Elsevier Editorial System(tm) for Remote Sensing of Environment
Manuscript Draft

Manuscript Number: RSE-D-14-01055R1

Title: System Vicarious Calibration for Ocean Color Climate Change Applications: Requirements for In Situ Data

Article Type: Original Research Paper

Keywords: Ocean Color, System Vicarious Calibration, Climate Data Records, Essential Climate Variables

Corresponding Author: Dr. Giuseppe Zibordi,

Corresponding Author's Institution: Joint Research Centre

First Author: Giuseppe Zibordi

Order of Authors: Giuseppe Zibordi; Giuseppe Zibordi; Mélin Frédéric; Voss Kenneth; Johnson B Carol; Franz A Bryan; Ewa Kwiatkowska; Huot Jean-Paul; Wang Menghua; David Antoine

System Vicarious Calibration for Ocean Color Climate Change Applications: Requirements for *In Situ* Data

G. Zibordi[1*], *F. Mélin*[1], *K.J. Voss*[2], *B.C. Johnson*[3], *B.A. Franz*[4], *E. Kwiatkowska*[5], *J.P. Huot*[6], *M. Wang*[7], *D. Antoine*[8, 9]

[1] European Commission, Joint Research Centre, Ispra, Italy;

[2] Physics Department, University of Miami, Coral Gables, FL, USA

[3] Sensor Science Division, National Institute of Standards and Technology, Gaithersburg, MD, USA

[4] Ocean Biology Processing Group, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD, USA

[5] EUMETSAT, Darmstadt, Germany

[6] European Space Agency, Noordwijk, The Netherlands

[7] National Oceanic and Atmospheric Administrations, Center for Satellite Applications and Research, College Park, MD, USA

[8] Sorbonne Universités, Université Pierre et Marie Curie, Paris 06, UMR 7093, Laboratoire d'Océanographie de Villefranche, Villefranche-sur-Mer, France

[9] Department of Imaging and Applied Physics, Remote Sensing and Satellite Research Group, Curtin University, Perth, WA, Australia

*Corresponding author: Tel. +39 0332 785902; *email address* giuseppe.zibordi@jrc.ec.europa.eu

ABSTRACT

System Vicarious Calibration (SVC) ensures a relative radiometric calibration to satellite ocean color sensors that minimizes uncertainties in the water-leaving radiance L_w derived from the top of atmosphere radiance L_T. This is achieved through the application of gain-factors, *g*-factors, to pre-launch absolute radiometric calibration coefficients of the satellite sensor corrected for temporal changes in radiometric sensitivity. The *g*-factors are determined by the ratio of simulated to measured spectral L_T values where the former are computed using: *i*. highly accurate *in situ* L_w reference measurements; and *ii*. the same atmospheric models and algorithms applied for the atmospheric correction of satellite data. By analyzing basic relations between relative uncertainties of L_w and L_T, and *g*-factors consistently determined for the same satellite mission using different *in situ* data sources, this work suggests that the creation of ocean color Climate Data Records (CDRs) should ideally rely on: *i*. one main long-term *in situ* calibration system (site and radiometry) established and sustained with the objective to maximize accuracy and precision over time of *g*-factors and thus minimize possible biases among satellite data products from different missions; and additionally *ii*. unique (i.e., standardized) atmospheric model and algorithms for atmospheric correction to maximize cross-mission consistency of data products at locations different from that supporting SVC. Finally, accounting for results from the study and elements already provided in literature, requirements and recommendations for SVC sites and field radiometric measurements are streamlined.

Keywords: Ocean Color, System Vicarious Calibration, Climate Data Record

1. Introduction

In recent decades, measurements of ocean color from earth-orbiting satellite sensors have demonstrated high value for a number of applications ranging from regional water quality assessment (e.g., Attila et al. 2013) to global climate change investigations (e.g., Behrenfeld et al. 2006). Confidence in results from these applications, however, depends on accuracy of the satellite-derived data products. The primary ocean color product is the spectral water-leaving radiance L_w , i.e., the radiance emerging from the sea that is retrieved from the total radiance L_T detected by the satellite, whose accuracy determines those of satellite-derived data products. These include the spectral distribution of the normalized water-leaving radiance L_{WN} (i.e., the water-leaving radiance that would occur with no atmosphere, the sun at the zenith and at the mean sun-earth distance) or of the equivalent remote sensing reflectance R_{RS}, applied to determine geophysical quantities such as the near-surface chlorophyll-a concentration (*Chla*).

Early accuracy requirements for satellite ocean color data products generally refer to the work of Gordon and Clark (1981), Gordon et al. (1983) and Gordon (1987). By considering oligotrophic waters, they indicated a 5% uncertainty for L_w in the blue spectral region to allow for the determination of *Chla* concentration with a 35% maximum uncertainty. Subsequently, spectrally independent uncertainties of 5%, with a 1% inter-band uncertainty, were included among the objectives of the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) mission (Hooker et al. 1992). These target uncertainties were later retained for successive missions and have become a science requirement for the ocean color community.

Achievement of the spectrally independent 5% uncertainty target in satellite-derived L_w is mostly challenged by the accuracy of the absolute radiometric calibration of satellite optical sensors and additionally by uncertainties in the quantification of the large atmospheric perturbations affecting L_T. In particular, current uncertainties of approximately 2-3% (Butler et al. 2007, Eplee et al. 2011, Esposito et al. 2004) in the absolute radiometric calibration of satellite sensors in the visible spectral region and the additional uncertainties affecting the atmospheric correction process generally larger than a few percent (IOCCG 2010), may lead to large differences among multi-mission L_w data (Zibordi et al. 2006).

These limitations can be resolved through the so-called System Vicarious Calibration (SVC) that determines vicarious adjustment gain-factors *g* (hereafter *g*-factors) for absolute radiometric calibration coefficients of satellite sensors (Gordon 1998) through simulation of top-of-atmosphere L_T using: *i*. highly accurate *in situ* L_w reference measurements; and *ii*. the same atmospheric models and algorithms as applied for the atmospheric correction of satellite data. The *g*-factors, determined by the ratio of simulated to measured spectral L_T values, aim at minimizing the combined effects of: *i*. uncertainties due to the absolute pre-flight radiometric calibration and characterization of the satellite sensor after correction for sensitivity change with time; and *ii*. inaccuracy of the models and algorithms applied in the atmospheric correction to determine L_w from L_T. Clearly, the SVC process allows the determination of L_w with the lowest uncertainty when satellite observation conditions are equivalent to those characterizing the data applied for the calculation of *g*-factors (i.e., when mean biases affecting the SVC and the regular atmospheric correction processes are identical and fully compensate each other). It must be emphasized that the system nature of SVC requires re-computing *g*-factors after any change in the models or algorithms applied for the atmospheric correction, or any significant change in instrument absolute or temporal calibration knowledge.

By considering uncertainty requirements for satellite-derived L_w applicable for the construction of Climate Data Records (CDRs), which serve as core climate benchmark observations, the present work investigates the calibration needs for L_T with the objective of discussing requirements for *in situ* L_w data suitable for the determination of *g*-factors.

2. System Vicarious Calibration Requirements

Vicarious calibration broadly refers to the indirect calibration of satellite sensors through simulation of top-of-atmosphere data (Koepke 1982). Generic vicarious calibration methods based on atmospheric models and algorithms different from those applied for the operational data processing cannot reduce absolute uncertainties in derived radiometric calibration factors below a few percent (IOCCG, 2013).

This may lead to very large uncertainties in satellite-derived L_w (see §2.1). Consequently, unlike SVC that minimizes uncertainties in retrieved L_w (Gordon 1998), generic vicarious calibration methods are best applied for the quality check of pre-launch absolute radiometric calibrations of satellite ocean color sensors.

In view of supporting the discussion on accuracy and precision needs for *g*-factors from SVC, the following subsections will review: *i*. requirements for the construction of CDRs from satellite-derived L_w; and *ii*. legacy requirements for *in situ* L_w reference measurements.

2.1 Requirements for CDRs of L_w

CDRs of Essential Climate Variables (ECVs) are intended to support climate change investigations through time-series of core benchmark observations with enough accuracy to allow the detection of long-term trends embedded in large natural variations (Leroy et al. 2008).

Requirements for the generation of a CDR of satellite-derived L_{WN} from L_w (WMO 2011), which is the fundamental satellite ocean color ECV, include:

1. Radiometric uncertainty lower than 5% in the blue and green spectral regions (downscaled with respect to the spectrally independent 5% uncertainty target listed among the objectives of several ocean color missions);
2. Stability better than 0.5% over a decade.

The requirement on uncertainty is essential to understand climate-driven processes and changes, while the requirement on stability is essential to confidently determine long-term changes or trends (Ohring et al. 2005).

As already anticipated, the strict requirement of 5% maximum uncertainty for L_w determined from L_T at relevant wavelengths, requires the application of SVC. While this need is commonly accepted by the satellite ocean color community, the accuracy and precision required for *g*-factors for different missions supporting the creation of CDRs appears less consolidated.

To strictly address such a need, the relationship linking uncertainties in L_w and L_T is hereafter investigated through the use of the measurement equation. Specifically, in the absence of atmospheric gaseous absorption and sun glint and foam perturbations, the top-of-atmosphere radiance L_T can be related to L_w through the following simplified model

$$L_T = L_R + L_A + L_w t_d \quad (1)$$

where L_R and L_A indicate the Rayleigh and aerosol atmospheric radiance contributions, and t_d is the diffuse atmospheric transmittance that varies with atmospheric path-length and constituents. By assuming the values of L_R and L_A are exactly determined for any given observation condition, following the Guide to the Expression of Uncertainty in Measurement (JCGM, 2008) Zibordi and Voss (2014) provided equations relating absolute uncertainties of L_T, $u(L_T)$, to those of L_w, $u(L_w)$, and also linking relative uncertainties $u(L_T)/L_T$ to $u(L_w)/L_w$. In agreement with their work, $u(L_T)$ and $u(L_T)/L_T$ are given by

$$u(L_T) = u(L_w) t_d \quad (2)$$

and

$$\frac{u(L_T)}{L_T} = \frac{u(L_w)}{L_w} t_d \frac{L_w}{L_T} . \quad (3)$$

For the purpose of this study centered on SVC, the uncertainties related to the atmospheric correction process do not influence the determination of L_w because both SVC and the atmospheric correction rely on the same robust atmospheric models and algorithms. Thus, to a first approximation Eq. 2 indicates that the absolute uncertainties $u(L_T)$ and $u(L_w)$ are solely related by the factor t_d. Differently, Eq. 3 shows that relative uncertainties $u(L_T)/L_T$ and $u(L_w)/L_w$ are additionally related by the ratio L_w/L_T. Because of this, while the relation between absolute uncertainties only slightly varies with the atmospheric optical properties through t_d, the dependence between relative uncertainties is highly variable with both marine and atmospheric optical properties, which affect the term $t_d \cdot L_w/L_T$. Thus, while satellite-derived L_w may

exhibit similar absolute uncertainties for data collected over different water types, the corresponding relative uncertainties may largely differ as a function of L_w and L_T. Considering that requirements for satellite ocean color CDRs are provided in relative terms (e.g., see Ohring et al. 2005 and WMO 2011), the following analysis only focuses on relative uncertainties.

Rearranging Eq. 3 as a function of $u(L_T)/L_T$, for which a realistic spectrally independent radiometric uncertainty of 2% is assumed together with an ideal value of t_d=1, $u(L_w)/L_w$ would be approximately 20%, 40% and 200% for L_w/L_T equal to 0.10, 0.05 and 0.01, respectively. These uncertainty values, which may tentatively refer to blue, green and red wavelengths in oligotrophic waters, show the impossibility of meeting science requirements when only relying on current absolute radiometric calibration uncertainties, even assuming an exact quantification of the atmospheric perturbations.

Conversely, the application of Eq. 3 assuming t_d=1 and a spectrally independent uncertainty of 5% for L_w, implies values of $u(L_T)/L_T$ as low as 0.5%, 0.25% and 0.05% for L_w/L_T equal to 0.10, 0.05 and 0.01, respectively. These values provide an estimate for the required spectral uncertainties of absolute radiometric calibrations for satellite ocean color sensors and further confirm that: *i*. even assuming that the uncertainties in $u(L_w)/L_w$ due to atmospheric correction are negligible, the sole uncertainties currently affecting in-flight absolute radiometric calibration are an impediment to meet ocean color science requirements for CDRs; and that *ii*. SVC is the only viable alternative to overcome limitations due to uncertainties in absolute radiometric calibration and atmospheric correction. It is additionally observed that, even accounting for future developments in absolute radiometric calibration, that are expected to considerably reduce uncertainties (Cramer et al. 2013 ans Levick et al. 2014), SVC will still remain an essential component of any ocean color mission to minimize effects of inaccurate atmospheric corrections.

2.2 Legacy Requirements for SVC sites and data

Early indications on the appropriateness of SVC sites for global missions (mostly derived from Gordon (1998)) include:

i. Cloud free, very clear and maritime atmosphere with aerosol optical thickness $\tau_a < 0.1$ in the visible, which maximizes the potential number of satellite and *in situ* coincident data (i.e., matchups) and additionally optimizes the performance of the atmospheric correction process;

ii. Horizontally uniform L_w over spatial scales of a few kilometers to increase the comparability between satellite and *in situ* data at different spatial resolutions;

iii. Mesotrophic waters to minimize the effects of *in situ* L_w measurement uncertainties in the blue spectral region (this requirement has been considered less stringent with respect to the previous two, leding to consider oligotrophic waters as an appropriate alternative);

iv. Coincident aerosol measurements to assess the atmospheric correction process.

In situ L_w data applicable for SVC are expected to have low uncertainty through the application of state-of-the-art instrumentation, data reduction and quality assurance/control. Indications, mostly derived from Clark et al. (2002), include the need for:

i. Hyper-spectral measurements to cover any ocean color spectral band regardless of its center-wavelengths and spectral responses;

ii. Fully characterized *in situ* radiometers to minimize uncertainties and allow their comprehensive quantification;

iii. Traceability of data to the International System of Units (SI) to ensure consistency with community shared measurement methods and standards.

Also, in the case of global data products contributing to the construction of CDRs, SVC should be applied using *in situ* L_w from measurement sites representative of the most common satellite observation conditions, i.e., the world oceans. The determination of regional *g*-factors has also been proposed for areas exhibiting unique optical features (Franz et al. 2007). It is, however, recognized that this solution is mostly intended to support local applications where accurate *in situ* L_w data exist.

Ultimately, the limited number of highly accurate *in situ* data and their high costs challenge SVC at large. This has generated debates on the suitability of a number of data sources for SVC and also motivation for various studies to explore legacy requirements. These studies have produced a number of

203 *g*-factors for the same satellite sensor relying on equivalent versions of the atmospheric correction code,
204 but using L_w from different sources. As will be shown later, results offer the great opportunity to
205 investigate differences among actual *g*-factors in view of discussing implications for the creation of
206 CDRs.
207
208
209 **3. Literature data**
210 Among *in situ* systems specifically designed to support SVC for satellite ocean color sensors, only two
211 ensured almost continuous data collection across a number of satellite ocean color missions. These are: *i*.
212 the Marine Optical Buoy (MOBY) developed by the National Oceanic and Atmospheric Administration
213 (NOAA) and the National Aeronautics and Space Administration (NASA) for the SeaWiFS and the
214 Moderate Resolution Imaging Spectroradiometer (MODIS) (Clark et al. 1997); and *ii*. the Buoy for the
215 Acquisition of a Long-Term Optical Time Series (Bouée pour L'acquisition de Séries Optiques à Long
216 Terme, BOUSSOLE), developed for the Medium Resolution Imaging Spectrometer (MERIS) by the
217 Laboratoire d'Océanographie de Villefranche (LOV) in collaboration with a number of agencies
218 (Antoine at al. 2008a).
219 Aside from MOBY and BOUSSOLE (Eplee et al. 2001, Franz et al. 2007, Bailey et al. 2008), a
220 number of alternative data sources were considered for SVC of SeaWiFS data (see Table 1). These
221 included *in situ* data sets obtained by combining measurements from a variety of instruments and
222 reduction schemes (Bailey et al. 2008), data from specific coastal areas commonly applied for regional
223 investigations (Mélin and Zibordi 2010), as well as modeled data (Werdell et al. 2007). Derived *g*-
224 factors, consistently determined by applying the scheme detailed in Franz et al. (2007) and the SeaWiFS
225 Data Analysis System (SeaDAS, Fu et al. 1998) software package (version 5), are summarized in Table
226 2.
227 In agreement with Franz et al. (2007) and with specific reference to SeaWiFS center-wavelengths, *g*-
228 factors are assumed fixed and equal to unity at 865 nm, while the value at 765 nm is computed by

imposing a pure maritime aerosol model for locations in the oligotrophic gyres of the southern hemisphere. Spectral *g*-factors in the visible, which are those listed in Table 2, are successively determined from the average of individual factors computed imposing *in situ* reference water-leaving radiances as target values for the satellite-derived L_w. It is important to note that the averaging reduces the effects of random contributions to uncertainties in *g*-factors, but it does not remove the effects of any bias.

Recalling that unity *g*-factors indicate no correction, the values in Table 2 exhibit high consistency with differences generally within a few tenths of percent. The standard deviation, σ_g, gives an indication of the precision affecting the SVC process as mostly resulting from *in situ* radiometer stability or varying observation conditions. It is noted that the number of matchups used for SVC in all cases is larger than the approximate 40 estimated by Franz et al. (2007) to determine sufficiently precise *g*-factors for SeaWiFS using MOBY data. However, it is expected that such a number, implicitly referred to SeaWiFS-MOBY matchups, may change when considering observation conditions different from those offered by the MOBY site or satellite sensor performances different from those of SeaWiFS.

General elements on the various data sources utilized for the determination of the *g*-factors listed in Table 2 are summarized in the following sub-sections.

Table 1. General elelments on the various sources utilized for SVC of SeaWiFS data: measurement method, spectral features and site location (see text for additional details).

Data Source	**L_w Method**	**Spectral Features**	**Site**
MOBY	In-water, fixed depths	Hyper-spectral	Pacific Ocean (Hawaii)
MOBY-MS	In-water, fixed depths	Reduced resolution	Pacific Ocean (Hawaii)
BOUSSOLE	In-water, fixed depths	Multi-spectral	Ligurian Sea
NOMAD	Various	Various	Various
AAOT	Above-water	Multi spectral	Adriatic Sea
HOT-ORM	Modeled	User definable	Pacific Ocean (Hawaii)
BATS-ORM	Modeled	User Definable	Atlantic Ocean (Bermuda)

Table 2. Values of *g*-factors (*g*) and related standard deviations (σ_g) determined for SeaWiFS at its center-wavelengths. *N* indicates the number of matchups used for their determination, and *Y* the approximate number of measurement years.

	Wavelength		**412**		**443**		**490**		**510**		**555**		**670**	
Data Source	***Y***	***N***	***g***	σ_g	***g***	σ_g	***g***	σ_g	***g***	σ_g	***g***	σ_g	***g***	σ_g
[(1)] MOBY	7	166	1.0368	0.009	1.0132	0.009	0.9918	0.008	0.9982	0.009	0.9993	0.009	0.9729	0.007
[(1)] MOBY-MS	7	166	1.0401	0.009	1.0136	0.009	0.9949	0.008	0.9937	0.009	0.9958	0.009	0.9691	0.007
[(1)] BOUSSOLE	3	46[(4)]	1.0402	0.005	1.0129	0.027	0.9961	0.033	1.0015	0.031	1.0007	0.021	0.9672	0.006
[(1)] NOMAD	7	64	1.0395	0.013	1.0135	0.013	0.9967	0.014	0.9962	0.017	0.9989	0.013	0.9693	0.009
[(2)] AAOT	5	99	1.0425	0.012	1.0143	0.014	0.9969	0.018	0.9977	0.019	1.0034	0.022	0.9819	0.020
[(3)] HOT-ORM	7	176	1.0300	0.015	1.0086	0.012	0.9879	0.009	0.9979	0.008	1.0046	0.009	0.9718	0.006
[(3)] BATS-ORM	7	241	1.0345	0.018	1.0020	0.016	0.9814	0.013	0.9941	0.011	1.0016	0.011	0.9731	0.006

[(1)] Bailey et al. (2008), [(2)] Mélin and Zibordi (2010), [(3)] Werdell et al. (2007). [(4)] 5 matchups at 412 nm, only.

3.1 MOBY and MOBY-MS

Since 1997, MOBY has been deployed approximately 11 nautical miles from Lanai (Hawaii) in 1200 m water depth (Clark et al. 1997, Clark et al. 2002). The site was selected based on requirements for an ideal SVC location and accounting for the need to ensure economical and convenient access to shore facilities.

The main components of the MOBY system are: *i*. a spar buoy tethered to a moored buoy; and *ii*. a hyper-spectral radiometer operating in the 340-955 nm spectral region with 1 nm resolution, coupled via fiber optics to a number of radiance and irradiance collectors. These collectors ensure measurements of in-water downward irradiance and upwelling radiance at 1, 5 and 9 m depth. Above-water downward irradiance is additionally measured at 2.5 m above the sea surface. The MOBY radiometer system undergoes regular characterizations and calibrations to guarantee high accuracy and traceability of data to the US National Institute of Standards and Technology (NIST). Internal system sources allow daily monitoring of radiometric stability. By statistically combining uncertainty contributions including those related to the calibration source and its transfer, radiometric stability during deployments, and

environmental effects, Brown et al. (2007) showed the capability of reducing uncertainties to approximately 3% in the 412-666 nm spectral interval for upwelling radiance L_u used to determine L_w.

A total of 166 MOBY-SeaWiFS matchups fulfilling strict SVC criteria (Bailey and Werdell 2006, Franz et al. 2007) over a 7-year period, were applied by Bailey et al. (2008) to produce the SeaWiFS *g*-factors. Criteria for the inclusion of SeaWiFS data resulting from the average of L_w values from the 5×5 pixels centered at the MOBY site, are: no processing flag raised (e.g., indicating cloud contamination, glint perturbations, navigation problems or failure of the atmospheric correction); satellite viewing angle less than 56 degrees; sun zenith angle less than 70 degrees; *Chla* lower than 0.2 $\mu g\ l^{-1}$; τ_a in the near infrared lower than 0.15; and coefficient of variation less than 0.15 for L_{WN} in the blue-green spectral regions and τ_a in the near-infrared. It is anticipated that similar matchup selection criteria were applied to the other datasets included in this review.

The qualified matchups were constructed by convolving MOBY hyperspectral L_w data with the actual SeaWiFS spectral band responses. Bailey et al. (2008) also considered the parallel case of MOBY L_w averaged over 10 nm bandwidths with center-wavelengths corresponding to those of SeaWiFS. These *g*-factors, referred to as MOBY-Multispectral (MOBY-MS), provide the unique opportunity to look into changes only due to differences in spectral resolution. In fact the radiometer system and measurement conditions are exactly the same for both hyperspectral and derived multispectral data.

3.2 BOUSSOLE

BOUSSOLE, operated in the Ligurian Sea since 2003, has been deployed at approximately 32 nautical miles from the coast in 2440 m water depth and relies on a moored buoy optimized to maximize its vertical stability and minimize the shading effects of its superstructure (Antoine et al. 2008a). Optical instrumentation on the buoy includes 7-band commercial radiometers with 10 nm bandwidth in the 400-700 nm spectral region. In–water upwelling radiance, upward irradiance, and downward irradiance are measured with radiometers deployed at 4 and 9 m depths, while the downward irradiance is also

293 measured at 4 m above the sea surface. Spectrally independent uncertainty values of approximately 6%
294 have been declared for the normalized remote sensing reflectance determined from L_w (Antoine et al.
295 2008b). Since 2008, BOUSSOLE is also equipped with hyperspectral radiometers to measure the in–
296 water upwelling radiance and downward irradiance, and the above-water downward irradiance. Data
297 from these instruments, which are not part of this study, will be relevant for vicarious calibration
298 activities of future missions.

299 A significant difference characterizes the extrapolation methods applied to subsurface radiometric data
300 from MOBY and BOUSSOLE. While MOBY values are simply determined from the linear fit of log-
301 transformed radiometric measurements with respect to depth, BOUSSOLE sub-surface values result
302 from the propagation of the 4 and 9 m depth values to the surface through models. This latter data
303 reduction scheme, requiring estimates of *Chla*, takes into account Raman effects and the related
304 nonlinearity of the log-transformed radiometric measurements with depth. Differences between the linear
305 fits of log-transformed radiometric measurements and modelled values, are within a few percent at 412
306 nm but increase up to several tens percent at 670 nm (Antoine et al. 2008b).

307 BOUSSOLE data were also considered by Bailey et al. (2008) for the determination of SeaWiFS *g*-
308 factors. Specifically, 46 matchups were identified from approximately a 3-year data record by relaxing
309 the inclusion criteria on *Chla* (0.25 instead of 0.20 $\mu g\ l^{-1}$). However, only 5 matchups were available for
310 the 412 nm center-wavelength due to unavailability of the spectral band during some deployments.

311

312 ***3.3 NOMAD***

313 The NASA bio-Optical Algorithm Data set (NOMAD, Werdell and Bailey 2005) includes multi-site
314 and multi-source data resulting from the reprocessing and strict quality control of radiometric
315 measurements from the SeaWiFS Bio-Optical Archive and Storage System (SeaBASS). The variety of
316 measurement methods, instruments, calibration and also data reduction schemes, make it difficult to
317 assign well-defined uncertainties to the NOMAD radiometric data set.

318 The SeaWiFS *g*-factors determined from NOMAD (Bailey et al. 2008) were computed using 64
319 matchups fulfilling SVC selection criteria –out of a total of 1039. These field radiometry data result from
320 overall 3475 quality controlled measurements out of 15400 from 1350 field campaigns included in
321 SeaBASS. These numbers clearly indicate the difficulty of supporting SVC with *in situ* L_w data from
322 repositories constructed for applications more focused on validation and development rather than
323 vicarious calibration.
324
325 ***3.4 AAOT***
326 In contrast with MOBY and BOUSSOLE data, which are collected with systems specifically designed
327 to support SVC, time-series data from a number of globally distributed coastal sites established to
328 support satellite ocean color validation activities are accessible through the Ocean Color component of
329 the Aerosol Robotic Network (AERONET-OC, Zibordi et al. 2009). AERONET-OC field radiometers
330 perform multispectral L_w measurements at a number of ocean color bands with center-wavelengths in the
331 410-1020 nm spectral region and 10 nm bandwidth. Data collection, reduction and quality control rely
332 on standardized methods (Zibordi et al. 2009) assuring cross-site consistency to data products. Among
333 AERONET-OC sites, the Acqua Alta Oceanographic Tower (AAOT, often indicated as 'Venise'),
334 located in the northern Adriatic Sea at approximately 8 nautical miles from the main land, since 2003 has
335 provided an almost uninterrupted series of data largely applied for the validation of multi-mission ocean
336 color radiometric data (e.g., Zibordi et al. 2006, Mélin et al. 2011, Zibordi et al. 2012). Uncertainties of
337 5% in the blue-green and 8% in the red spectral regions have been quantified for the AAOT fully quality
338 assured normalized water-leaving radiance determined from L_w (Gergely and Zibordi 2014).
339 AERONET-OC data from the AAOT were used by Mélin and Zibordi (2010) for the determination of
340 regional SeaWiFS *g*-factors. Specifically, 99 qualified matchups were identified from a 5-year data set
341 by relaxing some selection criteria (e.g., accepting *Chla* up to 3 $\mu g\ l^{-1}$ and coefficient of variation up to
342 0.20 in the blue-green spectral region for satellite data). A particular effort was devoted to correct *in situ*
343 L_w spectra for the effects of differences in center-wavelengths with respect to SeaWiFS bands.

Results from the study give insight on the relevance of coastal vicarious calibration sites for regional investigations and additionally provide elements to evaluate their suitability for global applications. Still, the spatial and inter-annual variability of both atmospheric and water optical properties in the region do not support the selection of the AAOT as a SVC site for the creation of CDRs.

3.5 HOT-ORM and BATS-ORM

Ocean Reflectance Models (ORM) are an additional source of radiance spectra (Morel and Maritorena, 2001) expected to be of suitable accuracy for oceanic waters. Even though these models are mostly relevant for bio-optical investigations or as diagnostic tools, their usefulness for SVC has been investigated by Werdell et al. (2007) to verify their fitness for historical satellite ocean color sensors (i.e., CZCS and OCTS) for which an extensive time-series of *in situ* radiometric measurements do not exist.

The SeaWiFS *g*-factors determined using ORM methodology include those relying on the *Chla* time-series from the U.S. Joint Global Ocean Flux Study (JGOFS) Bermuda Atlantic Time-series Study (BATS) and Hawaiian Ocean Time-series (HOT). Specifically, ORM-BATS *g*-factors were determined using 241 matchups from 1998 to 2004, while ORM-HOT *g*-factors were computed for the same period with 176 matchups (Werdell et al. 2007). Comprehensive uncertainty estimates for modeled L_w were not provided.

4. Analysis and discussion

It shall be noted that the *g*-factors in Table 2 were determined with an earlier version of the SeaDAS processor (i.e., version 5) based on an atmospheric model and pre-launch absolute calibration factors (specifically at 412 nm) different from those currently in use. Because of this, the *g*-factors in Table 2 need to be considered outdated for present SeaWiFS data processing. Still, they are the result of a unique combination of investigations and remain a convenient data set to explore effects of differences among *g*-factors in the creation of CDRs. Making use of these data, the following analysis focuses on percent

differences between g-factors determined from MOBY data, g^{MOBY}, and those from other data sources, g, computed as

$$\Delta g = 100\ \frac{g - g^{MOBY}}{g^{MOBY}} \quad (4)$$

The choice of the g-factors from MOBY as the reference is justified by its ideal location (exhibiting oligotrophic waters and maritime aerosol, in addition to annual cycles of small amplitude) and an extensive characterization of field radiometers and careful examination of radiometric uncertainties. This choice, however, has not to be interpreted as implicitly advocating the use of MOBY for SVC of any satellite ocean color mission.

For completeness it is also mentioned that the HOT-ORM and BATS-ORM g-factors included in Table 2, were discussed by Werdell et al. (2007) with respect to the older MOBY g-factors determined by Franz et al. (2007) on the basis of 150 match-ups. Those g-factors exhibit spectrally averaged differences of -0.09% with respect to the more recent values by Bailey et al. (2008) used in the current analysis. Still, the changes in the values of Δg for HOT-ORM and BATS-ORM resulting from the application of the g-factors from Franz et al. (2007) instead of those from Bailey et al. (2008), does not affect the following discussion and conclusions.

Table 3. Relative percent differences Δg between SeaWiFS g-factors determined using Eq. 4 applied to data in Table 2. The values in bold indicate Δg exceeding ±0.3% in the blue-green spectral regions and ±0.1% in the red.

Wavelength	**412**	**443**	**490**	**510**	**555**	**670**
Data Source	Δg [%]	Δg [%]	Δg [%]	Δg [%]	Δg [%]	Δg [%]
MOBY-MS	**+0.32**	+0.04	**+0.31**	**-0.45**	**-0.35**	**-0.39**
BOUSSOLE	**+0.33**	-0.03	**+0.43**	**+0.33**	+0.14	**-0.59**
NOMAD	+0.26	+0.03	**+0.49**	-0.20	-0.04	**-0.37**
AAOT	**+0.55**	+0.11	**+0.51**	-0.05	**+0.41**	**+0.93**
HOT-ORM	**-0.66**	**-0.45**	**-0.39**	-0.03	**+0.53**	**-0.11**
BATS-ORM	-0.22	**-1.11**	**-1.05**	**-0.41**	+0.23	+0.02

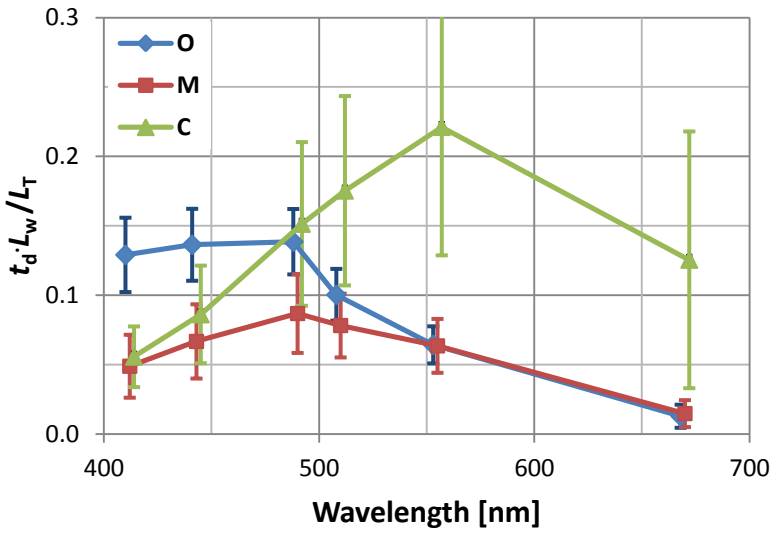
The Δg values in Table 3 from the same data source (i.e., inter-band) or across data sources (i.e., intra-band) are generally lower than ±0.5%.

At a first scrutiny, the values of Δg determined for the AAOT and HOT-ORM appear to slightly differ from those determined for a more ideal site like BOUSSOLE or from a very large pool of data like NOMAD. Also interesting are the values of Δg determined for MOBY-MS, which clearly indicate the appreciable effects of non-matching spectral bands or SeaWiFS out-of-band responses, and consequently the importance of *in situ* hyperspectral L_w data.

Excluding HOT-ORM and BATS-ORM, the values of Δg exhibit high intra-band consistency between 412 and 490 nm, while they show a larger spread between 510 and 670 nm. Excluding a few spectral values from HOT-ORM (i.e., 412 nm), BATS-ORM (i.e, 443 and 490 nm) and AAOT (i.e., 670 nm), Δg is generally lower than ±0.5% for all the data sources.

In view of more quantitatively investigating differences in g-factors, Eq. 3 is applied to compute $u(L_T)/L_T$ as a function of $u(L_w)/L_w$ accounting for actual mean spectral values of the term $t_d \cdot L_w/L_T$ determined using 1997-2010 SeaWiFS data for three different locations: the MOBY site in the Pacific Ocean with mean satellite-derived *Chla* of 0.08±0.02 μg l^{-1} representing oligotrophic waters (O); the BOUSSOLE site in the Ligurian Sea with mean *Chla* of 0.36±0.37 μg l^{-1} representing mesotrophic waters (M); and the AAOT coastal site in the northern Adriatic Sea with mean *Chla* of 1.74±1.40 μg l^{-1} representing coastal waters moderately dominated by sediments (C). When considering all three water types (see Fig. 1), $t_d \cdot L_w/L_T$ exhibits a large range of mean values spanning from approximately 0.07-0.14 at 490 nm, 0.06-0.22 at 555 nm and 0.01-0.12 at 670 nm. These differences are mostly due to site dependent changes in L_w and L_A, both contributing to L_T (see Eq. 1).

As already stated in §2.1, the following analysis assumes the uncertainties related to the atmospheric correction process do not affect the determination of satellite-derived L_w because of the use of the same atmospheric models and algorithms for SVC and for atmospheric correction.



414

415 Figure 1. Spectral values of $t_d \cdot L_w/L_T$ for oligotrophic (O), mesotrophic (M) and coastal (C) waters.
416 Mean values and standard deviations σ (indicated by the vertical error bars), result from the analysis of
417 814, 1487 and 1045 SeaWiFS data extractions, respectively. The center-wavelengths between spectra
418 have been shifted by ±2 nm to increase readability.

419

420 Fig. 2 summarizes results from the application of Eq. 3 using identical spectrally independent relative
421 uncertainties for *in situ* L_w (i.e., 5%). The derived values of $u(L_T)/L_T$ exhibit a significant spectral
422 dependence and, as expected, are smaller when $t_d \cdot L_w/L_T$ is smaller (i.e., in correspondence with the
423 lower values of L_w). Specifically, the lowest $u(L_T)/L_T$ are observed for mesotrophic waters with values
424 included in the range of approximately 0.2-0.5% in the blue-green spectral regions, and dropping below
425 0.1% at 670 nm. The values observed for the oligotrophic waters are higher in the blue spectral region
426 with values approaching 0.7%. In agreement with the higher values of L_w, $u(L_T)/L_T$ computed for the
427 coastal waters reach 1.1% at 555 nm and 0.6% at 670 nm. It is mentioned that differences in the
428 observation conditions at the various sites or in the spectral values of $u(L_w)/L_w$, may lead to $u(L_T)/L_T$
429 different from those presented in Fig. 2. Additionally, the relative combined uncertainty value of L_T
430 determined from a number N of *in situ* L_w data obtained with equivalent observation conditions would

decrease with respect to the value of $u(L_T)/L_T$ from an individual L_w due to the statistical averaging of the random component of uncertainties.

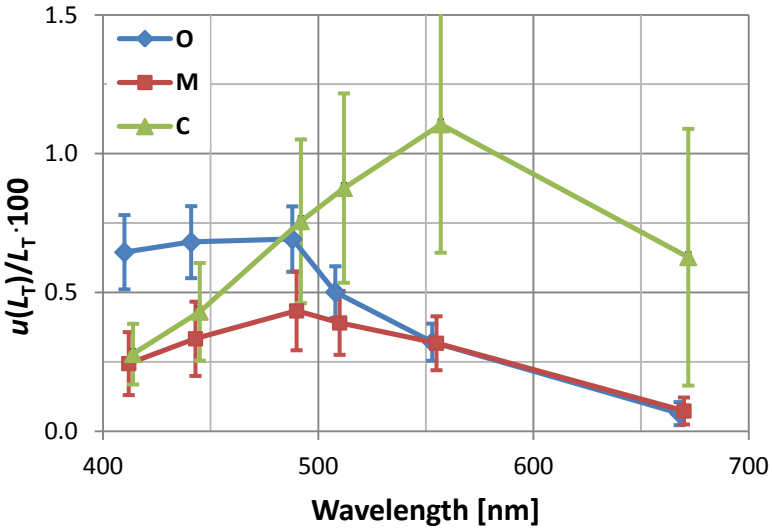


Figure 2. Relative uncertainties $u(L_T)/L_T$ determined assuming a spectrally independent 5% uncertainty value for L_w with the mean values of $t_d \cdot L_w/L_T$ given in Fig. 1 for different water types: oligotrophic (O), mesotrophic (M) and coastal (C). The vertical bars refer to values determined with $t_d \cdot L_w/L_T \pm \sigma$.

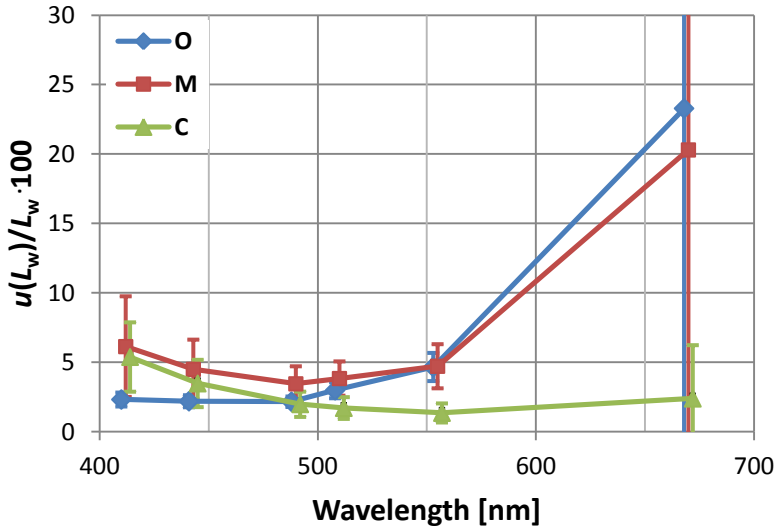


Figure 3. Relative uncertainties $u(L_w)/L_w$ determined assuming a spectrally independent 0.3% uncertainty value for L_T and the mean values of $t_d \cdot L_w/L_T$ given in Fig. 1 for different water types: oligotrophic (O), mesotrophic (M) and coastal (C). The vertical bars refer to values determined with $t_d \cdot L_w/L_T \pm \sigma$.

Rearranging Eq. 3, relative uncertainties in satellite-derived $L_{w,}$ can be investigated as a function of $u(L_T)/L_T$. By assigning a spectrally independent value of 0.3% to $u(L_T)/L_T$ (i.e., a value that occurs often for $|\Delta g|$ in Table 3), results displayed in Fig. 3 indicate that the 5% uncertainty requirement in satellite-derived L_w generally cannot be met in the red for oligotrophic and mesotrophic waters, and is challenging in the blue mostly at 412 nm for mesotrophic and coastal waters. Because of this, the 0.3% value assigned to $u(L_T)/L_T$, could be considered a rough upper threshold for the uncertainties of *g*-factors allowing to meet the 5% science requirement for $u(L_w)/L_w$ in the blue-green spectral regions. The same $u(L_w)/L_w$ values displayed in Fig. 3 also indicate that the application to different missions of *g*-factors determined with independent *in situ* data sources and exhibiting typical differences of 0.3% in the blue-green spectral regions with respect to the values obtained with an identical *in situ* data source, may introduce mission dependent biases of several percent in multi-mission CDRs. These biases would hinder stability requirements in satellite-derived products even when applying the same atmospheric correction code to the processing of data from different missions. This result is confirmed by practical assessments presented in Werdell et al (2007) showing that for deep waters Δg~0.3% may lead to biases of 4% in L_w at 555 nm.

In addition, the spectral differences affecting the values of Δg from the same data source or across data sources (see Table 3), may lead to significant spectral inconsistencies in CDRs. These inconsistencies (i.e., substantial inter-band spectral changes of Δg) would affect the capability of meeting the 1% inter-band uncertainty for L_w included in some mission objectives and likely the 3% stability requirement for an ECV like *Chla* (WMO 2011), which is commonly derived from spectral ratios of L_w.

A statistical index that can be of interest to discuss stability requirements for the construction of CDRs from different satellite missions, is the relative standard error of the mean (*RSEM*) of g-factors g determined from

$$RSEM = (\sigma_g/\mathrm{g})/\sqrt{N_y} \quad (5)$$

with σ_g standard deviation of g assumed invariant with time for each considered data source, and N_y the scaled number of match-ups per decade (i.e., N_y=10·N/Y where N is the number of actual matchups and Y the number of measurement years).

The scaling of the number of matchups over a decade, that forces the assumption of continuous availability of measurements for each *in situ* data source during the considered period, is only applied to facilitate the comparability of *RSEM* values for data which were available for a limited number of years at the time of this study. Nevertheless, continuous operation and delivery of measurements are required for any *in situ* SVC data source contributing to the creation of CDRs,

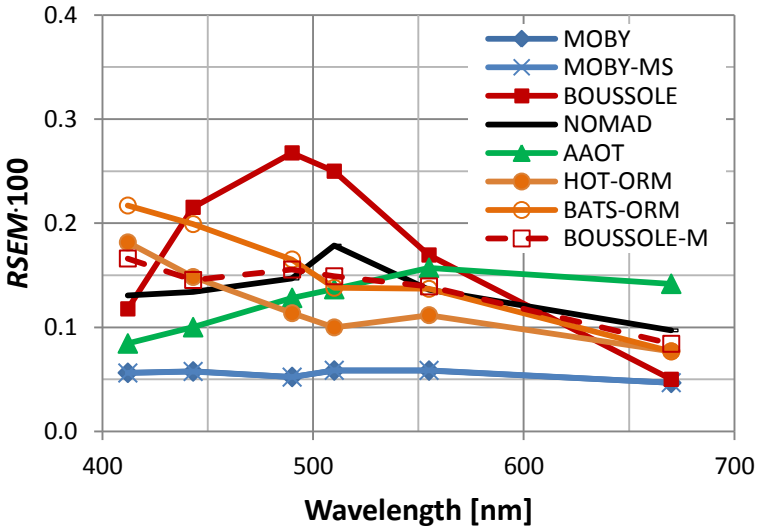


Figure 4. Plot of the standard percent error of the mean (*RSEM*) for the SeaWiFS g-factor given in Table 2 and additionally for MERIS g-factors determined with BOUSSOLE data (i.e., BOUSSOLE-M).

In view of supporting such a discussion on stability requirements through actual numbers, Fig. 4 displays *RSEM* values computed using the data in Table 2.

The notably low values of *RSEM* determined with the MOBY and MOBY-MS data suggest high measurement precision likely explained by very stable measurement conditions, systematic calibration and characterization of field radiometers, robust quality assessment of field measurements and quality control of data products. The higher *RSEM* values resulting from the other data sources are likely explained by a number of factors including (but not restricted to): *i*. measurement conditions perturbed by time-dependent changes in the marine and atmospheric optical properties or observation geometry; *ii*. instability of the *in situ* measurement system when challenged by environmental perturbations during deployments (e.g., bio-fouling) or by variable performance of radiometer systems operated during successive deployments, or even by different measurement methods when considering a combined data set; *iii* or a relatively small of number of matchups N_y per decade.

The large *RSEM* values determined for BOUSSOLE, which refer to field radiometric measurements performed during the early deployment phase of the buoy system, are due to large σ_g and a relatively small number of matchups. Successive improvements in quality assurance and control of the radiometric measurements have led to a great reduction of σ_g. This is shown by the BOUSSOLE-M *RSEM* values also displayed in Fig. 4, and computed applying recent σ_g of *g*-factors determined for the Medium Resolution Imaging Spectrometer (MERIS). These updated values of σ_g, which refer to a 7-year measurement period, vary between 0.006 and 0.012 with *N* ranging from 15 to 42.

Overall, the previous findings suggest that any element affecting reproducibility of measurements and observation conditions with time, and thus challenging the precision of *in situ* reference measurements, should be minimized to lessen perturbations affecting the random component of uncertainties for *g*-factors and thus the stability requirement for CDRs resulting from the combination of multi-mission satellite-derived data. In addition, frequent swaps of radiometer systems exhibiting similar measurement uncertainties should be considered an important best practice. In fact, the measurement uncertainties

would average over the number of deployments occurring during each satellite mission. This is expected to increase the probability of achieving equivalent precision for *g*-factors applicable to the processing of satellite data from independent missions.

To conclude, the 0.5% stability requirement over a decade (WMO 2011) entails maximum uncertainties of approximately 0.05, 0.025 and 0.005% in *g*-factors, assuming generic values of 0.10, 0.05 and 0.01 for the term $t_d \cdot L_w / L_T$. These uncertainties are comparable to the *RSEM* values determined for MOBY in the blue-green spectral regions over a period of approximately 10 years, while they are significantly lower than those determined from the other *in situ* data sources (see Fig. 4). This result further indicates: *i*. the need for long-term highly consistent *in situ* data applicable to SVC in view of minimizing any appreciable perturbation that may affect the determination of *g*-factors over time for different or successive satellite missions; and *ii*. caution in using data from sole or multiple sources, which may refer to measurement conditions difficult to reproduce for different missions.

Additionally, the application of mission-independent atmospheric models and algorithms for the atmospheric correction process is critical.

5. Summary and Recommendations

SVC does not literally lead to the absolute radiometric calibration of the satellite sensor. Rather, assuming equivalent observation conditions characterizing both SVC and atmospheric correction processes, SVC forces the determination of satellite-derived L_w with an uncertainty comparable to that of the *in situ* reference L_w applied for the indirect calibration process. This is achieved through vicarious adjustment gain-factors (i.e., *g*-factors), which are applied to the top of atmosphere radiances L_T after full instrument calibration (e.g., following pre-launch absolute calibration and characterization, and additionally, corrections for temporal changes in radiometric sensitivity as determined through the sensor-specific on-orbit calibration system).

The investigation presented in this work highlights that the relative uncertainty that may affect *g*-factors, to a first approximation depends on the term $t_d \cdot L_w/L_T$ and on the uncertainties affecting *in situ* L_w data. This finding and differences among *g*-factors determined for the SeaWiFS spectral bands using various data sources, but relying on the same atmospheric models and atmospheric correction algorithms, provide suggestions on the suitability of *in situ* L_w data sources for SVC devoted to support the construction of CDRs. Specifically, when considering the blue and green center-wavelengths commonly applied for the determination of *Chla*, satellite-derived L_w resulting from the application of *g*-factors differing by as little as 0.3% can result in spectral biases close to 5%. These biases are several times higher than the 0.5% target stability value per decade indicated for satellite ocean color data products expected to contribute to CDRs. Thus, in view of avoiding inconsistencies in long-term data records resulting from the combination of satellite products from multiple missions, a careful evaluation of sites and *in situ* measurements supporting SVC is needed. In particular, the determination of *g*-factors by combining match-ups from multiple sites, which is often a viable solution to shorten the otherwise long time needed to accumulate a relatively large number of matchups satisfying early mission needs or mission-specific objectives, has to be regarded as a potential source of artifacts for CDRs. In fact, even assuming equivalent uncertainties for *in situ* data from different sources and a single atmospheric correction code, differences in *g*-factors may result from a diverse performance of the atmospheric correction process at different sites due to differences in satellite observing geometries or marine and atmospheric optical properties. Further, differences in the performance of various *in situ* radiometer systems may also affect the accuracy and precision of *g*-factors through those of the *in situ* L_w data and thus also affect the stability requirements of CDRs.

In view of defining strategies for the upcoming satellite ocean color missions, the previous findings and considerations suggest that the creation of ocean color CDRs should ideally rely on: *i*. one main long-term *in situ* calibration system (site and radiometry) established and sustained with the objective to maximize accuracy and precision over time of *g*-factors and thus minimize possible biases among satellite data products from different missions; and *ii*. unique (i.e., standardized) atmospheric models and

557 algorithms for atmospheric corrections to maximize cross-mission consistency of data products at
558 locations different from that supporting SVC.

559 Accounting for results from this study and any element already provided in literature, it is expected
560 that an ideal ocean color SVC site should meet the following general requirements:

561 1. Located in a region chosen to maximize the number of high-quality matchups by trading off factors
562 such as best viewing geometry, sun-glint avoidance, low cloudiness, and additionally set away from
563 any continental contamination and at a distance from the mainland to safely exclude any adjacency
564 effect in satellite data;
565 2. Exhibiting known or accurately modeled optical properties coinciding with maritime atmosphere and
566 oligotrophic/mesotrophic waters, to represent the majority of world oceans and minimize relative
567 uncertainties in computed *g*-factors;
568 3. Characterized by high spatial homogeneity and small environmental variability, of both atmosphere
569 and ocean, to increase precision of computed *g*-factors.

570 Any field radiometer system supporting SVC should rely on advanced *in situ* measurement
571 technologies, data reduction methods and quality assurance/control schemes to minimize relative
572 standard uncertainties in *in situ* L_w to within state-of-the-art values. In particular, uncertainty target
573 values should be 3-4% in the blue-green spectral regions and, even though not relevant for GCOS,
574 tentatively below 5% in the red, with inter-band uncertainties lower than 1% . In particular, accounting
575 for findings from this study and from literature and without advocating the adoption of any existing SVC
576 radiometry system, the fulfillment of the following wide-range requirements for *in situ* radiometric
577 measurements should be considered of utmost importance:

578 i. Hyperspectral field data with sub-nanometer resolution to allow system vicarious calibration of any
579 satellite ocean color sensor regardless of its center-wavelengths and spectral responses, and thus
580 ensure minimization of inter-band uncertainties;

ii. State-of-the-art absolute calibration traceable to National Metrology Institutes (i.e., tentatively with target standard calibration uncertainty lower than 2% for radiance and stability better than 0.5% per deployment) and comprehensive characterizations of radiometers in terms of linearity, temperature dependence, polarization sensitivity and stray light effects, in view of minimizing measurement uncertainties and allowing for accurate determinations of uncertainty budgets;

iii. Application of quality assurance/control schemes minimizing effects of measurement perturbations like those (when applicable) due to infrastructure shading, radiometer self-shading, wave perturbations, bio-fouling, and additionally scheduling regular checks of *in situ* systems and frequent swap of radiometers, as best practice to maximize long-term accuracy and precision of *in situ* reference radiometric data;

iv. Data rate ensuring generation of matchups for any satellite ocean color mission with time differences appropriate to minimize variations in bi-directional effects due to changes in sun zenith and daily fluctuations in the vertical distribution of phytoplankton.

In addition to requirements for establishing an ideal SVC site and generating *in situ* radiometric data with the needed accuracy and precision, the supplementary capability of continuously characterizing both the atmospheric (e.g., τ_a) and water (e.g., inherent) optical properties would provide additional important elements for the quality assurance of matchups applicable to determine *g*-factors.

It is reminded that strategies for the construction of CDRs also suggest establishing and maintaining secondary *in situ* long-term systems with performance equivalent to the main one in terms of data accuracy, precision and measurement conditions. This recommendation is enforced by the fundamental need to allow for redundancy ensuring fault-tolerance to SVC and additionally to provide optimal means for continuous verification and validation of satellite primary data products including the capability to accurately investigate systematic effects induced by different observation conditions (i.e., viewing and illumination geometry, atmosphere and water types).

It is finally mentioned that the need to standardize the atmospheric correction process for multi-mission data contributing to CDRs is a requirement as relevant as the availability of *in situ* data from one

607 ideal SVC site. This operational need, however, should not be seen as an impediment to further advance
608 atmospheric models and atmospheric correction algorithms.
609
610
611 **Acknowledgments**

612 The authors would like to express appreciation to the International Ocean Color Coordination Group
613 (IOCCG) for promoting a number of initiatives stimulating discussions on satellite ocean color system
614 vicarious calibration which encouraged this study.

615 This work is dedicated to the memory of Dennis Clark whose commitment to Ocean Color led to the
616 design, implementation and long-term operation of MOBY.

617

618 The contents of this paper reflect the views of the authors and should not be interpreted as an official
619 statement of policy, decision, or position on behalf of any of the organizations mentioned.

6. References

Antoine, D., Guevel, P. , Deste, J. F. , Bécu, G. Louis, F., Scott, A. J. & Bardey, P. (2008a). The "BOUSSOLE" buoy-a new transparent-to-swell taut mooring dedicated to marine optics: design, tests, and performance at sea. *J. Atmos. Oceanic Technol.*, 25**,** 968–989 (2008a).

Antoine, D., D'Ortenzio, F., Hooker, S. B., Bécu, G., Gentili, B., Tailliez, D. & Scott, A. J. (2008b). 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). *J. Geophys. Res.*, 113(C7), C07013.

Attila, J., Koponen, S., Kallio, K., Lindfors, A., Kaitala, S. & Ylöstalo, P. (2013). MERIS Case II water processor comparison on coastal sites of the northern Baltic Sea. *Remote Sens. Environ.*, 128, 138–149.

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

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

Behrenfeld, M. J., O'Malley, R. T., Siegel, D. A., McClain, C. R., Sarmiento, J. L., Feldman, G. C. Milligan, A. J., Falkowski, P. G., Letelier, R. M. & Boss, E. S. (2006). Climate-driven trends in contemporary ocean productivity. *Nature*, 444.7120: 752–755.

Brown, S. W., Flora, S. J., Feinholz, M. E., Yarbrough, M. A., Houlihan, T., Peters, D., Kim, K. Y. S., Mueller, J. L., Johnson, B. C. & Clark, D. K. (2007). The Marine Optical BuoY (MOBY) radiometric calibration and uncertainty budget for ocean color satellite sensor vicarious calibration. In SPIE Conference Proceedings *Remote Sensing*, pp. 67441M–67441M. International Society for Optics and Photonics.

Butler, J.J., Johnson, B. C., Rice, J. P., Brown, S. W. & Barnes, R. A. (2007). Validation of radiometric standards for the laboratory calibration of reflected-solar Earth-observing satellite instruments. In SPIE Conference Proceedings *Earth Observing Systems*, pp. 667707-667707. International Society for Optics and Photonics.

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

Clark, D. K., Feinholz, M. E., Yarbrough, M. A., Johnson, B. C., Brown, S. W., Kim, Y. S. & Barnes, R. A. (2002). Overview of the radiometric calibration of MOBY. In SPIE Proceedings *Earth Observing Systems* VI, 4483, 64–76.

Clark, D. K., Yarbrough, M. A., Feinholz, M., Flora, S., Broenkow, W., Kim, Y. S., Johnson, B. C., Brown, S. W., Yuen, M. & Mueller, J. L. (2003). A Radiometric Buoy for Performance Monitoring and Vicarious Calibration of Satellite Ocean Color Sensors: Measurement and Data Analysis Protocols," In *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation*. J.L. Mueller, G. S. Fragion and V. R. Mc Clain (Eds.). NASA/TM-2003-211621/Rev4-Vol. IV, Greenbelt, MD.

668 Cramer, C. E., Lykke, K. R., Woodward, J. T. & Smith, A. W. (2013). Precise Measurement of Lunar
669 Spectral Irradiance at Visible Wavelengths," *J. Res. Natl. Inst. Stan.*, 118, 396–402.
670
671 Eplee, R. E., Robinson, W. D., Bailey, S. W., Clark, D. K., Werdell, P. J., Wang, M., Barnes R. A. &
672 McClain, C. R. (2001). Calibration of SeaWiFS. II. vicarious techniques. *Appl. Opt.*, 40, 6701–6718.
673
674 Eplee Jr, R. E., Sun, J.-Q., Meister, G., Patt, F. S., Xiong, X. & McClain, C. R. (2011). Cross calibration
675 of SeaWiFS and MODIS using on-orbit observations of the Moon. *Appl. Opt.* 50, 120-133.
676
677 Esposito, J. A., Xiong, X., Wu, A., Sun, J. & Barnes, W. L. (2004). MODIS reflective solar bands
678 uncertainty analysis. In SPIE Conference Proceedings *Earth Observing Systems*, pp. 448-458.
679 International Society for Optics and Photonics.
680
681 Franz, B. A., Bailey, S. W., Werdell, P. J. & McClain, C. R. (2007). Sensor-independent approach to the
682 vicarious calibration of satellite ocean color radiometry. *Appl. Opt.*, 46, 5068–5082.
683
684 Fu, G., Scheiber, B. D., Settle, K. J., Darzi, M., McClain, C. R. & Arrigo, K. R. (1996). SeaDAS: A
685 processing package for ocean color satellite imagery. In Proceedings *Twelfth Int. Conf. on Interactive*
686 *Inform. and Processing Systems for Meteorol., Oceanogr., and Hydrol*, pp. 451-456.
687
688 Gergely, M. & Zibordi, G. (2014). Assessment of AERONET L_{WN} uncertainties. *Metrologia* 51, 40–47.
689
690 Gordon, H. R. (1987). Calibration requirements and methodology for remote sensors viewing the ocean
691 in the visible. *Remote Sens. Environ.* 22, 103–126.
692

693 Gordon, H. R. (1998). In orbit calibration strategy for ocean color sensors. *Remote Sens. Environ*. 63,
694 265– 278.
695
696 Gordon, H. R. & Clark, D. K. (1981). Clear water radiances for atmospheric correction of coastal zone
697 color scanner imagery," *Appl. Opt.* 20, 4175–4180.
698
699 Gordon, H. R., Clark, D. K., Brown, J. W., Brown, O. B., Evans, R. H. & Broenkow, W. W. (1983).
700 Phytoplankton pigment concentrations in the Middle Atlantic Bight: Comparison of ship
701 determinations and CZCS estimates. *Appl. Opt.* 22, 20–36 (1983).
702
703 Hooker, S. B., Esaias, W. E., Feldman, G. C., Gregg, W. W., McClain, C. R. (1982). An overview of
704 SeaWiFS and ocean color. *NASA Tech. Memo.* 1992–104566, vol. 1, S.B. Hooker and E.R. Firestone
705 (Eds.), NASA Goddard Space Flight Center, Greenbelt, MD.
706
707 IOCCG (2010). *Atmospheric Correction for Remotely-Sensed Ocean-Colour Products.* Wang, M. (Ed.),
708 Reports of the International Ocean-Colour Coordinating Group, No.10, pp. 78, IOCCG, Dartmouth,
709 Canada.
710
711 IOCCG (2013). *In-flight Calibration of Satellite Ocean-Colour Sensors*. Frouin, R. *(*Ed.*)*, Reports of the
712 International Ocean-Colour Coordinating Group, No.14, pp. 106, IOCCG, Dartmouth, Canada.
713
714 Joint Committee for Guides in Metrology (JCGM), (2008). *Evaluation of Measurement Data—Guide to*
715 *the Expression of Uncertainty in Measurement*, JCGM 100:2008.
716
717 Koepke, P. (1982). Vicarious satellite calibration in the solar spectral range by means of calculated
718 radiances and its application to Meteosat. *Appl. Opt*. 21, 2845–2854.

719

720 Leroy, S. S., Anderson, J. G. & Ohring, G. (2008). Climate signal times and constrains on climate
721 benchmark accuracy requirements. *J. Climate* 21, 841–846.

722

723 Levick, A. P., Greenwell, C. L., Ireland, J. Woolliams, E. R., Goodman, T. M., Bialek, A., & Fox N. P.
724 (2014). Spectral radiance source based on supercontinuum laser and wavelength tunable bandpass
725 filter: the spectrally tunable absolute irradiance and radiance source. *Appl. Opt.* 53, 3508-3519.

726

727 Mélin, F. & Zibordi, G. (2010). Vicarious calibration of satellite ocean color sensors at two coastal
728 sites." *Appl. Opt.* 49, 798–810.

729

730 Mélin, F., Zibordi, G., Berthon, J.-F., Bailey, S., Franz, B., Voss, K. J., Flora, S. & Grant, M. (2011).
731 Assessment of MERIS reflectance data as processed with SeaDAS over the European Seas. *Opt.*
732 *Express*, 19, 25657–25671.

733

734 Morel, A. & Maritorena, S. (2001). Bio-optical properties of oceanic waters: A reappraisal. *J. Geophys.*
735 *Res.*, 106 (C4) 7163–7180.

736

737 Ohring, G., Wielicki, B., Spencer, R., Emery, B. & Datla, R. (2005). Satellite instrument calibration for
738 measuring global climate change: Report of a workshop. *B. Am. Meteorol. Soc*. 86, 1303–1313.

739

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

Werdell, P. J., Bailey, S. W., Franz, B. A., Morel, A. & McClain, C. R. (2007). On-orbit vicarious calibration of ocean color sensors using an ocean surface reflectance model. *Appl. Opt.* 46, 5649-5666 (2007).

World Meteorological Organization (WMO), (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 (2010 Update)," World Meteorological Organization, Report GCOS – 154.

Zibordi, G., Mélin, F. & Berthon, J.-F. (2006). Comparison of SeaWiFS, MODIS, and MERIS radiometric products at a coastal site. *Geophys. Res. Lett.* L06617, doi:10.1029/2006GL025778.

Zibordi, G., Holben, B. N., Slutsker, I., Giles, D., D'Alimonte, D., Mélin, F., Berthon, J.-F., Vandemark, D., Feng, H., Schuster, G., Fabbri, B. E., Kaitala, S. & Seppälä, J. (2009). AERONET-OC: a network for the validation of ocean color primary radiometric products. *J. Atmos. Ocean. Technol*. 26, 1634–1651.

Zibordi, G. & Voss, K. J. (2014). In situ optical radiometry in the visible and near infrared." In *Optical Radiometry for Ocean Climate Measurement*, G. Zibordi, C. Donlon and A. Parr (Eds.), Academic Press.

Captions

Figure 1. Spectral values of $t_d \cdot L_w/L_T$ for oligotrophic (O), mesotrophic (M) and coastal (C) waters. Mean values and standard deviations σ (indicated by the vertical error bars), result from the analysis of 814, 1487 and 1045 SeaWiFS data extractions, respectively. The center-wavelengths between spectra have been shifted by ±2 nm to increase readability.

Figure 2. Relative uncertainties $u(L_T)/L_T$ determined assuming a spectrally independent 5% uncertainty value for L_w with the mean values of $t_d \cdot L_w/L_T$ given in Fig. 1 for different water types: oligotrophic (O), mesotrophic (M) and coastal (C). The vertical bars refer to values determined with $t_d \cdot L_w/L_T \pm \sigma$.

Figure 3. Relative uncertainties $u(L_w)/L_w$ determined assuming a spectrally independent 0.3% uncertainty value for L_T and the mean values of $t_d \cdot L_w/L_T$ given in Fig. 1 for different water types: oligotrophic (O), mesotrophic (M) and coastal (C). The vertical bars refer to values determined with $t_d \cdot L_w/L_T \pm \sigma$.

Figure 4. Plot of the standard percent error of the mean (*RSEM*) for the SeaWiFS *g*-factor given in Table 2 and additionally for MERIS *g*-factors determined with BOUSSOLE data (i.e., BOUSSOLE-M).

Figure 1

Click here to download Figure: Figure 1.eps

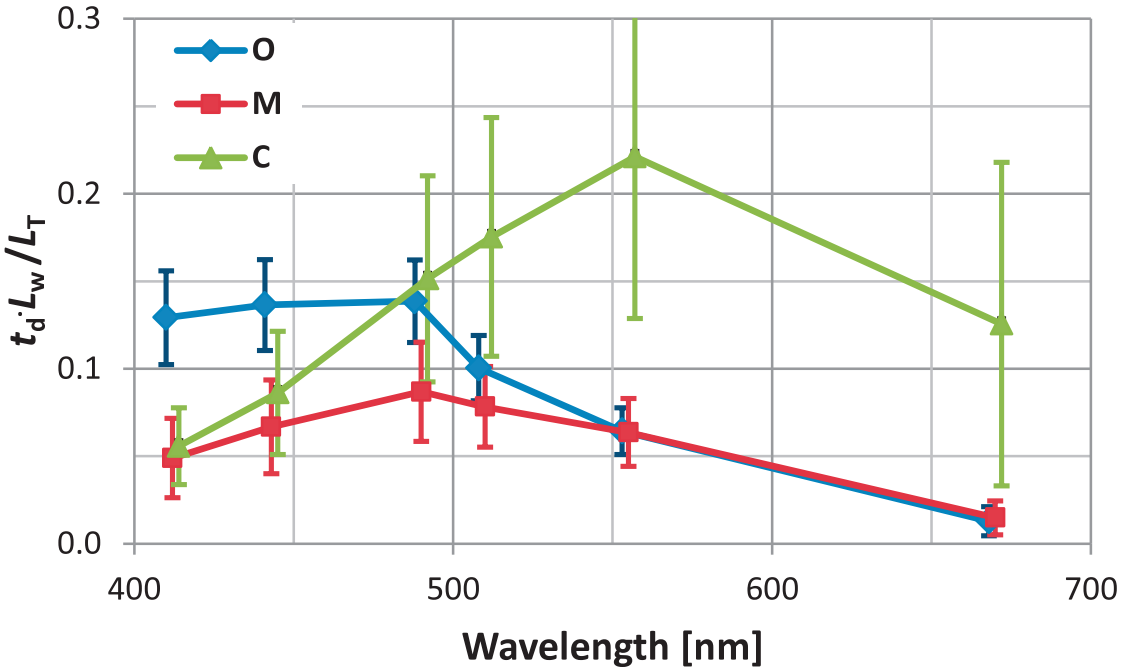


Figure 2

Click here to download Figure: Figure 2.eps

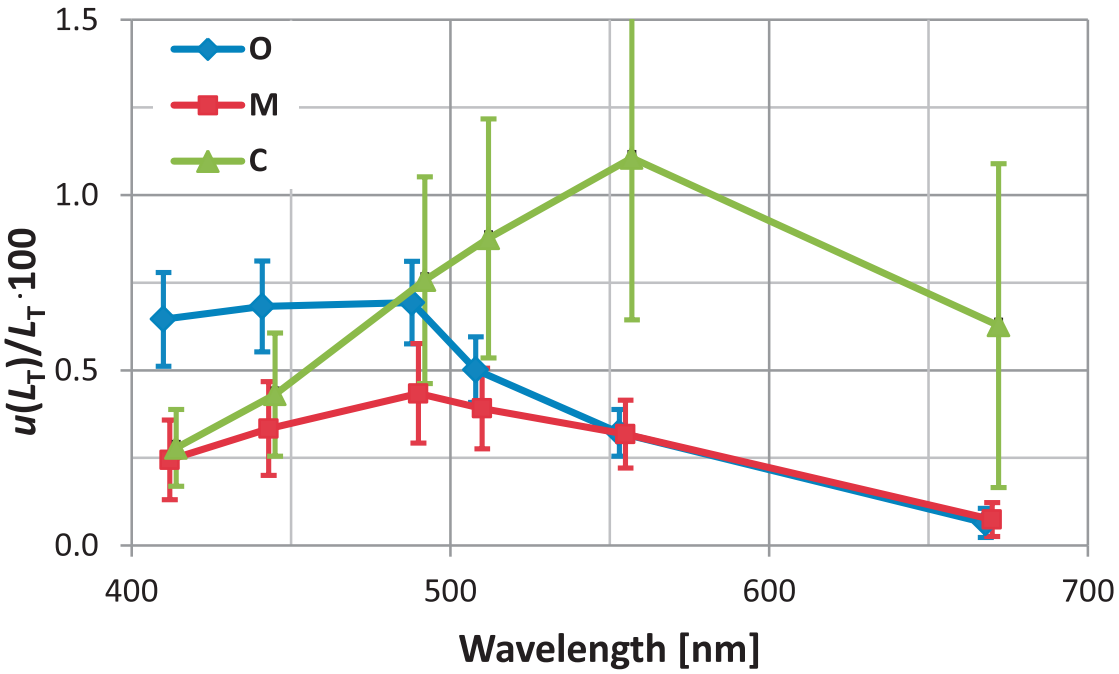


Figure 3

Click here to download Figure: Figure 3.eps

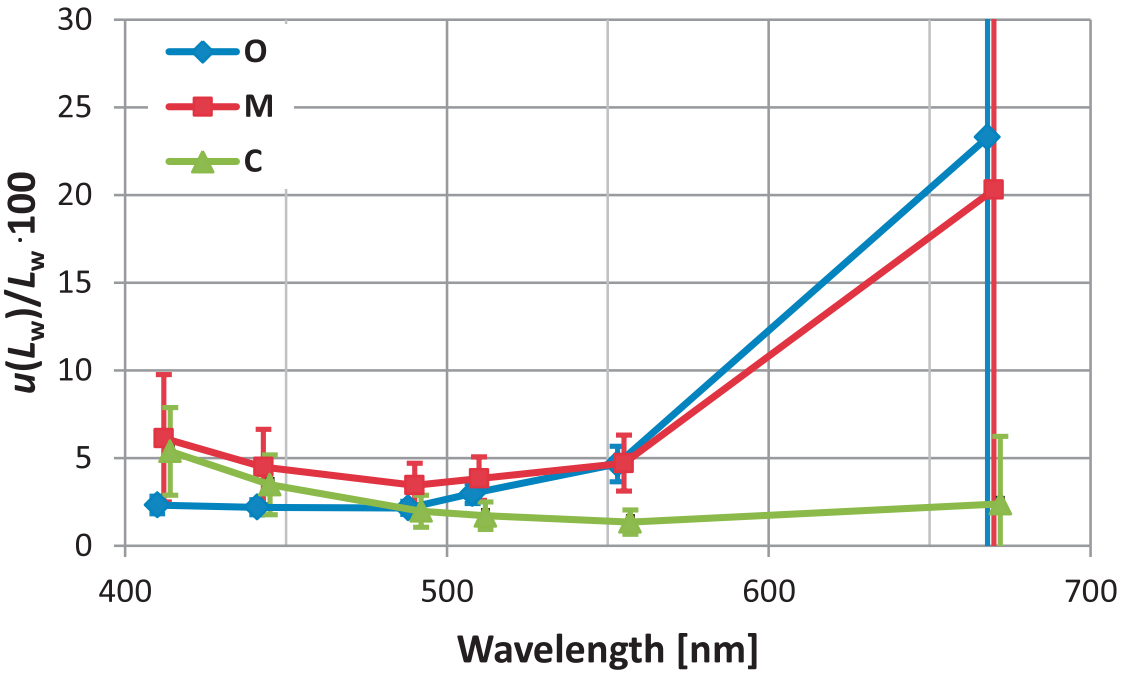


Figure 4

Click here to download Figure: Figure 4.eps