Universiteit Leiden
The Netherlands

Accessible remote sensing of water
Burggraaff, O.

5 | Biases from incorrect reflectance convolution

Reflectance, a crucial earth observation variable, is converted from hyperspectral to multi-spectral through convolution. This is done to combine time series, validate instruments, and apply retrieval algorithms. However, convolution is often done incorrectly, with reflectance itself convolved rather than the underlying (ir)radiances. Here, the resulting error is quantified for simulated and real multispectral instruments, using 18 radiometric data sets ($N = 1799$ spectra). Biases up to 5% are found, the exact value depending on the spectrum and band response. This significantly affects extended time series and instrument validation, and is similar in magnitude to errors seen in previous validation studies. Post-hoc correction is impossible, but correctly convolving (ir)radiances prevents this error entirely. This requires publication of original data alongside reflectance.

5.1   Introduction

Reflectance, the spectral fraction of light reflected by a surface, is an essential earth observation (EO) variable. It forms the basis for data products such as chlorophyll and suspended matter in water [62, 227, 232], and canopy cover and biomass on land [415, 416]. As such, it is a routine data product for EO satellites, including NASA's Landsat and ESA's Sentinel programmes, and in situ radiometers.

Spectral data are divided into two categories, namely multispectral and hyperspectral. Multispectral instruments observe in several broad, discrete wavelength bands. Examples include the Moderate Resolution Imaging Spectrometer (MODIS) and the Visible Infrared Imaging Radiometer Suite (VIIRS), but also in situ instruments including unmanned aerial vehicles (UAVs) and even smartphones [281]. Conversely, hyperspectral instruments provide continuous wavelength coverage with a fine spectral resolution. Examples include the TriOS RAMSES, Sea-Bird HyperOCR, and ASD FieldSpec field-going spectroradiometers, as well as the Ocean Color Instrument (OCI) due to fly on the Plankton, Aerosol, Cloud, ocean Ecosystem (PACE) mission. Hyperspectral data have a finer spectral sampling and, typically, resolution and thus contain more information than multispectral ones, but depending on the instrument design, often collect less light in each band, giving a worse signal-to-noise ratio.

Since the current EO landscape is a mixture of both types, it is often desirable to convert data between the two, typically from hyper- to multispectral. Three common use cases for this process exist, namely combining time series, instrument validation, and retrieval algorithms.

The first use case is merging and extending time series using different sensors. Long-term, high temporal resolution time series are necessary to study fundamental biogeochemical processes and long-term effects [195, 220] such as climate change [417]. Current efforts focus on merging multispectral time series, on the radiance or reflectance level [91, 418, 419], achieving relative errors on reflectance <5% [91, 419], or on the end product level [220, 420]. Future efforts will focus on extending multispectral time series with new hyperspectral sensors, for example extending MODIS/VIIRS aerosol optical depth (AOD) series with OCI (PACE) data [421]. This is done by converting hyperspectral data to the multispectral sensor's bands, to simulate what the latter would have measured. However, calibration differences and sensor characterisation imperfections can introduce significant biases, for example up to 0.10 AOD for OCI-MODIS/VIIRS [421].

The second use case is the validation of multispectral (often satellite) data using in situ hyperspectral sensors. This is done by comparing simultaneous match-up measurements from both instruments [152]. Validation is done on all products, including normalised radiance [422], reflectance [415, 418, 423, 424], and derived products such as chlorophyll [220, 223] and inherent optical properties (IOPs) [424]. Similar validation is done for in situ multispectral sensors, such as UAVs [425] and smartphones [121, 276]. Vicarious calibration similarly involves comparing match-up data, but aimed at determining satellite gain factors [113]. Since vicarious calibration is performed on (normalised) radiance rather than reflectance, it is outside the scope of this work, though a brief discussion is given in Section 5.2.4.

The third use case is the application of multispectral retrieval algorithms to hyperspectral data. Such algorithms are commonly based on the ratio between spectral bands and are thus called band-ratio algorithms. For example, band-ratio algorithms relating chlorophyll to Sentinel-2A (S2A) Multispectral Instrument (MSI) bands have been developed for Vietnamese [221] and Estonian [62] lakes, the latter with a mean standard error in chlorophyll-a of 5%. While derived on multispectral data, such algorithms are also applied to hyperspectral

data, both to derive products and for validation, requiring a spectral conversion. Differences between these converted data and in situ data have been found [426], which may be due in part to incorrect treatment of radiometry. It should be noted that instead of a spectral conversion, often only the central band wavelength reflectance is used [223, 227].

Converting hyperspectral data to multispectral bands is commonly, though not exclusively (see Section 5.2.2), termed spectral convolution. An in-depth description is provided in Section 5.2, but in short, the hyperspectral data are multiplied by the spectral response function (SRF) of the multispectral band and the product is integrated. This is done for quantities including radiance [427, 428], optical thickness [428, 429], IOPs [424, 430], vegetation indices [416], and reflectance [415, 430].

However, reflectance is often convolved incorrectly. As shown in Section 5.2, hyperspectral reflectance cannot simply be convolved to simulate what a multispectral instrument would observe. Instead, the numerator and denominator, (ir)radiances, should be convolved separately and then divided. This error occurs frequently in the literature, for example in [62, 121, 221, 227, 276, 415, 416, 418, 420, 423–425, 431, 432], with few works convolving radiances before division [223, 426, 433].

This work quantifies the error induced by incorrect spectral convolution of reflectance in each of the three use cases, for a variety of synthetic and real instruments using 18 archival data sets totaling $N = 1799$ spectra. To narrow the scope, this work focuses on remote sensing of ocean colour. However, the principles and methods apply broadly to any fractional quantity, including other reflectances (soil, vegetation), attenuation coefficients, and degree of polarisation, as well as spatial convolution [434]. While the existence of this error has been pointed out previously [430, 433, 435] and quantified at $\lesssim 1\%$ for a single data set and sensor [430], a large-scale quantitative assessment has not yet been published.

This work fits into a wider field of EO error analysis. Recent efforts include investigations into the out-of-band response of EO sensors [115], the impact of differing spectral [433] and spatial [434] resolutions on satellite match-up analyses, and the impact of hyperspectral SRFs having a non-zero bandwidth [436]. On the experimental side, significant efforts have gone into glint removal in above-water radiometry [120, 206, 209] and rigorous characterisation of instrumental [281, 425] and methodological [152, 435] uncertainties. A broad, in-depth overview of uncertainties in ocean colour data is provided in the recently published International Ocean Color Coordination Group (IOCCG) report number 18 [88].

Section 5.2 describes the theoretical background of reflectance and spectral convolution. Section 5.3 describes the data used in this work and the method for quantifying the convolution error. Results are presented in Section 5.4. Finally, Section 5.5 contains a discussion of the results and conclusions.

5.2 Theoretical background

5.2.1 Reflectance

Reflectance R is the ratio of upwelling over downwelling (ir)radiance. Radiance $L(\lambda, \theta, \phi)$ is the radiant energy per wavelength λ propagating in a direction (θ, ϕ), in W m^{-2} nm^{-1} sr^{-1}, while irradiance $E(\lambda)$ is L integrated over a solid angle, in W m^{-2} nm^{-1}. The units of R depend on which ratio is taken. Since this work deals only with wavelength dependence, (θ, ϕ) terms are dropped henceforth for clarity.

Different reflectances can be defined by dividing different (ir)radiances. Examples include the bi-directional radiance reflectance [415], the non-directional irradiance reflectance [432], and the uni-directional remote sensing reflectance R_{rs} used in ocean colour [120, 209]. As defined in Equation (5.1), R_{rs} is the ratio of water-leaving radiance L_w [120] over downwelling irradiance E_d [97], in units of sr^{-1}. This work focuses on R_{rs}, but the same mathematics apply to any reflectance.

$$R_{rs}(\lambda) = \frac{L_w(\lambda)}{E_d(\lambda)} \tag{5.1}$$

5.2.2   Spectral convolution

Multispectral data are simulated from hyperspectral data through spectral convolution. As shown in Equation (5.2), this involves multiplying the hyperspectral data $L(\lambda)$ by the multispectral band SRF $S_B(\lambda)$, integrating the result over all wavelengths in the band ($\int_{\lambda \in B} d\lambda$), and normalising by the effective bandwidth. In this work, convolved quantities are denoted by a bar, such as $\bar{L}(B)$ in Equation (5.2). In practice, spectral convolution is often a sum over discrete L and S_B data. The convolution process is shown graphically in Figure 5.1.

$$\bar{L}(B) = \frac{\int_{\lambda \in B} L(\lambda) S_B(\lambda) d\lambda}{\int_{\lambda \in B} S_B(\lambda) d\lambda} \tag{5.2}$$

Convolving hyperspectral data is really an approximation, due to the finite spectral resolution of hyperspectral sensors. As derived in Appendix 5.A, this method is valid if the full width at half maximum (FWHM) of the multispectral band is at least double that of the hyperspectral sensor.

Nomenclature

Various names for this process are used in the literature, including convolution or convolving [97, 152, 276, 415, 416, 418–420, 423, 425, 427, 432, 433, 435, 436], SRF-weighting [429], simulation [426], and band-averaging [121, 281, 424, 427, 428, 430]. Since it is the most common term, 'spectral convolution' is used in this work. However, it should be noted that this term may instead refer to smoothing the spectrum with a kernel [99]. Finally, since neither process involves transforming the SRF, both are actually cross-correlations rather than convolutions.

5.2.3   Reflectance convolution

Just as the hyperspectral remote sensing reflectance $R_{rs}(\lambda)$ is $L_w(\lambda)$ over $E_d(\lambda)$, the convolved $\bar{R}_{rs}(B)$ is $\bar{L}_w(B)$ over $\bar{E}_d(B)$. Both are calculated as in Equation (5.2) and then divided, as shown in Equation (5.3). Convolving (ir)radiances to calculate a band-average reflectance will be referred to in this work as working in *radiance space* or *L-space*, and the result as $\bar{R}_{rs}^L(B)$. Mathematically, this is the correct method for convolving R_{rs} to simulate multispectral data.

$$\bar{R}_{rs}^L(B) = \frac{\bar{L}_w(B)}{\bar{E}_d(B)} = \frac{\int_{\lambda \in B} L_w(\lambda) S_B(\lambda) d\lambda}{\int_{\lambda \in B} E_d(\lambda) S_B(\lambda) d\lambda} \tag{5.3}$$
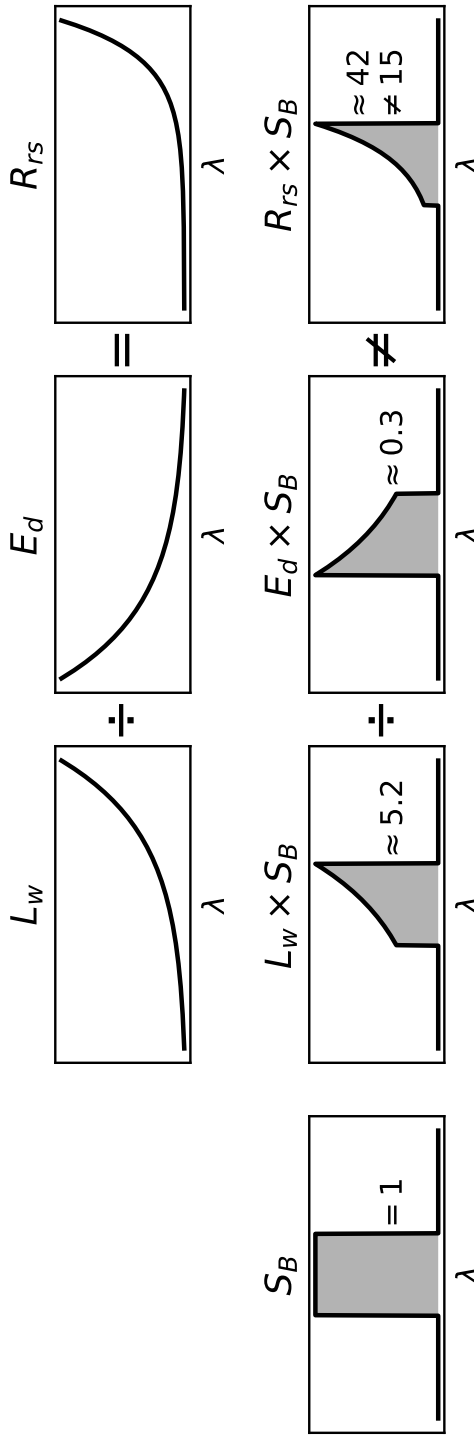
Figure 5.1: Demonstration of spectral convolution with a boxcar SRF S_B (bottom left). Each column shows the convolution of a spectrum (L_w, E_d, R_{rs}) with S_B, the shaded area showing the integral as calculated with Equation (5.2). The $\times$ symbol indicates a spectral convolution. The band-average value of R_{rs} calculated in R-space is 42, significantly different from the correct (L-space) value of $5.2/0.3 \approx 15$. Axis labels and units are dropped for clarity.

Instead, one might simply convolve R_{rs} itself. This will be referred to as working in *reflectance space* or *R-space* and the result as $\bar{R}_{rs}^R(B)$. The expression for $\bar{R}_{rs}^R(B)$ is given in Equation (5.4).

$$\bar{R}_{rs}^R(B) = \frac{\int_{\lambda \in B} R_{rs}(\lambda) S_B(\lambda) \mathrm{d}\lambda}{\int_{\lambda \in B} S_B(\lambda) \mathrm{d}\lambda} = \frac{\int_{\lambda \in B} \frac{L_w(\lambda)}{E_d(\lambda)} S_B(\lambda) \mathrm{d}\lambda}{\int_{\lambda \in B} S_B(\lambda) \mathrm{d}\lambda} \tag{5.4}$$

Working in *R*-space is incorrect, as shown in Figure 5.1 and the following example. First, let the SRF $S_B(\lambda)$ be a boxcar response of 1 for $0 \le \lambda \le 1$ and 0 elsewhere. Then all integrals need only be evaluated for those wavelengths and the SRF bandwidth is 1. Second, let $L_w(\lambda) = e^{e\lambda}$ and $E_d(\lambda) = e^{-e\lambda}$. Such spectra are not physical but demonstrate the mathematical principles well. As shown in Equations (5.5) and (5.6), $\bar{R}_{rs}^L(B) \approx 15$ and $\bar{R}_{rs}^R(B) \approx 42$ differ significantly.

$$\bar{R}_{rs}^L(B) = \frac{\int_0^1 e^{e\lambda} \mathrm{d}\lambda}{\int_0^1 e^{-e\lambda} \mathrm{d}\lambda} \approx \frac{5.2}{0.3} \approx 15 \tag{5.5}$$

$$\bar{R}_{rs}^R(B) = \frac{\int_0^1 \frac{e^{e\lambda}}{e^{-e\lambda}} \mathrm{d}\lambda}{1} = \int_0^1 e^{2e\lambda} \mathrm{d}\lambda \approx 42 \tag{5.6}$$

5.2.4   General rule

Convolution is a useful tool, but the order of operations is not always intuitive. A general rule of thumb can be used, which applies to any kind of convolution (spectral or spatial) when converting high- to low-resolution (spectral or spatial) data. For other purposes, such as smoothing, reflectance itself can be transformed.

As a rule of thumb, only quantities the lower-resolution sensor would observe can be convolved. This includes the at-sensor (ir)radiance (in physical units [435]) but not reflectance and derived products. Propagation of in situ radiances, through surfaces when measured underwater [120] or through the atmosphere for vicarious calibration [113], must occur prior to convolution to accurately simulate the radiance at a multispectral sensor. Simplifications may be necessary [115, 428, 437] but should be mathematically justified. Finally, hyperspectral upwelling radiance L_u, measured in- or above-water, should be converted to L_w [120] before convolution when comparing it to multispectral L_w.

5.3   Methods

Archival data sets containing (ir)radiance and reflectance data were used to test the principles described in Section 5.2 and quantify the errors resulting from working in *R*-space rather than *L*-space. All analysis was done using custom Python scripts, available from GitHub[18].

5.3.1   Radiometric data

18 archival radiometric data sets were used [37, 438–451], totaling $N = 1799$ spectra. Data were sourced from the SeaWiFS Bio-optical Archive and Storage System (SeaBASS) [241]

---

[18]https://github.com/burggraaff/reflectance_convolution

and PANGAEA[19]. Only data sets including either the original radiometric data or R_{rs} and L_w or E_d were used. In most cases, the given R_{rs} and E_d were used and $L_w = R_{rs}E_d$ was reconstructed. This reduced the amount of post-processing, such as glint removal, that was necessary. All spectra used in the further analysis are shown in Figure 5.2. An overview of the data and post-processing is provided in Appendix 5.B.

Small imperfections in the resulting data, such as residual atmospheric bands in R_{rs} (Figure 5.2), are no problem. For this work, it is only necessary to obtain a set of realistic spectra, not to determine IOPs. Negative R_{rs} were removed since they are not physical but instead the result of measurement error or over-correction of glint; this is no problem for the same reason.
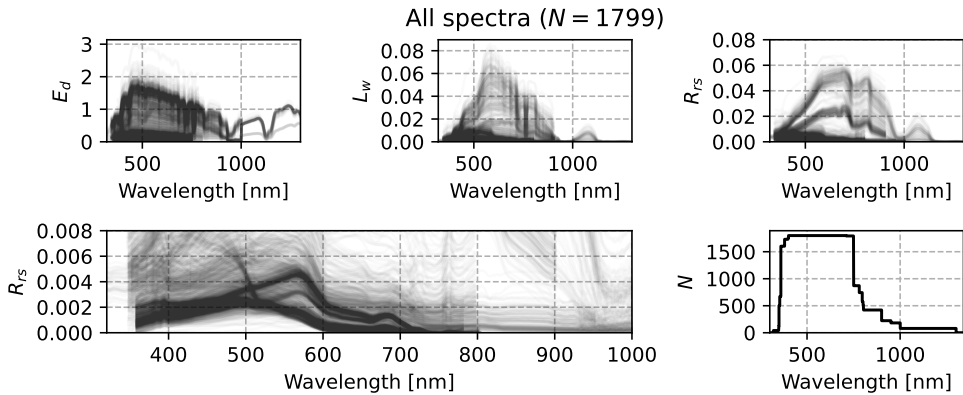

Figure 5.2: All radiometric data used in this work, in SI units. $N(\lambda)$ is the number of spectra that include wavelength λ. The bottom left panel is a zoom on the top right one. Individual spectra are plotted with high transparency.

5.3.2 Spectral convolution

Spectral convolution was implemented in the custom Python library described above. The radiometric data were interpolated to the SRF wavelengths. If the radiometric data and SRF wavelengths did not overlap fully, the convolution was only done if the integral of the SRF over the non-overlapping wavelengths was ≤5% of its total integral. The integration was done using the SciPy implementation of Simpson's rule in the `integrate.simps` function [398].

In each experiment, data were convolved in both L- and R-space, and the resulting reflectances were compared in absolute and relative terms. The absolute difference is $\Delta\bar{R}_{rs} = \bar{R}_{rs}^R(B) - \bar{R}_{rs}^L(B)$, meaning a positive $\Delta\bar{R}_{rs}$ corresponds to an overestimation in R-space. The relative difference was normalised to $\bar{R}_{rs}^L(B)$, and set to 0% if $\bar{R}_{rs}^L(B) = 0$ sr^{-1}. All spectra were treated separately, enabling a statistical analysis of the difference on varying input spectra. Due to the finite spectral resolution of the hyperspectral data, some data sets could not be convolved with some multispectral SRFs (see Appendix 5.A).

---

[19]https://pangaea.de/

5.3.3   Spectral response functions

Synthetic

The dependence of $\Delta \bar{R}_{rs}$ on band location and width was investigated by generating various synthetic boxcar and Gaussian filters. Both are common approximations of real SRFs [416, 430]. Boxcars were evaluated on wavelengths with an non-zero response, Gaussians on wavelengths from 320–800 nm. For both, a 0.1 nm step size was used to properly sample narrow bands. Central wavelengths between 330–809 nm (1 nm steps) and FWHMs of 6–65 nm (1 nm steps) were used, representative of real multispectral instruments (Section 5.3.3).

Real instruments

The behavior of $\Delta \bar{R}_{rs}$ for real multispectral instruments, namely eleven satellite instruments and three low-cost sensors, was also investigated. A selection of these is shown in Figure 5.3. Panchromatic bands were not used as they are intended for spatial sharpening, not reflectance measurements. Only bands fitting the radiometric data (within 320–1300 nm) were used. The satellite instruments were the Enhanced Thematic Mapper Plus (ETM+) aboard Landsat 7 [452], Operational Land Imager (OLI) aboard Landsat 8 [453], Coastal Zone Color Scanner (CZCS), Sea-Viewing Wide Field-of-View Sensor (SeaWiFS), MODIS aboard Aqua and Terra, Medium Resolution Imaging Spectrometer (MERIS), VIIRS aboard Suomi NPP [454], MSI aboard S2A/B [455], and the Ocean and Land Colour Instrument (OLCI) aboard Sentinel-3A/B (S3A/B) [456]. These are all commonly used to measure R_{rs}. The low-cost sensors were one UAV, the DJI Phantom Pro 4, and two smartphones, the iPhone SE and Samsung Galaxy S8 [281]. Such sensors have become popular in their own right as they can provide radiance data, if radiometrically calibrated [121, 281, 425], but also serve as proxies for new cubesat sensors such as the Planet Labs RapidEye and Dove series.

The radiometric response and SRF may be affected by mechanical and electronic effects, including satellite launch and sensor drift, as well as by viewing angle and electronic crosstalk. Using up-to-date calibration data from the instrument developer negates these problems. Here, the SRFs recommended by instrument developers or in literature were used, representative of what is done in the wider literature.
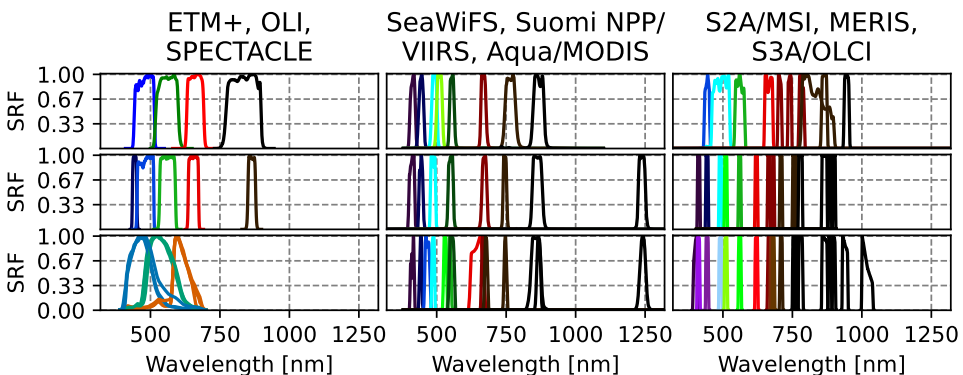

Figure 5.3: Selected spectral response functions (SRFs) of real sensors used in this work, labelled from top to bottom. For this plot, each SRF was normalised to a maximum of 1.

5.3.4   Retrieval algorithm propagation

Finally, the error induced by R-space convolution was propagated through several retrieval algorithms. These were the polynomial OCx chlorophyll-*a* (Chl-a) algorithms [220] for MODIS (OC6, OC3), SeaWiFS (OC4), MERIS (OC4), VIIRS (OC3), and CZCS (OC3), the exponential Ha+17 S2A/MSI Chl-a algorithm [221], and the polynomial Lymburner+16 (LL+16) OLI total suspended matter (TSM) algorithm [420]. These are representative of most multispectral retrieval algorithms in the literature, which differ only in bands used or coefficient values.

Equation (5.7) describes OCx, with [Chl-a] in mg m^{-3}, a_i instrument-specific empirical coefficients, and λ_B, λ_G the instrument's blue and green bands. The Ha+17 algorithm is given in Equation (5.8), with B3, B4 the R_{rs} in the respective S2A/MSI bands. The LL+16 algorithm is given in Equation (5.9), with G, R the R_{rs} in the OLI Green and Red bands, and TSM in mg L^{-1}.

$$\log_{10}([\text{Chl-a}]) = a_0 + \sum_{i=1}^{4} a_i \left[\log_{10}\left(\frac{R_{rs}(\lambda_B)}{R_{rs}(\lambda_G)} \right) \right]^i \tag{5.7}$$

$$[\text{Chl-a}] = 0.80 \exp\left(0.35 \frac{\text{B3}}{\text{B4}} \right) \tag{5.8}$$

$$\text{TSM} = 3957 \left(\frac{G + R}{2} \right)^{1.6436} \tag{5.9}$$

For each input spectrum, both $\bar{R}_{rs}^L$ and $\bar{R}_{rs}^R$ were propagated through each algorithm and the results were compared, analogous to the $\bar{R}_{rs}$ comparison described in Section 5.3.2.

5.4   Results

5.4.1   Simulated instruments

The reflectance convolution error $\Delta\bar{R}_{rs}$ was calculated for the synthetic SRFs described in Section 5.3.3. As an example, Figure 5.4 shows $\Delta\bar{R}_{rs}$ as a function of central wavelength λ_c and FWHM for the *seaswir-a* (see Table 5.1) data. The sign and magnitude of the error depend on the input spectrum. For example, the local minima around 400 and 520 nm correspond to local maxima in the derivative E_d spectrum $dE_d/d\lambda$. Similarly, the local maxima at 480 nm correspond to a local minimum in $dE_d/d\lambda$. Furthermore, the magnitude of $\Delta\bar{R}_{rs}$ increases with wider FWHMs. This is expected since E_d, L_w, and R_{rs} are less spectrally flat over a wider spectral range [430].

Rather than a random error around a median of 0, the difference is a systematic bias in either direction. This is especially clear in Figure 5.4 at $\lambda_c \leq 460$ nm. Being a bias, it needs to be corrected rather than simply incorporated into an error budget. This will be discussed in Section 5.5.

Similar trends were found in the other data sets and with the Gaussian SRFs. For the latter, the λ_c–$\Delta\bar{R}_{rs}$ relation was similar to boxcars with the same FWHM, but larger in magnitude and smoother. This is due to the Gaussian wings covering more of the spectrum than the boxcar's sharp edges. For example, for $\lambda_c = 420$ nm, $\Delta\bar{R}_{rs} = (-1.5\pm0.2)\%$ for a 30 nm boxcar and $(-3.8 \pm 0.4)\%$ for a 30 nm Gaussian, error bars indicating the 5%–95% range, for the
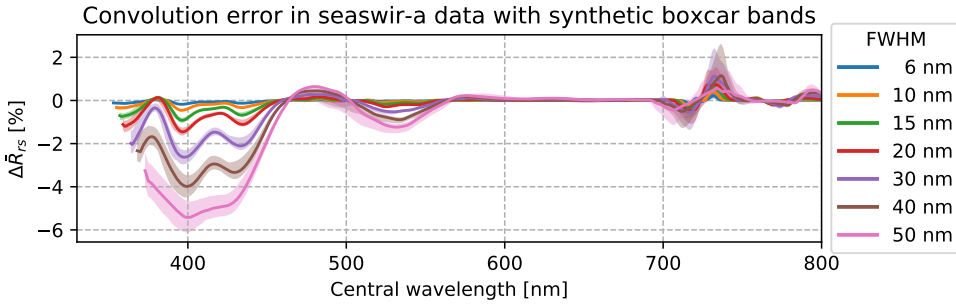
Figure 5.4: Relative reflectance convolution error (Section 5.3.2) in the *seaswir-a* data with boxcar filters of varying FWHM, as a function of central wavelength λ_c. Lines indicate the median error for each filter, shaded areas the 5%–95% range.

seaswir-a data. Finally, the same boxcar filter applied to the *tarao* data gave $\Delta\bar{R}_{rs}(420 \text{ nm}) = (+0.01 \pm 0.02)\%$. This value is much smaller since the *tarao* spectra are smoother than the *seaswir-a* ones; a similar trend was seen across all data sets. These differences highlight the importance of determining this error for each filter and data set, as an ensemble correction is impossible.

5.4.2   Real instruments

$\Delta\bar{R}_{rs}$ was also calculated using the real SRFs described in Section 5.3.3. For example, Figure 5.5 shows the distribution of $\Delta\bar{R}_{rs}$ across all data for the five OLI bands. As with the synthetic sensors, $\Delta\bar{R}_{rs}$ is typically a bias in one direction rather than a random error and its magnitude and sign depend on the input spectrum. For example, in the OLI Blue band $\Delta\bar{R}_{rs} > 0$ for 77% (1380/1799) of spectra while in Green $\Delta\bar{R}_{rs} < 0$ for 80% (1444/1799). Furthermore, a similar trend for larger errors with wider bands was seen, for example in the OLI Green band ($\lambda_c = 562$ nm, FWHM = 57 nm) $\Delta\bar{R}_{rs} = (-0.2^{+0.4}_{-0.9})\%$ while in the similar S3A/OLCI Oa6 band ($\lambda_c = 560$ nm, FWHM = 10 nm) $\Delta\bar{R}_{rs} = (-0.00^{+0.03}_{-0.05})\%$. No significant differences were found between paired instruments such as S3A/OLCI and S3B/OLCI. Some multispectral band-data set combinations are technically invalid (Appendix 5.A); however, these need not be excluded from these overall statistics, as they do not affect the observed trends.

Comparing the convolution error between data sets, as in Figure 5.6 for the OLI Green band, again revealed significant differences. Depending on the data, $\Delta\bar{R}_{rs}$ was a systematic underestimation (*tarao* $\Delta\bar{R}_{rs} = (-0.7 \pm 0.2)\%$), overestimation (*seaswir-r* $\Delta\bar{R}_{rs} = (+0.2 \pm 0.1)\%$), or a random error around 0 (*orinoco* $\Delta\bar{R}_{rs} = (+0.1^{+0.3}_{-0.5})\%$). This is similar to what was observed in Section 5.4.1 and again shows that the error must be quantified separately for each filter and data set.

Low-cost sensors

Finally, the SPECTACLE low-cost sensors [281] are particularly interesting due to their broad bands. The convolution error in their RGB bands, using all data, is shown in Figure 5.7. Interestingly, $\Delta\bar{R}_{rs}$ was largest in the relatively narrow R bands, possibly due to the shapes of the input spectra or the multi-peaked SRFs [281]. Overall, the large magnitude of $\Delta\bar{R}_{rs}$
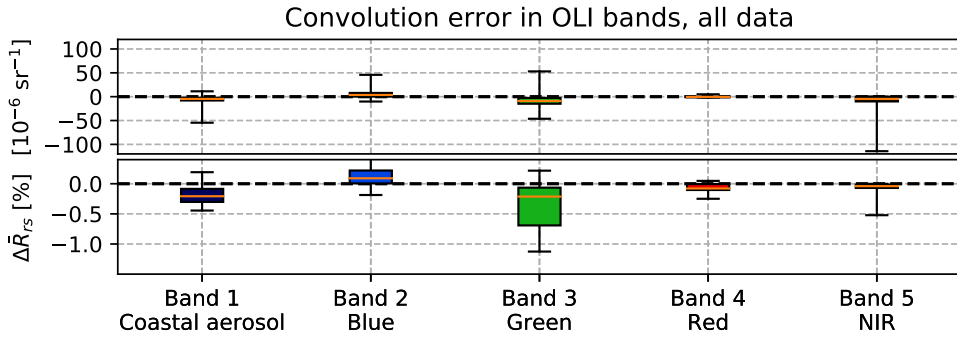
Figure 5.5: Reflectance convolution error in the OLI bands using all data. The boxplots show the error distribution across the data, with an orange median line, first–third quartile boxes, and 5%–95% percentile whiskers. Outliers are not shown for clarity.
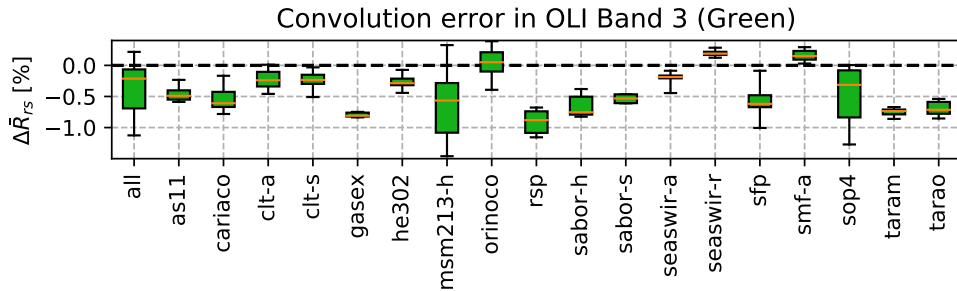

Figure 5.6: Reflectance convolution error in the OLI green band for each data set. The boxplots represent the error distribution within each data set as in Figure 5.5.

(down to −5% in the R bands) highlights the importance of correct spectral convolution for these sensors.

5.4.3   Retrieval algorithms

Finally, the reflectance convolution error was propagated through the retrieval algorithms described in Section 5.3.4. The results for the *smf-a* data set are shown in Figure 5.8. As in the previous sections, the propagated error in Chl-a and TSM was a bias of a few percent. Its sign varied by data set and by algorithm; for example, for the *seaswir-a* data, VIIRS OC3 underestimated Chl-a (ΔChl-a $= -1.4^{+0.7}_{-0.3}\%$) while CZCS OC3 overestimated it (ΔChl-a $= +0.8^{+0.3}_{-0.1}\%$). The magnitude of the error was consistently on the percent level for all data sets and algorithms. These results are representative for most band-ratio algorithms, as discussed in Section 5.3.4.

5.5   Discussion & conclusions

In this work, the effects of incorrectly convolving reflectance when simulating multispectral data (Section 5.2) were investigated. While this error has been pointed out previously [430,
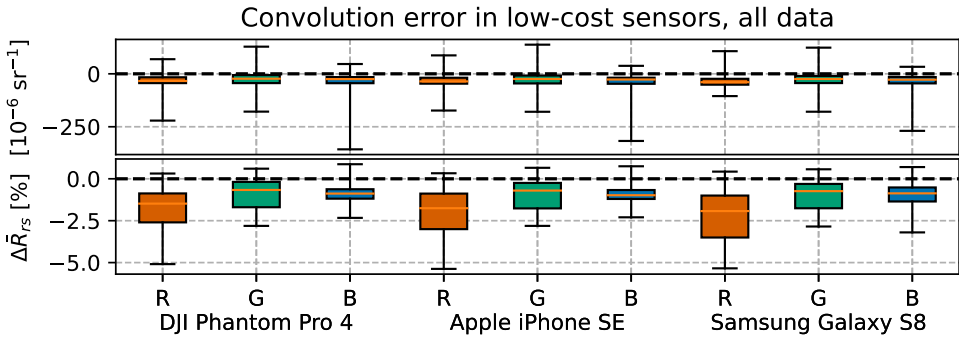
Figure 5.7: Reflectance convolution error in the SPECTACLE sensor RGB bands using all data. The boxplots represent the error distribution within each data set as in Figure 5.5.
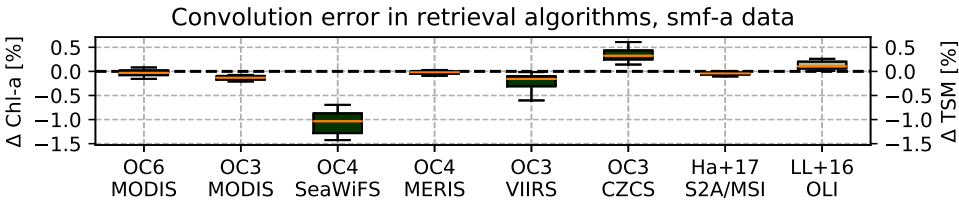

Figure 5.8: Propagated convolution error in the Chl-a (OCx, Ha+17) and TSM (LL+16) retrieval algorithms (Section 5.3.4), using the *smf-a* data set. The boxplots represent the error distribution within each data set as in Figure 5.5.

433, 435], it still commonly occurs in the literature (see Section 5.1). Only one quantitative analysis was found, in which for one data set and one sensor the difference was found to be ≲1% and neglected [430]. However, this result cannot be generalised to all data sets and sensors, as shown in this work.

Significant errors, up to several percent, in the remote sensing reflectance ($\Delta \bar{R}_{rs}$) were found for all data sets and sensor bands (Sections 5.4.1 and 5.4.2). The error was largest near features in the input spectra, particularly peaks in the derivative of E_d ($dE_d/d\lambda$), and for sensors with wide FWHMs, especially low-cost sensors (Section 5.4.2). For example, in the narrow (FWHM ≈ 10 nm) OLCI bands, $|\Delta \bar{R}_{rs}| \lesssim 0.1\%$, while in the wide (FWHM > 50 nm) R bands of low-cost sensors, $|\Delta \bar{R}_{rs}| > 5\%$ for >5% of the spectra. Furthermore, the magnitude and sign of $\Delta \bar{R}_{rs}$ differed significantly between data sets due to varying spectral shapes.

Since uncertainty requirements are typically ±5% for satellite-derived R_{rs}, and even stricter for validation data [88], errors on this scale are significant. Moreover, the error was typically a bias, causing a systematic over- or underestimation of $\bar{R}_{rs}$ and derived products. Preventing such biases is crucial to obtain representative data [88]. Finally, the convolution error is important simply due to its prevalence in the literature [62, 121, 221, 227, 276, 415, 416, 418, 420, 423–425, 431, 432].

If not prevented, the convolution error will create dubious patterns in combined time series. Depending on the data set and sensor, the convolution error is similar to or larger than errors found in existing band-shifting algorithms for combining multispectral time series [91, 419]. With the launch of PACE, for which time series extension is a primary goal [421],

this effect must be accounted for to achieve desired uncertainty requirements [88, 433].

Similarly, incorrect reflectance convolution in instrument validation leads to systematic over- or undercorrections. For example, reflectances from the HydroColor smartphone app have been validated using WISP [121] and HyperSAS [276] data, convolved in R-space, finding significant errors and biases. Biases of -9.5×10^{-4} to $+1.3 \times 10^{-4}$ sr^{-1} were found in the WISP comparison; in Section 5.4.2, the convolution error caused biases on the order of 10^{-4} sr^{-1} for 5%–14% of spectra, varying per band. Errors in the HyperSAS comparison were on the percent level, similar to the errors up to 5% found in Section 5.4.2. Interestingly, in both studies the convolved data underestimated the multispectral data, as would be expected from the negative biases found in this work. This suggests that the convolution error may have contributed a significant part of the error in both studies. However, a direct comparison is difficult due to differing input spectra, as shown in Figure 5.6, and band responses. Thus, the error in these cases cannot definitively be attributed to incorrect convolution. Additionally, many other factors causing significant errors in low-cost sensor data are known [281].

This importance for validation also applies to satellites. For example, in [424], systematic underestimations up to 1% were found in band-average $\bar{R}_{rs}$ (compared to hyperspectral R_{rs}) convolved in R-space with the OLI, MSI, and ETM+ SRFs. This is similar to, and may be explained by, the reflectance convolution error found for these sensors in Section 5.4.2 and shown in Figures 5.5 and 5.6. The same study found no significant errors in convolved VIIRS and OLCI reflectance, agreeing with the correlation between FWHM and error demonstrated in Section 5.4.1.

Conversely, the effects on retrieval algorithms are minor. The convolution error in Chl-a and TSM algorithms (Section 5.4.3) was on the percent level. Since errors in satellite-retrieved Chl-a can be up to 500% [232], a bias of a few percent can safely be neglected. Typical TSM errors are less extreme but still significantly larger than the ≤1% found here [227, 420]. While only a few algorithms were tested, as discussed in Section 5.3.4, these results are representative for most band-ratio algorithms. While many studies opt to use only the central band wavelength, not the full SRF [223, 227], in which case the convolution error does not occur, comparing narrow- and wide-band data that way introduces similar problems, described in [430].

Prevention of the convolution error is straight-forward while post-hoc correction is not. As explained in Section 5.2, simply convolving (ir)radiances instead of reflectance prevents the error from occurring, and is the only mathematically correct procedure. Of course this requires the original data to be available, which is not always true. Post-hoc correction is impossible since the error is highly variable across different sensors and data sets. When lacking original data, the reported uncertainty may simply be increased by a few percentage points [430] but this fails to account for systematic biases. An estimate may be made, for example by reconstructing L_w from a reported R_{rs} and simulated E_d, but this introduces further assumptions.

To this end, it is recommendable that published data sets, intended for satellite validation, contain not only products such as reflectance but also the raw data, at-sensor (ir)radiance data, and calibration data. This way, the convolution error can be avoided. Furthermore, it would greatly increase the amount of data available for other studies requiring radiometric data, such as those into glint removal [206]. Finally, publication of original data, as well as sensor characteristics, allows for traceability, which is crucial for quality control [88].

While this work focused on the remote sensing reflectance R_{rs} using ocean colour data, the principles and conclusions are broadly applicable. A general rule of thumb on convolution

practice is given in Section 5.2.4. In short, the principles outlined in this work are relevant to the simulation of low-resolution data from high-resolution data. This includes all types of reflectance, as well as other divisional quantities such as attenuation coefficients and degree of polarisation. Furthermore, it includes all types of convolution, including spectral and spatial. In all cases, a correct order of operations is crucial to prevent systematic errors. Any simplifications should be justified mathematically, not made at whim.

5.A   Validity of spectral convolution

Consider a source radiance spectrum $L_{src}(\lambda)$, observed by a hyperspectral sensor with N bands. The spectral radiance arriving at the sensor $L(\lambda)$ is the product of the source spectrum and any atmospheric effects. However, the data used in this work were recorded in or only a few meters above the source, so atmospheric effects can safely be ignored:

$$L(\lambda) \approx L_{src}(\lambda) \tag{5.10}$$

The hyperspectral sensor records the radiance in bands $h = 1, 2, ..., N$, with central wavelengths λ_h. Each band h has its own SRF $S_h(\lambda)$, and the radiance recorded in band h, L_h, is the spectral convolution of the at-sensor spectrum $L(\lambda)$ with $S_h(\lambda)$. The integral is evaluated over all wavelengths; for clarity, this is not written explicitly in this section. The denominator in Equation (5.11) corrects for the spectral (or 'quantum') efficiency of the sensor in band h.

$$L_h = \frac{\int L(\lambda)S_h(\lambda)\mathrm{d}\lambda}{\int S_h(\lambda)\mathrm{d}\lambda} \tag{5.11}$$

The resulting spectrum measured by the hyperspectral sensor, $L_H(\lambda)$, consists of the individual band spectra:

$$L_H(\lambda) = L_1, L_2, ..., L_N = \frac{\int L(\lambda)S_1(\lambda)\mathrm{d}\lambda}{\int S_1(\lambda)\mathrm{d}\lambda}, \frac{\int L(\lambda)S_2(\lambda)\mathrm{d}\lambda}{\int S_2(\lambda)\mathrm{d}\lambda}, ..., \frac{\int L(\lambda)S_N(\lambda)\mathrm{d}\lambda}{\int S_N(\lambda)\mathrm{d}\lambda} \tag{5.12}$$

S_h affects $L(\lambda)$ in two ways. The first is to lower it due to the spectral efficiency of the sensor. This is described by an overall SRF $S_H(\lambda)$; dividing the data by $S_H(\lambda)$ corrects for this. The second effect is to smoothen the data: since in practice S_h is never a delta function, band h records not only the radiance at its central wavelength λ_h but also at other wavelengths where $S_h(\lambda) > 0$.

The smoothening can be described as a cross-correlation ($\star$) between the observed radiance $S_H(\lambda)L(\lambda)$ and a bandwidth function G. In reality, each band will have a slightly different G_h, for example due to stray light; however, for simplicity, here G is assumed to be the same for all bands. Then $L_H(\lambda)$ can be described as in Equation (5.13).

$$L_H(\lambda) = \frac{(S_H L) \star G}{S_H}(\lambda) \tag{5.13}$$

Now consider a multispectral band M with SRF $S_M(\lambda)$. Following the same logic, Equation (5.11) gives the radiance recorded in band M, L_M, as in Equation (5.14).

$$L_M = \frac{\int L(\lambda) S_M(\lambda) \mathrm{d}\lambda}{\int S_M(\lambda) \mathrm{d}\lambda} \tag{5.14}$$

However, when simulating multispectral data from hyperspectral data, the original radiance $L(\lambda)$ is not available. Instead, the recorded hyperspectral radiance $L_H(\lambda)$ is used. This means that in practice, one does not calculate L_M as in Equation (5.14) but an approximation L_M^H, as in Equation (5.15).

$$L_M^H = \frac{\int L_H(\lambda) S_M(\lambda) \mathrm{d}\lambda}{\int S_M(\lambda) \mathrm{d}\lambda} = \frac{\int \frac{(S_H L) \star G}{S_H}(\lambda) S_M(\lambda) \mathrm{d}\lambda}{\int S_M(\lambda) \mathrm{d}\lambda} \tag{5.15}$$

The approximation $L_M^H \approx L_M$ holds in two cases. The first is if $L_H(\lambda) \approx L(\lambda)$, that is if Equation (5.16) holds. From information theory it follows that this is true if G is significantly narrower than typical features in $L(\lambda)$.

$$\frac{(S_H L) \star G}{S_H}(\lambda) \approx L(\lambda) \tag{5.16}$$

The second case where $L_M^H \approx L_M$ holds is when the multispectral band M is significantly wider than G and typical features in $L(\lambda)$. Then, any radiance redistributed from λ_h to surrounding wavelengths in $L_H(\lambda)$ is still captured in the integral $\int L_H(\lambda) S_M(\lambda) \mathrm{d}\lambda$, and the value of the integral is the same.

E_d has narrow line features, as does L_w by extension. Hyperspectral (ir)radiance sensors typically undersample these features [97, 120], so Equation (5.16) does not hold in practice.

However, the second case does hold, if M is significantly wider than the hyperspectral band (or G). The Nyquist-Shannon theorem provides, to first order, a requirement: $L_M^H \approx L_M$ if the FWHM of M is at least twice that of the hyperspectral data. Then, hyperspectral data $L_H(\lambda)$ adequately approximate the original radiance $L(\lambda)$ for spectral convolution purposes.

5.B Radiometric data

Table 5.1 lists the radiometric data sets used in this work. Some of these contain unphysical data due to measurement errors, environmental effects, and instrumental problems [206, 431, 437]. This appendix describes how the data were filtered and homogenised before processing.

First, all spectra were converted to SI units. Second, negative R_{rs} values were clipped to 0 if $-10^{-4} < R_{rs}(\lambda) < 0$ as this is within typical measurement errors; spectra with any $R_{rs}(\lambda) \leq -10^{-4}$ were removed wholly. For *as11*, 5 spectra with negative R_{rs} were removed. For *cariaco*, 230 spectra missing R_{rs} and 11 missing E_d values were removed, as were 64 spectra with negative and 1 spectrum with unphysically high (>0.8) R_{rs}. For *clt-a*, 6 spectra with negative R_{rs} were removed. For *clt-s*, the spectra within 3-minute windows suggested in the accompanying documentation were averaged and L_w was calculated from L_u and L_t following the Mobley protocol [209]; 19 spectra with missing and 36 with negative R_{rs} were removed. For *gasex*, wavelengths $\lambda > 710$ nm were removed due to incomplete data. For *he302*, 3 spectra with $R_{rs}(800 \text{ nm}) \geq 0.003$ were removed as outliers; the original authors noted the difficulty in normalising these data [206]. For *msm213-h*, L_w was used to reconstruct E_d; 179 spectra with missing data were removed, as were spectra with unphysically large jumps in E_d, namely 21 with $|E_d(\lambda_1) - E_d(\lambda_2)| \geq 0.2$ and $|E_d(\lambda_2) - E_d(\lambda_3)| \geq 0.2$ and 1

Label	Reference	Location	Instrument	N	Spectra Range	FWHM	Step
as11	[438]	Arabian Sea	HyperTSRB	8	398 – 798	10	3.3
cariaco	[439]	Caribbean Sea	SpectraScan PR650	58	380 – 780	8	4
clt-a	[440]	Chesapeake Bay	ASD	56	350 – 1000	3	2
clt-s	[440]	Chesapeake Bay	HyperSAS	161	354 – 796	10	2
gasex	[441]	South Atlantic Ocean	HyperTSRB	7	352 – 710	10	2
he302	[442]	North Sea	RAMSES	41	320 – 950	10	5
msm213-h	[443]	North Atlantic Ocean	HyperPro	56	360 – 800	10	5
orinoco	[444]	Caribbean Sea, Orinoco River	SpectraScan PR650	70	380 – 780	8	4
rsp	[445]	Atlantic Ocean, Pacific Ocean	HyperPro	27	349 – 802	10	3.3
sabor-h	[446]	Atlantic Ocean	HyperPro	12	349 – 802	10	3.3
sabor-s	[446]	Atlantic Ocean	HyperSAS	6	358 – 749	10	1
seaswir-a	[37]	Gironde, La Plata	ASD	80	350 – 1300	3	1
seaswir-r	[37]	Gironde, La Plata, Scheldt	RAMSES	198	350 – 900	10	2.5
sfp	[447]	Florida Strait	SVC GER 1500	16	399 – 801	3	1.5
smf-a	[448]	Gulf of Mexico	ASD	46	350 – 1000	3	2
sop4	[449]	North Sea, Skagerrak	RAMSES	870	360 – 750	10	5
taram	[450]	Mediterranean Sea	HyperPro	42	348 – 803	10	3.3
tarao	[451]	Global	HyperPro	45	401 – 750	10	3.3

Table 5.1: Summary of radiometric data sets used in this work. Spectral range, FWHM, and step are in nanometers, the former rounded to the nearest integer. The number of spectra N is the number after post-processing and removal of bad data.

5

with $|E_d(\lambda_1) - E_d(\lambda_2)| \geq 0.35$, with λ_1, λ_2, λ_3 subsequent wavelengths. For *orinoco*, 1 spectrum with negative R_{rs} was removed. For *sabor-s*, which contains polarised and unpolarised spectra, only the latter were used and the wavelength range was clipped to 358–749 nm because of incomplete and noisy data elsewhere. For *seaswir-a*, the provided plaque radiance L_d was used to calculate $E_d = \pi L_d$ [209], assuming a plaque reflectance of $R_g \approx 1$ [431]; no units were given for these data, so the resulting E_d spectra were divided by 10^5 to be in line with the others. 'Water reflectance' R_w was provided instead of R_{rs}; comparing [431] and [209] showed that $R_{rs} = R_w/\pi$. Finally, 35 spectra with missing E_d and 39 with negative R_{rs} were removed. For *seaswir-r*, R_w was similarly converted to R_{rs} and 2 spectra with negative R_{rs} were removed. For *sop4*, the provided L_u and L_s were used to calculate R_{rs} following the Mobley protocol [209] for simplicity [206]. The R_{rs} spectra were then normalised by subtracting $R_{rs}(750\,\text{nm})$ [209] and the results used to re-calculate L_w. Next, 885 spectra with unphysical $\max(E_d(\lambda)) < 0.01$ were removed and the wavelength range cropped to 360–750 nm to remove noisy data. Spectra with unphysical features were then removed, namely 23 with $|R_{rs}(\lambda_1) - R_{rs}(\lambda_2)| \geq 0.005$ and $|R_{rs}(\lambda_2) - R_{rs}(\lambda_3)| \geq 0.005$ and 3 with $E_d(400\,\text{nm}) - E_d(405\,\text{nm}) > 0.01$; finally, 289 spectra with negative R_{rs} were removed. The remaining data sets (*sabor-h*, *sfp*, *rsp*, *taram*, and *tarao*) required no post-processing. For *sfp*, only the mean R_{rs} spectra were used.

Funding

Acknowledgements

5

5