Sensitivity of Satellite Ocean Color Data to System Vicarious Calibration of the Long Near Infrared Band

Brian B. Barnes, Chuanmin Hu, Sean W. Bailey, and Bryan A. Franz

***Abstract*— Satellite ocean color missions require accurate system vicarious calibrations (SVC) to retrieve the relatively small remote-sensing reflectance (R_{rs}, sr^{-1}) from the at-sensor radiance. However, the current atmospheric correction and SVC procedures do not include calibration of the "long" near infrared band (NIR_L—869 nm for MODIS), partially because earlier studies, based primarily on simulations, indicate that accuracy in the retrieved R_{rs} is insensitive to moderate changes in the NIR_L vicarious gain (g). However, the sensitivity of ocean color data products to $g(NIR_L)$ has not been thoroughly examined. Here, we first derive 10 SVC "gain configurations" (vicarious gains for all visible and NIR bands) for MODIS/Aqua using current operational NASA protocols, each time assuming a different $g(869)$. From these, we derive a suite of ∼1.4E6 unique gain configurations with $g(869)$ ranging from 0.85 to 1.2. All MODIS/A data for 25 locations within each of five ocean gyres were then processed using each of these gain configurations. Resultant time series show substantial variability in dominant $R_{rs}(547)$ patterns in response to changes in $g(869)$ (and associated gain configurations). Overall, mean $R_{rs}(547)$ values generally decrease with increasing $g(869)$, while the standard deviations around those means show gyre-specific minima for $0.97 < g(869) < 1.02$. Following these sensitivity analyses, we assess the potential to resolve $g(869)$ using such time series, finding $g(869) = 1.025$ most closely comports with expectations. This approach is broadly applicable to other ocean color sensors, and highlights the importance of rigorous cross-sensor calibration of the NIR_L bands, with implications on consistency of merged-sensor data sets.**

***Index Terms*— MODIS/Aqua, ocean color, ocean gyres, SeaWiFS, vicarious calibration.**

TABLE I
DEFINITIONS OF FREQUENTLY USED ACRONYMS AND SYMBOLS

Abbreviation	Symbol
SVC	System Vicarious Calibration
NIR_L, NIR_S	"Long" and "Short" wavebands in the Near Infrared
VIS	Visible wavelengths
MOBY	Marine Optical Buoy
OBPG	NASA Ocean Biology Processing Group
SeaDAS, L2GEN	NASA software, module for processing data
L1A, GEO, L1B, L2	Intermediate data processing levels
XCAL	File containing instrument calibration coefficients
g	Vicarious calibration coefficient; gain; g-factor
F2007	Gain configurations calculated using [5] (Table II)
ρ, R_{rs}	Reflectance, Remote sensing reflectance
C_a	Chlorophyll-a concentration
CDOM	Colored dissolved organic matter concentration
L_t	Total at-sensor radiance
L_w	Water-leaving radiance
L_a	Aerosol contributions to L_t
ε	$L_a(NIR_S)/L_a(NIR_L)$
α	Ångström exponent
τ_a	Aerosol optical thickness (AOT)
BRDF	Bidirectional reflectance distribution function
OOB	Out of band
$\bar{x}_{Rrs}$, σ_{Rrs}	Mean, standard deviation of all $R_{rs}(547)$ data within a time series
x_t, y_t	Measured, modeled $R_{rs}(547)$ at time t
i	Interval (days) of calculated moving means
σ_{mod}	Standard deviation of y_{ti} for a particular model

I. INTRODUCTION

SATELLITE ocean color measurements are critical in assessing global ocean conditions, with a sampling frequency impossible to reach using traditional *in situ* methods. Among the key challenges in processing and interpreting satellite ocean color data are ensuring that satellite-measured signals, particularly at-sensor total radiance (L_t, mW cm^{-2} μm^{-1} sr^{-1}), are properly calibrated and that the retrieved ocean remote sensing reflectance (R_{rs}, sr^{-1}) or water leaving radiance (L_w, mW cm^{-2} μm^{-1} sr^{-1}) are validated (see Table I for list of symbols and acronyms). "Calibration" in this context includes a combination of procedures designed to: 1) ensure stability of the satellite data record in space, time, and with varying sensor satellite geometry, and 2) minimize differences between satellite measurements and "known" quantities, including satellite observations of solar or lunar radiance and *in situ* R_{rs} measurements of ocean targets. This comparison to *in situ* R_{rs} is accomplished through a process called "system vicarious calibration" (SVC, see Section II-B), with uncertainties in SVC being propagated throughout an instrument's entire measurement data set—potentially leading to cross-sensor discrepancies.

As such, an enormous amount of effort has been directed toward ensuring calibration of satellite-measured L_t and validation of the derived data products. SVC, in particular, is critical in assuring accurate retrievals of L_w for two reasons: 1) L_w for most ocean waters only contributes a small portion

Manuscript received February 3, 2020; revised April 28, 2020; accepted June 2, 2020. This work was supported by the National Aeronautics and Space Administration (NASA) under Grant NNX16AQ71G. *(Corresponding author: Brian B. Barnes.)*

Brian B. Barnes and Chuanmin Hu are with the College of Marine Science, University of South Florida, St. Petersburg, FL 33701 USA (e-mail: bbarnes4@usf.edu).

Sean W. Bailey and Bryan A. Franz are with the NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA.

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2020.3000475

of L_t [1], thus a 1% uncertainty in L_t may thus translate to 10% or higher uncertainties in L_w; and 2) the retrieval of L_w from L_t is through a process called atmospheric correction, which is based on radiative-transfer simulations [2] that may contain uncertainties. SVC includes the process to force the satellite-measured L_t to agree with the estimated L_t using the same radiative-transfer procedure as the atmospheric correction [2]–[6]. Therefore, uncertainties in the atmospheric correction may be compensated, at least to first order. Numerous validation efforts for satellite-derived R_{rs} products show strong agreement between satellite/*in situ* [7]–[14] and satellite/satellite [15]–[20] data, indicating high fidelity of both the SVC and the atmospheric correction procedures used to process the measured L_t.

Despite these successes, some assumptions within currently operational atmospheric correction and SVC procedures require further assessment. In particular, for practical reasons, the SVC coefficient (termed "gain" or "g-factor," g) for the "long" near infrared (NIR_L) band [$g(NIR_L)$] is taken as 1.0, while in reality it could be slightly different. This band is expected to be "correctly" calibrated, within acceptable uncertainties, from prelaunch measurements. Radiative-transfer simulations on point data by Wang and Gordon [21] suggest that calibration errors in this NIR_L band of up to 5% in either direction, after subsequent SVC of other bands, have little impact on the retrieved R_{rs} in visible bands. To our knowledge, however, the impacts of this assumption on actual satellite data have yet to be rigorously assessed.

Several additional SVC efforts have been conducted using *in situ* measurements and atmospheric models that differ from those used for operational SVC. For example, using data derived from AERONET-OC (Aerosol Optical Network with additional ocean color components), both Mélin and Zibordi [22] and Hlaing *et al.* [23] independently performed SVC of the NIR_L band for extant ocean color sensors, deriving $g(NIR_L)$ which widely deviated from 1.0 for three separate satellite sensors. As these works were primarily intended to understand coastal atmospheric correction functioning and limitations, only specific coastal aerosol models were considered, meaning these $g(NIR_L)$ may not be applicable to open ocean targets. Also, the aerosol models [24] used in Mélin and Zibordi [22] differ from those within current operational processing [25]. Switching to these models, as was done in NASA calibration update R2009.0, improved satellite aerosol radiance retrievals [22]. Additional efforts to calibrate NIR_L bands using measurements of sky radiance distributions via sun photometers [26]–[28] typically include uncertainties which are too large (>2%) for SVC.

In early 2015, the NASA Ocean Biology Processing Group (OBPG) used SeaWiFS (Sea-viewing Wide Field of View Sensor onboard the OrbView-2 satellite) data to perform a preliminary investigation of the potential impact of changing $g(NIR_L)$ on global time series (https://oceancolor.gsfc.nasa.gov/analysis/global/st117_st115/; https://oceancolor.gsfc.nasa.gov/analysis/global/st116_st115/). Briefly, the entire SeaWiFS mission data set was processed three times, each time assuming a different prelaunch calibration of the NIR_L band (865nm for SeaWiFS): 0.95, 1.0 (default), and 1.05.

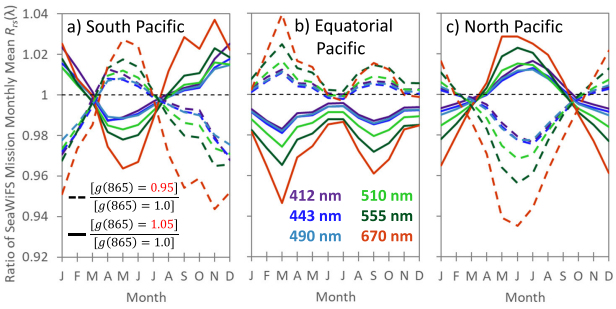


Fig. 1. Ratios between mission-long monthly SeaWiFS $R_{rs}(\lambda)$ means derived with three $g(865)$ from (a) South Pacific (30–40S, 179–140W), (b) Equatorial Pacific (10S–10N, 179–140W), and (c) North Pacific (30–40N, 179–140W). The processing that derived these R_{rs} data sets differed only in the NIR_L gain used as a starting point for SVC, either 0.95 (numerator for dotted lines), 1.05 (numerator for solid lines), or 1.0 (denominator for all lines). Color indicates waveband (nm). All other gains also varied following SVC. Note the seasonally-, radiometrically-, and geographically-dependent disagreement between these data sets. For these calculations, year-specific monthly binned files for which both test cases [e.g., $g(865) = 0.95$ and $g(865) = 1.0$, dotted lines] contained valid data were used to generate overall mission monthly means. The ratios of these mission monthly means thus include only 4 km bins with common data at monthly resolution.

From these data, two comparisons were performed: 0.95 vs 1.0 and 1.0 vs 1.05 (Fig. 1). In each comparison, mission monthly means encompassing several regions (Fig. 1 shows the South, Equatorial, and North Pacific oceans) were calculated. The ratio between the means as calculated from these data sets shows systematic disagreements (Fig. 1).

Overall, Fig. 1 indicates that, in practice, a 5% change in $g(NIR_L)$ can indeed lead to large-scale changes in the derived R_{rs}. It is noted that differences between these monthly means are generally within ±2–3% for most bands, similar to residual uncertainties reported by Wang and Gordon [21]. Thus, while the NIR_L calibration assumption (i.e., insensitivity of visible R_{rs} to ±5% errors in NIR_L calibration) appears valid within acceptable uncertainties for individual point data, errors in $g(NIR_L)$ on large scale, aggregated data sets may be introducing substantial seasonally- and geographically-dependent biases.

At the time of this writing, however, the SeaWiFS data used in Fig. 1 are somewhat dated, as they correspond to NASA reprocessing version 2010.1. It is unknown whether such disagreements exist using the current calibration and processing versions, or in other satellite data sets. Unfortunately, validation-quality *in situ* R_{rs} measurements are few and sparsely distributed in global oceans, with matchups often having higher uncertainties than the ±2%–3% observed in Fig. 1 [29], [30]. This means that large-scale assessment of the NIR_L calibration assumption cannot be completed using satellite/*in situ* matchups.

Nevertheless, potential systematic biases introduced by the NIR_L calibration assumption could complicate our ability to capture climactic variability in ocean gyres—topics that are only accessible using long-term satellite ocean color time series [31]–[33]. As such, the objective of this work is to more extensively investigate the sensitivity of global satellite data to changes in $g(NIR_L)$. To accomplish this, we investi-

gated large-scale R_{rs} time series in ocean gyres from Moderate Resolution Imaging Spectroradiometer onboard Aqua (MODIS/A), in response to changes in $g(\mathrm{NIR_L})$. From these data, we asses the potential to derive an "optimal" $g(\mathrm{NIR_L})$ for future data processing. Ocean gyres were desirable for these analyses due to 1) spatiotemporal homogeneity as compared to coastal oceans; 2) likely validity of the black pixel assumption in atmospheric correction [4]; and 3) primarily oceanic atmospheric aerosols.

II. Background of Satellite Data Processing

A. Satellite Sensor Calibration

The NASA OBPG is tasked with operational processing of MODIS data (as well as data from other satellite ocean color sensors), which includes efforts to ensure calibration of the sensor itself. Calibration, in this sense, refers to the combination of "instrument calibration" and SVC. The former is a broad term that includes prelaunch calibration and ongoing efforts to ensure within-sensor consistency. Such within-sensor consistency characterizations are necessary to correct changes in sensor responsiveness resulting from differences in sensor temperature, detector (MODIS has ten independent detectors), mirror side, and scan angle. Temporal degradation of the sensor is also corrected through instrument calibration. Instrument calibration coefficients are thus time-, detector-, mirror-side-, and scan-angle-specific, and are applied (multiplied) to Level-1A measured at-sensor radiance values to create Level-1B calibrated at-sensor total radiance (L_t) data.

Details on the OBPG SVC and atmospheric correction system, as necessary to understand the former, are extensively documented in other works [2]–[6], [34]. Here, we provide a very brief summary as a basis to explain the rationale and approach of this work.

B. Atmospheric Correction and SVC

The overarching goal of atmospheric correction is to derive water-leaving radiance (L_w) from satellite-measured L_t in the equation

$$L_t = \left[L_r + L_a + TL_g + tL_{wc} + tL_w\right] t_g f_p \tag{1}$$

where L_r, L_a, L_g, and L_{wc} are radiance contributions from molecular (Rayleigh) scattering, aerosol scattering (including Rayleigh-aerosol multiple scattering), glint, and white caps, respectively. T and t are Rayleigh-aerosol direct and diffuse transmittance from the surface to the sensor, respectively, while t_g is transmittance due to bidirectional gaseous absorption, and f_p is a correction parameter for polarization. Although all of these parameters are spectrally dependent, wavelength notation is omitted for simplicity. For SVC work, pixels with L_g contributions are excluded. Most other parameters (L_r, L_{wc}, t_g, f_p) are well characterized using models and ancillary data [4], [5], [35]–[38], leaving L_a as the primary unknown in solving for L_w, as t can be derived once L_a is known. For very clear ocean targets, it is assumed that water absorbs all near-infrared (NIR) light, thus $L_w(\mathrm{NIR}) = 0$ and $L_a(\mathrm{NIR})$ can be directly calculated. The ratio (ε, unitless) between L_a at two NIR bands ($\mathrm{NIR_S}$ and $\mathrm{NIR_L}$—short and long, respectively) is used to select an aerosol model from among several aerosol types using precalculated look up tables (LUTs). Such selected aerosol models [24], [25] are subsequently used to extrapolate L_a from the NIR to the visible bands (VIS), leading to derivation of $L_w(\mathrm{VIS})$. This process is carried out for each satellite pixel.

For the VIS bands, the reverse process of atmospheric correction is used in SVC: estimating L_t from *in situ* L_w using the same radiative-transfer procedure as in the atmospheric correction. The estimated L_t is compared with the satellite-measured L_t, with the average of their ratio being termed a "vicarious gain" (g). In practice, for SVC of the VIS bands, this process is carried out using many L_w measurements from the Marine Optical Buoy (MOBY), located at a clear water site 20 km west of Lanai, Hawaii, USA. For each band, the vicarious gain is used to adjust satellite-measured L_t before atmospheric correction for the entire mission. As such, SVC includes correction of biases in the instrument calibration and the combined instrument-atmospheric correction system.

Operational SVC of the NIR bands, however, is not performed using MOBY. Recall that during atmospheric correction, $L_a(\mathrm{VIS})$ must first be estimated to propagate $L_w(\mathrm{VIS})$ to $L_t(\mathrm{VIS})$. Since $L_a(\mathrm{VIS})$ is calculated from $L_t(\mathrm{NIR})$, $g(\mathrm{NIR})$ must be established before SVC of the VIS bands. Instead, the approach to vicariously calibrate these two NIR bands uses satellite data from extremely oligotrophic and remote ocean waters, where it is assumed that $L_w(\mathrm{NIR}) = 0$ [4] and that the atmospheric aerosols result only from oceanic processes, meaning the aerosol type is known and well modeled. Finally, it is assumed that $g(\mathrm{NIR_L}) = 1.0$, implying that the instrument calibration for this band is "correct," or that errors in this calibration do not appreciably contribute to uncertainties in $R_{rs}(\mathrm{VIS})$ [21]. Thus, knowing the aerosol type and geometry, the "correct" ε should also be known, and $g(\mathrm{NIR_s})$ is the coefficient which forces the satellite ratio to match expected values.

Because the SVC gains are used for the entire satellite mission, extreme care is taken to ensure that the satellite and *in situ* data used in SVC efforts are of the highest quality. Nevertheless, *in situ* R_{rs} often contain uncertainties at least encompassing the target accuracy for satellite retrievals [29], [30]. Current OBPG gains (calibration 2018.0) for the MODIS/A 412-869nm bands range from 0.9791 to 1.0197, with a mean ($\bar{x}$) $\pm$ standard deviation (σ) of 1.0029 ± 0.0117 and $g(869) = 1.0000$.

III. Approach

Our approach to determine the impact of changing $g(\mathrm{NIR_L})$ uses the current OBPG atmospheric correction procedure, and relies on time series patterns of R_{rs} (547) in ocean gyres. We selected R_{rs} (547) as a target for these analysis due, partially, to a presumed relationship with C_a (chlorophyll-a concentration, in mg $\mathrm{m^{-3}}$). Dynamics of many in-water properties (Fig. 2) in these regions have been extensively studied using satellite data [31]–[33] with $C_a \ll 0.25$ mg $\mathrm{m^{-3}}$ and generally showing smooth and regular seasonal patterns. Analyzing C_a directly, however, would require a study design

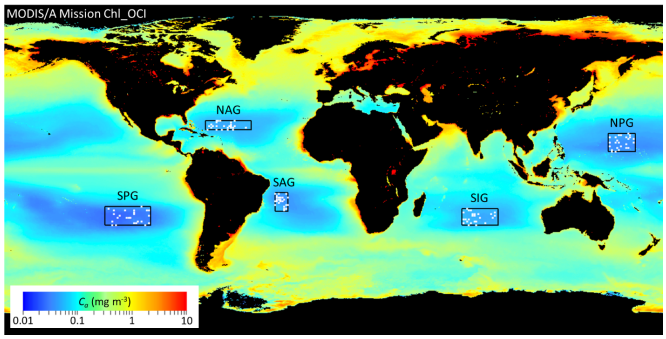


Fig. 2. MODIS/A chlor_a [39] mission composite (https://oceancolor.gsfc.nasa.gov/l3/) overlain with black squares delineating the five regions assessed in this study. Within each region, white dots indicate 25 randomly selected stations from which 17 years of data were processed.

TABLE II
GEOGRAPHIC INFORMATION FOR THE FIVE STUDIED OLIGOTROPHIC REGIONS (MODIFIED FROM [42])

Region	Notation	Longitude Range (°E)	Latitude Range (°N)	Area (10^6 km^2)
South-Sargasso Sea	NAG	-70, -45	22, 27	1.404
Mariana Islands Zone	NPG	150, 165	10, 20	1.787
South Indian Gyre	SIG	70, 90	-30, -21	2.004
Easter Island Zone	SPG	-125, -100	-30, -20	2.794
Brazilian Atlantic Gyre	SAG	-32, -25	-22.5, -12.5	0.823

that assessed five wavebands. This is because prior calibration of the 748 and 869 bands is required to propagate L_a(NIR) to the VIS during atmospheric correction, while R_{rs} (443, 547, and 667) are required to calculate C_a [39]. For this research, the 547 (green) band is preferable to bands in the blue (<500 nm) and red (>600 nm) wavelengths bands as R_{rs}(blue) is more strongly influenced by in-water constituents, especially colored dissolved organic matter (CDOM) [40], which may have different (or offset) seasonality than C_a [31]. R_{rs}(red) shows much higher relative uncertainties than R_{rs}(green) [41]. It is noted that as g(547) is a direct offset parameter for R_{rs}(547), it is possible for many different gain configurations of g(547, 748, and 869) to yield the same R_{rs}(547) for a single pixel, meaning it is necessary to perform this assessment on a multitude of satellite pixels.

A. Study Regions

Five gyre regions were selected for analysis, corresponding to those identified in Morel *et al.* [42] as "the clearest ocean waters" (see Fig. 2, Table II). In the upper layers of ocean gyres, C_a dynamics, as observable by satellites, are primarily driven by wind and temperature, whereby winter mixing and summer warming cause deepening and shoaling of the mixed layer, respectively, leading to higher C_a in winter than in summer. The biological mechanism for the higher wintertime C_a, in response to wintertime mixing, is either an increase in nutrient availability [32], [33], [43] or photoacclimation of phytoplankton to a reduced light regime [43]–[45]. Another forcing on C_a may include horizontal nutrient transfer via

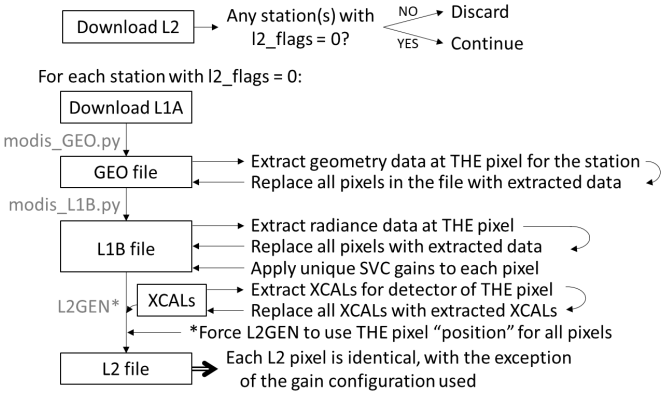


Fig. 3. Flowchart describing satellite data processing, highlighting modifications to intermediate files (squares) and source code (*). Specific SeaDAS modules are listed in gray.

Ekman transport [46]. While the thermodynamically-linked biological response likely does not substantially vary within these ocean gyre centers, Ekman forcings may be more localized to the edges of the study areas, potentially leading to intragyre spatial differences in C_a and thus R_{rs}(547).

Within each region, 25 locations (hereafter termed "stations") were randomly selected from the suite of integer latitude and longitude pairs. Integers were used here to ensure at least one degree latitude or longitude (~110 km) between any two stations. All MODIS/A L2 data from these five regions and spanning the years 2002–2018 (inclusive) were downloaded at Level-2 from NASA OBPG archives (https://oceancolor.gsfc.nasa.gov). Granules were processed as described below if they contained pixels at one or more of the stations which were not identified by any Level-2 Processing flags. These geographically "nearest" pixels were identified by first finding the swath center which passed closest to the station of interest, then identifying the single pixel within that swath nearest to the station.

B. Processing Software

In this work, we sought to perform millions of implementations of the "standard" OBPG atmospheric correction routine for each of a relatively small number of pixels. Such "standard" processing is traditionally done using the L2GEN software within the SeaDAS distribution from NASA OBPG. However, running millions of implementations of L2GEN on each full granule that contained a pixel of interest, where each implementation would lead to a 17-year time series of daily observations, was computationally impossible. Thus, we modified the L2GEN inputs and source code (SeaDAS version 7.5) to allow for the same pixel to be processed many times within a single L2GEN implementation (see flowchart in Fig. 3).

Specifically, once granules were identified using L2 data (Section III-A), the associated Level-1A (L1A) file was downloaded and processed to create a GEO file. All relevant geographic, time, and geometry data fields within this GEO file were modified to match the data from the pixel of interest, resulting in homogenous arrays. The L1A file and the modified

TABLE III

FRANZ *et al.* [5]-DERIVED VICARIOUS GAIN CONFIGURATIONS (HEREAFTER "F2007"). THESE TEN CONFIGURATIONS WERE USED TO FORMULATE 1.4E6 GAIN CONFIGURATIONS, AND EACH CONFIGURATION WAS USED TO DERIVE 17-YEAR $R_{rs}(547)$ TIME SERIES FOR ALL GYRE STATIONS

$g(412)$	$g(443)$	$g(469)$	$g(488)$	$g(531)$	$g(547)$	$g(555)$	$g(645)$	$g(667)$	$g(678)$	$g(748)$	$g(859)$	$g(869)$
0.9699	0.978	0.9919	0.9757	0.9734	0.9691	0.962	0.96	0.9452	0.9404	0.9156	0.8901	0.87
0.9711	0.9796	0.9939	0.9781	0.9768	0.973	0.966	0.9668	0.9526	0.9482	0.9255	0.9039	0.884
0.9748	0.9852	1.0008	0.9863	0.9893	0.9872	0.9806	0.9923	0.9806	0.9774	0.9632	0.9572	0.938
0.9757	0.9863	1.0021	0.9878	0.9912	0.9893	0.9829	0.9959	0.9846	0.9817	0.9687	0.9653	0.946
0.977	0.9876	1.0038	0.9897	0.9941	0.9927	0.9864	1.0026	0.9922	0.9898	0.9794	0.9812	0.962
0.9791	0.9908	1.0079	0.9945	1.0017	1.0014	0.9954	1.0197	1.0109	1.0095	1.0053	1.0184	1*
0.98	0.9923	1.0099	0.9971	1.006	1.0065	1.0007	1.0297	1.0217	1.0209	1.02	1.0399	1.022
0.9804	0.993	1.0112	0.9986	1.0087	1.0096	1.004	1.0367	1.0299	1.0299	1.0327	1.0594	1.042
0.9811	0.9944	1.0133	1.0014	1.0139	1.0157	1.0104	1.0502	1.0456	1.0465	1.0563	1.0961	1.08
0.9812	0.9948	1.0138	1.0022	1.0167	1.0192	1.0143	1.0605	1.0572	1.0587	1.0743	1.1249	1.11

*Current default gain configuration

GEO file were then processed to Level-1B (L1B), which was then modified, akin to the GEO file modifications, to replicate the L_t values of the pixel of interest across the entire granule for all bands. This process resulted in L1B files for which the geometry, geolocation, time, mirror side, and calibrated radiance in all pixels were the same.

For all VIS and NIR bands, artificial "gains" were applied within the L1B files. The specific gains varied for each pixel as described in Section III-C. In practice, to match L2GEN gain application, scaled integers in the L1B file were converted to reflectance (ρ) using the appropriate scale factors and offsets, which were subsequently converted to radiance (L_t) via

$$L_t(\lambda) = \rho(\lambda) * \frac{F_0(\lambda) * (1/d_{es})^2}{\pi} \quad (2)$$

where F_0 is the extraterrestrial solar irradiance [47], and d_{es} is the Earth–Sun distance at the time of the MODIS measurement. Gains were multiplied to L_t, and these L_t values were converted back to ρ, rescaled as integers, and inserted into the modified L1B files. To retain the precision of L_t values as calculated in L2GEN, this rescaling required modification of the HDF (Hierarchical Data Format) file attributes that define scaling coefficients for the integer data within the L1B files.

Within L2GEN, there are four main "positional" elements which are required for assigning the correct cross-calibration and polarization correction coefficients: sensor band, pixel number, detector number, and mirror side. Sensor band required no modification in this context, while mirror side for all pixels had been adjusted to match that of the pixel of interest within the GEO file. Detector number and pixel number are automatically assigned during L2GEN processing. Therefore, for each L1B file, we modified the XCAL files used as ancillary inputs to L2GEN so that all detector-specific coefficients matched those of the detector which collected the pixel of interest. Additionally, as pixel number is used in the arithmetic for application of XCAL and polarization coefficients, we modified the L2GEN source code and recompiled the software such that pixel number could be fixed to an exogenously input value, which we set to match the pixel number for the pixel of interest. This modified L2GEN was then run on the modified L1B and GEO files, using the modified XCAL files. Otherwise, default L2GEN settings were used, with the exception of all gains being set to 1.0 (recall that artificial gains had already been applied). After completion of the L2GEN processing, the end result was thus an L2 "granule" whereby the same original satellite pixel had been processed millions of times with slightly varying spectral gain coefficients. It is noted that for the condition which used the current default gain configuration, this procedure exactly reproduced the corresponding data in NASA OBPG L2 files.

C. OBPG SVC Gains

While most 2018.0 MODIS/A gain coefficients are within 1%–2% of the instrument calibration (i.e., $0.98 < g < 1.02$), for this study we assessed a much larger range of possible $g(\mathrm{NIR_L})$ values ($0.85 < g < 1.2$). Recall that a combination of g(547, 748, and 869) are directly required to construct an $R_{rs}(547)$ value. With OBPG gains currently reported to 0.0001 precision, testing the entire suite of possibilities for all three bands independently would yield ~4E10 configurations, which would be computationally impractical. As such, we constrained the gain combinations tested to those which would be possible using the current MOBY-based SVC procedure. Specifically, the standard OBPG SVC procedure [5] was performed 10 times, each starting with an arbitrarily selected $g(869)$ value. According to the process summarized in Section II-B, for each of the ten $g(869)$ coefficients, $g(748)$ was derived from ocean gyre targets as the value which forced ε to match values expected from the predetermined aerosol model. Using the combination of those two gains to calculate L_a(VIS), g(VIS) were then calculated according to comparisons with MOBY data. All 10 such derivations used the exact same gyre data to derive $g(748)$ and MOBY data to derive g(VIS). However, the final number of pixels included in the derived g(VIS and $\mathrm{NIR_S}$) decreased with increasing $g(869)$, as did the standard deviation about these g(VIS) determinations. Both are due to differential determinations of l2_flags for quality control of SVC matchups amongst the different $g(869)$ conditions. The additionally excluded pixels in the higher $g(869)$ condition were primarily at the extremes of the data spread, which slightly decreased the SVC standard deviation. Nevertheless, data quantity for all MOBY-based gain determinations ($N > 600$), however, included many more data points than the minimum (~40) suggested by Franz *et al.* [5], and were evenly distributed seasonally and annually. Table III

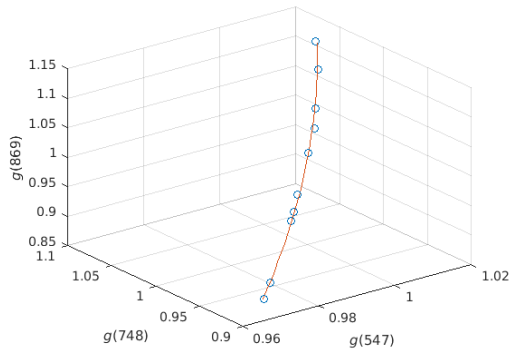


Fig. 4. F2007 MODIS/A SVC configurations (blue circles) and 3-D polynomial regression line (orange). Each gain configuration results from SVC starting with an arbitrarily selected $g(869)$. All gain configurations used in this work ($N \sim 1.4E6$) are within 0.0078 of this line, with 0.0001 increments.

TABLE IV

SECOND-ORDER POLYNOMIAL REGRESSION COEFFICIENTS $[y = \beta_2 x^2 + \beta_1 x + \beta 0]$ AND R^2 DESCRIBING THE RELATIONSHIPS BETWEEN $g(547)$ AND GAINS FOR ALL OTHER BANDS, AS DERIVED USING F2007 GAIN CONFIGURATIONS

Band Center	β_2	β_1	β_0	R^2
412	-2.486	5.179	-1.715	1.000
443	-2.709	5.731	-2.033	1.000
469	-2.403	5.229	-1.820	1.000
486	-2.310	5.133	-1.830	1.000
531	-0.520	1.901	-0.380	1.000
547	0.000	1.000	0.000	1.000
555	0.404	0.238	0.352	1.000
645	7.574	-13.089	6.532	0.999
667	9.655	-17.005	8.358	0.997
678	10.785	-19.127	9.349	0.997

lists the suite of SVC gain configurations, hereafter termed "F2007" (Franz *et al.*, 2007; [5]) gain configurations.

Recognizing that the intervals between $g(869)$ values included in the F2007 data set may obscure results in our sensitivity analyses, we sought to approximate F2007 data for assessment of $g(\mathrm{NIR_L})$ from 0.85 to 1.2 at the current precision of operational SVC (0.0001). Fortunately, F2007 $g(869)$ and $g(748)$ for these 10 configurations showed an extremely strong relationship. This was expected, given how these gains were derived. Simple linear regression between $g(869)$ and $g(748)$ indicated a nearly directly dependent relationship ($R^2 = 0.9992$) with slope and intercept of 0.66 and 0.34. Both $g(869)$ and $g(748)$ also showed a strong relationship with $g(547)$. Plotted together (Fig. 4), these gain configurations appeared to be best represented by a three-dimensional second-order polynomial. As such, the suite of gain configurations tested in the subsequent work were selected as those within a linear distance of 0.0078 from this polynomial regression line, with $g(869)$ varying from 0.85 to 1.2 in 0.0001 increments. This threshold for deviation from the polynomial regression line was determined as the maximum residual among the F2007 gain configurations. Polynomial coefficients against a placeholder variable (d, with $-0.142 < d < 0.150$) for this regression line in X, Y, and Z space [g(547, 748, and 869, respectively)] were (−0.219, 0.175, 0.998), (−0.096, 0.545, 0.995), and (0.110, 0.820, 0.984). For example, $g(547) = -0.219d^2 + 0.175\mathrm{d} + 0.998$.

To a lesser extent, $R_{\mathrm{rs}}(547)$ retrievals are also impacted by SVC gains for wavebands other than $\mathrm{NIR_S}$ and $\mathrm{NIR_L}$. Such impacts can be manifested through correction for "out-of-band" (OOB) [48]–[50] and Bidirectional Reflectance Distribution Function (BRDF) [50]–[53] effects. For both of these parameters, an initial C_a estimate, calculated using the 443-, 547-, and 667-nm bands, is used to determine the spectral dependence of the correction. For large deviations from F2007 spectral gains [e.g., if $g(547) = 0.85$ and g(other VIS) $= 1.0$, see Table III], these effects on derived $R_{\mathrm{rs}}(547)$ can be on the order of several percent or more. Thus, for the remaining VIS bands, second-order polynomial functions were fit between OBGP-derived $g(547)$ and gains for all other bands. The coefficient of determination (R^2) for these regressions were all above 0.995 (Table IV). For each $g(547)$, these regression coefficients were used to estimate gains for all other bands. All told, the gains applied at L1B included: 1) for wavebands at 547, 748, and 869: a suite of incrementally-defined gains, with a step of 0.0001, within the 3D F2007 envelope (Fig. 4); and 2) for all other wavebands: gains derived from polynomial regression from $g(547)$.

D. Time Series Analyses

Seventeen-year $R_{\mathrm{rs}}(547)$ time series for each of the approximately 1.4 million gain configurations were derived at each of the 125 gyre stations, excluding any pixels with sensor zenith $>40°$ [15]. For each gyre and gain configuration, $R_{\mathrm{rs}}(547)$ data from the 25 sample stations were combined into composite gyre- and gain-configuration-specific time series. Several methods were employed to remove data of questionable quality. Through this process, to maintain the same number of data points for all data sets, any pixel identified as "questionable" in ANY gain-configuration data set was removed from ALL data sets. For example, outliers were identified by comparing each data point to the corresponding gyre- and gain-combination specific $\bar{x}$ and σ. Any pixel which exceeded 3σ from this overall gyre- and gain-configuration specific $\bar{x}$ was excluded from all analyses, which removed 3%–4% of all data per gyre. Additionally, the Level-2 flags product, which is uniquely determined for each pixel as processed with each gain configuration, was used to identify and remove potentially low-quality data. Any pixel which was identified by any of the Level-2 processing flags as calculated using any gain configuration was removed from analysis for all gain configurations (~24% of data). Overall, this left 5274, 2999, 5818, 5077, and 6275 pixels in each time series for NAG, NPG, SIG, SPG, and SAG, respectively, with data from the 25 individual stations being combined for subsequent analyses.

For each gyre and gain configuration, moving means were calculated for each date of data

$$y_{ti} = \frac{1}{N_{ti}} \sum_{-i}^{i} x_{t+i} \tag{3}$$

where x is MODIS-derived $R_{\mathrm{rs}}(547)$ and y_{ti} is the moving mean $R_{\mathrm{rs}}(547)$ for a given date (t) and moving mean interval (i), with i being half the moving mean span (i.e., $i = 15.5$ for a 31-day moving mean span). As such, i describes the number of days before and after t which are included in the calculation of y_{ti}, while N_{ti} is the number of pixels between $t - i$ and $t + i$. These moving means (y_{ti}) were calculated for all instances with valid data (i.e., where $N_{ti} > 0$), thus t only exists for "sample-days," defined as dates for which valid pixels were found within a given gyre. For the 10 F2007 gain configurations (Table III), all odd integer spans between 1 and 365 days were tested. Results for a subset of these (spans = 7, 15, 31, 61, 91, 181, 365) were calculated for all other gain configurations. Due to data from 25 gyre stations being combined in calculation of y_{ti} (3) and subsequent equations, in many cases multiple values for x were collected for the same t. In such instances, we retain the original data resolution (e.g., $x_{t,1}$, $x_{t,2}$, ..., are all considered separately, and all share the same y_{t} value), for all subsequent calculations, except as specifically noted. This means that y_{t} for dates with multiple x are more heavily weighted in the calculations. However, this also means that interday variability is retained, as needed to quantify uncertainties associated with combining data from the 25 stations within each gyre.

For each "moving mean model," defined as the 17-year y_{ti} time series—unique to each gain configuration, gyre, and span, we calculated R^2 from the Sum of Squared Residuals (SSR) and Total Sum of Squares (SST) as $R^2 = 1 - (\mathrm{SSR}/\mathrm{SST})$, as

$$\mathrm{SST} = \sum (x_{\mathrm{t}} - \bar{x}_{R\mathrm{rs}})^2 \quad \text{and} \quad \mathrm{SSR} = \sum (x_{\mathrm{t}} - y_{\mathrm{t}})^2 \qquad (4)$$

with $\bar{x}_{R\mathrm{rs}}$ representing the mean of all $R_{\mathrm{rs}}(547)$ data points within the 17-year time series for that gain configuration, gyre, and moving mean span. Mean absolute deviation (MAD) was also used to access residuals/deviations of the satellite data from the moving means as

$$\mathrm{MAD} = \sum |x_{\mathrm{t}} - y_{\mathrm{t}}|. \qquad (5)$$

Additionally, lag-1 sample-day autocorrelation coefficients (i.e., comparing y_{t} to y_{t+1}) were calculated for each moving mean model as a metric of temporal coherence, returning the two-tail pairwise Pearson's linear correlation coefficient (r) and associated p-value (p). Due to temporal gaps in the data, y_{t} and y_{t+1} do not necessarily correspond to subsequent days. While this is not optimal for autocorrelation coefficient calculation, the same satellite pixels were used for all gain configurations, meaning the temporal gaps did not differ for analyses comparing various gain configurations. Also, contrary to all other analyses, autocorrelation coefficient calculations were considered using one y_{t} per t, regardless of the number of valid x collected on that date. This is required as identical y_{t}, which result from multiple x values on the same t, would artificially increase the autocorrelation coefficient.

Several descriptive statistics were calculated for visualization of general trends, including overall mean and standard deviation of all $R_{\mathrm{rs}}(547)$ data points within each 17-year time series ($\bar{x}_{R\mathrm{rs}}$ and $\sigma_{R\mathrm{rs}}$, respectively; in sr^{-1}), as well as the standard deviation of y_{ti} from each moving mean (σ_{mod}; in sr^{-1}). "Global" MAD were calculated for each gain configuration using the combined data from all gyres within (5).

IV. Results

A. Effect of Changing $g(NIR_L)$

For illustration, results are shown first for a tiny subset of the data (Fig. 5), namely NAG $R_{\mathrm{rs}}(547)$ time series data as processed using three of the F2007 gain configurations, overlain by four moving mean models (spans of 7, 31, 91, and 365 days). In all cases, longer moving mean spans resulted in more stable looking time series, with the 365-day models being nearly flat. For the NAG (Fig. 5) 31- and 91-day models, time series from all three gain configurations show a largely seasonal pattern at a period of 1 year, with $R_{\mathrm{rs}}(547)$ higher in late (boreal) summer and lower in late winter. Deviations from this pattern typically manifest as plateaus or dips during the late summer peak, but are much more common in the time series calculated using $g(869) = 0.87$ than those with $g(869) = 1.11$. For these three time series, $\bar{x}_{R\mathrm{rs}}$ decreases slightly with an increasing $g(869)$ (1.6% decrease from $g(869) = 0.87$ to $g(869) = 1.00$; 3.8% decrease from $g(869) = 1.00$ to $g(869) = 1.11$).

Overlaying time series as created using different gain combinations on the same plot (Fig. 6), some general patterns emerge. First, for all gyres, the amplitude of the seasonality in $R_{\mathrm{rs}}(547)$ is larger for time series processed with larger $g(869)$. The greatest agreement appears to be during the summer season (boreal for NAG and NPG, austral for SAG, SPG, and SIG). Additionally, the exact timing of peaks differs according to the gain configuration used in processing, particularly in the NAG, where time series processed with smaller $g(869)$ show later timing for the seasonal maxima. Other peaks (e.g., NAG in late 2011; NPG in early 2005; SAG in mid-2003) are observed in time series as processed using some gain configurations, but not others (Fig. 6). It is also noted that the general shape of the time series can differ according to gain configuration. For example, all SIG time series show a prominent austral summertime $R_{\mathrm{rs}}(547)$ peak, but the time series processed with $g(869) = 0.87$ and 0.946 also show a secondary wintertime peak. In the SPG, data processed with $g(869) = 0.87$ is largely flat—showing no readily apparent regular inter or intraannual patterns. With $g(869) = 1.00$, a single peak becomes apparent in most austral summers, while the time series processed with $g(869) = 1.11$ shows a very regular seasonal oscillation.

Placing Figs. 5 and 6 in the context of results all 1.4E6 gain configurations, overall trends in the summary statistics for the time series are shown in Fig. 7. In particular, $\bar{x}_{R\mathrm{rs}}$ [Fig. 7(a)] shows a general decrease with an increasing $g(869)$. For the NAG, NPG, and SAG, the $\bar{x}_{R\mathrm{rs}}$ increased slightly for gain configurations with $g(869)$ between 0.85 and 0.95, then decreased steadily with an increasing $g(869)$ [Fig. 7(b)]. In the other two gyres (SPG and SIG), this decrease was more substantial, spanning the range of $g(869)$ tested. Among the gain configurations tested, the variability in $\bar{x}_{R\mathrm{rs}}$ for individual gyres varied between 12.8% (NPG) and 19.4% (SIG), calculated as the unbiased percent difference between highest and lowest observed value. Standard deviations about these means ($\sigma_{R\mathrm{rs}}$)

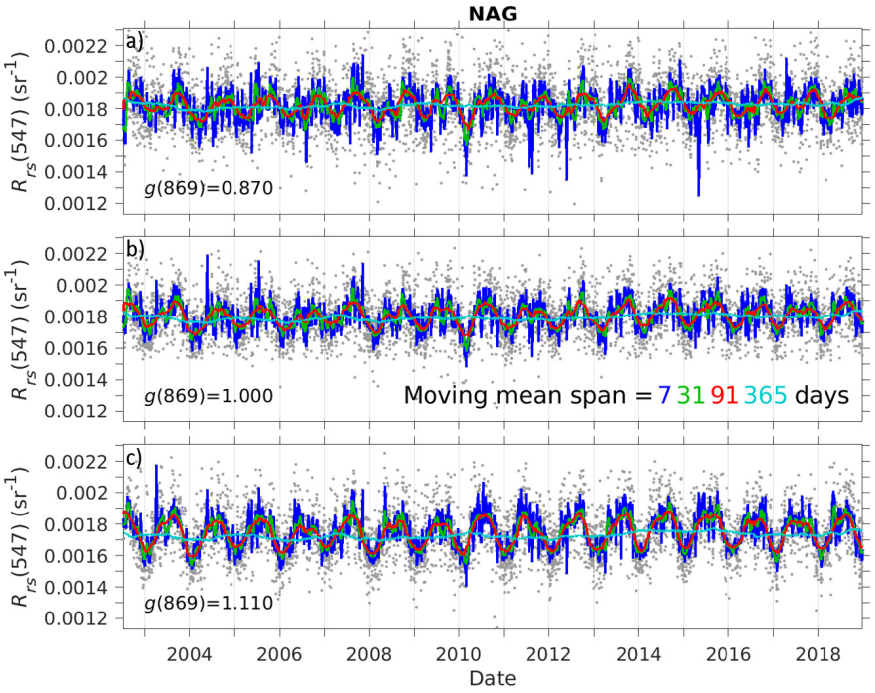


Fig. 5. Time series of $R_{rs}(547)$ (sr^{-1}) aggregated from the 25 North Atlantic Gyre (NAG) stations (gray dots) as processed using three different F2007 gain configurations. These gain configurations were determined using MOBY and Franz *et al.* (2007), with initial $g(869)$ set at (a) 0.87, (b) 1.00 (current OBPG default), and (c) 1.11 (rows 1, 6, and 10, respectively, in Table II). Moving means overlay the data, with span length indicated by color: 7 (dark blue), 31 (green), 91 (red), and 365 (light blue) days.

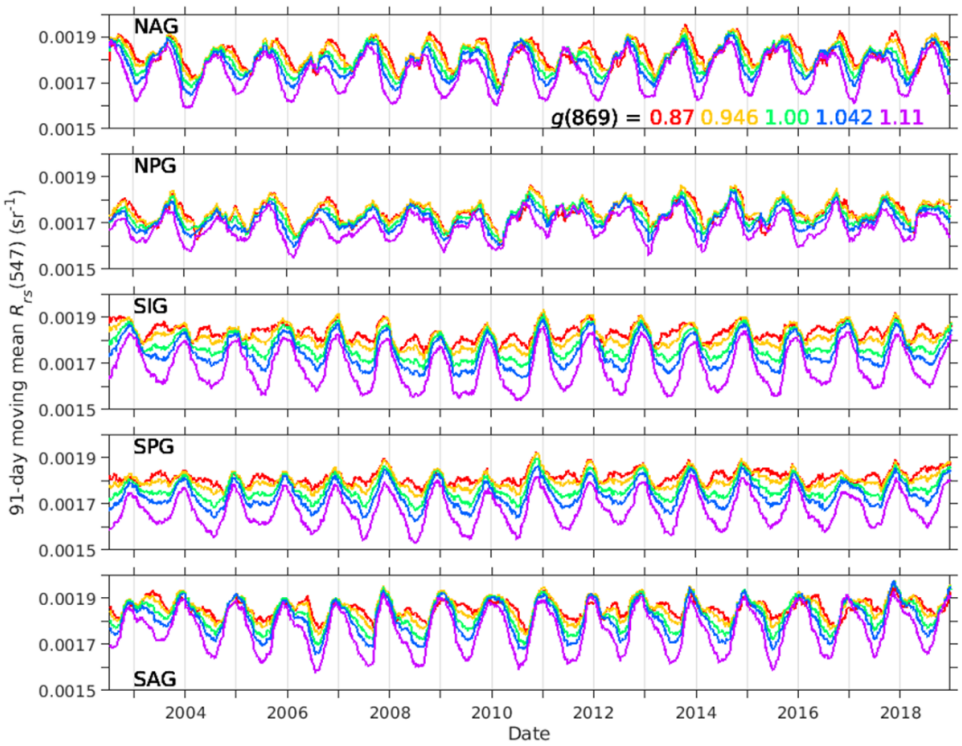


Fig. 6. 91-day moving means for the $R_{rs}(547)$ time series in the 5 ocean gyres (top to bottom: NAG, NPG, SIG, SPG, SAG) as processed using 5 different F2007 gain combinations: $g(869) = 0.87$ (red), 0.946 (orange), 1.00 (green), 1.042 (blue), and 1.11 (purple).

also showed variability with changing $g(869)$. Fig. 7(b) shows σ_{Rrs} was highest for the most extreme values of $g(869)$, with the exact locations of the minima differing by gyre (roughly between 0.97 and 1.02).

It is noted that the spread of data points in Fig. 7 and many subsequent figures results from multiple gain configurations sharing the same $g(869)$, as was necessary to accommodate uncertainties in the 3-D polynomial fit between

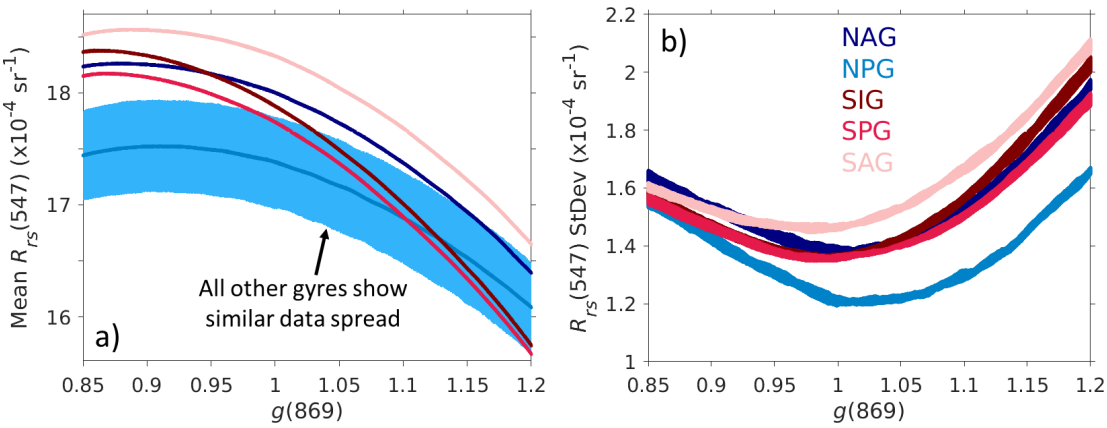


Fig. 7. Gyre- and gain-configuration- specific (a) $\bar{x}_{Rrs}$ [mean $R_{rs}(547)$ for each 17-year time series], (b) σ_{Rrs} [$R_{rs}(547)$ standard deviation]. Color indicates gyre: NAG (navy), NPG (blue), SIG (maroon), SPG (red), and SAG (pink). Note that each "line" in the panels represents ~1.4E6 individual data points. Due to overlap, in (a), the full data spread is only shown for NPG. The center of the data spread is shown for all other gyres, with the data spread for these gyres being similar to that for NPG.

g(547, 748, and 869) (i.e., the F2007 "envelope"). For most statistical measures, the variability associated with gain configurations sharing an individual g(869) value is quite small when compared to that seen across gyres or along the range of g(869) tested. Nevertheless, if all gain configurations had been calculated using the standard SVC procedure [5], then each g(869) would correspond to a single gain configuration.

B. Comparison to Moving Mean Models

While these summary statistics (Fig. 7) demonstrate clear impacts of changing g(869) on $R_{rs}(547)$ retrievals, it is impossible to infer much about the shape or fit of these time series using these statistics alone. For example, a large σ_{Rrs} could result from widely scattered data around a mean that is stable in time, or from data tightly clustered around a harmonic oscillation. The former condition is regularly seen for gain configurations with low (<0.9) g(869), while the latter is seen when $g(869) > \sim 1.1$ [Figs. 5–7(b)]. As such, we calculated moving mean models (3) and associated statistics for each time series. In these analyses and subsequent discussion, we broadly define "better" models as having: 1) higher temporal stability of the moving mean values (y_t, quantified as lag-1 sample-day autocorrelation); and 2) low residuals from that model (low MAD, high R^2). Both of these assumptions are based on the premise that, in ocean gyre centers, any change in $R_{rs}(547)$ should occur gradually rather than abruptly (i.e., nature is not noisy, and thus data should not "jump"). Autocorrelation coefficient captures temporal coherence of low-frequency $R_{rs}(547)$ variations—depending on the span, this can include weekly, monthly, seasonal, and/or annual variations. MAD and R^2 quantify higher frequency variability, at the level of individual samples. This model framework is primarily based on a logic argument of temporal coherence in gyres, but is necessary as (1) current knowledge of ocean gyre $R_{rs}(547)$ dynamics is derived from satellite-based assessments, which (2) are themselves impacted by the initial g(869) (Figs. 1, 5, and 6). Thus, given an appropriate moving mean span and absent confounding factors (Section V-B.), we operationalize "better" time series as those with minimal residuals around a smoother moving mean. Another important consideration here is that all metrics are calculated using common pixels. Thus, a change in MAD between two different gain configurations reflects a stronger or weaker fit to moving means as calculated from the same pixels.

Obviously, the value used as the model interval [i = span/2; (3)] will impact the calculated moving mean values (y_{it}) and subsequent model statistics. This is intuitive for moving means, as small spans (<7 days) should result in more temporally unstable models that strongly fit the individual data points (Fig. 5, dark blue), while large spans (>300 days) should yield nearly flat models which reflect little of the variability in individual measurements (Fig. 5, light blue). This interplay, across the entire range of i tested, is presented in Fig. 8 for three different F2007 gains. As expected, for all gyres and gain configurations, R^2 and lag-1 sample-day autocorrelation coefficient show an inverse relationship. Comparing Fig. 8(a)–(c), it is noted that model R^2 and autocorrelation coefficient both appear to increase with an increasing g(869). Average data quantity [mean of N_{it}; Fig. 8(d)] for most gyres is roughly equivalent to the span. This means that any given mean (y_t) calculated with a seven-day span will include an average of seven $R_{rs}(547)$ values (and 31 values in 31 days, etc.). There are, however, many dates (t) with multiple measurements ($x_{t,1}$ $x_{t,2}$, ...), and many dates with no measurements. N_{it} for NPG is approximately two-thirds that seen in the other gyres.

Although we performed all subsequent analyses using seven different i values, in the following section we show results only for a span of 31 days. Most other spans, except 365 days, showed very similar results, and are discussed in Section V-B. It is noted [Fig. 8(a)–(c)] that the autocorrelation coefficient has largely reached the maximum value (~0.99 for all gyres except NPG) for a span of 31 days, thus there are diminishing returns in temporal stability of the models achieved by further increasing the span. Additionally, R^2 for spans of ~31 days are fairly low (0.1–0.4), but continue decrease for longer spans.

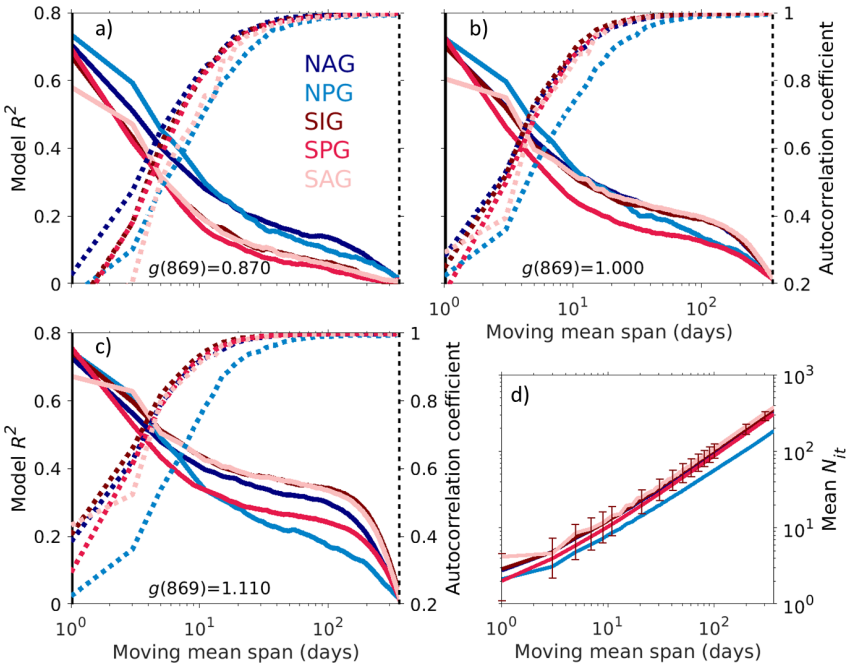


Fig. 8. (a)–(c) Response of model fit (R^2, left axes, solid lines) and model temporal coherence (lag-1 sample-day autocorrelation coefficient, right axes, dotted lines) to changes in the model span (x-axes, note log scale). Panels a–c show results for separate F2007 gain configurations, with initial $g(869)$ set at (a) 0.87, (b) 1.00 current OBPG default), and (c) 1.11 (rows 1, 6, and 10, respectively, in Table III). Mean data quantity for these calculations shown in (d), with error bars indicating $\pm 1\sigma(N_{it})$ for SPG [all other gyres showed similarly sized $\sigma(N_{it})$]. Gyre results shown independently: NAG (navy), NPG (blue), SIG (maroon), SPG (red), and SAG (pink). Axes scales for (a)–(c) are identical.

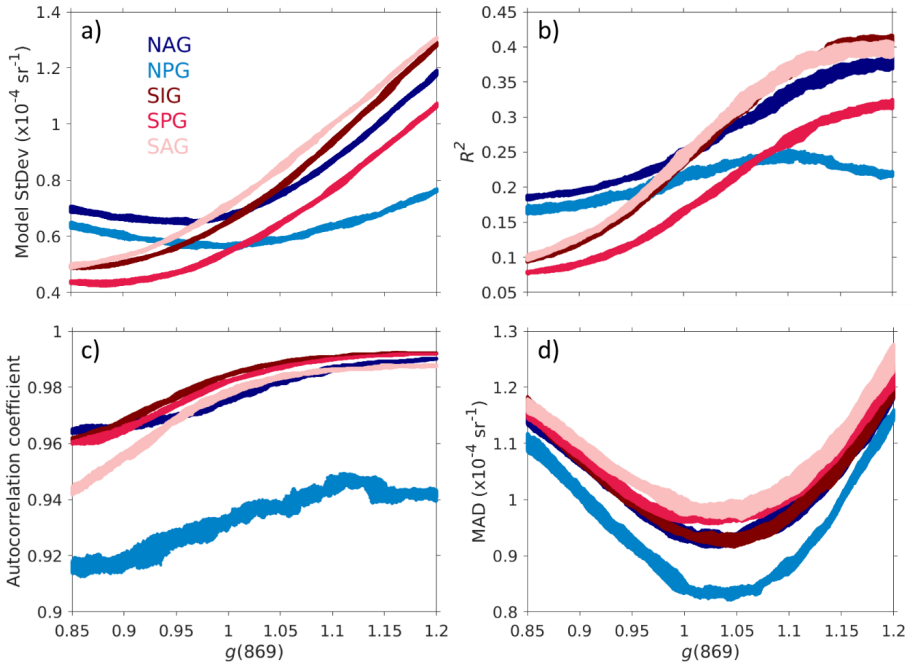


Fig. 9. Gyre- and gain-configuration- specific characteristics of the 31 day moving mean models and associated residuals: (a) standard deviation of the 31 day moving means [σ_{mod}], (b) model R^2, (c) autocorrelation coefficient, and (d) MAD. Color indicates gyre: NAG (navy), NPG (blue), SIG (maroon), SPG (red), and SAG (pink). Note that each "line" these panels represents ~1.4E6 individual data points.

As with the overall summary statistics, model statistics also show substantial sensitivity to changes in $g(869)$, with responses occasionally varying between gyres (Fig. 9). For SIG, SPG, and SAG, moving mean values (y_t) showed a narrower variability (i.e., lower σ_{mod}) for gain configurations with lower $g(869)$, increasing steadily with an increasing $g(869)$ [Fig. 9(a)]. The northern hemisphere gyres, however, show minima in σ_{mod} for $g(869) \approx 0.95 - 1.0$. R^2 increased with increasing $g(869)$ for all gyres except NPG, for which R^2 peaked for gain configurations with $g(869) \approx 1.1$ [Fig. 9(b)].

Lag-1 sample-day autocorrelation coefficient increased slightly with increasing $g(869)$, indicating increased temporal coherence of y_t [Fig. 9(c)]. For $g(869) > 1.05$, most gyres show little to no additional improvement in autocorrelation coefficient. As autocorrelation coefficient is sensitive to the variability in y_t, it is possible that some of the increases in autocorrelation coefficient (especially for $g(869) > \sim 1.05$) are an artifact of increasing σ_{mod}. NPG autocorrelation coefficient was much lower than that for all other gyres. Finally, MAD for all gyres shows minima for gain configurations with $g(869) \approx 1.01$–1.06 [Fig. 9(d)]. MAD changes with $g(869)$ around these minima are quite small (i.e., the troughs in Fig. 9(d) are flat).

V. Discussion

In this article, we attempted to determine the impact of changes in $g(\text{NIR}_\text{L})$ on MODIS-derived time series in ocean gyre centers. This was accomplished through investigating the extent to which MODIS $R_{\text{rs}}(547)$ time series in these gyre centers responds to 1.4E6 different combinations of $g(869)$, $g(748)$, and $g(547)$. The results show substantial sensitivity of the observed ocean gyre dynamics to changes in $g(869)$, with changes observed in all parameters measured, including the overall mean and standard deviation of 17-year $R_{\text{rs}}(547)$ time series [Fig. 7(a) and (b)]. Even the shape of observed trends can vary within a single gyre (Figs. 5 and 6). As a defining example, in the SIG, $R_{\text{rs}}(547)$ time series change from an apparent annual cycle with two peaks per annum [$g(869) \approx 0.85$] to an annual seasonal oscillation peaking in the boreal winter [$g(869) \approx 1.1$], dependent only on the prelaunch calibration of $g(869)$.

It is important to highlight similarities in results between the MODIS/A analyses presented here and the SeaWiFS assessment performed by NASA, a portion of which is summarized in Fig. 1. These two sensors greatly differ in their methods of data collection, and the approaches for data integration toward assessment of $g(869)$ sensitivity between the previous and current studies are quite different. Nevertheless, both analyses show sensitivity in $R_{\text{rs}}(\text{VIS})$ associated with the calibration of the NIR_L band, leading to apparent differences that are geographically- and temporally-dependent.

For the suite of gain configurations tested, satellite-derived ocean gyre $R_{\text{rs}}(547)$ means ($\bar{x}_{R\text{rs}}$) varied by up to nearly 20% [Fig. 7(b)], calculated as the unbiased percent difference between the highest and lowest $\bar{x}_{R\text{rs}}$. This indicates that uncertainties for overall gyre-specific means associated with $g(\text{NIR}_\text{L})$ can be quite large; recall that target uncertainty for satellite ocean color R_{rs} is 5% [54]. Obviously, these percent change statistics calculated for a smaller range of initial $g(869)$ would show reduced variability in overall mean and standard deviation associated with $g(869)$ changes. Indeed, considering only $0.95 < g(869) < 1.05$, gyre specific $\bar{x}_{R\text{rs}}$ varied in the range of 6%–8%. Additionally, these values are potentially inflated by the study design, in that multiple gain configurations share a single $g(869)$. Considering only F2007 with $0.946 \leq g(869) \leq 1.042$, $\bar{x}_{R\text{rs}}$ varied between 2% and 4%.

At first glance, these changes in $\bar{x}_{R\text{rs}}$ appear incompatible with the overall MOBY-based calibration procedure. For example, since $g(547)$ were calculated using the same MOBY data for all 10 of the F2007 gain configurations, then $R_{\text{rs}}(547)$ should not show variability according to $g(869)$. Clearly, there exists a disconnect between the data set used in this study and the MOBY and/or gyre data sets as used for SVC. We noted that all gyres studied show overall decreases in $\bar{x}_{R\text{rs}}$ with increasing $g(869)$ [Fig. 7(a)]. Additionally, $\bar{x}_{R\text{rs}}$ for SeaWiFS data (as presented in Fig. 1, and processed using an entirely different methodology) also changes with $g(865)$. For the Equatorial Pacific, $\bar{x}_{R\text{rs}}$ increased by 1% when switching from $g(865) = 0.95$ to $g(865) = 1.0$, and decreased by $\sim$2% when switching from 1.0 to 1.05. For those analyses, the 0.95 and 1.05 conditions are not directly comparable due to different data used in the two comparisons. In recent years, there has been some discussion as to the costs and benefits of using a single site (e.g., MOBY) for SVC of the VIS bands [55], [56]. In particular, inherent inconsistencies of atmospheric correction over different solar viewing geometry, or relating to the prevailing atmospheric optical properties, would lead to different SVC gains for different stations [56]. Perhaps the responsiveness of $\bar{x}_{R\text{rs}}$ to $g(869)$, as seen in the current study, reflects a relative variety in solar viewing geometries or atmospheres within the current data set when compared to the MOBY data set. As the largest sensitivity in $R_{\text{rs}}(547)$ to $g(869)$ changes appear in wintertime (Fig. 6), solar zenith may be particularly relevant.

To investigate manifestations of varying gain combinations on individual satellite data products and atmospheric correction parameters/intermediates, we processed several complete MODIS/A granules using each of the F2007 gains. Obviously, the direct impacts of varying these gains were universally observed. For example, $L_t(\lambda)$ varied proportionally to $g(\lambda)$. Almost all other parameters and atmospheric correction intermediates investigated showed substantial variability in response to gain configurations that was not consistent. Indeed, this variability often changed signs within individual granules as processed using different gain configurations. For any given satellite pixel, considering a more restrictive range of $g(869)$ than used in the current study [i.e., $g(869) = 1.0 \pm 5\%$], the aerosol models [25] selected were typically identical, regardless of the gain configuration used in processing. Nevertheless, ε and thus the Ångström exponent (α, unitless) did show slight variability within granules in response to varying gain configurations. BRDF(547) correction also varies very slightly (roughly $\pm 0.1\%$ within a single granule). Since geometry is unchanged in these comparisons, this impact results from differences in the initial C_a value used to derive the spectral shape of the BRDF correction. This impact is not particularly surprising, as the initial C_a value is calculated after atmospheric correction, but exemplifies the difficulty in disentangling the multiple precursors that ultimately impact any individual $R_{\text{rs}}(547)$ retrieval. Overall, calculating the unbiased percent difference for individual pixels as processed using two different gain configurations with only moderate (<4%) changes in $g(869)$, some pixels showed relative stability in $R_{\text{rs}}(547)$ (i.e., $\pm 2\%$), while others showed much larger impacts ($\pm 10\%$).

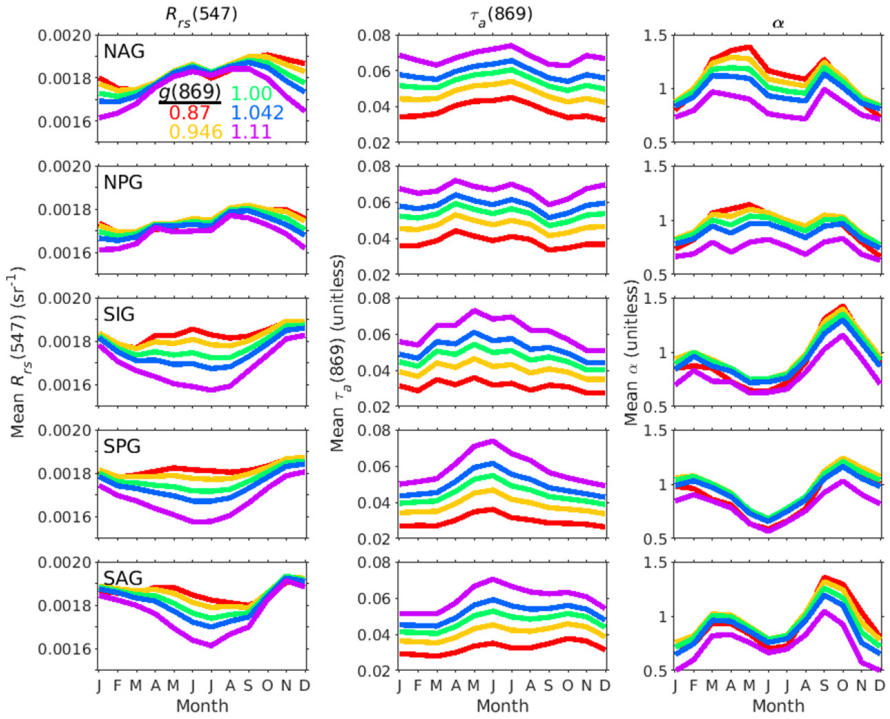


Fig. 10. Monthly climatological means for (left column) $R_{\rm rs}(547)$, (middle column) $\tau_a(869)$, and (right column) α within each of the 5 gyres (top to bottom: NAG, NPG, SIG, SPG, SAG), as processed using 5 F2007 gain configurations: with $g(869) = 0.87$ (red), 0.946 (orange), 1.00 (green), 1.042 (blue), and 1.11 (purple).

On a larger scale, the overall magnitude of changes in summary statistics (i.e., $\bar{x}_{R\rm rs}$, MAD, etc.) described in this work are minor. For example, a 4% decrease in $\bar{x}_{R\rm rs}$ corresponds to only $\sim$0.0007 sr^{-1}. Changes in MAD with $g(869) = 1.0 \pm 5\%$ are on the order of $\sim$0.0001 sr^{-1}. Each summary statistic, however, represents thousands of individual pixels, each of which responds differently to $g(869)$ changes. Indeed, scatterplots of all pixels as processed using different gain configurations reveal surprisingly low R^2 between these data sets, which supports the findings of our granule-based assessment in the previous paragraph. For example, comparing pixels as processed using the F2007 gain configuration with $g(869) = 0.946$ to those processed with $g(869) = 1.042$, R^2 for $R_{\rm rs}(547)$ ranges from 0.68 (NPG) to 0.75 (SAG). Thus, a 4% decrease in $\bar{x}_{R\rm rs}$ is not manifested for every pixel, but is revealed only in the aggregation of many points.

Indeed, looking at aggregated monthly climatological means (Fig. 10), a potential reason for some of the sensitivity to $g(869)$ can be elucidated. As first noted in Fig. 6, summertime $R_{\rm rs}(547)$ showed less sensitivity to $g(869)$ than wintertime data [Fig. 10 (left)]. For all seasons, the derived aerosol optical thickness at 869nm [AOT(869) or $\tau_a(869)$] shows increases with an increasing $g(869)$. This is unsurprising considering how it is calculated, using $L_a(869)$ (see [34, Fig. 1]). In wintertime, α shows little variability according to $g(869)$, while summertime α decreases with increasing $g(869)$. Very basically, the effect of these two trends is as follows: in summertime, the impact of thicker derived atmospheres [higher $\tau_a(869)$ with an increasing $g(869)$] on $R_{\rm rs}(547)$ is modulated by a decrease in the slope of aerosol effects from the NIR to VIS [lower α with an increasing $g(869)$]. In summertime, no such decrease in α is observed with an increasing $g(869)$, meaning L_a(VIS) increases and $R_{\rm rs}$(VIS) decreases with an increasing $g(869)$.

As such, we believe that the overarching findings described above can only be achieved through statistical analyses of large-scale, long-term time-series data, as opposed to using statistics of point-based validations or simulations. Indeed, through simulations, Wang and Gordon [21] showed that accuracy of retrieved $R_{\rm rs}(443)$ for clear waters can meet the mission goal of <5% uncertainties as long as $g(\rm NIR_L)$ is 0.9–1.1. Using the approaches described here, this study instead shows that even a small change in $g(\rm NIR_L)$ can result in different time-series shapes and uncertainty distributions within ocean gyres. These results do not directly challenge the findings of Wang and Gordon [21], but suggest that those uncertainties associated with $g(\rm NIR_L)$ may not be uniformly distributed in space or time. We therefore believe that more constraints may be needed to define future mission goals.

A. Identifying an "optimal" g(869)

Knowing that $R_{\rm rs}(547)$ time series are sensitive to the $g(869)$ used to derive them, and that these patterns differ by gyre, a natural extension to this work is to consider which gain configuration(s) produce(s) time series that most closely match expectations. Based on the premise that $R_{\rm rs}(547)$ changes slowly in ocean gyres instead of "jumping" day to day, we thus parameterized an 'optimal' time series as one which (1) closely tracks a moving mean which is (2) gradually

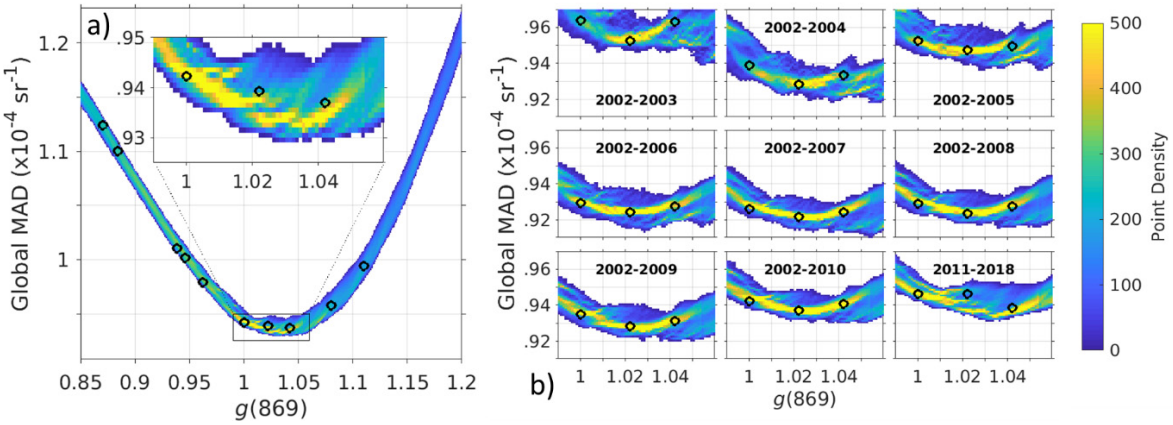


Fig. 11. Global MAD (including data from all gyres) shown with color indicating density of gain configurations with the indicated value. Black circles specify values for the 10 F2007 gain configurations. In (b), the inset region from (a) is replicated for various temporal subsets of the data.

changing, operationalized as low MAD and high autocorrelation coefficient, respectively. For the former, note that model R^2 is not as diagnostic as MAD, since SST (4) for a temporally homogenous $R_{rs}(547)$ time series would be low, yielding a low R^2 irrespective of the actual model fit. SSR itself is also not used to quantify residuals, as the squaring procedure gives more weight to higher $R_{rs}(547)$ values.

For all gyres, the autocorrelation coefficient is maximal for gain configurations with $g(869) > \sim 1.1$, with little variation for $g(869) > 1.0$. As noted, since autocorrelation coefficient can be impacted by the variability in y_t [similar to R^2; see (4)], and σ_{mod} increases steadily in this range [Fig. 9(a)], these small increases in autocorrelation coefficient for $g(869) > 1.0$ are potentially not diagnostic. Within each gyre, MAD shows a low shallow trough for gain configurations with $g(869) \sim 1.01$–1.06 [Fig. 9(d)]. "Global" MAD [gyre-integrated; calculated for each gain configuration using (5), but using data from all gyres] shows a minimum for gain configurations with $g(869) \sim 1.025$ [Fig. 11(a)]. As with gyre-specific MAD, this trough is very flat. Global MAD for the 10 F2007 gain configurations listed in Table III are marked in Fig. 11(a). Note that these ten data points are not perfectly well represented by the highest density of all data points. Nevertheless, from these parameters combined (and considering assumptions and limitations in Section V-B.), we estimate that $g(869)$ between roughly 1.01 and 1.05 produces time series which are potentially "optimal."

B. Assumptions, Uncertainties, and Limitations

In addition to the main assumption defining expected gyre $R_{rs}(547)$ dynamics as measured from satellites—that $R_{rs}(547)$ changes gradually—several other assumptions are embedded in this work and our interpretation of the results. The first concerns data quantity, including the stations and year ranges used. Within this work, we attempted to elucidate gyre seasonal patterns using a subset of data points (gyre specific N range from 2999—6275). For mission-long monthly climatologies, N for such gyres (black squares in Fig. 3) are on the order of 10^7 pixels for SeaWiFS or 10^8 pixels for MODIS and VIIRS/SNPP (Visible Infrared Imaging Radiometer Suite onboard Suomi-NPP). Additionally, we used individual pixels with nadir 1 km spatial resolution, as opposed to integrating data from multiple pixels (e.g., a 3 × 3 pixel box) as is more common in calibration and validation exercises [7]. Thus, can the subset used in our study well represent the gyres as a whole? We calculated gyre-specific $R_{rs}(547)$ monthly climatologies from this subset using NASA default gains and compared them to the OBPG-derived mission monthly climatologies, available from NASA OBPG (https://oceancolor.gsfc.nasa.gov). While there are some small differences between the two, nearly all are within 5%, and no seasonal or hemispheric trends were observed. Of particular note, monthly mean data used in the current analyses are slightly higher, likely due to the more stringent quality control methods used in this article than in the development of the OBPG mission monthly climatologies.

To further investigate the data quantity requirements for adequately capturing gyre dynamics, and to assess if the observed results are specific to the time periods studied, we also repeated the analyses using various subsets of our full data set. First, we trimmed the data set into approximately half (2002–2010 compared to 2011–2018, inclusively), with results of both subsets closely mirroring results of the entire data set. Notably, global MAD for both subsets was also minimal for $g(869)$ between 1.01 and 1.05 [Fig. 11(b)].

Restricting the dataset further, we replicated the analyses using data from incrementally increasing numbers of years, starting with 2002–2003, inclusive, which represents 1.5 years of data [Fig. 11(b)]. As such, we also attempted to determine the minimum number of years of satellite data needed to arrive at a stable "optimal" gain. Somewhat surprisingly, results for data sets including only a few years of data were remarkably similar to those for the entire data set—for all time periods tested, global MAD showed minima for the largest density of data points for $1.01 < g(869) < 1.05$ [Fig. 11(b)].

Even assuming sufficient data quantity to appropriately capture long-term gyre variability, we noted that a substantial difference in data quantity between NPG ($N = 2999$) and all other gyres ($N = 5077$ – 6275). The MAD statistic itself

is normalized by the number of data points, thus findings based on this metric should not be impacted by cross-gyre differences in data quantity. The autocorrelation coefficient for NPG is lower and more variable than all other gyres for all gain configurations [Fig. 9(c)], perhaps due to this data quantity discrepancy. As such, we repeated these analyses, randomly excluding all but 2999 data points for each gyre. Results were essentially unchanged for all parameters assessed except autocorrelation coefficient, which showed similar trends to the full data set, but was reduced (by roughly 5% for all affected gyres) and more variable for all gyres. Nevertheless, even with equal numbers of data points in each gyre, autocorrelation coefficient was still lower and more variable for NPG than the other gyres. Perhaps the conditions which reduce NPG data quantity (e.g., higher overall cloud cover) also affect the scatter observed in the remaining data.

Relatedly, Fig. 8 shows maximal R^2 of $\sim$0.7 at a span of 1 day ($i = 0.5$ days), such that all means were constructed using data from only one date. This R^2 value indicates 30% of the intraday variability in a gyre is unexplained by a daily running mean. Potential reasons include: 1) sensor noise; 2) atmospheric correction errors; or 3) real intragyre spatial differences in $R_{\rm rs}(547)$ for a given day (i.e., between the 25 points within each gyre). Absent errors in instrument calibration, sensor noise should be randomly distributed in space and time. For item 3), if $R_{\rm rs}(547)$ at specific stations within these gyre centers are independent (e.g., not responding to the same environmental forcings), then it is not appropriate to combine the individual station data into composite gyre time series. To test this point, we repeated all analysis, assessing each of the 125 stations independently. Nearly all $\bar{x}_{R\rm rs}$ for individual stations were within 1% of the corresponding overall gyre mean. Of note, one SPG station did show $\bar{x}_{R\rm rs}$ nearly 8% higher than that for the entire gyre, however replicating all analyses without this station only minimally affected overall results. For individual stations, $\sigma_{R\rm rs}$ was obviously much higher ($\pm$25%) than that for the merged-station time series. Due to additional noise in individual stations resulting from a reduced number of data points relative to the merged data sets, we considered the 91-day models for this analysis. In doing so, we noted that individual stations were very similar to those of the full data set in both magnitude and trend relative to $g(869)$.

As mentioned (item 2 in the previous paragraph), atmospheric correction uncertainties may lead to differences between individual $R_{\rm rs}(547)$ retrievals and corresponding moving mean values. Recall that NIR data are used to select an appropriate atmospheric aerosol model, with L_a(VIS) then calculated through extrapolation from L_a(NIR) [2]–[6]. In this study, data from different gyres were analyzed independently, and only common pixels across all gain configurations were considered. Thus, in this context, atmospheric correction errors are only relevant if they manifest differently (1) within individual gyres [i.e., resulting in intragyre spatial differences in $R_{\rm rs}(547)$], or (2) between gain configurations. For the former, our investigation of individual stations indicated large-scale intragyre stability in this study, in both magnitude and trend. For the latter, atmospheric aerosol characteristics do, indeed, vary according to the gain combination used in satellite data processing. Specifically, in the aggregate, $\tau_a(869)$ increased with an increasing $g(869)$, while α decreased with an increasing $g(869)$, primarily in the summer (Fig. 10). With current default processing [$g(869) = 1.00$], there have been some efforts to validate these products. For MODIS/A validated against global coastal and oceanic AERONET (Aerosol Robotic Network) data, Mélin *et al.* [57] found uncertainties in $\tau_a(869)$ on the order of 26%, with +9% bias. SeaWiFS monthly $\tau_a(865)$ data validated at the Bermuda AERONET station showed root mean square difference (RMSD) on the order of 0.002 [25]. This RMSD is roughly 25% of the overall data mean. In coastal waters near China, Cui *et al*. [58] found MODIS $\tau_a(869)$ uncertainties of 47%. Without replicating these validation exercises, we note that switching from the current default gain configuration for the F2007 gain configuration with $g(869) = 1.022$ resulted in an overall mean increase in $\tau_a(869)$ of 5.4% and a decrease in α of 1.6%.

Recall that the preceding results have been presented in the context of a 31-day moving mean even though we tested spans of 7, 15, 31, 61, 91, 181, and 365 days. Gyre-specific MAD for all spans tested were nearly identical in trend relative to $g(869)$ [Fig. 9(d)]. The one exception was for the 365-day span, which unsurprisingly mirrored the overall data standard deviation [$\sigma_{R\rm rs}$, Fig. 7(b)]. Relative to the 31-day span, MAD magnitude was $\sim$10% higher for the 181-day span, $\sim$10% lower for the seven-day span, and almost equal (within 5%) for the 15-, 61-, and 91-day spans. Differences seen in gyre-specific MAD between 31-day span and all other tested spans were similarly observed for the global MAD parameter. Autocorrelation coefficient for spans of 7- to 91-days showed the same trends relative to $g(869)$ as those seen in Fig. 9(c), with higher spans showing little response in autocorrelation coefficient according to $g(869)$ changes. The autocorrelation coefficient values for spans of 7- to 91-days were always $>$0.995. Similarly, R^2 patterns relative to $g(869)$ were nearly identical for the 7- to 181-day spans. As seen in Fig. 7(a)–(c), an increase in span length corresponded to magnitude decreases in R^2 and increases in autocorrelation coefficient.

Despite these sources of error and potentially confounding factors, from these results we estimate that the "optimal" $g(\rm NIR_L)$ for MODIS/A is $\sim$1.025. Precision about this value, however, is quite low, as both individual-gyre MAD and global MAD show little variability for gain configurations with $g(869)$ between $\sim$1.01 and 1.05 (Fig. 11). Thus, *sensu* Gordon and Wang (2002), we estimate that $R_{\rm rs}(547)$ time series errors in ocean gyres are insensitive to $g(869)$ errors of $\pm$1%–2%. This is not to say that specific $R_{\rm rs}(547)$ retrievals do not differ according to $g(869)$ in this range, or that overall descriptive $R_{\rm rs}(547)$ statistics for particular gyres are stable for all $g(869)$ between $\sim$1.01 and 1.05. Indeed, for example, all gyres show particularly sharp declines in mean $R_{\rm rs}(547)$ in this range [Fig. 7(a)]. Nevertheless, for MODIS data, the current approach is unable to definitively identify an "optimal" $g(\rm NIR_L)$ within $\pm$2%, but $g(\rm NIR_L)$ $\sim$1.025 appears to be an optimal choice within the 1.01–1.05 range.

C. Implications

Recalling that changes in SeaWiFS $g(\mathrm{NIR_L})$ resulted in biases that were hemispherically- and seasonally-dependent (Fig. 1), and observing similar effects via the current analyses, it follows that cross-sensor inconsistencies in $\mathrm{NIR_L}$ calibration could potentially contribute to known cross-sensor discrepancies in retrieved R_{rs} [59]–[61]. Even in the absence of a "perfect" $g(\mathrm{NIR_L})$ for any sensor (Section V-A.), this work highlights the need to ensure $\mathrm{NIR_L}$ agreement between satellite platforms, perhaps via cross-calibration using simultaneous nadir overpasses [20], [62], [63].

Looking within the MODIS/A data set alone, processing of the entire mission using more 'optimal' vicarious gains determined here [e.g., Line 7 of Table III with $g(869) = 1.022$] as compared with the current default processing (with $g(869) = 1.0$), should yield gyre time series with slightly higher temporal coherence [Fig. 9(c)] and lower data scatter [Fig. 9(d)] about a moving mean model with a larger variance [Fig. 9(a)]. Because the impact of climate variability on oceanic C_{a} is subtle and slow [64], [65], small improvements in data product quality may lead to noticeable changes in trend analyses.

VI. Conclusion

Determining SVC gains for the $\mathrm{NIR_L}$ bands of ocean color sensors within acceptable uncertainties has been problematic in the past. Within this study, such difficulties are made more salient, as extensive time-series analyses of MODIS/A $R_{\mathrm{rs}}(547)$ retrievals over ocean gyres reveal sensitivity of the observed R_{rs} to changes in $g(\mathrm{NIR_L})$. Cross-sensor calibration of $\mathrm{NIR_L}$ bands may be one mechanism to reduce any potential artifacts in merged-sensor data sets resulting from the sensitivity of $R_{\mathrm{rs}}(\mathrm{VIS})$ to $g(\mathrm{NIR_L})$. In a larger sense, many results presented here (Figs. 1, 5–7, and 9) highlight uncertainties within the field of satellite ocean color on the exact nature of $R_{\mathrm{rs}}(547)$ dynamics in ocean gyres. Absent an independent assessment of such dynamics, we find that $g(\mathrm{NIR_L}) \sim 1.025$ results in time series which are most 'plausible,' according to our operationalization of a logical argument that "nature does not jump." This $g(869)$ is not too far away from the current 1.0 but can lead to more coherent temporal patterns in ocean gyre $R_{\mathrm{rs}}(547)$ and lower model uncertainties as compared with the current defaults. Likewise, although the change is small, potential implications exist for assessments of ocean gyre dynamics in response to climate variability.

References

[1] H. R. Gordon and A. Y. Morel, *Remote Assessment of Ocean Color for Interpretation of Satellite Visible Imagery: A Review*. New York, NY, USA: Springer-Verlag, 1983.

[2] H. R. Gordon, "Atmospheric correction of ocean color imagery in the Earth Observing System era," *J. Geophys. Res., Atmos.*, vol. 102, no. D14, pp. 17081–17106, Jul. 1997.

[3] H. R. Gordon, "Calibration requirements and methodology for remote sensors viewing the ocean in the visible," *Remote Sens. Environ.*, vol. 22, no. 1, pp. 103–126, Jun. 1987.

[4] 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.*, vol. 33, no. 3, pp. 443–452, 1994.

[5] B. A. Franz, S. W. Bailey, P. J. Werdell, and C. R. McClain, "Sensor-independent approach to the vicarious calibration of satellite ocean color radiometry," *Appl. Opt.*, vol. 46, no. 22, pp. 5068–5082, Aug. 2007.

[6] S. W. Bailey, B. A. Franz, and P. J. Werdell, "Estimation of near-infrared water-leaving reflectance for satellite ocean color data processing," *Opt. Express*, vol. 18, no. 7, pp. 7521–7527, 2010.

[7] B. B. Barnes, J. P. Cannizzaro, D. C. English, and C. Hu, "Validation of VIIRS and MODIS reflectance data in coastal and oceanic waters: An assessment of methods," *Remote Sens. Environ.*, vol. 220, pp. 110–123, Jan. 2019.

[8] F. Mélin, G. Zibordi, and J.-F. Berthon, "Assessment of satellite ocean color products at a coastal site," *Remote Sens. Environ.*, vol. 110, no. 2, pp. 192–215, Sep. 2007.

[9] D. Antoine *et al.*, "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.*, vol. 113, no. C7, Jul. 2008, Art. no. C07013.

[10] G. Zibordi, J.-F. Berthon, F. Mélin, D. D'Alimonte, and S. Kaitala, "Validation of satellite ocean color primary products at optically complex coastal sites: Northern adriatic sea, northern baltic proper and Gulf of Finland," *Remote Sens. Environ.*, vol. 113, no. 12, pp. 2574–2591, Dec. 2009.

[11] S. Maritorena, O. H. F. d'Andon, A. Mangin, and D. A. Siegel, "Merged satellite ocean color data products using a bio-optical model: Characteristics, benefits and issues," *Remote Sens. Environ.*, vol. 114, no. 8, pp. 1791–1804, Aug. 2010.

[12] V. Brando, J. Lovell, E. King, D. Boadle, R. Scott, and T. Schroeder, "The potential of autonomous ship-borne hyperspectral radiometers for the validation of ocean color radiometry data," *Remote Sens.*, vol. 8, no. 2, p. 150, Feb. 2016.

[13] S. Hlaing *et al.*, "Evaluation of the VIIRS ocean color monitoring performance in coastal regions," *Remote Sens. Environ.*, vol. 139, pp. 398–414, Dec. 2013.

[14] M. Wang *et al.*, "VIIRS ocean color products: A progress update," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 5848–5851.

[15] B. B. Barnes and C. Hu, "Dependence of satellite ocean color data products on viewing angles: A comparison between SeaWiFS, MODIS, and VIIRS," *Remote Sens. Environ.*, vol. 175, pp. 120–129, Mar. 2016.

[16] B. B. Barnes and C. Hu, "Cross-sensor continuity of satellite-derived water clarity in the Gulf of Mexico: Insights into temporal aliasing and implications for long-term water clarity assessment," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 1761–1772, Apr. 2015.

[17] C. Hu, B. Barnes, L. Qi, and A. Corcoran, "A harmful algal Bloom of Karenia brevis in the Northeastern Gulf of Mexico as revealed by MODIS and VIIRS: A comparison," *Sensors*, vol. 15, no. 2, pp. 2873–2887, Jan. 2015.

[18] C. Hu and C. Le, "Ocean color continuity from VIIRS measurements over Tampa bay," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 5, pp. 945–949, May 2014.

[19] R.-R. Li *et al.*, "Inter-comparison between VIIRS and MODIS radiances and ocean color data products over the chesapeake bay," *Remote Sens.*, vol. 7, no. 2, pp. 2193–2207, Feb. 2015.

[20] S. Uprety, C. Cao, X. Xiong, S. Blonski, A. Wu, and X. Shao, "Radiometric intercomparison between suomi-NPP VIIRS and Aqua MODIS reflective solar bands using simultaneous nadir overpass in the low latitudes," *J. Atmos. Ocean. Technol.*, vol. 30, no. 12, pp. 2720–2736, Dec. 2013.

[21] M. Wang and H. R. Gordon, "Calibration of ocean color scanners: How much error is acceptable in the near infrared?" *Remote Sens. Environ.*, vol. 82, nos. 2–3, pp. 497–504, Oct. 2002.

[22] F. Mélin and G. Zibordi, "Vicarious calibration of satellite ocean color sensors at two coastal sites," *Appl. Opt.*, vol. 49, no. 5, p. 798, 2010.

[23] S. Hlaing, A. Gilerson, R. Foster, M. Wang, R. Arnone, and S. Ahmed, "Radiometric calibration of ocean color satellite sensors using AERONET-OC data," *Opt. Express*, vol. 22, no. 19, 2014, Art. no. 23385.

[24] 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*. Hanscom Air Force Base, MA, USA: Air Force Geophysics Laboratory, 1979.

[25] Z. Ahmad *et al.*, "New aerosol models for the retrieval of aerosol optical thickness and normalized water-leaving radiances from the SeaWiFS and MODIS sensors over coastal regions and open oceans," *Appl. Opt.*, vol. 49, no. 29, pp. 5545–5560, 2010.

[26] H. R. Gordon and T. Zhang, "How well can radiance reflected from the ocean–atmosphere system be predicted from measurements at the sea surface?" *Appl. Opt.*, vol. 35, no. 33, p. 6527, 1996.
[27] N. Martiny, R. Frouin, and R. Santer, "Radiometric calibration of SeaWiFS in the near infrared," *Appl. Opt.*, vol. 44, no. 36, p. 7828, 2006.
[28] R. Santer and N. Martiny, "Sky-radiance measurements for ocean-color calibration-validation," *Appl. Opt.*, vol. 42, no. 6, pp. 896–907, 2003.
[29] 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.*, vol. 102, nos. 1–2, pp. 12–23, May 2006.
[30] S. B. Hooker and S. Maritorena, "An evaluation of oceanographic radiometers and deployment methodologies," *J. Atmos. Ocean. Technol.*, vol. 17, no. 6, pp. 811–830, Jun. 2000.
[31] Z. Lee *et al.*, "Time series of bio-optical properties in a subtropical gyre: Implications for the evaluation of interannual trends of biogeochemical properties," *J. Geophys. Res.*, vol. 115, no. C9, pp. 1–13, 2010.
[32] S. R. Signorini, B. A. Franz, and C. R. McClain, "Chlorophyll variability in the oligotrophic gyres: Mechanisms, seasonality and trends," *Frontiers Mar. Sci.*, vol. 2, pp. 1–11, Feb. 2015.
[33] S. R. Signorini and C. R. McClain, "Subtropical gyre variability as seen from satellites," *Remote Sens. Lett.*, vol. 3, no. 6, pp. 471–479, Nov. 2012.
[34] A. Ibrahim, B. A. Franz, Z. Ahmad, and S. W. Bailey, "Multi-band atmospheric correction algorithm for ocean color retrievals," *Frontiers Earth Sci.*, vol. 7, p. 116, May 2019.
[35] M. Wang, "The Rayleigh lookup tables for the SeaWiFS data processing: Accounting for the effects of ocean surface roughness," *Int. J. Remote Sens.*, vol. 23, no. 13, pp. 2693–2702, Jan. 2002.
[36] K. D. Moore, K. J. Voss, and H. R. Gordon, "Spectral reflectance of whitecaps: Their contribution to water-leaving radiance," *J. Geophys. Res., Oceans*, vol. 105, no. C3, pp. 6493–6499, Mar. 2000.
[37] R. Frouin, M. Schwindling, and P.-Y. Deschamps, "Spectral reflectance of sea foam in the visible and near-infrared: *In situ* measurements and remote sensing implications," *J. Geophys. Res.*, vol. 101, no. C6, 1996, Art. no. 14361.
[38] 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.*, vol. 44, no. 26, p. 5524, 2005.
[39] C. Hu, Z. Lee, and B. Franz, "Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference," *J. Geophys. Res., Oceans*, vol. 117, no. C1, pp. 1–25, Jan. 2012.
[40] A. Bricaud, A. Morel, and L. Prieur, "Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domains," *Limnology Oceanogr.*, vol. 26, no. 1, pp. 43–53, Jan. 1981.
[41] C. Hu *et al.*, "Dynamic range and sensitivity requirements of satellite ocean color sensors: Learning from the past," *Appl. Opt.*, vol. 51, no. 25, pp. 6045–6062, Sep. 2012.
[42] A. Morel, H. Claustre, and B. Gentili, "The most oligotrophic subtropical zones of the global ocean: Similarities and differences in terms of chlorophyll and yellow substance," *Biogeosciences*, vol. 7, no. 10, pp. 3139–3151, Oct. 2010.
[43] R. M. Letelier, R. R. Bidigare, D. V. Hebel, M. Ondrusek, C. D. Winn, and D. M. Karl, "Temporal variability of phytoplankton community structure based on pigment analysis," *Limnol. Oceanogr.*, vol. 38, no. 7, pp. 1420–1437, Nov. 1993.
[44] C. D. Winn *et al.*, "Seasonal variability in the phytoplankton community of the North Pacific subtropical gyre," *Global Biogeochem. Cycles*, vol. 9, no. 4, pp. 605–620, Dec. 1995.
[45] A. Mignot, H. Claustre, J. Uitz, A. Poteau, F. D'Ortenzio, and X. Xing, "Understanding the seasonal dynamics of phytoplankton biomass and the deep chlorophyll maximum in oligotrophic environments: A bio-argo float investigation," *Global Biogeochem. Cycles*, vol. 28, no. 8, pp. 856–876, Aug. 2014.
[46] R. G. Williams and M. J. Follows, "The Ekman transfer of nutrients and maintenance of new production over the North Atlantic," *Deep Sea Res. I, Oceanographic Res. Papers*, vol. 45, nos. 2–3, pp. 461–489, Feb. 1998.
[47] G. Thuillier *et al.*, "The solar spectral irradiance from 200 to 2400 nm as measured by the SOLSPEC spectrometer from the atlas and Eureca missions," *Solar Phys.*, vol. 214, no. 1, pp. 1–22, 2003.
[48] W. L. Barnes, T. S. Pagano, and V. V. Salomonson, "Prelaunch characteristics of the Moderate Resolution Imaging Spectroradiometer (MODIS) on EOS-AM1," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 4, pp. 1088–1100, Jul. 1998.
[49] H. R. Gordon, "Remote sensing of ocean color: A methodology for dealing with broad spectral bands and significant out-of-band response," *Appl. Opt.*, vol. 34, no. 36, p. 8363, 1995.
[50] 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.*, vol. 40, no. 3, p. 343, 2001.
[51] A. Morel and B. Gentili, "Diffuse reflectance of oceanic waters: Its dependence on sun angle as influenced by the molecular scattering contribution," *Appl. Opt.*, vol. 30, no. 30, p. 4427, 1991.
[52] A. Morel and B. Gentili, "Diffuse reflectance of oceanic waters II bidirectional aspects," *Appl. Opt.*, vol. 32, no. 33, p. 6864, 1993.
[53] A. Morel and B. Gentili, "Diffuse reflectance of oceanic waters III implication of bidirectionality for the remote-sensing problem," *Appl. Opt.*, vol. 35, no. 24, p. 4850, 1996.
[54] S. B. Hooker and W. E. Esaias, "An overview of the SeaWiFS project," *Eos, Trans. Amer. Geophys. Union*, vol. 74, no. 21, pp. 241–246, 1993.
[55] G. Zibordi and F. Mélin, "An evaluation of marine regions relevant for ocean color system vicarious calibration," *Remote Sens. Environ.*, vol. 190, pp. 122–136, Mar. 2017.
[56] G. Zibordi *et al.*, "System vicarious calibration for ocean color climate change applications: Requirements for *in situ* data," *Remote Sens. Environ.*, vol. 159, pp. 361–369, Mar. 2015.
[57] F. Mélin, M. Clerici, G. Zibordi, B. N. Holben, and A. Smirnov, "Validation of SeaWiFS and MODIS aerosol products with globally distributed AERONET data," *Remote Sens. Environ.*, vol. 114, no. 2, pp. 230–250, Feb. 2010.
[58] T. Cui *et al.*, "Assessment of satellite ocean color products of MERIS, MODIS and SeaWiFS along the East China coast (in the Yellow Sea and East China Sea)," *ISPRS J. Photogramm. Remote Sens.*, vol. 87, pp. 137–151, Jan. 2014.
[59] S. Djavidnia, F. Mélin, and N. Hoepffner, "Comparison of global ocean colour data records," *Ocean Sci.*, vol. 6, no. 1, pp. 61–76, Jan. 2010.
[60] F. Mélin, G. Sclep, T. Jackson, and S. Sathyendranath, "Uncertainty estimates of remote sensing reflectance derived from comparison of ocean color satellite data sets," *Remote Sens. Environ.*, vol. 177, pp. 107–124, May 2016.
[61] C. Zhang *et al.*, "Bridging between SeaWiFS and MODIS for continuity of chlorophyll–a concentration assessments off Southeastern China," *Remote Sens. Environ.*, vol. 102, nos. 3–4, pp. 250–263, Jun. 2006.
[62] N. Pahlevan, Z. Lee, J. Wei, C. B. Schaaf, J. R. Schott, and A. Berk, "On-orbit radiometric characterization of OLI (Landsat-8) for applications in aquatic remote sensing," *Remote Sens. Environ.*, vol. 154, pp. 272–284, Nov. 2014.
[63] N. Pahlevan *et al.*, "Landsat 8 remote sensing reflectance (Rrs) products: Evaluations, intercomparisons, and enhancements," *Remote Sens. Environ.*, vol. 190, pp. 289–301, Mar. 2017.
[64] W. W. Gregg and M. E. Conkright, "Decadal changes in global ocean chlorophyll," *Geophys. Res. Lett.*, vol. 29, no. 15, pp. 20-1–20-4, Aug. 2002.
[65] D. Antoine, "Bridging ocean color observations of the 1980s and 2000s in search of long-term trends," *J. Geophys. Res.*, vol. 110, no. C6, pp. 1–22, 2005.

Brian B. Barnes received the B.S. degree in zoology and psychology from the University of Florida, Gainesville, FL, USA, in 2004, the M.S. degree in marine science from the College of William and Mary, Virginia Institute of Marine Science, Gloucester Point, VA, USA, in 2009, and the Ph.D. degree in marine science from the University of South Florida, St. Petersburg, FL, USA, in 2013.

He is a Research Associate with the Optical Oceanography Laboratory, College of Marine Science, University of South Florida. His research interests include calibration and validation of ocean color satellite data toward enhanced coastal research and monitoring.

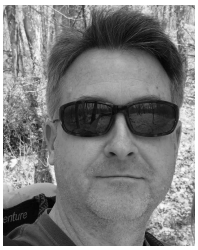
Chuanmin Hu received the B.S. degree in physics from the University of Science and Technology of China, Hefei, China, in 1989, and the Ph.D. degree in physics (environmental optics) from the University of Miami, FL, USA, in 1997.

He is a Professor of optical oceanography with the University of South Florida, St. Petersburg, FL, USA, who also directs the Optical Oceanography Lab. He uses laboratory, field, and remote sensing techniques to study marine algal blooms (harmful and nonharmful, macroalgae and microalgae), oil spills, coastal and inland water quality, and global changes. He has authored and coauthored more than 270 refereed articles, many of which have been highlighted on journal covers and by AGU and NASA.

Dr. Hu led the group effort to establish a Virtual Buoy System (VBS), an Integrated Redtide Information System (IRIS), and a Sargassum Watch System (SaWS) to monitor water quality and algal blooms. From 2009 to 2014, he served as a Topical Editor on ocean optics and ocean color remote sensing at *Applied Optics*. From 2015 to 2017, he served as a Chief Editor for *Remote Sensing of Environment*.

Sean W. Bailey received the B.Sc. degree in marine science and biology from the University of Miami, FL, USA, in 1992, and the M.Sc. degree in biological oceanography from the University of Southern Mississippi, Hattiesburg, MS, USA, in 1997.

In 1997, he joined the NASA Goddard Space Flight Center, Greenbelt, MD, USA, as a Contractor, working for what is currently the Ocean Biology Processing Group. In 2015, he became a Civil Servant with NASA, where he serves as the Deputy Manager for the Ocean Biology Distributed Active Archive Center. His research interests are focused on bio-optical remote sensing of the oceans, specifically with the on-orbit calibration of satellite-based sensors and the validation of their data products.



Bryan A. Franz has supported research and operations for satellite remote sensing of Ocean Biology and Biogeochemistry, NASA, including the development of science data products from the Sea-viewing Wide Field-of-view Sensor (SeaWiFS), Moderate Resolution Imaging Spectroradiometer (MODIS), Medium Resolution Imaging Spectrometer (MERIS), Visible and Infrared Imaging Radiometer Suite (VIIRS), and other satellite-based remote sensing radiometers, since 1996. Since 2009, he has served as the Ocean Discipline Leader for the MODIS Science Team and the combined MODIS and VIIRS Science Team, and he is leading the development of the Science Data Segment (SDS) for the Plankton, Aerosol, Cloud, Ocean Ecosystem (PACE) Mission. He is a Research Scientist and an Assistant Chief for Science Research with the Ocean Ecology Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA. His work at NASA has focused on standardization of atmospheric correction and bio-optical algorithms, vicarious calibration, and instrument cross-calibration techniques to derive consistent, multimission time-series of ocean optical and biogeochemical properties.