Sensor-independent approach to the vicarious calibration of satellite ocean color radiometry

Bryan A. Franz,[1,2,*] Sean W. Bailey,[1,3] P. Jeremy Werdell,[1,4] and Charles R. McClain[1,5]

[1]Ocean Biology Processing Group, 614.8, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA
[2]Science Applications International Corporation, 10260 Campus Point Dr., San Diego, California 92121, USA
[3]Futuretech Corp., 7307 Hanover Pkwy., Suite D, Greenbelt, Maryland 20770, USA
[4]Science Systems and Applications, Inc., 10210 Greenbelt Rd., Suite 600, Lanham, Maryland 20706, USA
[5]National Aeronautics and Space Administration, 614.8, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

*Corresponding author: bryan.franz@gsfc.nasa.gov

Received 5 February 2007; revised 2 April 2007; accepted 3 April 2007;
posted 4 April 2007 (Doc. ID 79697); published 9 July 2007

The retrieval of ocean color radiometry from space-based sensors requires on-orbit vicarious calibration to achieve the level of accuracy desired for quantitative oceanographic applications. The approach developed by the NASA Ocean Biology Processing Group (OBPG) adjusts the integrated instrument and atmospheric correction system to retrieve normalized water-leaving radiances that are in agreement with ground truth measurements. The method is independent of the satellite sensor or the source of the ground truth data, but it is specific to the atmospheric correction algorithm. The OBPG vicarious calibration approach is described in detail, and results are presented for the operational calibration of SeaWiFS using data from the Marine Optical Buoy (MOBY) and observations of clear-water sites in the South Pacific and southern Indian Ocean. It is shown that the vicarious calibration allows SeaWiFS to reproduce the MOBY radiances and achieve good agreement with radiometric and chlorophyll *a* measurements from independent *in situ* sources. We also find that the derived vicarious gains show no significant temporal or geometric dependencies, and that the mission-average calibration reaches stability after ~20–40 high-quality calibration samples. Finally, we demonstrate that the performance of the vicariously calibrated retrieval system is relatively insensitive to the assumptions inherent in our approach.

OCIS codes: 010.0010, 280.0280, 120.0120, 010.4450, 120.5630.

1. Introduction

Satellite ocean color data records provide the research community with a means of studying the Earth's marine biosphere on spatial and temporal scales unattainable via conventional *in situ* methods. The sea-viewing wide field-of-view sensor (SeaWiFS [1]) and moderate resolution imaging spectroradiometer flying on the Aqua spacecraft (MODIS-Aqua [2]), for example, have supplied global marine bio-optical data since 1997 and 2002, respectively. The community relies on these data to support studies ranging from regional ecosystem monitoring to the development of global climate records.

Space-based ocean color sensors measure the radiance exiting the top of the atmosphere (TOA) at a number of discrete wavelengths, λ, generally spanning the visible and near infrared (NIR) spectral regime. An atmospheric correction algorithm (e.g., Gao *et al.* [3], Antoine and Morel [4], Gordon and Wang [5]) is required to retrieve the portion of that TOA radiance signal, $L_t(\lambda)$, that is associated with radiance upwelled from beneath and transmitted through the sea surface. The desired uncertainties on this water-leaving radiance retrieval, $L_w(\lambda)$, however, cannot be achieved through instrument calibration and characterizations alone [6]. For example, the prelaunch calibration uncertainties for SeaWiFS

0003-6935/07/225068-15$15.00/0

are approximately 3% of the $L_t(\lambda)$ signal [7]. In a typical open-ocean scenario of oligotrophic waters, where the absorption of blue light is minimal, $L_w(\lambda)$ contributes ~10% to the total signal at the TOA in the blue-green spectral regime (i.e., 412 to 555 nm). As such, the 3% uncertainty in the prelaunch calibration approaches 30% on $L_w(\lambda)$, which is well above the stated goal of 5% for the L_w retrieval at 443 nm, $L_w(443)$ [8]. To retrieve water-leaving radiances with sufficient fidelity for climate and ecosystem research, satellite ocean color sensors require additional on-orbit calibration [6].

The Ocean Biology Processing Group (OBPG) at NASA Goddard Space Flight Center is responsible for the operational processing of ocean products from SeaWiFS, MODIS, and other ocean color capable sensors. In the vicarious calibration process developed by the OBPG, multiplicative correction factors are derived that force the instrument response at each sensor wavelength, in combination with the atmospheric correction algorithm, to retrieve expected values of $L_w(\lambda)$. During operational data processing, these gain factors are applied to $L_t(\lambda)$, effectively updating the prelaunch and onboard instrument calibration to account for characterization errors or undetermined postlaunch changes in response, as well as any systematic bias associated with the atmospheric correction algorithm. In this paper, the operational vicarious calibration approach employed by the OBPG is described in detail. Following that, we present results for the calibration of the SeaWiFS mission and investigate the sensitivity of those results to various assumptions within our approach.

2. Approach

The vicarious calibration of a satellite-borne ocean color radiometer, as defined here, is a system calibration that includes the instrument, which measures $L_t(\lambda)$, and the processing algorithm that is required to remove the atmospheric signal from the observed radiance and retrieve $L_w(\lambda)$. The goal is to adjust the sensor + algorithm system response as needed to maximize the agreement between remotely sensed water-leaving radiance retrievals and the expected water-leaving radiance. That expectation is typically based on *in situ* measurements of upwelling radiance at the sea surface, but it can also be based on models or regional climatologies, or even retrievals from another remote sensor. While the OBPG currently uses $L_w(\lambda)$ derived from the Marine Optical Buoy (MOBY) [9] for the visible-band vicarious calibration, the methodology described here does not presume anything about the heritage of the ground truth $L_w(\lambda)$ targeted for calibration.

The vicarious calibration is performed after the instrument calibration. While a discussion of instrument-level calibration is beyond the scope of this document, it is assumed that every effort has been made to incorporate prelaunch characterization of instrument responsivities, temperature sensitivities, and optical sensitivities (e.g., mirror-side differences, scan-angle dependencies) into the measurement of $L_t(\lambda)$. Furthermore, it is assumed that any temporal degradation of the instrument response has been corrected based on the on-board calibration system (e.g., lunar and solar calibrations). However, some residual error in instrument calibration is expected, particularly due to changes in the response subsequent to laboratory testing but preceding orbital operations. The vicarious calibration is therefore used to correct for any error due to instrument calibration deficiencies, as well as systematic bias in the processing algorithm.

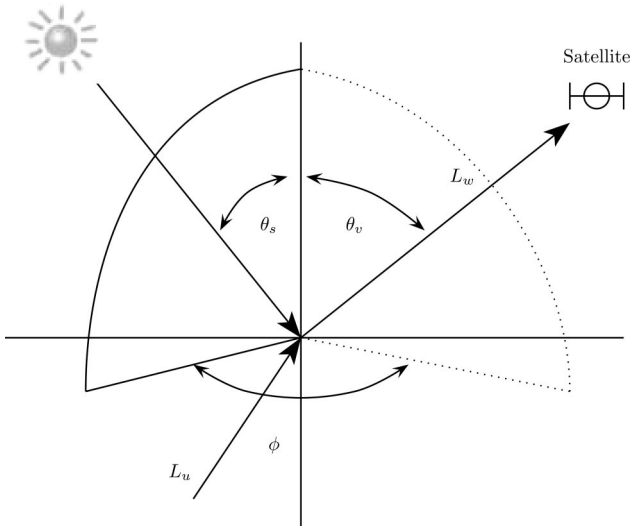

Fig. 1. Geometry definition showing radiant path vector from the Sun to a location on the Earth's surface, path of upwelling radiance (L_u), and path of water-leaving radiance (L_w) from surface to satellite sensor. θ_s, θ_v, and ϕ are the solar and sensor view zenith angles and the relative azimuth angle, respectively.

The operational processing algorithm takes $L_t(\lambda)$ as input and produces $L_w(\lambda)$ as output, and uncertainties in that process vary with surface and atmospheric conditions as well as viewing and solar radiant-path geometries (θ_s, θ_v, and ϕ: Fig. 1). The vicarious calibration is effectively an inversion of the forward processing algorithm, wherein a known water-leaving radiance, $L_w{}^t(\lambda)$, is the input and predicted TOA radiance, $L_t{}^t(\lambda)$, is the output (where the superscript t is introduced to indicate targeted or predicted values). The ratio of predicted to observed TOA radiance is the vicarious gain: the correction factor that, when applied to $L_t(\lambda)$, would force the system to yield $L_w{}^t(\lambda)$. If the calibration inversion is performed for a series of $[L_t(\lambda), L_w{}^t(\lambda)]$ match-up pairs, the resulting set of gains can be averaged for each sensor wavelength, and any geometry or time-dependent variabilities due to uncharacterized instrument response or poor algorithm performance will manifest as uncertainties on that mean gain. In this scenario, the vicarious calibration provides a mechanism for adjusting the system performance to minimize the average difference between expected water-leaving radiances (e.g., *in situ* measurements)

and retrieved water-leaving radiances, for nominal geometries and environmental conditions.

A. Atmospheric Correction

To describe the vicarious calibration process in detail, it is useful to review the components of the atmospheric correction algorithm. The standard algorithm employed within the OBPG is based on the work of Gordon and Wang [5], with additional modifications [10,11]. Equation (1) describes the process through which contributions from $L_w(\lambda)$, which is the quantity we wish to retrieve, and other surface and atmospheric sources are combined in the operational atmospheric correction model to form $L_t(\lambda)$, i.e.,

$$L_t(\lambda) = [L_r(\lambda) + L_a(\lambda) + t_{d_v}(\lambda)L_f(\lambda) + t_{d_v}(\lambda)L_w(\lambda)]t_{g_v}(\lambda)t_{g_s}(\lambda)f_p(\lambda). \quad (1)$$

In Eq. (1), $L_r(\lambda)$, $L_a(\lambda)$, and $L_f(\lambda)$ represent the radiance contributions associated with air molecules (Rayleigh scattering) [12,13], aerosols (including Rayleigh-aerosol interactions) [5], and surface whitecaps or sea foam [14–16], respectively. The $t_{d_v}(\lambda)$ term accounts for diffuse transmittance along the sensor view path from the surface to the satellite, while $t_{g_s}(\lambda)$ and $t_{g_v}(\lambda)$ account for losses due to gaseous absorption along the radiant paths from the Sun to the surface and the surface to the sensor, respectively [5,17]. The remaining term, $f_p(\lambda)$, is a correction for instrument response to the polarization of the observed radiance [18,19]. All terms in Eq. (1) and subsequent equations are summarized with relevant references in Table 1, and spectral dependence is hereafter implied unless required for clarity. Note also that the additional component of L_t associated with direct specular reflection of the Sun from the surface to the sensor is not included here, as satellite observations for which Sun glint contamination has been predicted are excluded from the OBPG calibration.

The primary unknowns in Eq. (1) are L_w and L_a, with the latter also affecting t_{d_v}. The remaining terms are computed *a priori* or reliably estimated given the radiant-path geometries. In the majority of ocean waters, where absorption in the NIR is very strong and reflectance is weak, the contribution of L_w to L_t can be assumed negligible or reliably estimated [20], and L_a(NIR) can be directly retrieved. With the Gordon and Wang atmospheric correction algorithm, if L_a is known at two NIR wavelengths then the aerosol type and concentration can be determined and combined with associated models to retrieve L_a in all bands [5]. With L_a fully determined, L_w can then be retrieved at all visible wavelengths by rearrangement of Eq. (1).

The retrieved L_w are subsequently normalized to the conditions of a nonattenuating atmosphere with the Sun directly overhead at a distance of 1 AU [21], as shown in Eq. (2). Here, μ_s is the cosine of the solar zenith angle (θ_s, Fig. 1), f_s adjusts for changes in Earth–Sun distance [22], f_b is a surface bidirectional reflectance correction [23–25], and f_λ accounts for spectral band-pass effects [26]. The spectral distribution of this normalized water-leaving radiance, $L_{wn}(\lambda)$, is the fundamental quantity we wish to determine, as it forms the basis for derived products, such as chlorophyll a concentration (C_a) or water absorption and scattering coefficients.

$$L_{wn}(\lambda) = L_w/(\mu_s f_s t_{d_s} f_b f_\lambda) \quad (2)$$

B. Vicarious Calibration

In operational processing, we use Eqs. (1) and (2) to retrieve L_{wn} from L_t. In the vicarious calibration, Eqs. (3)–(5), we reverse the process to retrieve $L_t^{\,t}$ for a specified $L_{wn}^{\,t}$ (referred to herein as the target value).

Table 1. Glossary of Symbols

Symbol	Description	References
λ	Sensor wavelength	[1,2]
L_t, $L_t^{\,t}$	TOA radiance, observed or (t) predicted	[1,2]
L_r	Radiance due to Rayleigh scattering from air molecules	[12,13]
L_a	Radiance due to scattering by aerosols, including Rayleigh–aerosol interactions	[5]
L_f	Radiance associated with whitecaps (foam) on the sea surface	[14–16]
L_w, $L_w^{\,t}$	Water-leaving radiance, retrieved or (t) targeted	[21]
L_{wn}, $L_{wn}^{\,t}$	Normalized water-leaving radiance, retrieved or (t) targeted	[21,23–25]
t_{g_s}, t_{g_v}, $t_{g_s}^{\,t}$	Transmittance due to gaseous absorption (e.g., ozone) for solar path (s) and sensor view path (v)	[5]
t_{d_s}, t_{d_v}, $t_{d_s}^{\,t}$	Rayleigh-aerosol diffuse transmittance for solar path (s) and sensor view path (v)	[5,17]
f_p	Polarization correction factor	18,19]
f_s, $f_s^{\,t}$	Earth–Sun distance correction	[22]
f_b, $f_b^{\,t}$	Bidirectional reflectance correction	[23–25]
f_λ, $f_\lambda^{\,t}$	Band-pass adjustment to L_{wn} or $L_{wn}^{\,t}$	[10,26]
θ_s, θ_v	Zenith angles for solar path (s) and sensor view path (v)	Fig. 1
μ_s, $\mu_s^{\,t}$	Cosine of solar zenith angle	Fig. 1
g_i	Vicarious gain for calibration sample (date and location) i	
$\bar{g}$	Mean vicarious gain	

As shown, g_i is the multiplicative correction to L_t that will force the instrument–algorithm system to yield the target value of ${L_{wn}}^t$ for observation case (or calibration sample) i. This direct approach differs from early SeaWiFS calibration efforts (i.e., Eplee *et al.* [7]), where both the visible and NIR vicarious gains were determined by an iterative comparison between satellite-retrieved and target L_{wn}.

$$L_{wn}{}^t(\lambda) = L_w{}^t/\left(\mu_s{}^t f_s{}^t t_{d_s}{}^t t_{g_s}{}^t f_b{}^t f_\lambda{}^t\right) \tag{3}$$

$$L_t{}^t(\lambda) = \left[t_{d_v} L_w{}^t(\mu_s t_{d_s} f_s f_b f_\lambda) + t_{d_v} L_f + L_r + L_a{}^t\right] t_{g_v} t_{g_s} f_p \tag{4}$$

$$g_i(\lambda) = L_t{}^t(\lambda)/L_t(\lambda) \tag{5}$$

The terms on the right-hand-side denominator of Eq. (3) may differ from those in Eq. (2) due to differences in the solar- and view-path geometries between the satellite retrieval and the target value. If ${L_w}^t$ is derived from *in situ* observations, for example, the view zenith angle is generally 0° (as in the case of a profiling radiometer or mooring), but the satellite view zenith angle of that same location can range from 0° to 60°. Furthermore, if the satellite observation and target ${L_w}^t$ are collected at different times of day, the solar zenith angle difference must be accounted for in all terms.

An additional term, ${t_{g_s}}^t$, has also been added to Eq. (3) to normalize for gaseous transmittance losses at the target. This term does not appear in the satellite normalization of L_w, Eq. (2), because it was already applied to the total observed radiance at the satellite via Eq. (1). The total atmospheric transmittance from the Sun to the surface is the product of gaseous transmittance, t_{g_s}, and diffuse transmittance, t_{d_s}, where the latter is a combination of Rayleigh and aerosol contributions and is therefore dependent on the aerosol properties. As such, ${t_{d_s}}^t$ represents an additional unknown for the calibration target. The total transmittance for the target, ${t_{d_s}}^t {t_{g_s}}^t$, can be obtained from contemporaneous *in situ* observations, such as a cosine collector or Sun photometer if such measurements are available. Alternatively, the satellite-retrieved aerosol properties can be used in combination with the solar zenith angle of the target and the atmospheric correction models to derive ${t_{d_s}}^t$ and ${t_{g_s}}^t$. As our purpose is to calibrate the instrument–algorithm system, the use of satellite-retrieved aerosol properties is advantageous in that it ensures that the L_w and ${L_w}^t$ are normalized with a common atmosphere. If the times of observation for L_w and ${L_w}^t$ are close, the relative effect will be small. However, if ${t_{d_s}}^t$ is determined from an *in situ* sensor, any error in that measurement will contribute additional uncertainty to the target ${L_{wn}}^t$. In practice, the OBPG uses the atmospheric properties retrieved from the satellite to determine the total atmospheric transmittance for the target via Eq. (6), whereby we simply adjust the retrieved transmittance terms for the slight difference in solar-path length between the satellite and target observation times, i.e.,

$$\left({t_{d_s}}^t {t_{g_s}}^t\right) = \exp\left[-\ln(t_{d_s} t_{g_s})\mu_s/\mu_s{}^t\right]. \tag{6}$$

Two other quantities in Eqs. (2) and (3) merit further discussion. Considering that all radiance terms in Eq. (1) are computed for the full relative spectral response of each sensor band-pass [27], the f_λ term converts the resulting full-band L_w to a nominal center wavelength value, effectively removing residual out-of-band response [26]. For the target, if ${L_w}^t$ is measured over a narrow band-pass at the nominal center wavelengths of the satellite sensor, then $f_\lambda{}^t = 1$. In contrast, if ${L_w}^t$ is obtained from a hyperspectral instrument (e.g., MOBY), then ${L_w}^t(\lambda)$ can be convolved with the relative spectral response of the sensor to be calibrated, and the spectral band-pass correction terms can be dropped from Eqs. (3) and (4). In the general case, $f_\lambda{}^t$ is used to shift ${L_w}^t$ to the band-pass of the sensor to be calibrated.

Finally, the bidirectional reflectance correction, f_b, accounts for radiant-path geometry dependencies in L_w due to anisotropy of the near-surface light field, which is a function of the absorption and scattering properties of the water column and its constituents. The algorithm employed by the OBPG to estimate f_b uses the concentration of C_a as a proxy for these inherent optical properties [23], where C_a is computed from the spectral distribution of L_{wn} via the operational algorithm (e.g., OC4 [28] for SeaWiFS and OC3M [28] for MODIS-Aqua). Note that this process requires iteration, as L_{wn} is required to determine f_b, and f_b is required to determine L_{wn}. For the inverse case, $f_b{}^t$ can be computed using C_a measurements associated with the calibration target (e.g., *in situ* fluorometry), or using ${L_w}^t$ as input into the operational C_a algorithm when target C_a is not available (as is the case for MOBY).

3. Calibration of the Near-infrared Bands

The vicarious calibration approach presented in Eqs. (3)–(5) requires that the aerosol radiances associated with the calibration target, ${L_a}^t$, are known for each spectral band. While it is possible to utilize additional target measurements of aerosol properties in the calibration process [29], reliable, simultaneous, and co-located measurements of aerosol properties and water-leaving radiances are not widely available from *in situ* sources. Lacking ground truth, the OBPG makes a number of assumptions to determine the aerosol contribution. Our approach takes advantage of the fact that, with the Gordon and Wang algorithm, the determination of L_a and t_{d_v} in the visible bands is exclusively dependent on the observed radiances in two NIR bands. Thus, a two step strategy is employed wherein we first calibrate the NIR bands and then fix that calibration and utilize the retrieved aerosol properties to reduce Eq. (4) to one unknown, the ${L_w}^t$, in each visible band. Here we describe the strategy employed to calibrate the NIR bands.

We begin with two simplifying assumptions: (1) that the water-leaving radiance in the two NIR bands is truly negligible, and (2) that the instrument calibration of the longest NIR wavelength (NIR_L, e.g., 865 and 869 nm for SeaWiFS and MODIS, respectively) is perfect, such that $g_i(NIR_L) = 1$ for all i. The first assumption is valid if the location of our calibration target is in highly oligotrophic waters, as pure sea water is a near perfect absorber in the NIR spectral regime. With $L_w(NIR) = 0$, Eqs. (1) and (4) reduce to Eqs. (7) and (8) presented next. The only significant unknown in these equations is L_a (t_{d_v} for L_f is estimated *a priori* based on Rayleigh scattering alone, in both the forward and inverse processes), so the uncalibrated estimation of L_a can be directly retrieved for the two NIR bands. Furthermore, when combined with the second assumption, and recognizing that $L_a{}^t = L_a$ for $\lambda = NIR_L$, $L_a{}^t(NIR_L)$ is fully determined for any clear-water calibration target.

$$L_t(\mathrm{NIR}) = [L_r + L_a + t_{d_v} L_f] t_{g_v} t_{g_s} f_p \tag{7}$$

$$L_t{}^t(\mathrm{NIR}) = [L_r + L_a{}^t + t_{d_v} L_f] t_{g_v} t_{g_s} f_p \tag{8}$$

The remaining objective is to calibrate the shorter of the two NIR wavelengths (NIR_S, e.g., 765 and 748 nm for SeaWiFS and MODIS, respectively). In the Gordon and Wang [5] algorithm, it is the ratio of the aerosol radiances in the two NIR channels that determines the aerosol type, where various aerosol types are represented by a suite of Shettle and Fenn [30] aerosol models. If the aerosol type is known, the associated model can be used in combination with the retrieved $L_a(NIR_L)$ to predict L_a in the shorter NIR wavelength, $L_a{}^t(NIR_S)$. This process is completely independent of the visible calibration, so the location of the NIR calibration site need not be coincident with that of the visible bands. We desire atmospheric conditions wherein the aerosol type is generally stable and predictable. We therefore select locations that are far from land or active volcanic islands, where the dominant aerosols result from purely maritime processes (i.e., sea salt and water vapor). Such open ocean locations also tend to satisfy our requirement for oligotrophic waters, where Eqs. (7) and (8) are strictly valid. The OBPG currently uses two sites for the NIR band calibration (Fig. 2): the South Pacific Gyre (SPG) and the Southern Indian Ocean (SIO). These locations are known to exhibit some of the clearest and most homogenous ocean waters on Earth [31,32], with stable aerosol conditions that are representative of noncoastal maritime atmospheres [33]. For each cloud-free glint-free observation of the target location, we assume an aerosol type that is consistent with that site, and we use the corresponding model to compute $L_a{}^t(NIR_S)$. Putting this into Eqs. (5) and (8), the vicarious gain for the shorter NIR band, $g_i(NIR_S)$, can be determined.

A set of NIR gains are computed for various observation dates of the target location(s) over the satellite mission lifespan, and these $g_i(NIR_S)$ are then averaged to determine the mission mean vicarious gain, $\bar{g}(NIR_S)$. In practice, the OBPG actually performs each individual calibration match-up on a 15×15 pixel area centered on the calibration target, and $g_i(NIR_S)$ is computed as the spatial average of up to 225 pixel gains [thus $\bar{g}(NIR_S)$ is the temporal average of a set of spatial averages]. Both the spatial and temporal averages are computed using the mean of the semi-interquartile range (MSIQR), which is defined as the simple average of the data within the 25th to 75th percentiles) to minimize the effects of spurious outliers. With the NIR calibration established and fixed for all time and space, the Gordon and Wang [5] atmospheric correction process can be operated to determine $L_a{}^t(\lambda)$ for all λ.

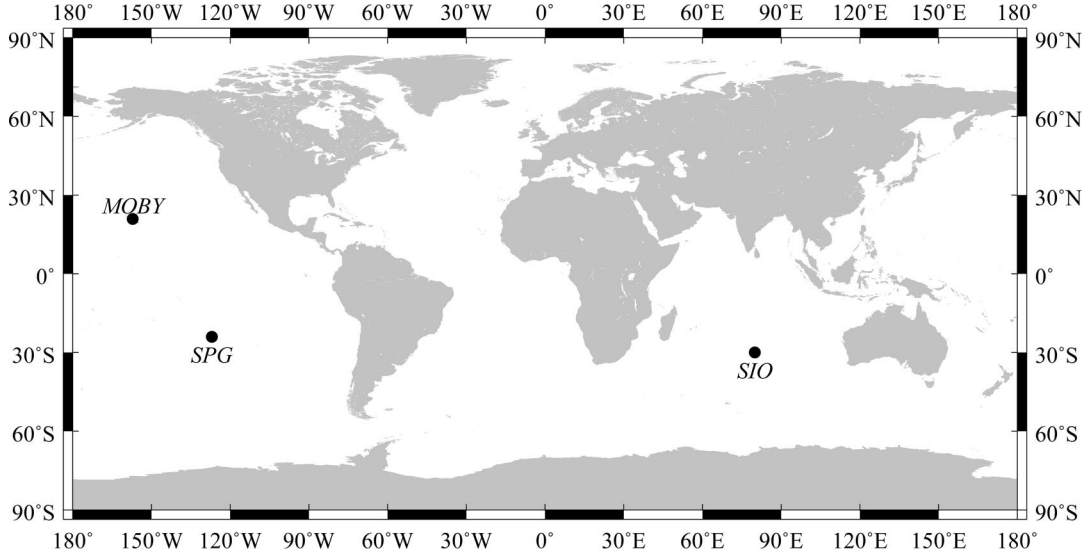

Fig. 2. Map showing location of the vicarious calibration sites used operationally by the OBPG for the calibration of SeaWiFS and other ocean color sensors. The MOBY site is used for the calibration of the visible wavelengths. The locations in the SPG and SIO are used for the vicarious calibration of the NIR wavelength(s).

4. Calibration of the Visible

The previous discussion attempted to maintain a degree of generality, as our approach to vicarious calibration is specific to the atmospheric correction algorithm but independent of the particular satellite sensor or the source of the calibration target data. In this section, we discuss the specifics of using a hyperspectral *in situ* radiometer to calibrate the visible channels of a space-based ocean color sensor. MOBY has been continuously moored at a site approximately 20 km west of the island of Lanai, Hawaii since late 1996, providing high-quality water-leaving radiance measurements for the vicarious calibration of SeaWiFS, MODIS, and various international ocean color sensors.

A. MOBY Data Set

MOBY measures upwelling radiance over the spectral range from 340–955 nm with 0.6-nm spectral resolution. The water-leaving radiance data from MOBY is provided to the OBPG by the MOBY Operations Team (MOT) [34], after convolution of the hyperspectral signal with the full spectral response of each band on the satellite sensor. Thus, the MOBY measurements represent the expected full-band response of the satellite sensor, and the f_λ and f_λ^t can therefore be dropped from Eqs. (3) and (4), respectively. MOBY is designed with three radiometric detectors positioned on separate arms that radiate from the central axis, allowing measurement of the upwelling radiance, L_u, at water depths of 1, 5, and 9 m. In the calculation of L_w^t, the L_u are extrapolated to the sea surface, which requires knowledge of the spectral attenuation of light through the water column. For MOBY, diffuse upwelling attenuation coefficients can be determined from the L_u measurements at any pair of discrete depths. In the processing performed by the MOT, L_u measured at the 1-m arm is extrapolated to the surface using the attenuation coefficient derived from the 1- and 5-m arms, and L_u measured at the 5-m arm is extrapolated to the surface using the attenuation coefficient derived from the 5- and 9-m arms. Only the L_w^t derived from the 1-m arm are used by the OBPG in the vicarious calibration process, with the second measurement utilized for quality screening.

The operational scenario for MOBY is to collect data at the approximate overpass time of each satellite supported by the MOT. Thus, for every potential satellite observation of the waters around MOBY, a contemporaneous *in situ* observation should exist. The number of satellite to *in situ* match-ups is reduced through quality screening of both the *in situ* measurements and the satellite observations. The MOBY measurements are initially quality controlled by the MOT [34] to eliminate such cases as instrument malfunction or significant buoy tilt, but the OBPG also excludes from the calibration process any measurements where radiances extrapolated from the 1- and 5-m arms are not within 5%. This additional quality control helps to eliminate calibration target data for days when the water properties or atmospheric conditions varied. Considering that the satellite may view MOBY at a slightly different time of day (a full MOBY measurement cycle requires approximately 20 minutes), such geophysical variabilities will tend to increase the error in the vicarious calibration. MOBY also measures surface irradiance via a cosine collector on the buoy. If the measured surface irradiance differs from a clear-sky irradiance model [35] by more than 10%, the measurement is not included in the calibration data set. This screens out cases where the MOBY measurements were obtained under variable illumination conditions.

B. Satellite to *In Situ* Match-Up Process

Given a set of quality screened *in situ* measurements with known location and time, a match-up process is performed to locate contemporaneous satellite sensor observations. This match-up process and subsequent quality screening is nearly identical to that described in Bailey and Werdell [36] for *in situ* validation of satellite sensor retrievals. Briefly, a regional extract of the satellite data is generated over the calibration site for each target date. The extract area is a 5×5 pixel box, where the pixel size is the native resolution of the satellite sensor. If any of the pixels within the 5×5 box are flagged by the atmospheric correction code as being contaminated by land, clouds, cloud shadows, or stray light, or if any pixel is flagged for navigation problems or fails atmospheric correction, the scene is excluded from further consideration. Furthermore, if the mean C_a retrieval for the scene is greater than 0.2 mg m^{-3}, the retrieved aerosol optical thickness in the NIR is greater than 0.15, θ_v is greater than 56°, or θ_s is greater than 70°, the scene is excluded. For each extract scene that passes all exclusion criteria, a gain is computed for each of the 25 pixels via Eqs. (4) and (5). A final gain for the scene, $g_i(\lambda)$, is then computed as the MSIQR over the distribution of the 25 pixel gains. The MSIQR provides a final level of statistical outlier rejection to minimize effects such as subpixel-scale clouds or undetected straylight in the individual gains.

Due to the extensive quality screening of the *in situ* and satellite observations, and particularly the frequency of cloudy conditions and Sun glint contamination, a relatively small number of the potential match-ups to MOBY are ultimately utilized in the calibration of the satellite. As an example, SeaWiFS observes MOBY at least once every other day. Allowing for rough seas and instrument malfunction, the MOT has provided 1450 contemporaneous MOBY match-ups for SeaWiFS calibration over 9 years. Of those 1450 opportunities, 150 (or approximately 10%) passed the satellite data screening process. For the MODIS sensors, which lack sensor tilt capabilities and are thus more susceptible to Sun glint contamination losses, the return is even lower. The individual $g_i(\lambda)$ from this select set of high-quality match-ups are aggregated via the MSIQR to derive the mission-

Table 2. SeaWiFS Vicarious Gain Coefficients

λ	412	443	490	510	555	670	765	865	N[a]
$\bar{g}$	1.0377	1.014	0.9927	0.9993	1.000	0.9738	0.9720	1.000	150 (97)
σ^b	0.009	0.009	0.008	0.009	0.008	0.007	0.010	0.0	
$S_E{}^c$	0.0007	0.0007	0.0007	0.0007	0.0007	0.0006	0.0011	0.0	

[a]Number of gain samples, g_i, used to compute the mean gain, $\bar{g}$, for $\lambda < 765$ ($\lambda = 765$).
[b]Standard deviation of the distribution of g_i about $\bar{g}$.
[c]Standard error on the mean, $\bar{g}$, computed as σ/sqrt(N).

averaged vicarious gain, $\bar{g}(\lambda)$, for each visible wavelength of the sensor.

C. Application to SeaWiFS

The SeaWiFS mission provides an ideal case study for the OBPG vicarious calibration process, as there now exists a time-series of quality screened SeaWiFS-MOBY match-ups spanning nearly a decade. In addition to increasing the confidence in the mean vicarious gain, the large data set allows for the investigation of residual dependencies with time- or radiant-path geometry. The calibration of the SeaWiFS NIR band at 765 nm was derived using the SIO and SPG sites with an assumed maritime aerosol type [30] at 90% relative humidity, and the visible bands were subsequently calibrated to MOBY $L_w{}^t$. The mean vicarious gain derived for each sensor band is provided in Table 2. While the derived gains for all bands are relatively small at less than 4% (relative to unity), there is a spectral dependence from 3.77% high at 412 nm to 2.8% low of at 765 nm that will be addressed in the next section. The standard deviation of the distribution of $g_i(\lambda)$ about $\bar{g}(\lambda)$ is consistently at or below 1%, thus yielding a standard error on $\bar{g}(\lambda)$, S_E, of ~0.1% at all wavelengths.

The $\bar{g}(\lambda)$ and individual $g_i(\lambda)$ are plotted as a function of mission time for several important bands in Fig. 3. The reduced match-up sampling of the visible

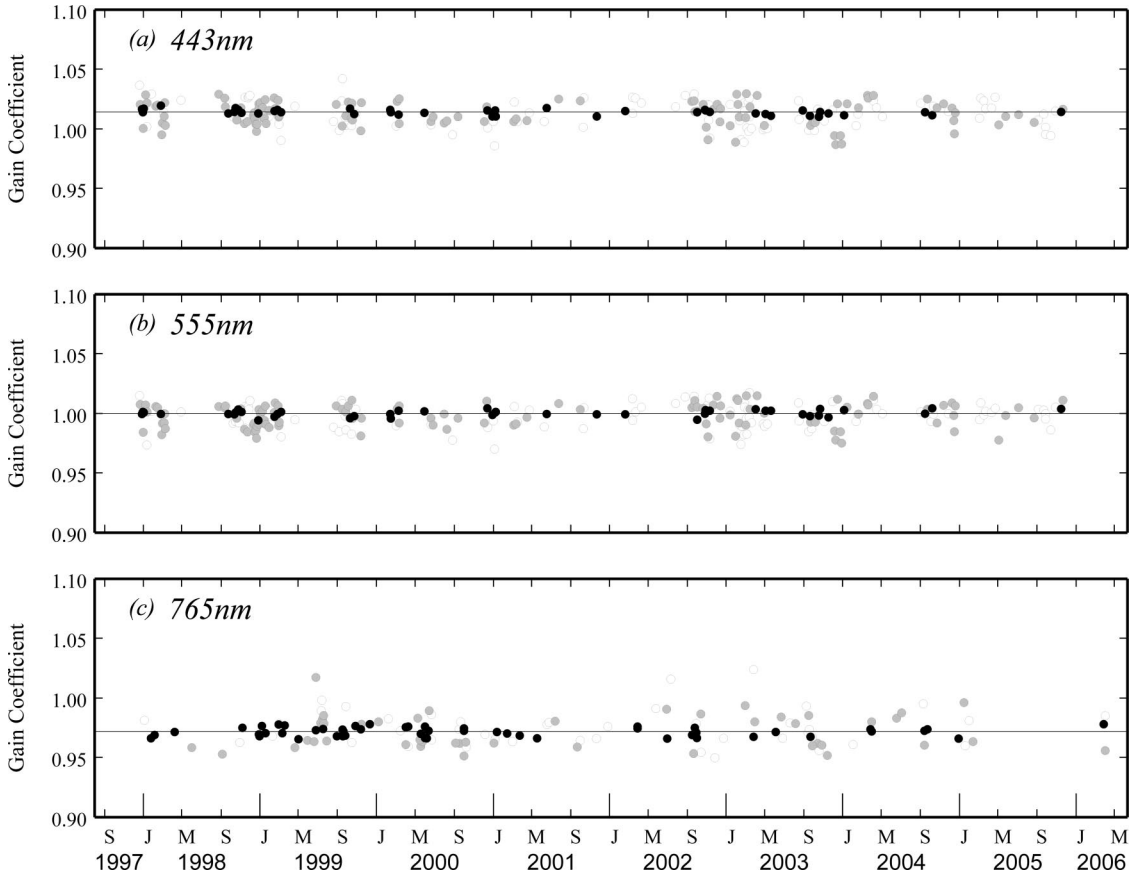

Fig. 3. Vicarious gains derived for SeaWiFS bands at 443, 555, and 765 nm based on calibration samples spanning the mission lifetime from September 1997 to March 2006. The individual calibration gains (circles) are distributed around the mission mean gain line, which is constant for all time. The filled circles are the gains that passed the quality screening process, with the grey and black fill used to distinguish the cases that fell outside or within the semi-interquartile range, respectively. The J, M, and S labels indicate January, May, and September, respectively.

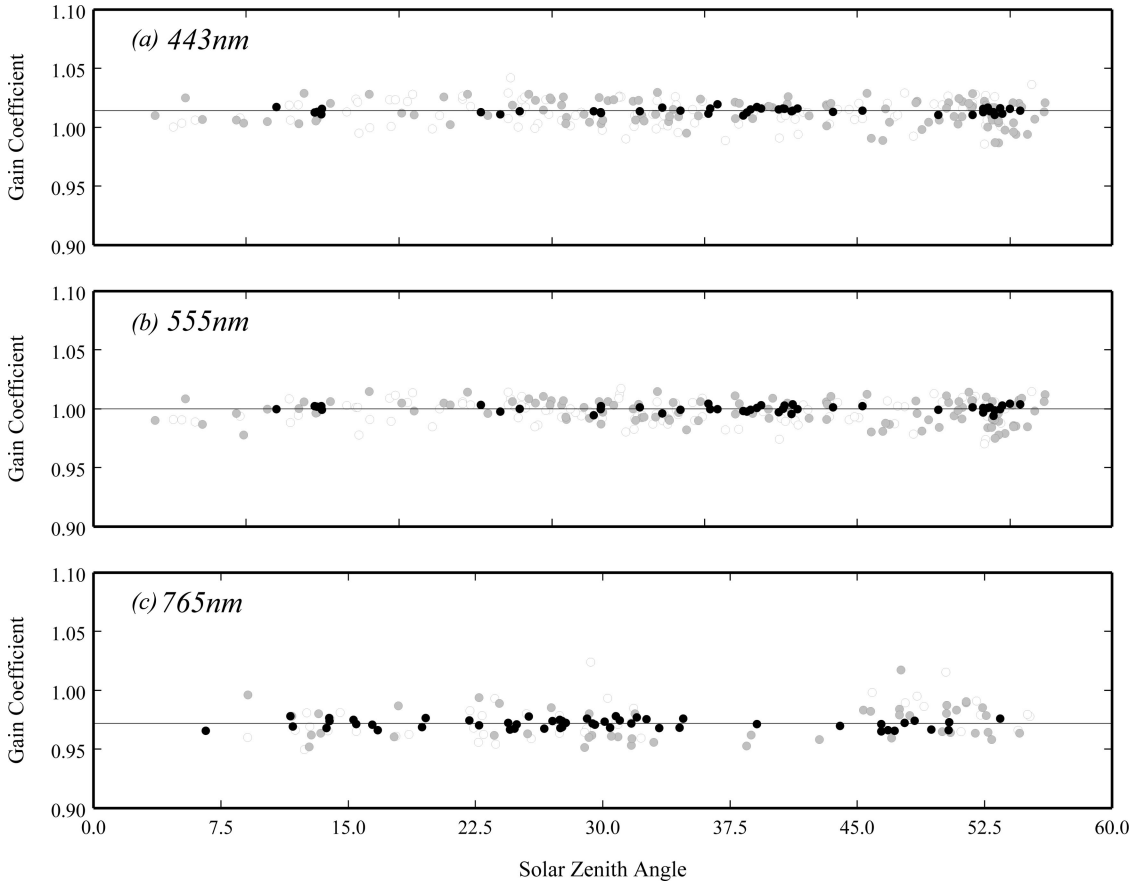

Fig. 4. Same data as in Fig. 3, but plotted as a function of the solar zenith angle associated with the time, date, and location of each calibration sample.

gains in the March (M) to September (S) periods is due to satellite retrieval losses for Sun glint and quality screening for seasonally elevated aerosol optical thickness. The match-up sampling for the NIR_S band at 765 nm [Fig. 3(c)] is more uniform because there are two NIR calibration sites (Fig. 2), which increases temporal sampling density, and the SIO site is at a latitude where Sun glint is less significant. The plots show that $g_i(\lambda)$ remains relatively stable as a function of time, both long-term and seasonally, which corroborates the temporal calibration of the instrument [37] and suggests consistency among the MOBY deployments (the MOT alternately deploys two mooring platforms for several months at a time). Similarly, $g_i(\lambda)$ is consistent over the range of solar and satellite view zenith angles associated with satellite observation of the target location (Figs. 4 and 5, respectively). While not evident in these trends, variations with geometry would suggest problems with the atmospheric correction algorithm or bidirectional reflectance estimations (f_b and $f_b{}^t$). Furthermore, variations in the sensor view angle may indicate uncorrected variability in the instrument's response with the scan angle, and variations with solar geometry might arise from complexities in the *in situ* determination of $L_w{}^t$ under certain sky conditions (e.g., instrument self-shading by the mooring platform or wave focusing and defocusing in high light conditions). If present, such systematic effects would tend to increase variability in $g_i(\lambda)$ and uncertainty in $\bar{g}(\lambda)$.

The extended time-series of the SeaWiFS mission also provides an opportunity to look at the change in the mean vicarious gain with sample size. While the OBPG only changes the operational vicarious calibration in coordination with a major reprocessing of the mission data set, this analysis can provide an indication as to the number of quality-screened calibration samples that may be required by future missions to achieve a stable mission-averaged gain. To mitigate the effect of any temporal instabilities specific to the SeaWiFS instrument or MOBY, we selected match-up cases at random rather than in time-order, growing the sample size one case at a time and recomputing $\bar{g}(\lambda)$ at each step. As with the vicarious calibration process, this analysis was performed independently for the NIR and visible-band calibrations, such that the change in visible-band $\bar{g}(\lambda)$ with sample size is derived for the case of a fixed NIR calibration. Figure 6 shows that the mean vicarious gain converged to within 0.1% of the mission-averaged value after ~30, 40, and 20 samples for the 443-, 555-, and 765-nm bands, respectively. The stepped behavior in the convergence is due to discrete outlier rejection associated with the MSIQR-averaging technique, but $\bar{g}(\lambda)$ fol-

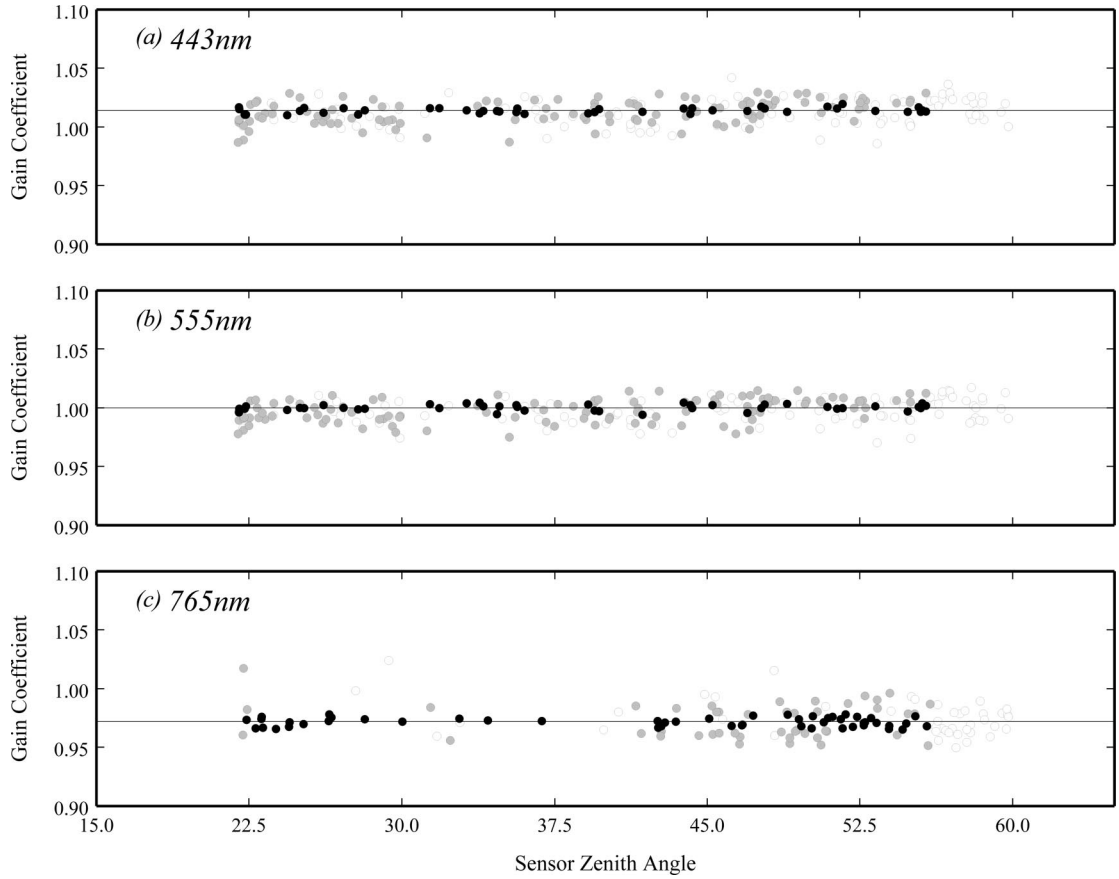

Fig. 5. Same data as in Fig. 3, but plotted as a function of the sensor view zenith angle. The SeaWiFS instrument scans from east to west over a scan angle range of +/−56°, but the sensor is also tilted 20° in the along track direction, so the resulting view zenith angle ranges from a minimum of 20° to a maximum of ~72°, but observations beyond 60° are discarded.

lows the expected change from high variability at low sample sizes to a near constant value at higher sample sizes. Once the vicarious calibration stabilizes, additional calibration measurements only serve to reduce the uncertainty in $\bar{g}(\lambda)$, as represented by the error bars indicating the standard error on the mean.

Given the average return of approximately 15 quality screened calibration samples per year, Figure 6 suggests that a sensor with similar observing capabilities to SeaWiFS may require 2–3 years to achieve a stable vicarious calibration in all spectral bands when a single target location is utilized for truth data. This assumes, however, that the inherent variability of the vicarious calibration system is similar to that of the SeaWiFS–MOBY system. Future satellite sensors would likely have higher signal-to-noise response, which would reduce the variability in the calibration gains and therefore reduce the number of samples required for calibration convergence, assuming that such sensors are temporally and geometrically characterized as well as the SeaWiFS instrument (e.g., Figs. 3–5). Furthermore, while a detailed analysis of the uncertainty and stability of MOBY relative to other potential calibration sources is beyond the scope of this analysis, the inherent variability of the vicarious calibration system is clearly dependent on the quality and stability of the target data.

D. Verification of Calibration

To estimate first-order uncertainties in the calibrated retrievals, we incorporate $\bar{g}(\lambda)$ into the OBPG satellite data product validation system [36] and calculate radiometric match-up statistics for the satellite retrieved L_{wn} and the *in situ* observations used to derive Table 2. Given the idealized nature of the calibration site such an analysis does not represent the error in a typical satellite retrieval of L_{wn}; however, the results presented in Table 3 do provide an indication of uncertainties associated with both the gain determination and the validation process. As expected, the satellite-to-*in situ* mean ratios and biases approach unity and zero, respectively, which demonstrates that the calibration was executed properly. Variability between the satellite and *in situ* L_{wn}, as estimated by the median absolute percent difference (MPD), is 1–2% in the blue-green spectral range, which is within the SeaWiFS radiometric accuracy goal [8] of 5%. The decrease in the correlation coefficient, r^2, with wavelength is caused by the limited dynamic range in green to red radiances exiting the spatially and temporally stable oligotrophic waters around MOBY. This is a natural result of the calibration site selection. Spatial homogeneity in the optical properties is an asset because the satellite sensor

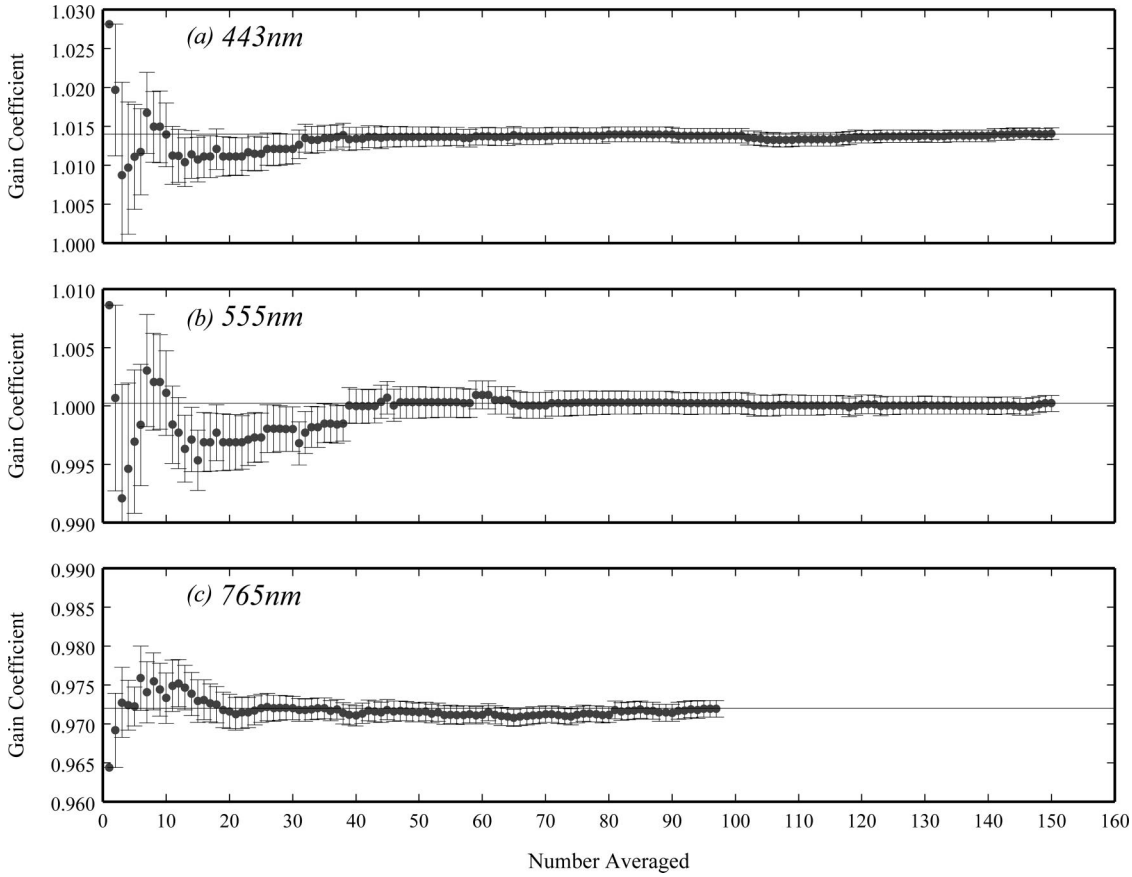

Fig. 6. Mean vicarious gains, $\bar{g}(\lambda)$, derived for SeaWiFS bands at 443, 555, and 765 nm based on calibration samples spanning the mission lifetime from September 1997 to March 2006. Individual gains from the mission-long set of calibration match-ups were randomly sampled, growing the sample set one case at a time and averaging to show the effect of increasing sample size on $\bar{g}(\lambda)$. Vertical error bars show the standard error on the mean at each sample size.

samples a much larger area (e.g., 1 km^2) than the typical *in situ* measurement, and we require the calibration target to be representative of the satellite sample. Oligotrophic conditions are also desirable as they are generally associated with temporally stable water constituents, such that small time differences between the satellite observation and the *in situ* observation are less of a concern.

The comparatively poor results for the 670-nm channel in Table 2 (i.e., $r^2 = 0.5$, slope = 4) are not unexpected, as the attenuation of pure sea water at wavelengths longward of 650 nm leads to water-leaving radiances very close to zero in clear water. The radiance ratios are therefore extremely sensitive to small differences between the field measurements and satellite retrievals, and the negligible dynamic range makes regression correlations meaningless. These problems are exacerbated by the relatively low signal-to-noise ratio of the SeaWiFS 670-nm band, as well as the general difficulty of acquiring accurate in-water measurements at that wavelength [38].

Table 3. Verification of Vicarious Calibration at MOBY

	Ratio[a]	MPD[b]	r^2	Slope[c]	Bias[d]
$L_{wn}(412)$	1.002	1.220	0.965	1.039	0.0033
$L_{wn}(443)$	1.003	1.184	0.956	1.009	0.0066
$L_{wn}(490)$	1.001	1.189	0.889	0.937	−0.0011
$L_{wn}(510)$	1.003	1.030	0.839	0.967	0.0014
$L_{wn}(555)$	0.998	2.270	0.731	1.284	0.0006
$L_{wn}(670)$	1.100	25.64	0.501	3.985	0.0011

[a]Median ratio of satellite to *in situ* L_{wn}.
[b]Median percent difference.
[c]Linear regression slope of satellite versus *in situ* L_{wn}.
[d]Mean bias [=Σ(satellite − *in situ*)/*N*].

For a more general validation of the calibration and atmospheric correction system, a similar match-up analysis was performed using globally distributed *in situ* measurements from the SeaWiFS Bio-optical Archive and Storage System (SeaBASS) [39]. Following the discussion in Bailey and Werdell [36], we restrict the match-up analysis to ground measurements collected in deep ocean water (depth greater than 1000 m), and we exclude measurements from MOBY. As expected, the agreement shown for the deep-water global match-up analysis presented in the left column of Table 4 is degraded relative to the idealized validation against the calibration scenes. Also shown in Table 4 is the comparison between SeaWiFS and *in situ* C_a, where the *in situ* C_a has been derived through high performance liquid chromatography and the satellite retrievals were derived using

Table 4. Validation of Vicarious Calibration Against Deep-Water *In Situ* Measurements

	Ratio[a]	MPD[a]	r^2	N[b]
$L_{wn}(412)$	1.002	11.8	0.930	188
$L_{wn}(443)$	0.950	15.5	0.873	318
$L_{wn}(490)$	0.942	12.2	0.817	318
$L_{wn}(510)$	0.957	10.6	0.579	164
$L_{wn}(555)$	0.968	14.8	0.827	318
$L_{wn}(670)$	1.347	64.7	0.595	306
C_a	0.994	26.1	0.875	149

[a]As defined in Table 3.
[b]Number of satellite to *in situ* match-up cases.

Table 5. Sensitivity of Deep-Water Validation to No Vicarious Calibration

	$\bar{g}(\lambda)$	Ratio[a]	MPD[a]	r^2	N^b
$L_{wn}(412)$	1.0000	0.245	80.0	0.861	54
$L_{wn}(443)$	1.0000	0.447	55.4	0.799	111
$L_{wn}(490)$	1.0000	0.760	25.7	0.772	111
$L_{wn}(510)$	1.0000	0.753	24.7	0.665	45

[a]As defined in Table 3.
[b]Number of satellite-to-*in situ* match-up cases.

the OC4v4 algorithm [28]. The satellite retrieval of C_a introduces additional uncertainties associated with the bio-optical algorithm that are independent of the accuracy of L_{wn} retrievals, but C_a is a critical product to be derived from ocean color missions so the sensitivity to vicarious calibration is relevant. The C_a algorithm is based on water-leaving reflectance ratios between two wavelengths (typically 443 and 555 nm in deep water), and thus it tends to absorb retrieval biases in L_{wn} that are spectrally independent and accentuate biases that are spectrally dependent. For our deep-water case, it is reassuring to see that the validation results show a mean ratio very close to unity; and, although the MPD is relatively large at 26%, it is within the uncertainties of the C_a algorithm derivation [28].

For the L_{wn} comparisons, the MPD of 10%–15% in the blue-green regime is significant relative to the SeaWiFS radiometric accuracy goal [8], but this is due to many factors, including error in the field measurements. Furthermore, given the limitation to deep water, where concentrations of C_a and other optically active constituents tend to be low and stable, the previous discussion regarding the degraded correlation with increasing wavelength applies here as well. Again, we refer to Bailey and Werdell [36] for a more complete discussion of the validation process and interpretation, but we present the results here to serve as a baseline for the analyses to be discussed in the next section.

5. Sensitivity Analyses

The deep-water validation can also be used to test the sensitivity of our vicarious calibration approach to various assumptions. Given that we are making relatively small changes (<4%) to the TOA radiances, one might question the need for any vicarious calibration. To assess this question, we ran the same deep water *in situ* validation analysis with no vicarious gains applied to the SeaWiFS observations, such that we are relying on the instrument prelaunch and onboard calibration capabilities alone. Comparing the results in Table 5 to those previously discussed in Table 4, the mean ratios for L_{wn} in the blue-green spectral regime are biased low relative to field measurements by 25% at 490 nm to 75% at 412 nm. Furthermore, C_a retrievals are degraded from a mean ratio near unity to a 25% bias low, indicating that the relative spectral distribution in the absence of vicarious calibration is skewed. Also note that the number of match-ups included in the statistics is significantly reduced for Table 5 relative to Table 4. This occurs because, without vicarious calibration, more than 50% of the original match-up cases failed to obtain a valid retrieval. Interestingly, the $L_{wn}(670)$ retrieval is actually improved relative to the fully calibrated case.

As previously discussed, the vicarious calibration adjusts for errors in both the instrument calibration and the atmospheric correction. Furthermore, the primary uncertainty in the atmospheric correction is the determination of contributions from aerosols, so perhaps it is only necessary to calibrate the NIR bands that govern the aerosol determination. For SeaWiFS, this calibration results in a reduction of the 765-nm radiances by approximately 3% [$\bar{g}(765) = 0.9720$]. The deep-water validation results presented in Table 6 demonstrate that the NIR vicarious calibration significantly improves agreement in retrieved C_a and $L_{wn}(\lambda)$ relative to field measurements. The only exception is the dramatically degraded agreement at $L_{wn}(670)$, which just serves to emphasize the relative sensitivity of this band to changes in the retrieval process. In the context of the Gordon and Wang atmospheric correction algorithm [5], the reduction in the 765-nm response relative to the 865-nm response has the effect of reducing the spectral slope of the selected aerosol models, which leads to reduced aerosol radiance estimates for the visible wavelengths. Referring to Eq. (1), the resulting increase in $L_{wn}(\lambda)$ brings the satellite retrievals in

Table 6. Sensitivity of Deep-Water Validation to NIR only Calibration; $\bar{g}(765) = 0.9720$

$\bar{g}(\lambda)$	Ratio[a]	MPD[a]	r^2	N^b
1.0000	0.595	40.6	0.915	188
1.0000	0.779	23.5	0.866	318
1.0000	1.002	11.3	0.816	318
1.0000	0.964	10.7	0.571	164
1.0000	0.965	15.0	0.822	318
1.0000	3.565	256.0	0.572	306
	0.984	26.1	0.872	144

[a]As defined in Table 3.
[b]Number of satellite-to-*in situ* match-up cases (a common set was used for Tables 6–10 and for the analysis presented in Table 4).

closer agreement with the field measurements, and the significant improvement in the C_a validation ratio suggests that the calibrated aerosol model selection yields a more accurate spectral distribution for $L_{wn}(\lambda)$. Still, comparing the results of Table 6 with those from Table 4 demonstrates that further improvement is achieved through vicarious calibration of the visible bands. In particular, the elevation of the 412- and 443-nm response [$\bar{g}(412) = 1.0377$ and $\bar{g}(443) = 1.014$] and depression of the 670-nm response [$\bar{g}(670) = 0.9738$] dramatically improved the validation agreement in those bands. It is worth noting that the magnitude and direction of the vicarious calibration adjustments and the associated improvement in the validation results appears to corroborate the findings of Barnes and Zalewski [40], in which a spectrally dependent error in the prelaunch calibration of SeaWiFS was identified.

Finally, in the vicarious calibration of the NIR bands it was assumed that the instrument calibration of the NIR_L band at 865 nm was sufficient for accurate ocean color retrievals [$\bar{g}(865) = 1.0$], and that the aerosols at our NIR calibration sites in the South Pacific and southern Indian Oceans (Fig. 2) can be characterized by a maritime aerosol type [30] at 90% relative humidity (M90 model). The assumption of unity for the vicarious calibration of the longest NIR wavelength has been widely examined by previous investigators. For example, Wang and Gordon [41] demonstrated through simulation that a radiometric calibration of the 865-nm band to within 10% is sufficient for accurate water-leaving radiance retrievals from SeaWiFS. Various studies have estimated the radiometric calibration for the SeaWiFS 865-nm band to be between 7% low [42] and 8% high [43]. Using aerosol information derived from CIMEL Sun photometers, Franz *et al.* [29] determined the vicarious calibration gain to be approximately 0.96, suggesting that the SeaWiFS 865-nm observed radiances are overestimated by 4%. In Tables 7 and 8 we examine the sensitivity of the deep-water validation results to changes in the 865-nm calibration of +4% and −4%, respectively. The vicarious calibrations of the 765-nm band and the visible bands were rederived for consistency in each case. The deep-water validation results do not differ significantly from those presented in Table 4, from which we conclude that the atmospheric correction is relatively insensitive to the calibration of the NIR_L band, which is consistent with the expectation of Wang and Gordon [41].

To test the assumption on aerosol type, we repeated the 765-nm vicarious calibration using two alternative aerosol models: a purely oceanic aerosol at 99% relative humidity (O99) and a maritime aerosol at 50% relative humidity (M50). These aerosol types represent a range of aerosol size distributions that would be expected in a typical open-ocean location [33], with Ångstrom exponents of −0.1 and 0.5, respectively. For the same aerosol radiance retrieval at 865 nm, the M50 model will yield a higher aerosol radiance in the visible, while the O99 model will yield a lower aerosol radiance. The spectral slope of the M90 model (Ångstrom exponent of 0.2) will generally fall between that of the O99 and M50 models. Tables 9 and 10 show the deep-water validation results for the M50 and O99 calibration assumptions [$\bar{g}(765) = 0.9835$ and $\bar{g}(765) = 0.9594$, respectively with $\bar{g}(865) = 1.0$ and visible gains rederived accordingly]. Again, the results compare favorably to those of Table 4, suggesting that the performance of the satellite-based ocean color retrieval process is relatively insensitive to the aerosol model assumption used in our vicarious calibration, at least for open-ocean conditions where maritime aerosols dominate and aerosol concentrations are relatively low (i.e.,

Table 8. Sensitivity of Deep-Water Validation to −4% 865-nm Calibration; $\bar{g}(765) = 0.9420$

$\bar{g}(\lambda)$	Ratio[a]	MPD[a]	r^2	N[b]
1.0337	1.005	11.0	0.932	188
1.0090	0.954	15.2	0.876	318
0.9856	0.940	12.4	0.822	318
0.9908	0.953	11.2	0.591	164
0.9885	0.958	14.0	0.837	318
0.9527	1.376	57.0	0.604	306
	0.987	24.7	0.875	149

[a]As defined in Table 3.
[b]Number of satellite-to-*in situ* match-up cases (a common set was used for Tables 6–10 and for the analysis presented in Table 4).

Table 7. Sensitivity of Deep-Water Validation to Calibration Assumptions +4% 865-nm Calibration; $\bar{g}(765) = 1.0007$

	$\bar{g}(\lambda)$	Ratio[a]	MPD[a]	r^2	N[b]
$L_{wn}(412)$	1.0405	0.998	11.8	0.927	188
$L_{wn}(443)$	1.0177	0.953	14.8	0.871	318
$L_{wn}(490)$	0.9982	0.944	12.5	0.812	318
$L_{wn}(510)$	1.0059	0.965	10.8	0.573	164
$L_{wn}(555)$	1.0099	0.977	15.3	0.822	318
$L_{wn}(670)$	0.9933	1.416	61.4	0.596	306
C_a		0.980	26.2	0.873	149

[a]As defined in Table 3.
[b]Number of satellite-to-*in situ* match-up cases (a common set was used for Tables 6–10 and for the analysis presented in Table 4).

Table 9. Sensitivity of Deep-Water Validation to M50 765-nm Calibration; $\bar{g}(765) = 0.9835$

	$\bar{g}(\lambda)$	Ratio[a]	MPD[a]	r^2	N[b]
$L_{wn}(412)$	1.0467	0.995	10.8	0.931	188
$L_{wn}(443)$	1.0245	0.955	16.0	0.879	318
$L_{wn}(490)$	1.0050	0.934	13.4	0.826	318
$L_{wn}(510)$	1.0131	0.953	10.6	0.608	164
$L_{wn}(555)$	1.0164	0.961	14.3	0.844	318
$L_{wn}(670)$	0.9915	1.399	53.5	0.599	306
C_a		1.007	26.0	0.871	149

[a]As defined in Table 3.
[b]Number of satellite-to-*in situ* match-up cases (a common set was used for Tables 6–10 and for the analysis presented in Table 4).

Table 10. Sensitivity of Deep-Water Validation to O99 765-nm Calibration; $\bar{g}(765) = 0.9594$

$\bar{g}(\lambda)$	Ratio[a]	MPD[a]	r^2	N[b]
1.0285	0.999	12.1	0.925	188
1.0031	0.957	14.2	0.866	318
0.9793	0.947	12.2	0.804	318
0.9840	0.968	10.6	0.559	164
0.9825	0.967	14.9	0.805	318
0.9544	1.341	63.8	0.594	306
	0.943	25.5	0.872	149

[a]As defined in Table 3.
[b]Number of satellite-to-*in situ* match-up cases (a common set was used for Tables 6–10 and for the analysis presented in Table 4).

aerosol optical thickness generally less than 0.3 at 500 nm [33]).

6. Conclusions

For satellite-borne ocean color sensors, we defined the vicarious calibration as the set of gain factors that, when applied to the observed TOA radiances, force the combined instrument response and atmospheric correction process to reproduce the expected water-leaving radiance at the sea surface. The vicarious calibration approach we described is independent of the source of the expected water-leaving radiance, but it is strictly specific to the atmospheric correction algorithm and relative to the instrument calibration.

We showed that the operational vicarious calibration of SeaWiFS, targeting two sites in the South Pacific and southern Indian Ocean to calibrate the NIR bands and MOBY measurements to calibrate the visible bands, enables the satellite sensor + algorithm system to accurately reproduce the water-leaving radiances from MOBY and significantly improves the quality of ocean color retrievals relative to globally distributed open-ocean field measurements. Furthermore, the calibration gains were shown to be stable over time and over the range of viewing and solar geometries encountered at the calibration sites, suggesting that the instrument prelaunch characterization, temporal calibration, and atmospheric correction model are correctly compensating for changes in the instrument and variations associated with radiant-path geometry. In addition, the mean vicarious gains were found to stabilize to within 0.1% of the final values after approximately 20–40 high-quality calibration samples, which provides some indication as to the extent of on-orbit calibrations that may be required by future missions to achieve similar quality in ocean color retrievals.

We also demonstrated that the ocean color retrievals are relatively insensitive to the NIR calibration (i.e., both the 765- and 865-nm bands of SeaWiFS) that governs the determination of aerosol contributions. This occurs because the vicarious calibration of the visible bands will tend to compensate for bias and spectral skew associated with the aerosol radiance retrievals, provided that the calibration process is internally consistent. In the OBPG approach, the visible bands are calibrated relative to the NIR bands, with the shorter NIR band calibrated relative to the longer NIR band, and all vicarious calibrations are performed relative to the exact same atmospheric correction algorithm and instrument calibration used for the retrieval of normalized water-leaving radiance from observed TOA radiance. It follows that, if the atmospheric correction process is altered in any way, the vicarious calibration must be regenerated to maintain consistency. In fact, the OBPG routinely rederives the vicarious calibration from the same set of MOBY and NIR targets when testing the performance of algorithm changes such as alternate aerosol model suites or modified bidirectional reflectance formulations. Similarly, the vicarious calibration is intimately tied to the instrument calibration. For example, it is likely that the OBPG will update the SeaWiFS prelaunch calibration prior to the next major reprocessing, based on the findings of Barnes and Zalewski [40], and the vicarious calibration will have to be recomputed relative to that revised instrument calibration. It follows that the vicarious gains produced by the OBPG for ocean color processing of SeaWiFS, MODIS, and other ocean color sensors are only valid for the operational atmospheric correction algorithm and instrument calibration.

For all ocean color sensors supported by the OBPG, the operational vicarious calibrations were derived using open-ocean calibration targets and typical maritime atmospheric conditions. Considering that the vicarious calibration will absorb systematic bias in the atmospheric correction algorithm, it should be recognized that a set of fixed multiplicative factors cannot adjust for algorithm deficiencies over the full range of oceanic and atmospheric conditions and radiant-path geometries associated with global observations, as any error in the atmospheric correction model is not likely to be fractionally constant relative to the TOA radiances. As the observation conditions deviate from the nominal conditions and geometries of the vicarious calibration targets, the ability of the vicarious calibration to compensate for algorithm error will be diminished. For example, in some coastal and inland waters, the typical aerosol may not be well represented by the operational aerosol model suite or model selection process. Similarly, these more complex cases may trigger components of the atmospheric correction algorithm that are rarely exercised over open oceans, such as the correction for nonzero L_w(NIR) in turbid or eutrophic waters [20]. As such, until perfect algorithms exist, it may be necessary to perform regionally specific vicarious calibrations to obtain accurate ocean color retrievals in coastal and inland waters. The vicarious gains derived by the OBPG are designed to optimize the fidelity of ocean color measurements in the deep oligotrophic and mesotrophic waters that comprise the vast majority of the world oceans.

This work was funded by NASA Earth Observing Systems (EOS)/MODIS and NASA Ocean Biogeo-

chemistry Programs, and we are grateful for the support of the OBPG Staff. We also acknowledge the MOBY Operations Team, and D. Clark of the National Oceanic and Atmospheric Administration for the development and maintenance of MOBY and for processing the MOBY measurements utilized for the vicarious calibration of SeaWiFS.

References

1. C. R. McClain, M. L. Cleave, G. C. Feldman, W. W. Gregg, S. B. Hooker, and N. Kuring, "Science quality seawifs data for global biosphere research," Sea Technol. **39,** 10–16 (1998).
2. W. E. Esaias, M. R. Abbot, I. J. Barton, O. B. Brown, J. W. Campbell, K. L. Carder, D. K. Clark, R. H. Evans, F. E. Hoge, H. R. Gordon, W. M. Balch, R. Letelier, and P. J. Minnett, "An overview of MODIS capabilities for ocean science observations," IEEE Trans. Geosci. Remote Sens. **36,** 1250–1265 (1998).
3. B. C. Gao, M. J. Montes, Z. Ahmad, and C. O. Davis, "Atmospheric correction algorithm for hyperspectral remote sensing of ocean color from space," Appl. Opt. **39,** 887–896 (2000).
4. D. Antoine and A. Morel, "A multiple scattering algorithm for atmospheric correction of remotely sensed ocean colour (meris instrument): principle and implementation for atmospheres carrying various aerosols including absorbing ones," Int. J. Remote Sens. **20,** 1875–1916 (1999).
5. H. R. Gordon and M. Wang, "Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: A preliminary algorithm," Appl. Opt. **33,** 443–452 (1994).
6. H. R. Gordon, "In-orbit calibration strategy for ocean color sensors," Remote Sens. Environ. **63,** 265–278 (1998).
7. R. E. Eplee Jr., W. D. Robinson, S. W. Bailey, D. K. Clark, P. J. Werdell, M. Wang, R. A. Barnes, and C. R. McClain, "Calibration of SeaWiFS. II. vicarious techniqes," Appl. Opt. **40,** 6701–6718 (2001).
8. S. B. Hooker, W. E. Esaias, G. C. Feldman, W. W. Gregg, and C. R. McClain, "An overview of SeaWiFS and ocean color," NASA Tech. Memo. 104566, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD (1992).
9. D. K. Clark, H. R. Gordon, K. J. Voss, Y. Ge, W. W. Broenkow, and C. Trees, "Validation of atmospheric correction over the oceans," J. Geophys. Res. **102,** 17209–17217 (1997).
10. F. S. Patt, R. A. Barnes, R. E. Eplee Jr., B. A. Franz, W. D. Robinson, G. C. Feldman, S. W. Bailey, J. Gales, P. J. Werdell, M. Wang, R. Frouin, R. P. Stumpf, R. A. Arnone, J. R. W. Gould, P. M. Martinolich, V. Ransibrahmanakul, J. E. O'Reilly, and J. A. Yoder, "Algorithm updates for the Fourth SeaWiFS data reprocessing," NASA Tech. Memo. 206892, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD (2003).
11. B. A. Franz, P. J. Werdell, G. Meister, S. W. Bailey, R. E. Eplee, G. C. Feldman, E. Kwiatkowska, C. R. McClain, F. S. Patt, and D. Thomas, "The continuity of ocean color measurements from SeaWiFS to MODIS" **5882,** 1–13 (2005).
12. M. Wang, "The Rayleigh lookup tables for the SeaWIFS data processing: accounting for the effects of ocean surface roughness," Int. J. Remote Sens. **23,** 2693–2702 (2002).
13. M. Wang, "A refinement for the Rayleigh radiance computation with variation of the atmospheric pressure," Int. J. Remote Sens. **26,** 5651–5653 (2005).
14. H. R. Gordon and M. Wang, "Influence of oceanic whitecaps on atmospheric correction of SeaWiFS," Appl. Opt. **33,** 7754–7763 (1994).
15. R. Frouin, M. Schwindling, and P. Y. Dechamps, "Spectral reflectance of sea foam in the visible and near infrared: in situ measurements and remote sensing implications," J. Geophys. Res. **101,** 14361–14371 (1997).
16. K. D. Moore, K. J. Voss, and H. R. Gordon, "Spectral reflectance of whitecaps: their contribution to water-leaving radiance," J. Geophys. Res. **105,** 6493–6499 (2000).
17. M. Wang, "Atmospheric correction of ocean color sensors: computing atmospheric diffuse transmittance," Appl. Opt. **38,** 451–455 (1999).
18. H. R. Gordon, T. Du, and T. Zhang, "Atmospheric correction of ocean color sensors: analysis of the effects of residual instrument polarization sensitivity," Appl. Opt. **36,** 6938–6948 (1997).
19. G. Meister, E. J. Kwiatkowska, B. A. Franz, F. S. Patt, G. C. Feldman, and C. R. McClain, "Moderate-resolution imaging spectroradiometer ocean color polarization correction," Appl. Opt. **44,** 5524–5535 (2005).
20. R. P. Stumpf, R. A. Arnone, J. R. W. Gould, P. M. Martinolich, and V. Ransibrahmanakul, "A partially coupled ocean-atmosphere model for retrieval of water-leaving radiance from seawifs in coastal waters," NASA Tech. Memo. 206892, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD (2003).
21. H. R. Gordon, "Clear water radiances for atmospheric correction of coastal zone color scanner imagery," Appl. Opt. **20,** 4175–4180 (1981).
22. *Astronomical Almanac* (U.S. Government Printing Office, 1984).
23. A. Morel, D. Antoine, and B. Gentilli, "Bidirectional reflectance of oceanic waters: accounting for Ramen emission and varying particle scattering phase function," Appl. Opt. **41,** 6289–6306 (2002).
24. H. R. Gordon, "Normalized water-leaving radiance: revisiting the influence of surface roughness," Appl. Opt. **44,** 241–248 (2005).
25. M. Wang, "Effects of ocean surface reflectance variation with solar elevation on normalized water-leaving radiance," Appl. Opt. **45,** 4122–4128 (2006).
26. M. Wang, B. A. Franz, R. A. Barnes, and C. R. McClain, "Effects of spectral bandpass on SeaWiFS-retrieved near-surface optical properties of the ocean," Appl. Opt. **40,** 343–348 (2001).
27. H. R. Gordon, "Remote sensing of ocean color: a methodology for dealing with broad spectral bands and significant out-of-band response," Appl. Opt. **34,** 8363–8374 (1995).
28. J. E. O'Reilly, S. Maritorena, D. A. Siegel, M. C. O'Brien, D. Toole, B. G. Mitchell, M. Kahru, F. P. Chavez, P. Strutton, G. F. Cota, S. B. Hooker, C. R. McClain, K. L. Carder, F. Muller-Karger, L. Harding, A. Magnuson, D. Phinney, G. F. Moore, J. Aiken, K. R. Arrigo, R. Letelier, and M. Culver, "Ocean color chlorophyll a algorithms for SeaWiFS, OC2, and OC4: Version 4," NASA Tech. Memo. 206892, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD (2000).
29. B. A. Franz, E. J. Ainsworth, and S. W. Bailey, "SeaWiFS vicarious calibration: an alternative approach utilizing simultaneous in situ observations of oceanic and atmospheric optical properties," NASA Tech. Memo. 209982, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD (2001).
30. E. P. Shettle and R. W. Fenn, *Models for the Aerosols of the Lower Atmosphere and the Effects of Humidity Variations on their Optical Properties,* vol. ADA085951 of *Environmental Research Papers* (Air Force Geophysics Laboratory Hanscom AFB, 1979).
31. A. Morel, B. Gentili, H. Claustre, M. Babin, A. Bricaud, J. Ras, and F. Tieche, "Optical properties of the 'clearest' natural waters," Limnol. Oceanogr. **52,** 217–229 (2007).
32. B. Fougnie, P. Henry, A. Morel, D. Antoine, and F. Montagner, "Identification and characterization of stable homogeneous

oceanic zones: climatology and impact on in-flight calibration of space sensors over Rayleigh scattering," in *Proceedings of Ocean Optics XVI* (Oceanography Society, 2002).
33. A. Higurashi, T. Nakajima, B. N. Holben, A. Smirnov, R. Frouin, and B. Chatenet, "A study of global aerosol optical climatology with two-channel avhrr remote sensing," J. Clim. **13,** 2011–2027 (2000).
34. J. L. Mueller, D. K. Clark, V. S. Kuwahara, G. Lazin, S. Brown, G. S. Fargion, M. A. Yarbrough, M. Feinholz, S. Flora, W. Broenkow, Y. S. Kim, B. C. Johnson, M. Yuen, P. G. Strutton, T. D. Dickey, M. R. Abbott, R. M. Letelier, M. R. Lewis, S. McLean, F. P. Chavez, J. R. M. A. Barnard, A. Subramaniam, D. Manov, X. Zheng, J. L. W. Harding, R. A. Barnes, and K. R. Lykke, "Ocean optics protocols for satellite ocean color sensor validation, Revision 4: Special topics in ocean optics protocols and appendices," NASA Tech. Memo. 2004-21621, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD (2003).
35. R. Frouin, D. W. Ligner, and C. Gautier, "A simple analytical formula to compute clear sky total and photosynthetically available solar irradiance at the ocean surface," J. Geophys. Res. **94,** 9731–9742 (1989).
36. S. W. Bailey and P. J. Werdell, "A multi-sensor approach for the on-orbit validation of ocean color satellite data products," Remote Sens. Environ. **102,** 12–23 (2006).
37. R. A. Barnes, R. E. Eplee, F. S. Patt, H. H. Kieffer, T. C. Stone, G. Meister, J. J. Butler, and C. R. McClain, "Comparison of SeaWiFS measurements of the moon with the U.S. geological survey lunar model," Appl. Opt. **43,** 5838–5854 (2002).
38. J. L. Mueller, A. Morel, R. Frouin, C. Davis, R. Arnone, K. Carder, Z. Lee, R. G. Steward, S. B. Hooker, C. D. Mobley, S. McLean, B. Holben, M. Miller, C. Pietras, K. D. Knobelspiesse, G. S. Fargion, J. Porter, and K. Voss, "Ocean optics protocols for satellite ocean color sensor validation, Revision 4, Volume III: Radiometric measurements and data analysis protocols," NASA Tech. Memo. 2003-21621, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD (2003).
39. P. J. Werdell, S. W. Bailey, G. Fargion, C. Pietras, K. Knobelspiesse, G. Feldman, and C. McClain, "Unique data repository facilitates ocean color satellite validation," EOS Trans. Am. Geophys. Union **84** (2003) p. 379 and 385.
40. R. A. Barnes and E. F. Zalewski, "Reflectance-based calibration of seawifs II. conversion to radiance," Appl. Opt. **42,** 1648–1660 (2003).
41. M. Wang and H. R. Gordon, "Calibration of ocean color scanners: how much error is acceptable in the near infrared?" Remote Sens. Environ. **82,** 497–504 (2002).
42. N. Martiny, R. Frouin, and R. Santer, "Radiometric calibration of SeaWiFS in the near infrared," Appl. Opt. **44,** 7828–7844 (2005).
43. R. Green and T. G. Chrien, "High altitude measurements of radiance at high spectral and spatial resolution for SIMBIOS sensor calibration, validation and intercomparisons," NASA Tech. Memo. 208645, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD (1999).