AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Oceans

RESEARCH ARTICLE

10.1002/2016JC012126

A system to measure the data quality of spectral remote-sensing reflectance of aquatic environments

Jianwei Wei[1], Zhongping Lee[1], and Shaoling Shang[2]

[1]Optical Oceanography Laboratory, School for the Environment, University of Massachusetts Boston, Boston, Massachusetts, USA, [2]State Key Laboratory of Marine Environmental Science, Xiamen University, Xiamen, China

Key Points:
- A QA system is developed for spectral remote-sensing reflectance
- The system consists of a reference and a score metric
- It is applicable to both remotely sensed and in situ ocean color data

Correspondence to:
J. Wei,
jianwei.wei@umb.edu

Citation:
Wei, J., Z. Lee, and S. Shang (2016), A system to measure the data quality of spectral remote-sensing reflectance of aquatic environments, *J. Geophys. Res. Oceans*, *121*, 8189–8207, doi:10.1002/2016JC012126.

Received 5 JUL 2016
Accepted 19 OCT 2016
Accepted article online 25 OCT 2016
Published online 11 NOV 2016

Abstract Spectral remote-sensing reflectance (R_{rs}, sr^{-1}) is the key for ocean color retrieval of water bio-optical properties. Since R_{rs} from in situ and satellite systems are subject to errors or artifacts, assessment of the quality of R_{rs} data is critical. From a large collection of high quality in situ hyperspectral R_{rs} data sets, we developed a novel quality assurance (QA) system that can be used to objectively evaluate the quality of an individual R_{rs} spectrum. This QA scheme consists of a unique R_{rs} spectral reference and a score metric. The reference system includes R_{rs} spectra of 23 optical water types ranging from purple blue to yellow waters, with an upper and a lower bound defined for each water type. The scoring system is to compare any target R_{rs} spectrum with the reference and a score between 0 and 1 will be assigned to the target spectrum, with 1 for perfect R_{rs} spectrum and 0 for unusable R_{rs} spectrum. The effectiveness of this QA system is evaluated with both synthetic and in situ R_{rs} spectra and it is found to be robust. Further testing is performed with the NOMAD data set as well as with satellite R_{rs} over coastal and oceanic waters, where questionable or likely erroneous R_{rs} spectra are shown to be well identifiable with this QA system. Our results suggest that applications of this QA system to in situ data sets can improve the development and validation of bio-optical algorithms and its application to ocean color satellite data can improve the short-term and long-term products by objectively excluding questionable R_{rs} data.

1. Introduction

Remote-sensing reflectance (R_{rs}, units: sr^{-1}) is defined as the ratio of water-leaving radiance (L_w, units: μW cm^{-2} sr^{-1} nm^{-1}) to downwelling irradiance just above the surface (E_s, units: μW cm^{-2} nm^{-1}). R_{rs} is a critical optical property for deriving water's optical and biogeochemical properties that include chlorophyll *a* concentration [*O'Reilly et al*., 1998], colored dissolved organic material (CDOM) absorption coefficient and particulate backscattering coefficient [*IOCCG*, 2006; *Lee et al*., 2002; *Mannino et al*., 2008; *Wei and Lee*, 2015], etc. Reliable retrieval and appropriate interpretation of these properties, to the first order, demand accurate R_{rs} data.

R_{rs} cannot be directly measured in the field or obtained remotely, rather derived from two properties (L_w and E_s) obtained independently. Errors in R_{rs} data from field measurements can be related to instrument platforms, strategies of deployment, data processing as well as the ambient environments [*Bailey et al*., 2008; *Hooker et al*., 2002; *Mueller et al*., 2003; *Toole et al*., 2000; *Zibordi et al*., 2002]. The in-water approach measures the in-water downwelling irradiance (E_d, units: μW cm^{-2} nm^{-1}) and upwelling radiance (L_u, units: μW cm^{-2} sr^{-1} nm^{-1}) separately, and then the in-water radiance is propagated to right above the water surface to obtain the water-leaving radiance. The known sources of measurement uncertainty could originate from instrument calibrations [*Bailey et al*., 2008; *Wei et al*., 2012], instrument self-shading effect [*Gordon and Ding*, 1992], wave-focusing effects [*Wei et al*., 2014; *Zibordi et al*., 2004], propagation of $L_u(z)$ to L_w [*Wei et al*., 2015], and reflection or shading noise from the ship hull. On the other hand, R_{rs} data from the above-water approach are likely subject to uncertainties due to sea surface reflection [*Lee et al*., 2010a; *Mobley*, 1999], bidirectional reflectance distribution function (BRDF) [*Lee et al*., 2011; *Morel and Gentili*, 1996; *Voss and Morel*, 2005], sky cloudiness, sun glints, and white caps [*Gordon and Wang*, 1994a]. The skylight-blocked approach (SBA) [*Lee et al*., 2013; *Tanaka and Sasaki*, 2006] can directly measure L_w (then R_{rs}), but to some degree it is also subject to self-shading error. In addition to these systematic errors, there are likely unidentified errors and uncertainties or artifacts during field observations resulted from uncontrollable field environment.

Obtaining accurate R_{rs} spectra from space is even more challenging. The quality of R_{rs} spectra derived from a satellite ocean color sensor is significantly impacted by the efficacy of atmospheric corrections [*Gordon and Wang*, 1994b; *IOCCG*, 2010]. For example, at the top of atmosphere (TOA) about 90% of the total radiance received by a satellite sensor is contributed from the atmosphere. Small errors in the atmospheric correction will result in large errors in the derived R_{rs}. In coastal regions, atmospheric correction of satellite ocean color images is a daunting task because of the complex atmospheric and water properties [*IOCCG*, 2010].

Methods to quality control (QC) R_{rs} data are diverse and system dependent [*Bailey and Werdell*, 2006; *McClain et al*., 1992; *Ruddick et al*., 2005; *Werdell and Bailey*, 2005; *Zibordi et al*., 2009b]. The in situ R_{rs} measurements usually go through certain QC steps during data conversion (e.g., absolute calibration) and subsequent data postprocessing (e.g., tilt filtering and data binning); these processes are often embedded in the processing software provided by the instrument manufacturers. The difficulty is, however, that environmental disturbance to the measurements and artifacts cannot be fully accounted for in these postprocessing steps. When compiling the NASA bio-Optical Marine Algorithm Dataset (NOMAD), *Werdell and Bailey* [2005] defined and utilized rejection criteria to eliminate abnormal and spurious data. For instance, the observed spectral surface irradiances were compared with modeled clear sky values, and those stations were considered questionable and discarded when the in situ value exceeded the modeled value by more than 33%. *Ruddick et al*. [2005] discussed the QC of above-water R_{rs} measurements at near-infrared (NIR) bands based on the spectral similarity in the NIR domain. Their method is more applicable to coastal turbid waters and does not provide further information on the quality of R_{rs} spectra in the visible domain. In processing the AERONET-OC data, *Zibordi et al*. [2009b] proposed a three-level scheme for the normalized water-leaving radiance (L_{wn}, units: $\mu W\ cm^{-2}\ sr^{-1}\ nm^{-1}$) with thresholds including, e.g., $L_{wn}(\lambda) > -0.01\ \mu W\ cm^{-2}\ nm^{-1}\ sr^{-1}$ to ensure an exclusion of large negative values; $L_{wn}(412) < L_{wn}(443)$ as commonly met in coastal waters and $L_{wn}(1020) < 0.1\ \mu W\ cm^{-2}\ nm^{-1}\ sr^{-1}$ to exclude reflecting obstacles along the optical path between the instrument and the water surface; and then inspection of the spectral consistency using statistical methods [*D'Alimonte and Zibordi*, 2006]. A QC procedure was also developed for satellite ocean color data [*McClain et al*., 1992]. This method consists of a system of processing flags for Level-2 products. A total of 31 flags are currently employed and applied pixel by pixel, including ATMFAIL, STRAYLIGHT, HIGLINT, COASTZ, etc. [*Robinson et al*., 2003]. These flags use predetermined thresholds. For example, the glint flag is activated when the normalized sun glint reflectance is found exceeding 0.005. Pixels with stray light contamination from adjacent, bright sources such as coasts and clouds are flagged in Level-2 products and masked in Level-3 products. Further, to validate satellite R_{rs} against in situ data, *Bailey and Werdell* [2006] introduced a scheme by considering the spatial uncertainty, coincidence determination, viewing and solar zenith angles, and statistical confidence of the mean pixel values. To reduce the R_{rs} errors from satellite sensor digitization noise, *Hu et al*. [2005] recommended the median R_{rs} values over a box of 3×3 pixels be obtained to reduce uncertainties in satellite ocean color products.

A common practice to evaluate the accuracy of the R_{rs} data obtained from remote-sensing platforms is to compare R_{rs} data with in situ measurements wavelength-by-wavelength. Statistical measures (e.g., R^2 and relative error, etc.) are often obtained as an indicator of the accuracy of the R_{rs} data obtained remotely. Such an approach characterizes the overall quality of the remotely obtained R_{rs} data and provides an indication of the systematic performance of the instrument and data-processing system for each spectral band [*Bailey and Werdell*, 2006; *Zibordi et al*., 2009a]. This way of evaluation treats the R_{rs} property similarly as those of biogeochemical properties such as the chlorophyll *a* concentration, and considers R_{rs} at each wavelength as an independent property. Such scatter-plots or linear regressions wavelength by wavelength cannot tell the quality of each R_{rs} spectrum, however. Note that for remote-sensing inversions [*IOCCG*, 2000, 2006], it seldom uses R_{rs} at one band to derive in-water properties, rather it uses multiple bands or the entire R_{rs} spectrum. Therefore, it is important and imperative to measure objectively the quality of each R_{rs} spectrum obtained from any platforms.

In this study, a novel system is developed for objective quality assurance (QA) of an individual R_{rs} spectrum. Based on in situ R_{rs} data obtained from coastal waters and clear oceanic waters, we define the domain of variability of the spectral shapes and amplitudes of R_{rs} spectra in different water types using a unique optical water type classification scheme. A score metric ranging from 0 (lowest quality) to 1 (highest quality) is further devised to measure the quality of any target R_{rs} spectrum. Examples of applications of this QA system to synthetic, in situ, and satellite data are also presented.

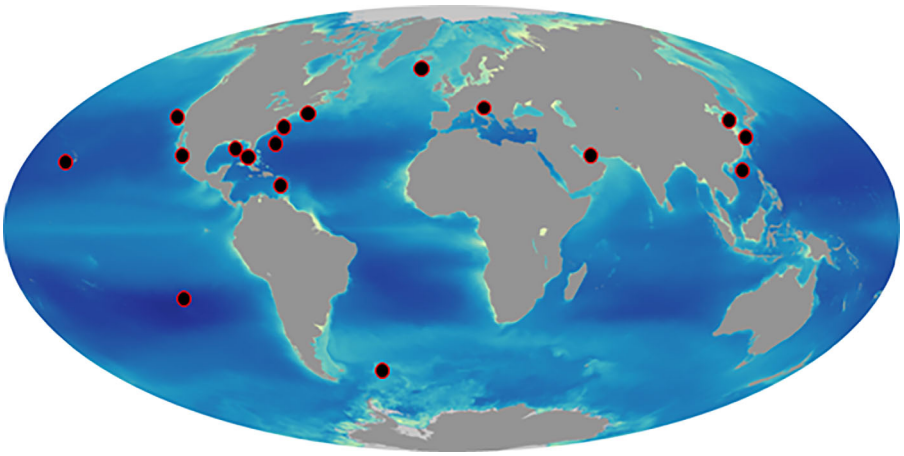

Figure 1. Locations of field observation of remote-sensing reflectance.

2. Development of the QA system

2.1. In Situ Measurements

In situ hyperspectral R_{rs} spectra were collected from a wide range of marine environments, with chlorophyll *a* concentration ([CHL], units: mg m^{-3}) varying from as low as 0.02 mg m^{-3} in the subtropical gyres to tens of mg m^{-3} in the coastal regions (Figure 1). The data were retrieved by above-water approach, SBA [*Lee et al*., 2013] and depth-profiling approach, with a wavelength range of 400–800 nm and 10 nm resolution. In total, there are 958 R_{rs} spectra adopted in subsequent analyses. The R_{rs} data from the three approaches account for about 83%, 11%, and 6%, respectively, of the whole data sets. The aggregate data of these measured hyperspectral R_{rs} spectra are illustrated in Figure 2. In this study, R_{rs} of nine visible wavelengths (412, 443, 488, 510, 531, 547, 555, 667, and 678 nm) are extracted from the hyperspectral database, and are referred to as the "reference" data hereafter. It is noted that many of these wavelengths are already incorporated in heritage and currently operational ocean color satellite sensors including the Sea-viewing Wide Field-of-view Sensor (SeaWiFS), Moderate-resolution Imaging Spectroradiometer (MODIS), Medium Resolution Imaging Spectrometer (MERIS), Visible Infrared Imaging Radiometer Suite (VIIRS), and Landsat 8.

2.1.1. R_{rs} Data From Above-Water Approach

The primary data for this R_{rs} database were obtained from the Gulf of Mexico, Adriatic Sea, South China Sea, North Pacific Gyre, Mississippi River plume, Yangtze River plume, as well as optically shallow waters in the West Florida Shelf and around Key West (Florida) (Figure 1). The range of [CHL] is from 0.02 to 80 mg m^{-3}; the range of suspended sediment matter is from the sensors' lower detection limit up to ~100 g m^{-3}; the range of bottom depth is from ~1 m to >1000 m with substrates of seagrass, sand, and coral reef. The instruments used are all calibrated commercial radiometers from ultraviolet bands to near-infrared bands including Spectrix (400–850 nm), GER (350–1050 nm), and RAMSES (320–950 nm) and have a fine pixel dispersion of ~2–3 nm with half-value-bandwidth of ~1–1.5 nm. The data-processing scheme follows the NASA ocean optics protocol [*Mueller et al*., 2003] and the spectral optimization scheme of *Lee et al.* [2010a] for refinement.

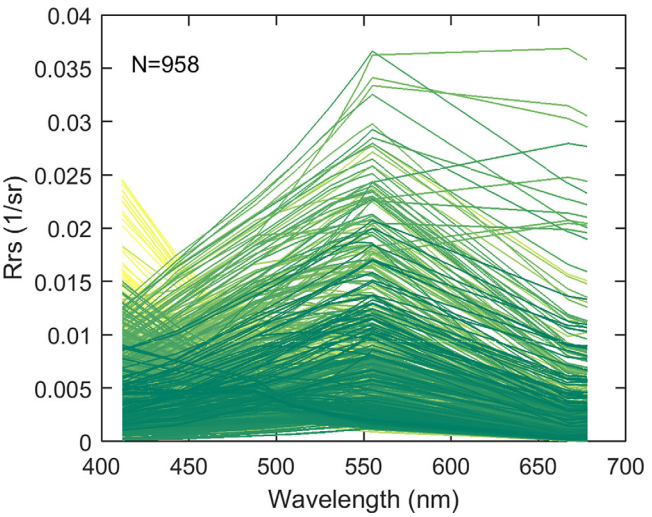

Figure 2. Spectra of remote-sensing reflectance used in this study.

2.1.2. R_{rs} Data From Floating Instrument With the SBA

Another source of hyperspectral R_{rs} was obtained from a floating system incorporating the SBA [*Lee et al*., 2013] in coastal waters including Massachusetts Bay, northern Gulf of Mexico, and the Caribbean Sea (south of Puerto Rico) (Figure 1). The SBA system, as its name implies, consists of a hyperspectral radiometric profiler (350–800 nm, Satlantic Inc., Halifax, Canada) and a skylight-blocking apparatus attached to the radiance sensor. It is thus able to directly measure L_w in a continuous manner and provides a time series of R_{rs} and allows a reliable evaluation of the R_{rs} uncertainty. The time series measurements of R_{rs} spectra were further inspected based on the R_{rs} data at a single band of 698 nm following the procedure of *Wei et al*. [2015]. Basically, a density function was first determined for the $R_{rs}(698)$ time series, and those deviating by $\pm 30\%$ from the first mode of the density function were then removed. This procedure can effectively eliminate those data potentially contaminated by sea surface reflected light when the cone is suspended above the surface and/or immersion of the radiance sensor head within the water at high seas. After filtering, the median R_{rs} spectrum is derived for the remaining R_{rs} data and considered as true R_{rs} for that water body. For the waters encountered and the spectral range (400–700 nm) considered, the instrument self-shading error is found quite small (less than 5%) based on *Gordon and Ding* [1992].

2.1.3. R_{rs} Data From Hyperspectral Profiler

A third source of hyperspectral R_{rs} data was retrieved from the North Pacific Gyre and the South Pacific Gyre using a free-fall hyperspectral profiler (HyperPro, Satlantic Inc., Halifax, Canada). The data-processing procedure follows the NASA ocean optics protocols [*Mueller et al*., 2003]. Specifically, the depth profiles of upwelling radiance are propagated to right above the water surface to derive the water-leaving radiance L_w, using the determined diffuse attenuation coefficient for upwelling radiance, and the upwelling radiance transmittance [*Austin*, 1974; *Wei et al*., 2015]. Then the mean remote-sensing reflectance is derived as the ratio of L_w to E_s. Because of extreme clarity, the self-shading effect of this system for the 400–700 nm range is negligible in these waters.

2.2. Clustering of Optical Water Types

The reference R_{rs} spectra were first normalized by their respective root of sum of squares (RSS),

$$nR_{rs}(\lambda)=\frac{R_{rs}(\lambda)}{\left[\sum_{i=1}^{N} R_{rs}(\lambda_i)^2\right]^{1/2}} \tag{1}$$

where the index N represents the total number of wavelengths, varying from 1 to 9 and λ_i corresponds to the wavelengths of 412, 443, 488, 510, 531, 547, 555, 667, and 678 nm. The nR_{rs} spectra vary over the range between 0 and 1, while it retains the "shapes" pertaining to the original R_{rs} spectra, i.e., the band ratios of $nR_{rs}(\lambda)$ remain the same as $R_{rs}(\lambda)$.

The number of data clusters k was evaluated using the gap method [*Tibshirani et al*., 2001]. The gap value is defined as:

$$GAP_n(k)=E_n^*[\log(W_k)]-\log(W_k) \tag{2}$$

where n is the sample size, k is the number of clusters being evaluated, and W_k is the pooled within-cluster dispersion measurement, with

$$W_k=\sum_{r=1}^{k}\frac{1}{2n_r}D_r \tag{3}$$

where n_r is the number of data points in cluster r, and D_r is the sum of the pair-wise distances for all points in cluster r. The expected value $E_n^*[\log(W_k)]$ is determined by Monte Carlo sampling from a reference distribution, and $\log(W_k)$ is computed from the sample data. According to the gap method, the optimum cluster number of the nR_{rs} data is determined as 23. Interestingly, this number is nearly the same as that of Forel-Ule water type classes developed 100 years ago [*Arnone et al*., 2004].

The unsupervised method, *K*-means clustering technique, was further used to group the nR_{rs} spectra. *K*-means clustering, or Lloyd's algorithm [*Lloyd*, 1982], is an iterative, data-partitioning algorithm that assigns n observations to exactly one of k clusters defined by the centroids, where k is chosen before the algorithm

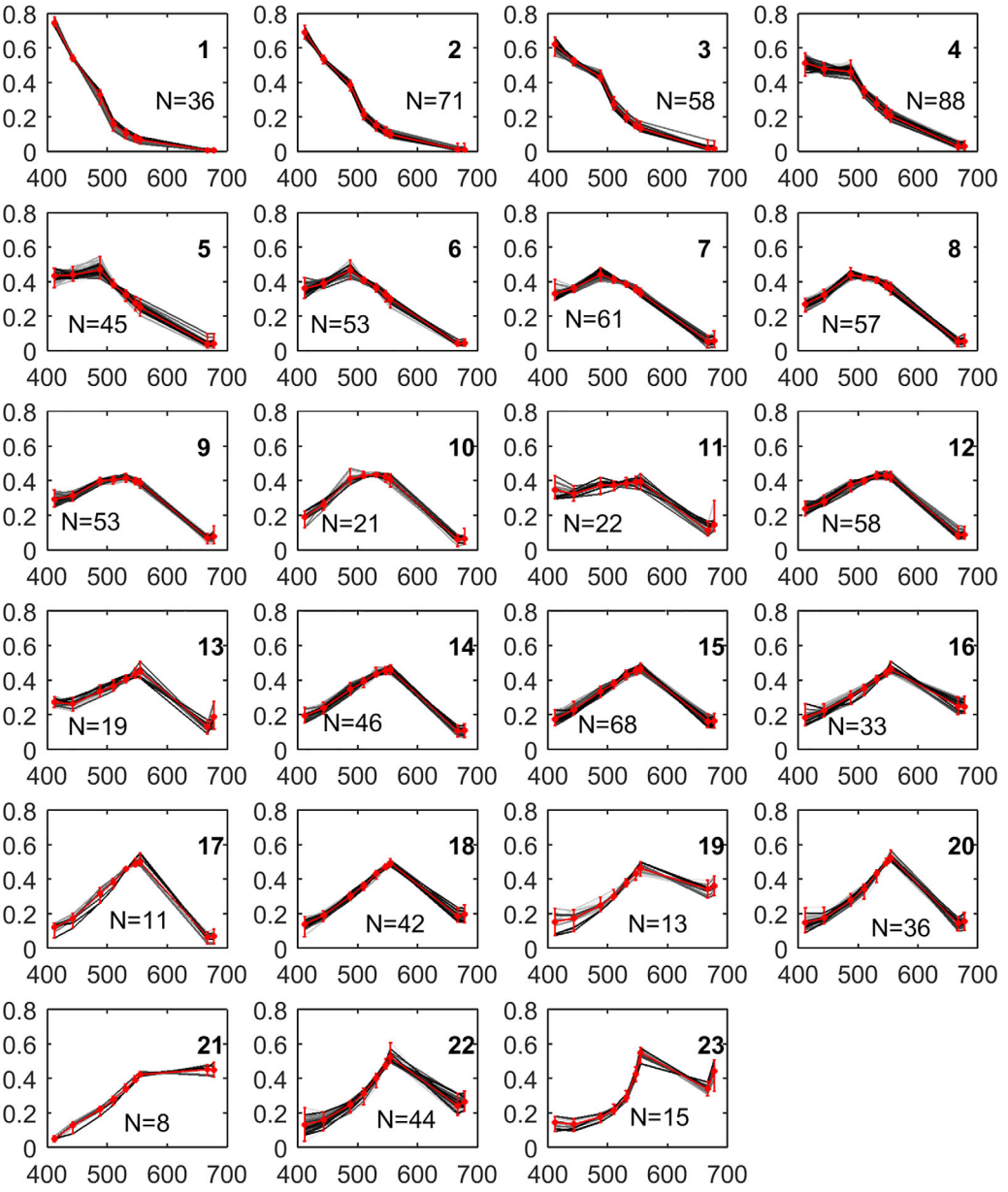

Figure 3. Clustering of 23 optical water types. The gray-black curves represent the normalized R_{rs} spectra classified into each water type. The mean R_{rs} spectra are highlight in red, with the error bars characterizing the maximum and minimum values. And in each water type, the total numbers of available R_{rs} spectra are also given.

starts (as shown above, $k = 23$). We note that the cosine distance was used in the clustering, which is one minus the cosine of the angle between the nR_{rs} spectrum x (a reference) and c (a target),

$$d(x,c) = 1 - \frac{xc'}{\sqrt{(xx')(cc')}} \tag{4}$$

Each centroid is the mean of the nR_{rs} spectra in that cluster, after normalizing those points to unit Euclidean length. The error function to be minimized is the total within-cluster sum of squares.

The nR_{rs} spectra as clustered into the 23 optical water types are illustrated in Figure 3, with the centroids of nR_{rs} spectra highlighted for each water type. Within each water type, the nR_{rs} spectra are very similar to each other and tightly distributed about the centroids. The centroid nR_{rs} spectra of the 23 water types are illustrated in Figure 4 (with values tabulated in Table 1). These "mean" reference spectra represent a large

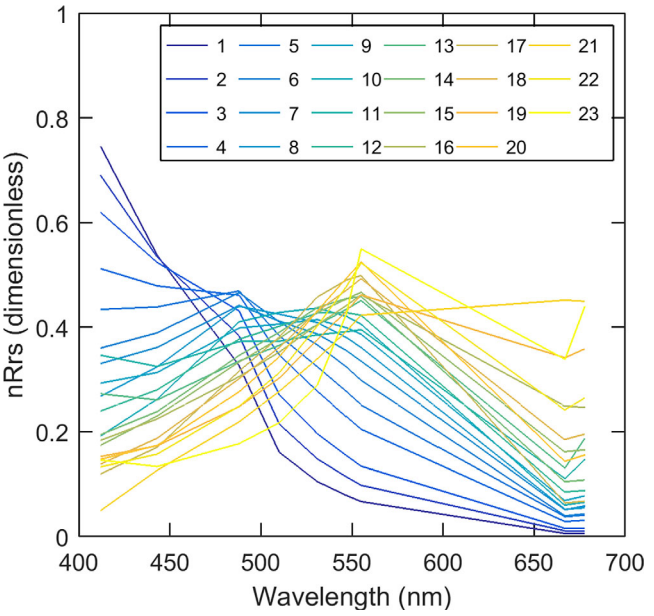

Figure 4. Mean normalized remote-sensing reflectance spectra characteristics of 23 water types.

range of waters with [CHL] varying from 0.02 mg m^{-3} to tens mg m^{-3}. Corresponding to Figure 3, the upper limits and lower limits of all nR_{rs} spectra at discrete bands have been identified for each water type and are tabulated in appendix (see Appendix A Tables A1 and A2). The identified nR_{rs} mean spectra, together with the upper boundary spectra and lower boundary spectra form the core of the QA system and will be used in the following sections to perform the QA of an individual R_{rs} spectrum.

It is necessary to emphasize that the above clustering procedure is focused on the spectral shapes, and is distinct from other exercises of optical water type clustering which often used the Euclidean distance of R_{rs} spectra [*Le et al.*, 2011; *Mélin and Vantrepotte*, 2015; *Moore et al.*, 2009, 2001]. The Euclidean distance, including transformed and weighted forms such as Mahalanobis and likelihood distances, is inherently insensitive to the shapes of the spectral pattern [*Sohn et al.*, 1999].

2.3. Water Type Matchup and Comparison

To evaluate or quantitatively measure the data quality of a target reflectance spectrum, R^*_{rs}, a four-step procedure is developed:

Step 1 is to match up $R^*_{rs}(\lambda')$ with $nR_{rs}(\lambda)$ with regard to the wavelengths. If R^*_{rs} has more spectral bands than that of nR_{rs}, we will only choose the same wavelengths with nR_{rs} for further analysis. If R^*_{rs} has fewer

Table 1. Mean Normalized Remote-Sensing Reflectance Spectra (nR_{rs}) for the 23 Optical Water Types

	Wavelength (nm)									
Water Type	412	443	488	510	531	547	555	667	678	CHL (mg m^{-3})[a]
1	0.738	0.535	0.335	0.169	0.112	0.084	0.072	0.007	0.007	0.06
2	0.677	0.534	0.394	0.225	0.156	0.120	0.104	0.011	0.010	0.10
3	0.608	0.521	0.436	0.280	0.204	0.161	0.140	0.016	0.017	0.16
4	0.510	0.478	0.462	0.348	0.279	0.230	0.206	0.029	0.031	0.35
5	0.430	0.436	0.472	0.386	0.326	0.278	0.253	0.038	0.041	0.52
6	0.363	0.387	0.458	0.408	0.368	0.328	0.304	0.042	0.047	0.76
7	0.309	0.355	0.451	0.419	0.392	0.356	0.335	0.048	0.052	0.94
8	0.276	0.315	0.415	0.415	0.414	0.394	0.378	0.062	0.067	1.51
9	0.349	0.335	0.391	0.386	0.387	0.382	0.378	0.090	0.118	1.63
10	0.228	0.275	0.383	0.407	0.430	0.427	0.420	0.079	0.082	2.35
11	0.291	0.276	0.342	0.367	0.401	0.424	0.437	0.129	0.181	3.28
12	0.187	0.241	0.342	0.382	0.427	0.450	0.461	0.147	0.151	3.80
13	0.173	0.220	0.342	0.393	0.447	0.462	0.464	0.093	0.096	4.61
14	0.188	0.235	0.319	0.363	0.412	0.445	0.463	0.215	0.214	4.77
15	0.143	0.191	0.306	0.365	0.434	0.472	0.492	0.170	0.180	6.57
16	0.181	0.200	0.261	0.307	0.365	0.410	0.437	0.359	0.374	7.25
17	0.174	0.203	0.283	0.334	0.399	0.446	0.472	0.272	0.280	7.07
18	0.142	0.169	0.279	0.349	0.439	0.498	0.525	0.121	0.131	10.41
19	0.050	0.126	0.219	0.277	0.340	0.392	0.423	0.452	0.449	10.81
20	0.117	0.153	0.258	0.324	0.412	0.477	0.515	0.243	0.259	12.28
21	0.163	0.175	0.249	0.308	0.400	0.490	0.544	0.190	0.217	16.08
22	0.111	0.135	0.226	0.292	0.385	0.463	0.511	0.310	0.329	17.57
23	0.145	0.133	0.176	0.215	0.286	0.423	0.548	0.341	0.449	34.59

[a]Median chlorophyll *a* concentration.

wavelengths than the nR_{rs} spectra (i.e., $N(\lambda') < 9$), a subset of $nR_{rs}(\lambda')$ and associated upper boundary spectra $nR_{rs}^{U}(\lambda')$ and lower boundary spectra $nR_{rs}^{L}(\lambda')$ will be extracted first for λ'.

Step 2 is the normalization of R_{rs}^{*} spectra following equation (1). For the case of $N(\lambda') < 9$, the new $nR_{rs}(\lambda')$ spectra will be rescaled through the normalization procedure of equation (1) so that the RSS of $nR_{rs}(\lambda')$ is equal to one. Further, the new upper and lower boundary spectra $nR_{rs}^{U}(\lambda')$ and $nR_{rs}^{L}(\lambda')$ will also be rescaled by the newly rescaled $nR_{rs}(\lambda')$ spectra as below,

$$nR_{rs}^{U}(\lambda) = \frac{nR_{rs}^{U}(\lambda)}{\left[\sum_{i=1}^{N} nR_{rs}(\lambda_i)^2\right]^{1/2}}$$
$$nR_{rs}^{L}(\lambda) = \frac{nR_{rs}^{L}(\lambda)}{\left[\sum_{i=1}^{N} nR_{rs}(\lambda_i)^2\right]^{1/2}} \quad (5)$$

Step 3 is to assign a water type to the target spectrum by comparing it with the reference nR_{rs} spectra. The spectral similarity between the target spectrum nR_{rs}^{*} and reference spectra nR_{rs} are estimated using a spectral angle mapper (SAM) [*Kruse et al.*, 1993],

$$\cos\alpha = \frac{\sum_{i=1}^{N}\left[nR_{rs}^{*}\cdot nR_{rs}\right]}{\sqrt{\sum_{i=1}^{N}\left[nR_{rs}^{*}(\lambda_i)\right]^2\sum_{i=1}^{N}\left[nR_{rs}(\lambda_i)\right]^2}} \quad (6)$$

where α is the angle formed between the reference spectrum nR_{rs} and the normalized target spectrum nR_{rs}^{*}. As a spectral classifier, SAM is able to determine the spectral similarity by treating them as vectors in a space with dimensionality equal to the number of bands, N. The water type of the target spectrum nR_{rs}^{*} is identified as the one with the largest cosine values (equivalent to the smallest angles).

Step 4 is the computation of QA scores by comparing the target spectrum nR_{rs}^{*} with the upper and lower boundaries (nR_{rs}^{U} and nR_{rs}^{L}) of the corresponding water type. The number of wavelengths where nR_{rs}^{*} falling within the boundaries is counted, and used to derive the total score (C_{tot}) for the nR_{rs}^{*} spectrum,

$$C_{tot} = \frac{C(\lambda_1) + C(\lambda_2) + \cdots + C(\lambda_N)}{N} \quad (7)$$

where C_i is the wavelength-specific score, with N the total number of wavelengths for both R_{rs}^{*} and R_{rs}^{ref}. At wavelength λ_i, for example, if $nR_{rs}^{*}(\lambda_i)$ is found beyond either the upper ($nR_{rs}^{U}(\lambda_i)$) or lower ($nR_{rs}^{L}(\lambda_i)$) boundary of nR_{rs}, a score of 0 will be assigned to this wavelength, i.e., $C(\lambda_i) = 0$; otherwise, $C(\lambda_i) = 1$. As suggested by equation (7), the total score C_{tot} will vary within the range of [0, 1]. A higher score indicates higher data quality.

To account for the measurement uncertainty and possible data-processing errors and likely insufficient data coverage, the original upper boundary and lower boundary are slightly modified by ±0.5%, $nR_{rs}^{U} = nR_{rs}^{U} \times (1 + 0.005)$ and $nR_{rs}^{L} = nR_{rs}^{L} \times (1 - 0.005)$, respectively. Note that this added range of 0.5% is one order of magnitude smaller than the projected accuracy for radiance measurement [*Hooker et al.*, 1992].

3. Evaluation of the QA System

3.1. Synthesized R_{rs} Spectra

3.1.1. Measurement-Error-Free Data

The R_{rs} spectra were synthesized with a database of inherent optical properties (IOP) presented in the International Ocean Color Coordinating Group report [*IOCCG*, 2006]. The synthetic IOP data cover a wide range of aquatic environments and have been extensively used for model development and validation [*Chen et al.*, 2015; *Salama and Stein*, 2009; *Wang et al.*, 2005; *Wei and Lee*, 2013, 2015; *Wei et al.*, 2016]. The remote-sensing reflectance was simulated by Hydrolight 5.1 [*Mobley and Sundman*, 2008]. The inelastic scattering

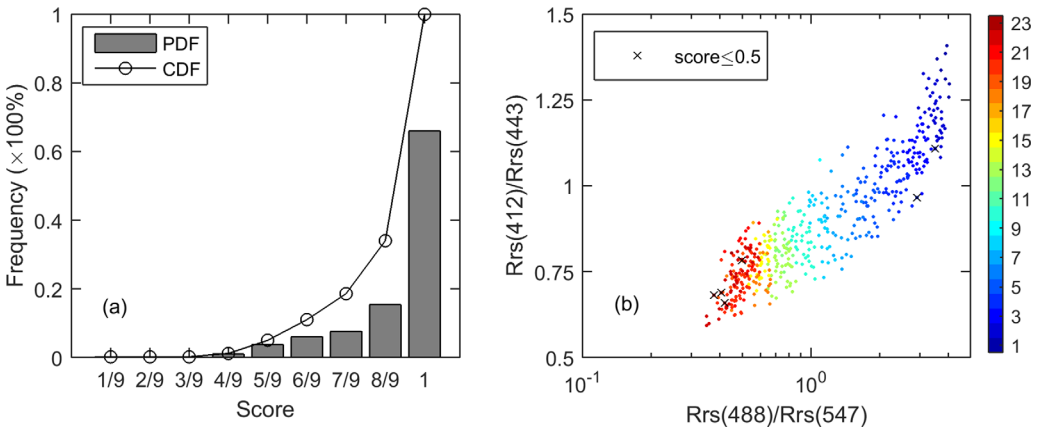

Figure 5. (a) Frequency distribution of the scores assigned to the simulated R_{rs} data. The scores are given in fractional numbers with the numerators referring to the numbers of good-quality wavebands. (b) Ratio distribution of $R_{rs}(412)/R_{rs}(443)$ against $R_{rs}(488)/R_{rs}(547)$. The corresponding 23 water types are denoted in colors and the spectra identified with low QA scores ($\leq$0.5) are denoted with cross symbols in black.

including Raman scattering and fluorescence was invoked and included in the model runs. For chlorophyll fluorescence, a fluorescence efficiency of 0.02 was adopted by default in Hydrolight. The chlorophyll *a* specific absorption spectra $a^*_{ph}(\lambda)$ were derived as the ratio of $a_{ph}(\lambda)$/[CHL]. Clear skies were assumed with the solar zenith angles at 30°. Sea surface roughness was relatively mild with wind speed of 5 m/s. The scattering phase function of particles was assumed to follow Fournier-Forand model, with changing backscattering ratios. The water absorption coefficient from 400 to 550 nm was adopted from *Lee et al*. [2015]; while the measurements by *Pope and Fry* [1997] were used for the other wavelengths. The scattering coefficients of pure seawater were provided by *Morel* [1974].

The data quality of simulated R_{rs} spectra at nine wavelengths (412, 443, 488, 510, 531, 547, 555, 667, and 678 nm) is assessed with the-above designed score metric model. As shown in Figure 5a, 65% of the spectra are identified with the highest quality with scores of 1. Another 18% of the spectra are scored as 8/9, meaning the individual R_{rs} spectrum at one wavelength has gone beyond the domain defined by our metrics. There are another 10% of the R_{rs} data found out of the range at two wavelengths. Figure 5b further characterizes the statistics of simulated R_{rs} spectra in terms of the two spectral ratios, $R_{rs}(412)/R_{rs}(443)$ and $R_{rs}(488)/R_{rs}(547)$. The IOCCG data represent a large range of waters varying from water types 1 to 23. And the relatively low QA scores (in this case, scores $\leq$0.5) are found with the spectra in extremely turbid waters and a few others in relatively clear oceanic waters. These low-score data points can be partly explained by the limited number of reference spectra used in particular water types in the model development, such as water types of 17, 19, 21, and 23 (Figure 3).

3.1.2. Error-Disturbed Data

Since the errors in R_{rs} data are hardly predictable, we only discussed two simple cases for assessment of the performance of the QA system. First, we synthesized random-error-disturbed R_{rs} spectra. The wavelengths were randomly chosen from the nine wavelengths of 412, 443, 488, 510, 531, 547, 555, 667, and 678 nm. Then the percentage errors were randomly set to $\delta = 5\%$, 10%, 20%, 30%, or 40%, which are assumed spectrally the same for all chosen wavelengths. So the error-disturbed spectra R^*_{rs} at wavelengths λ_i was represented as $R_{rs}(\lambda_i)^* = R_{rs}(\lambda_i) \times (1 + \delta)$. The data quality of R^*_{rs} was then examined with the score metric model and is illustrated in Figure 6a. When the R_{rs} spectra are only subjected to small errors (5%, in this case), very high scores ($\geq$8/9) are found for 90% of the spectra; 70% of the R_{rs} spectra have found with the highest scores of 1. As expected, with the increasingly larger errors being added more R_{rs} spectra show low scores. It is also noted that a small portion of R_{rs} spectra are found assigned to the highest scores of 1 even though the percentage errors are as high as up to 30% and 40%; this reflects the rare situation where the R_{rs} spectrum is subjected to same percentage errors at all nine wavelengths.

Unlike the random errors considered above, the R_{rs} spectra measured from space and in situ are often subject to spectrally dependent errors [*Bailey and Werdell*, 2006; *Hooker et al*., 2002; *Hu et al*., 2013; *Voss et al*., 2010; *Zibordi et al*., 2009a]. In blue oceanic waters, for example, the percentage errors of R_{rs} are more

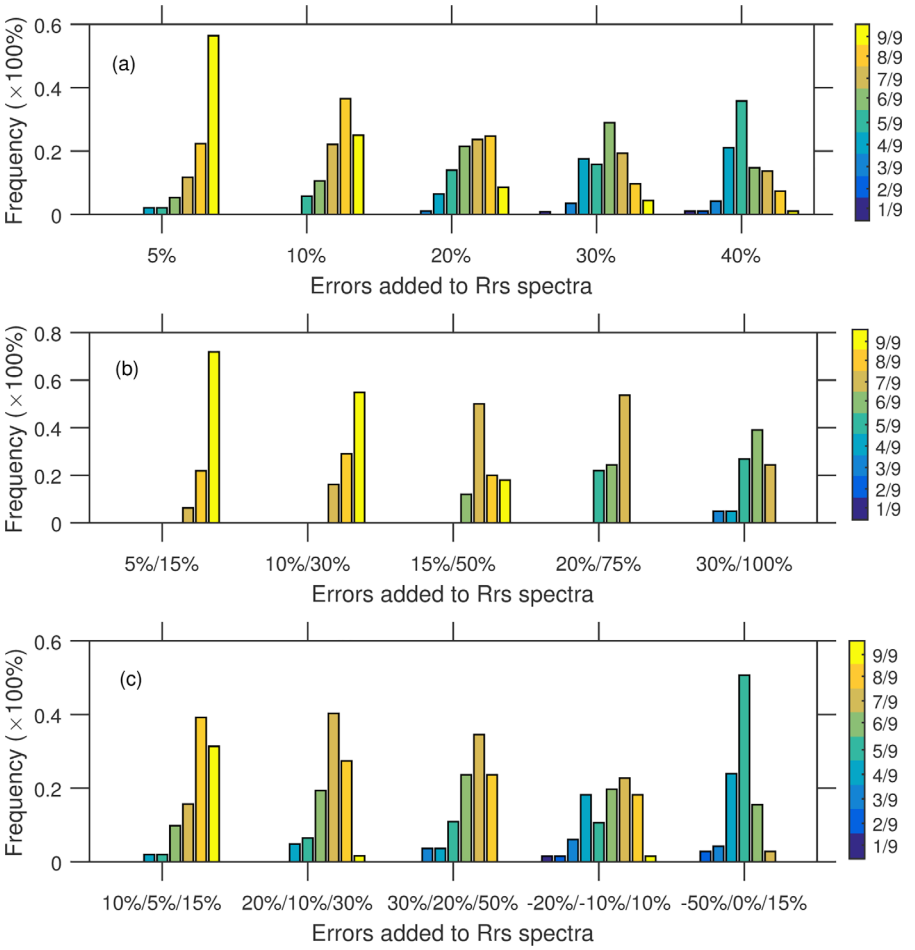

Figure 6. Frequency distribution of the scores for simulated error-disturbed R_{rs} spectra. (a) Same errors are added to randomly selected wavelengths. (b) Blue water R_{rs} spectra with different errors added to shorter wavelengths (412, 443, 488, 510, and 530 nm) and longer wavelengths (547, 555, 667, and 678 nm), separately; each pair of errors are denoted in the format of fractions in the plot. (c) Similar to (b) but for green and yellow water R_{rs} spectra; three errors are added to shorter wavelengths (412, 443, and 488 nm), intermediate wavelengths (510, 531, and 547 nm) and longer wavelengths (555, 667, and 678 nm), separately.

significant in longer wavelengths because of their extremely small reflectance values [*Hu et al.*, 2013]. Likewise, the percentage errors of R_{rs} spectra in green waters tend to be larger at shorter wavelengths. We synthesized new R_{rs} spectra with spectrally related errors. For the data belonging to water types 1–7, a relatively smaller error (5%, 10%, 15%, 20%, or 30%) and a larger error (correspondingly, 15%, 30%, 50%, 75%, or 100%) are universally added to the shorter wavelengths (412, 443, 488, 510, and 530 nm) and longer wavelengths (547, 555, 667, and 678 nm), separately. For other water types, we added a relatively larger error to the shorter wavelengths (412, 443, and 488 nm), a relatively small error to the intermediate wavelengths (510, 531, and 547 nm), and relatively large but different error to the longer wavelengths (555, 667, and 678 nm). Figure 6b illustrates the frequency distribution of the scores for the blue waters (in this case, water types 1–7), which has shown an increasing number of low-score spectra identified with increasing errors being added to R_{rs} spectra. Figure 6c describes the data quality for the green and yellow waters; the same conclusion can also be reached. In particular, the last group of frequency distribution in Figure 6c mimics the satellite ocean color in coastal waters when the absorbing aerosols are present. For such cases, the model of score metrics works effectively in identifying them by finding low scores.

3.2. Satellite Measured R_{rs} Spectra

The MODIS Aqua ocean color images were selected and retrieved from NASA ocean color data archive, which were calibrated and processed by Ocean Biology Processing Group (OBPG) with the latest processing (R2014). The R_{rs} spectra at six wavelengths of 412, 443, 488, 531, 547, and 667 nm were further obtained on

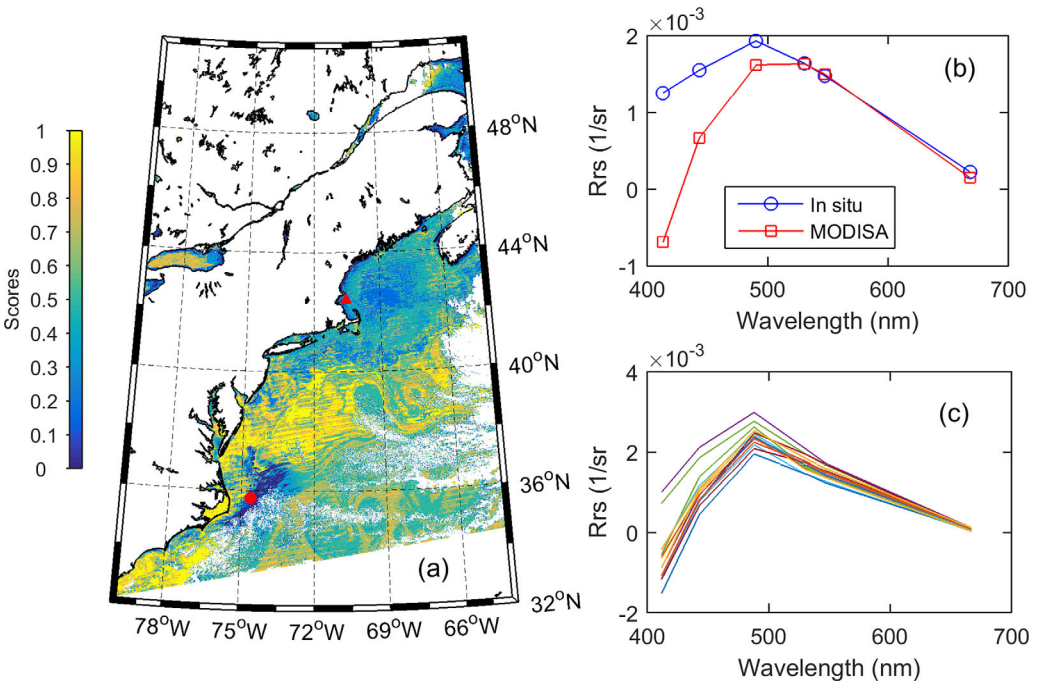

Figure 7. Quality evaluation of the MODISA ocean color data in US East Coasts and Mid-Atlantic Bight. (a) Scores of the satellite ocean color data (September 2015, A2015258180000.L2_LAC). (b) Comparison of the satellite and in situ ocean color matchup in Massachusetts Bay (in situ station is indicated by a "▲" symbol in Figure 7a; position: 42.4049°N, −70.5469°W; sampling time: 17:07 UTC). (c) MODISA measured "bad" R_{rs} spectra at the low-score pixels (indicated by "●" in Figure 7a).

a per-pixel basis. Quality of the R_{rs} spectrum of each pixel was then quantified by the score metrics, and examples are provided below.

3.2.1. Level-2 Ocean Color on the U.S. East Coasts

An ocean color image collected over the northeastern U.S. coasts on 15 September 2015 was retrieved from OBPG, and the sky was clear. As would be expected, the distribution of the QA scores reveals that the quality of R_{rs} data is very much location dependent (Figure 7a). The offshore waters in the Mid-Atlantic Bight and Southern Atlantic Bight are generally characterized by high QA scores, except for those near the shorelines. From the New York Bight, Gulf of Maine to St. Lawrence River and further north, however, the R_{rs} QA scores are generally lower than 0.5, suggesting questionable R_{rs} data for at least half of the evaluated spectral bands. In particular, the water masses in the center of the Gulf of Maine are of very low QA scores (less than 0.3), indicating highly questionable R_{rs} data for this ocean color image. Indeed, these low-score pixels are generally flagged in Level-2 ocean color products. For this particular image, the flags of "production failure (PRODFAIL)" are generally evoked and assigned to the pixels in the vast majority of the waters within the Gulf of Maine.

On the same day, R_{rs} was measured with the SBA system in the southwest of Gulf of Maine to validate the satellite ocean color data (Figure 7a). Comparison of the matchup data indicates that the MODIS Aqua R_{rs} spectrum (which has a QA score of 0.5) deviates from in situ matchup R_{rs} spectrum (which has a QA score of 1) increasingly from 488 nm toward 412 nm (Figure 7b). This finding confirms and validates the low QA score of the satellite R_{rs} spectrum. Note that the MODIS Aqua R_{rs} has apparently been underestimated at the shorter wavelengths, with negative values found at 412 nm, but is fairly reasonable at green and red bands. Such discrepancies suggest a failure of atmosphere correction for these measurements. The Gulf of Maine is generally clear and not subject to significant contribution of suspended sediments. A likely cause for the failure of atmospheric correction is the frequent occurrence of absorbing aerosols in the New England region [*Sierau et al.*, 2006].

It is interesting, and surprising, that the distribution of the QA score of this MODIS Aqua image is found related to the mesoscale eddy systems of the Gulf Stream: the clear blue waters in the Sargasso Sea are observed with relatively lower QA scores (<0.7), with extremely low QA score pixels (<0.1) found for a large

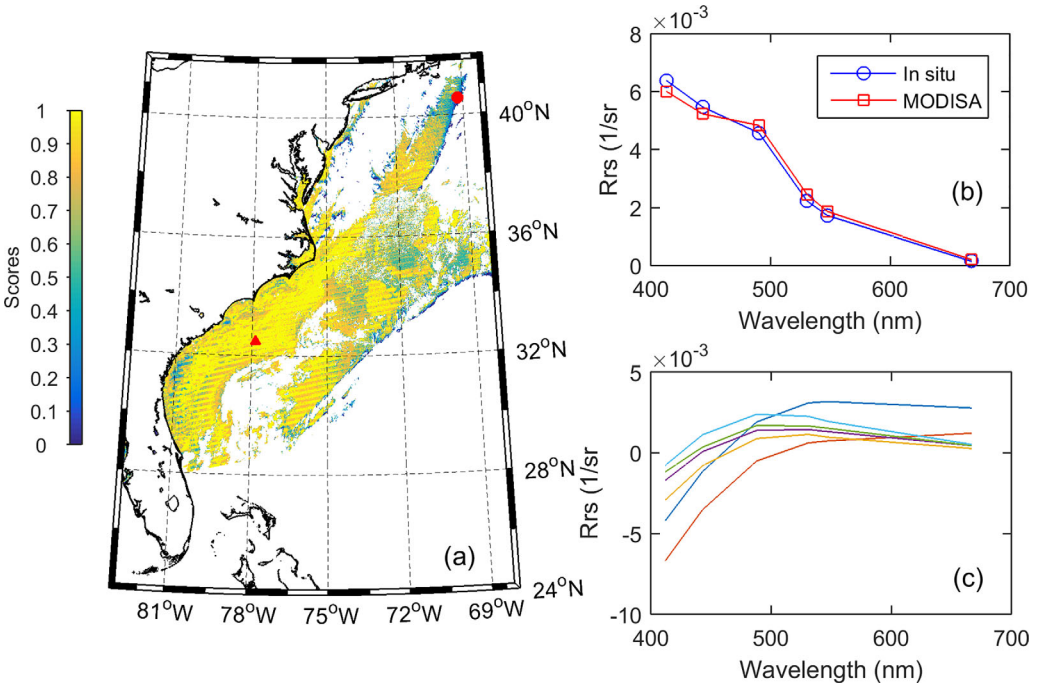

Figure 8. Quality evaluation of the MODISA ocean color data in US East Coasts and the Northwest Atlantic Ocean. (a) Scores of the satellite ocean color data (11 December 2015, A2015345180500.L2_LAC). (b) Comparison of the satellite ocean color and in situ data measured at a station indicated by "▲" symbol in Figure 8a (Station number: 20; position: 32.4944°N, −77.8581°W; sampling time: 17:04 UTC). (c) MODISA measured "bad" R_{rs} spectra from the pixels indicated by "●" in Figure 8a.

water mass stretching northeastward in the east of North Carolina (Figure 7a). These pixels are flagged in the Level-2 products as perturbation of moderate glints. We retrieved and plotted the R_{rs} spectra from this region and presented Figure 7c, and found that the R_{rs} spectra at shorter wavelengths are generally underestimated, likely a result of overcorrection of the atmospheric contributions due to glint effects.

Another MODIS Aqua ocean color image was obtained for east U.S coastal and offshore waters on 11 December 2015. The QA score map is shown in Figure 8a. The score distribution has shown very good data quality for most of the pixels in this scene. The coastal waters in west of the Atlantic Ocean are generally assigned to very high scores (>0.9), including lower Chesapeake Bay. In situ R_{rs} spectrum was obtained on the same day at a deep water location (off South Carolina) using the SBA system. A comparison of satellite R_{rs} with in situ R_{rs} confirms the good quality of the satellite data for this pixel (Figure 8b).

On the other hand, low-score pixels are also easily found along the edges of cloud patches (the clouds are masked out and shown in white, north of 28°N in the image). Most of these pixels at the edges are flagged as "product failure (PRODFAIL)" in the Level-2 products. According to the QA score distribution, however, the pixels with a distance more than a few pixels away from the cloud edges still show very low QA scores (<0.5). Figure 8c shows the R_{rs} spectra of these pixels, where the R_{rs} values at the blue wavelengths are obviously underestimated while R_{rs}(670) is likely overestimated.

3.2.2. Level-3 Ocean Color Data in Global Oceans

As examples, the QA metrics was also applied to global Level-3 MODISA daily (A2002188_L3m_9km) and monthly R_{rs} data (A20021822002212_L3m_9km) (Figure 9). Note all questionable pixels as flagged in Level-2 products have been removed before the Level-3 composites. As a result, the percentage of pixels with high scores in the Level-3 R_{rs} images is generally very high. Further, the frequency of high-score pixels in monthly product is slightly higher than that of daily product (for brevity, the figures are not shown here), which is a result of the Level-3 data binning procedure.

The satellite R_{rs} measurements in the majority of global oceans have QA scores close to 1, with the highest scores generally found in the center of the subtropical gyres. Not only the deep oceans but most coastal waters retained high QA scores for these Level-3 products. Exceptions do exist, however. Relatively lower QA scores (<0.6) are found in the Mediterranean Sea and in the equatorial upwelling regions in west of

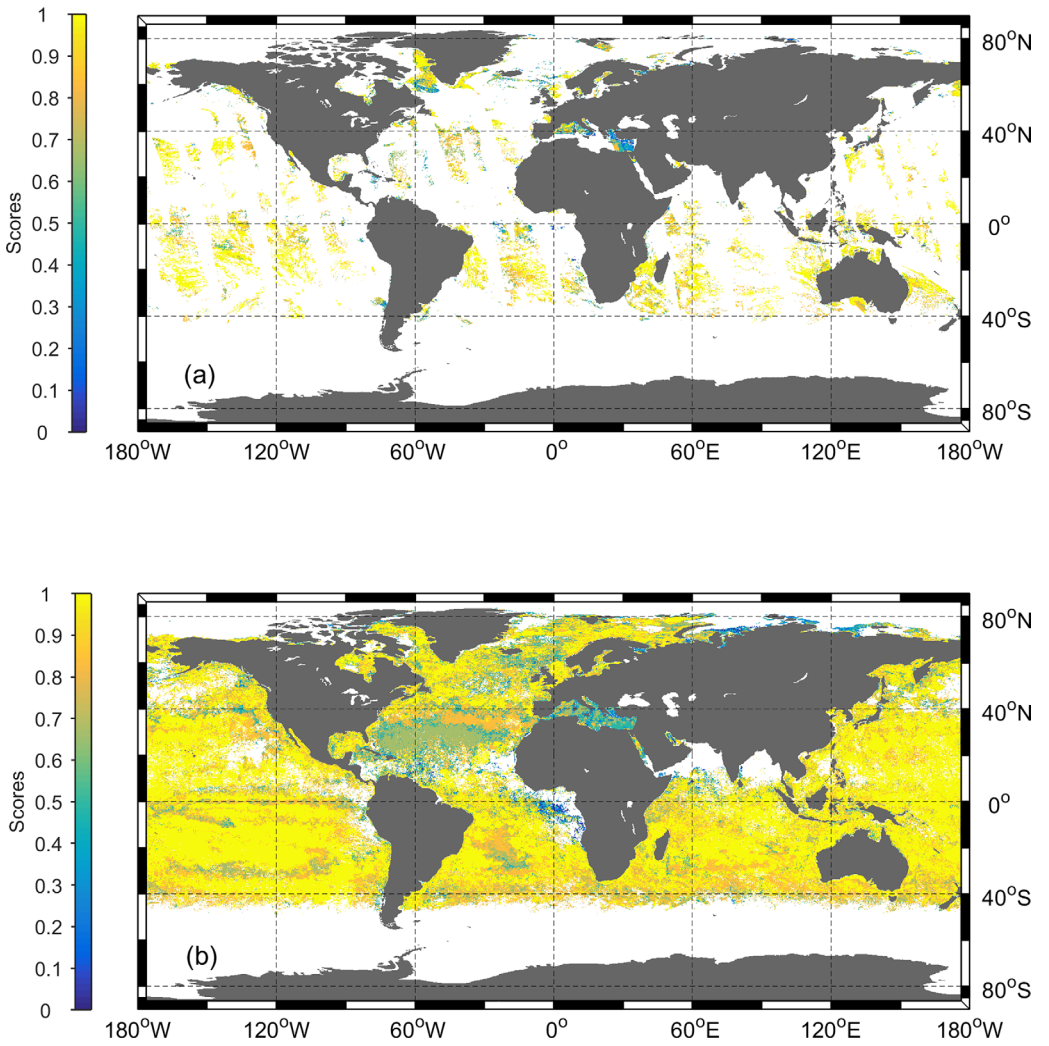

Figure 9. Scores of the MODISA R_{rs} data in global oceans. (a) Level-3 daily data (A2002188_L3m_9km). (b) Level-3 monthly data (A20021822002212_L3m_9km) for July 2002.

Africa (see Figure 9). Surprisingly, the extremely clear oceanic waters in the North Atlantic Gyre (NAG) do not show high QA scores either. It should be emphasized this phenomenon of low-score Level-3 R_{rs} data in NAG and the equatorial upwelling region in west of Africa persists almost throughout the year (for brevity, the figures for other months are omitted here). This low score could be due to underrepresentation of R_{rs} of NAG in the reference database, where the spectral shapes of these R_{rs} might be different from that of other oligotrophic waters because of likely impacts of the Saharan dusts [e.g., *Claustre et al.*, 2002]. It is necessary to carry out detailed in situ measurements over this region to pinpoint the exact reasons of the low QA scores of these clear waters.

3.3. In Situ Measured R_{rs} Spectra

The NOMAD data set was compiled from a large online data depository SeaBASS [*Werdell and Bailey*, 2005], and is a widely used for bio-optical model development and validations [*Hu et al.*, 2012; *O'Reilly et al.*, 1998; *Szeto et al.*, 2011]. Currently the database consists of more than 4000 measurements. For testing and application purposes, we only chose the R_{rs} spectra with available measurements at wavelengths of 412, 443, 488, 555, and 667 nm, and obtained a total of 2358 R_{rs} spectra (Figure 10). Among these spectra, about one-third was measured in clear oceanic waters with [CHL] < 0.25 mg m^{-3}; the rest were from turbid coastal waters. Two-thirds of the spectra were measured by the in-water approach, the rest was obtained from the above-water approach. After applying the QA system to these selected R_{rs}, we found low-quality R_{rs} spectra

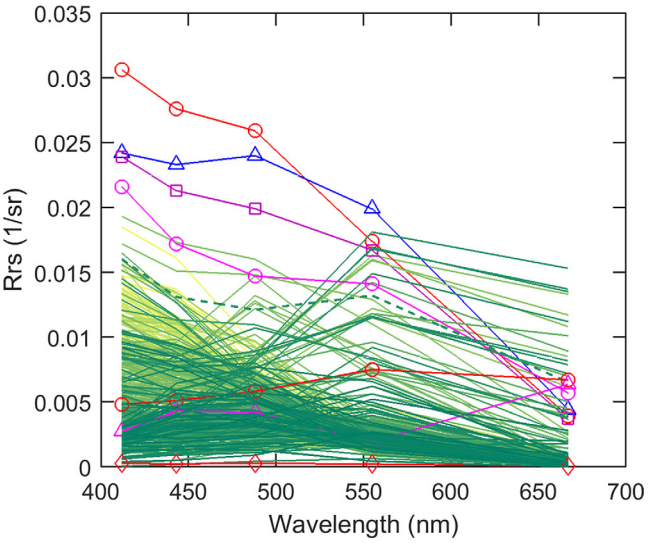

Figure 10. R_{rs} spectra extracted from NOMAD database for QA testing ($n = 2358$). The spectra highlighted with symbols (circles, diamonds, square, and triangles) exemplify the bad R_{rs} spectra present in NOMAD.

accidentally included in the NOMAD data set. For illustration, a few questionable R_{rs} spectra (QA scores < 0.5) are highlighted in Figure 10. In addition, we found that the QA scores are not dependent on instruments or [CHL], and R_{rs} with low QA appeared in every ocean basins (For brevity, the distribution map is not shown).

The inclusion of low-quality R_{rs} data in NOMAD may impact the related ocean color algorithm development and subsequent model validations. For example, if we only use the R_{rs} spectra with the highest scores (QA score = 1) for the development of [CHL] algorithm, the coefficients of the new algorithm (here, called OC3m-QA) will be quite different than the original OC3m algorithm, with improved model accuracy (see Figure 11).

4. Discussion

4.1. Characteristics of the New Water Type Classification

The score metrics developed in this study relies on cosine-based classification scheme, which groups the R_{rs} spectra with similar spectral shapes. This optical classification is distinctive from other practices [*Le et al*., 2011; *Mélin and Vantrepotte*, 2015; *Moore et al*., 2009, 2001]. The spectral angle mapper [*Kruse et al*., 1993] is widely used in land remote-sensing community and its extension to ocean color application deserves more discussion.

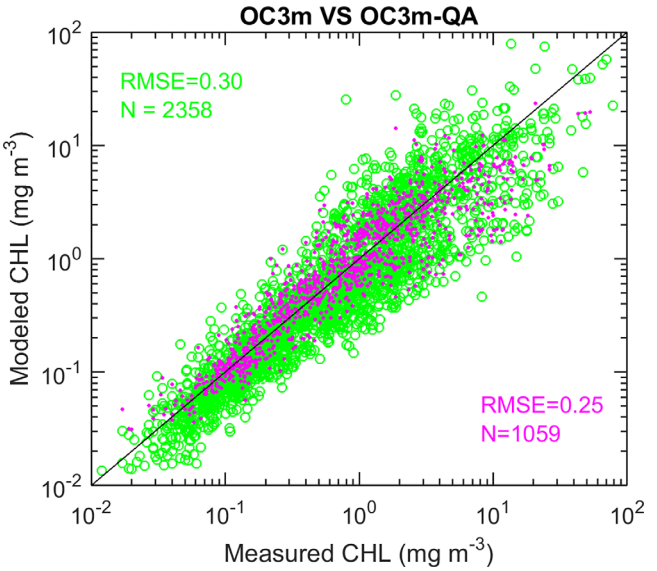

Figure 11. Comparison of estimated [CHL] with in situ measured [CHL]. The original (not quality controlled) data are compared to OC3m modeled data (in green open circles). The quality controlled data (in pink dots, with the highest scores) are compared to OC3m-QA estimated data. OC3m-QA has new model coefficients developed from these highest-score data (N = 1059, excluding those from the Arctic and Southern Oceans) only. The OC3m-QA has the new model coefficients as 0.2053, −2.0159, 1.1159, −1.3885, and 0.5474.

First, it suppresses the disturbances of absolute amplitudes of R_{rs} spectra to the clustering, and only takes in account the spectral shapes of R_{rs}. A direct consequence is that only the waters with similar "water color" can be grouped together. For example, the blue water is by no accident combined with green or yellow color waters because of their contrasting R_{rs} spectral shapes (Figure 3).

Second, the normalization in equation (1) basically rescales the R_{rs} spectra to the domain of (0,1), rendering the nR_{rs} spectra comparable with respect to their "amplitudes." Moreover, the normalization retains the same spectral ratios for nR_{rs} with the corresponding R_{rs} spectra, while at the same time has constrained the clustered nR_{rs} spectra within a very narrow range. To illustrate, the coefficient of variation (CV) of nR_{rs} and R_{rs} is compared for each

Table 2. Coefficient of Variation (%) of the nR_{rs} Spectra and Corresponding R_{rs} Spectra for Each Water Type

	$nR_{rs}(\lambda)$				$R_{rs}(\lambda)$			
Water Type	412 nm	443 nm	555 nm	667 nm	412 nm	443 nm	555 nm	667 nm
1	3	2	21	86	25	24	13	77
2	3	1	11	43	19	18	18	51
3	4	2	11	55	21	20	20	78
4	6	4	8	34	27	26	25	46
5	7	4	8	38	37	38	44	96
6	8	3	5	20	43	44	44	48
7	8	4	4	29	60	60	60	66
8	9	5	4	21	61	63	63	64
9	9	4	6	29	82	95	91	145
10	13	6	4	25	60	66	64	87
11	11	8	6	16	108	119	103	129
12	13	7	3	14	58	61	58	65
13	18	11	3	25	39	40	39	44
14	13	8	3	11	42	46	39	46
15	18	9	2	10	52	54	48	53
16	15	11	4	6	105	111	108	111
17	14	9	4	7	86	92	84	91
18	22	13	3	24	55	51	41	50
19	21	17	3	5	58	49	47	46
20	25	15	4	10	52	53	44	48
21	23	14	4	13	51	46	40	46
22	28	16	5	8	67	71	74	74
23	20	18	6	7	46	46	32	36

water type in Table 2. The CV's of nR_{rs} spectra are much smaller than the CV's of the corresponding R_{rs} spectra in the visible domain. In blue waters, such a tendency is more evident at short wavelengths because of higher R_{rs} values in this spectral domain. In contrast, it is more obvious at longer wavelengths in the green or yellowish waters.

4.2. Applications to Satellite Ocean Color Validations

A common practice in evaluating the satellite ocean color data quality is through comparison with in situ data set; with the latter assumed as the "ground truth" [*Bailey and Werdell*, 2006] by default. Graphically, the ocean color data matchups can be illustrated in a manner of scatter-plots. Various error statistics can be derived for the matchups, such as the relative percentage error (ε), unbiased absolute percentage error (δ), and the root-mean-square error (RMSE). These estimators convey important information on the overall quality of various ocean color observations [*IOCCG*, 2006; *Zibordi et al*., 2009a]. However, the practices of scatter-plots alone do not explicitly convey the quality of an individual R_{rs} spectrum. The limitation is further illustrated in Figure 12 with R_{rs} spectra. The newly developed QA system complements this common validation practice in ocean color products, and assesses the data quality for individual R_{rs} spectra.

The ocean color data measured by different sensors have different number of spectral bands and different central wavelengths, including MODIS Aqua (412, 443, 488, 531, 547, 667, and 678 nm), SNPP VIIRS (412, 443, 488, 555, and 667 nm), SeaWiFS (412, 443, 488, 510, 531, 555, and 667 nm), MERIS (412, 443, 488, 510, 555, 667, and 678 nm), and Landsat 8 (443, 482, 561, and 655 nm). The QA score system developed here includes nine spectral bands, which have more bands than any of these satellite ocean color sensors. It is important to evaluate whether the developed QA system with nine wavelengths can be reliably applied to R_{rs} data from sources with lower number of spectral bands. To test this impact, we extracted four subsets of R_{rs} data from the reference data set used in section 2, according to different wavelengths of MODIS Aqua, SeaWiFS, VIIRS and Landsat 8 (for Landsat 8, we assumed 443, 488, 555, and 667 nm as the central wavelengths). We first tested the repeatability of identified water types. It is found that the water type clustering is fairly accurate, even though different spectral bands are employed. The statistics of the scores determined for every sensor are presented in Figure 13. For each case, about 90% of the evaluation spectra are assigned to high QA scores of $>$0.8. The best performance is found with SeaWiFS data, then MODIS Aqua, VIIRs and Landsat 8 data (Figure 13b). This is probably related to the fact that SeaWiFS has the highest number of wavelengths, while the simulated Landsat 8 data only have four bands being used in the QA system. Note that in the comparisons we ignored the potential effects due to the differences in the spectral

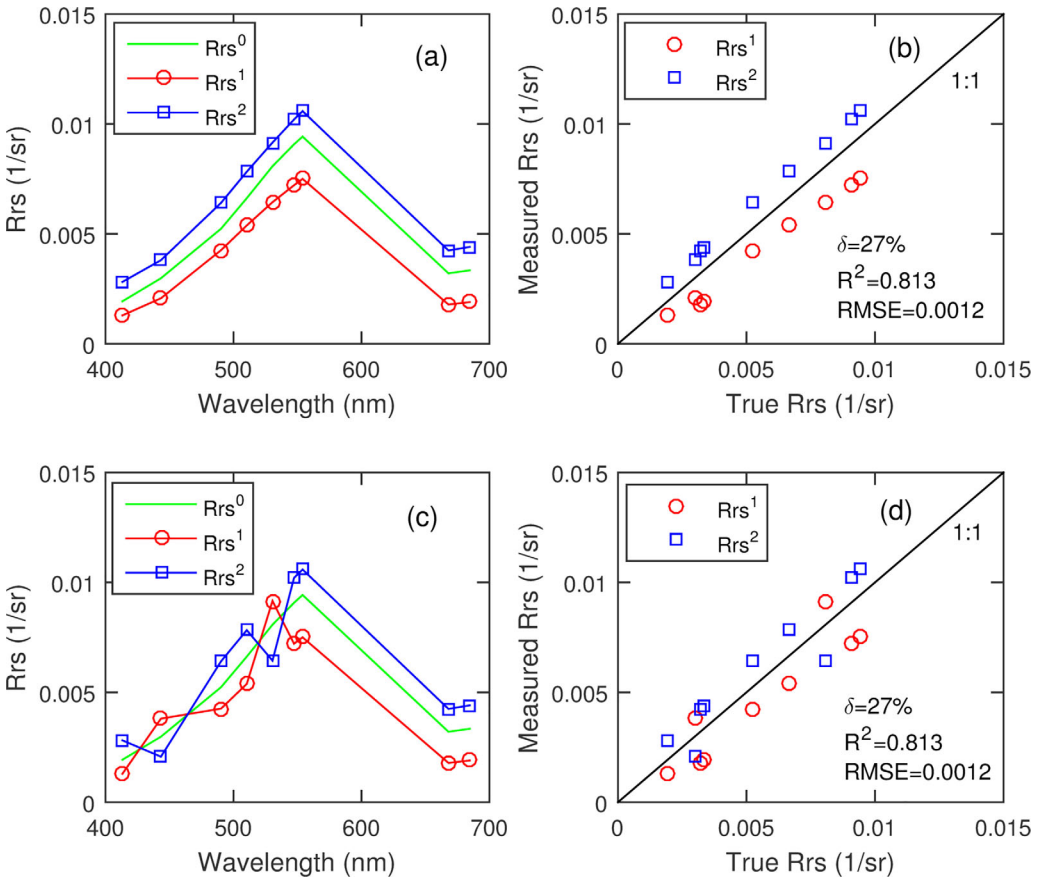

Figure 12. Application and limitation of scatter-plot practice in evaluating the data quality of R_{rs} spectra. The R_{rs}^0 spectrum (in green) represents true values. The R_{rs}^1 and R_{rs}^2 spectra refer to two example measurements. The scatter-plot of Figure 12b represents the measurement scenario in Figure 12a, where the R_{rs}^1 and R_{rs}^2 spectra retain the same shapes as R_{rs}^0. Figure 12c represents another situation when the R_{rs}^1 and R_{rs}^2 spectra are erroneously measured (for brevity, we simply switched R_{rs} values at some bands between R_{rs}^1 and R_{rs}^2 spectra). Interestingly, the resulted new scatter-plot in Figure 12d is exactly the same with Figure 12b, with the same error statistics.

bandwidth (FWHM) of satellite sensors. Landsat 8 has a much wider bandwidth (15–60 nm) than the other ocean color satellite sensors (10–20 nm) in the visible domain.

On the other hand, for hyperspectral data, a modification or expansion of this QA system can be readily applicable.

4.3. Bio-Optical Implications and Significance

The remote-sensing reflectance is probably the most commonly used apparent optical property (AOP) in ocean color remote sensing. For example, the widely used band ratio [*O'Reilly et al.*, 1998] or band difference [*Hu et al.*, 2012] algorithms for [CHL] rely on R_{rs} data at both blue and green wavelengths. Semianalytical bio-optical algorithms require the R_{rs} spectrum at more wavelengths as inputs [*Ciotti et al.*, 1999; *IOCCG*, 2000, 2006; *Wei and Lee*, 2015]. Thus, the QA of every individual R_{rs} spectrum is fundamental for accurate remote sensing of water-column properties. When questionable or erroneous R_{rs} spectra are involved, the bio-optical retrievals from the ocean color will mostly likely be undermined. Our preliminary analysis demonstrates the significance of R_{rs} data quality to estimates of [CHL] (Figures 10 and 11). Errors or uncertainties of R_{rs} measurements will propagate to the retrievals of inherent optical retrievals as well [*Lee et al.*, 2010b].

The "low" data quality of MODIS R_{rs} found in the North Atlantic Gyre and in the Mediterranean Sea (Figure 9) may be partly related to an insufficient representativeness of R_{rs} spectra in the database for these waters. Indeed, the oligotrophic waters in North Atlantic and marginal seas could be optically different than that in the Pacific or South Atlantic, because the former are likely subject to the impacts of dusts transported from Sahara in the easterly trades and Asian dusts from the west [*Uno et al.*, 2009]. The absorptive fine mineral particles present in surface waters can attenuate the blue light, addition to the contribution of

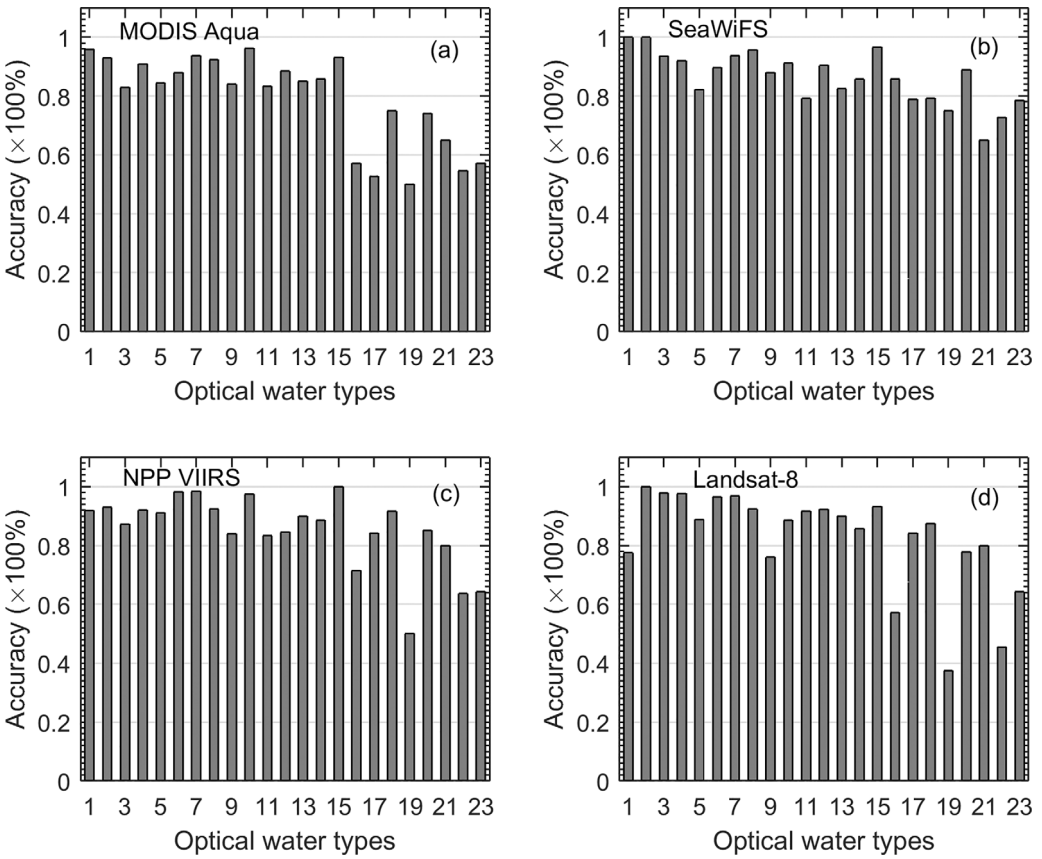

Figure 13. Testing results for the accuracy of score metrics model when applied to satellite ocean color data. (a) MODIS Aqua: 412, 443, 488, 531, 547, and 667 nm; (b) SeaWiFS: 412, 443, 488, 510, 531, 555, and 667 nm; (c) VIIRS: 412, 443, 488, 555, and 667 nm; (d) Landsat 8 (assumed 443, 488, 555, and 667 nm).

phytoplankton absorption and CDOM and detritus [*Claustre et al.*, 2002]. On the other hand, the presence of absorptive particles in the atmosphere poses a challenge to atmospheric correction, resulting in underestimated R_{rs} at blue bands in these waters [*Schollaert et al.*, 2003]. The latter situation could have occurred in the satellite retrieved ocean color data. A comprehensive investigation with accurate field measurement of R_{rs} is required to completely understand the low QA scores of MODIS R_{rs} in those waters.

Nevertheless, it is imperative to QC and quality assure an individual R_{rs} spectrum measured from both in situ and satellite platforms, in order to produce multisensor optical and biogeochemical products with high quality.

5. Conclusions

We have developed a QA system to objectively measure the quality of each individual R_{rs} spectrum. The QA system is based on classification of R_{rs} spectral similarity determined by the cosine distance, which is further measured by the upper and lower boundaries obtained from a wide range of field observations. This scheme and the spectral constraints together constitute a robust and quantitative QA system. Basically the metric system provides information of likely reliability of a target R_{rs} spectrum. If a score of 0 is reached, it indicates the target R_{rs} is highly questionable, although it could be true for that specific environment. The performance of this QA scheme was tested with synthesized, in situ measured, and satellite obtained R_{rs} spectra and has shown great reliability and applicability, where "bad" or highly questionable R_{rs} spectra were well identified.

The QA metric system does not require a priori knowledge of the optical properties of waters under study, and it is applicable to multispectral or hyperspectral R_{rs} data obtained from any platform. Since it is based on in situ R_{rs} measurements from a wide range of water types, the innovative QA system can be universally used for QC of ocean color data from global oceans.

Because all field measurements are discrete, it is likely that the database used here does not cover every water types and/or there are situations where the range of R_{rs} variability goes beyond the domains defined here. Such a limitation can be updated or revised when more high-quality in situ measurements are available. A MATLAB® script is made available (http://oceanoptics.umb.edu/score_metric/) to facilitate the evaluation and refinement of the score metrics. Nevertheless, this QA scheme provides an easily applicable system to quantitatively evaluate the quality of individual R_{rs} spectra.

APPENDIX A

Table A1. Upper Boundary of the Normalized Remote-Sensing Reflectance Spectra for the 23 Water Types

OWT	412	443	488	510	531	54750	555	667	678
1	0.780	0.559	0.367	0.203	0.138	0.109	0.096	0.046	0.047
2	0.711	0.555	0.424	0.254	0.182	0.141	0.126	0.028	0.027
3	0.646	0.540	0.471	0.322	0.243	0.197	0.173	0.067	0.062
4	0.570	0.515	0.528	0.374	0.312	0.265	0.240	0.062	0.062
5	0.478	0.488	0.548	0.418	0.352	0.314	0.301	0.099	0.098
6	0.423	0.416	0.506	0.427	0.390	0.358	0.345	0.065	0.071
7	0.362	0.386	0.485	0.439	0.413	0.378	0.360	0.090	0.096
8	0.328	0.343	0.464	0.449	0.441	0.418	0.412	0.094	0.140
9	0.429	0.369	0.434	0.413	0.412	0.403	0.410	0.166	0.175
10	0.283	0.318	0.471	0.451	0.451	0.454	0.452	0.128	0.125
11	0.360	0.319	0.373	0.400	0.427	0.451	0.477	0.170	0.284
12	0.253	0.287	0.374	0.405	0.439	0.475	0.507	0.183	0.188
13	0.235	0.253	0.392	0.424	0.473	0.486	0.488	0.128	0.134
14	0.263	0.263	0.350	0.382	0.429	0.461	0.507	0.262	0.276
15	0.202	0.219	0.333	0.381	0.448	0.493	0.521	0.203	0.224
16	0.230	0.224	0.296	0.339	0.382	0.432	0.465	0.393	0.419
17	0.232	0.244	0.316	0.355	0.415	0.463	0.503	0.302	0.313
18	0.202	0.204	0.309	0.376	0.455	0.522	0.560	0.163	0.170
19	0.066	0.147	0.236	0.296	0.367	0.415	0.439	0.479	0.493
20	0.159	0.184	0.296	0.356	0.429	0.500	0.571	0.290	0.293
21	0.235	0.237	0.293	0.336	0.443	0.515	0.605	0.241	0.286
22	0.159	0.167	0.251	0.318	0.408	0.482	0.573	0.351	0.383
23	0.180	0.167	0.198	0.233	0.310	0.452	0.578	0.379	0.509

Table A2. Lower Boundary of the Normalized Remote-Sensing Reflectance Spectra for the 23 Water Types

OWT	412	443	488	510	531	547	555	667	678
1	0.709	0.512	0.271	0.119	0.073	0.053	0.044	0.002	0.002
2	0.638	0.509	0.364	0.198	0.132	0.100	0.084	0.003	0.003
3	0.553	0.497	0.412	0.246	0.179	0.140	0.119	0.007	0.007
4	0.436	0.438	0.419	0.310	0.241	0.193	0.169	0.010	0.011
5	0.365	0.390	0.417	0.366	0.287	0.232	0.202	0.016	0.015
6	0.307	0.360	0.405	0.387	0.347	0.297	0.272	0.029	0.028
7	0.251	0.315	0.415	0.403	0.373	0.334	0.306	0.016	0.021
8	0.195	0.266	0.375	0.386	0.390	0.371	0.345	0.023	0.025
9	0.295	0.316	0.367	0.362	0.359	0.352	0.341	0.058	0.066
10	0.131	0.234	0.336	0.381	0.407	0.390	0.376	0.022	0.032
11	0.247	0.240	0.311	0.345	0.366	0.370	0.377	0.085	0.118
12	0.148	0.207	0.302	0.336	0.409	0.425	0.427	0.110	0.115
13	0.092	0.161	0.313	0.375	0.423	0.438	0.436	0.024	0.023
14	0.158	0.200	0.265	0.311	0.382	0.427	0.438	0.154	0.179
15	0.066	0.149	0.273	0.334	0.418	0.455	0.466	0.135	0.143
16	0.156	0.161	0.226	0.282	0.356	0.394	0.417	0.328	0.332
17	0.137	0.176	0.252	0.310	0.388	0.418	0.437	0.244	0.243
18	0.058	0.116	0.249	0.321	0.419	0.480	0.499	0.050	0.054
19	0.032	0.080	0.183	0.246	0.324	0.378	0.411	0.417	0.409
20	0.036	0.096	0.218	0.293	0.395	0.464	0.490	0.204	0.217
21	0.107	0.141	0.199	0.246	0.347	0.464	0.508	0.149	0.171
22	0.073	0.098	0.200	0.249	0.330	0.450	0.485	0.264	0.292
23	0.093	0.095	0.146	0.194	0.265	0.382	0.485	0.301	0.383

Acknowledgments
This study is funded by the National Oceanic and Atmospheric Administration (NOAA) JPSS VIIRS Ocean Color Cal/Val Project (Lee), the National Key Research and Development Program of China (2016YFC1400900, Shang), the National Aeronautic and Space Administration (NASA) Ocean Biology and Biogeochemistry and Water and Energy Cycle Programs (Lee), and the Chinese Ministry of Science and Technology (2016YFA0601201, Shang), and the NSF-China (41576169, Shang). We are indebted to Giuseppe Zibordi (Joint Research Center, Italy), Chuanmin Hu (University of South Florida), Marlon Lewis (Dalhousie University, Canada) for sharing in situ radiometric measurements and many investigators who submitted data to SEABASS (data used in this effort are downloadable from http://oceanoptics.umb.edu/score_metric). We appreciate the fine work of Kelly Luis in the correction of the English text. We thank an anonymous reviewer and Giuseppe Zibordi for constructive comments which have improved the manuscript.

References

Arnone, R. A., A. M. Wood, and R. W. Gould Jr. (2004), Science box: The evolution of optical water mass classification, *Oceanography*, *17*(2), 14–15.

Austin, R. W. (1974), The remote sensing of spectral radiance from below the ocean surface, in *Optical Aspects of Oceanography*, edited by N. G. Jerlov and E. Steemann Nielsen, pp. 317–344, Academic, New York.

Bailey, S. W., and P. J. Werdell (2006), A multi-sensor approach for the on-orbit validation of ocean color satellite data products, *Remote Sens. Environ.*, *102*(1–2), 12–23, doi:10.1016/j.rse.2006.01.015.

Bailey, S. W., S. B. Hooker, D. Antoine, B. A. Franz, and P. J. Werdell (2008), Sources and assumptions for the vicarious calibration of ocean color satellite observations, *Appl. Opt.*, *47*(12), 2035–2045.

Chen, J., T. Cui, and W. Quan (2015), A neural network-based four-band model for estimating the total absorption coefficients from the global oceanic and coastal waters, *J. Geophys. Res.*, *120*, 36–49, doi:10.1002/2014JC010461.

Ciotti, Á. M., J. J. Cullen, and M. R. Lewis (1999), A semi-analytical model of the influence of phytoplankton community structure on the relationship between light attenuation and ocean color, *J. Geophys. Res.*, *104*(C1), 1559–1578.

Claustre, H., A. Morel, S. B. Hooker, M. Babin, D. Antoine, K. Oubelkheir, A. Bricaud, K. Leblanc, B. Queguiner, and S. Maritorena (2002), Is desert dust making oligotrophic waters greener?, *Geophys. Res. Lett.*, *29*(10), 1469, doi:10.1029/2001GL014056.

D'Alimonte, D., and G. Zibordi (2006), Statistical assessment of radiometric measurements from autonomous systems, *IEEE Trans. Geosci. Remote Sens.*, *44*(3), 719–728, doi:10.1109/TGRS.2005.862505.

Gordon, H. R., and K. Ding (1992), Self-shading of in-water optical instruments, *Limnol. Oceanogr.*, *37*(3), 491–500.

Gordon, H. R., and M. Wang (1994a), Influence of oceanic whitecaps on atmospheric correction of ocean-color sensor, *Appl. Opt.*, *33*, 7754–7763.

Gordon, H. R., and M. Wang (1994b), Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: A preliminary algorithm, *Appl. Opt.*, *33*(3), 443–452.

Hooker, S. B., W. E. Esaias, G. C. Feldman, W. W. Gregg, and C. R. McClain (1992), An overview of SeaWiFS and Ocean Color, *NASA Tech. Memo. Rep. 10456*, 24 pp., NASA Goddard Space Flight Cent., Greenbelt, Md.

Hooker, S. B., G. Lazin, G. Zibordi, and S. McLean (2002), An evaluation of above- and in-water methods for determining water-leaving radiances, *J. Atmos. Oceanic Technol.*, *19*(4), 486–515.

Hu, C., F. E. Müller-Karger, C. Taylor, K. L. Carder, C. Kelble, E. Johns, and C. Heil (2005), Red tide detection and tracing using MODIS fluorescence data: A regional example in SW Florida coastal waters, *Remote Sens. Environ.*, *97*, 311–321.

Hu, C., Z. P. Lee, and B. Franz (2012), Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference, *J. Geophys. Res.*, *117*, C01011, doi:10.1029/2011JC007395.

Hu, C., L. Feng, and Z. P. Lee (2013), Uncertainties of SeaWiFS and MODIS remote sensing reflectance: Implications from clear water measurements, *Remote Sens. Environ.*, *133*(2), 168–182.

IOCCG (2000), Remote sensing of ocean colour in coastal, and other optically-complex, waters, *Int. Ocean-Colour Coord. Group Rep. 3*, 140 pp, IOCCG, Dartmouth, Nova Scotia, Canada.

IOCCG (2006), Remote sensing of inherent optical properties: Fundamentals, tests of algorithms, and applications, *Int. Ocean-Colour Coord. Group Rep. 5*, 126 pp., IOCCG, Dartmouth, Nova Scotia, Canada.

IOCCG (2010), Atmospheric correction for remotely-sensed ocean color products, *Int. Ocean-Colour Coord. Group Rep. 10*, 78 pp., IOCCG, Dartmouth, Nova Scotia, Canada.

Kruse, F. A., A. B. Lefkoff, J. B. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. Goetz (1993), The spectral image processing system (SIPS)—Interactive visualization and analysis of imaging spectrometer data, *Remote Sens. Environ.*, *44*, 145–163.

Le, C., Y. Li, Y. Zha, D. Sun, C. Huang, and H. Zhang (2011), Remote estimation of chlorophyll a in optically complex waters based on optical classification, *Remote Sens. Environ.*, *115*(2), 725–737.

Lee, Z. P., K. L. Carder, and R. Arnone (2002), Deriving inherent optical properties from water color: A multi-band quasi-analytical algorithm for optically deep waters, *Appl. Opt.*, *41*(27), 5755–5772.

Lee, Z. P., Y.-H. Ahn, C. Mobley, and R. Arnone (2010a), Removal of surface-reflected light for the measurement of remote-sensing reflectance from an above-surface platform, *Opt. Express*, *18*(25), 26,313–26,342.

Lee, Z. P., R. Arnone, C. Hu, P. J. Werdell, and B. Lubac (2010b), Uncertainties of optical parameters and their propagations in an analytical ocean color inversion algorithm, *Appl. Opt.*, *49*(3), 369–381.

Lee, Z. P., K. Du, K. J. Voss, G. Zibordi, B. Lubac, R. Arnone, and A. Weidemann (2011), An inherent-optical-property-centered approach to correct the angular effects in water-leaving radiance, *Appl. Opt.*, *50*(19), 3155–3167.

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

Lee, Z. P., J. Wei, K. Voss, M. Lewis, A. Bricaud, and Y. Huot (2015), Hyperspectral absorption coefficient of "pure" seawater in the range of 350–550 nm inverted from remote sensing reflectance, *Appl. Opt.*, *54*(3), 546–558.

Lloyd, S. (1982), Least squares quantization in PCM, *IEEE Trans. Inform. Theory*, *28*(2), 129–137, doi:10.1109/TIT.1982.1056489.

Mannino, A., M. E. Russ, and S. B. Hooker (2008), Algorithm development and validation for satellite-derived distributions of DOC and CDOM in the US Middle Atlantic Bight, *J. Geophys. Res.*, *113*, C07051, doi:10.1029/2007JC004493.

McClain, C. R., W. E. Esaias, W. Barnes, B. Guenther, D. Endres, S. B. Hooker, B. G. Mitchell, and R. Barnes (1992), Calibration and validation plan for SeaWiFS, in *NASA Tech. Memo. 104566*, vol. 3, edited by S. B. Hooker and E. R. Firestone, 41 pp., NASA Goddard Space Flight Cent., Greenbelt, Md.

Mélin, F., and V. Vantrepotte (2015), How optically diverse is the coastal ocean?, *Remote Sens. Environ.*, *160*, 235–251, doi:10.1016/j.rse.2015.01.023.

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

Mobley, C. D., and L. K. Sundman (2008), *Hydrolight 5 and Ecolight 5 User's Guide*, 99 pp., Sequoia Scientific, Inc., Bellevue, Wash.

Moore, T. S., J. W. Campbell, and H. Feng (2001), A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms, *IEEE Trans. Geosci. Remote Sens.*, *39*(8), 1764–1776.

Moore, T. S., J. W. Campbell, and M. D. Dowell (2009), A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product, *Remote Sens. Environ.*, *113*(11), 2424–2430.

Morel, A. (1974), Optical properties of pure water and pure sea water, in *Optical Aspects of Oceanography*, edited by N. G. Jerlov and E. Steemann Nielsen, 494 pp., Academic, New York.

Morel, A., and B. Gentili (1996), Diffuse reflectance of oceanic waters. III: Implications of bidirectionality for the remote-sensing problem, *Appl. Opt.*, *35*(24), 4850–4862.

Mueller, J. L., G. S. Fargion, and C. R. McClain (2003), Radiometric measurements and data analysis protocols, *Ocean Opt. Protoc. for Satell. Ocean Color Sensor Validation Rep. NASA/TM-2003-21621/Rev-Vol III*, 78 pp., NASA Goddard Space Flight Cent., Greenbelt, Md.

O'Reilly, J., S. Maritorena, B. G. Mitchell, D. Siegel, K. L. Carder, S. Garver, M. Kahru, and C. McClain (1998), Ocean color chlorophyll algorithms for SeaWiFS, *J. Geophys. Res.*, *103*, 24,937–24,953.

Pope, R. M., and E. S. Fry (1997), Absorption spectrum (380–700 nm) of pure water. II. Integrating cavity measurements, *Appl. Opt.*, *36*(33), 8710–8723.

Robinson, W. D., B. A. Franz, F. S. Patt, S. W. Bailey, and P. J. Werdell (2003), Masks and flags updates, in Algorithm Updates for the Fourth SeaWiFS Data Reprocessing, edited by S. B. Hooker and E. R. Firestone, pp. 34–40, NASA Goddard Space Flight Cent., Greenbelt, Md.

Ruddick, K., V. De Cauwer, and B. Van Mol (2005), Use of the near infrared similarity reflectance spectrum for the quality control of remote sensing data, *Proc. SPIE*, *5885*, doi:10.1117/12.615152.

Salama, M. S., and A. Stein (2009), Error decomposition and estimation of inherent optical properties, *Appl. Opt.*, *48*(26), 4947–4962, doi:10.1364/AO.48.004947.

Schollaert, S. E., J. A. Yoder, J. E. O'Reilly, and D. L. Westphal (2003), Influence of dust and sulfate aerosols on ocean color spectra and chlorophyll a concentrations derived from SeaWiFS off the U.S. east coast, *J. Geophys. Res.*, *108*(C6), 3191, doi:10.1029/2000JC000555.

Sierau, B., D. S. Covert, D. J. Coffman, P. K. Quinn, and T. S. Bates (2006), Aerosol optical properties during the 2004 New England Air Quality Study–Intercontinental Transport and Chemical Transformation: Gulf of Maine surface measurements—Regional and case studies, *J. Geophys. Res.*, *111*, D23S37, doi:10.1029/2006JD007568.

Sohn, Y., E. Moran, and F. Gurri (1999), Deforestation in North-Central Yucatan(1985–1995)—Mapping secondary succession of forest and agricultural land use in Sotuta using the cosine of the angle concept, *Photogramm. Eng. Remote Sens.*, *65*, 947–958.

Szeto, M., P. J. Werdell, T. S. Moore, and J. W. Campbell (2011), Are the world's oceans optically different?, *J. Geophys. Res.*, *116*, C00H04, doi:10.1029/2011JC007230.

Tanaka, A., and H. Sasaki (2006), Alternative measuring method for water-leaving radiance using a radiance sensor with a domed cover, *Opt. Express*, *14*(8), 3099–3105.

Tibshirani, R., G. Walther, and T. Hastie (2001), Estimating the number of clusters in a data set via the gap statistic, *J. R. Stat. Soc. Ser. B*, *63*, part 2, 411–423.

Toole, D. A., D. A. Siegel, D. W. Menzies, M. J. Neumann, and R. C. Smith (2000), Remote-sensing reflectance determinations in the coastal ocean environment: Impact of instrumental characteristics and environmental variability, *Appl. Opt.*, *39*(3), 456–469.

Uno, I., K. Eguchi, K. Yumimoto, T. Takemura, A. Shimizu, M. Uematsu, Z. Liu, Z. Wang, Y. Hara, and N. Sugimoto (2009), Asian dust transported one full circuit around the globe, *Nat. Geosci.*, *2*(8), 557–560.

Voss, K. J., and A. Morel (2005), Bidirectional reflectance function for oceanic waters with varying chlorophyll concentrations: Measurements versus predictions, *Limnol. Oceanogr. Methods*, *50*(2), 698–705.

Voss, K. J., S. McLean, M. R. Lewis, C. Johnson, S. Flora, M. Feinholz, M. Yarbrough, C. C. Trees, M. S. Twardowski, and D. K. Clark (2010), An example crossover experiment for testing new vicarious calibration techniques for satellite ocean color radiometry, *J. Atmos. Oceanic Technol.*, *27*(10), 1747–1759.

Wang, P., E. Boss, and C. S. Roesler (2005), Uncertainties of inherent optical properties obtained from semianalytical inversions of ocean color, *Appl. Opt.*, *44*(19), 4074–4085.

Wei, J., and Z. P. Lee (2013), Model of the attenuation coefficient of daily photosynthetically available radiation in the upper ocean, *Methods Oceanogr.*, *8*, 56–74.

Wei, J., and Z. P. Lee (2015), Retrieval of phytoplankton and color detrital matter absorption coefficients with remote sensing reflectance in an ultraviolet band, *Appl. Opt.*, *54*(4), 636–649.

Wei, J., R. Van Dommelen, M. R. Lewis, S. McLean, and K. J. Voss (2012), A new instrument for measuring the high dynamic range radiance distribution in near-surface sea water, *Opt. Express*, *20*(24), 27,024–27,038.

Wei, J., M. R. Lewis, R. Van Dommelen, C. J. Zappa, and M. S. Twardowski (2014), Wave-induced light field fluctuations in measured irradiance depth profiles: A wavelet analysis, *J. Geophys. Res.*, *119*, 1344–1364, doi:10.1002/2013JC009572.

Wei, J., Z. P. Lee, M. Lewis, N. Pahlevan, M. Ondrusek, and R. Armstrong (2015), Radiance transmittance measured at the ocean surface, *Opt. Express*, *23*(9), 11,826–11,837.

Wei, J., Z. P. Lee, M. Ondrusek, A. Mannino, M. Tzortziou, and R. Armstrong (2016), Spectral slopes of the absorption coefficient of colored dissolved and detrital material inverted from UV-visible remote sensing reflectance, *J. Geophys. Res.*, *121*, 1953–1969, doi:10.1002/2015JC011415.

Werdell, P. J., and S. W. Bailey (2005), An improved bio-optical data set for ocean color algorithm development and satellite data product validation, *Remote Sens. Environ.*, *98*, 122–140.

Zibordi, G., S. B. Hooker, J. F. Berthon, and D. D'Alimonte (2002), Autonomous above-water radiance measurements from an offshore platform: A field assessment experiment, *J. Atmos. Oceanic Technol.*, *19*(5), 808–819.

Zibordi, G., D. D'Alimonte, and J. F. Berthon (2004), An evaluation of depth resolution requirements for optical profiling in coastal waters, *J. Atmos. Oceanic Technol.*, *21*, 1059–1073.

Zibordi, G., J.-F. Berthon, F. Mélin, D. D'Alimonte, and S. Kaitala (2009a), Validation of satellite ocean color primary products at optically complex coastal sites: Northern Adriatic Sea, Northern Baltic Proper and Gulf of Finland, *Remote Sens. Environ.*, *113*, 2574–2591.

Zibordi, G., et al. (2009b), AERONET-OC: A network for the validation of ocean color primary products, *J. Atmos. Oceanic Technol.*, *26*(8), 1634–1651, doi:10.1175/2009JTECHO654.1.