Remote Sensing of Environment 177 (2016) 107–124

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Uncertainty estimates of remote sensing reflectance derived from comparison of ocean color satellite data sets

F. Mélin [a,*], G. Sclep [a,b], T. Jackson [c], S. Sathyendranath [d]

[a] European Commission, Joint Research Centre (JRC), Institute for Environment and Sustainability (IES), TP270, via Fermi 2749, Ispra 21027, Italy
[b] European Commission, Publications Office of the European Union, 2 rue Mercier, Luxemburg L-2985, Luxembourg
[c] Plymouth Marine Laboratory (PML), Prospect Place, The Hoe, Plymouth PL1 3DH, UK
[d] National Centre for Earth Observation, Plymouth Marine Laboratory (PML), Prospect Place, The Hoe, Plymouth PL1 3DH, UK

ARTICLE INFO

Article history:
Received 12 May 2015
Received in revised form 25 January 2016
Accepted 5 February 2016
Available online 22 February 2016

Keywords:
Ocean color
Remote sensing reflectance
Uncertainty

ABSTRACT

Assigning uncertainty to ocean-color satellite products is a requirement to allow informed use of these data. Here, uncertainty estimates are derived using the comparison on a 12th-degree grid of coincident daily records of the remote-sensing reflectance R_{RS} obtained with the same processing chain from three satellite missions, MERIS, MODIS and SeaWiFS. The approach is spatially resolved and produces σ, the part of the R_{RS} uncertainty budget associated with random effects. The global average of σ decreases with wavelength from approximately 0.7–0.9 10^{-3} sr^{-1} at 412 nm to 0.05–0.1 10^{-3} sr^{-1} at the red band, with uncertainties on σ evaluated as 20–30% between 412 and 555 nm, and 30–40% at 670 nm. The distribution of σ shows a restricted spatial variability and small variations with season, which makes the multi-annual global distribution of σ an estimate applicable to all retrievals of the considered missions. The comparison of σ with other uncertainty estimates derived from field data or with the support of algorithms provides a consistent picture. When translated in relative terms, and assuming a relatively low bias, the distribution of σ suggests that the objective of a 5% uncertainty is fulfilled between 412 and 490 nm for oligotrophic waters (chlorophyll-*a* concentration below 0.1 mg m^{-3}). This study also provides comparison statistics. Spectrally, the mean absolute relative difference between R_{RS} from different missions shows a characteristic U-shape with both ends at blue and red wavelengths inversely related to the amplitude of R_{RS}. On average and for the considered data sets, SeaWiFS R_{RS} tend to be slightly higher than MODIS R_{RS}, which in turn appear higher than MERIS R_{RS}. Biases between mission-specific R_{RS} may exhibit a seasonal dependence, particularly in the subtropical belt.

1. Introduction

To fulfill the potential of ocean-color remote sensing, uncertainty estimates should be assigned to satellite products. Meeting this requirement should help enhance user confidence, define the range of possible applications of data products and ultimately allow informed use of the data. Uncertainty values are needed to operate data assimilation in ecological and climate models (Cummings et al., 2009; Gregg, 2008), perform the merging of different data sets (Pottier et al., 2006) or to allow trend analysis in climate research (Vantrepotte & Mélin, 2011). Considering that water-leaving radiance, or equivalently remote-sensing reflectance (R_{RS}), and chlorophyll-a concentration (Chl*a*) are listed among the Essential Climate Variables (ECV) (GCOS, 2011; Bojinski et al., 2014), climate applications are particularly relevant and different approaches have been developed to assess series of data products (Mélin & Franz, 2014). The requirement to have uncertainty estimates associated with the satellite products distributed to the user community is now included in some projects or mission planning (Donlon et al., 2012; Hollmann et al., 2013).

Much of the knowledge gathered about uncertainties of satellite ocean color products is the result of validation activities comparing satellite data with field measurements. Even though required as a major element of any satellite mission (McClain, Feldman, & Hooker, 2004), validation is limited in terms of spatial and/or temporal distribution. Some time series exist for a few locations but the spatial distribution of the validation data points is very uneven, with a sparse coverage of the open ocean regions even when large in situ Chl*a* data bases are used (Gregg & Casey, 2004). This relative scarcity of validation data is even more apparent for R_{RS} products (Bailey & Werdell, 2006). In order to extend validation results to scales associated with satellite data, several techniques have been proposed, mostly for derived products (Chl*a* or inherent optical properties, IOPs) (Mélin & Franz, 2014). For instance error propagation exercises can be conducted with bio-optical algorithms (Lee, Arnone, Hu, Werdell, & Lubac, 2010), and the effect of uncertainties in the algorithm parameters can be

* Corresponding author.

http://dx.doi.org/10.1016/j.rse.2016.02.014

estimated by running the inversion multiple times (Wang, Boss, & Roesler, 2005). Non-linear inversion of bio-optical models through the minimization of a cost function can provide an inversion confidence (e.g., Maritorena, Hembise Fanton d'Andon, Mangin, & Siegel, 2010; Salama, Dekker, Su, Mannaerts, & Verhoef, 2009; Werdell et al., 2013).

For R_{RS}, less development has come to fruition. Actually, some of the techniques applied to estimate uncertainties for bio-optical algorithms rely on fixed values for the uncertainty of R_{RS}. On the other hand, the work by Salama & Stein (2009) distinguishes different contributions to the uncertainty associated with IOPs, including that due to the input R_{RS}. Some atmospheric correction schemes rely on the minimization of cost functions (Chomko & Gordon, 2001; Stamnes et al., 2003) and would therefore be amenable to the calculation of inversion confidence estimates as explored with bio-optical algorithms. A Bayesian methodology for atmospheric correction (Frouin & Pelletier, 2015) produces an uncertainty estimate characterizing the inversion process. Exercises of accuracy analysis (e.g., Bulgarelli & Zibordi, 2003; Bulgarelli, Mélin, & Zibordi, 2003; Neukermans et al., 2009) can also provide valuable insights on atmospheric correction performance to support uncertainty assessments. But further developments are certainly needed to produce robust estimates of the uncertainty in R_{RS}.

In that context, the comparison of coincident satellite data through collocation analysis has the potential to support the characterization of their uncertainties. Collocation techniques have been applied to geophysical data sets in a variety of disciplines, oceanography (e.g., Caires & Sterl, 2003; Stoffelen, 1998; O'Carroll, Eyre, & Saunders, 2008), atmospheric sciences (Toohey & Strong, 2007), hydrology (Roebeling, Wolters, Meirink, & Leijnse, 2012) or the study of terrestrial vegetation or moisture (Fang, Wei, Jiang, & Scipal, 2012; D'Odorico et al., 2014; Scipal, Holmes, de Jeu, Naeimi, & Wagner, 2008). As for ocean color, an application dealt with the characterization of uncertainty values associated with Chl*a* (Mélin, 2010). Albeit not a substitute for the actual validation of R_{RS} satellite products with field data, the large amount of matching data from different missions represents a spatially-resolved reservoir of precious information. By comparing coincident satellite retrievals, this paper aims at adopting the approach previously applied to Chl*a* (Mélin, 2010) to estimate uncertainty terms associated with R_{RS}. Specifically, the study focuses on the zero-mean random error of the uncertainty budget and the biases existing between satellite data. The method is first introduced before describing the results. Then these are compared with independent estimates obtained by other routes.

2. Comparison of satellite data

This section presents the approach followed to compare coincident satellite data and how the comparison can be used to derive uncertainty estimates (Mélin, 2010). Let us consider two ensembles of N coincident satellite derived values associated with the same quantity, $(x_i)_{i=1,N}$ and $(y_i)_{i=1,N}$. They can be compared with several metrics, e.g., describing their differences (in units of x and y) and relative differences (in %) (Mélin & Franz, 2014), including the mean absolute relative difference $|\psi|$ and the mean relative difference (or relative bias) ψ:

$$|\psi| = \frac{1}{N}\sum_{i=1}^{N}\frac{2|y_i - x_i|}{y_i + x_i} \tag{1}$$

$$\psi = \frac{1}{N}\sum_{i=1}^{N}\frac{2(y_i - x_i)}{y_i + x_i}. \tag{2}$$

In these metrics, the numerator is divided by the average of the two values being compared. This avoids arbitrarily selecting one product as the value of reference, while numerically it has the advantage of preventing cases where only the denominator is close to zero (if x and y are positive as imposed in this study, see Section 3.1). Additionally, the mean difference (or bias) can be expressed simply as:

$$\delta = \frac{1}{N}\sum_{i=1}^{N}(y_i - x_i) = \overline{y} - \overline{x} \tag{3}$$

where the overbar symbol indicates an average value.

Considering scatter plots comparing R_{RS} data obtained from different missions (e.g., Kahru, Lee, Kudela, Manzano-Sarabia, & Mitchell, 2015; Mélin, Zibordi, & Djavidnia, 2009; Vanhellemont, Neukermans, & Ruddick, 2014), a linear model appears an appropriate framework to compare two R_{RS} data sets. We assume that x and y can be modeled linearly as a function of a reference state r and zero-mean random errors, ε and ζ, respectively:

$$x_i = r_i + \epsilon_i \tag{4}$$

$$y_i = \alpha + \beta r_i + \zeta_i \tag{5}$$

where α and β are, respectively, the additive and multiplicative biases between x and y. Another assumption is that ε and ζ are uncorrelated and independent of r. The reference state r only serves to relate the two data sets being analyzed. Ideally, r should then be related to the true state in order to complete the uncertainty budget.

The aim of the analysis is to derive an estimate of the standard deviations σ_ε and σ_ζ of the terms ε and ζ. In this mathematical framework, these σ terms can be seen as the random error associated with x and y. Working with a real and finite set of satellite data, the error actually cannot be known, and the estimates of the σ terms should be considered as a measure of the uncertainty due to random effects (GUM, 2008). The determination of σ_ε and σ_ζ can be done with the following development. First, variance and covariance terms for x and y, σ_x^2, σ_y^2 and σ_{xy}, can be written as:

$$\sigma_x^2 = \sigma_r^2 + \sigma_\epsilon^2 \tag{6}$$

$$\sigma_{xy} = \beta\sigma_r^2 \tag{7}$$

$$\sigma_y^2 = \beta^2\sigma_r^2 + \sigma_\zeta^2 \tag{8}$$

using the fact that covariance terms containing ϵ or ζ are null. Further, eliminating the variance of r, σ_r, within this set of three equations leads to the following system of two equations (Toohey & Strong, 2007):

$$\sigma_\epsilon^2 = \sigma_x^2 - \frac{1}{\beta}\sigma_{xy} \tag{9}$$

$$\sigma_\zeta^2 = \sigma_y^2 - \beta\sigma_{xy}. \tag{10}$$

This system contains three unknowns: β and the variance of the random terms σ_ε^2 and σ_ζ^2 that are the quantities of interest.

To remove the ambiguity of the system, the ratio of the standard deviations associated with ε and ζ is considered known, defining the following quantity: $\lambda = \sigma_\zeta/\sigma_\varepsilon$. Knowing λ, β can be computed from the variance and covariance of x and y. For instance by multiplying Eq. (9) by λ^2, the term on its right-hand side becomes equal to that of Eq. (10), leading, after multiplication by β, to the 2nd-degree equation:

$$\beta^2\sigma_{xy} + \left(\lambda^2\sigma_x^2 - \sigma_y^2\right)\beta - \lambda^2\sigma_{xy} = 0. \tag{11}$$

β being the slope of linear (model II) regression, the solution is written with the plus sign (Legendre & Legendre, 1998):

$$\beta = \frac{\sigma_y^2 - \lambda^2\sigma_x^2 + \sqrt{\left(\sigma_y^2 - \lambda^2\sigma_x^2\right)^2 + 4\lambda^2\sigma_{xy}^2}}{2\sigma_{xy}}. \quad (12)$$

Eq. (11) can also be expressed in terms of $1/\beta$, leading to an equivalent way of writing the solution for β:

$$1/\beta = \frac{\sigma_x^2 - \sigma_y^2/\lambda^2 + \sqrt{\left(\sigma_y^2/\lambda^2 - \sigma_x^2\right)^2 + 4\sigma_{xy}^2/\lambda^2}}{2\sigma_{xy}}. \quad (13)$$

The substitution of $1/\beta$ and β into Eqs. (9) and (10), respectively, leads to the solutions for σ_ε and σ_ζ:

$$\sigma_\epsilon^2 = \frac{\sigma_x^2 + \sigma_y^2/\lambda^2 - \sqrt{\left(\sigma_y^2/\lambda^2 - \sigma_x^2\right)^2 + 4\sigma_{xy}^2/\lambda^2}}{2} \quad (14)$$

$$\sigma_\zeta^2 = \frac{\lambda^2\sigma_x^2 + \sigma_y^2 - \sqrt{\left(\sigma_y^2 - \lambda^2\sigma_x^2\right)^2 + 4\lambda^2/\sigma_{xy}^2}}{2}. \quad (15)$$

Most of the results presented in the paper are constructed with λ assumed equal to 1 (in that case, $\sigma = \sigma_\varepsilon = \sigma_\zeta$), which implies that the different products are taken to have the same level of random error in their error budget. This assumption is further discussed in the next section and tested with different values of λ in Section 4.1. Based on the accumulation of data for a given location, the σ estimates are spatially resolved but the capacity of the approach to provide temporal information is very limited, being dependent on the number of coincident data and their distribution in time. With the considered ocean color data sets, the study computed multi-annual values as well as seasonal estimates (see Section 4.2).

It is stressed that in the framework described above, σ_ε and σ_ζ are not associated with a comparison metric but are intrinsic properties of the data sets x or y, respectively, even though by construction these σ terms are mathematically dependent upon the matching data set (Eqs. (14) and (15)). This is why the σ terms are the main focus of this study since they provide an independent estimate of part of the uncertainty budget associated with the considered data sets. For completeness, results will also be shown for the comparison metrics, δ, $|\psi|$ and ψ.

3. Data sets

3.1. Satellite data sets

Remote-sensing reflectance daily data were obtained from the National Aeronautics and Space Administration (NASA) archive of Level-3 gridded data from the Sea-viewing Wide Field-of-view Sensor (SeaWiFS, McClain et al., 1998), the Moderate Resolution Imaging Spectroradiometer (MODIS, Esaias et al., 1998) onboard the Aqua platform, and the Medium Resolution Imaging Spectrometer (MERIS, Rast, Bezy, & Bruzzi, 1999). The three missions are treated with a common strategy for calibration and processed with the same standard algorithms (Franz, Bailey, Werdell, & McClain, 2007), with data associated with reprocessing versions 2010.0, 2013.1 and 2012.1, respectively. The so-called binned satellite products are arranged on a sinusoidal grid with a spatial resolution of 12th-degree (for SeaWiFS) and 24th-degree (for MODIS and MERIS). The SeaWiFS Level-3 data are based on Global Area Coverage (GAC) imagery, made up of full-resolution pixels (1.1-km at nadir) subsampled every fourth pixel every fourth scan line (see Mélin, Steinich, Gobron, Pinty, & Verstraete, 2002). With respect to MODIS and MERIS data, for which all pixels are fed into the Level-3 binning process, the SeaWiFS Level-3 data are thus characterized by a scarcer information content. How this difference in Level-3 construction specifically affects their uncertainties would require a dedicated study that is not undertaken here.

The Quasi-Analytical Algorithm (QAA) (version 5, Lee, Carder, & Arnone, 2002; Lee, Lubac, Werdell, & Arnone, 2009) was applied to these data to produce fields of inherent optical properties (IOPs), namely phytoplankton absorption coefficient a_{ph}, absorption due to colored detrital matter (CDM) a_{cdm} (itself due to chromophoric dissolved organic matter (CDOM) and non-pigmented particles), and backscattering coefficient of particles b_{bp}.

The comparison of R_{RS} data from different satellite missions is hampered by the different set of wavelengths associated with their respective sensors (Table 1, Mélin & Franz, 2014). Band-shifting schemes have been proposed for validation analysis (Zibordi, Berthon, Mélin, D'Alimonte, & Kaitala, 2009a) or for the purpose of merging (Mélin et al., 2011a), where band-shifting refers to the process of correcting for band differences between different multi-spectral data sets. In the current study, a generic band-shifting approach that has recently been proposed and assessed (Mélin & Sclep, 2015) was operated to express R_{RS} values from one mission onto the wavelengths of another mission. The scheme first relies on the application of the QAA to produce IOPs (a_{ph}, a_{cdm} and b_{bp}) at 443 nm. Using characteristic spectral shapes, these IOPs are expressed on the wavelengths of interest and the algorithm is operated in a forward mode to generate R_{RS} at the target wavelengths. The uncertainty incurred by band-shifting is estimated in most cases at well below 5% of the reflectance value (with averages of typically 1–2%), especially for open ocean regions (Mélin & Sclep, 2015).

This scheme was applied to express the MODIS and MERIS R_{RS} values at the SeaWiFS bands, and the MODIS R_{RS} at the MERIS bands, so that the data between these missions could be readily compared (Table 1). The reported statistics are all provided at the SeaWiFS bands for a common comparison. Results are well consistent if another set of wavelengths is selected, for instance if the MERIS bands are used to compare MERIS and MODIS data (not shown). The last pre-processing step before comparison was to average the MERIS and MODIS data onto the SeaWiFS 12th-degree grid to allow a direct match-up of the grid points (this was done if at least 3 higher-resolution values were valid within the 12th-degree SeaWiFS bin). The three comparison pairs are referred to as MoSe, SeMe and MeMo (MODIS/SeaWiFS, SeaWiFS/MERIS, and MERIS/MODIS, respectively).

The collocation analysis relies on periods of overlap between missions. The period of reference, for which results are given unless specified otherwise, is selected as 2003–2007, when the three missions functioned in an optimal way without any data gaps. To test the effect of additional years, calculations were also conducted over 2003–2011 (9 years) for the comparison MERIS/MODIS and for a composite period made of July–December 2002, 2003–2007, April–June 2008, January–March 2009, and 2010 when SeaWiFS is involved (for a total of 7 years of data). This composite period was constructed to account for the data gaps in the SeaWiFS record while assuring an even representation of all calendar months to avoid affecting the results with a possible seasonal dependence.

Table 1
Nominal center-wavelengths (in nm) associated with each satellite mission, SeaWiFS (SWF), MODIS (MOD) and MERIS (MER), and with the comparison between pairs of mission.

SWF	MOD	MER	Comparison
412	412	412	412
443	443	443	443
490	488	490	490
510	531	510	510
555	547	560	555
670	667	665	670

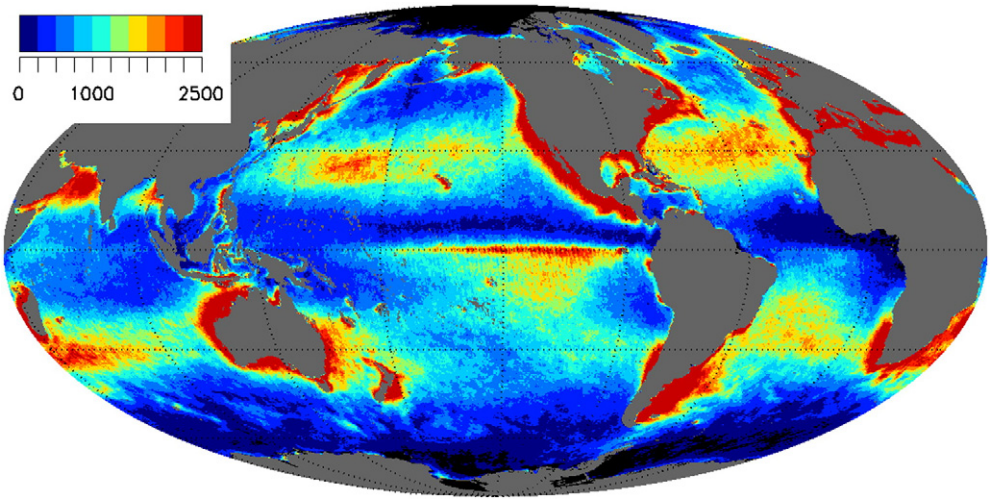

Fig. 1. Number of coincident daily matching pairs between SeaWiFS and MODIS accumulated over the period 2003–2007 in 3rd-degree macro-bins. Among macro-bins with a number of matching pairs larger than 0, it exceeds 2500 (maximum of the color scale) in 4% of the cases.

The first step of the analysis was the identification on a daily basis of all matching data pairs with valid values (i.e., positive R_{RS}) for the two missions to be compared (negative R_{RS} were excluded since they might restrict the use of the QAA and of statistical calculations such as Eqs. (1) and (2); furthermore, a negative value might suggest a retrieval of lower quality by the atmospheric correction). The spatial distribution of coincident data being very variable (see Fig. 1), all daily matching pairs from each spatial bin were accumulated into macro-bins of 4×4 bins in order to increase the statistical basis of the calculations. So the comparison statistics were obtained on a 3rd-degree grid but it is stressed that the matching pairs were compared at the original resolution without prior averaging. A specific analysis was conducted using the MERIS and MODIS products at their original resolution: in that case, the matching pairs are obtained at a 24th-degree resolution and the macro-bins contain 8×8 original bins. Statistics were further reduced on the basis of regional averages computed on biogeographic provinces that constitute a partition of the world's ocean (Longhurst, 2006) or on ensembles of these provinces.

The performance of the atmospheric correction appears very similar for the three missions when evaluated in terms of aerosol optical thickness with globally distributed field data (Mélin, Zibordi, & Holben, 2013; Mélin, Zibordi, Carlund, Holben, & Stefan, 2013), which is consistent with the fact that the satellite data sets were produced with the same calibration strategy and atmospheric correction scheme (Franz et al., 2007). Furthermore, validation statistics obtained for R_{RS} over a more restricted set of locations are as well consistent across missions, especially when results are derived at the same site for the three missions (Antoine et al., 2008; Mélin & Franz, 2014; Mélin et al., 2011b; Mélin, Zibordi, & Berthon, 2012; Zibordi et al., 2009a). So at least for these data sets, the assumption of λ close to 1 appears appropriate but departures from 1 will be analyzed and discussed (see Sections 4 to 6).

3.2. Field data sets

Field data were used to provide validation statistics to be compared with the results of the collocation analysis. Validation statistics from published studies were updated with similar match-up selection protocols, band-shifting procedures and statistical analyses (Mélin & Franz, 2014; Mélin et al., 2011b; Zibordi et al., 2009a; Zibordi, Mélin, Berthon, & Talone, 2014).

Three sites are part of the Ocean Color component of the Aerosol Robotic Network (AERONET-OC, Zibordi et al., 2006a, 2009b): the Acqua Alta Oceanographic Tower (AAOT), a coastal site located in the northern Adriatic Sea (Zibordi, Mélin, & Berthon, 2006b), and two sites located in the Baltic Sea, the Gustaf Dalen Lighthouse Tower (GDLT) along the coast of Sweden (Zibordi, Strömbeck, Mélin, & Berthon, 2006c) and the Helsinki Lighthouse Tower (HLT) in the Gulf of Finland. From an optical point of view, these sites can be characterized as having moderately sediment-dominated waters (AAOT) and CDOM-dominated absorbing waters (GDLT and HLT), but they actually show a variety of optical conditions (Zibordi et al., 2009a). A fourth site is the Marine Optical BuoY (MOBY) (Clark et al., 1997) operating in deep oligotrophic waters offshore Lanai (Hawaii). A common feature shared by these sites is that they have produced a large body of accurate measurements of R_{RS} data (Brown et al., 2007; Gergely & Zibordi, 2014) through the operation of autonomous systems. That has been translated to a varying extent into a significant set of match-up points for validation analysis allowing meaningful statistics to be derived.

In the comparison between field observations and satellite retrievals, the comparison framework introduced in Section 2 (Eqs. (4) and (5)) can be used, with $(x_i)_{i=1,N}$ and $(y_i)_{i=1,N}$ being the field data and the satellite match-up data, respectively. Differently from Section 2, the uncertainties associated with field measurements are considered known. For the three AERONET-OC sites, the selected uncertainty values are those presented by Gergely & Zibordi (2014) in units of R_{RS} (sr^{-1}), while a relative uncertainty of 5% (12.5% in the red) is assumed for the MOBY data (Brown et al., 2007) that is multiplied by the average of the data used for validation in order to express the uncertainty in sr^{-1}. These uncertainties are assumed to provide estimates of σ_ε (considering that they are not affected by a bias term). Knowing σ_ε, the system made of Eqs. (9) and (10) yields σ_ξ:

$$\sigma_\xi^2 = \sigma_y^2 - \frac{\sigma_{xy}^2}{\sigma_x^2 - \sigma_\epsilon^2}. \tag{16}$$

In Section 5, this value will be referred to simply as σ obtained through validation (noted σ^v).

Validation studies often report two comparison metrics, the root-mean-square (RMS) difference Δ and the average bias δ (Eq. (3)). From these, another comparison metric not affected by biases can be computed, the unbiased RMS difference Δ_u between satellite and field values of R_{RS} written as:

$$\Delta_u = \left(\frac{1}{N} \sum_{i=1}^{N} (y_i - x_i + \overline{x} - \overline{y})^2 \right)^{1/2} = \sqrt{\Delta^2 - \delta^2} \tag{17}$$

where N is the number of match-ups between field data $(x_i)_{i=1,N}$ and satellite values $(y_i)_{i=1,N}$. Values of Δ_u will also be presented in the analysis since it is available from literature and from some projects such as the Ocean Colour Climate Change Initiative (see Section 5). For completeness, relationships between σ and Δ_u are provided, derived by substituting y_i as a function of x_i (Eqs. (4) and (5)) into Eq. (17), and considering the relationship between σ_x and σ_r (Eq. (6)):

$$\Delta_u^2 = (\beta-1)^2\sigma_r^2 + \sigma_\epsilon^2 + \sigma_\zeta^2 \quad (18)$$

$$= (\beta-1)^2\sigma_x^2 + \beta(2-\beta)\sigma_\epsilon^2 + \sigma_\zeta^2. \quad (19)$$

These relationships indicate that Δ_u is larger than σ_ζ, with the two values converging when σ_ε associated with field data is small and β is close to 1.

It is recalled that the MOBY site is used for vicarious calibration for the three missions, so that by construction the bias between satellite and field data is minimized at this location (the vicarious calibration process adjusts the calibration coefficients applied to the instrument to minimize the difference between satellite R_{RS} obtained with the atmospheric correction and selected field observations; Franz et al., 2007; Zibordi et al., 2015). But again, considering that bias terms are not included in the metrics considered (σ or Δ_u), this is not a limitation of the study.

4. Results

The distribution of matching pairs for the comparison MODIS/SeaWiFS (MoSe) as obtained within the reference period (2003–2007) is shown in Fig. 1 on the 3rd-degree grid used for the analysis. The number of match-ups is found to be greatest around some mid-latitude continental land masses, whereas regions characterized by a frequent cloud coverage, like the Pacific and Atlantic intertropical convergence zones, show a low number of match-ups. The frequency of match-ups also tends to decrease in high-latitude areas. This distribution is variable in time too; for instance the Arabian Sea is associated with a high number of match-ups but their frequency is actually much reduced in boreal summer during the summer monsoon regime. The average number of match-ups for each 3rd-degree macro-bin is equal to N = 980 (standard deviation, s.d., 813) out of a theoretical maximum of approximately 29,000 (the number of days within the 5-year period multiplied by 16, the number of 12th-degree bins within the macro-bins). The number of match-ups is rather evenly distributed through the seasons (on average, N = 244, 260, 288 and 248, for boreal winter, spring, summer and fall, respectively). When the extended period (2002–2010) is considered, N = 1320 with the same spatial distribution. Patterns are as well similar for the other comparisons, SeaWiFS/MERIS (SeMe, N = 649, s.d. 441) and MERIS/MODIS (MeMo, N = 594, s.d. 458), albeit of a smaller amplitude due to the reduced coverage by MERIS. For the comparison performed with the original 24th-degree resolution of the MERIS and MODIS products, the average number of matching pairs within 3rd-degree macro-bins reaches 2743.

For each macro-bin, the statistics defined in Section 2 are computed if the number of match-ups is at least 50. This threshold is a compromise between covering at least 95% of the domain of analysis (for the three comparison pairs over 2003–2007) and preserving a sizeable match-up data set for statistical analysis. Statistics are presented in Figs. 2 to 8 (obtained with $\lambda = 1$ for Figs. 2, 4, 5 and 8, unless specified otherwise). Fig. 2 shows the global average of σ terms, δ, $|\psi|$ and ψ for the three comparison pairs.

4.1. Distribution of σ

First focusing on σ, an introduction to the analysis is required. As described in Section 2, σ is an intrinsic property of a data set as introduced by Eqs. (4) and (5), but in practice the method of collocation implies that it is mathematically dependent upon another data set (as seen from Eqs. (14) and (15)). So for any pair of data sets A and B, the analysis provides $\sigma_{A|B}$ (i.e., σ of A when given B) and $\sigma_{B|A}$ (i.e., σ of B when given A), linked by a factor λ. If $\lambda = 1$, the two estimates collapse into one value (see global averages on Fig. 2). The availability of three missions, and therefore of three comparison pairs, leads to two σ estimates for each mission, for instance $\sigma_{Se|Mo}$ and $\sigma_{Se|Me}$ for SeaWiFS, that can be seen as two realizations of the procedure put in place through Eqs. (4) to (15). The computation of their difference allows an evaluation of the uncertainty that can be associated with σ. This uncertainty will be further explored by varying λ away from 1, and computing the difference with respect to the reference case $\lambda = 1$, as well as the coefficient of variation (CV, standard deviation divided by the mean) associated with specific sets of simulations performed with different values of λ. Tests will be run with λ equal to 1.3 and 1.5 (and their inverse), values that appear conservative considering validation results (see Section 5.1). This sensitivity analysis on λ presented in the next sub-section is completed by tests on the match-up data set, the time period, and the spatial resolution. The presentation of the geographical variability associated with σ is presented in a following sub-section.

4.1.1. Sensitivity analysis of σ estimates

The three curves obtained with $\lambda = 1$ offer a coherent picture for σ, with a global average value decreasing with wavelength from 0.7–0.9 10^{-3} sr^{-1} at 412 nm to 0.05–0.1 10^{-3} sr^{-1} at the red band (Fig. 2). However, the three estimates obtained with $\lambda = 1$ (in that case $\sigma_{Se|Mo} = \sigma_{Mo|Se}$, $\sigma_{Se|Me} = \sigma_{Me|Se}$, and $\sigma_{Mo|Me} = \sigma_{Me|Mo}$) are not identical. The two estimates obtained for SeaWiFS from the pairs MoSe and SeMe differ by approximately ±30% from 412 to 555 nm, and 42% at 670 nm. On the other hand, $\sigma_{Mo|Se}$ and $\sigma_{Mo|Me}$ are very close to each other from 490 to 670 nm (from ±3% to 7%), but $\sigma_{Mo|Me}$ (equal to $\sigma_{Me|Mo}$) becomes higher at 443 nm (difference of 16%) and even higher at 412 nm (difference of 29%) where it becomes equal to $\sigma_{Me|Se}$ that is the highest value along the full spectral interval. The relative difference between the σ resulting from the comparison pairs SeMe and MeMo otherwise increases with wavelength, from 13% at 443 nm, 24% at 490 nm, approximately 35% at 510 and 555 nm, and 49% at 670 nm.

For each comparison pair, Fig. 2 also displays the cases $\lambda = 1.3$ and $\lambda = 1/1.3$, yielding six pairs of dotted curves (pairs of curves are often overlaid since, for sensors A and B, $\sigma_{A|B}(\lambda)$ tends to be close to $\sigma_{B|A}(1/\lambda)$). For each mission, relative differences with respect to the case $\lambda = 1$ are computed for the four cases (e.g., comparing $\sigma_{Se|Mo}(\lambda = 1.3)$ and $\sigma_{Se|Mo}(\lambda = 1/1.3)$ with respect to $\sigma_{Se|Mo}(\lambda = 1)$, and $\sigma_{Se|Me}(\lambda = 1.3)$ and $\sigma_{Se|Me}(\lambda = 1/1.3)$ with respect to $\sigma_{Se|Me}(\lambda = 1)$). For SeaWiFS, the difference is approximately 12% in modulus between 412 and 490 nm (from −15% for $\sigma_{Se|Me}(\lambda = 1/1.3)$ to +12.3% for $\sigma_{Se|Mo}(\lambda = 1.3)$), 18% at 510 and 555 nm, and +26% ($\lambda = 1.3$) and −21% ($\lambda = 1/1.3$) at 670 nm. Between 412 and 490 nm, results are similar for MERIS (slightly higher, 14%) and for MODIS. For MERIS, the difference for $\sigma_{Me|Mo}$ is approximately 17% at 510 and 555 nm, and 20% at 670 nm, but below 10% (in modulus) for $\sigma_{Me|Se}$ (between −10% and +6% at 510 and 555 nm, +2% and −3% at 670 nm). For MODIS, differences are found below 10% (in modulus) for all four comparisons between 510 and 670 nm (between −10% and +2%).

A further sensitivity analysis on λ is shown in Fig. 3 with six complete cases built assuming that one mission has σ lower or higher than σ associated with the other two missions that are assumed equal, which leads to five lines per graph ($\lambda \neq 1$ for two comparison pairs, $\lambda = 1$ for the third pair). This is conducted with fairly high and low values of λ (1.5 and 1/1.5). So, when applied to SeaWiFS, the cases are characterized by the assumption $\sigma_{Se|Me} = \sigma_{Se|Mo}$ with $\sigma_{Se|Me} = \lambda\sigma_{Me|Se}$, $\sigma_{Se|Mo} = \lambda\,\sigma_{Mo|Se}$ (implying $\sigma_{Mo|Me} = \sigma_{Me|Mo}$) with $\lambda = 1/1.5$ (Fig. 3a) and $\lambda = 1.5$ (Fig. 3b). This is repeated for MODIS and MERIS. All the curves appear in the same interval as those shown for $\lambda = 1$ in Fig. 2. When computing the same statistics as for the cases $\lambda = 1.3$ or 1/1.3

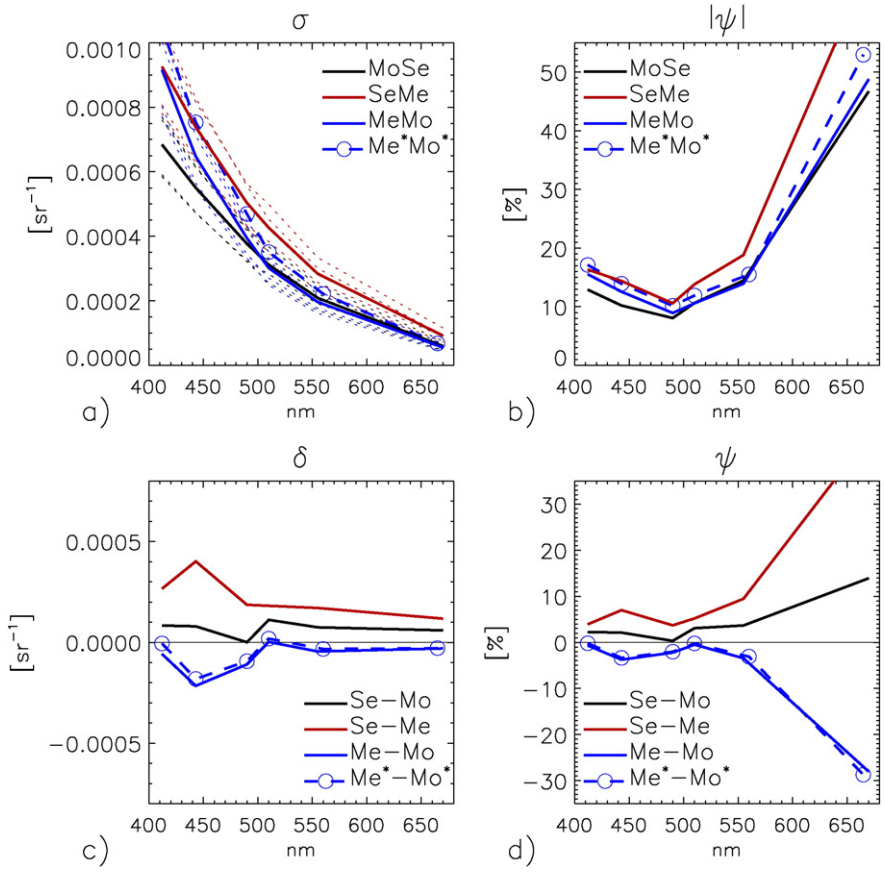

Fig. 2. Global average of comparison statistics for the three mission pairs, expressed as a) σ, b) $|\psi|$, c) δ and d) ψ. For each comparison pair in a), dotted curves are obtained with λ equal to 1.3 and 1/1.3. In the case of the comparison pair MeMo, the star '*' indicates results obtained with a 24th-degree spatial resolution.

(i.e., relative differences with respect to σ obtained with $\lambda = 1$), differences are logically found higher. Differences are mostly around 15–20% in absolute value (from −24.4% to −20.1% and from +13.4% to +19.9%) between 412 and 490 nm, around 26% for $\sigma_{Se|Mo}$, $\sigma_{Se|Me}$ and $\sigma_{Me|Mo}$, and below 17% for $\sigma_{Mo|Se}$, $\sigma_{Mo|Me}$ and $\sigma_{Me|Se}$ (from −17.3% to +9.8%) at 510 and 555 nm, 30–40% for $\sigma_{Se|Mo}$, $\sigma_{Se|Me}$ and $\sigma_{Me|Mo}$, and below 12% for $\sigma_{Mo|Se}$, $\sigma_{Mo|Me}$ and $\sigma_{Me|Se}$ (from −11.9% to +5.9%) at 670 nm.

For each sensor, statistics are computed for the set of six simulations associated with λ equal to 1/1.5, 1 and 1.5. For SeaWiFS, the coefficient of variation CV is approximately 23% between 412 and 490 nm, 29% at 510 and 555 nm, and 40% at 670 nm. It is fairly similar for MERIS, increasing from 20% at 412 nm to 25% at 555 nm, and 30% at 670 nm, while CV is lower for MODIS, 23% at 412 nm, 19% at 443 nm, 16% at 490 nm, 12% at 510 and 555 nm, and 8% at 670 nm.

Part of the differences between σ estimates can be due to the fact that the R_{RS} matching pairs differ from one comparison to the other. This was evaluated in the following way: for a subset of bins, the calculations applied to the pair MoSe (with $\lambda = 1$) were repeated 100 times with a randomly selected half of the comparison data set. The coefficient of variation among the results of this boot-strapping exercise is of the order of 10–15% between 412 and 555 nm (decreasing with wavelength), and 5% at 670 nm. Another test is conducted to assess the impact of the overlap period over the results: by performing the calculations over an extended period (2002–2010 for MoSe and SeMe, 2003–2011 for MeMo), the results are almost identical to those obtained with the reference period 2003–2007, with differences usually not exceeding 2% at all bands, i.e., 0.01–0.02 10^{-3} sr^{-1} in the blue (maximum of 4.5%, or 0.03 10^{-3} sr^{-1}, at 412 nm for MoSe).

Finally an analysis was made to test the influence of the spatial resolution on the results. From the analysis performed with the 24th-degree resolution between MERIS and MODIS (with $\lambda = 1$), the average σ is increased by approximately 16% at all wavelengths (with increments decreasing from 0.14 10^{-3} to 0.01 10^{-3} sr^{-1} from 412 to 670 nm) with respect to the 12th-degree case (see Fig. 2a). Conducting the analysis with a different resolution might lead to conflicting effects: on one hand, σ could decrease with a higher resolution since the match between the two products is more accurate spatially and there are more matching pairs; on the other hand, averaging the MERIS and MODIS products to a lower resolution would tend to reduce the noise associated with the data, and thus σ. Results suggest that the latter effect is slightly dominating.

Making the synthesis of these various analyses, it appears that the uncertainties associated with σ can be evaluated at 20–30% between 412 and 555 nm, and 30–40% at 670 nm. From the result of the boot-strapping exercise, approximately a third to half of the uncertainty (from 412 to 555 nm) could be due to the selection of the R_{RS} matching data for each comparison pair. These uncertainties on σ may not be equally distributed around the value obtained with $\lambda = 1$. Indeed, Fig. 2a indicates that σ_{Me} associated with MERIS might actually be higher than that resulting from the case $\lambda = 1$, since $\sigma_{Me|Se}$ is higher than the other estimates across the spectrum while $\sigma_{Me|Mo}$ is also higher in the blue. Conversely, σ_{Mo} might be lower since σ tends to be lower when MODIS is part of the comparison pair (Fig. 2a). This is confirmed when σ_{Me} is assumed to be 1.5 times σ_{Se} and σ_{Mo} (Fig. 3f) or when σ_{Mo} is assumed to be 1.5 times lower than σ_{Se} and σ_{Me} (Fig. 3c): in these cases, the two σ estimates associated with each mission (associated with a specific symbol in Fig. 3) are closer to each other than when $\lambda = 1$ for the three comparison pairs. In spite of the potential afforded

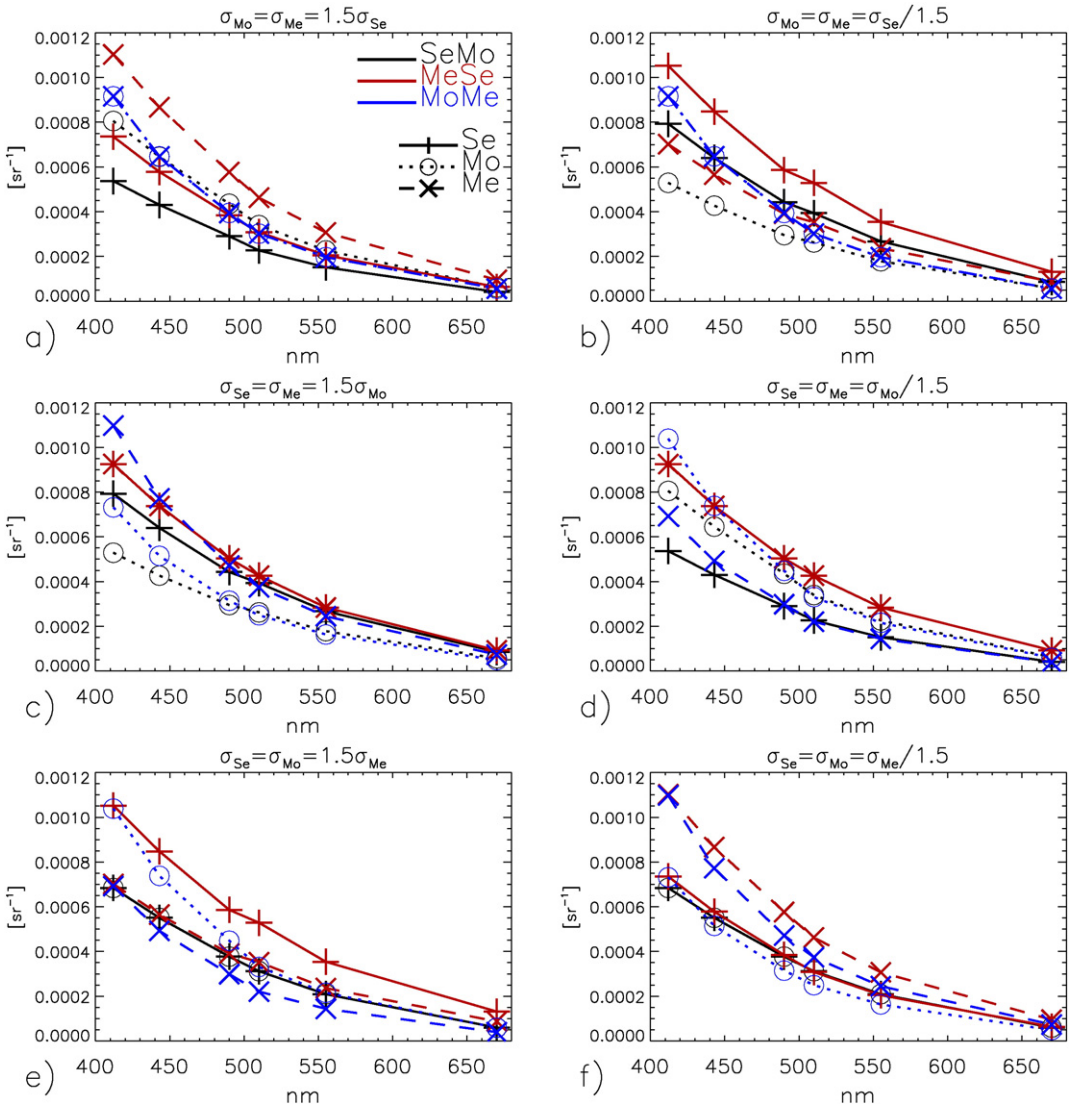

Fig. 3. Global average σ for the three mission pairs. Each graph is associated with a specific case built assuming that one mission has σ lower (a,c,e) or higher (b,d,f) than the value of σ associated with the other two missions by a factor $\lambda = 1.5$. In each graph, the color refers to the comparison pair, whereas the symbol refers to the mission. For instance, black (blue) dotted lines with open circles represent σ estimates obtained for MODIS from the comparison with SeaWiFS (MERIS) data.

by collocation statistics, the assumption of a fixed λ is admittedly a limitation of the current approach, and more work is needed to better characterize the relative values of the uncertainties associated with each mission and how they vary with wavelength, location and season. More sensitivity tests on the value of λ could help in determining the most probable values for the σ spectrum associated with each mission.

4.1.2. Spatial and temporal variability of σ

The spatial distribution of σ is illustrated in Fig. 4 (for MoSe, λ 1) at two selected wavelengths (443 and 555 nm) that are representative of all the bands in terms of relative spatial distribution. Similar spatial patterns are obtained for SeMe and MeMo but with slightly different amplitudes (not shown). Overall, the distribution of σ does not show a strong spatial variability, with most values in the interval 0.5–0.8 10^{-3} sr^{-1} at 443 nm, but some patterns can be noticed. Some regions of high σ are associated with a relatively small number of match-ups (Fig. 1) but this is by no means a general feature, as seen in the Arabian Sea or some coastal regions (e.g., in the Black Sea or the Gulf of Mexico). Relatively high σ are found in the northern mid-latitudes, along the equatorial band (or slightly off the Equator in the eastern Pacific), and the Southern Ocean, a region of high wind speeds generating white caps (Salisbury, Anguelova, & Brooks, 2014). Conversely, the subtropical gyres or some high-latitude regions like the Hudson Bay or the Baltic Sea show relatively low values of σ.

Regional averages of σ are illustrated in Fig. 5 (still for the MoSe pair) and confirm the fairly small spatial variability of σ: regional averages display a similar amplitude with the same spectral shape, and most are within one standard deviation of the global average, with the standard deviation representing 15% of the average value at 412 and 443 nm. Considering the 56 Longhurst provinces, only 5 (for 443 and 490 nm) to 10 (for 670 nm) are characterized by a regional σ beyond one standard deviation of the global average, with the Baltic Sea, the Red Sea, the China Sea, the Indian coastal waters, and the Guianas coastal province (Longhurst, 2006) being the most affected. The ratio between maximum and minimum province averages of σ is between 1.7 and 2 from 412 to 510 nm, and is 3.2 and 6.8 at 555 and 670 nm, respectively. This increase in relative excursion of σ for the green and red wavelengths is explained by σ values being closer to 0. Among the represented regional estimates, only σ for the Southern Ocean and the Northern Indian Ocean exceeds 0.75 10^{-3} sr^{-1}, whereas the Baltic Sea's σ is distinctly small from 412 to 490 nm (0.53 10^{-3} sr^{-1} at 412 nm). Beyond 490 nm, the lowest values are found in regions with

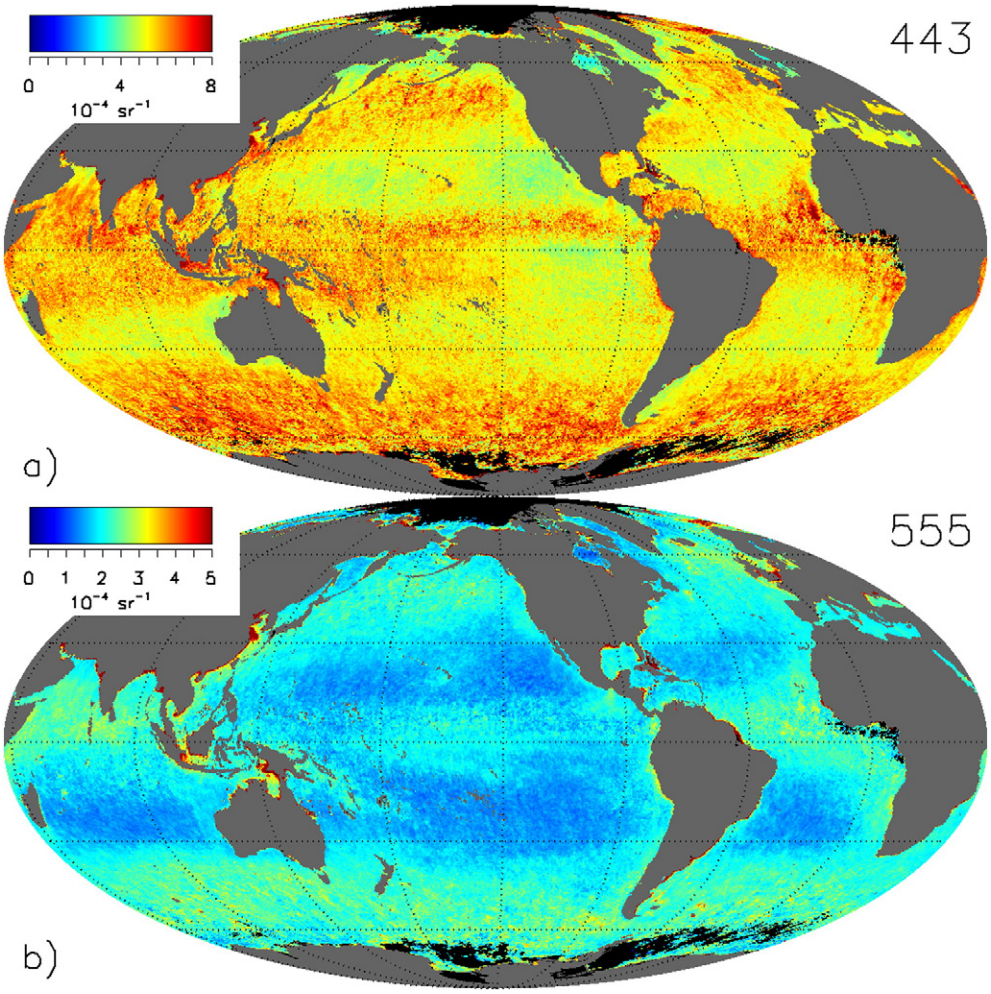

Fig. 4. Distribution of σ for the comparison SeaWiFS/MODIS at a) 443 nm, and b) 555 nm.

the lowest Chl*a* (Fig. 5a), particularly in the subtropical gyres as confirmed by Fig. 4. Seasonal variations of σ also appear very small: in terms of global average, all four seasonal averages are in a envelope not exceeding $\pm 5\%$ (with a maximum excursion between the four seasons of 0.05 10^{-3} sr^{-1} at 412 nm), a result that mostly holds for all regions (not shown).

4.2. Comparison statistics

For completeness, comparison statistics are presented. Before describing the results for the comparison metrics introduced in Section 2, the values obtained for β, which is the slope of linear (model II) regression, can be mentioned. From 412 to 490 nm, β is found close to 1, with global median values from 0.90 (regression of SeaWiFS data with respect to MERIS at 412 nm) to 1.11 (regression of MERIS data with respect to MODIS at 490 nm). Departures from 1 are more pronounced at higher wavelengths, for instance global median values of 1.48 and 3.0 at 555 and 670 nm, respectively, associated with the regression of MERIS data with respect to MODIS. These median values vary greatly by region depending on the dynamic range of R_{RS}, from 0.95 to 3.9 at 555 nm, and from 0.92 to 7.6 at 670 nm, if the Longhurst provinces are analyzed. At global scale, these statistics are dominated by the oligotrophic waters that cover a large part of the global ocean, and for which the natural variability of R_{RS} at the green and red wavelengths is small, potentially leading to values of β that can be noticeably different from 1.

The mean absolute relative difference $|\psi|$ is fairly similar for the three comparison pairs (Fig. 2b), with $|\psi|$ between 12.9% and 16.3% at 412 nm. These statistics are virtually identical for the comparisons MeMo conducted at the 12th-degree and 24th-degree resolutions, a result that holds for ψ and δ (see Fig. 2). $|\psi|$ shows larger spatial variations than σ (Fig. 5b), particularly in the blue and red, and in terms of spectral shape. Considering the 56 Longhurst provinces, more than 20 (up to 24) are associated with a regional $|\psi|$ beyond one standard deviation of the global average at 412, 443, 490 and 670 nm, while this is the case for 9 and 12 provinces at 510 and 555 nm, respectively. The ratio between maximum and minimum province averages of $|\psi|$ varies from 2 (at 510 nm) to 11 (at 412 nm). In general, $|\psi|$ spectra are characterized by a U shape, with a minimum in the interval 510–555 nm, in the range 10–15%. Two end-members can be distinguished: *i*) oligotrophic waters, as represented in Fig. 5b by the Mediterranean Sea, tropical provinces and subtropical gyres (with the lowest median Chl*a* of 0.07 mg m^{-3}, Fig. 5), with $|\psi|$ around 10% between 412 and 490 nm (with the minimum for $|\psi|$ of approximately 6.5% from 412 to 490 nm in the subtropical gyres), and high values at 555 and 670 nm; *ii*) coastal regions or marginal seas characterized by turbid waters, as exemplified by the Black and Baltic Seas (with the highest Chl*a*), with high values of $|\psi|$ in the blue (exceeding 30% at 412 nm) and relatively low values at 670 nm. Between these two end-members, there can be some variations, like for the northern Indian Ocean showing high $|\psi|$ across the spectrum. Overall, $|\psi|$ tends to be high when R_{RS} is low and vice-versa: in oligotrophic waters, R_{RS} is maximum at 412 nm and decreases

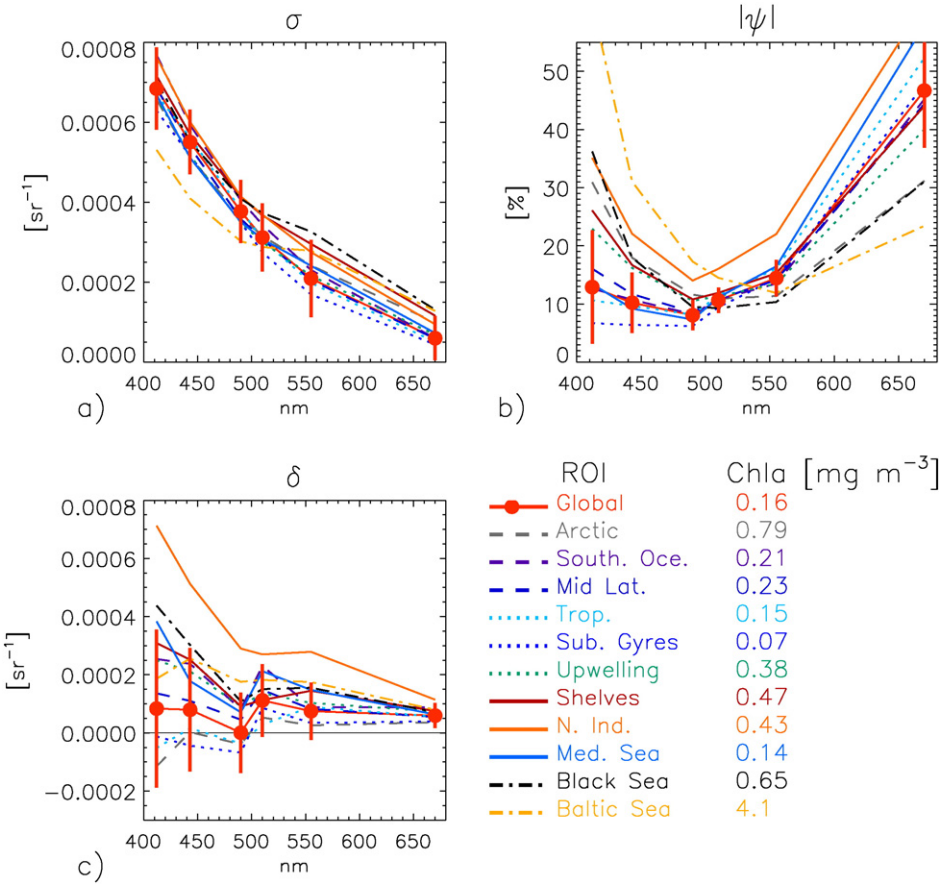

Fig. 5. Regional averages of comparison statistics for the pair SeaWiFS/MODIS, expressed as a) σ, b) $|\psi|$, and c) δ. Averages are for the global ocean (with vertical bars, ± 1 standard deviation), Arctic Ocean, Southern Ocean, Mid-latitude provinces, Tropical provinces, subtropical gyres, upwelling provinces, ocean shelves, northern Indian Ocean, Mediterranean, Black and Baltic Seas. Multi-annual median Chl*a* is indicated for each ensemble.

to very low levels at 670 nm, whereas turbid waters often show low values at 412 nm, particularly in absorbing waters like those found in the Baltic Sea (Berthon & Zibordi, 2010), and relatively higher values at green and red bands (Mélin & Vantrepotte, 2015).

The mean relative difference (or relative bias) ψ shows different results for the three comparisons (Fig. 2d). For the pair MoSe, the globally-averaged ψ is below +3.7% between 412 and 555 nm (+13.9% at 670 nm). In the interval 412–555 nm, it is comparable for the pair MeMo with an opposite sign (i.e., MERIS R_{RS} being lower than MODIS R_{RS}) with ψ between −0.7% and −3.8%, but it is much higher (in modulus) at 670 nm (−28%). The largest values (in modulus) are found for the pair SeMe, with ψ ranging from +3.7% to +9.5% below 555 nm, and reaching +45% at 670 nm (i.e., SeaWiFS R_{RS} higher than MERIS R_{RS}).

A picture consistent with ψ is given by δ (bias, Fig. 2c) even though the spectral shape is different. Global statistics are much affected by oligotrophic to mesotrophic waters, for which R_{RS} tends to be relatively high at blue wavelengths and low in the red. The average δ at 670 nm is of the order of 10^{-4} sr^{-1} or below, which can be translated into large relative differences at this wavelength (Fig. 2d), while δ of 0.2–0.3 10^{-3} sr^{-1} is associated with fairly small ψ at blue wavelengths. The maximum of δ is found at 443 nm, +0.4 10^{-3} sr^{-1} for the pair SeMe.

Figs. 6 and 7 illustrate the spatial variability of δ at 443 and 555 nm for the pairs MoSe and SeMe, respectively (comparable patterns can be seen for ψ), the former being complemented by the regional δ averages of Fig. 5c. At 443 nm (Fig. 6a), SeaWiFS R_{RS} appear mostly larger than MODIS R_{RS}, particularly in the Southern Ocean, the tropical Atlantic or the northern Indian Ocean. The subtropical gyres and the Arctic basin are the only regions where δ is negative, actually between 412 and 490 nm (Fig. 5c). At 555 nm, most of the ocean has δ lower than 0.2 10^{-3} sr^{-1} (in modulus) (Fig. 6b), with again the tropical Atlantic or the northern Indian Ocean characterized by higher values. Consistently with Fig. 2c, δ for the pair SeMe is mostly positive (i.e., SeaWiFS R_{RS} higher) in the global ocean at all bands (see Fig. 7 for 443 and 555 nm), with few isolated features of negative bias. Values are particularly high in equatorial regions and the Southern Ocean. For the pair MeMo, the distribution of δ can be deduced from Figs. 6 and 7; it shows mostly negative values (i.e., MERIS R_{RS} lower), except at 412 and 510 nm, bands for which regions with positive δ can be noticed and whose global average is actually close to 0 (Fig. 2c).

Differently from σ, bias statistics show more seasonal variations at least for some regions. Fig. 8 shows examples of such variations for δ of the pair MoSe averaged over selected regions. Estimates for a specific season are constructed by accumulating matching pairs for days in December–February for boreal winter, March–May for spring, June–August for summer and September–November for fall. A clear feature is observed for the subtropical gyres, here exemplified by the North Atlantic Tropical province (NATR) and the South Pacific Subtropical Gyre (SPSG) (see Longhurst, 2006, for the exact definition of these provinces): the bias tends to be relatively high, i.e., positive and higher than the multi-annual average, in the regional winter (boreal winter for NATR, boreal summer for SPSG), and low (negative) in the regional summer. This result actually applies to most of the subtropical belt. Seasonal variations are also found in smaller regions. For instance, in the Black Sea (Fig. 8c), δ appears relatively high in summer (SeaWiFS R_{RS} higher than MODIS R_{RS}), and low in winter (closer to null). In the Arabian Sea (Fig. 8d), there are not enough data for a robust estimate of δ in summer because of the scarce coverage by standard satellite products

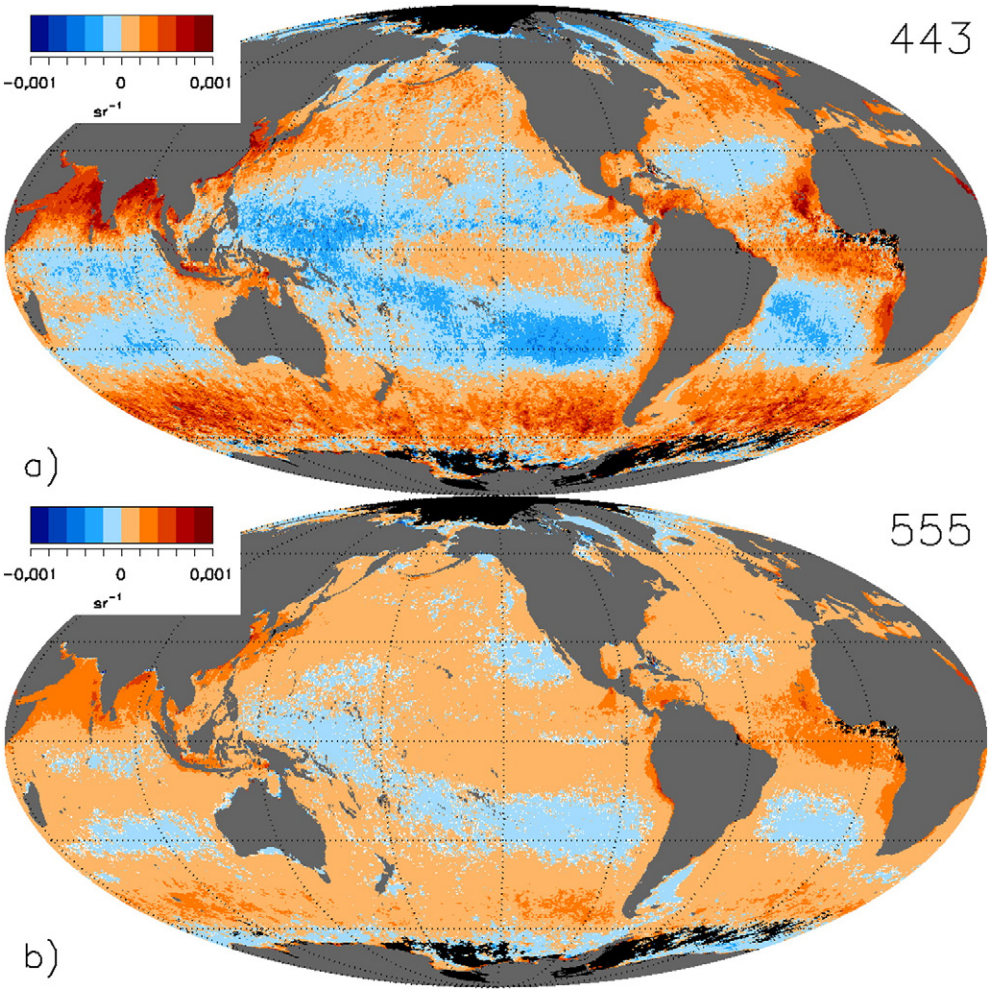

Fig. 6. Distribution of δ for the comparison SeaWiFS/MODIS at a) 443 nm, and b) 555 nm. Positive values mean that SeaWiFS R_{RS} is higher than MODIS R_{RS}.

during the summer monsoon. In winter, δ is higher than the overall average whereas it is lower (but still positive) in spring. Results obtained for δ are consistent with those displayed by ψ (not shown). Various factors can explain why comparison statistics might evolve with the season, as the performance of the atmospheric correction is influenced by the geometry of illumination and by the annual cycle of the atmospheric constituents and of the water content. Similar variations have been noticed in previous studies comparing optical properties associated with different missions (e.g., Mélin, 2011; Mélin et al., 2009) or in terms of validation statistics (Mélin, Zibordi, & Berthon, 2007; Zibordi, Mélin, & Berthon, 2012); the fact that they can occur between missions processed with the same atmospheric correction scheme suggests different sensitivities to these factors from one mission to the other.

5. Comparison with other uncertainty estimates

The main focus of this study is the term referred to as σ, the part of the R_{RS} uncertainty budget associated with random effects. Satellite products should be accompanied by uncertainty estimates but in turn the quality of these estimates needs to be assessed. The first way to conduct such an assessment is through a comparison with validation statistics, i.e., resulting from the comparison of satellite products with field data (Section 5.1). As anticipated in Section 3.2, σ obtained by collocation of satellite products can be compared with a σ^v term derived through the same mathematical framework. The un-biased RMS difference Δ_u will also be considered for comparison with other estimates.

Moreover, the σ term can be compared with uncertainty estimates derived by other approaches. Here, two other techniques are considered. The first one was developed in the framework of the European Space Agency (ESA) Ocean Colour Climate Change Initiative (CCI, Sathyendranath et al., 2016) that is including uncertainty estimates for R_{RS} in its data stream. Matching field data from a large data base of optical measurements with daily Level-3 R_{RS} products, validation statistics were derived and analyzed across a suite of optical water types (or classes) built from the satellite R_{RS} products. A similar ensemble of field data compared with Level-2 satellite imagery was used to define validation statistics per optical water type by Moore, Campbell, & Feng (2015), where the types were constructed from field observations (Moore, Campbell, & Dowell, 2009). Other investigations have also expressed validation results as a function of optical classes (Goyens, Jamet, & Schroeder, 2013; Mélin & Franz, 2014). CCI uncertainty estimates are obtained at global scale by computing for each daily bin the weighted average of validation statistics associated with the relevant optical classes, where the weights are the class memberships determined for each class in a way similar to the approach proposed by Moore et al. (2009) for Chl*a*. For instance, if R_{RS} appears to fully belong to a specific class (membership of 1), it is assumed that its uncertainty is well represented by the related validation results obtained with the match-ups belonging to that class. Considering that the CCI uncertainty estimates are traceable to validation statistics, σ is here compared with the unbiased RMS difference Δ_u derived from CCI data and computed according to Eq. (17) (for clarity the notation Δ_u^{CCI} is introduced). Uncertainty estimates

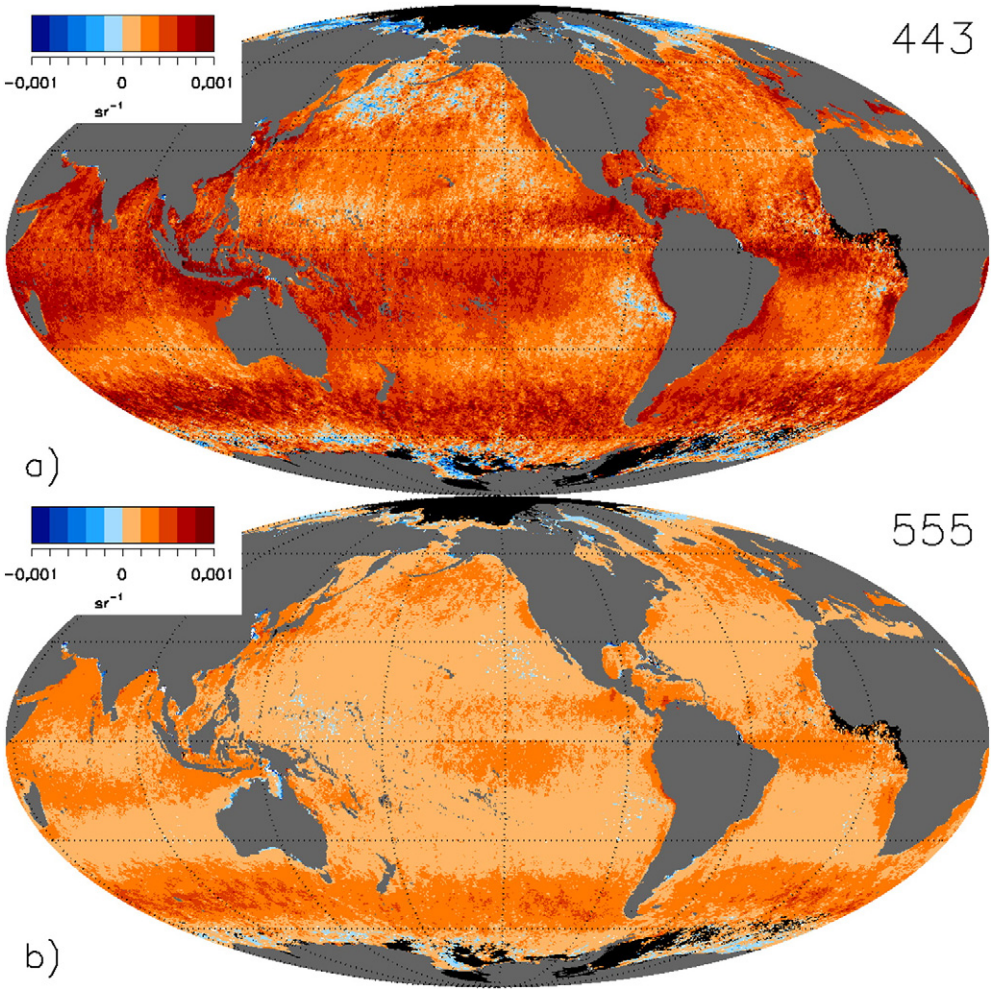

Fig. 7. Same as Fig. 6 for the comparison SeaWiFS/MERIS. Positive values mean that SeaWiFS R_{RS} is higher than MERIS R_{RS}.

are delivered for the SeaWiFS bands on a 24th-degree grid as multi-annual averages.

Another approach has been proposed to derive uncertainty estimates for SeaWiFS and MODIS R_{RS} (Hu, Feng, & Lee, 2013). It relies on Chl*a* bio-optical algorithms with results applicable to oligotrophic waters (Chl*a* lower than 0.25 mg m^{-3}). The concept is to compare the outputs of two Chl*a* algorithms, the OC4v6 blue–green band-ratio formulation (O'Reilly et al., 2000) (working with 443 nm in oligotrophic waters) and a color index (CI) three-band subtraction method (Hu, Lee, & Franz, 2012). When both outputs are in close agreement, the associated R_{RS} is assumed to be affected by a low uncertainty (and for practical purposes error-free), while this uncertainty increases as the algorithm outputs diverge. Practically, for a given Chl*a* the reference R_{RS} is taken as the average of the R_{RS} spectra producing two Chl*a* estimates differing by less than 5%. Uncertainty estimates, provided as a function of Chl*a*, are derived as the standard deviation Ω of the R_{RS} deviations from the reference R_{RS} in areas of the North Atlantic and South Pacific subtropical gyres (for more details about this approach, see Hu et al., 2013).

These comparisons are made with respect to σ estimates obtained with λ equal to 1, keeping in mind the uncertainties on σ discussed in Section 4.1.1.

5.1. Comparison at validation sites

This section compares uncertainty estimates at specific validation sites. The comparison of σ with validation statistics is possible for sites with a significant number of match-ups such as those introduced in Section 3.2. In practice, σ as well as the unbiased RMS difference Δ_u obtained from the validation analysis (Section 3.2) for the three missions are compared with σ found for the macro-bin associated with the site location for the three comparison pairs, and with Δ_u^{CCI} obtained from the CCI project (Fig. 9). Here each site is associated with three estimates σ^c obtained from the comparisons MoSe, MeMo, and SeMe (noted σ^c_{MoSe}, σ^c_{MeMo}, and σ^c_{SeMe}, respectively) with λ taken equal to 1, and three estimates σ^v from the validation statistics of SeaWiFS, MODIS, and MERIS (noted σ^v_{Se}, σ^v_{Mo}, and σ^v_{Me}, respectively). For clarity, it is stressed that the σ^v values are computed at the center wavelengths associated with each mission (Table 1). The number of match-ups between sites and missions is highly variable, 545 at MOBY and 461 at AAOT but as low as approximately 30 at the Baltic sites for SeaWiFS, between 148 (GDLT) and 553 (AAOT) for MODIS, and between 23 (HLT) and 163 (MOBY) for MERIS. Moreover, Δ_u^{CCI} values are extracted for the site location from multi-annual averaged maps.

At the MOBY site (Fig. 9a), the various σ estimates appear comparable, with σ^c obtained from comparison usually lower than σ^v obtained from validation. For instance the ratio of the three σ^c with respect to σ^v_{Se} varies between 0.76 and 1.05 for wavelengths between 412 and 555 nm. This ratio is between 0.55 and 0.80 at 670 nm, but it varies from 0.93 to 1.36 if σ^v_{Mo} is taken as the reference (σ^v_{Se} being noticeably higher than σ^v_{Mo} at that band). The CCI value Δ_u^{CCI} is fairly comparable with σ estimates for wavelengths above 500 nm but appears noticeably higher in the blue. Differences are however reduced if Δ_u^{CCI} is compared to Δ_u^v obtained from validation: for instance the ratio between Δ_u^{CCI} and $\Delta^v_{u,Se}$ varies between 0.99 (at 555 nm) and 1.28 (at 443 nm). Indeed,

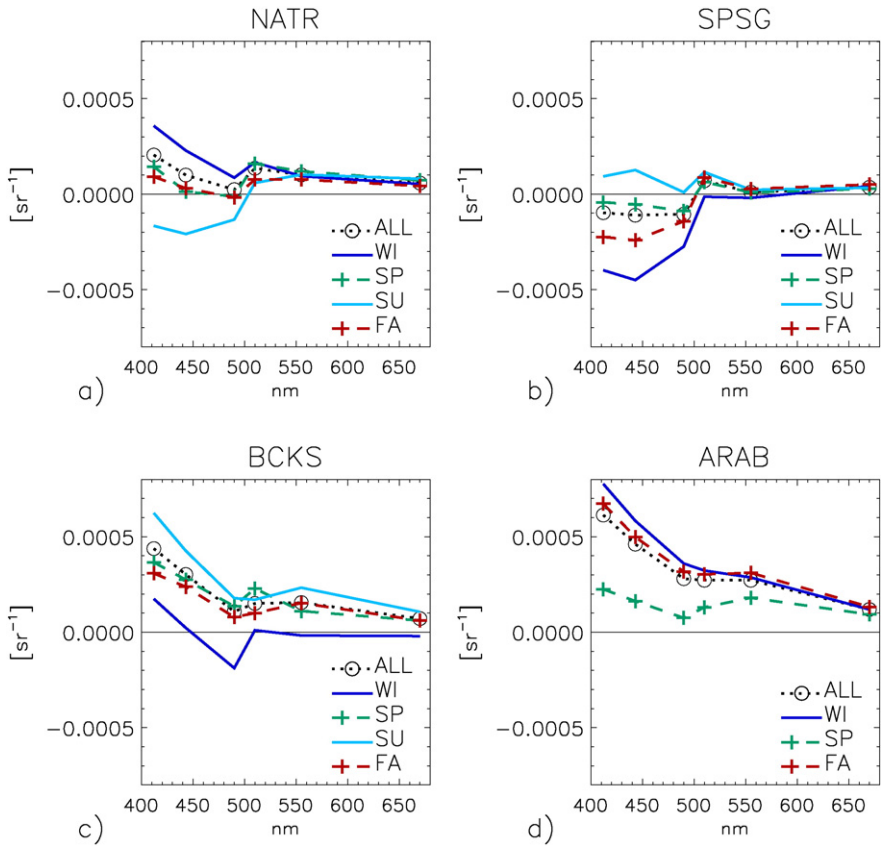

Fig. 8. Regional average of δ for the comparison SeaWiFS/MODIS over the a) North Atlantic Tropical province, b) South Pacific Subtropical Gyre, c) Black Sea, and d) Arabian Sea. The multi-annual average (2003–2007) is shown by a dotted line with circles, while seasonal values are given for boreal winter (WI), spring (SP), summer (SU) and fall (FA). Positive values mean that SeaWiFS R_{RS} is higher than MODIS R_{RS}.

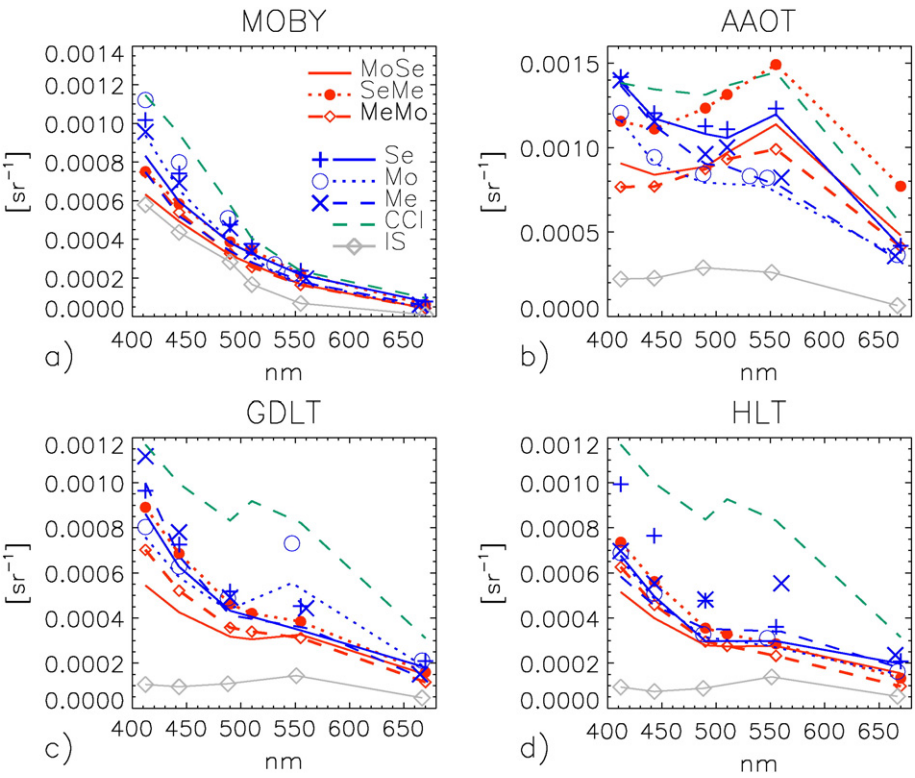

Fig. 9. For 4 validation sites, σ estimates obtained for the three comparison pairs (red lines), and from validation statistics for each mission (blue lines). Separate symbols, '+', 'o' and 'x', provide the same validation statistics in terms of Δ_u for SeaWiFS, MODIS and MERIS, respectively. Δ_u^{CCI} values are shown with dashed green lines. Statistics are given for the sites a) MOBY, b) AAOT, c) GDLT and d) HLT. The gray curves with diamonds are uncertainty estimates for the corresponding field measurements (see text). In a), the curve SeMe is mostly overlaid by the line Se, while the line MeMo is mostly overlaid by the line MoSe.

from validation results, σ^v is systematically lower than Δ_u^v, by approximately 20–25% between 412 and 443 nm, 10–15% between 510 and 560 nm, and 2% in the red. Fig. 9a also displays an estimate of the uncertainty associated with the field data, by multiplying the relative uncertainty documented for the site (12.5% in the red, 5% otherwise; Brown et al., 2007) by the average R_{RS} computed over the match-up data. The ratio between this uncertainty and σ^c_{MoSe} (i.e., the lowest σ estimate) varies from approximately 0.9 between 412 and 490 nm, and is equal to 0.61, 0.41 and 0.24 at 510, 555 and 670 nm, respectively. This shows that a 5% uncertainty is approximately reached only for wavelengths below 500 nm at this site.

Results found at the northern Adriatic coastal site AAOT are also comparable for the σ estimates obtained through comparison or validation, with as much variation between the three σ^cs as between the three σ^vs (Fig. 9b). The highest values are associated with the validation results σ^v at 412 nm while the maximum in the green is found for σ^c_{SeMe}. Except σ^v_{Mo} and σ^v_{Me}, the σ spectra as well as Δ_u^{CCI} show a local maximum at the green band. For that site, the values of σ^v and Δ_u^v obtained from validation are actually close, differing by usually less than 5%. Differently than for MOBY, the uncertainty associated with the field data (equivalent to 5%, except for the red, 7%; Gergely & Zibordi, 2014) appears much lower (by a factor at least 3) than the σ estimates.

For the Baltic sites (Fig. 9c and d) it is worth remembering that the number of match-ups obtained for SeaWiFS (29 and 31 at GDLT and HLT, respectively) and MERIS (43 and 23 at GDLT and HLT, respectively) is fairly low. With the exception of a relatively high σ^v_{Mo} at 547 nm, the various σ estimates are well consistent at GDLT (Fig. 9c). For instance, the ratio between σ^c_{MoSe} and σ^v_{Se} varies from approximately 0.7 between 412 and 490 nm to 0.94 at 555 nm and 0.83 at 670 nm. At HLT (Fig. 9d), the agreement appears still better, with the ratio between σ^c_{MoSe} and σ^v_{Mo} being around 0.8–0.9 from 412 to 490 nm, around 1 in the green, and 1.03 at 670 nm. For both Baltic sites, Δ_u^{CCI} is noticeably higher than the other estimates even considering the Δ_u^v obtained from validation. The difference between the values of σ^v and Δ_u^v obtained from validation varies widely, with the former being 4% to 39% lower. Being derived from a comparison with field data (including those from the four sites considered here), CCI estimates reflect validation statistics obtained for specific optical water types, so that to find a coherence between Δ_u^{CCI} and Δ_u^v is not surprising. By the same token, validation results obtained with other field data sets may contribute large Δ_u values for the optical water types most frequently encountered at GDLT and HLT, explaining why fairly high Δ_u^{CCI} are found at the Baltic sites. As for AAOT, the uncertainty associated with the Baltic field data (equivalent to at least 5%) are lower than σ by a factor 2 or 3 at the green and red bands, or a factor 5 in the blue.

The availability of validation statistics for the three missions allows the comparison of their respective σ^v. From 412 to 490 nm (neglecting the impact associated with the difference between 488 and 490 nm), the ratio between σ^v of the three missions is mostly in the interval 0.8–1.2, with a ranking that is not systematic: for instance MODIS σ^v_{Mo} is higher than for SeaWiFS and MERIS at the MOBY site, but it is lower at AAOT. There are a few exceptions, σ^v_{Se} being higher than σ^v_{Mo} by a factor 1.3–1.35 at 443 and 490 nm at AAOT, and σ^v_{Me} being higher than σ^v_{Mo} by a factor 1.3 at 412 nm at GDLT. At 510 nm, σ^v_{Me} is lower than σ^v_{Se} by a factor 0.85 at MOBY and AAOT (that band is not available for the Baltic sites). Regardless of differences in wavelengths, σ^v at the green band are generally very close, but σ^v_{Se} is significantly higher than for MODIS and MERIS by a factor of approximately 1.5 at AAOT (Fig. 9b), while σ^v_{Mo} is higher at GDLT (Fig. 9c). Neglecting the difference in wavelength in the red (665 to 670 nm), the ratio between σ^vs is in the interval 0.75–1.3, except σ^v_{Se} that is higher than σ^v_{Mo} by a factor 1.68 at MOBY. While keeping in mind that these

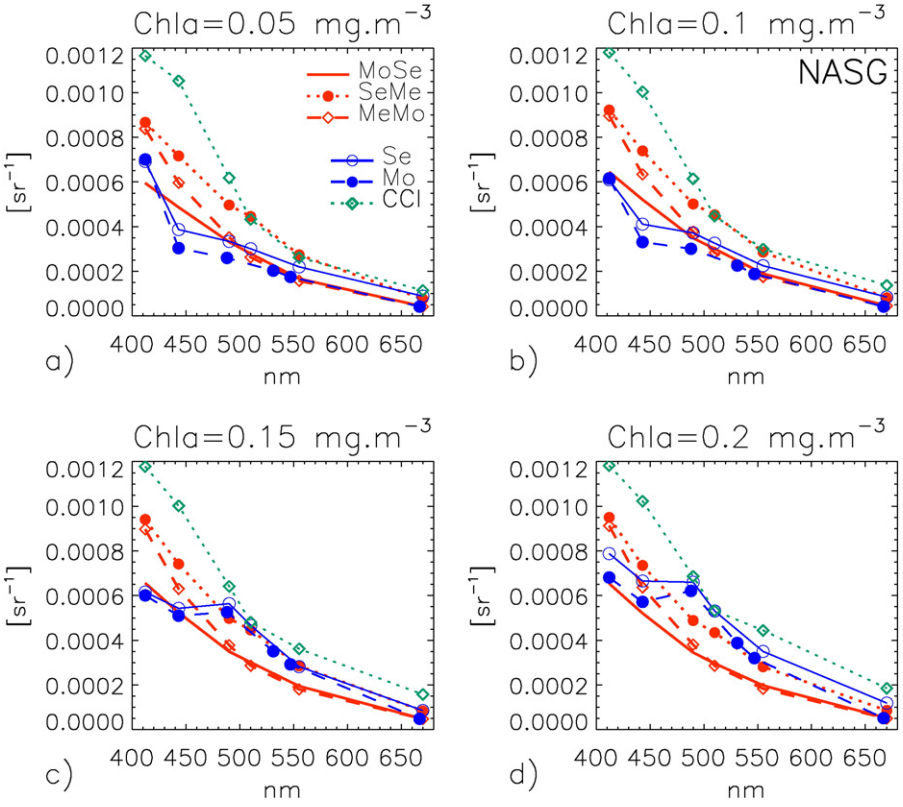

Fig. 10. σ spectra obtained from the three comparison pairs over the North Atlantic subtropical gyre (red lines with three different line styles). Associated uncertainty estimates Ω from Hu et al. (2013) are provided for MODIS (dashed blue lines with filled circles) and SeaWiFS (blue lines with circles). Δ_u^{CCI} are shown with green dotted lines with diamonds. Statistics are given for Chl*a* equal to a) 0.05, b) 0.1, c) 0.15 and d) 0.2 mg m^{-3}. For wavelengths above 500 nm, the line MeMo is mostly overlaid by the line MoSe.

validation statistics were computed from different sets of match-ups, they shed some light on the values that can be associated with λ and show that the sensitivity analysis performed on λ (Section 4.1.1) covers a realistic interval.

5.2. Comparison with estimates over oligotrophic waters

To allow a comparison with the uncertainty estimated by the approach of Hu et al. (2013), an analysis is conducted over the North Atlantic and South Pacific subtropical gyres (NASG and SPSG, respectively). For comparison with the Ω estimates (noted Ω_{Se} and Ω_{Mo} for SeaWiFS and MODIS, respectively), σ is averaged for the two regions over several Chl*a* intervals around 0.05 (0.025–0.075), 0.1 (0.075–0.125), 0.15 (0.125–0.175) and 0.2 (0.175–0.225) mg m $^{-3}$. The Chl*a* values are taken from a SeaWiFS annual average. A similar analysis is conducted with the CCI estimates Δ_u^{CCI}. The resulting spectra are compared in Figs. 10 and 11.

For Chl*a* of 0.05 and 0.1 mg m $^{-3}$, the largest differences between the various estimates are found for blue wavelengths (Figs. 10a, b, 11a and b), with Δ_u^{CCI} being highest. For both regions, the Ω values are actually close to σ derived from the MoSe comparison (i.e., the lowest σ) with the ratio Ω to σ_{MoSe} mostly in the interval 0.7–1.5, except in the red with $\Omega_{Se}/\sigma_{MoSe}$ equal to 2.4 in SPSG for Chl*a* of 0.1 mg m $^{-3}$, and $\Omega_{Se}/\sigma_{MoSe}$ around 2 in NASG. In SPSG at 0.05 mg m $^{-3}$, Ω is lower and higher than σ_{MoSe} at 443 and 490 nm, respectively. Actually, Ω is often characterized by a local maximum at 490 nm that is not seen in the σ or Δ_u^{CCI} spectra.

For 0.15 and 0.2 mg m $^{-3}$ (Figs. 10c, d, 11c and d), Ω values are higher (as increasing with Chl*a*; see also Fig. 7 in Hu et al., 2013) and closer to σ derived from the SeMe comparison, at least between 490 and 670 nm. At 412 and 443 nm, and for Chl*a* of 0.15 mg m $^{-3}$, Ω is close to σ from the MoSe comparison in NASG (Fig. 10c), and close to σ from the MeMo and SeMe comparisons in SPSG (Fig. 11c). For Chl*a* of 0.2 mg m $^{-3}$ in NASG, Ω at these wavelengths is between σ found for the MoSe and MeMo comparisons (Fig. 10d). For Chl*a* of 0.15 mg m $^{-3}$, Δ_u^{CCI} is close to Ω and σ_{SeMe} between 490 and 555 nm, but is the highest estimate at 412, 443 and 670 nm. It is also the highest across the spectrum for Chl*a* of 0.2 mg m $^{-3}$, approximately a factor of 2 with respect to σ_{MoSe}. This difference can not be explained by the relationship between validation statistics expressed as Δ_u and σ (Eq. (18)): for oligotrophic waters, the results obtained at MOBY suggest a difference not exceeding 25% (Section 5.1). Computed for similar water types (oligotrophic waters and fixed Chl*a* intervals), Δ_u^{CCI} estimates associated with the two regions NASG and SPSG are logically very similar.

6. Discussion

This section discusses the results obtained for σ. A first point of discussion is associated with the method itself, specifically in its definition of the value assumed for λ. The results based on $\lambda = 1$ show that the σ estimates obtained from the three comparison pairs are not identical, albeit comparable (Fig. 2a), which indicates that this assumption is not necessarily valid everywhere and for each band. Finding a fixed λ value that would reconcile the σ estimates appears challenging: the global average of σ associated with the MeMo pair is close to the estimate from the MoSe pair from 490 to 670 nm but close to the estimate from the SeMe comparison at 412 nm, so that a common λ value for all bands may not always apply, without mentioning possible spatial and temporal variations (where these can partly be due to the variations in the availability of matching pairs). Validation results (Antoine et al., 2008; Mélin et al., 2011b, 2012; Mélin & Franz, 2014; Zibordi et al., 2009a) suggest a certain range of possible λ values, typically found in the interval from 0.8 to 1.2 (Section 5.1). Recently, using a large

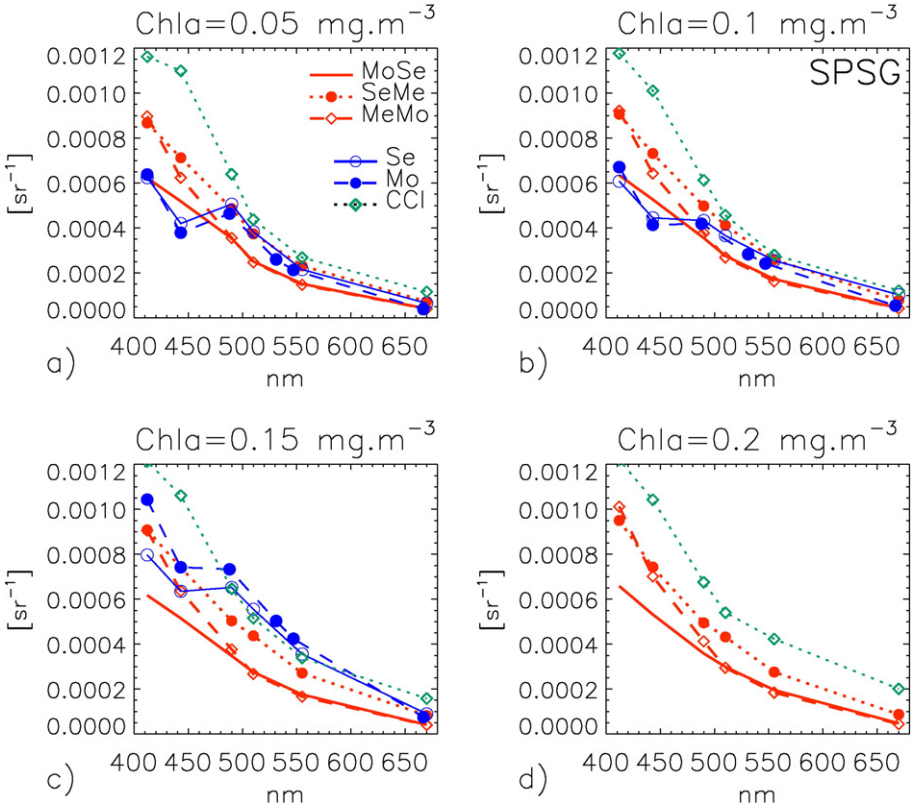

Fig. 11. σ spectra obtained from the three comparison pairs over the South Pacific subtropical gyre (red lines with three different line styles). Associated uncertainty estimates Ω from Hu et al. (2013) are provided for MODIS (dashed blue lines with filled circles) and SeaWiFS (blue lines with circles). Δ_u^{CCI} are shown with green dotted lines with diamonds. Statistics are given for Chl*a* equal to a) 0.05, b) 0.1, c) 0.15 and d) 0.2 mg m^{-3}. For wavelengths above 500 nm, the line MeMo is mostly overlaid by the line MoSe.

match-up data base, Moore et al. (2015) presented consistent results for SeaWiFS and MODIS: the validation statistics expressed in terms of Δ_u show a ratio close to 1 at comparable wavelengths (412 to 490 nm, from 0.98 to 1.1), and 1.48 in the red (SeaWiFS Δ_u higher). Altogether, validation statistics support the sensitivity analysis that has been conducted with λ values of 1, 1.3 and 1.5 (and their inverse). The differences seen between the estimates obtained with $\lambda = 1$ (across the three comparison pairs) as well as the variations observed by varying the value of λ provide an evaluation of the uncertainties that can be associated with σ, approximately 20–30% between 412 and 555 nm, and 30–40% at 670 nm. Boot-strapping tests indicate that a third or half of these uncertainties (between 412 and 555 nm) can be related to the selection of the R_{RS} matching pairs. One way to solve the issue of the value of λ is to use a triple collocation approach where matching pairs are required for the three missions contemporaneously but the disadvantage is that the number of match-ups would drop drastically. An approach assessing the variability of R_{RS} within adjoining pixels in areas of low natural variability could also help in determining the relative values of σ for the different missions.

The estimates of σ obtained from the collocation exercise were confronted with estimates obtained through other avenues, showing a general coherence with some differences. Part of these can be explained by the fact that the different approaches relied on data sets that differ in terms of time period or spatial scale. Furthermore, they were obtained with very different methods and each uncertainty estimate may represent a different facet of the uncertainty budget. The first comparison was with σ obtained from validation analyses. Overall, the various σ estimates are well coherent at the four sites considered (Fig. 9). σ estimates can also be compared with the unbiased RMS difference Δ_u obtained from validation but this term can be noticeably higher (20% or more) with respect to σ since it can be impacted by the slope of linear regression between the two data sets and by the uncertainties associated with the field data (Eq. (18)).

The σ estimates were compared as well with Δ_u obtained from the CCI project whereby validation statistics are extended at global scale through an optical classification scheme. Δ_u^{CCI} compared well with the other estimates at the validation sites MOBY and AAOT but yielded higher values at the Baltic sites. For subtropical gyres (Figs. 10 and 11), Δ_u^{CCI} appeared as the highest estimates particularly at 412 and 443 nm. Again part of these differences is due to the fact that Δ_u is higher than σ by construction (Eq. (18)); actually, Δ_u^{CCI} is closer to Δ_u^{ν} derived from the validation analyses. Moreover, CCI values are provided on a 24th-degree grid, so that Δ_u^{CCI} could be divided by 2 ($\sqrt{4}$) by averaging 4 values to express it at a 12th-degree resolution. In that case, Δ_u^{CCI} would be remarkably in line with σ for the MoSe comparison (e.g., going from 1.2 10^{-3} sr^{-1} to 0.6 10^{-3} sr^{-1} at 412 nm, Figs. 10 and 11). On the other hand, this might not be applicable, considering that validation statistics characterizing each water type very likely would not have differed by a factor of 2 if the CCI match-up exercise had been conducted on a 12th-degree grid. Based on the results obtained with the MeMo comparison conducted at both resolutions, a resolution degraded from 24th-degree to 12th-degree resulted in a decrease in σ of approximately 16%. This raises the question of how uncertainties behave across the small scale variability of satellite ocean color products (Doney, Glover, McCue, & Fuentes, 2003), how their estimates should be handled across different spatial resolutions, and how the difference in scale between satellite products and point field measurements is accounted for when uncertainty estimates are traceable to field data.

The third approach considered is derived from the comparison of outputs from two Chl*a* algorithms (Hu et al., 2013). The resulting estimates Ω are only available for oligotrophic regions with Chl*a* lower than 0.25 mg m^{-3} that represent a sizeable portion of the world ocean. Here again, there appeared to be a general coherence between σ and the model-derived Ω, expressed as the standard deviation of R_{RS} departures from an error-free R_{RS} determined by comparing the algorithm outputs. This error-free R_{RS} is intended as a value of reference and may differ from the theoretical true R_{RS}. So by construction, Ω does not include any contribution that might arise from biases existing between the satellite R_{RS} and the actual value. In that respect, its description of the uncertainty budget is incomplete, in the same way as σ's is. For completeness, it is mentioned that the Ω formulation as a function of Chl*a* was obtained with Level-2 satellite data (i.e., with pixel-size resolution). Tests performed by Hu et al. (2013) suggest that similar results would have been obtained with degraded resolutions typical of Level-3 products.

Next is discussed the σ distribution with respect to biases and mission objectives. Again, in the mathematical framework introduced in Section 2, σ is the part of the R_{RS} uncertainty budget associated with random effects, i.e., not affected by biases. The uncertainty budget should be completed by the systematic difference between reference state r and true state (that is only attainable with in situ data). Some space was also granted to results describing the bias δ between satellite products, which has some relevance when discussing the overall uncertainty budget. Indeed, the bias between R_{RS} from two missions can be seen as a lower estimate of half of the actual bias that would affect at least one of these products. So if δ between missions is any indicator of the actual bias (i.e., with respect to true values), then typical amplitudes of 0.1–0.2 10^{-3} sr^{-1} (Fig. 2c) represent 10–20% of the σ value in the blue. Considering the comparison pair MoSe (MODIS/SeaWiFS), the multi-annual average bias δ is found within $\pm\sigma$ for at least 96% of the macro-bins for wavelengths between 412 and 510 nm, 94% at 555 nm and 51% at 670 nm. These numbers are very similar for the other two pairs (i.e., around 95%) with some exceptions (lower for the pair SeMe at 555 and 670 nm, 86% and 19%, respectively, but higher at 670 nm for the pair MeMo, 75%).

It is also worth comparing σ with mission objectives (McClain et al., 1992) or GCOS (GCOS, 2011) requirements (5% for blue and green wavelengths), while bearing in mind that conclusions reached in relation with σ are subject to the σ uncertainties recorded above. The statistics provided below are derived with match-up pairs expressed on a 12th-degree grid and, considering the analysis performed at two different resolutions with the MeMo data set, they might be slightly degraded at the 4-km spatial resolution associated with the GCOS objectives. Insofar as the σ component exhibits a small variability (both in time and space) particularly for the blue bands, the uncertainty expressed in relative terms associated with R_{RS} varies at first order as the inverse of R_{RS}. Incidentally, this would support the definition of mission objectives in (radiometric) reflectance units. Here, the comparison between σ and mission objectives is done through computing the average ratio of σ (pair MoSe) and the multi-annual (2003–2007) average of R_{RS} (taking the SeaWiFS record as a reference). The global average of $\sigma/\bar{R_{RS}}$ is in the interval 7–11% between 412 and 555 nm, while it is lower for Chl*a* below 0.3 mg m^{-3} (7% between 412 and 490 nm) (Table 2). The results are again compatible with validation statistics for open ocean waters: the validation exercises performed at MOBY (Section 5.1) yield a

Table 2
Average ratio (expressed in %) between σ and the multi-annual (2003–2007) average of R_{RS} at the SeaWiFS wavelengths. σ is that associated with the SeaWiFS/MODIS pair, and the average R_{RS} is that of SeaWiFS. Statistics are given for various intervals of Chl*a* (in mg m^{-3}), with the last four columns associated with the intervals 0.025–0.075, 0.075–0.125, 0.125–0.175 and 0.175–0.225. The first line of statistics provides the percentage of the domain of analysis corresponding to the Chl*a* interval (in %).

Chl*a*	All	≤0.1	≤0.3	>0.3	0.05	0.1	0.15	0.2
Area [%]	100	30.5	77.7	21.6	20.2	18.7	16.2	12.2
412	9.8	4.9	7.4	18.3	4.5	6.5	8.3	9.8
443	8.5	5.2	7.1	13.5	4.8	6.5	8.0	8.9
490	7.2	5.6	6.5	9.6	5.3	6.3	7.0	7.4
510	9.0	8.0	8.7	10.2	7.8	8.7	9.2	9.3
555	11.4	11.1	11.5	11.1	10.9	11.7	11.9	11.8
670	22.3	23.0	22.3	22.6	23.3	22.4	21.7	21.3

mean absolute relative difference between field and satellite data of approximately 7% between 412 and 490 nm for the three missions, and Fig. 9a shows that at these wavelengths σ is indeed close to the field data uncertainty expressed in sr $^{-1}$. Another validation activity performed in Mediterranean oligotrophic waters found differences of 8% to 15% in that spectral range (Zibordi, Berthon, Mélin, & D'Alimonte, 2011) (another study in the Ligurian Sea showed differences above 20% at 412 nm for SeaWiFS and MODIS; Antoine et al., 2008). The ratio $\sigma/\bar{R}_{RS}$ tends to increase with Chl*a* in a way remarkably consistent with the results of Hu et al. (2013) (see their Fig. 6), with a value of 5–6% for Chl*a* of 0.1 mg m $^{-3}$. Therefore both studies reach the conclusion that the mission objective of a 5% accuracy is fulfilled at the blue bands (412–490 nm) for oligotrophic waters defined as Chl*a* ≤ 0.1-mg m $^{-3}$, whereas this figure is more like 5–10% for Chl*a* between 0.1 and 0.3 mg m $^{-3}$. Differently from the blue spectral domain, for green and red bands the ratio $\sigma/\bar{R}_{RS}$ appears relatively stable, at least within the broad categories included in Table 2. Indeed when considering coastal regions, both σ and the average R_{RS} tend to increase. For coastal sites, the σ estimates appear far from the field data uncertainty, equivalent to at least 5% (Fig. 9b,c,d). By construction, the σ terms have a zero average, which means that over suitable spatial and/or temporal scales, the uncertainties that are discussed here tend to be canceled out. On the other hand, this discussion is still relevant considering that GCOS requirements are expressed for a daily sampling associated with a 4-km resolution (GCOS, 2011).

7. Conclusions

This work described a collocation approach producing an uncertainty estimate σ representing the part of the R_{RS} uncertainty budget associated with random effects, i.e., not inclusive of biases. The σ distributions compared well with independent estimates obtained from field data or derived from algorithms. It is again stressed that σ, albeit resulting from a comparison of two data sets, is an intrinsic property of each data set. σ estimates are spatially resolved with a spatial resolution that can be that of the satellite products, even though degrading the resolution allows for more robust statistical calculations. Conversely, the approach has a very limited capacity to provide temporally-resolved information even if seasonal averages were produced; in general the temporal discrimination that is attainable ultimately depends on the amount of matching pairs available for analysis. But considering its very small variations in time, the global distribution of σ (Figs. 4 or 5a) appears as a lower-bound estimate of the R_{RS} uncertainty that is readily applicable to existing missions, keeping in mind that the bias could account for an additional contribution, e.g., representing 10–20% of the σ value in the blue. The fact that σ generally shows small variations with respect to uncertainty estimates expressed in relative terms (in %) further suggests that uncertainty goals should be expressed in radiometric units. Moreover, when applied to several sensor pairs and with different λ values, the results could be further exploited to better constrain the most probable value of the σ spectrum relative to each mission.

Using R_{RS} data expressed on a common 12th-degree grid, the global average of σ has been found to decrease with wavelength from approximately 0.7–0.9 10 $^{-3}$ sr $^{-1}$ at 412 nm to 0.05–0.1 10 $^{-3}$ sr $^{-1}$ at the red band. The spectral shape of σ actually reflects that of R_{RS} if oligotrophic waters are considered. However, the shape of σ is mostly conserved in coastal waters or marginal seas (Fig. 5a) for which R_{RS} spectra show different features (Mélin & Vantrepotte, 2015). Taking its global average (pair MoSe), the spectral dependence of σ is well reproduced by a power law $\lambda^{-4.3}$ ($r^2 = 0.983$, RMS difference between σ and fitted value of 0.03 10 $^{-3}$ sr $^{-1}$), which is actually close to the spectral dependence of the Rayleigh optical thickness (Bodhaine, Wood, Dutton, & Slusser, 1999). Considering that the spectral shape of σ is also reminiscent of the spectral character of the top-of-atmosphere radiance as well as the Rayleigh and aerosol radiances, this suggests the obvious conclusion that the amplitude of σ is strongly dependent on the calibration of the sensor and the performance of the atmospheric correction.

The study was completed by the description of other statistical indicators describing the differences between mission-specific data, δ, $|\psi|$ and ψ. These differences are fairly significant even though all three missions considered were processed with the same strategy for calibration and atmospheric correction. They may stem from a variety of sources (IOCCG, 2007): differences in sensor design, functioning and spectral characteristics may propagate through the processing chain, the atmospheric correction may respond differently to a different geometry of observation and illumination or varying aerosol types, and products may be affected by spatial mismatches associated with the mapping of pixels with different sizes and shapes. Moreover the overpass times are separated by a few hours, which may generate real differences, particularly in tidal regions (Wang et al., 2013), while in some areas little variability has been shown for reflectance fields (Zibordi et al., 2006b; Mélin & Zibordi, 2007). Generally, in most cases, it still appears challenging to confidently attribute differences between products to the natural variability occurring between subsequent overpass times with respect to the effect on the atmospheric correction occasioned by changes in the geometry of illumination and the atmospheric content (aerosols, clouds) or simply by noise. Fully documenting and explaining the variations of these quantities as a function of comparison pairs, space, season and wavelength is out of the scope of this paper, but some results of general validity were provided that can be summarized as follows:

- comparison statistics of relative differences ($|\psi|$ and ψ), display a larger spatial variability than the σ distribution, particularly at blue and red bands,
- the mean absolute relative difference $|\psi|$ shows a characteristic U-shape whose tilt changes with water type, with relatively low (high) values for blue wavelengths and high (low) values in the red in oligotrophic (turbid) waters, i.e., with a mirror evolution with respect to the R_{RS} amplitude,
- bias metrics may show significant seasonal variations (including changes in sign) for some regions such as the subtropical waters,
- on average, SeaWiFS R_{RS} tend to be slightly higher than MODIS R_{RS}, which in turn appear higher than MERIS R_{RS}.

Based on past experience with previous versions of the data sets, it is likely that major results of the paper (amplitude and spectral shape of σ, generally less spatial variability of σ with respect to $|\psi|$, U-shape of $|\psi|$ changing with optical water types) will not be strongly affected by future mission reprocessing efforts, unless a significant leap forward in terms of algorithms and/or instrument characterization is achieved. The last point listed above, associated with the bias distributions, may be the most susceptible to revisions since changes associated with reprocessing can affect the sign of δ particularly if δ is small (see example with MERIS in Zibordi, Mélin, Berthon, & Canuti, 2013).

The capacity to routinely associate R_{RS} distributions with a robust estimate of uncertainty is a field in its infancy but new developments are likely to be spurred by operational requirements. For instance a project such as ESA's CCI is including uncertainty estimates for R_{RS} in its data stream (Sathyendranath et al., 2016). In that context, collocation of coincident records from various missions appears as a relatively untapped resource in ocean-color science. This study is also the first attempt at quantitatively comparing spatially-resolved uncertainty estimates obtained from various methods. Even though developed independently with very different principles, a consistent picture of uncertainty estimates of R_{RS} emerged, but more work is needed to develop new approaches and to compare their different results, as well as to investigate how uncertainty estimates vary with mission, time, location and wavelength.

Collocation can also be used with a slightly different perspective if σ is considered known for a well-characterized mission, instead of assuming that the ratio of the σ terms associated with two missions (i.e., λ) is known. The system of Eqs. (9) and (10) can then be solved to yield σ for another, less characterized, mission. Considering that data bases of comparison between missions are likely to grow much faster than validation data sets when a new mission is launched, this approach also has a promising potential to contribute to the definition of the uncertainties of R_{RS} derived from a new satellite sensor in its early mission stages, and could be easily included in the validation strategies of new missions.

Obviously collocation relies on overlaps between missions. Of the three missions studied here, two are terminated and one is well past its expected life time, making it urgent to have additional missions joining MODIS and the Visible/Infrared Imager Radiometer Suite (VIIRS, (Schueler et al., 2002)) in orbit.

Acknowledgments

This work is a contribution to the Ocean Colour Climate Change Initiative (OC-CCI) of the European Space Agency. It is also a contribution to the activities of the National Centre for Earth Observation of the Natural Environmental Research Council, UK. G. Zibordi is warmly thanked for all the efforts leading to the availability of the AERONET-OC data and for comments on the manuscripts. The MOBY program, currently supported by NOAA, and K. Voss and K. Hughes are sincerely thanked for the availability of the data used for validation statistics. Acknowledgments are due to ESA and NASA for the distribution of satellite data. Thanks also go to the reviewers for thorough comments that really helped improve the manuscript.

References

Antoine, D., d'Ortenzio, F., Hooker, S.B., Bécu, G., Gentili, B., Tailliez, D., & Scott, A.J. (2008). Assessment of uncertainty in the ocean reflectance determined by three satellite ocean color sensors (MERIS, SeaWiFS and MODIS-A) at an offshore site in the Mediterranean Sea (BOUSSOLE project). *Journal of Geophysical Research*, *113*, C07013. http://dx.doi.org/10.1029/2007JC004472.

Bailey, S.W., & Werdell, P.J. (2006). A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Remote Sensing of Environment*, *102*, 12–23.

Berthon, J. -F., & Zibordi, G. (2010). Optically black waters in the northern Baltic Sea. *Geophysical Research Letters*, *37*, L09605. http://dx.doi.org/10.1029/2010GL043227.

Bodhaine, B.A., Wood, N.B., Dutton, E.G., & Slusser, J.R. (1999). On Rayleigh optical depth calculations. *Journal of Atmospheric and Oceanic Technology*, *16*, 1854–1861.

Bojinski, S., Verstraete, M., Peterson, T.C., Richter, C., Simmons, A., & Zemp, M. (2014). The concept of Essential Climate Variables in support of climate research, applications, and policy. *Bulletin of the American Meteorological Society*, *95*, 1431–1443.

Brown, S.W., Flora, S.J., Feinholz, M.J., Yarbrough, M.A., Houlihan, T., Peters, D., ... Clark, D.K. (2007). The Marine Optical BuoY (MOBY) radiometric calibration and uncertainty budget for ocean color satellite sensor vicarious calibration. In R. Meynart, S.P. Neeck, H. Shimoda, & S. Habib (Eds.), *Proceedings SPIE. Sensors, Systems, and Next-Generation Satellites XI, 6744*. (pp. 67441M).

Bulgarelli, B., & Zibordi, G. (2003). Remote sensing of ocean colour: Accuracy assessment of an approximate atmospheric correction method. *International Journal of Remote Sensing*, *24*, 491–509.

Bulgarelli, B., Mélin, F., & Zibordi, G. (2003). SeaWiFS-derived products in the Baltic Sea: Performance analysis of a simple atmospheric correction algorithm. *Oceanologia*, *45*, 655–677.

Caires, S., & Sterl, A. (2003). Validation of ocean wind and wave data using triple collocation. *Journal of Geophysical Research*, *108*, 3098. http://dx.doi.org/10.1029/2002JC001491.

Chomko, R.M., & Gordon, H.R. (2001). Atmospheric correction of ocean color imagery: Test of the spectral optimization algorithm with the Sea-viewing Wide Field-of-View Sensor. *Applied Optics*, *40*, 2973–2984.

Clark, D.K., Gordon, H.R., Voss, K.J., Ge, Y., Broenkow, W., & Trees, C. (1997). Validation of atmospheric correction over the oceans. *Journal of Geophysical Research*, *102*(D14), 17209–17217.

Cummings, J., Bertino, L., Brasseur, P., Fukumori, I., Kamachi, M., Martin, M.J., ... Weaver, A. (2009). Ocean data assimilation systems for GODAE. *Oceanography*, *22*, 96–109.

D'Odorico, P., Gonsamo, A., Pinty, B., Gobron, N., Coops, N., Mendez, E., & Schaepman, M.E. (2014). Intercomparion of fraction of absorbed photosynthetically active radiation products derived from satellite data over Europe. *Remote Sensing of Environment*, *142*, 141–154.

Doney, S.C., Glover, D.M., McCue, S.J., & Fuentes, M. (2003). Mesoscale variability of Sea-viewing Wide Field-of-view Sensor (SeaWiFS) satellite ocean color: Global patterns and spatial scales. *Journal of Geophysical Research*, *108*, 3024. http://dx.doi.org/10.1029/2001JC000843.

Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M. -H., Féménias, P., Frerick, J., ... Sciarra, R. (2012). The Global Monitoring for Environment and Security (GMES) Sentinel-3 mission. *Remote Sensing of Environment*, *120*, 37–57.

Esaias, W.E., Abbott, M.R., Barton, I., Brown, O.B., Campbell, J.W., Carder, K.L., ... Minnett, P.J. (1998). An overview of MODIS capabilities for ocean science observations. *IEEE Transactions on Geoscience and Remote Sensing*, *36*, 1250–1265.

Fang, H., Wei, S., Jiang, C., & Scipal, K. (2012). Theoretical uncertainty analysis of global MODIS, CYCLOPES, and GLOBCARBON LAI products using a triple collocation method. *Remote Sensing of Environment*, *124*, 610–621.

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

Frouin, R., & Pelletier, B. (2015). Bayesian methodology for inverting satellite ocean-color data. *Remote Sensing of Environment*, *159*, 332–360.

GCOS (2011). Systematic observation requirements for satellite-based data products for climate, 2011 update: Supplemental details to the satellite-based component of the Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC. *GCOS, 154*. World Meteorological Organization.

Gergely, M., & Zibordi, G. (2014). Assessment of AERONET-OC L. uncertainties. *Metrologia*, *51*, 40–47.

Goyens, C., Jamet, C., & Schroeder, T. (2013). Evaluation of four atmospheric correction algorithms for MODIS-Aqua images over constrasted coastal waters. *Remote Sensing of Environment*, *131*, 63–75.

Gregg, W.W. (2008). Assimilation of SeaWiFS ocean chlorophyll data into a three-dimensional global ocean model. *Journal of Marine Systems*, *69*, 205–225.

Gregg, W.W., & Casey, N.W. (2004). Global and regional evaluation of the SeaWiFS chlorophyll data set. *Remote Sensing of Environment*, *93*, 463–479.

GUM (2008). Guide to the expression of uncertainty in measurement. *Joint Committee for Guides in Metrology* (134 pp.).

Hollmann, R., Merchant, C.J., Saunders, R., Downy, C., Buchwitz, M., Cazenave, A., ... Wagner, W. (2013). The ESA Climate Change Initiative — Satellite data records for essential climate variables. *Bulletin of the American Meteorological Society*, *94*, 1541–1552.

Hu, C., Feng, L., & Lee, Z. -P. (2013). Uncertainties of SeaWiFS and MODIS remote sensing reflectance: Implications from clear water assessments. *Remote Sensing of Environment*, *133*, 168–182.

Hu, C., Lee, Z. -P., & Franz, B.A. (2012). Chlorophyll *a* algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference. *Journal of Geophysical Research*, *117*, C01011. http://dx.doi.org/10.1029/2011JC007395.

IOCCG (2007). Ocean Colour Data Merging. In W.W. Gregg, J. Aiken, E. Kwiatkowska, S. Maritorena, F. Mélin, H. Murakami, S. Pinnock, & C. Pottier (Eds.), *Reports of the International Ocean-Colour Coordinating Group, 5*. Dartmouth, Canada: IOCCG (65pp).

Kahru, M., Lee, Z. -P., Kudela, R.M., Manzano-Sarabia, M., & Mitchell, B.G. (2015). Multi-satellite time series of inherent optical properties in the California Current. *Deep-Sea Research Part II*, *112*, 91–106.

Lee, Z. -P., Arnone, R.A., Hu, C., Werdell, P.J., & Lubac, B. (2010). Uncertainties of optical parameters and their propagations in an analytical ocean color inversion algorithm. *Applied Optics*, *49*, 369–381.

Lee, Z. -P., Carder, K.L., & Arnone, R.A. (2002). Deriving inherent optical properties from water color: A multiband quasi-analytical algorithm for optically deep waters. *Applied Optics*, *41*, 5755–5772.

Lee, Z. -P., Lubac, B., Werdell, P.J., & Arnone, R.A. (2009). An update of the Quasi-Analytical Algorithm (QAA_v5). *Technical Report, International Ocean-Colour Coordinating Group* (http://www.ioccg.org/groups/software.html).

Legendre, P., & Legendre, L. (1998). *Numerical Ecology* (2nd ed.). Elsevier (853 pp).

Longhurst, A.R. (2006). *Ecological geography of the sea*. Academic Press (560 pp.).

Maritorena, S., Hembise Fanton d'Andon, O., Mangin, A., & Siegel, D.A. (2010). Merged satellite ocean color data products using a bio-optical model: Characteristics, benefits and issues. *Remote Sensing of Environment*, *114*, 1791–1804.

McClain, C.R., Cleave, M.L., Feldman, G.C., Gregg, W.W., Hooker, S.B., & Kuring, N. (1998). Science quality SeaWiFS data for global biosphere research. *Sea Technology*, *39*, 10–16.

McClain, C.R., Esaias, W.E., Barnes, W., Guenther, B., Endres, D., Hooker, S.B., ... Barnes, R. (1992). SeaWiFS calibration and validation plans. In S.B. Hooker, & E.R. Firestone (Eds.), *NASA Technical Memorandum 1992–104566*. Maryland: NASA-GSFC, Greenbelt (3, 42 pp).

McClain, C.R., Feldman, G.C., & Hooker, S.B. (2004). An overview of the SeaWiFS project and strategies for producing a climate research quality global ocean bio-optical time series. *Deep-Sea Research Part II*, *51*, 5–42.

Mélin, F. (2010). Global distribution of the random uncertainty associated with satellite derived Chla. *IEEE Geoscience and Remote Sensing Letters*, *7*, 220–224.

Mélin, F. (2011). Comparison of SeaWiFS and MODIS time series of inherent optical properties for the Adriatic Sea. *Ocean Science*, *7*, 351–361.

Mélin, F., & Franz, B.A. (2014). Assessment of satellite ocean colour radiometry and derived geophysical products. In G. Zibordi, C. Donlon, & A. Parr (Eds.), *Optical Radiometry for Oceans Climate Measurements. Experimental Methods in the Physical Sciences, vol. 47*, . Academic Press (722 pp.,chapter 6.1).

Mélin, F., & Sclep, G. (2015). Band-shifting for ocean color multi-spectral reflectance data. *Optics Express*, *23*, 2262–2279.

Mélin, F., & Vantrepotte, V. (2015). How optically diverse is the coastal ocean? *Remote Sensing of Environment*, *160*, 235–251.

Mélin, F., & Zibordi, G. (2007). Optically based technique for producing merged spectra of water-leaving radiances from ocean color remote sensing. *Applied Optics*, *46*, 3856–3869.

Mélin, F., Steinich, C., Gobron, N., Pinty, P., & Verstraete, M.M. (2002). Optimal merging of LAC and GAC data from SeaWiFS. *International Journal of Remote Sensing*, *23*, 801–807.
Mélin, F., Vantrepotte, V., Clerici, M., D'Alimonte, D., Zibordi, G., Berthon, J. -F., & Canuti, E. (2011a). Multi-sensor satellite time series of optical properties and chlorophyll *a* concentration in the Adriatic Sea. *Progress in Oceanography*, *91*, 229–244.
Mélin, F., Zibordi, G., & Berthon, J. -F. (2007). Assessment of satellite ocean color products at a coastal site. *Remote Sensing of Environment*, *110*, 192–215.
Mélin, F., Zibordi, G., & Berthon, J. -F. (2012). Uncertainties in remote sensing reflectance from MODIS-Terra. *IEEE Geoscience and Remote Sensing Letters*, *9*, 432–436.
Mélin, F., Zibordi, G., Berthon, J. -F., Bailey, S.W., Franz, B.A., Voss, K.J., ... Grant, M. (2011b). Assessment of MERIS reflectance data as processed by SeaDAS over the European Seas. *Optics Express*, *19*, 25657–25671.
Mélin, F., Zibordi, G., Carlund, T., Holben, B.N., & Stefan, S. (2013b). Validation of SeaWiFS and MODIS Aqua/Terra aerosol products in coastal regions of European marginal seas. *Oceanologia*, *55*, 27–51.
Mélin, F., Zibordi, G., & Djavidnia, S. (2009). Merged series of normalized water leaving radiances obtained from multiple satellite missions for the Mediterranean Sea. *Advances in Space Research*, *43*, 423–437.
Mélin, F., Zibordi, G., & Holben, B.N. (2013a). Assessment of the aerosol products from the SeaWiFS and MODIS ocean color missions. *IEEE Geoscience and Remote Sensing Letters*, *10*, 1185–1189.
Moore, T.S., Campbell, J.W., & Dowell, M.D. (2009). A class-based approach to characterizing and mapping the uncertainty of MODIS ocean chlorophyll product. *Remote Sensing of Environment*, *113*, 2424–2430.
Moore, T.S., Campbell, J.W., & Feng, H. (2015). Characterizing the uncertainties in spectral remote sensing reflectance for SeaWiFS and MODIS-Aqua based on global in situ matchup data sets. *Remote Sensing of Environment*, *159*, 14–27.
Neukermans, G., Ruddick, K., Bernard, E., Ramon, D., Nechad, Deschamos, & P.-Y (2009). Mapping total suspended matter from geostatioanry satellites: A feasibility study with SEVIRI in the southern North Sea. *Optics Express*, *17*, 14029–14052.
O'Carroll, A.G., Eyre, J.R., & Saunders, R.W. (2008). Three-way error analysis between AATSR, AMSR-E, and in-situ sea surface temperature observations. *Journal of Atmospheric and Oceanic Technology*, *25*, 1197–1207.
O'Reilly, J.E., Maritorena, S., Siegel, D.A., O'Brien, M.C., Toole, D.A., Mitchell, B.G., ... Culver, M. (2000). Ocean color chlorophyll a algorithms for SeaWiFS, OC2, and OC4: Version 4. In S.B. Hooker, & E.R. Firestone (Eds.), *NASA Technical Memorandum 2000–206892, 20, 9–23*. Greenbelt, Maryland: NASA-GSFC.
Pottier, C., Garçon, V., Larnicol, G., Sudre, J., Schaeffer, P., & Le Traon, P. -Y. (2006). Merging SeaWiFS and MODIS/Aqua ocean color data in North and Equatorial Atlantic using weighted averaging and objective analysis. *IEEE Transactions on Geoscience and Remote Sensing*, *44*, 3436–3451.
Rast, M., Bezy, J.L., & Bruzzi, S. (1999). The ESA Medium Resolution Imaging Spectrometer MERIS — A review of the instrument and its mission. *International Journal of Remote Sensing*, *20*, 1681–1702.
Roebeling, R.A., Wolters, E.L.A., Meirink, J.F., & Leijnse, H. (2012). Triple collocation of summer precipitation retrievals from SEVIRI over Europe with gridded rain gauge and weather radar data. *Journal of Hydrometeorology*, *13*, 1552–1566.
Salama, M.S., & Stein, A. (2009). Error decomposition and estimation of inherent optical properties. *Applied Optics*, *48*, 4947–4962.
Salama, M.S., Dekker, A.G., Su, Z., Mannaerts, C.M., & Verhoef, W. (2009). Deriving inherent optical properties and associated inversion-uncertainties in the Dutch lakes. *Hydrology and Earth System Sciences*, *13*, 1113–1121.
Salisbury, D.J., Anguelova, M.D., & Brooks, I.M. (2014). Global distribution and seasonal dependence of satellite-based whitecap fraction. *Geophysical Research Letters*, *41*, 1616–1623. http://dx.doi.org/10.1002/2014GL059246.
Sathyendranath, S., et al. (2016). Creating an ocean-colour time series for use in climate studies. *The experience of the ocean-colour Climate Change Initiative. submit.*
Schueler, C.F., Clement, J.E., Ardanuy, P.E., Welsch, C., DeLuccia, F., & Swenson, H. (2002). NPOESS VIIRS sensor design overview. *Proceedings SPIE 4483, Earth Observing Systems VI* (10.1117/12.453451).
Scipal, K., Holmes, T., de Jeu, R., Naeimi, V., & Wagner, W. (2008). A possible solution for the problem of estimating the error structure of global soil moisture data sets. *Geophysical Research Letters*, *35*, L24403. http://dx.doi.org/10.1029/2008GL035599.
Stamnes, K., Li, W., Yan, B., Eide, H., Barnard, A., Pegau, W.S., & Stamnes, J.J. (2003). Accurate and self-consistent ocean color algorithms: Simultaneous retrieval of aerosol optical properties and chlorophyll concentrations. *Applied Optics*, *42*, 939–951.
Stoffelen, A. (1998). Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. *Journal of Geophysical Research*, *103*, 7755–7766.
Toohey, M., & Strong, K. (2007). Estimating biases and error variances through the comparison of coincident satellite measurements. *Journal of Geophysical Research*, *112*, D13306. http://dx.doi.org/10.1029/2006JD008192.
Vanhellemont, Q., Neukermans, G., & Ruddick, K. (2014). Synergy between polar-orbiting and geostationary sensors: Remote sensing of the ocean at high spatial and high temporal resolution. *Remote Sensing of Environment*, *146*, 49–62.
Vantrepotte, V., & Mélin, F. (2011). Inter-annual variations in the SeaWiFS global chlorophyll a concentration (1997–2007). *Deep Sea Research, Part I*, *58*, 439–441.
Wang, M., Ahn, J.H., Jiang, L., Shi, W., Son, S., Park, Y. -J., & Ryu, J. -H. (2013). Ocean color products from the Korean Geostationary Ocean Color Imager (GOCI). *Optics Express*, *21*, 3835–3849.
Wang, P., Boss, E.S., & Roesler, C. (2005). Uncertainties of inherent optical properties obtained from semi-analytical inversions of ocean color. *Applied Optics*, *44*, 4074–4085.
Werdell, P.J., Franz, B.A., Bailey, S.W., Feldman, G.C., Boss, E., Brando, V.E., ... Mangin, A. (2013). A generalized ocean color inversion for retrieving marine inherent optical properties,. *Applied Optics*, *52*, 2019–2037.
Zibordi, G., Berthon, J. -F., Mélin, F., & D'Alimonte, D. (2011). Cross-site consistent in-situ measurements for satellite ocean color applications: the BiOMaP radiometric dataset. *Remote Sensing of Environment*, *115*, 2104–2115.
Zibordi, G., Berthon, J. -F., Mélin, F., D'Alimonte, D., & Kaitala, S. (2009a). Validation of satellite ocean color primary products at optically complex coastal sites: Northern Adriatic Sea, northern Baltic Proper, Gulf of Finland. *Remote Sensing of Environment*, *113*, 2574–2591.
Zibordi, G., Holben, B.N., Hooker, S.B., Mélin, F., Berthon, J. -F., Slutsker, I., ... Al Mandoos, A. (2006a). A network for standardized ocean color validation measurements. *Eos, Transactions American Geophysical Union*, *87*(30), 293–297.
Zibordi, G., Holben, B.N., Slutsker, I., Giles, D., D'Alimonte, D., Mélin, F., ... Seppälä, J. (2009b). AERONET-OC: A network for validation of ocean color primary products. *Journal of Atmospheric and Oceanic Technology*, *26*, 1634–1651.
Zibordi, G., Mélin, F., & Berthon, J. -F. (2006b). A time series of above-water radiometric measurements for coastal water monitoring and remote sensing product validation. *IEEE Geoscience and Remote Sensing Letters*, *3*, 120–124.
Zibordi, G., Mélin, F., & Berthon, J. -F. (2012). Intra-annual variations of biases in remote sensing primary ocean color products at a coastal site. *Remote Sensing of Environment*, *124*, 627–636.
Zibordi, G., Mélin, F., Berthon, J. -F., & Canuti, E. (2013). An assessment of MERIS ocean color products for European seas. *Ocean Science*, *9*, 521–533.
Zibordi, G., Mélin, F., Berthon, J. -F., & Talone, M. (2014). In situ autonomous radiometry measurements for satellite ocean color validation in the western Black Sea. *Ocean Science Discussions*, *11*, 3003–3034.
Zibordi, G., Mélin, F., Voss, K., Johnson, B.C., Franz, B.A., Kwiatkowska, E., ... Antoine, D. (2015). System vicarious calibration for ocean color climate change applications: Requirements for in situ data. *Remote Sensing of Environment*, *159*, 361–369.
Zibordi, G., Strömbeck, N., Mélin, F., & Berthon, J. -F. (2006c). Tower-based radiometric observations at a coastal site in the Baltic Proper. *Estuarine, Coastal and Shelf Science*, *69*, 649–654.