



OPEN ACCESS

EDITED BY

Zhigang Cao,
Chinese Academy of Sciences (CAS), China

REVIEWED BY

Agnieszka Bialek,
National Physical Laboratory, United Kingdom
Sundarabalan Balasubramanian,
Geo-Sensing and Imaging Consultancy, India

*CORRESPONDENCE

Minwei Zhang,
✉ minwei.zhang@nasa.gov

RECEIVED 21 July 2025

ACCEPTED 23 September 2025

PUBLISHED 10 October 2025

CITATION

Zhang M, Ibrahim A, Franz BA, Sayer AM,
Werdell PJ and McKinna LI (2025) Practical
aspects of providing pixel-level spectral R_{rs}
error covariance in satellite ocean
color products.
Front. Remote Sens. 6:1670390.
doi: 10.3389/frsen.2025.1670390

COPYRIGHT

© 2025 Zhang, Ibrahim, Franz, Sayer, Werdell
and McKinna. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Practical aspects of providing pixel-level spectral R_{rs} error covariance in satellite ocean color products

Minwei Zhang^{1,2*}, Amir Ibrahim², Bryan A. Franz²,
Andrew M. Sayer^{2,3}, P. Jeremy Werdell² and Lachlan I. McKinna^{2,4}

¹Science Application International Corp., McLean, VA, United States, ²Ocean Ecology Laboratory, Goddard Space Flight Center, National Aeronautics and Space Administration, Greenbelt, MD, United States, ³GESTAR II, University of Maryland Baltimore County, Baltimore, MD, United States, ⁴Go2Q Pty Ltd., Sunshine Coast, QLD, Australia

We previously established a derivative-based approach to generate a pixel-level spectral error covariance matrix in satellite-retrieved remote sensing reflectance, ΣR_{rs} . However, one practical issue is the delivery of the products without increasing the file size by an order of magnitude or more, considering that for N sensor spectral bands, there are $N \times (N+1)/2$ covariance matrix elements to be specified at each pixel. The issue becomes more pertinent for hyperspectral imaging spectroradiometers such as the Ocean Color Instrument (OCI) on NASA's Plankton, Aerosol, Cloud, ocean Ecosystem mission (PACE), which has 286 bands, resulting in ~40,000 unique elements in ΣR_{rs} per pixel that would lead to a ~60 GB Level-2 file for one 5-min granule. As a first step to tackle the issue, we took OCI and Moderate Resolution Imaging Spectroradiometer (MODIS) data to explore the possibility of approximating ΣR_{rs} using a third-degree polynomial, thereby decreasing the memory overhead to $4 \times N$ numbers. We found that ΣR_{rs} derived from the polynomial fitting matches well with the original value, with the difference smaller than 5%. We then compared the relative uncertainty in two derived ocean color data products (chl_a and $K_d(490)$) calculated using the original fully computed ΣR_{rs} and then using the polynomial model approximation for ΣR_{rs} , finding the absolute difference between the two approaches to be smaller than 0.5%. These evaluations suggest the polynomial approximation of ΣR_{rs} is suitable without degrading the scientific quality. By including the coefficients derived from polynomial fitting instead of the full error covariance matrix, a typical 5-min Level-2 file for OCI decreases from ~60 GB to a more practical ~1.7 GB.

KEYWORDS

error covariance, ocean color, PACE, OCI, remote sensing reflectance, MODIS

1 Introduction

Due to the capability of obtaining a long time series of data at a global scale, satellite remote sensing has been widely used to study Earth System science and global change (Dutkiewicz et al., 2019; Cael et al., 2023). As a product fundamental to deriving a host of bio-optical and biogeochemical properties of the water column, ocean remote sensing reflectance (R_{rs} , sr^{-1}) has been routinely generated from satellite radiometry and distributed by NASA's Ocean Biology Distributed Active Archive Center (OB.DAAC) for more than 3 decades. While R_{rs} and properties derived from it have been validated against ground

truth, there is no routinely generated pixel-level uncertainty product. Although the uncertainty in water-leaving reflectance at the pixel level is provided for the Ocean and Land Colour Instrument (OLCI) onboard Sentinel-3 (Lamquin et al., 2013), it is underestimated, as only sensor noise is included. It should be noted that the European Space Agency's Ocean Color Climate Change Initiative program (OC_CCI) does provide pixel-level uncertainty in R_{rs} (Jackson et al., 2017), computed using the weighted average of the uncertainty within optical water classes associated with that pixel. The uncertainty of each class is derived from validation against ground truth, which has an issue with representing the pixel-level uncertainty, as described by Zhang et al. (2022). Furthermore, this approach could not capture spatiotemporal variations in some key drivers of uncertainty in R_{rs} , such as geometry and atmospheric turbidity. Without associated uncertainty, ocean color data products are incomplete measurements, limiting the interpretation of subtle changes in global ocean biology and biogeochemistry. Uncertainty in ocean color products has been recommended by the Group on Earth Observations (GEO) and the International Ocean Colour Coordinating Group (IOCCG) within the framework of quality assurance for earth observation and should be required by any satellite mission (Fox, 2010; IOCCG, 2019).

Historically, the ocean color community has relied on comparison with collocated *in situ* data to estimate uncertainty in ocean color products via “matchup” validation analyses. While the validation generates reasonable confidence in those products, it has some limitations (Zhang et al., 2022). In addition, while an analytical framework for estimating standard uncertainties in derived biogeochemical data products has been developed (McKinna et al., 2019), implementing this operationally requires pixel-level knowledge of standard uncertainties in $R_{rs}(\lambda)$ and the error covariance between R_{rs} at different bands. Because of these limitations, we developed a derivative-based approach to calculate the pixel-level standard uncertainty in $R_{rs}(\lambda)$ retrieved from the Moderate-Resolution Imaging Spectroradiometer (MODIS) aboard Aqua using a multiple scattering epsilon (MSEPS) atmospheric correction algorithm (Ahmad and Franz, 2015), which has been used recently to complete a multi-mission ocean color reprocessing for heritage sensors by the Ocean Biology Processing Group (OBPG) at NASA. The $R_{rs}(\lambda)$ uncertainty products generated by Zhang et al. (2022) and Zhang et al. (2024) were demonstrated to be reliable through evaluation using the Monte Carlo method as well as closure analysis with the statistics derived from validation.

While uncertainty can be readily generated, one practical issue is the delivery of all spectral uncertainty products without increasing data file sizes by an order of magnitude or more. A spectral error covariance matrix, $\sum R_{rs}$, is necessary. This matrix contains the variance in $R_{rs}(\lambda_i)$ at the i th band as the diagonal elements and the covariance between R_{rs} at the i th and j th band, denoted as $u(R_{rs}(\lambda_i), R_{rs}(\lambda_j))$, as the off-diagonal elements. We note the off-diagonal terms in $\sum R_{rs}$ are not negligible (Zhang et al., 2024) and must be considered when more than one band is used to derive data products, for example, the band ratio algorithms for chlorophyll-*a* pigment concentration (chl_a , mg/m³) or the diffuse attenuation coefficient for downwelling irradiance at 490 nm ($K_d(490)$, m⁻¹).

Taking MODIS, for example, with $N = 10$ sensor spectral bands spanning 412 nm–678 nm, $\sum R_{rs}$ has a dimension of 10×10 .

Considering that $\sum R_{rs}$ is symmetric, a total of 55 ($N \times (N+1)/2$) numbers must be stored at each pixel. The issue becomes more pertinent as we move forward to hyperspectral sensors with hundreds of bands, such as the Ocean Color Instrument (OCI), with 286 bands ranging from ultraviolet to shortwave infrared wavelengths on NASA's Plankton, Aerosol, Cloud, ocean Ecosystem mission (PACE) (<https://pace.oceansciences.org/about.htm>), which launched in February 2024. It is unfeasible to deliver tens of thousands of $\sum R_{rs}$ values at each pixel, given the broad (~2,600 km) swath and moderate (1.2 km at the sub-satellite point) spatial footprint. As a first step toward delivering uncertainty products without unmanageably increasing the data file size, we investigate in this study the spectral variation of $\sum R_{rs}$ retrieved from MODIS and OCI data using the MSEPS algorithm, with the goal of decreasing the dimensionality needed to reproduce it well enough for downstream scientific use.

2 Satellite data

Level-1A (uncalibrated) data of MODIS aboard the Aqua satellite and Level-1B (calibrated and geolocated) data of OCI from the PACE mission were downloaded from the Ocean Biology Distributed Active Archive Center (OB.DAAC). Level-1A MODIS data were processed to Level-1B using the SeaDAS software package (<https://seadas.gsfc.nasa.gov>) and the latest instrument calibration coefficients, which are also distributed by OB.DAAC. $R_{rs}(\lambda)$ was derived by applying the MSEPS atmospheric correction algorithm to L1B data, and the uncertainty in $R_{rs}(\lambda)$ was then calculated using the derivative-based approach established by Zhang et al. (2022). Here, we used MODIS data (i) over the Marine Optical BuoY (MOBY) (Clark et al., 2003) during 2002–2019; (ii) over the U.S. East Coast on 28 January 2012; and (iii) globally during 3–10 December 2019. We used OCI data over the U.S. East Coast on 13 May 2024.

3 Methodology

3.1 MSEPS atmospheric correction

We used MSEPS to calculate $R_{rs}(\lambda)$ from spectral top-of-atmosphere (TOA) radiances $L_t(\lambda)$, which can be expressed as Equation 1 (the spectral term, λ , is omitted for brevity):

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

where $t_g(\lambda)$ is two-way gas transmittance, $L_r(\lambda)$ (mW.cm⁻².μm⁻¹.sr⁻¹) is the radiance from scattering by air molecules in the absence of aerosol, $L_a(\lambda)$ (mW.cm⁻².μm⁻¹.sr⁻¹) is the radiance from scattering by aerosols that also accounts for interactions between air molecules and aerosol scattering, $L_f(\lambda)$ (mW.cm⁻².μm⁻¹.sr⁻¹) is the radiance from scattering by surface whitecaps, $L_g(\lambda)$ (mW.cm⁻².μm⁻¹.sr⁻¹) is the sun glint, and $t_{vr}(\lambda)$ and $t_v(\lambda)$ are the diffuse transmittance from surface to sensor with the former only including molecular scattering and the latter including molecular and aerosol scattering. $T(\lambda)$ is the beam transmittance for the view path. $L_f(\lambda)$ is calculated using wind

speed based on an empirical model (Stramska and Petelski, 2003). A sun glint coefficient is calculated using a statistical model (Cox and Munk, 1954), from which $L_g(\lambda)$ can be derived (Wang and Bailey, 2001). The calculation of $L_a(\lambda)$ is detailed by Ahmad and Franz (2015). After removing $L_r(\lambda)$, $L_a(\lambda)$, $L_t(\lambda)$, and $T(\lambda)$ $L_g(\lambda)$ from $L_t(\lambda)$, water-leaving radiance $L_w(\lambda)$ ($\text{mW}\cdot\text{cm}^{-2}\cdot\mu\text{m}^{-1}\cdot\text{sr}^{-1}$) can be derived, and $R_{rs}(\lambda)$ is then calculated by normalizing $L_w(\lambda)$ to downwelling irradiance (the spectral term, λ , is omitted for brevity):

$$R_{rs} = (L_t/t_g - L_r - TL_g - t_{vr}L_f - L_a)f_b / (t_v F_0 t_s \cos(\theta_s)) \quad (2)$$

where f_b is the bidirectional reflectance correction calculated from the model presented by Morel et al. (2002) using chl_a , $t_s(\lambda)$ is the spectral diffuse transmittance from the Sun to the surface, θ_s is the solar zenith angle, and $F_0(\lambda)$ ($\text{mW}\cdot\text{cm}^{-2}\cdot\mu\text{m}^{-1}$) is the spectral extraterrestrial solar irradiance accounting for the Earth–Sun distance. F_0 for OCI is calculated from the solar irradiance measurements from the Total and Spectral Solar Irradiance Sensor-1 (TSIS-1) (Coddington et al., 2021). F_0 for MODIS is from Thuillier et al. (2004).

3.2 Error covariance in R_{rs}

While the approach for calculating the error covariance matrix terms $u(R_{rs}(\lambda_i), R_{rs}(\lambda_j))$ is detailed by Zhang et al. (2024), we briefly recap here, taking MODIS as an example. It is based on a derivative approach. Two variables y_1 and y_2 are function of variables x , z , expressed by Equation 3:

$$\begin{cases} y_1 = f_1(x_1, x_2, \dots, x_n) \\ y_2 = f_2(z_1, z_2, \dots, z_n) \end{cases} \quad (3)$$

error covariance between y_1 and y_2 , denoted by $u(y_1, y_2)$, following the convention of (IOCCG, 2019), can be derived from Equation 4:

$$u(y_1, y_2) = \left(\frac{\partial f_1}{\partial x_1}, \frac{\partial f_1}{\partial x_2}, \dots, \frac{\partial f_1}{\partial x_n} \right) \begin{bmatrix} u(x_1, z_1) & \dots & u(x_1, z_n) \\ \vdots & \ddots & \vdots \\ u(x_n, z_1) & \dots & u(x_n, z_n) \end{bmatrix} \begin{pmatrix} \frac{\partial f_2}{\partial z_1} \\ \frac{\partial f_2}{\partial z_2} \\ \vdots \\ \frac{\partial f_2}{\partial z_n} \end{pmatrix}^T \quad (4)$$

where $u(x_i, z_j)$ represents the error covariance between variables x_i and z_j , $\frac{\partial f_1}{\partial x_i}$ and $\frac{\partial f_2}{\partial z_j}$ are the partial derivatives of y_1 with respect to x_i and of y_2 with respect to z_j , and T is the transpose operator. $u(y_1, y_2)$ turns to variance when f_1 and f_2 are same, denoted by $u^2(y_1)$. $u(y_1)$ is the standard uncertainty (JCGM, 2011). Based on Equations 2,4, $u(R_{rs}(\lambda_i), R_{rs}(\lambda_j))$ can be calculated from Equations 5–7:

$$u(R_{rs}(\lambda_i), R_{rs}(\lambda_j)) = \frac{\partial R_{rs}(\lambda_i)}{\partial X_i} C_{R_{rs}} \left(\frac{\partial R_{rs}(\lambda_j)}{\partial X_j} \right)^T \quad (5)$$

where

$$\frac{\partial R_{rs}(\lambda_i)}{\partial X_i} = \left(\frac{\partial R_{rs}(\lambda_i)}{\partial L_{rfc}(\lambda_i)}, \frac{\partial R_{rs}(\lambda_i)}{\partial L_{rfc}(\text{NIR})}, \frac{\partial R_{rs}(\lambda_i)}{\partial chl_a}, \frac{\partial R_{rs}(\lambda_i)}{\partial \tau'_a(869)}, \frac{\partial R_{rs}(\lambda_i)}{\partial rh} \right) \quad (6)$$

and

$$C_{R_{rs}} = \begin{bmatrix} u(L_{rfc}(\lambda_i), L_{rfc}(\lambda_j)) & u(L_{rfc}(\lambda_i), L_{rfc}(\text{NIR})) & 0 & 0 & 0 \\ u(L_{rfc}(\text{NIR}), L_{rfc}(\lambda_j)) & u^2(L_{rfc}(\text{NIR})) & 0 & 0 & 0 \\ 0 & 0 & u^2(chl_a) & 0 & 0 \\ 0 & 0 & 0 & u^2(\tau'_a(869)) & 0 \\ 0 & 0 & 0 & 0 & u^2(rh) \end{bmatrix} \quad (7)$$

X_i is a vector of the uncertainty variables that are used to propagate the uncertainty through the atmospheric correction procedure, which is defined as Equation 8:

$$X_i = [L_{rfc}(\lambda_i), L_{rfc}(\text{NIR}), chl_a, \tau'_a(869), rh] \quad (8)$$

where NIR corresponds to the two near-infrared bands used for MSEPS, τ'_a is the aerosol optical thickness, and rh is the relative humidity. L_{rfc} is defined as Equation 9 (the spectral term, λ , is omitted for brevity):

$$L_{rfc} = L_t/t_g - L_r - t_{vr}L_f \quad (9)$$

3.3 Polynomial fitting of spectral $u(R_{rs}(\lambda_i), R_{rs}(\lambda))$

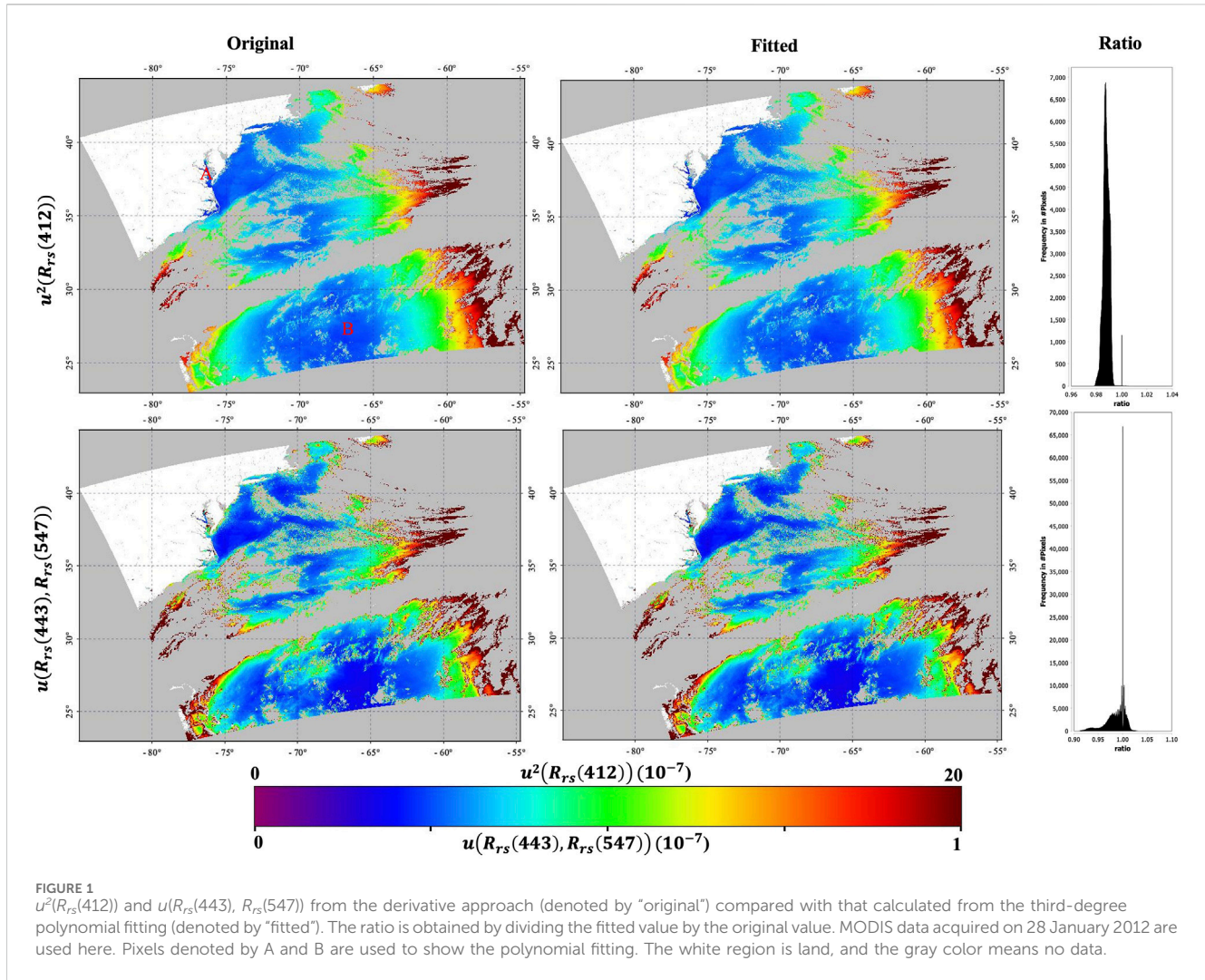
Spectral $u(R_{rs}(\lambda_i), R_{rs}(\lambda_j))$ are analyzed by fitting $u(R_{rs}(\lambda_i), R_{rs}(\lambda))$ as a function of wavelength. Here, λ_i and λ represent a specific wavelength and all the wavelengths, respectively. For example, $u(R_{rs}(412), R_{rs}(\lambda))$ is fitted as a function of MODIS bands in the wavelength range of 412–678 nm. Only visible wavelengths are considered in this study. We noted previously that the covariance was often spectrally smooth and decayed as a function of wavelength difference (Zhang et al., 2024). As a result, we experimented using a third-degree polynomial:

$$u(R_{rs}(\lambda_i), R_{rs}(\lambda)) = a + b\lambda + c\lambda^2 + d\lambda^3 \quad (10)$$

where λ is the central wavelength in units of μm . As the matrix is symmetric, we do not “re-use” wavelengths when fitting polynomials as this will lead to an asymmetry. There are, therefore, ten points used to fit the coefficients for 412 nm ($\lambda = 412, 443, 469, 488, 531, 547, 555, 645, 667$, and 678), nine for 443 nm, eight for 469 nm, and so on. This means that there is no need to fit coefficients for wavelengths of 555 nm or longer, considering that there are four or fewer covariances to use for this band, which is equal to or less than the four coefficients needed for the fitting (so it is more efficient to reproduce the matrix exactly). The same fitting method is also applied to OCI. The coefficients a , b , c , and d in Equation 10, derived by means of least-square fitting, are stored in the Level-2 file for each band λ_i .

3.4 Uncertainty in chl_a and $K_d(490)$

Variance in chl_a (denoted by $u^2(chl_a)$), calculated from the NASA standard algorithm (Hu et al., 2019; O'Reilly and Werdell, 2019), and in $K_d(490)$, denoted by $u^2(K_d(490))$, calculated from the algorithm presented by NASA (2023) is used here to evaluate the effect from polynomial fitting. Specifically, $u^2(chl_a)$ and $u^2(K_d(490))$ are calculated twice. First, they are calculated using the $\sum R_{rs}$ generated from the derivative approach, and then they are



calculated using the polynomial model for $\sum R_{rs}$. The difference in $u^2(chl_a)$ and $u^2(K_d(490))$ calculated from these two approaches is used to evaluate the effect of polynomial fitting. Based on the derivative approach, $u^2(chl_a)$ and $u^2(K_d(490))$ are calculated from Equation 11:

$$\left. \begin{aligned} u^2(chl_a) &= \sum_{i=1}^4 \left(\frac{\partial chl_a}{\partial R_{rs}(\lambda_i)} \right)^2 u^2(R_{rs}(\lambda_i)) \\ &+ 2 \sum_{i=1}^3 \sum_{j=i+1}^4 \frac{\partial chl_a}{\partial R_{rs}(\lambda_i)} \frac{\partial chl_a}{\partial R_{rs}(\lambda_j)} u(R_{rs}(\lambda_i), R_{rs}(\lambda_j)) + u_{mc}^2 \\ u^2(K_d(490)) &= \sum_{i=1}^2 \left(\frac{\partial K_d(490)}{\partial R_{rs}(\lambda_i)} \right)^2 u^2(R_{rs}(\lambda_i)) \\ &+ 2 \frac{\partial K_d(490)}{\partial R_{rs}(\lambda_1)} \frac{\partial K_d(490)}{\partial R_{rs}(\lambda_2)} u(R_{rs}(\lambda_1), R_{rs}(\lambda_2)) + u_{mk}^2 \end{aligned} \right\} \quad (11)$$

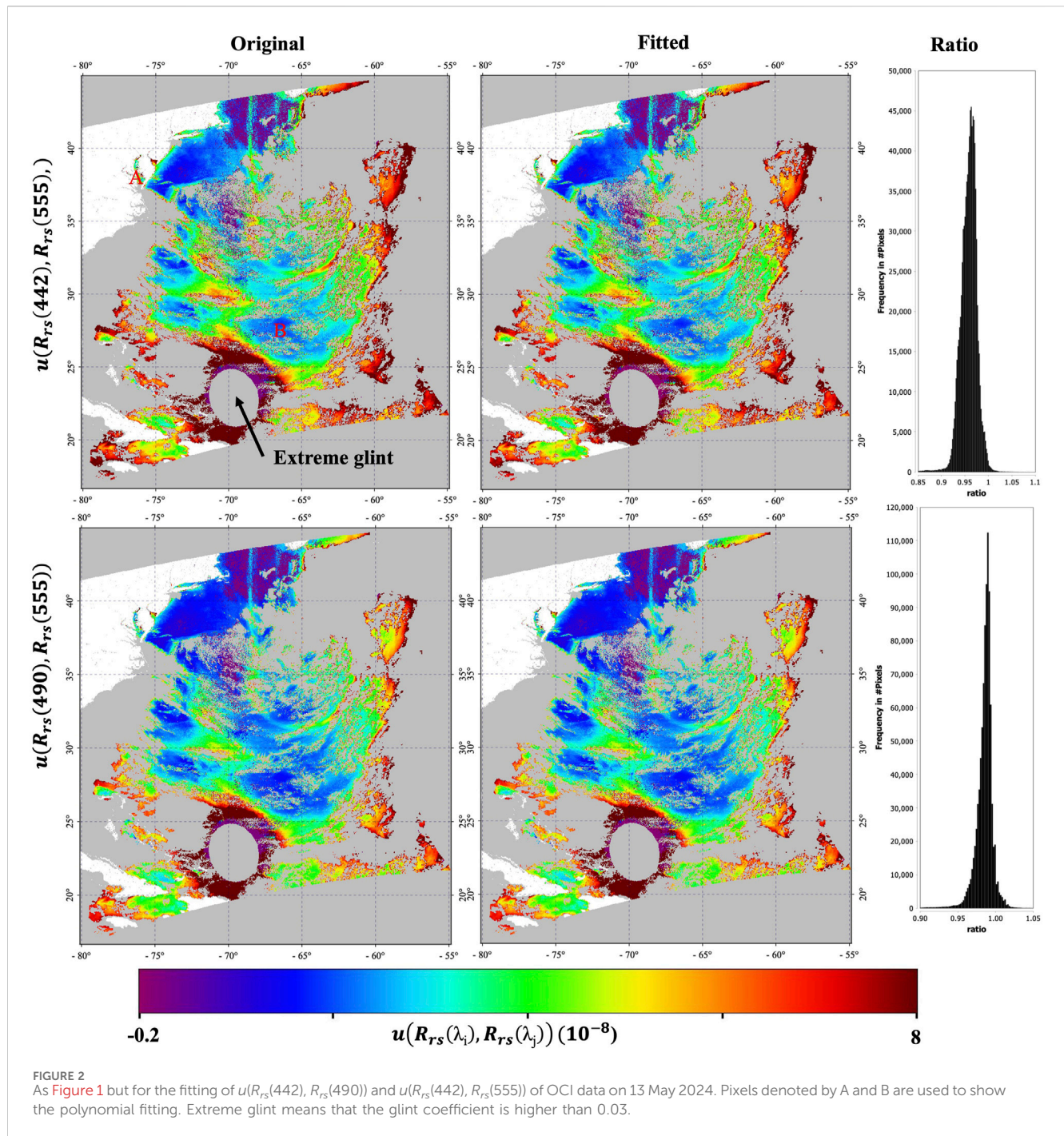
MODIS bands at 443 nm, 488 nm, 547 nm, and 667 nm are used in the standard chl_a algorithm. Bands at 488 nm and 547 nm are used to calculate $K_d(490)$. u_{mc} and u_{mk} are the model fitting uncertainty for chl_a and $K_d(490)$, calculated by multiplying the mean relative uncertainty by chl_a or $K_d(490)$. A mean relative uncertainty of 13% was derived by averaging the relative difference between model-

fitted and *in situ* chl_a over all the data points that are used to fit the model coefficients in Hu et al. (2019). A mean relative uncertainty of 10% was used for $K_d(490)$ in this study.

4 Results

4.1 Effect of the fitting on $u(R_{rs}(\lambda_i), R_{rs}(\lambda))$

Figure 1 shows an example using a MODIS Aqua scene of the North Atlantic Ocean adjacent to the U.S. East Coast. In the example, we compare $u^2(R_{rs}(412))$ and $u(R_{rs}(443), R_{rs}(547))$ calculated from the derivative approach and that derived from third-degree polynomial fitting. Values are higher at the edge than at the center of the swath. These higher values result from the high TOA signal due to the longer path length, which leads to higher sensor and forward model uncertainty. As a blue-to-green ratio of $R_{rs}(\lambda)$ is used to calculate chl_a through the traditional ocean chlorophyll-type (OCx) algorithm (O'Reilly and Werdell, 2019), we chose to show the error covariance between $R_{rs}(\lambda_i)$ at the blue and green bands to evaluate the effect of modeling $u(R_{rs}(\lambda_i), R_{rs}(\lambda))$ with the polynomial. Like Figures 1, 2 shows an example using OCI data



over the same geographic domain. We can see from Figures 1, 2 that $u(R_{rs}(\lambda_j), R_{rs}(\lambda_i))$ derived from the fitting matches well with that derived from the derivative approach, with the ratio in the range of 0.95–1.05. They show a similar spatial variation pattern.

Figure 3 further shows the comparison between $u(R_{rs}(\lambda_j), R_{rs}(\lambda_i))$ calculated from the derivative approach and that derived from third-degree polynomial fitting at the global scale. $u(R_{rs}(\lambda_j), R_{rs}(\lambda_i))$ is relatively higher in coastal and high-latitude regions. Like Figure 2, the third-degree polynomial fitted $u(R_{rs}(\lambda_j), R_{rs}(\lambda_i))$ matches very well with the original one, showing a similar spatial pattern, with mean ratios (standard deviations) of 0.989 (0.002), 0.99

(0.015), and 0.83(1.866) for $u(R_{rs}(412))$, $u(R_{rs}(443), R_{rs}(547))$, and $u(R_{rs}(443), R_{rs}(667))$, respectively. While the distribution is broader for $u(R_{rs}(443), R_{rs}(667))$, the absolute values are small, so this is a ratio of very small numbers.

While Figures 1–3 show that fitted $u(R_{rs}(\lambda_j), R_{rs}(\lambda_i))$ has a similar spatial pattern as the original fully computed one, Figure 4 shows a time series $u(R_{rs}(\lambda_j), R_{rs}(\lambda_i))$ derived from the fitting compared with the original over MOBY during 2002–2010, which shows consistent temporal variations, although $u(R_{rs}(443), R_{rs}(667))$ exhibits some bias. Considering the small value, such bias should not affect the uncertainty in the downstream products too much.

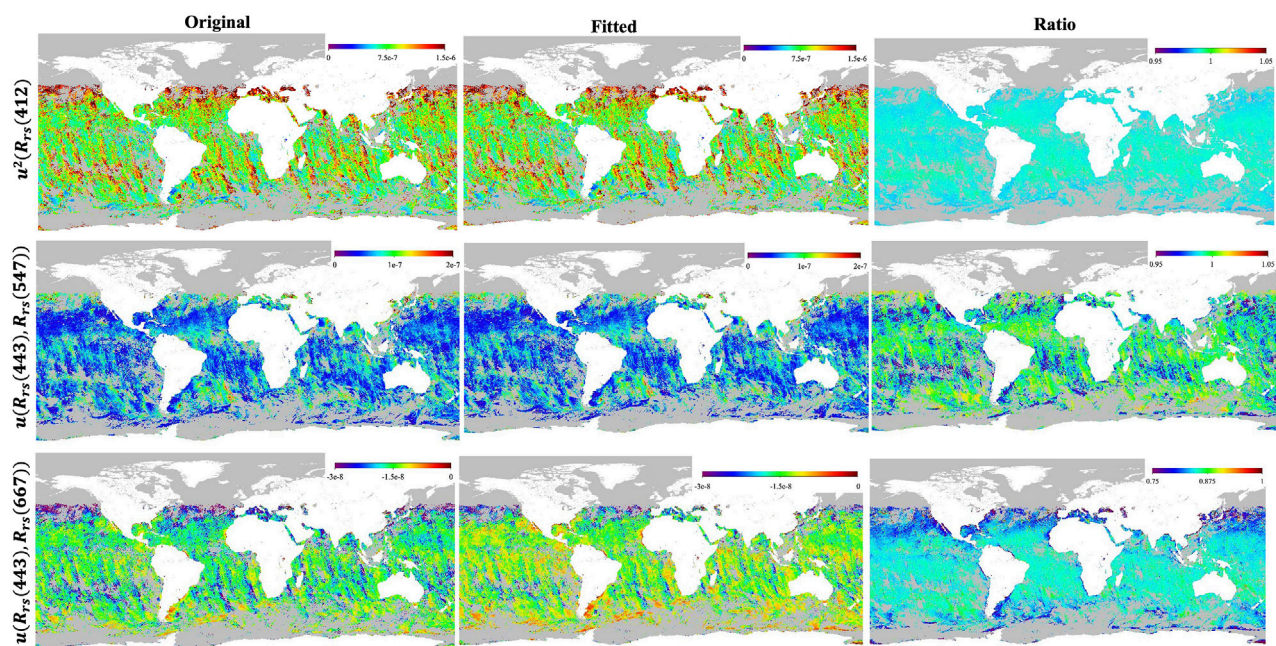


FIGURE 3
As Figure 1, but at a global scale using MODIS data during 3–10 December 2019.

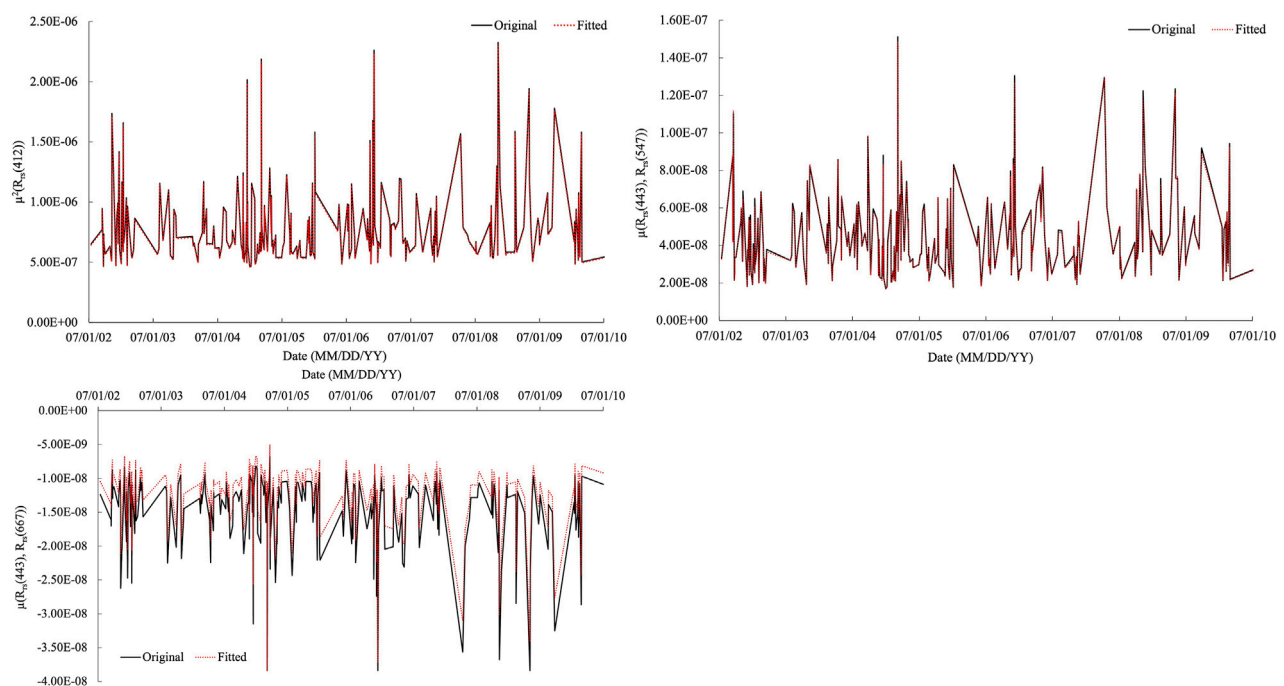


FIGURE 4
 $u(R_{rs}(\lambda_i))$, $R_{rs}(\lambda_i)$ derived from the polynomial fitting compared with that derived from the derivative approach based on MODIS data over MOBY during 2002–2010.

4.2 Fitting of $u(R_{rs}(\lambda_i), R_{rs}(\lambda))$ as a function of wavelength

Figure 5 shows two examples of the fitting of $u(R_{rs}(412), R_{rs}(\lambda))$ and $u(R_{rs}(531), R_{rs}(\lambda))$ versus wavelength over two pixels (denoted

by A and B in Figure 1, with pixel A over coastal waters near the Chesapeake Bay and pixel B over open ocean waters of the North Atlantic Ocean. Like Figures 5, 6 shows examples of the fitting of $u(R_{rs}(413), R_{rs}(\lambda))$ and $u(R_{rs}(530), R_{rs}(\lambda))$ for OCI over two pixels (denoted by A and B in Figure 2). The high R-squared value (>0.99)

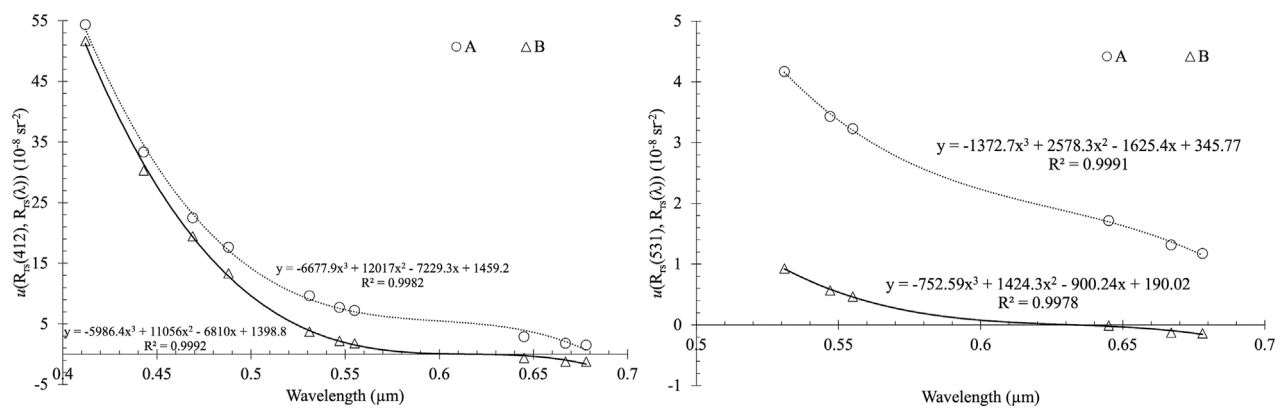


FIGURE 5
Third-degree polynomial fitting of (A) $u(R_{rs}(412), R_{rs}(\lambda))$ and (B) $u(R_{rs}(531), R_{rs}(\lambda))$ versus wavelength over two pixels denoted by A and B in Figure 1. R^2 is the R-squared value for the fitting.

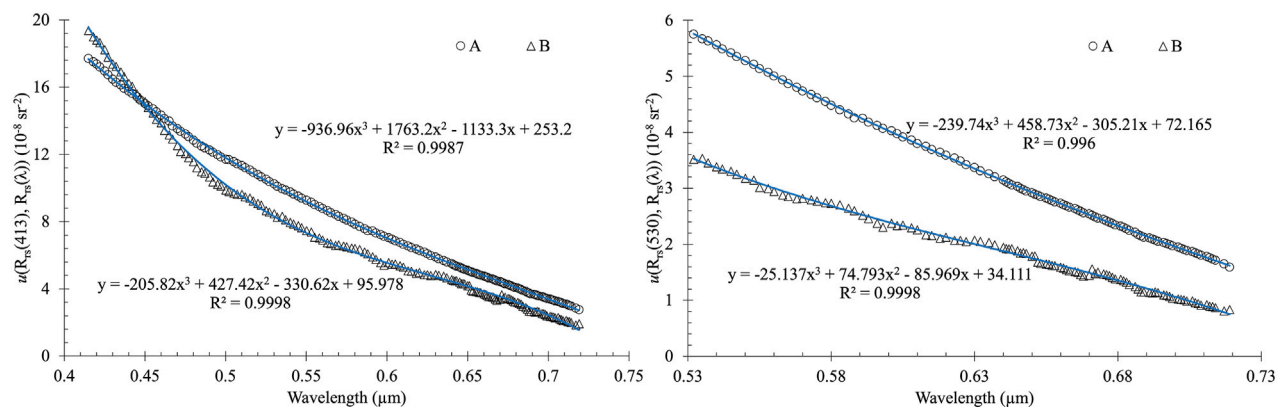


FIGURE 6
As Figure 5, but for OCI data over two pixels denoted by A (coastal) and B (open ocean) in Figure 2.

indicates the reliability of the fitting. Note the comparatively higher error covariance over coastal waters than over clear waters across visible wavelengths.

4.3 Effect on uncertainty in chl_a and $K_d(490)$

$R_{rs}(\lambda)$ is fundamental to deriving a host of bio-optical and biogeochemical properties of the water column. Taking chl_a and $K_d(490)$ as examples, the effect of using the polynomial approximation for error covariance in $R_{rs}(\lambda)$ is evaluated on $u^2(chl_a)$ and $u^2(K_d(490))$. Specifically, $u^2(chl_a)$ and $u^2(K_d(490))$ calculated using $u(R_{rs}(\lambda_j), R_{rs}(\lambda_i))$ derived from the derivative approach are compared with those calculated using $u(R_{rs}(\lambda_j), R_{rs}(\lambda_i))$ derived from polynomial fitting. We can see from Figure 7 that the relative uncertainty in chl_a and $K_d(490)$, denoted by $\delta(chl_a)$ ($u(chl_a) \times 100/chl_a$) and $\delta(K_d(490))$ ($u(K_d(490)) \times 100/K_d(490)$), calculated using original and polynomial fitted error covariance show similar spatial variation, with the absolute difference smaller than 0.5%.

5 Conclusion

We assessed how well a polynomial fitting can reproduce the full spectral error covariance in R_{rs} from MODIS and OCI data, with the goal of decreasing the Level-2 file size and therefore facilitating efficient delivery of those products. We found that a third-degree polynomial works well for data processed using the MSEPS algorithm, which is the standard algorithm used in R_{rs} retrieval for all global multispectral ocean color missions processed by NASA. We first evaluated the fitting by comparing the values numerically to the original error covariance source data, finding ratio close to 1 (typically from 0.95 to 1.05). We also found that using the fitted approximation of the covariance matrix to estimate the uncertainty of the downstream products chl_a and $K_d(490)$ led to only a small absolute difference of the relative uncertainty that was less than 0.5%. Thus, we conclude that the accuracy of the approximation appears sufficient for most data user purposes.

By including the coefficients derived from polynomial fitting rather than the full $\sum R_{rs}$, a standard Level-2 file will decrease from ~231 MB to ~165 MB for a 5-min granule shown in Figure 1. The

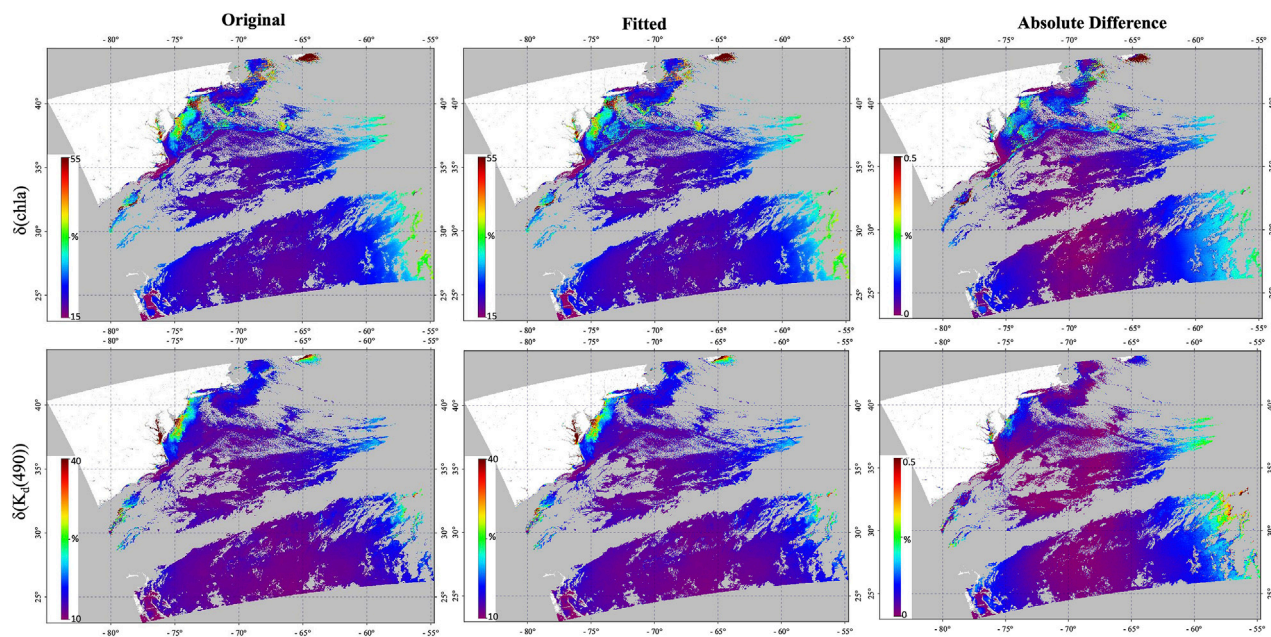


FIGURE 7
 $\delta(chl_a)$ and $\delta(K_d(490))$ calculated using the original and the polynomial fitted error covariance in $R_{rs}(\lambda)$ retrieved from MODIS data shown in Figure 1. The absolute difference is between $\delta(chl_a)$ and $\delta(K_d(490))$ calculated in those two ways.

standard Level-2 file contains τ_a (869), Ångström exponent, $R_{rs}(\lambda)$ at wavelengths from 412–678 nm, chl_a , $K_d(490)$, particulate inorganic carbon (pic), particulate organic carbon (poc), instantaneous photosynthetically available radiation (ipar), normalized fluorescence line height (nflh), and photosynthetically available radiation (par). Taking the OCI 5-min granule in Figure 2 as an example, the Level-2 file would be ~60 GB, including full $\sum R_{rs}$, which decreases to ~1.7 GB by including the polynomial fitting coefficients.

Another benefit of using the polynomial model to estimate $R_{rs}(\lambda)$ error covariance terms comes into play during operational data processing. As highlighted in this study, the full set of sensor bands is not always used in bio-optical algorithms (e.g., two for $K_d(490)$ and four for chl_a). The approach described can thus facilitate dynamic computation of only the $\sum R_{rs}$ terms required by an algorithm, thereby reducing system memory overheads and improving computational efficiency.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

MZ: Conceptualization, Data curation, Investigation, Validation, Writing – original draft, Writing – review and editing. AI: Writing – review and editing. BF: Conceptualization, Funding acquisition, Supervision, Writing – review and editing. AS: Writing – review and editing. PW: Funding acquisition, Writing – review and editing. LM: Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. NASA Terra and Aqua Senior Review for MODIS algorithm maintenance and the NASA PACE Project.

Acknowledgments

We are grateful to NASA's Ocean Biology Distributed Active Archive Center for making available all the *in situ* data and MODIS, OCI data used in our analysis.

Conflict of interest

Author MZ was employed by Science Application International Corp. Author LM was employed by Go2Q Pty Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Ahmad, Z., and Franz, B. A. (2015). Ocean color retrieval using multiple-scattering epsilon values. *Int. Ocean Color Sci. Meet.*
- Cael, B. B., Bisson, K., Boss, E., Dutkiewicz, S., and Henson, S. (2023). Global climate-change trends detected in indicators of ocean ecology. *Nature* 619, 551–554. doi:10.1038/s41586-023-06321-z
- Clark, D., Murphy, M., Yarbrough, M., Feinholz, M., Flora, S., Broenkow, W., et al. (2003). MOBY, A radiometric buoy for performance monitoring and vicarious calibration of satellite ocean color sensors: measurement and data analysis protocols. *SeaWiFS Postlaunch Tech. Rep. Ser.* Available online at: <https://www.nist.gov/publications/moby-radiometric-buoy-performance-monitoring-and-vicarious-calibration-satellite-ocean> (Accessed December 5, 2024).
- Coddington, O. M., Richard, E. C., Harber, D., Pilewskie, P., Woods, T. N., Chance, K., et al. (2021). The TSIS-1 hybrid solar reference spectrum. *Geophys. Res. Lett.* 48, e2020GL091709. doi:10.1029/2020GL091709
- Cox, C., and Munk, W. (1954). Measurement of the roughness of the Sea surface from photographs of the Sun's glitter. *J. Opt. Soc. Am.* 44, 838–850. doi:10.1364/JOSA.44.000838
- Dutkiewicz, S., Hickman, A. E., Jahn, O., Henson, S., Beaulieu, C., and Monier, E. (2019). Ocean colour signature of climate change. *Nat. Commun.* 10, 578. doi:10.1038/s41467-019-08457-x
- Fox, N. (2010). A guide to expression of uncertainty of measurements. *GEO*. Available online at: https://qa4eo.org/docs/QA4EO-QAEO-GEN-DQK-006_v4.0.pdf (Accessed September 10, 2025).
- Hu, C., Feng, L., Lee, Z., Franz, B. A., Bailey, S. W., Werdell, P. J., et al. (2019). Improving satellite global chlorophyll a data products through Algorithm Refinement and data recovery. *J. Geophys. Res. Oceans* 124, 1524–1543. doi:10.1029/2019JC014941
- IOCCG (2019). "Uncertainties in Ocean colour remote sensing," in *International Ocean Colour Coordinating Group*. Dartmouth, Canada. Available online at: <https://ioccg.org/wp-content/uploads/2020/01/ioccg-report-18-uncertainties-rr.pdf> (Accessed September 10, 2025).
- Jackson, T., Sathyendranath, S., and Mélin, F. (2017). An improved optical classification scheme for the Ocean Colour Essential climate Variable and its applications. *Remote Sens. Environ.* 203, 152–161. doi:10.1016/j.rse.2017.03.036
- JCGM (2011). Evaluation of measurement data — supplement 2 to the "expression of uncertainty in measurement" -Extension to any number of output quantities.
- Lamquin, N., Mangin, A., Mazeran, C., Bourg, B., Bruniquel, V., and D'Andon, O. F. (2013). Pixel-by-pixel uncertainty propagation in OLCI clear water branch. Available online at: https://user.eumetsat.int/s3/eup-strap-media/pdf_s3_atbd_olci_l2_un_e9ee7d9041.pdf (Accessed September 11, 2025).
- McKinna, L. I. W., Cetinić, I., Chase, A. P., and Werdell, P. J. (2019). Approach for propagating radiometric data uncertainties through NASA Ocean color algorithms. *Front. Earth Sci.* 7, 176. doi:10.3389/feart.2019.00176
- Morel, A., Antoine, D., and Gentili, B. (2002). Bidirectional reflectance of oceanic waters: accounting for Raman emission and varying particle scattering phase function. *Appl. Opt.* 41, 6289–6306. doi:10.1364/AO.41.006289
- NASA (2023). Diffuse attenuation coefficient for downwelling irradiance at 490 nm (Kd). Available online at: <https://oceancolor.gsfc.nasa.gov/resources/atbd/kd/>.
- O'Reilly, J. E., and Werdell, P. J. (2019). Chlorophyll algorithms for ocean color sensors - OC4, OC5 and OC6. *Remote Sens. Environ.* 229, 32–47. doi:10.1016/j.rse.2019.04.021
- Stramska, M., and Petelski, T. (2003). Observations of oceanic whitecaps in the north polar waters of the Atlantic. *J. Geophys. Res. Oceans* 108, 2002JC001321–a. doi:10.1029/2002JC001321
- Thuillier, G., Floyd, L., Woods, T. N., Cebula, R., Hilsenrath, E., Hersé, M., et al. (2004). Solar irradiance reference spectra for two solar active levels. *Adv. Space Res.* 34, 256–261. doi:10.1016/j.asr.2002.12.004
- Wang, M., and Bailey, S. W. (2001). Correction of sun glint contamination on the SeaWiFS ocean and atmosphere products. *Appl. Opt.* 40, 4790–4798. doi:10.1364/AO.40.004790
- Zhang, M., Ibrahim, A., Franz, B. A., Ahmad, Z., and Sayer, A. M. (2022). Estimating pixel-level uncertainty in ocean color retrievals from MODIS. *Opt. Express* 30, 31415–31438. doi:10.1364/OE.460735
- Zhang, M., Ibrahim, A., Franz, B. A., Sayer, A. M., Werdell, P. J., and McKinna, L. I. (2024). Spectral correlation in MODIS water-leaving reflectance retrieval uncertainty. *Opt. Express* 32, 2490–2506. doi:10.1364/OE.502561