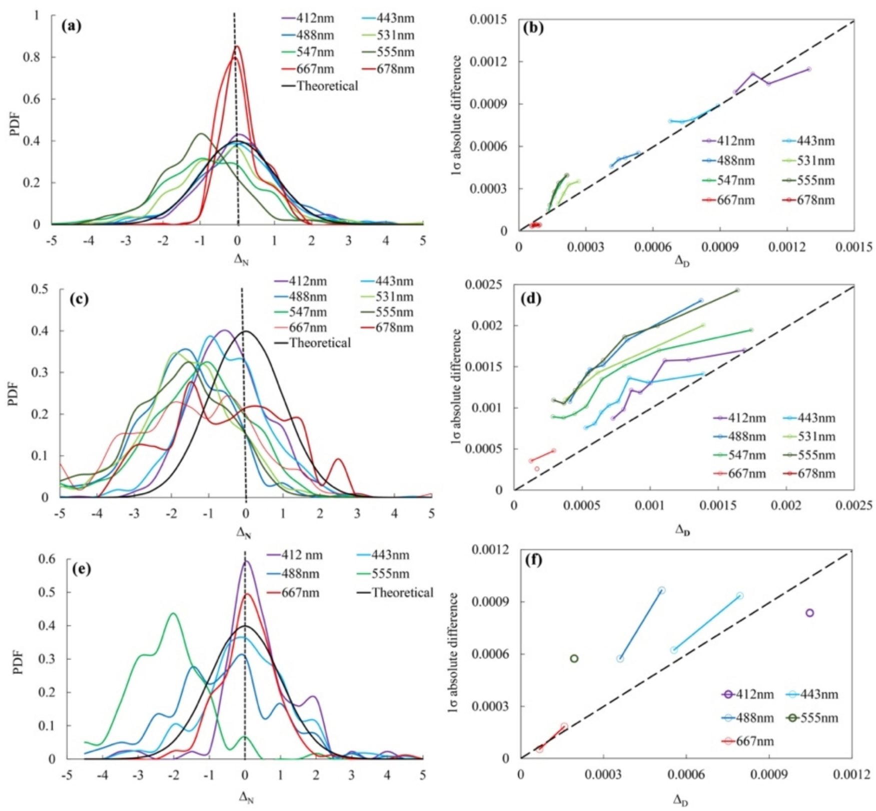

Fig. 7. Evaluation of $u_c(R_{rs})$ using matchup comparison between MODIS retrieved and in situ R_{rs} at MOBY (1st row), AAOT(2nd row), and BOUSSOLE (3rd row). The left column shows the PDF of uncertainty-normalized difference, with the black line representing theoretical Gaussian distribution with mean 0 and variance 1. The right column shows the Δ_D versus 1σ absolute difference between retrieved and in situ R_{rs}; the 1:1 line is dashed.

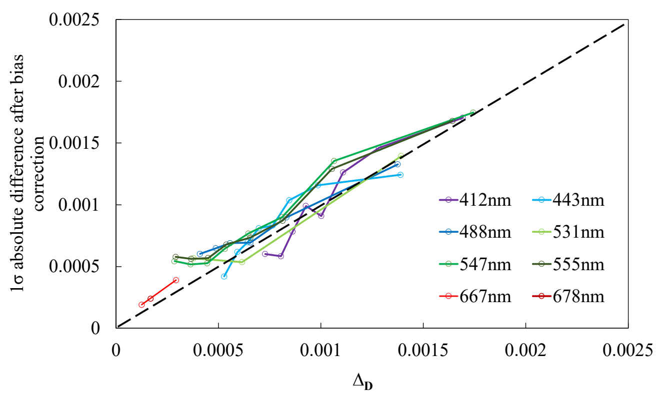

Fig. 8. As Fig. 7d but subtracting mean R_{rs} bias at AAOT before calculating the absolute difference.

Table 3. Number of matchups between MODIS retrieved and in situ R_{rs} used in Fig. 7.

	412	443	488	531	547	555	667	678
MOBY	470	470	470	470	470	470	466	470
AAOT	863	887	698	369	899	814	222	65
BOUSSOLE	88	231	240			119	163	

3.2.3. Comparison with uncertainty estimates from other studies

Hu et al. calculated uncertainty in R_{rs} as the standard deviation of the difference between MODIS retrieved and reference R_{rs} [13]. The reference R_{rs} for a given chl-a level (with ±2% range to find enough pixels for statistical analysis) is derived by averaging all the R_{rs} that produce chl-a from two algorithms matching within 5%, where one of the algorithms has been shown to be highly resistant to spectrally correlated bias in R_{rs}. Fig. 9 shows the comparison between $u_c(R_{rs})$ derived from the derivative method and the uncertainty presented by Hu et al. MODIS data over the North Atlantic and South Pacific subtropical gyres during Dec. 3-10, 2019 are used to calculate $u_c(R_{rs})$. We can see from Fig. 9b that uncertainty from these two approaches both show a higher value for chl-a level of 0.05 mg/m^3 than that for chl-a level of 0.03 mg/m^3. While uncertainty from these two approaches show a similar spectral pattern that is decreasing with wavelength, uncertainty derived from the derivative method is higher than that from Hu et al. at bands 412 nm and 443 nm. These two compare reasonably well at 488 nm, 531 nm, and 547 nm. The lower uncertainty in Hu et al. is likely primarily due to the uncertainty in the reference R_{rs} and spatial/temporal variations between specific R_{rs} and reference R_{rs}, which are not accounted for. The uncertainty in Hu et al. only captures the model uncertainty and noise, not the instrument systematic uncertainties. Differences between the AC algorithms (MSEPS for this study vs. GW94 for Hu et al.) and the assumptions in the derivative method could also contribute to the difference in the uncertainty from these two approaches.

Another approach has been presented by [14], which uses coincident R_{rs} data between different satellite missions and between satellite missions and in situ measurements. Fig. 10 shows the comparison of the uncertainty derived from averaging $u_c(R_{rs})$ over all the matchups described in Section 3.2.2 with that presented by Mélin et al. at MOBY and AAOT. While Fig. 10 shows that these two compare reasonably well at MOBY, uncertainty derived from the derivative method tends to be lower than that from Mélin et al. especially at AAOT. The higher value from Mélin et al. may be partly due to the contribution from the spatiotemporal variation between MODIS retrieved and in situ R_{rs}. The difference may also result from different AC algorithms used for generating R_{rs} (again MSEPS vs. GW94) and the assumptions in the uncertainty estimate techniques between those studies. Please note from Fig. 10b the lower $u_c(R_{rs})$ from the derivative method than uncertainty in in situ R_{rs} [48], which means that $u_c(R_{rs})$ is probably underestimated. The underestimation may result from the forward model uncertainty that is estimated at MOBY which is representative of open ocean. However, the forward model uncertainty is likely larger in coastal waters as AAOT than that in open ocean due to the complexity in atmosphere (e.g., presence of absorbing aerosol) and water optical properties (e.g., bidirectional reflectance correction).

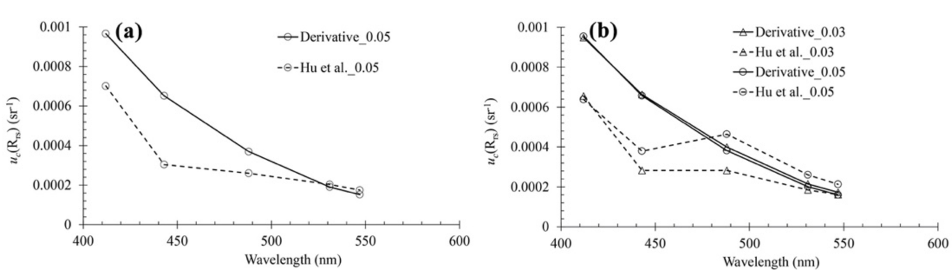

Fig. 9. Comparison between $u_c(R_{rs})$ and uncertainty estimates from [13] over (a) North Atlantic subtropical gyre, and (b) South Pacific subtropical gyre. The numbers in the legend refer to chl-a. Derivative_0.05 is calculated by averaging $u_c(R_{rs})$ over all the pixels in the region with chl-a in the range of 0.05×(1±2%). The other values are from Table 3 in Hu et al. MODIS data from Dec. 3-10, 2019 are used to calculate mean $u_c(R_{rs})$ for derivative method.

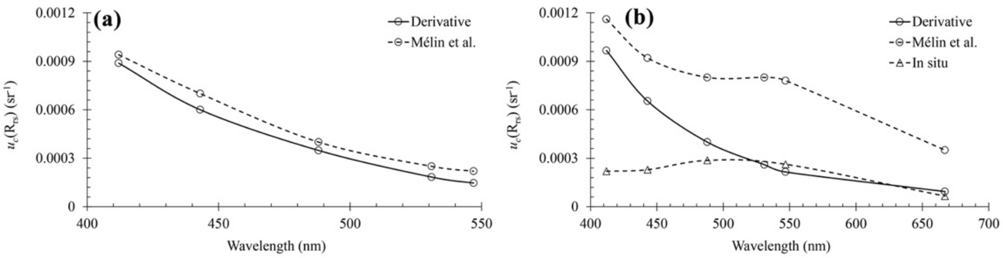

Fig. 10. Comparison between $u_c(R_{rs})$ and the uncertainty estimates from Mélin et al. at (a) MOBY, and (b) AAOT. $u_c(R_{rs})$ is derived by averaging over all the matchups described in Section 3.2.2. Uncertainty values for Mélin et al. are estimated from Fig. 9 in [14]. Uncertainty in in situ R_{rs} at AAOT is from [48].

3.3. Global $u_c(R_{rs})$ maps

Fig. 11 shows 8-day global $u_c(R_{rs})$ and δ at 412, 443, 488, 531, and 547 nm, as well as chl-a calculated with the OCI algorithm [51]. δ at 412, 443, and 488 nm show a similar spatial pattern to chl-a. This spatial pattern results from chl-a absorption. The low R_{rs} due to chl-a absorption results in a high δ over waters with high chl-a and vice versa for waters with low chl-a. The increased atmospheric turbidity could also increase the $u_c(R_{rs})$ in coastal regions. This spatial pattern is not obvious at bands 531 nm and 547 nm as these bands are only weakly dependent on chl-a. $u_c(R_{rs})$ doesn't show as much spatial variability as δ, which is also presented by [14]. It should be noted that the 8-day $u_c(R_{rs})$ is simply the average uncertainty in each bin over that period and does not represent the uncertainty in an 8-day (Level-3) R_{rs} mean. Fig. 12 shows the cumulative distribution function (CDF) of δ over clear water pixels (chl-a ≤ 0.1 mg/m^3) with valid data in Fig. 11. Overall, around 7.3%, 17.0%, and 35.2% of all the valid pixels with clear waters have δ ≤ 5% at bands 412 nm, 443 nm, and 488 nm respectively, which is a common goal of ocean color retrievals for clear waters [52]. Those percentage numbers are different from the conclusion reached by [13, 14] that the goal of 5% is fulfilled at blue bands over clear waters. The difference is primarily due to the methods used to generate $u_c(R_{rs})$. Those methods have different assumptions. The difference could also be due to the different satellite data used to generate the R_{rs} uncertainty. While MODIS global data during Dec. 3-10, 2019 are used in this study, Hu et al. (2013) use data over the North Atlantic and South Pacific subtropical gyres in 2006 and Melin et al. (2016) use global data during 2003-2007.

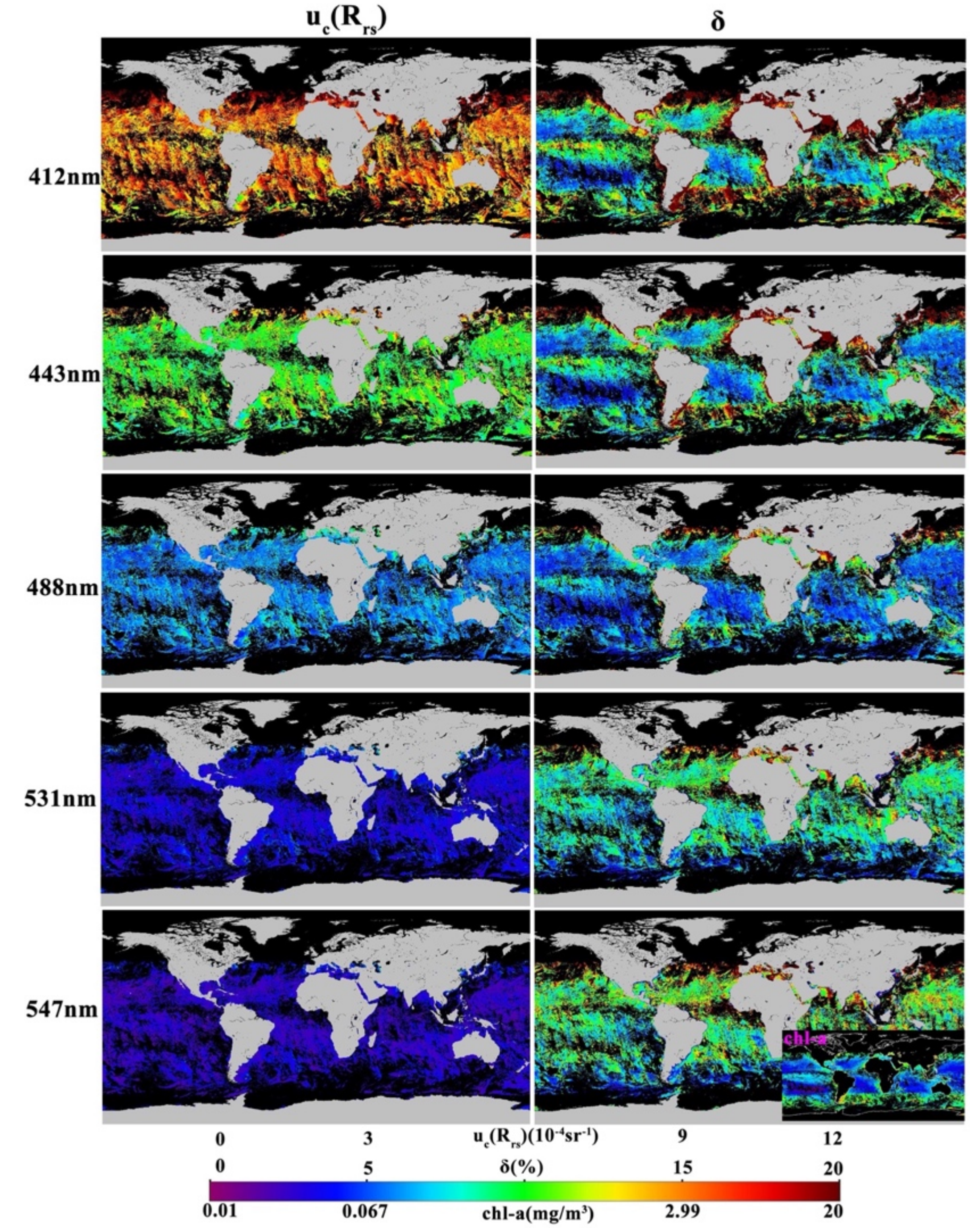

Fig. 11. 8-day R_{rs} uncertainty in both absolute (left column) and relative term (right column), calculated from the derivative method, and chl-a calculated using R_{rs} retrieved using MSEPS from MODIS data during Dec. 3-10, 2019. Gray means land and black means no valid data.

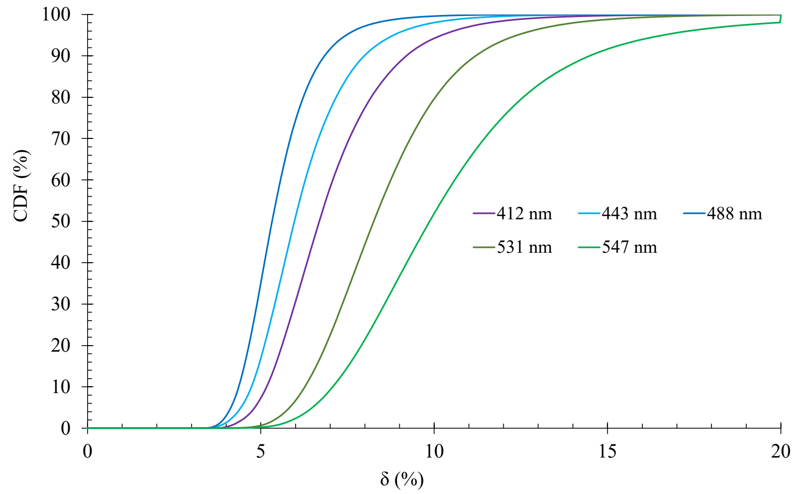

Fig. 12. Cumulative distribution function (CDF) of δ at bands 412 nm, 443 nm, 488 nm, 531 nm, and 547 nm for all the pixels with valid data and with chl-a ≤ 0.1 mg/m^3 in Fig. 11.

4. Discussions and Conclusions

We present and perform an initial evaluation of a derivative-based method to calculate the uncertainty in R_{rs} retrieved from MSEPS atmospheric correction algorithm. Distinct from the (diagnostic) uncertainty products derived from statistics of validation against in situ data, which represent an overall summary uncertainty for an entire dataset, this (prognostic) method estimates a pixel-level uncertainty. It accounts for uncertainty sources including instrument random noise, instrument systematic uncertainty, and forward model uncertainty.

We first assessed the derivative method by comparing estimates considering only instrument random noise with Monte Carlo analysis, which showed reasonable (within 10% on average) spatial and spectral agreement. We then performed a deeper closure analysis, comparing MODIS $u_c(R_{rs})$ against statistical analysis of matchups between MODIS R_{rs} retrievals and coincident in situ measurements at MOBY, AAOT, and BOUSSOLE, while also accounting for uncertainties in in situ measurements and effects of spatial and temporal sampling differences. The closure analysis demonstrates the capability of the derivative method at characterizing the relative magnitude and spectral dependence of R_{rs} uncertainty. However, the uncertainty is systematically overestimated or underestimated at some wavelengths and sites, showing the need for a better understanding of the uncertainty model and contributions from in situ data and spatial/temporal variation.

$u_c(R_{rs})$ presented above includes multiple uncertainty sources, which may raise questions about the contribution from each source. MODIS scene from Fig. 6 is used to examine the contribution from instrument random noise, instrument systematic uncertainty and forward model uncertainty to $u_c(R_{rs})$. Results indicate that instrument random noise is generally a much smaller contribution than either instrument systematic or forward model uncertainty sources. It is not trivial to further disentangle the instrument systematic and forward model uncertainty due to considerable spectral covariance between the terms.

While sensor noise is reasonably well understood, the other sources involve simplifications and assumptions including uncertainty in L_t, calibration for band 869 nm, and MOBY L_w. For the uncertainty in L_t, only sensor noise and systematic error are accounted for in this study, but there could be other sources, e.g., structured errors[53] that haven't been quantified. The calibration uncertainty on MODIS 869-nm band is assumed to be 2% in this study. The effect from this assumption is assessed by calculating the mean ratio and standard deviation of $u_c(R_{rs})$ assuming a 2.5% vs. 2% calibration uncertainty for that band over all the valid pixels in Fig. 6. The resulting changes in uncertainty are 1.20± 0.06 at band 443 nm and 1.40± 0.07 at band 547 nm. The uncertainty in MOBY L_w is estimated at between 2.3%-4.4% in the blue-red wavelengths using L_u uncertainty presented by [46] and the environmental uncertainty (personal communication with Kenneth J Voss). The effect of uncertainty in MOBY L_w is evaluated by comparing $u_c(R_{rs})$ derived using those values with that derived using a 5% constant uncertainty. The mean ratio of the former to the latter (and standard deviation) of $u_c(R_{rs})$ over all the valid pixels in Fig. 6 is 1.28± 0.051 at band 443 nm and 1.11±0.20 at band 547 nm. The effect is more significant at 443 nm than at 547 nm, due to the small $L_w(547)$ at MOBY.

While the evaluations using MC and validation results indicate the derivative method established in this study can provide reasonable $u_c(R_{rs})$, some issues need further investigation, including the need for more specific quantitative knowledge of the uncertainty in ancillary data, calibration at the 869-nm band, uncertainty in in situ measurements at MOBY. Forward model uncertainty is affected by the uncertainty in in situ L_w at MOBY, which is significant at blue bands. Despite this, the method shows significant progress towards providing useful pixel-level R_{rs} uncertainty estimates and can be updated as our knowledge of the contributing terms improves.

Appendix A. Calculation of partial derivative of R_{rs}

As described in Section 2.3, aerosol calculation starts with $L_a(748)$ and $L_a(869)$ (Eq.(9)), from which L_a at all bands are derived using MSEPS. During the AC process, the partial derivative of $L_a(\lambda)$, $t_v(\lambda)$, $t_s(\lambda)$, and $\tau_a(\lambda)$ with respect to $L_{rfc}(NIR)$, $\tau'_a(869)$, chl-a, and rh can be derived, denoted by $\frac{\partial L_a(\lambda)}{\partial L_{rfc}(NIR)}$, $\frac{\partial L_a(\lambda)}{\partial \tau'_a(869)}$, $\frac{\partial L_a(\lambda)}{\partial chl_a}$, and $\frac{\partial L_a(\lambda)}{\partial rh}$ (the same notation is adopted for the derivative of $t_v(\lambda)$, $t_s(\lambda)$, and $\tau_a(\lambda)$ by replacing $L_a(\lambda)$). Then, the partial derivatives of R_{rs} with respect to $L_{rfc}(NIR)$, $\tau'_a(869)$, chl-a, and rh can be derived:

$$\begin{aligned}
\frac{\partial R_{rs}(\lambda)}{\partial L_{rfc}(NIR)} &= \frac{-\partial L_a(\lambda)}{\partial L_{rfc}(NIR)} f_b(\lambda) \big/ [t_v(\lambda) t_s(\lambda) F_0(\lambda) cos\theta_s] - \frac{\partial t_s(\lambda)}{\partial L_{rfc}(NIR)} \\
&[L_{rfc}(\lambda) - TL_g(\lambda) - L_a(\lambda)] f_b(\lambda) \big/ [t_s^2(\lambda) t_v(\lambda) F_0(\lambda) cos\theta_s] - \frac{\partial t_v(\lambda)}{\partial L_{rfc}(NIR)} \\
&[L_{rfc}(\lambda) - TL_g(\lambda) - L_a(\lambda)] f_b(\lambda) \big/ [t_v^2(\lambda) t_s(\lambda) F_0(\lambda) cos\theta_s]
\end{aligned} \tag{A1a}$$

$$\begin{aligned}
\frac{\partial R_{rs}(\lambda)}{\partial \tau'_a(869)} &= \frac{-\partial L_a(\lambda)}{\partial \tau'_a(869)} f_b(\lambda) \big/ [t_v(\lambda) t_s(\lambda) F_0(\lambda) cos\theta_s] - \frac{\partial t_s(\lambda)}{\partial \tau'_a(869)} \\
&[L_{rfc}(\lambda) - TL_g(\lambda) - L_a(\lambda)] f_b(\lambda) \big/ [t_s^2(\lambda) t_v(\lambda) F_0(\lambda) cos\theta_s] - \frac{\partial t_v(\lambda)}{\partial \tau'_a(869)} \\
&[L_{rfc}(\lambda) - TL_g(\lambda) - L_a(\lambda)] f_b(\lambda) \big/ [t_v^2(\lambda) t_s(\lambda) F_0(\lambda) cos\theta_s] - \\
&\frac{\partial TLg(\lambda)}{\partial \tau'_a(869)} f_b(\lambda) \big/ [t_v(\lambda) t_s(\lambda) F_0(\lambda) cos\theta_s]
\end{aligned} \tag{A1b}$$

$$\begin{aligned}
\frac{\partial R_{rs}(\lambda)}{\partial chl_a} &= \frac{-\partial L_a(\lambda)}{\partial chla} f_b(\lambda) \big/ [t_v(\lambda) t_s(\lambda) F_0(\lambda) cos\theta_s] - \frac{\partial t_s(\lambda)}{\partial chl_a} \\
&[L_{rfc}(\lambda) - TL_g(\lambda) - L_a(\lambda)] f_b(\lambda) \big/ [t_s^2(\lambda) t_v(\lambda) F_0(\lambda) cos\theta_s] - \frac{\partial t_v(\lambda)}{\partial chl_a} \\
&[L_{rfc}(\lambda) - TL_g(\lambda) - L_a(\lambda)] f_b(\lambda) \big/ [t_v^2(\lambda) t_s(\lambda) F_0(\lambda) cos\theta_s] + \frac{\partial f_b(\lambda)}{\partial chl_a} \\
&[L_{rfc}(\lambda) - TL_g(\lambda) - L_a(\lambda)] \big/ [t_v(\lambda) t_s(\lambda) F_0(\lambda) cos\theta_s]
\end{aligned} \tag{A1c}$$

$$\frac{\partial R_{rs}(\lambda)}{\partial rh} = \frac{-\partial L_a(\lambda)}{\partial rh} f_b(\lambda) \Big/ [t_v(\lambda)t_s(\lambda)F_0(\lambda)cos\theta_s] - \frac{\partial t_s(\lambda)}{\partial rh}$$
$$[L_{rfc}(\lambda) - TL_g(\lambda) - L_a(\lambda)] f_b(\lambda) \Big/ [t_s^2(\lambda)t_v(\lambda)F_0(\lambda)cos\theta_s] - \frac{\partial t_v(\lambda)}{\partial rh}$$
$$[L_{rfc}(\lambda) - TL_g(\lambda) - L_a(\lambda)] f_b(\lambda) \Big/ [t_v^2(\lambda)t_s(\lambda)F_0(\lambda)cos\theta_s] \quad \text{(A1d)}$$

Based on Eq.(7), the partial derivative of $R_{rs}(\lambda)$ with respect to $L_{rfc}(\lambda)$ can be calculated as

$$\frac{\partial R_{rs}(\lambda)}{\partial L_{rfc}(\lambda)} = f_b(\lambda) / [t_v(\lambda)t_s(\lambda)F_0(\lambda)cos\theta_s] \quad \text{(A2)}$$

Table A1. Glossary of symbols

Symbol	Description	Unit
λ	Wavelength	nm
VIS	Visible bands	nm
NIR	Near infrared bands	nm
SWIR	Shortwave infrared bands	nm
L_t	top-of-atmosphere (TOA) radiance received by the sensor	$mW.cm^{-2}.\mu m^{-1}.sr^{-1}$
L_t^t	Predicted L_t in the vicarious calibration	$mW.cm^{-2}.\mu m^{-1}.sr^{-1}$
L_w	Water-leaving radiance	$mW.cm^{-2}.\mu m^{-1}.sr^{-1}$
L_{wn}	Normalized water-leaving radiance	$mW.cm^{-2}.\mu m^{-1}.sr^{-1}$
L_r	Radiance from air molecular scattering	$mW.cm^{-2}.\mu m^{-1}.sr^{-1}$
L_a	Radiance from aerosol scattering	$mW.cm^{-2}.\mu m^{-1}.sr^{-1}$
L_f	Radiance from foam scattering	$mW.cm^{-2}.\mu m^{-1}.sr^{-1}$
L_g	Sun glint	$mW.cm^{-2}.\mu m^{-1}.sr^{-1}$
L_u	Up-welling radiance	$mW.cm^{-2}.\mu m^{-1}.sr^{-1}$
t_v	Diffuse transmittance for view path	Unitless
t_{vr}	Diffuse transmittance for view path without aerosol	Unitless
t_s	Diffuse transmittance for sun path	Unitless
T	Beam transmittance for view path	Unitless
t_g	Transmittance for solar and sensor view paths from gas	Unitless
τ_a	aerosol optical thickness	Unitless
ε	Epsilon	Unitless
F_0	Extraterrestrial solar irradiance corrected for earth-sun distance	$mW.cm^{-2}.\mu m^{-1}$
E_d	Downward irradiance	$mW.cm^{-2}.\mu m^{-1}$
R_{rs}	Remote sensing reflectance	sr^{-1}
R_{rs}^m	MODIS Retrieved R_{rs}	sr^{-1}
R_{rs}^f	In situ R_{rs}	sr^{-1}
ρ_t	TOA reflectance	Unitless
ρ_a	Aerosol multiple scattering reflectance	Unitless
ρ_{as}	Aerosol single scattering reflectance	Unitless
rh	Relative humidity	Unitless
ws	wind speed	m/s
σ	Standard deviation	
chl-a	Chlorophyll-a concentration	mg/m^3
u_c	Combined standard uncertainty	
u	Standard uncertainty	
δ	Relative uncertainty	%
Δ_D	Expected discrepancy of R_{rs}	sr^{-1}
Δ_N	Uncertainty-normalized difference	Unitless
θ_s	Solar zenith angle	Degree
θ_v	Sensor view zenith angle	Degree
χ	Sensor noise	$mW.cm^{-2}.\mu m^{-1}.sr^{-1}$
f_b	Bidirectional reflectance correction	Unitless
f_p	Polarization correction	Unitless
g	Vicarious calibration gain	Unitless
r	Band-to-band correlation coefficient between ρ_t	Unitless

Appendix B. Estimation of instrument systematic uncertainty and forward model uncertainty

In order to quantify instrument systematic uncertainty and forward model uncertainty, we need to go through the SVC process following the approach presented by [42]. Due to a stable aerosol loading and negligible L_w(NIR), South Pacific Gyre (SPG) region was selected to calibrate the 748-nm band. As the pixels used for calibration are required to be free of sun glint, the predicted $L_t^t(748)$ can be expressed as:

$$L_t^t(748) = \left[L_r(748) + L_a(748) + t_v(748)L_f(748)\right]t_g(748)f_p(748) \qquad \text{(B1)}$$

where f_p is polarization correction factor [54]. $L_a(748)$ is extrapolated from $L_{rfc}(869)$ using the aerosol model determined by the time series of aerosol measurements over SPG. The partial derivatives of $L_a(748)$ and $t_v(748)$ with respect to $L_{rfc}(869)$ can be derived during the extrapolation of $L_{rfc}(869)$ to $L_a(748)$, denoted by $\frac{\partial L_a(748)}{\partial L_{rfc}(869)}$ and $\frac{\partial t_v(748)}{\partial L_{rfc}(869)}$, from which the partial derivative of $L_t^t(748)$ can be expressed as:

$$\frac{\partial L_t^t(748)}{\partial L_{rfc}(869)} = t_g(748)f_p(748)[\frac{\partial L_a(748)}{\partial L_{rfc}(869)} + L_f(748)\frac{\partial t_v(748)}{\partial L_{rfc}(869)}] \qquad \text{(B2a)}$$

The partial derivative of $L_t^t(748)$ with respect to $L_r(748)$, $L_f(748)$, and $t_g(748)$ can be written as:

$$\frac{\partial L_t^t(748)}{\partial L_r(748)} = t_g(748)f_p(748) \qquad \text{(B2b)}$$

$$\frac{\partial L_t^t(748)}{\partial L_f(748)} = t_v(748)t_g(748)f_p(748) \qquad \text{(B2c)}$$

$$\frac{\partial L_t^t(748)}{\partial t_g(748)} = [L_r(748) + L_a(748) + t_v(748)L_f(748)]f_p(748) \qquad \text{(B2d)}$$

Combining Eq.(B2) with $u_c(L_{rfc}(869))$, $u_c(L_r(748))$, $u_c(L_f(748))$, and $u_c(t_g(748))$, the uncertainty in $L_t^t(748)$ can be derived, denoted by $u_c(L_t^t(748))$. It should be noted that the uncertainty in $L_t(869)$ used to calculate $u_c(L_{rfc}(869))$ only include sensor noise, which is a limitation of this approach considering the possibility of a systematic error component in $L_t(869)$. A vicarious calibration gain sample (g_i) can be derived from:

$$g_i(748) = \frac{L_t^t(748)}{L_t(748)} \qquad \text{(B3)}$$

where $L_t(748)$ is the measured value with the uncertainty coming from sensor noise. Combining the partial derivative of g_i with respect to $L_t^t(748)$ and $L_t(748)$ derived from Eq.(B3) with $u_c(L_t^t(748))$ and $u(L_t(748))$, the uncertainty in g_i can be derived, denoted by $u_c(g_i(748))$.

Based on g(748) calculated by averaging all $g_i(748)$, VIS are vicariously calibrated using in situ L_w at MOBY. The uncertainty in L_w is derived from the uncertainty in L_u presented by [46] and the environmental uncertainty during the propagation of L_u to L_w. Pixels used for calibration are again required to be free from sun glint, so $L_t^t(VIS)$ can be expressed as:

$$L_t^t(\mathrm{VIS}) = \left[L_r(\mathrm{VIS}) + L_a(\mathrm{VIS}) + t_v(\mathrm{VIS})L_f(\mathrm{VIS}) + t_v(\mathrm{VIS})L_w(\mathrm{VIS})\right]t_g(\mathrm{VIS})f_p(\mathrm{VIS}) \qquad \text{(B4)}$$

With the assumption of negligible L_w(NIR), L_a(NIR) are equal to L_{rfc}(NIR) and then applied to calculate $L_a(\lambda)$ using MSEPS, from which the partial derivative of $L_a(\lambda)$ and $t_v(\lambda)$ with respect to L_{rfc}(NIR) can be calculated, denoted by $\frac{\partial L_a(\mathrm{VIS})}{\partial L_{rfc}(NIR)}$ and $\frac{\partial t_v(\mathrm{VIS})}{\partial L_{rfc}(NIR)}$. The derivative of L_t^t(VIS) with respect to L_{rfc}(NIR) can be derived as:

$$\frac{\partial L_t^t(\mathrm{VIS})}{\partial L_{rfc}(NIR)} = \left[\frac{\partial L_a(\mathrm{VIS})}{\partial L_{rfc}(NIR)} + \left(L_w(\mathrm{VIS}) + L_f(\mathrm{VIS})\right)\frac{\partial t_v(\mathrm{VIS})}{\partial L_{rfc}(NIR)}\right] t_g(\mathrm{VIS}) f_p(\mathrm{VIS}) \quad \text{(B5a)}$$

The derivative of $L_t^t(\lambda)$ with respect to $L_r(\lambda)$, L_f(VIS), L_w(VIS), t_g(VIS) can be expressed as:

$$\frac{\partial L_t^t(VIS)}{\partial L_r(VIS)} = t_g(VIS) f_p(\mathrm{VIS}) \quad \text{(B5b)}$$

$$\frac{\partial L_t^t(VIS)}{\partial L_f(VIS)} = \frac{\partial L_t^t(VIS)}{\partial L_w(VIS)} = t_v(VIS) t_g(VIS) f_p(\mathrm{VIS}) \quad \text{(B5c)}$$

$$\frac{\partial L_t^t(VIS)}{\partial t_g(VIS)} = [L_r(VIS) + L_a(VIS) + t_v(VIS) L_f(VIS) + t_v(VIS) L_w(VIS)] f_p(\mathrm{VIS}) \quad \text{(B5d)}$$

Combining Eq.(B5) with $u_c(L_{rfc}(\mathrm{NIR}))$, $u_c(L_r(\mathrm{VIS}))$, $u_c(L_f(\mathrm{VIS}))$, and $u(L_w(\mathrm{VIS}))$, the uncertainty in $L_t^t(VIS)$ can be calculated, denoted by $u_c(L_t^t(VIS))$. Using the same approach as that for 748 nm, g_i(VIS) and $u_c(g_i(\mathrm{VIS}))$ can be derived. Standard deviation, σ, is derived from all g_i at each band. The standard error (SE) is derived from:

$$SE = \frac{\sigma}{\sqrt{N}} \quad \text{(B6)}$$

where N is the number of vicarious calibration gain samples, which are 221 and 63 for 748 nm and VIS bands respectively. The forward model uncertainty is derived from subtracting in quadrature $u_c(g)$ and *SE* from σ. $u_c(g)$ is derived by averaging those N $u_c(g_i)$.

Appendix C. Correlation between ρ_t

As the 748-nm band is calibrated against the 869-nm band, and VIS bands are calibrated by NIR bands, covariance exist between ρ_t(VIS) and ρ_t(NIR) and between ρ_t(748) and ρ_t(869). Following the approach presented by [18], the correlation coefficients between ρ_t(VIS/748) and ρ_t(869) and between ρ_t(VIS) and ρ_t(748) are calculated using vicariously calibrated ρ_t over a box of 5×5 pixels in the SPG (26.5°-27.5°S, 124.5-123.5°W) with very clear waters. A valid box requires that all pixels are free of level 2 ocean color flags indicating processing problems (https://oceancolor.gsfc.nasa.gov/atbd/ocl2flags/) and the coefficient of variation of R_{rs} at 412 nm, 443 nm, and 488 nm is smaller than 1% (i.e. spatial stability). Using those 25 ρ_t, correlation coefficients between different bands can be derived. Table C1 lists the mean correlation coefficients over around 500 MODIS granules during 2002-2019. The correlation coefficients are used to calculate the covariance $u(L_t(\lambda_i), L_t(748))$, $u(L_t(\lambda_i), L_t(869))$, and $u(L_t(748), L_t(869))$ based on Eq.(12), which should be included when calculating $u_c(R_{rs}(\lambda_i))$. It should be noted here that the correlation between ρ_t is assumed to be a good approximation of inter-band error correlation.

Table C1. Correlation coefficients between ρ_t at VIS and NIR bands.

r	412	443	488	531	547	555	667	678	748
748	0.51	0.59	0.74	0.85	0.87	0.81	0.97	0.97	1.0
869	0.45	0.53	0.68	0.80	0.82	0.76	0.94	0.95	0.97

Funding. NASA Terra and Aqua Senior Review for MODIS algorithm maintenance and the NASA PACE Project.

Acknowledgements. We acknowledge NASA's Ocean Biology Distributed Active Archive Center for making available all the in situ data and MODIS data used in our analysis. We also thank the in situ data collection teams (and Principal Investigators) from MOBY (K. Voss), AAOT (G. Zibordi), and BOUSSOLE (D. Antoine) for the collection, processing, and quality control of those datasets. We further wish to acknowledge K. Voss for helpful

discussion and insight into the uncertainties on the MOBY measurements.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. M. Defoin-Platel and M. Chami, "How ambiguous is the inverse problem of ocean color in coastal waters?," Journal of Geophysical Research: Oceans **112**(2007).
2. D. Antoine, F. d'Ortenzio, S. B. Hooker, G. Bécu, B. Gentili, D. Tailliez, and A. J. Scott, "Assessment of uncertainty in the ocean reflectance determined by three satellite ocean color sensors (MERIS, SeaWiFS and MODIS-A) at an offshore site in the Mediterranean Sea (BOUSSOLE project)," Journal of Geophysical Research: Oceans **113**(2008).
3. IOCCG, "Uncertainties in Ocean Colour Remote Sensing," (International Ocean Colour Coordinating Group, Dartmouth, Canada, 2019).
4. F. Mélin, P. Colandrea, P. D. Vis, and S. E. Hunt, "Sensitivity of Ocean Color Atmospheric Correction to Uncertainties in Ancillary Data: A Global Analysis With SeaWiFS Data," IEEE Transactions on Geoscience and Remote Sensing **60**, 1-18 (2022).
5. P. De Vis, F. Mélin, S. E. Hunt, R. Morrone, M. Sinclair, and B. Bell, "Ancillary Data Uncertainties within the SeaDAS Uncertainty Budget for Ocean Colour Retrievals," Remote Sensing **14**, 497 (2022).
6. N. Fox, "A guide to expression of uncertainty of measurements " (GEO, 2010).
7. C. D. Mobley, "Estimation of the remote-sensing reflectance from above-surface measurements," Applied Optics **38**, 7442-7455 (1999).
8. S. B. Hooker and S. Maritorena, "An Evaluation of Oceanographic Radiometers and Deployment Methodologies," Journal of Atmospheric and Oceanic Technology **17**, 811-830 (2000).
9. F. Mélin, "From Validation Statistics to Uncertainty Estimates: Application to VIIRS Ocean Color Radiometric Products at European Coastal Locations," Frontiers in Marine Science **8**(2021).
10. S. W. Bailey and P. J. Werdell, "A multi-sensor approach for the on-orbit validation of ocean color satellite data products," Remote Sensing of Environment **102**, 12-23 (2006).
11. M. Zhang, C. Hu, J. Cannizzaro, M. G. Kowalewski, and S. J. Janz, "Diurnal changes of remote sensing reflectance over Chesapeake Bay: Observations from the Airborne Compact Atmospheric Mapper," Estuarine, Coastal and Shelf Science **200**, 181-193 (2018).

12. T. Jackson, S. Sathyendranath, and F. Mélin, "An improved optical classification scheme for the Ocean Colour Essential Climate Variable and its applications," Remote Sensing of Environment **203**, 152-161 (2017).
13. C. Hu, L. Feng, and Z. Lee, "Uncertainties of SeaWiFS and MODIS remote sensing reflectance: Implications from clear water measurements," Remote Sensing of Environment **133**, 168-182 (2013).
14. F. Mélin, G. Sclep, T. Jackson, and S. Sathyendranath, "Uncertainty estimates of remote sensing reflectance derived from comparison of ocean color satellite data sets," Remote Sensing of Environment **177**, 107-124 (2016).
15. J. Concha, A. Mannino, B. Franz, and W. Kim, "Uncertainties in the Geostationary Ocean Color Imager (GOCI) Remote Sensing Reflectance for Assessing Diurnal Variability of Biogeochemical Processes," Remote Sensing **11**, 295 (2019).
16. Z. Lee, R. Arnone, C. Hu, P. J. Werdell, and B. Lubac, "Uncertainties of optical parameters and their propagations in an analytical ocean color inversion algorithm," Applied Optics **49**, 369-381 (2010).
17. L. I. W. McKinna, I. Cetinić, A. P. Chase, and P. J. Werdell, "Approach for Propagating Radiometric Data Uncertainties Through NASA Ocean Color Algorithms," Frontiers in Earth Science **7**(2019).
18. N. Lamquin, A. Mangin, C. Mazeran, B. Bourg, V. Bruniquel, and O. F. D'Andon, "OLCI L2 Pixel-by-Pixel Uncertainty Propagation in OLCI Clean Water Branch," (ESA, 2013), p. 51.
19. 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," International Journal of Remote Sensing **20**, 1875-1916 (1999).
20. B. A. Franz and E. M. Karaköylü, "PACE OCI Signal to Noise Performance Requirement: Assessment and Verification Approach for Ocean Color Science," (Goddard Space Flight Cente, Maryland, 2018).
21. C. Mobley, J. Werdell, B. Franz, Z. Ahmad, and S. Bailey, "*Atmospheric Correction for Satellite Ocean Color Radiometry*," (NASA Goddard Space Flight Cente, Maryland,2016).
22. H. R. Gordon and M. Wang, "Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: a preliminary algorithm," Applied Optics **33**, 443-452 (1994).
23. D. B. Gillis, J. H. Bowles, M. J. Montes, and W. J. Moses, "Propagation of sensor noise in oceanic hyperspectral remote sensing," Optics Express **26**, A818-A831 (2018).
24. B.-C. Gao, M. J. Montes, Z. Ahmad, and C. O. Davis, "Atmospheric correction algorithm for hyperspectral

remote sensing of ocean color from space," Applied Optics **39**, 887-896 (2000).

25. A. Ibrahim, B. A. Franz, Z. Ahmad, and S. W. Bailey, "Multiband Atmospheric Correction Algorithm for Ocean Color Retrievals," Frontiers in Earth Science **7**(2019).

26. Z. Ahmad and B. A. Franz, "Uncertainty in aerosol model characterization and its impact on ocean color retrievals," (Goddard Space Flight Center, Maryland, 2018).

27. Z. Ahmad and B. A. Franz, "Ocean color retrieval using multiple-scattering epsilon values," in *International Ocean Color Science Meeting 2015,* (2015).

28. D. K. Clark, H. R. Gordon, K. J. Voss, Y. Ge, W. Broenkow, and C. Trees, "Validation of atmospheric correction over the oceans," Journal of Geophysical Research: Atmospheres **102**, 17209-17217 (1997).

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

30. D. Antoine, P. Guevel, J.-F. o. Dest?, G. B?cu, F. Louis, A. J. Scott, and P. Bardey, "The ?BOUSSOLE? Buoy?A New Transparent-to-Swell Taut Mooring Dedicated to Marine Optics: Design, Tests, and Performance at Sea," Journal of Atmospheric and Oceanic Technology **25**, 968-989 (2008).

31. M. Wang, "A refinement for the Rayleigh radiance computation with variation of the atmospheric pressure," International Journal of Remote Sensing **26**, 5651-5566 (2005).

32. H. R. Gordon and M. Wang, "Influence of oceanic whitecaps on atmospheric correction of ocean-color sensors," Applied Optics **33**, 7754-7763 (1994).

33. D. Stramski and J. Piskozub, "Estimation of Scattering Error in Spectrophotometric Measurements of Light Absorption by Aquatic Particles from Three-Dimensional Radiative Transfer Simulations," Appl. Opt. **42**, 3634-3646 (2003).

34. C. Cox and W. Munk, "Measurement of the Roughness of the Sea Surface from Photographs of the Sun?s Glitter," J. Opt. Soc. Am. **44**, 838-850 (1954).

35. M. Wang and S. W. Bailey, "Correction of sun glint contamination on the SeaWiFS ocean and atmosphere products," Applied Optics **40**, 4790-4798 (2001).

36. Z. Ahmad, B. A. Franz, C. R. McClain, E. J. Kwiatkowska, J. Werdell, E. P. Shettle, and B. N. Holben, "New aerosol models for the retrieval of aerosol optical thickness and normalized water-leaving radiances from the SeaWiFS and MODIS sensors over coastal regions and open oceans," Applied Optics **49**, 5545-5560 (2010).

37. S. W. Bailey, B. A. Franz, and P. J. Werdell, "Estimation of near-infrared water-leaving reflectance for satellite

ocean color data processing," Optics Express **18**, 7521-7527 (2010).

38. A. Morel, D. Antoine, and B. Gentili, "Bidirectional reflectance of oceanic waters: accounting for Raman emission and varying particle scattering phase function," Applied Optics **41**, 6289-6306 (2002).

39. JCGM, "Evaluation of measurementdata — Guide to the expression of uncertainty in measurement," (2008).

40. M. Stramska and T. Petelski, "Observations of oceanic whitecaps in the north polar waters of the Atlantic," Journal of Geophysical Research: Oceans **108**, n/a-n/a (2003).

41. X. Xiong, J. Sun, X. Xie, W. L. Barnes, and V. V. Salomonson, "On-Orbit Calibration and Performance of Aqua MODIS Reflective Solar Bands," IEEE Transactions on Geoscience and Remote Sensing **48**, 535-546 (2010).

42. B. A. Franz, S. W. Bailey, P. J. Werdell, and C. R. McClain, "Sensor-independent approach to the vicarious calibration of satellite ocean color radiometry," Applied Optics **46**, 5068-5082 (2007).

43. X. Xiong, J. Sun, A. Wu, K.-F. Chiang, J. Esposito, and W. Barnes, *Terra and Aqua MODIS calibration algorithms and uncertainty analysis*, SPIE Remote Sensing (SPIE, 2005), Vol. 5978.

44. A. M. Sayer, Y. Govaerts, P. Kolmonen, A. Lipponen, M. Luffarelli, T. Mielonen, F. Patadia, T. Popp, A. C. Povey, K. Stebel, and M. L. Witek, "A review and framework for the evaluation of pixel-level uncertainty estimates in satellite aerosol remote sensing," Atmos. Meas. Tech. **13**, 373-404 (2020).

45. G. Zibordi, M. Talone, and F. Mélin, "Uncertainty Estimate of Satellite-Derived Normalized Water-Leaving Radiance," IEEE Geoscience and Remote Sensing Letters **19**, 1-5 (2022).

46. S. Brown, S. Flora, M. Feinholz, M. Yarbrough, T. Houlihan, D. Peters, Y. S. Kim, J. Mueller, B. C. Johnson, and D. Clark, *The marine optical buoy (MOBY) radiometric calibration and uncertainty budget for ocean color satellite sensor vicarious calibration*, SPIE Remote Sensing (SPIE, 2007), Vol. 6744.

47. A. Białek, V. Vellucci, B. Gentil, D. Antoine, J. Gorroño, N. Fox, and C. Underwood, "Monte Carlo–Based Quantification of Uncertainties in Determining Ocean Remote Sensing Reflectance from Underwater Fixed-Depth Radiometry Measurements," Journal of Atmospheric and Oceanic Technology **37**, 177-196 (2020).

48. M. Gergely and G. Zibordi, "Assessment of AERONET-OC LWN uncertainties," Metrologia **51**, 40-47 (2013).

49. G. Zibordi, J.-F. Berthon, F. Mélin, D. D'Alimonte, and S. Kaitala, "Validation of satellite ocean color primary producet at optically complex coastal sites: Northern Adriatic Sea, Northern Baltic Proper and Gulf of Finland," Remote Sensing of Environment **113** 2574-2591

(2009).

50. G. Zibordi and F. Mélin, "An evaluation of marine regions relevant for ocean color system vicarious calibration," Remote Sensing of Environment **190**, 122-136 (2017).

51. C. Hu, Z. Lee, and B. Franz, "Chlorophyll aalgorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference," Journal of Geophysical Research: Oceans **117**, n/a-n/a (2012).

52. S. B. Hooker, W. E. Esaias, G. C. Feldman, W. W. Gregg, and C. R. McClain, "An overview of SeaWiFS and ocean color," (Goddard Space Flight Center, Greenbelt, MD, 1992).

53. J. Mittaz, C. J. Merchant, and E. R. Woolliams, "Applying principles of metrology to historical Earth observations from satellites," Metrologia **56**, 032002 (2019).

54. G. Meister, E. J. Kwiatkowska, B. A. Franz, F. S. Patt, G. C. Feldman, and C. R. McClain, "Moderate-Resolution Imaging Spectroradiometer ocean color polarization correction," Applied Optics **44**, 5524-5535 (2005).