Journal:	*Applied Optics*
Article Title:	Exploring the limits for sky and sun glint correction of hyperspectral above-surface reflectance observations
Authors:	Philipp M. M. Groetsch, Robert Foster, and Alexander Gilerson
Accepted for publication:	14 February 2020
Final version published:	20 March 2020
DOI:	https://doi.org/10.1364/AO.385853

Exploring the limits for sky and sun glint correction of hyperspectral above-surface reflectance observations

PHILIPP M. M. GROETSCH,[1,*] ROBERT FOSTER,[2] ALEXANDER GILERSON[1]

[1]*The City College of New York, CUNY, New York, USA*
[2]*Remote Sensing Division, Naval Research Laboratory, Washington D.C., USA*
[*]*pgrotsch@ccny.cuny.edu*

Abstract: Above-surface radiance observations of water need to be corrected for reflections on the surface to derive reflectance. The three component glint model (3C) [Opt. Express **25**, A742 (2017)] was developed to spectrally resolve contributions of sky and sun glint to the surface-reflected radiance signal $L_r(\lambda)$, also for observations recorded at high wind speed and with fixed-position measurement geometries that frequently lead to significant sun glint contributions.

Performance and limitations of 3C are assessed for all relevant wind speeds, clear sky atmospheric conditions, illumination/viewing geometries, and sun glint contamination levels. For this purpose, a comprehensive set of $L_r(\lambda)$ spectra was simulated with a spectrally resolved sky radiance distribution model and Cox-Munk wave slope statistics. Reflectances were also derived from an extensive 4-year data set of continuous above-surface hyperspectral observations from the Long Island Sound Coastal Observatory (LISCO), allowing to corroborate 3C processing results from simulations and measurements with regards to sky and sun glint contributions.

Simulation- and measurement-derived $L_r(\lambda)$ independently indicate that spectral dependencies of the sky light distribution and sun glint contributions may not be neglected for observations recorded at wind speeds exceeding 4 m/s even for sun glint-minimizing measurement geometries (Sun-sensor azimuth angle $\Delta\phi = 90 - 135°$). These findings are in accordance with current measurement protocols for satellite calibration/validation activities. In addition, it is demonstrated that 3C is able to reliably derive water reflectance for wind speeds up to 8 m/s and $\Delta\phi > 20°$.

1. Introduction

Oceans and lakes are undergoing changes in their biogeochemical composition at an increasing pace due to effects of climate change and other human impacts [1]. An in-depth understanding of the underlying cause-and-effect chains is prerequisite to sustainable stewardship of these ecosystems [2]. Such an understanding crucially depends on timely and accurate observations of relevant biogeochemical parameters and their dynamics [3]. Satellite remote sensing is uniquely equipped for this task, observing the water color at ever-increasing spatio-temporal and spectral resolution [3]. However, calibration and validation of water color-derived data products from satellite imagery require high quality ground-based observations [4]. Also other applications, such as timely warnings of harmful algal blooms or regional water quality reporting, lean on ground-based observations, i.e. when overcast conditions prevail or when high spatial or temporal resolution requirements can not be met by satellite observations alone [5, 6].

However, accurately determining water color from near-surface observations, rather than from space, is a tricky task. Waves and ripples cause highly variable surface reflections of the sky and potentially the Sun, which must be accounted for in order to accurately derive spectral signatures of the water body [7]. In this context, it is convenient to express water color in terms of remote

sensing reflectance $R_{rs}(\lambda)$, which is defined as the water-leaving radiance $L_w(\lambda)$ normalized by the planar downwelling irradiance $E_d(\lambda)$ to account for variable illumination conditions. Unfortunately, it is not possible to directly measure $L_w(\lambda)$ from above water in free space (sky-light blocking approaches can mitigate this fundamental issue for floating instruments [8]), since surface-reflected radiance $L_r(\lambda)$ contributes to the observed total radiance $L_t(\lambda)$:

$$R_{rs}(\lambda) = \frac{L_t(\lambda)}{E_d(\lambda)} - \frac{L_r(\lambda)}{E_d(\lambda)}. \quad (1)$$

For a perfectly flat water surface, $L_r(\lambda)$ is simply the reflection of the sky radiance in the specular direction of the water-observing sensor, $L_s(\lambda)$, scaled by the Fresnel reflectance factor ρ_f. While ρ_f depends primarily on viewing zenith angle θ_v and the refractive index of water n_w, matters are more complicated for a wind-roughened water surface.

Diffuse sky radiance is spectrally distinct from direct sun light and several orders of magnitude lower in intensity [9]. In comparison to these contrasts, the diffuse sky light distribution offers little variation with regards to spectral composition and magnitude. If sun glint can be assumed negligible, the net effect of reflected diffuse sky radiance on individual wave facets may therefore be approximated with an effective diffuse sky reflectance factor ρ to scale $L_s(\lambda)$ observations [10]. Various approaches maintain $L_s(\lambda)$ as a spectral approximation for $L_r(\lambda)$ and focus instead on estimating ρ, e.g. [5, 11], with [10] as the probably most widely used scheme (further referred to as M99) providing values for the diffuse sky reflectance factor ρ for a range of wind speeds and viewing geometries:

$$R_{rs}(\lambda) = \frac{L_t(\lambda)}{E_d(\lambda)} - \rho \cdot \frac{L_s(\lambda)}{E_d(\lambda)} - \Delta, \quad (2)$$

with scalar reflectance offset Δ to account for residual sun glint and potential white cap contributions [8]. This offset can be determined for open ocean situations (case 1) from reflectance in the near-infrared (NIR) spectral region. For optically complex waters (case 2), where NIR reflectance can not be neglected, various approaches exist to approximate Δ, e.g. [8, 12]. For highly turbid waters, the shape of $R_{rs}(\lambda)$ in the NIR spectral region can be anticipated and used for quality control purposes [13, 14].

M99 recommends a specific viewing geometry to strike a balance between minimizing the impact of sun glint, which is not accounted for with the approach, and low $L_r(\lambda)$ contributions that are stable over a large range of illumination conditions. This "Mobley geometry", i.e. observing $L_t(\lambda)$ and $L_s(\lambda)$ at $\theta_v = 40°$ from nadir and zenith, respectively, with an azimuth difference between 90° and 135° (further referred to as $\Delta\phi = 90 - 135°$) to the Sun, endured two decades of scrutiny to be recommended also for the most demanding applications, i.e. satellite product calibration and validation [4, 15]. However, M99 and all other approaches that derive $R_{rs}(\lambda)$ in the context of Eq. (2) are built on the assumption of a spectrally isotropic diffuse sky radiance distribution, i.e. sky light only changes in intensity, not color [10, 16]. Several studies have pointed out that each wave facet of a wind-roughened sea surface reflects a different part of the sky and potentially the Sun, both of which deviate in intensity and spectral shape from $L_s(\lambda)$ [8, 9, 17–19]. A spectral dependency may therefore be assigned to the reflectance factor ρ to justify the use of $L_s(\lambda)$ in Eq. (2). This dependency is primarily a function of atmospheric composition and sea state [9], with polarization adding further complexity to its estimation [11, 17, 20]. Furthermore, for a given wind speed, M99 supplies only a single ρ value to remove surface glint from the highly variable water-leaving signal. For a typical measurement sequence, variability in $L_t(\lambda)$ thus directly translates to the uncertainty budget of retrieved $R_{rs}(\lambda)$. Averaging schemes have been proposed to mitigate this direct relationship, e.g. considering only the lowest 20% of a $L_t(\lambda)$ measurement series when processing to $R_{rs}(\lambda)$ with M99 [15] (further referred to as "L_t^{rel}"), assuming sun glint contamination for the rejected observations. While this empirical approach considerably reduces estimated uncertainty levels, it also artificially constricts Cox-Munk wave slope statistics [7] that are core to the M99 estimates for ρ, and was

thus restricted to wind speeds below 4 m/s in the context of high accuracy requirements (satellite calibration/validation) [15].

A range of natural phenomena in aquatic ecosystems exhibit distinct temporal patterns that can be revealed with consistent and continuous $R_{rs}(\lambda)$ observations, e.g. tidal effects [21]; BRDF patterns [22]; phytoplankton fluorescence yield variability [23]; phytoplankton photoprotective mechanisms [24]; phytoplankton functional types [25]; colored dissolved organic matter (CDOM) composition [26]. However, relying on established measurement protocols and correction schemes that artificially limit the viable range of relative viewing azimuth directions, such consistent diurnal $R_{rs}(\lambda)$ time series may often not be derived from variable- and especially fixed-direction radiometric data sets. Sampling from expansive platforms such as research vessels or moored coastal observatories poses additional challenges, e.g. constricted potential viewing azimuth directions [5], shading [27] and adjacency effects [28, 29], and altered wave statistics due to wind-shading or when observing underway. These and other factors typically lead to the rejection of vast amounts of valuable observations [5] during quality control, hampering full leverage of such data sets.

Recently, a processing scheme was proposed to mitigate the impact of spectrally variable water surface reflections on derived remote sensing reflectance. The "3C method" [18] comprises a full-spectrum inversion to spectrally separate the water signal from surface-reflected sky and sun glint components. This separation is carried out for each individual observation in a measurement sequence, which greatly increases the precision of derived $R_{rs}(\lambda)$ without resorting to empirical averaging schemes like L_t^{rel}. A byproduct of each 3C inversion is the optimization residual ϵ, which can be used in conjunction with precision thresholds to effectively flag measurement sequences that may not be adequately corrected with 3C. The scheme has been validated for a broad range of water bodies, illumination conditions, and measurement geometries [18, 30–32]. For ideal measurement conditions, i.e. clear skies, low wind speed, and optimized viewing geometries, 3C was shown to perform equally to established methodologies, based on a number of comparisons between reflectances derived from above- and sub-surface [18]. In addition, it was demonstrated that 3C allows the retrieval of relevant reflectance spectra also for measurements collected in sub-optimal conditions, i.e. scattered clouds, rough water surfaces, and Sun-facing measurement directions [18]. In clear sky conditions, 3C also allows to derive $R_{rs}(\lambda)$ when sky radiance observations are lacking, enabling straightforward processing of low-altitude airborne observations, i.e. from drones and planes [30]. These capabilities may greatly increase the amount of viable observations available to downstream applications, especially in the context of autonomous, high-frequency observations. However, a comprehensive assessment of error budgets attached to 3C-retrieved reflectances is still lacking, specifically with regards to sun-glint prone measurement geometries and high wind speeds, i.e. the conditions for which other established methodologies are not applicable.

In this study, we assess the conditions under which 3C is capable of reliably separating sky and sun glint from the water-leaving signal. Reliable parameter retrievals are both accurate and precise, i.e. their distribution is centered close to the true value and features low variance. Retrieval accuracy and precision are thus examined separately, based on simulations for which the true value is known, and field measurements for which realistic variances can be observed. In both assessments, we focus on observation geometry and wind speed as the main drivers for sky and sun glint contamination.

In a first step, $R_{rs}(\lambda)$ retrieval accuracy is estimated from a simulated validation data set of above-water $L_t(\lambda)$ observations. The simulations are representative for waters in the Western Long Island sound and cover all relevant clear sky atmospheric conditions, observation geometries, and water surface roughness levels.

In a second step, $R_{rs}(\lambda)$ retrieval precision is examined for an extensive 4-year data set of continuous above-surface hyperspectral observations from the Long Island Sound Coastal

Observatory (LISCO). During each cycle, measurements were conducted in burst mode for several minutes and thus yield a representative sample of $L_t(\lambda)$ variability. Precision of $R_{rs}(\lambda)$ retrievals was assessed for wind speeds ranging up to 10 m/s at all relevant observation geometries [33] and thus a wide range of sky and sun glint contamination.

In support of airborne and satellite applications, it is further discussed to which extent $L_s(\lambda)$ measurements are required and beneficial for accurate and precise $R_{rs}(\lambda)$ retrievals.

2. Materials and Methods

2.1. Simulated Data Set

A wind-roughened water surface is composed of wave facets that each reflect a different part of the sky, and potentially the Sun, into the solid angle that comprises the sensor field of view Ω_v:

$$L_t(\lambda, \Omega_v) = L_w(\lambda, \Omega_v) + \int_{2\pi} p(\Omega \rightarrow \Omega_v) \cdot \rho_f(\Omega \rightarrow \Omega_v) \cdot L_s(\lambda, \Omega)\, d\Omega, \tag{3}$$

with $\Omega \rightarrow \Omega_v$ indicating the solid angle of a sky patch reflected into Ω_v and p as the wave slope probability for that direction [9]. Wave statistics were calculated based on the slope distributions first described by [7], with updated Gram-Charlier series coefficients from [34]. Dependencies on wind direction are weak according to Cox-Munk wave statistics [7] and were not investigated here, i.e. wind was assumed to be directed towards the sun azimuth position in all simulations.

The spectral sky dome radiance distribution was calculated with an analytical model for sky [35] and sun light [36], further referred to as "HW12". HW12 is a fully analytical model evolved from earlier works by [37] with spectral output at 11 wavelengths ranging from the UV to the red (320 – 720 nm), accounting for all scattering orders, and ingesting ground albedo to describe adjacency effects. [38] compared 8 clear sky models and attested HW12 a generally good spectral fit to their Radtran reference implementation, however, with a scaling error that was later corrected for in the here applied model reference implementation (version 1.4a, link).

The sky light distribution was queried from the HW12 model for N randomly selected zenith (θ ranging from 0 to 90°) and azimuth angles (ϕ ranging from 0 to 180°), leveraging symmetries. The number of queries $N = 92270$ was chosen such that the area covered by each direction equals the solid angle of the sun disc $\Omega_{sun} = 6.8096 \cdot 10^{-5} sr$ to sample sun glint representatively. Sun disc radiance was calculated with a separate routine, sampling the complete disc as recommended by [36]. Input parameters for HW12 are sun zenith angle θ_{sun}, atmospheric turbidity, and bottom albedo. In this study, adjacency effects were neglected by setting bottom albedo to zero. Atmospheric turbidity as defined in HW12 is represented by empirical relationships between aerosol optical depth and slope with values relevant to this study between 2 and 10, representing very clear to slightly hazy skies [35].

A comprehensive database of surface reflectance spectra, $L_r(\lambda)/E_d(\lambda)$, was populated following Eq. (3) for a fixed viewing angle $\theta_v = 40°$ and parameter ranges as follows: atmospheric turbidity ranging from 2 to 10 (5 steps); θ_{sun} ranging from 0 to 75° (6 steps); $\Delta\phi$ ranging from 0 to 180° (13 steps); v_w ranging from 0 to 15 m/s (16 steps). Total downwelling irradiance $E_d(\lambda)$ was calculated by cosine-weighting individual radiance distributions in a numerical integration over the complete radiance distribution. The analytical HW12 model was chosen with runtime quality control in mind, which could not be achieved with comparatively slow ray-tracing or Monte Carlo-based approaches. HW12 is limited to multi spectral output, which is not considered a limitation here, as small-scale spectral absorption features are absent in $R_{sky}(\lambda) = L_s(\lambda)/E_d(\lambda)$ spectra evaluated here. The multi-spectral $L_r(\lambda)/E_d(\lambda)$ database that resulted from HW12 model runs was extrapolated to the relevant 350-900 nm range at 1 nm spectral resolution, following the $R_{sky}(\lambda)$ model inversion procedure described in [28] and neglecting adjacency effects, resulting

in highly satisfactory model fits for all considered situations (mean of absolute percent differences $|\psi| = 1.05 \pm 0.87\,\%$).

2.2. Measured Data Set

The Long Island Sound Coastal Observatory (LISCO) is located on an offshore platform at N40.57016°, W73.20030°. A Satlantic HyperSAS (Halifax, Canada) system had been installed from 2009 to 2014 at LISCO on top of a retractable tower at approximately 12 m above the water surface. Three spectrometers observed downwelling irradiance $E_d(\lambda)$, sky radiance $L_s(\lambda)$, and total water-leaving radiance $L_t(\lambda)$ in the wavelength range 305 to 905 nm with 180 equally spaced channels. The $L_t(\lambda)$ and $L_s(\lambda)$ radiance sensors had a 3° full-angle field-of-view and were oriented towards the West, observing water and sky at 40° from nadir and zenith, respectively. This fixed observation geometry caused the relative azimuth angle $\Delta\phi$ between Sun and radiance sensor to be variable and high sun glint contributions are generally to be expected towards the afternoon, depending on the time of the year [33]. Sensor calibrations were initially performed by Satlantic Inc. (Halifax, Canada) and regularly tracked at the CCNY Optical Remote Sensing Lab. The highly productive, phytoplankton and CDOM dominated waters [39] can safely be assumed optically deep in the vicinity of the platform and a water depth of approximately 15 m, depending on the tide. Further details on installation, setup, and calibration are briefly summarized below and an in-depth review of the data set and comparison to AERONET-OC observations recorded from the same platform can be found in [33].

Measurements were performed every 30 minutes during daytime, with all sensors recording for approximately 70 seconds in burst mode. Depending on the illumination conditions, 26 to 47 (25% and 75% percentiles, respectively) individual observations were acquired on average during each measurement cycle, which is considered a representative sample of variability in $L_t(\lambda)$. For the following analysis, 17294 measurement cycles during the period 1 January 2011 to 11 April 2014 were considered, amounting to a total of 758396 reflectance observations. All radiometric observations were screened for transmission errors, dark current and non-linearity corrected, and converted to physical units with the processing software Satlantic SatCon. All observations were then processed to remote sensing reflectance spectra with the 3C method [18] and subsequently quality controlled, as detailed in the following section.

Wind speed was derived from Copernicus Climate Change Service Information 2019, ECMWF ERA-5 reanalysis data set [40] and hourly wind speed components for 10 m above the ground were linearly interpolated in time to match measurement time stamps.

2.3. Reflectance Processing

The 3C method allows to separate signal contributions from water surface reflections and the water itself by means of full-spectrum inversion. 3C is applied here following the recommendations in [18], using the 3C Python implementation (link) that is publicly available under the LGPLv3 license. Remote sensing reflectance is derived from above-surface measurements of $L_t(\lambda)$, $L_s(\lambda)$, and $E_d(\lambda)$:

$$R_{rs}(\lambda) = \frac{L_t(\lambda)}{E_d(\lambda)} - \rho \cdot \frac{L_s(\lambda)}{E_d(\lambda)} - \Delta(\lambda), \tag{4}$$

with the term $\rho \cdot \frac{L_s(\lambda)}{E_d(\lambda)}$ representing a first-order correction for sky glint reflections on the water surface that is exact for a perfectly flat sea and $\rho = \rho_f$. All other symbols are listed in Table 1. For wind-roughened water surfaces, parts of the sky and potentially the Sun are also contributing to the observed water radiance [9]. This effect is accounted for with a spectrally resolved offset $\Delta(\lambda)$, which is derived for every measurement in a sequence during 3C model inversion. An analytical downwelling irradiance model [41] was adapted to spectrally resolve the offset parameter $\Delta(\lambda)$ in

Eq. (4):

$$\Delta(\lambda) = \quad \rho_{dd} \cdot \frac{E_{dd}(\lambda)}{E_d(\lambda)\pi} + \rho_{ds} \cdot \left[\frac{E_{dsr}(\lambda)}{E_d(\lambda)\pi} + \frac{E_{dsa}(\lambda)}{E_d(\lambda)\pi} \right], \tag{5}$$

with ρ_{dd} and ρ_{ds} as the reflectance factors for downwelling direct irradiance $E_{dd}(\lambda)$, downwelling Rayleigh-scattered diffuse irradiance $E_{dsr}(\lambda)$, and downwelling aerosol-scattered diffuse irradiance $E_{dsa}(\lambda)$. The bio-optical model currently implemented in 3C [42] was developed for optically complex waters and was initialized to reflect the findings of [39] for the Long Island Sound. Specifically, CDOM absorption slope was set to $S = 0.0147$ and suspended particulate matter (SPM) backscattering was modeled with a power law $b_b^{sm} = SPM \cdot 0.0042 \cdot (\lambda/\lambda_0)^{\gamma}$, with $\lambda_0 = 500\,nm$ and $\gamma = 0.23$, in contrast to the spectrally flat b_b^{sm} applied in [18]. Directionality effects due to variable viewing and sun zenith angles are implicitly taken into account when translating sub-surface reflectance to above-surface $R_{rs}(\lambda)$. Table 1 lists all model parameter starting values and search ranges, including the water constituent concentrations that were also initialized according to [39]. The constrained limited memory Broyden–Fletcher–Goldfarb–Shanno algorithm (L-BFGS-B) [43] was used to minimize the residual-sum-of-squares ϵ between observed $L_t(\lambda)/E_d(\lambda)$ and modeled $L_t^m(\lambda)/E_d^m(\lambda)$ with the bounding parameters listed in Table 1 and weighting function $W(\lambda)$. The residual ϵ was calculated as

$$\epsilon = \sum^{i} \left(\frac{L_t(\lambda_i)}{E_d(\lambda_i)} - \frac{L_t^m(\lambda_i)}{E_d^m(\lambda_i)} \right)^2 \cdot W(\lambda_i), \tag{6}$$

and the spectral weighting function $W(\lambda) = 1$, except for the UV to visible (350-500 nm) and NIR (800-900 nm) spectral regions where $W(\lambda) = 5$. The spectral region around the deep oxygen A-band rotational absorption lines were assigned a lower spectral weight $W(760-775\text{nm}) = 0.1$. The fit residual ϵ is calculated in conjunction with $W(\lambda)$, which is beneficial for rapid optimization convergence, however, it is a biased indicator of fit quality. In support of assessment and interpretation of fit residuals, root-mean-square-deviation RMSD were also calculated:

$$RMSD = \sqrt{\sum^{i} \left(\frac{L_t(\lambda_i)}{E_d(\lambda_i)} - \frac{L_t^m(\lambda_i)}{E_d^m(\lambda_i)} \right)^2}. \tag{7}$$

Chlorophyll a pigment fluorescence was accounted for in the bio-optical model with a fluorescence reflectance component $R_{rs}^{fl}(\lambda)$, based on [23, equation 7.36]:

$$R_{rs}^{fl}(\lambda) = L_{fl}(685\,nm) \cdot \exp(-4 \cdot \ln(2) \cdot [(\lambda - 685\,nm)/25]^2)/1.1\,Wm^{-2}nm^{-1}. \tag{8}$$

Fluorescence radiance $L_{fl}(685\,nm)$ was treated as a free parameter in the bio-optical model inversion. Accurate inversion of this parameter was enhanced by setting spectral weights $W(660 - 710\,nm) = 2$.

3. Results

The full-spectrum inversion of the simulated $L_t(\lambda)$ data set to $R_{rs}(\lambda)$ is presented first. Following, 3C processed $R_{rs}(\lambda)$ for four years of continuous observations at LISCO are summarized.

3.1. Inversion of simulated data set

The simulated $L_t(\lambda)$ validation data set was derived according to Eq. (1), based on a comprehensive set of $L_r(\lambda)$ spectra (see section 2.1) and twelve $R_{rs}(\lambda)$ spectra representing monthly means of quality controlled (Level L3, see Table 2 and accompanying text in Section 3.2) and spectrally smoothed (boxcar, 10 nm window size) reflectances at LISCO. Level L3 quality control restricts $\Delta\phi$

Table 1. Free model parameters (upper section) in 3C model optimization. Values marked with an asterisks are modified for the 3C parameterization making no use of $L_s(\lambda)$ observations, see text. The lower section lists further reoccurring parameter names and acronyms that are used in this manuscript. Units for radiances ($mWm^{-2}nm^{-1}sr^{-1}$) and irradiances ($mWm^{-2}nm^{-1}$) are abbreviated with [1] and [2], respectively.

Parameter	Acronym, [unit]	Start (min/max)
Concentration of chlorophyll-*a*	Chla, $[mg\,m^{-3}]$	7.1 (0.01/100)
Concentration of suspended particulate matter	SPM, $[g\,m^{-3}]$	2.9 (0/100)
Colored dissolved organic matter absorption at 440 *nm*	CDOM, $[m^{-1}]$	0.3 (0.1/1)
Fluorescence radiance (685 *nm*)	$L_{fl}(665\,nm)$, [1]	1 (0/3)
Air-water interface reflectance factor	ρ, [−]	ρ_f^* (0/ρ_f)
3C direct reflectance factor	ρ_{dd}, [−]	0.001 (0/4)
3C diffuse reflectance factor	ρ_{ds}, [−]	0.01 (-0.01/1)
Ångström exponent	α, [−]	1 (0/5)
Aerosol optical thickness (550 *nm*)	β, [−]	0.05 (0/1)
Wavelength	λ, $[nm]$	
Sun zenith angle	θ_{sun}, [°]	
Viewing zenith angle	θ_v, [°]	
Relative azimuth angle to Sun	$\Delta\phi$, [°]	
Wind speed	v_w, $[m\,s^{-1}]$	
Sky radiance	$L_s(\lambda)$, [1]	
Total radiance	$L_t(\lambda)$, [1]	
Water-leaving radiance	$L_w(\lambda)$, [1]	
Water surface-reflected radiance	$L_r(\lambda)$, [1]	
Downwelling vector irradiance	$E_d(\lambda)$, [2]	
Remote-sensing reflectance $L_w(\lambda)/E_d(\lambda)$	$R_{rs}(\lambda)$, $[sr^{-1}]$	
Sky reflectance $L_s(\lambda)/E_d(\lambda)$	$R_{sky}(\lambda)$, $[sr^{-1}]$	
Air-water interface reflectance $L_r(\lambda)/E_d(\lambda)$	$R_{refl}(\lambda)$, $[sr^{-1}]$	
Fresnel reflectance factor (unpolarized)	ρ_f, [−]	
Residual-sum-of-squares	ϵ, $[sr^{-1}]$	
Root-mean-square deviation	$RMSD$, $[sr^{-1}]$	
Normalized RMSD	$NRMSD$, [%]	
Percent Difference	ψ, [%]	

to 90-135°, for which 3C results can be considered equivalent to established methodologies [18]. Reference $R_{rs}(\lambda)$ spectra are presented in Figure 1. This simulated data set was then inverted to assess $R_{rs}(\lambda)$ retrieval performance. Two 3C parameterizations and a reference model were considered in the inversion:

1. 3C with $L_s(\lambda)$
   Standard 3C parameterization according to Eq. (4).

2. 3C without $L_s(\lambda)$
   Spectrally resolved 3C-offset $\Delta(\lambda)$ only, disregarding $L_s(\lambda)$ spectra by setting $\rho = 0$ in Eq. (4).

3. Reference model
   According to Eq. (2) and [8], i.e. scaling $L_s(\lambda)$ with scalar reflectance factor ρ and offset Δ, which are determined with bio-optical model inversion [8]. The same bio-optical $R_{rs}(\lambda)$ model and parameter bounds as in 3C were used in the inversion procedure.

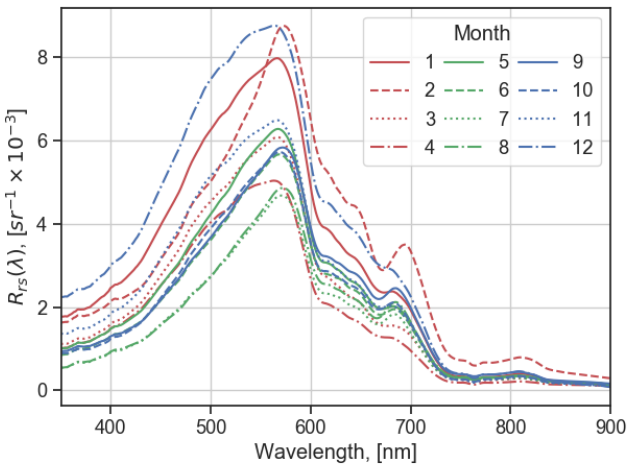

Fig. 1. Monthly averages of quality controlled (level L3, see Table 2) remote sensing reflectance $R_{rs}(\lambda)$ for the period 2011-2014 at the LISCO site. The optically complex site features a significant and variable offset in the near-infrared spectral region.

Two representative $R_{rs}(\lambda)$ inversion results are shown in Figure 2 for $\Delta\phi = 135°$ (Mobley geometry) and 20° (sun glint-prone geometry), featuring $L_r(\lambda)$ spectra that are dominated by sky and sun glint, respectively. Since $R_{rs}(\lambda)$ reference spectra were known, $\psi[R_{rs}(\lambda)]$ could be calculated for each inversion. Three spectral ranges were considered: ultra-violet (UV, 350-450 *nm*), visible (VIS, 450-750 *nm*), and near-infrared (NIR, 750-900 *nm*). The fit bias was calculated for each model inversion as the mean of signed percent differences $\psi[R_{rs}(\lambda)]$ between reference $R_{rs}(\lambda)$ and their modeled approximations $R^m_{rs}(\lambda)$:

$$\psi[R_{rs}(\lambda)] = \sum^{i} \frac{R_{rs}(\lambda_i) - R^m_{rs}(\lambda_i)}{R_{rs}(\lambda_i)}. \tag{9}$$

Figure 3 shows ψ as a function of wind speed and relative azimuth angle for each correction scheme and the UV, VIS, and NIR spectral regions. Both the reference model and 3C (w/ $L_s(\lambda)$) tend to overestimate $R_{rs}(\lambda)$ in all wavelength ranges, i.e. apply too low a glint correction for

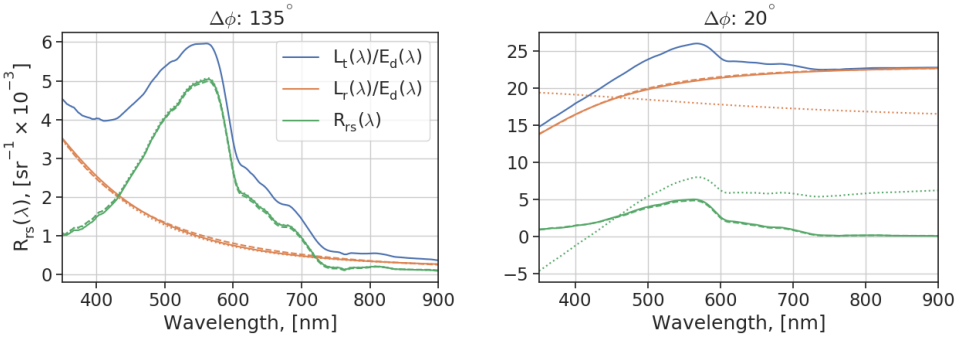

Fig. 2. Examples for inversion results of simulated $L_t(\lambda)/E_d(\lambda)$ spectra (continuous lines), using the 3C configuration that makes no use of $L_s(\lambda)$ spectra (dashed) and the reference model (dotted). Left panel: equal performance of both models at ideal relative azimuth angle $\Delta\phi = 135°$. Right panel: Sun-facing viewing geometry at $\Delta\phi = 20°$, causing considerable sun glint and large bias in reference model retrievals. The $R_{rs}(\lambda)$ reference spectrum is that of the month April at LISCO, see Figure 1. Simulations were carried out for a wind speed of 4 m/s and a sun zenith angle $\theta_{sun} = 30°$.

relative azimuth angles larger than approximately 45°. This trend increases with wind speed and is most pronounced in the UV. For the Mobley geometry ($\Delta\phi = 90 - 135°$) the reference model consistently yields $\psi > 7\%$ from 5 m/s in the UV, while 3C (w/ $L_s(\lambda)$) reaches that threshold only at 10 m/s or remains below it (w/o $L_s(\lambda)$). Small relative azimuth angles in combination with wind-roughened water surfaces lead to increasing sun glint contributions. The reference model has no means to correct for sun glint, leading to very large fit errors in all spectral ranges for $\Delta\phi < 45°$. Please note that ρ was determined in the optimization of the reference model, which reduced errors considerably compared to $\rho = \rho_f$ as suggested in [8] (results not shown). In sun glint-dominated conditions, $L_r(\lambda)$ can be several times the $L_w(\lambda)$ signal (i.e. see Figure 2), yet 3C (both schemes) performs surprisingly well at all wind speeds, i.e. $\psi = 0 - 2$ % except for the visible spectral range where ψ reaches 7 % for $\Delta\phi < 20°$. Trends in ψ with regards to sun zenith angle, and atmospheric conditions were generally small in comparison to dependencies on $\Delta\phi$ and wind speed, and have already been discussed exhaustively elsewhere [9]. However, in the ultra-violet spectral range, larger negative biases were observed for very clear skies than for hazy conditions. This is likely to be caused by high UV radiation levels due to Rayleigh-scattering at low aerosol concentrations. The 3C parameterization without $L_s(\lambda)$ resulted in a significant bias ($\psi > 10\,\%$) for observations recorded at $\theta_{sun} \geq 60°$, which should be excluded from processing in this configuration. No consistent trend was found with regards to $R_{rs}(\lambda)$ reference spectrum, indicating that the bio-optical model parameterization presented in section 2.3 is adequate to capture the range of $R_{rs}(\lambda)$ variability at LISCO (see Figure 1), including the non-zero offset in the near-infrared spectral region. It was further investigated whether 3C fit residuals (e.g. RMSD) may serve as a quantitative measure for $R_{rs}(\lambda)$ retrieval bias (e.g. ψ), however, correlations between RMSD and ψ for the 12 considered $R_{rs}(\lambda)$ spectra (Figure 1) were found to be inconsistent. Despite the lack of clear linear correlation, large residuals were generally indicative of large retrieval biases. This observation supports the use of 3C fit residual thresholds as a means of quality control, as further elaborated in the next section.

3.2. Inversion and quality control of LISCO data set

Four years (2011-2014) of radiometric observations at LISCO were processed with 3C (with and without $L_s(\lambda)$ observations) and subsequently filtered to adhere to a range of quality levels.

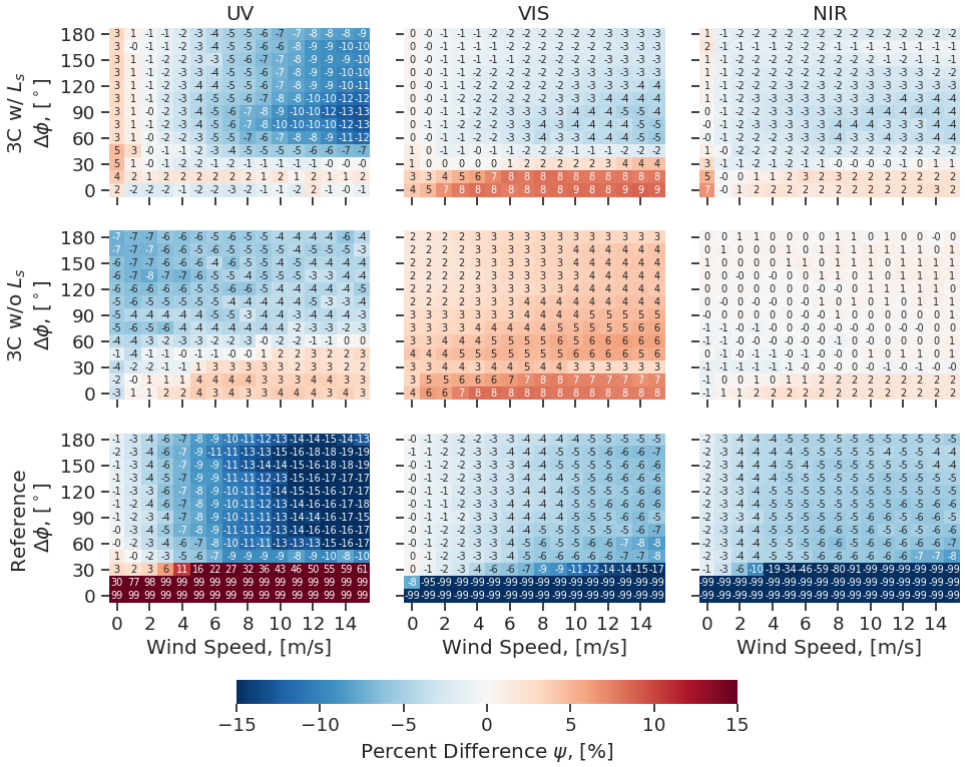

Fig. 3. Fit bias as a function of wind speed and relative viewing azimuth angle $\Delta\phi$. Bias is shown as signed percent differences $\psi[R_{rs}(\lambda)]$, averaged for atmospheric compositions, sun zenith angles, and reference reflectance spectra (see Figure 1) and were determined for three correction schemes (rows: 3C without using $L_s(\lambda)$ observations, 3C using $L_s(\lambda)$ observations, and the reference implementation) and spectral ranges (columns: UV [350-450 *nm*], VIS [450-750 *nm*], NIR [750-900 *nm*]). Values exceeding ±99 % were clipped.

Quality control was carried out on a measurement cycle basis, with each cycle consisting of a number of individual observations, which were collected in burst-mode and within several minutes. Expressions that were evaluated are listed in Table 2, along with percentages of affected measurement cycles for each flag. From these results, distinct quality levels were derived, which can be summarized as follows:

- minimal (L0): basic radiometric consistency checks and 3C model convergence
- intermediate (L1): L0 & restricted viewing geometry, limited variance on bio-optical parameter retrievals, representative $L_s(\lambda)$ observations
- intermediate/clear sky (L2): L1 & clear skies
- strict (L3): L2 & restricted relative azimuth $\Delta\phi = 90 - 135$.

Most quality control flags are based on empirical thresholds that were derived from parameter histogram analysis and manual spectrum inspection of the LISCO data set, however, several flags may require additional explanation:

Table 2. Observation cycles that raised individual quality control flags, in percent and for each quality control level (L0 to L3). Dashes (–) indicate flags that were not allowed to be raised to pass a given quality control level. Please refer to Section 3.2 for further explanations on quality levels. The last row lists the total counts and percentages of cycles that passed each quality control level. Symbols are according to Table 1, in addition to the coefficient of variation (c_v), standard deviation (σ), 5 %-percentile (P5), maximum value (^{max}), and oxygen-A absorption bands (O_2). Each observation cycle consists of several individual observations that had been recorded in rapid succession. Flags were raised if any observation in a cycle was evaluated as true with the respective expression, except if indicated with an asterisks (*), where the median was used. Wavelength-averaged radiometric parameters were used except where specific wavelength bands are provided.

Flag	**Expression**	**all**	**L0**	**L1**	**L2**	**L3**
few measurements	# Measurements < 10	0.21	–	–	–	–
high θ_{sun}	$\theta_{sun} > 80°$	19.21	2.38	–	–	–
low $\Delta\phi$	$\Delta\phi < 20°$	18.24	13.13	–	–	–
sub-optimal $\Delta\phi$	$\Delta\phi < 90°$ or $\Delta\phi > 135°$	82.98	78.23	72.85	70.49	–
negative $L_s(\lambda)$	$L_s(\lambda) < 0$	7.92	–	–	–	–
negative $L_t(\lambda)$	$L_t(\lambda) < 0$	18.82	–	–	–	–
negative $E_d(\lambda)$	$E_d(\lambda) < 0$	6.42	–	–	–	–
low light	$E_d(^{max}) < 0.1\ W/m^2nm$	24.66	–	–	–	–
clouds present	$R_{sky}(750\,nm) > 0.1\ sr^{-1}$	66.24	55.92	41.03	–	–
high variability $L_s(\lambda)$	$c_v(L_s) > 5\%$	35.22	30.99	21.71	–	–
high variability $E_d(\lambda)$	$c_v(E_d) > 5\%$	21.07	15.25	11.25	–	–
Sun in $L_s(\lambda)$	$R_{sky}(750\,nm) > 2\,/\,\pi$	10.63	4.00	–	–	–
high variability $R_{rs}(O_2)$	$\sigma(R_{rs}(O_2)) > 0.0003\ sr^{-1}$	59.06	43.14	27.32	–	–
high variability $R_{rs}(\lambda)$	$c_v(R_{rs}) > 10\%$	28.69	14.86	–	–	–
high variability Chla	$c_v(Chla) > 25\%$	42.38	28.72	–	–	–
high variability SPM	$c_v(SPM) > 25\%$	18.22	4.49	–	–	–
high variability CDOM	$c_v(CDOM) > 25\%$	26.22	12.73	–	–	–
high RMSD	RMSD > 0.002 sr^{-1} (*)	12.86	–	–	–	–
negative $R_{rs}(\lambda)$	P5($R_{rs}(< 750\,nm)$) < 0	11.74	–	–	–	–
	Cycles passed, [#]	*17294*	*11842*	*7149*	*3158*	*932*
	Cycles passed, [%]	*100*	*68.5*	*41.3*	*18.3*	*5.4*

i) The "high RMSD" flag was raised already at the first quality control level if the median RMSD exceeded 0.002 sr^{-1}, indicating that the 3C model optimization did not converge satisfactorily. This threshold is based on the median RMSD+3σ of all measurements that passed the minimal (L0) quality level with suspended "high RMSD" flag. While almost 13 % of all measurements were flagged by this metric, only 2.23 % and 1.30 % of L0 and L1 quality controlled measurements, respectively, were not also affected by other flags. Furthermore, L2 and L3 quality controlled observations were not affected at all. These findings suggest that 3C fit residuals are reliable measures to flag inappropriate measurements, but should be regarded as just one element in a comprehensive quality control scheme.

ii) More than 80 % of all observations were filtered out by the sub-optimal $\Delta\phi$-flag, i.e. enforcing the Mobley geometry. This makes the $\Delta\phi$-flag the single most important flag and is also the only constraint added from L2 to L3 where it still removes more than 70 % of otherwise fully quality controlled observations. This result clearly demonstrates the need for processing schemes like 3C that can reliably process observations also outside the constricting Mobley geometry.

iii) Only approximately a fifth of all observations qualified as clear sky conditions (L2), with more than 55 % of all L0 observations affected by clouds. It was shown previously that 3C can reliably retrieve $R_{rs}(\lambda)$ from overcast and scattered-cloud conditions [18]. For this reason we here focus on clear sky conditions.

iv) Standard deviation in derived L0 $R_{rs}(\lambda)$ at the Oxygen-A absorption bands at 760-775 *nm* ("high variability $R_{rs}(\lambda)(O_2)$") was significantly ($p \ll 1$ %) correlated with (mean) $L_t(\lambda)$ ($R^2 = 0.89$), derived $L_r(\lambda)$ ($R^2 = 0.63$), and RMSD ($R^2 = 0.57$), supporting its use as a quality indicator for intense sky and/or sun glint contamination.

Figure 4 shows two representative examples for 3C processed measurement cycles that passed different quality control levels. Highly variable $L_t(\lambda)$ observations are common to both examples, caused by elevated wind speed (6 m/s, left panel) and high sun glint contributions (right panel), respectively, and resembling the simulated cases shown in Figure 2. In both instances, 3C allowed to attribute the vast majority of this variability to surface reflections and thus to retrieve $R_{rs}(\lambda)$ with high precision. In the following section, we focus on the derived surface reflection component.

3.3. Variability in surface reflections

Variability inherent to the radiometric measurements as a function of wind speed is displayed in Figure 5 for the clear sky quality controlled (level L2, see Table 2) LISCO data set. Variability is presented as coefficient of variation c_v per measurement cycle, averaged over wind speed bins of 2 m/s. Variability for both $L_s(\lambda)$ and $E_d(\lambda)$ remain below 1 % for all considered wind speeds up to 10 m/s. The corresponding "high variability" quality control flag for both parameters capped c_v only from 5 % and is therefore not considered artificially constricting. Variability of $L_t(\lambda)$ increased with wind speed, exceeding the 5 % c_v mark from 4 m/s, and reaching to almost 15 % at 8 m/s, on average. Variability of derived $R_{rs}(\lambda)$ also increases with wind speed, however, at a much lower rate and remains, on average, below 5 % for wind speeds up to 8 m/s. Lacking $L_s(\lambda)$ information in the 3C inversion proved beneficial to retrieving consistent $R_{rs}(\lambda)$, with on average 30 % lower c_v values as compared to retrievals incorporating $L_s(\lambda)$ observations.

These results clearly show that 3C successfully attributed a large part of the observed variability in $L_t(\lambda)$ to reflections at the air-water interface. This is in contrast to look-up-table approaches that yield a constant correction factor for a given wind speed, i.e. M99, and thus directly translate variability in $L_t(\lambda)$ to measurement uncertainty in derived $R_{rs}(\lambda)$. An offset correction may mitigate this issue partially, however, it is not straightforward for optically complex waters like at LISCO that feature non-negligible NIR reflectance. For example, the scalar offset Δ in the reference model (Eq. (2)) had to be determined by bio-optical model inversion [8], essentially matching the processing complexity of 3C. Without offset correction, i.e. the original M99

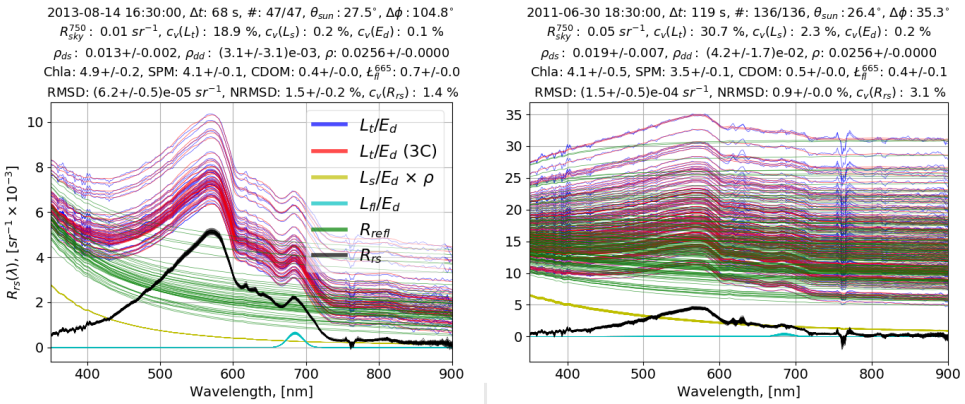

Fig. 4. Processing results of two typical measurement cycles recorded at LISCO on 14 August 2013, UTC 16:30 pm (left panel) and 30 June 2011, UTC 18:30 pm (right panel), with clear sky conditions and approximately 6 and 4 m/s wind speed, respectively. Yellow lines represent a first-order correction with $L_s(\lambda)$, scaled with reflectance factor ρ. Statistics above each panel characterize measurement conditions and were used in quality control (see Table 2). The measurement cycle in the left panel passed the highest quality level (L3), while the right cycle was classified L1 (see Table 2) due to sub-optimal $\Delta\phi$ angles and high variability in the Oxygen-A absorption bands. In both cases, $R_{rs}(\lambda)$ was retrieved with high precision (see $c_v(R_{rs})$). Units for fit parameter results were omitted for readability and are given in Table 1.

approach, much larger biases resulted for the reference model (not shown), which is specifically of concern for coastal and inland waters that feature low reflectance values in the blue and ultra-violet spectral regions. Exactly in those spectral regions 3C offers the largest improvements over the reference model, also when applied in conditions without major sun glint contributions.

Fit residuals (RMSD) for the 3C run with $L_s(\lambda)$ average to $2.95 \pm 3.97 \times 10^{-4}\ sr^{-1}$ (L0), $1.88 \pm 1.68 \times 10^{-4}\ sr^{-1}$ (L1), $1.32 \pm 1.07 \times 10^{-4}\ sr^{-1}$ (L2), and $1.49 \pm 1.21 \times 10^{-4}\ sr^{-1}$ (L3), and generally increase slightly with wind speed and towards smaller relative azimuth angles. The lowest average RMSD, $0.79 \pm 0.32 \times 10^{-4}\ sr^{-1}$, resulted for $\Delta\phi = 170 - 180°$ and L2 quality control. RMSDs of the 3C parameterization without $L_s(\lambda)$ behave similarly (not shown).

Derived $R_{refl}(\lambda)$ are shown in Figure 6 as a function of wind speed and relative azimuth angle. With increasing wind speed and decreasing relative azimuth angle, $R_{refl}(\lambda)$ rises in magnitude but flattens spectrally. This is a result of increasing sun glint in both cases, which is spectrally complementary to sky glint. Rayleigh scattering is most efficient at $\Delta\phi \approx 180$, which leads to pronounced UV to blue spectral contributions to $R_{refl}(\lambda)$ at those geometries.

The sky reflectance factor ρ was allowed to vary between 0 and 0.0256 (ρ_f) in the 3C model inversion that took $L_s(\lambda)$ observations into account. The rationale behind the upper bound ρ_f is that for a perfectly flat water surface, 3C (Eq. (4)) converges to Eq. (2) with $\Delta = 0$ and $\rho = \rho_f$. The spectrally resolved 3C-offset $\Delta(\lambda)$ may then account for deviations to this ideal situation when the water surface is wind-roughened, which makes the 3C parameters ρ_{ds} and ρ_{dd} interpretable as additional sky and sun glint contribution levels. For situations when $L_s(\lambda)$ is not spectrally representative at all, i.e. when sun glint contributes heavily, the model optimization can account for this by setting $\rho < \rho_f$. However, the optimization predominantly retained the default value of 0.0256, with zero being the second most frequent outcome. While this is true for all filter levels, the percentage of ρ values at 0.0256 is considerably higher for the clear sky filter levels L2 (98.5 %) and L3 (97.84 %), as compared to less restrictive levels L0 (74.6 %)

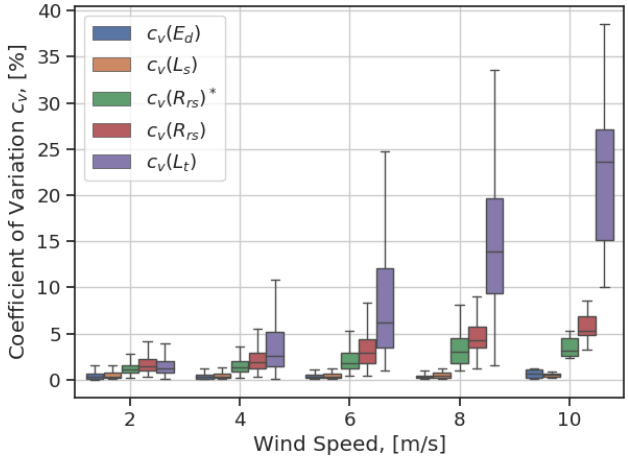

Fig. 5. Measurement cycle average coefficients of variation c_v for clear sky observations at LISCO (filter scheme L2, see Table 2) for the observed wind speed range (bins of 2 m/s) and for measured radiometric parameters downwelling irradiance $E_d(\lambda)$, sky radiance $L_s(\lambda)$, and total radiance $L_t(\lambda)$, as well as derived remote sensing reflectance $R_{rs}(\lambda)$. The asterisks (*) marks 3C model runs that did not make use of $L_s(\lambda)$ measurements. Boxes represent the c_v quartiles and whiskers extend 1.5 × the interquartile range.

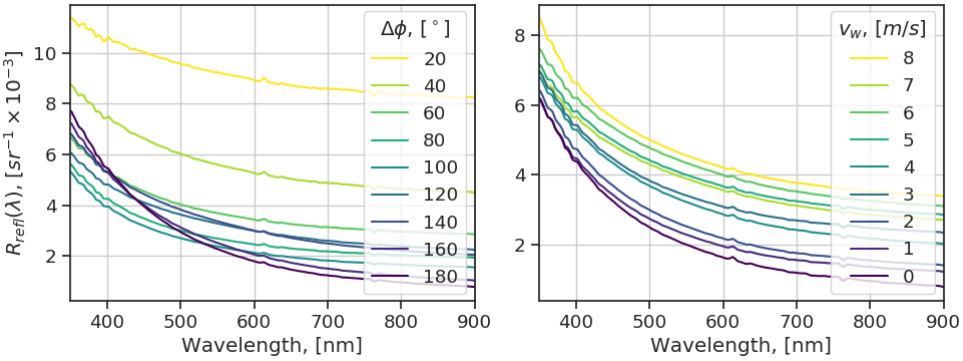

Fig. 6. Air-water interface reflectance $R_{refl}(\lambda)$ for all LISCO observations, averaged over relative azimuth angle $\Delta\phi$ (left panels) and wind speed v_w (right panels). Observations passed quality control levels L2 (i.e. $\Delta\phi \geq 20°$).

and L1 (78.28 %) that also include overcast conditions. This indicates that, on average, $L_s(\lambda)$ observations are less representative of surface reflected $L_r(\lambda)$ for scattered-cloud and overcast sky conditions, and that adding ρ as a bounded variable parameter to the optimization has no impact on the vast majority of clear sky observations.

Average direct and diffuse reflectance factors ρ_{dd} and ρ_{ds}, respectively, are presented in Figure 7 as a function of relative azimuth angle $\Delta\phi$ and wind speed. For the 3C model parameterization that accounted for $L_s(\lambda)$ observations, the diffuse reflection factor ρ_{ds} and ρ are presented jointly as the latter resulted predominantly in a constant value (ρ_f). This representation illustrates the compensating role of ρ_{ds} when omitting $L_s(\lambda)$ observations from the 3C processing. However, whenever $L_s(\lambda)$ observations are not spectrally representative of $L_r(\lambda)$, 3C has to compensate for this effect, which introduces significant differences to both reflectance factors resulting from 3C runs with and without $L_s(\lambda)$.

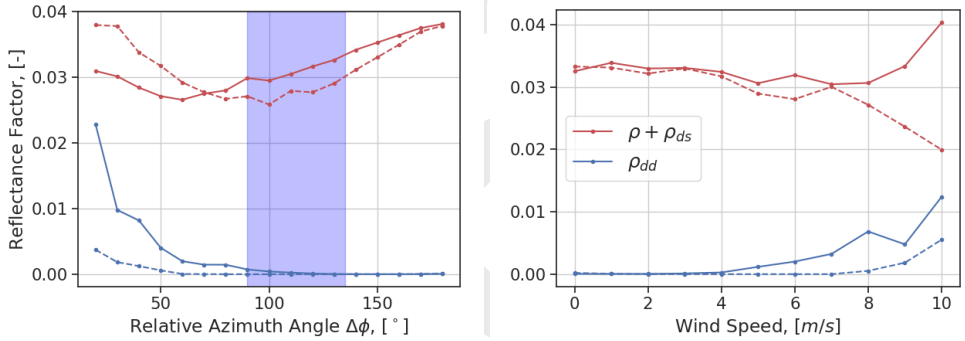

Fig. 7. Reflectance factors as a function of relative azimuth angle to the Sun $\Delta\phi$ (left panel) and wind speed (right panel) for filter scheme L2. Dashed lines indicate the 3C parameterization making no use of $L_s(\lambda)$ measurements. For the parameterization that uses $L_s(\lambda)$ observations, sky and diffuse reflectance factors ρ and ρ_{ds}, respectively, were combined for comparability (ρ = 0.0256 for the majority of observations). The Mobley geometry is indicated in blue shades for $\Delta\phi = 90 - 135°$ in the left panel.

As already anticipated from the simulation results in section 3.1, $\Delta\phi \approx 20°$ appears to be the tipping point below which sun glint is the dominating contribution to $L_r(\lambda)$ (on average for all considered θ_{sun}). Sun glint contributions were – on average – negligible within the $\Delta\phi = 90-135°$ range and low wind speeds, which supports the findings of [9]. A noticeable increase in ρ_{dd} was observed for $\Delta\phi < 90°$ and wind speeds above 4 m/s when processing with $L_s(\lambda)$. However, residual sun glint contributions were regularly observed also within the Mobley window, i.e. the standard deviation from the low median of $\rho_{dd} = 0.0001$ amounts to 0.07451. A ratio $\rho_{dd}/(\rho_{ds} + \rho) > 1/3$ can safely be assumed indicative for sun glint dominated $R_{refl}(\lambda)$ spectra, as direct downwelling irradiance amounts to approximately two thirds of total $E_d(\lambda)$, see i.e. [44], depending on atmospheric conditions and sun zenith angle. This ratio is exceeded for almost 9.6 % of all level L3 observations recorded at $\Delta\phi = 90 - 135°$. When considering only observations that are below the 20 %-percentile of $L_t(\lambda)$ in each measurement cycle, i.e. applying the L_t^{rel} filter scheme, still 9.3 % of the respective subset are identified as sun glint dominated. This indicates that L_t^{rel} is not particularly effective in removing sun glint dominated spectra, at least for observations that already adhere to the high L3 quality requirements. In addition, by selecting for the darkest $L_t(\lambda)$ observations in each sequence, L_t^{rel} also changes the mean spectral shape and magnitude of $L_r(\lambda)/E_d(\lambda)$. Compared to the mean $L_r(\lambda)/E_d(\lambda)$ spectrum per measurement cycle, this effect amounts to $\psi(350 - 900\,nm) = 9.1 \pm 9.2$ % ($\psi(350 - 450\,nm) = 4.3 \pm 4.9$ %,

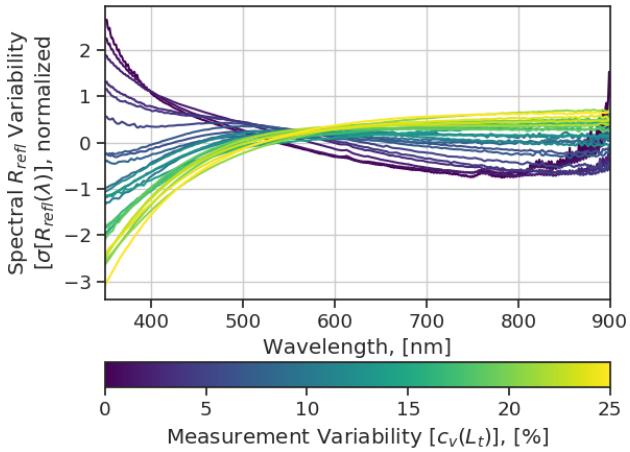

Fig. 8. Spectral shape of $R_{refl}(\lambda)$ variability (standard deviation of air-water interface reflectance $\sigma[R_{refl}(\lambda)]$) for all LISCO level L2 (i.e. $\Delta\phi \geq 20°$) observations, as a function of $L_t(\lambda)$ variability (coefficient of variation $c_v(L_t)$, 1 % binning). Spectra are normalized (z-score: subtraction of mean and division by standard deviation) to enable visual comparison of spectral shapes.

$\psi(450-750\,nm) = 8.8 \pm 9.0\,\%$, $\psi(750-900\,nm) = 12.8 \pm 12.8\,\%$) for level L3 observations. The impact of L_t^{rel} is further strongly dependent on wind speed with $\psi(350-900\,nm) < 10\,\%$ up to $v_w = 4\,m/s$. Regardless the cause, sun glint contamination introduces spectrally distinct variability to $L_t(\lambda)$ observations, which is illustrated in Figure 8: moving from low to high variability in $L_t(\lambda)$ ($c_v(L_t)$), the spectral shape of $\sigma[R_{refl}(\lambda)]$ transitions from resembling sky to sun reflectance. While Figure 8 shows this for level L2 LISCO observations ($\Delta\phi > 20°$), the same pattern can be observed for level L3 observations at $\Delta\phi = 90-135°$. This indicates that $c_v(L_t)$ and the shape of $\sigma[R_{refl}(\lambda)]$ are suitable indicators for sun glint contamination, and further illustrates why the reference model (Eq. (2)) with its spectral dependency fixed to $R_{sky}(\lambda)$ observations performs well for low variability conditions (see Figure 3), but fails once sun glint contributes more considerably.

4. Discussion and Conclusions

All existing approaches for the processing of in situ radiometric observations to $R_{rs}(\lambda)$ in the context of Eq. (2) are built on two fundamental assumptions: i) negligible sun glint and ii) spectrally isotropic sky radiance distributions. Measurement protocols were adjusted to accommodate for these assumptions, i.e observations are ideally to be collected at i) $\Delta\phi = 90-135°$ to minimize sun glint contributions [10] and ii) wind speeds below $4\,m/s$ to have waves facets reflect the sky in close vicinity to the $L_s(\lambda)$ observation [15]. However, in the context of the here considered LISCO data set, these constraints severely constrict the amount of observations available for analysis, i.e. less than 20 % adhere to $\Delta\phi = 90-135°$ (see Table 2, "sub-optimal $\Delta\phi$") and the majority was collected for wind speeds of 4 m/s and above (see Figure 5). The 3C model was developed to account for both sun glint and a spectrally variable sky radiance distribution, with the expressed aim to make more observations available for down-stream applications [18]. In the current work, we assessed to which extent these goals were reached by corroborating 3C inversions of simulated and measured above-surface radiometric observations.

Simulated $L_t(\lambda)$ spectra were inverted to derive retrieval bias in a realistic full-spectrum

inversion scenario for a range of representative $R_{rs}(\lambda)$ spectra at the LISCO site. Comparable biases were observed between the reference model and either 3C implementation for $\Delta\phi = 90-135°$ and wind speeds below 4 m/s, i.e. optimal conditions according to current measurement protocols for satellite calibration/validation activities. For observations outside of $\Delta\phi = 90-135°$, application of the reference model resulted in large errors, as expected due to sun glint contributions. Fit biases of approximately 5 % are in line with literature values for both theoretical uncertainty estimates and field validation experiments, e.g. [4, 18, 45]. When processing with 3C, the presented results clearly indicate that this level of uncertainty can be maintained also for wind speeds up to 8 m/s and relative viewing azimuth angles $\Delta\phi > 20°$. In the following, these findings and corresponding limitations are discussed in more detail.

Cox-Munk wave statistics are based on measurement series at a single location (Hawaii), however, have been scrutinized and tuned in the many decades since publication [34]. More accurate methodologies, e.g. [11], promise improved accuracies specifically for high solar zenith and viewing angles, which are not a focus in the current study. Coastal areas and bays like the Long Island Sound are likely to feature wave patterns that are not fully coherent with Cox-Munk slope statistics. This may specifically be an issue for look-up-table approaches like M99 that rely on accurate wave slope statistics, in contrast to bio-optical inversion schemes like 3C or [8].

The presented 3C processing results further confirm general applicability of 3C for coastal water observations, within the specifications of the applied bio-optical model [46]. For a reliable separation of sky and water signals with 3C, it is most relevant that the spectral shapes of all in-water components are accurate in the UV and NIR spectral bands, i.e. where biological processes have little impact on the reflectance spectrum. For example, spring bloom phytoplankton species composition at LISCO differs quite substantially from the rest of the year. The default specific phytoplankton absorption spectrum in 3C is not very suitable during that period and in the visible spectral bands. Processing with a more adequate phytoplankton parameterization for the spring period, however, had negligible impact on derived reflectances (results not shown). Still, striving for an optimal model fit is the best approach to assure robust and accurate reflectance retrievals. For this reason, 3C is not bound to a specific bio-optical model for its $R_{rs}(\lambda)$ component and researchers are encouraged to adjust and/or replace the currently implemented model if locally-tuned equivalents are available.

Other natural phenomena also manifest themselves in observed above-water radiometry and are not (yet) accounted for in 3C. For example, white caps introduce spectral reflectance features related to water absorption [47], and are likely contributing to rising fit residual averages with wind speed for the LISCO data set. White caps could be easily added to 3C as a spectral model component, supporting spectral decomposition by explaining a very distinctly shaped part of the observed variability in the NIR. While such an extension is out of scope for this publication, further research in this direction is highly encouraged. Spectral adjacency effects, too, have spectrally distinct effects on observed $R_{refl}(\lambda)$ in the NIR spectral region that could be accounted for in 3C [28] to support validation efforts for inland waters and coastal areas. Polarization of the downwelling light field is altered by interactions with the air-water interface and the water itself, which affects the magnitude of surface-reflected radiance [11, 17]. However, these effects are almost independent of wavelength in the here considered ranges and thus accountable for in the bio-optical model inversion. Minute wavelength dependencies in the water index of refraction n_w [9] were not considered.

The 3C fit residual ϵ proved as a robust indicator for failed retrievals, however, our findings do not support its use as a quantitative estimator for retrieval bias. Absolute values of ϵ depend on the magnitude of retrieved $R_{rs}(\lambda)$ and $L_r(\lambda)$, the spectral resolution and noise level of all radiometric input parameters, the parameterization of the 3C bio-optical model, and many other potential factors. We therefore refrain from stating an absolute threshold for ϵ to ensure model convergence for any particular retrieval, but rather recommend to analyze ϵ (or RMSD) histograms to find

suitable thresholds. For the here presented LISCO data set, the threshold RMSD=0.002 sr^{-1} (mean + 3σ) reliably separated obvious outliers while retaining the majority of spectra for further analysis.

With increasing wind speed, larger parts of the sky are sampled than covered by the field-of-view of a $L_s(\lambda)$ sensor. This issue could be mitigated with $L_s(\lambda)$ sensor field-of-views that vary with wind speed [17], e.g. derived from imaging spectrometer observations [19], or at least are optimized for typical conditions at a measurement site. Such an optimized $L_s(\lambda)$ field-of-view would be considerably larger than that of the $L_t(\lambda)$ sensor and depends on various parameters, i.e. average wind speed, instrument distance to the water, and integration times. For the LISCO data set, excluding $L_s(\lambda)$ from the 3C processing yielded consistently lower $c_v(R_{rs})$ at comparable fit residuals, indicating an improved separation of air-water interface and water signals and further substantiating the viability of this processing mode. Simulations indicated little dependency on wind speed for $R_{rs}(\lambda)$ when derived without $L_s(\lambda)$ observations, however, with considerable ($\psi \approx -7$ %) bias in the UV. Limiting application to sun zenith angles $\theta_{sun} \geq 60°$ reduced the average bias to levels comparable with the standard 3C parameterization. Sky reflectance $R_{refl}(\lambda)$ spectra may further contain spectral information in addition to the the components already considered in 3C (Rayleigh and aerosol scattering, direct sun light). In overcast conditions, theory predicts a spectrally flat $L_s(\lambda)/E_d(\lambda)$ ratio equal to $\pi^{-1}\ sr^{-1}$. Particularly in overcast conditions, atmospheric path lengths for light reaching a sky radiance sensor are typically different from the average path length of photons collected by a $E_d(\lambda)$ sensor, which introduces spectral absorption features into their ratio. Similarly, scattered cloud may or may not cover the $L_s(\lambda)$ sensor field-of-view and/or be reflected on the water surface [31]. Furthermore, for $R_{rs}(\lambda)$ observations nearby land, ground albedo can introduce adjacency effects to the NIR spectral range [28]. All these effects may not be accounted for when correcting $R_{rs}(\lambda)$ with $\Delta(\lambda)$ only. However, observations from airborne platforms like drones and and airplanes are typically acquired when clear sky conditions prevail and could therefore benefit from $L_s(\lambda)$-less $R_{rs}(\lambda)$ retrievals with 3C, despite the limitations. Removing the strict requirement for $L_s(\lambda)$ observations could enable low-cost reflectance instruments, i.e. in the context of citizen science or open-source projects. Last but not least, $L_s(\lambda)$-less processing is a precursor to application of 3C in the context of hyperspectral observations from space, e.g. as anticipated from NASA PACE [48].

Redundant $L_t(\lambda)$ observations, i.e. acquired in rapid succession (>2 Hz) over several minutes, are essential to estimate the uncertainty in derived $R_{rs}(\lambda)$. For such short periods of time, the water reflectance signal can usually be assumed constant. Therefore, if a correction methodology is able to attribute a large share of $L_t(\lambda)$ variability to air-water interface reflections, $R_{rs}(\lambda)$ retrieved from a series of redundant $L_t(\lambda)$ observations should match closely. For the LISCO data set, 3C achieved this goal with average $c_v(R_{rs}) < 5$ % for all wind speed bins (see Figures 5, 4). It should be noted that $c_v(R_{rs})$ is indicative of retrieval precision, which is required yet not sufficient to estimate uncertainty levels.

Sun glint contributions to $L_t(\lambda)$ are usually low when observing within the Mobley geometry constraints. This is partially because wave facets are rarely oriented such that they reflect the Sun disc, and partially because the Fresnel reflectance for these orientations is low. However, because yellow sun light is spectrally so profoundly contrasting blue sky light, even small sun glint contributions dominate the observed spectral variability of $R_{refl}(\lambda)$ (Figure 8). This is likely why our results suggest that averaging schemes like L_t^{rel} may be suitable for removing extreme outliers, yet had very little effect on already quality controlled observations and at the expense of artificially altering the magnitude and spectral composition of resulting $R_{refl}(\lambda)$. Such filtered $R_{refl}(\lambda)$ do not adhere to Cox-Munk wave statistics that are core to tabulated ρ values of e.g. the M99 approach. However, variance in $R_{refl}(\lambda)$ sequences can directly be estimated from $L_t(\lambda)/E_d(\lambda)$ observations (see Eq. (1)), without relying on elaborate processing schemes like

3C and assuming only a constant water signal during data acquisition. In essence, a $\sigma[L_t/E_d]$ spectrum that resembles the shape of $R_{sky}(\lambda)$ indicates negligible sun glint contamination. This recommends $L_t(\lambda)$ variance spectra as valuable assets for quality control purposes and we encourage researchers to leverage the wealth of information contained in redundant $L_t(\lambda)$ observations, rather than artificially constricting it.

In summary, we assessed $R_{rs}(\lambda)$ retrieval bias for 3C (Eq. (4)) and a reference model (Eq. (2)), based on representative $R_{rs}(\lambda)$ spectra at LISCO and a simulated $L_r(\lambda)$ data set. Our results confirm the restrictions placed on retrievals with the reference model, i.e. $\Delta\phi = 90 - 135°$ (Mobley geometry) and wind speeds below 4 m/s, to ensure $L_s(\lambda)$ observations to be spectrally representative of $L_r(\lambda)$. For 3C, comparable retrieval errors were obtained for $\Delta\phi \geq 20°$, i.e. only excluding directly Sun facing observations, wind speeds up to approximately 8 m/s, and all relevant sun zenith angles. These loosened restrictions allowed to confidently process the LISCO data set with 3C in its entirety, which was recorded at wind speeds predominantly above 4 m/s and with only about 20 % of all observations adhering to the Mobley geometry. Derived $R_{refl}(\lambda)$ spectra and 3C parameter results are consistent with simulations and corroborate the established operation boundaries, for which a high level of $R_{rs}(\lambda)$ retrieval precision ($c_v(R_{rs}) < 5$ %) was maintained.

Funding

National Aeronautics and Space Administration (NASA), Ocean Biology and Biochemistry program (NNX16AR46G); National Oceanic and Atmospheric Administration (NOAA): Joint Polar Satellite System Calibration/validation program and ESSRST Center (NA16SEC4810008).

Acknowledgments

This research was performed while Robert Foster held a NRC Research Associateship Award at the U.S. Naval Research Laboratory in Washington, D.C.

Disclosure

The authors declare no conflicts of interest.

References

1. H.-O. Pörtner, D. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama, and N. W. (eds.), "IPCC Special Report on the Ocean and Cryosphere in a Changing Climate," Tech. rep. (2019).
2. T. Platt, N. Hoepffner, V. Stuart, and C. Brown, "IOCCG Report Number 7: Why ocean colour? The societal benefits of ocean-colour technology," Tech. rep. (2008).
3. W. W. Gregg, "IOCCG Report Number 6: Ocean-Colour Data Merging," Tech. rep. (2007).
4. Ruddick, K.G., Voss, K., Boss, E., Castagna, A., Frouin, R., Gilerson, A., Hieronymi, M., Johnson, B.C., Kuusk, J., Z. Lee, and M. Ondrusek, "A Review of Protocols for Fiducial Reference Measurements of WaterLeaving Radiance for Validation of Satellite Remote-Sensing Data over Water," Remote. Sens. **11**, 2198 (2019).
5. S. G. Simis and J. Olsson, "Unattended processing of shipborne hyperspectral reflectance measurements," Remote. Sens. Environ. **135**, 202–212 (2013).
6. A. G. Dekker, N. Pinnel, P. Gege, X. Briottet, A. Court, S. Peters, K. R. Turpie, S. Sterckx, M. Costa, C. Giardino, V. E. Brando, F. Braga, M. Bergeron, T. Heege, and B. Pflug, "Feasibility Study for an Aquatic Ecosystem Earth Observing System," Tech. rep., Committee on Earth Observing Satellites (CEOS) (2017).
7. C. Cox and W. Munk, "Measurement of the Roughness of the Sea Surface from Photographs of the Sun's Glitter," J. Opt. Soc. Am. **44**, 838–850 (1954).
8. Z. Lee, Y.-h. Ahn, C. Mobley, and R. Arnone, "Removal of surface-reflected light for the measurement of remote-sensing reflectance from an above-surface platform," Opt. Express **18**, 171–182 (2010).
9. X. Zhang, S. He, A. Shabani, P.-W. Zhai, and K. Du, "Spectral sea surface reflectance of skylight," Opt. Express **25**, A1 (2017).
10. C. D. Mobley, "Estimation of the remote-sensing reflectance from above-surface measurements." Appl. Opt. **38**, 7442–55 (1999).

11. C. D. Mobley, "Polarized reflectance and transmittance properties of windblown sea surfaces," Appl. Opt. **54**, 4828–4849 (2015).
12. R. J. Gould, R. Arnone, and M. Sydor, "Absorption, Scattering, and Remote-Sensing Reflectance Relationships in Coastal Waters: Testing a New Inversion Algorithm," J. Coast. Res. **17**, 328–341 (2001).
13. K. G. Ruddick, V. De Cauwer, Y.-J. Park, and G. Moore, "Seaborne measurements of near infrared water-leaving reflectance: The similarity spectrum for turbid waters," Limnol. Oceanogr. **51**, 1167–1179 (2006).
14. K. Ruddick, V. De Cauwer, and B. Van Mol, "Use of the near infrared similarity reflectance spectrum for the quality control of remote sensing data," Remote. Sens. Coast. Ocean. Environ. **5885**, 588501 (2005).
15. G. Zibordi, "Experimental evaluation of theoretical sea surface reflectance factors relevant to above-water radiometry," Opt. Express **24**, A446 (2016).
16. C. D. Mobley, L. K. Sundman, and R. Road, "HYDROLIGHT 5 ECOLIGHT 5 USERS ' GUIDE," (2008).
17. R. Foster and A. Gilerson, "Polarized transfer functions of the ocean surface for above-surface determination of the vector submarine light field," Appl. Opt. **55**, 9476 (2016).
18. P. M. M. Groetsch, P. Gege, S. G. H. Simis, M. A. Eleveld, and S. W. M. Peters, "Validation of a spectral correction procedure for sun and sky reflections in above-water reflectance measurements," Opt. Express **25**, A742–A761 (2017).
19. C. Carrizo, A. Gilerson, R. Foster, A. Golovin, and A. El-Habashi, "Characterization of radiance from the ocean surface by hyperspectral imaging," Opt. Express **27**, 1750 (2019).
20. M. Hieronymi, "Polarized reflectance and transmittance distribution functions of the ocean surface," Opt. Express **24**, A1045 (2016).
21. W. Shi, M. Wang, and L. Jiang, "Tidal effects on ecosystem variability in the Chesapeake Bay from MODIS-Aqua," Remote. Sens. Environ. **138**, 65–76 (2013).
22. K. J. Voss and A. Morel, "Bidirectional reflectance function for oceanic waters with varying chlorophyll concentrations: Measurements versus predictions," Limnol. Oceanogr. **50**, 698–705 (2005).
23. A. Gilerson and Y. Huot, "Sun-induced chlorophyll-a fluorescence," in *Bio-optical modelling and remote sensing of inland waters,* D. R. Mishra, I. Ogashawara, and A. A. Gitelson, eds. (Elsevier, 2017), pp. 189–231.
24. J. Seppälä, P. Ylöstalo, and H. Kuosa, "Spectral absorption and fluorescence characteristics of phytoplankton in different size fractions across a salinity gradient in the Baltic Sea," Int. J. Remote. Sens. **26**, 387–414 (2005).
25. S. Sathyendranath, J. Aiken, and S. Alvain, "IOCCG Report Number 15: Phytoplankton functional types from Space." Tech. rep. (2014).
26. B. K. Grunert, C. B. Mouw, and A. B. Ciochetto, "Characterizing CDOM Spectral Variability Across Diverse Regions and Spectral Ranges," Glob. Biogeochem. Cycles **32**, 57–77 (2018).
27. G. Zibordi, J. P. Doyle, and S. B. Hooker, "Offshore tower shading effects on in-water optical measurements," J. Atmospheric Ocean. Technol. **16**, 1767–1779 (1999).
28. P. M. M. Groetsch, P. Gege, S. G. H. Simis, M. A. Eleveld, and S. W. M. Peters, "Variability of adjacency effects in sky reflectance measurements," Opt. Lett. **42**, 1–5 (2017).
29. B. Bulgarelli and G. Zibordi, "Seasonal Impact of Adjacency Effects on Ocean Color Radiometry at the AAOT Validation Site," IEEE Geosci. Remote. Sens. Lett. **15**, 488–492 (2018).
30. P. Gege, "A case study at Starnberger See for hyperspectral bathymetry mapping using inverse modeling," in *WHISPERS 2014,* (2014), pp. 1–4.
31. A. Göritz, S. A. Berger, P. Gege, H. P. Grossart, J. C. Nejstgaard, S. Riedel, R. Röttgers, and C. Utschig, "Retrieval of water constituents from hyperspectral in-situ measurements under variable cloud cover-A case study at Lake Stechlin (Germany)," Remote. Sens. **10** (2018).
32. V. P. Lance, "Report for Dedicated JPSS VIIRS Ocean Color Calibration / Validation Cruise May 2018," Tech. rep. (2018).
33. T. Harmel, A. Gilerson, S. Hlaing, A. Tonizzo, T. Legbandt, A. Weidemann, R. Arnone, and S. Ahmed, "Long Island Sound Coastal Observatory: Assessment of above-water radiometric measurement uncertainties using collocated multi and hyperspectral systems," Appl. Opt. **50**, 5842 (2011).
34. F. M. Bréon and N. Henriot, "Spaceborne observations of ocean glint reflectance and modeling of wave slope distributions," J. Geophys. Res. Ocean. **111** (2006).
35. L. Hosek and A. Wilkie, "An analytic model for full spectral sky-dome radiance," ACM Transactions on Graph. **31**, 1–9 (2012).
36. L. Hošek and A. Wilkie, "Adding a solar-radiance function to the Hošek-Wilkie skylight model," IEEE Comput. Graph. Appl. **33**, 44–52 (2013).
37. A. J. Preetham, P. Shirley, and B. Smits, "A practical analytic model for daylight," in *Proceedings of the SIGGRAPH annual conference on Computer graphics,* (1999), pp. 91–100.
38. E. Bruneton, "A Qualitative and Quantitative Evaluation of 8 Clear Sky Models," IEEE Transactions on Vis. Comput. Graph. **23**, 2641–2655 (2017).
39. D. A. Aurin, H. M. Dierssen, M. S. Twardowski, and C. S. Roesler, "Optical complexity in Long Island Sound and implications for coastal ocean color remote sensing," J. Geophys. Res. **115**, 1–18 (2010).
40. H. Hersbach, B. Bell, P. Berrisford, A. Horányi, J. M. Sabater, J. Nicolas, R. Radu, D. Schepers, A. Simmons, C. Soci, and D. Dee, "Global reanalysis: goodbye ERA-Interim, hello ERA5," ECMWF Newsl. pp. 17–24 (2019).
41. W. Gregg and K. Carder, "A simple spectral solar irradiance model for cloudless maritime atmospheres," Limonology

And Oceanogr. **35**, 1657–1675 (1990).
42. A. Albert and Mobley, "An analytical model for subsurface irradiance and remote sensing reflectance in deep and shallow case-2 waters." Opt. express **11**, 2873–90 (2003).
43. R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A Limited Memory Algorithm for Bound Constrained Optimization," SIAM J. on Sci. Comput. **16**, 1190–1208 (1995).
44. P. Gege, "Analytic model for the direct and diffuse components of downwelling spectral irradiance in water." Appl. Opt. **51**, 1407–1419 (2012).
45. G. Zibordi, K. Ruddick, I. Ansko, G. Moore, S. Kratzer, J. Icely, and A. Reinart, "In situ determination of the remote sensing reflectance: An inter-comparison," Ocean. Sci. **8**, 567–586 (2012).
46. A. Albert and C. Mobley, "An analytical model for subsurface irradiance and remote sensing reflectance in deep and shallow case-2 waters," Opt. Express **11**, 2873 (2009).
47. H. M. Dierssen, "Hyperspectral measurements, parameterizations, and atmospheric correction of whitecaps and foam from visible to shortwave infrared for ocean color remote sensing," Front. Earth Sci. **7**, 1–18 (2019).
48. P. J. Werdell, M. J. Behrenfeld, P. S. Bontempi, E. Boss, B. Cairns, G. T. Davis, B. A. Franz, U. B. Gliese, E. T. Gorman, O. Hasekamp, K. D. Knobelspiesse, A. Mannino, J. V. Martins, C. McClain, G. Meister, and L. A. Remer, "The plankton, aerosol, cloud, ocean ecosystem mission status, science, advances," Bull. Am. Meteorol. Soc. **100**, 1775–1794 (2019).