

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/281107714>

Spectral Information Adaptation and Synthesis Scheme for Merging Cross-Mission Ocean Color Reflectance Observations...

Article in IEEE Transactions on Geoscience and Remote Sensing · June 2015

DOI: 10.1109/TGRS.2015.2456906

CITATIONS

4

READS

127

3 authors, including:



[Kaixu Bai](#)

East China Normal University

15 PUBLICATIONS 30 CITATIONS

[SEE PROFILE](#)



[Ni-Bin Chang](#)

University of Central Florida

329 PUBLICATIONS 6,223 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Systems analysis applied to solid waste management [View project](#)

Spectral Information Adaptation and Synthesis Scheme for Merging Cross-Mission Ocean Color Reflectance Observations From MODIS and VIIRS

Kaixu Bai, Ni-Bin Chang, *Senior Member, IEEE*, and Chi-Farn Chen

Abstract—Obtaining a full clear view of coastal bays, estuaries, lakes, and inland waters is challenging with single satellite sensor observations due to cloud impacts. Cross-mission sensors provide the synergistic opportunity to improve spatial and temporal coverage by merging their observations; however, discrepancies originating from the instrumental, algorithmic, and temporal differences should be eliminated before merging. This paper presents the Spectral Information Adaptation and Synthesis Scheme (SIASS) for generating cross-mission consistent ocean color reflectance by merging 2012–2015 observations from Moderate Resolution Imaging Spectroradiometer and Visible Infrared Imaging Radiometer Suite over Lake Nicaragua in Central America, where the cloud impact is salient. The SIASS is able to not only eliminate incompatibilities for matchup bands but also reconstruct spectral information for mismatched bands among sensors. Statistics indicate that the average monthly coverage of a merged ocean color reflectance product over Lake Nicaragua is nearly twice that of any single-sensor observation. Results show that SIASS significantly improves consistency among cross-mission sensors by mitigating prominent discrepancies. In addition, reconstructed spectral information for those mismatched bands help preserve more spectral characteristics needed to better monitor and understand the dynamic aquatic environment. The final implementation of SIASS to map the chlorophyll-*a* concentration demonstrates the efficacy of SIASS in bias correction and consistency improvement. In general, SIASS can be applied to remove cross-mission discrepancies among sensors to improve the overall consistency.

Index Terms—Data merging, Moderate Resolution Imaging Spectroradiometer (MODIS), ocean color, remote sensing, Visible Infrared Imaging Radiometer Suite (VIIRS).

I. INTRODUCTION

REMOTE sensing in an aquatic environment is challenging due to the negative impacts from aerosols, sun glint, clouds, and other factors during data collection [1]–[3]. For this reason, obtaining full clear coverage of the interest area with single-sensor observation is difficult, particularly over the

tropical regions where dense cloud cover is frequent. Cross-mission satellite sensors with similar characteristics provide a synergistic opportunity to improve spatial and temporal coverage by coalescing multiple mission satellite observations into a single quality unified product, namely, data merging [4], [5]. Due to differences in instrument design and retrieval algorithms, the consistency or continuity among cross-mission sensors should be investigated before establishing a consistent data record, particularly when generating a long-term coherent data record [6], [7]. Considering different focuses in real-world applications, the most advisable merging scheme is to merge the primary radiometric products, such as the ocean color remote sensing reflectance (or normalized water-leaving radiance). These radiometric products form the baseline information for deriving all higher level data products to aid in environmental decision making, even those products that are not included in the current operational product list [8].

Accurate merging of primary radiometric data products must account for discrepancies resulting from instrumental (e.g., different sensor design, center wavelengths, and bandwidths), algorithmic (e.g., calibration, atmospheric correction, and aerosol models), and temporal (e.g., local overpassing time) differences among cross-mission sensor observations. To address spectral distortion due to mismatched center wavelengths, an optically based technique was applied to compute radiometric quantities at the desired center wavelengths by solving an inverse radiative transfer problem [9]. In addition, a semianalytical bio-optical ocean color data merging model was employed to produce global retrievals of three biogeochemically relevant variables using the normalized water-leaving radiance from Sea-Viewing Wide Field-of-View Sensor (SeaWiFS) and Moderate Resolution Imaging Spectroradiometer (MODIS) [10]. By applying the same bio-optical merging model, observations from SeaWiFS, MODIS, and Medium Resolution Imaging Spectrometer (MERIS) were merged to create a coherent long-time series of ocean color products under the NASA Ocean Color MEaSURES and the ESA GlobColour projects [11]. Statistics indicate that the merged products have better spatial and temporal coverage than each individual mission; the average daily global ocean coverage of merged data products is nearly twice that of any single mission observation. These benefits greatly promote the application potential of ocean color data merging approaches.

In addition to these analytical merging models, empirical methods were also introduced to merge multiple ocean color

Manuscript received March 24, 2015; revised May 18, 2015; accepted June 26, 2015.

K. Bai is with the Key Laboratory of Geographic Information Science, Ministry of Education, East China Normal University, Shanghai 200241, China.

N.-B. Chang is with the Department of Civil, Environmental, and Construction Engineering, University of Central Florida, Orlando, FL 32816 USA (e-mail: nchang@ucf.edu).

C.-F. Chen is with the Center for Space and Remote Sensing Research, National Central University, Zhongli 32001, Taiwan.

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

Digital Object Identifier 10.1109/TGRS.2015.2456906

sensor observations. A machine learning technique was used to eliminate discrepancies among chlorophyll-*a* (chlor-*a* hereafter) concentrations derived from MODIS and SeaWiFS to produce consistent daily global ocean color coverage [12]. Accurate results were obtained for low chlor-*a* concentrations, which suggested the potential of this method to eliminate incompatibility resulting from sensors, such as scan angle dependencies and seasonal and spatial trends in data. The noisy and highly dynamic nature of differences between cross-mission sensors, however, makes empirical methods difficult to extend this scheme onto new time and space domains different from those used for training in the machine learning process. Meanwhile, statistical methods such as the multilinear regression algorithm were also utilized to remove systematic biases among sensors by projecting the MODIS and SeaWiFS observations onto the *in situ* measurements. This correction scheme improved the consistency between cross-mission sensors with the salient effects at blue bands (i.e., 412 and 443 nm). Similar approaches were also used to create scientifically reliable ocean color climate data records [13], as well as post-launch radiometric recalibration [14].

The aforementioned approaches, either analytical or empirical, to some extent are all capable of reducing discrepancies among cross-mission sensors to improve consistency. Nevertheless, these methods still have some limitations. On one hand, the optical remote sensing techniques require solving a complex inverse radiative transfer problem. On the other hand the machine-learning-based methods need abundant relevant information for data retrieval in addition to the radiometric products, such as satellite viewing geometry, solar zenith angle, aerosol optical depth, ozone amount, and water vapor, which is not always available in association with radiometric products. Those machine-learning-based methods are highly dependent on the *in situ* measurements for cross calibration, whereas long-term *in situ* measurements are spatially scarce, which limits the broad applications of these machine-learning-based methods on a large scale.

At present, these *in situ* measurement-based cross-calibration schemes can only remove the systematic bias resulting from the instrumental and algorithmic differences among sensors. In other words, they cannot account for the location-dependent bias related to hydrodynamic factors. For example, the bias between two cross-mission sensors would not be the same at two different locations in one lake if one measurement is located in the outflow with strong advection of water mass, whereas the other is located in the center of lake with small dynamics. This bias is also difficult to measure due to the stochastic nature embedded in the temporal difference among sensors.

In regard to the merging of primary radiometric products, previous approaches only considered merging observations at the common bands, meaning that some valuable spectral information was abandoned during the merging process. For example, no observation is recorded at 531-nm wavelength by VIIRS, but observations are recorded in the MODIS product associated with this band. To preserve the unique spectral characteristics at this wavelength while merging VIIRS with MODIS, spectral information at 531 nm should be reconstructed from available neighboring bands in VIIRS observations [15]. Although band-shifting methods are able to address this issue,

it still requires to use the complex bio-optical model in some occasions [16].

In this paper, a new method called Spectral Information Adaptation and Synthesis Scheme (SIASS) was developed to overcome the barriers described above. The strength of SIASS is the development of a generalized scheme for bias correction between cross-mission sensors, relying on their common observations collected during the overlapped time periods. With the aid of SIASS, cross-mission discrepancies are removed without using any *in situ* measurements other than sensors common observations. Compared with previous methods, SIASS is able to not only eliminate incompatibilities between the common bands but also reconstruct spectral information for those mismatched bands among sensors. Practical implementation of the SIASS method was confirmed by applying it to merge cross-mission ocean color reflectance observations from MODIS and VIIRS over Lake Nicaragua during the time period of 2012–2015 in this study.

II. DATA AND METHODS

A. Study Area and Data Sources

Lake Nicaragua was chosen as the study area because of its location in a tropical region where dense clouds are present throughout the year (see Fig. 1). This lake is the largest freshwater lake in Central America, with an area of 8264 km². The two general seasons are the dry season from November to April and the wet season from May to October. Statistics show that dense cloud cover is mainly observed during the wet season, with an annual average cloud cover of nearly 70% [17]. As the lake has been considered as a future drinking water source by the Nicaragua government and several other Central American countries, monitoring biophysical parameters to characterize water quality conditions and pollution levels in this lake on a near-real-time basis is thus critical.

Traditional manual sampling of water quality parameters is labor-intensive and costly, as well as incapable of capturing vast spatial variability simultaneously. To provide instantaneous insight into the phenomena of interest with a large spatial coverage, daily recorded satellite ocean color observations should be used. Currently, several satellite sensors, mainly MODIS aboard the Terra (1999–present, MODIS-Terra hereafter) and the Aqua (2002–present, MODIS-Aqua hereafter), as well as VIIRS aboard the S-NPP (2011–present), are orbiting the Earth with a capability of monitoring the ocean color and inland waters (see Table I). These sensors map the Earth on a daily basis at a moderate resolution. The general performance and accuracy of ocean color products derived from MODIS-Aqua, MODIS-Terra, and VIIRS evaluated in the literature indicate comparable accuracies between the sensors at most wavelengths [18], [19]. To have sufficient samples for the present study, the common Level-2 ocean color reflectance observations over Lake Nicaragua collected from these three sensors during the 2012–2015 overlapped time period were used without flag filtering. The quality control measure was limited to remove all negative values from these reflectance products. These radiometric products have previously been radiometric and atmospheric corrected to remove remaining

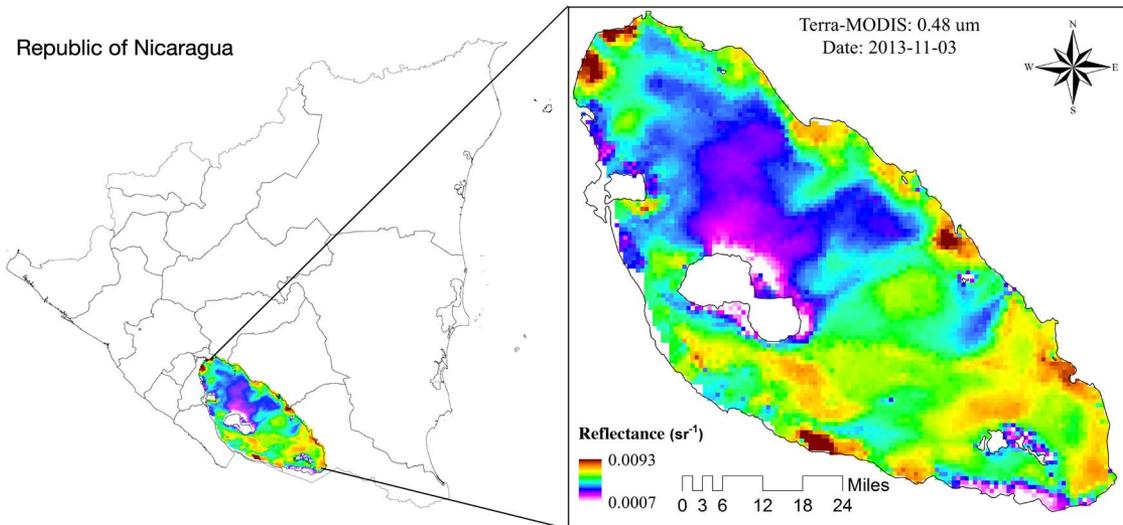


Fig. 1. Location of Lake Nicaragua.

TABLE I
CHARACTERISTICS OF MODIS AND VIIRS OCEAN COLOR REMOTE SENSING REFLECTANCE PRODUCTS. THE LEVEL-2 OCEAN COLOR REFLECTANCE DATA OF MODIS-TERRA, MODIS-AQUA, AND VIIRS USED IN THIS STUDY WERE ACQUIRED FROM THE NASA OCEAN BIOLOGY PROCESSING GROUP (OBPG)

| | MODIS | VIIRS |
|------------------------|--|--|
| Wavelengths (nm) | 412: 405-420 443: 438-448 469: 459-479 488: 483-493 531: 526-536 547: 546-556 555: 545-565 645: 620-670 667: 662-672 678: 673-683 | 410: 402-422 443: 436-454 486: 478-488 551: 545-565 671: 662-682 |
| Spatial Resolution (m) | 1,000 | 750 |
| Temporal Resolution | Daily | Daily |
| Local Crossing Time | MODIS-Terra: 10:30 AM MODIS-Aqua: 1:30 PM | 1:30 PM |
| Data Format | HDF | NetCDF |
| Data Access | NASA OBPG | NASA OBPG |

effects of solar orientation and atmospheric attenuation due to scattering and aerosols [1], [20], [21]. To unify the spatial resolution, VIIRS ocean color reflectance observations were resampled to the MODIS resolution (i.e., 1000 m) using the SeaWiFS Data Analysis System (SeaDAS) package (version 7.1) during reprojection processes. Both MODIS and VIIRS ocean color reflectance observations were mapped onto the UTM-16N projected coordinate system along with the WGS-84 geographic coordinate system.

B. SIASS Method

The main objective of SIASS is to remove cross-mission biases among sensors for better image fusion. Common ob-

servations between successive generations of sensors during the overlapped time period provide a synergistic opportunity for sensor intercomparisons [22]. With this advantage, biases across missions can be quantified and then removed to improve consistency among cross-mission sensor observations. An observation from one sensor can be calibrated to be consistent with the other sensor by removing the systematic bias and the associated location-dependent bias. Following this basic theory, the SIASS is designed to eliminate biases between cross-mission sensor observations for wavelengths in the synchronized bands while reconstructing spectral information for wavelengths of mismatched bands. Generally, for each pixel over a geographical grid (i.e., a geographical grid is a squared area to pinpoint any location on Earth with unique geographic information (latitude/longitude), and it is always referred to as a pixel in remote sensing) common to cross-mission sensors, the cross-mission ocean color reflectance bias at wavelength λ (denoted by $\Delta Rrs(\lambda)$ hereafter) is assumed to mainly consist of two portions, the sensor-dependent systematic bias ($\Delta Rrs^{SAT}(\lambda)$) and the location-dependent bias ($\Delta Rrs^{LCT}(\lambda)$), i.e.,

$$\langle \Delta Rrs(\lambda) \rangle = \langle \Delta Rrs^{SAT}(\lambda) \rangle + \langle \Delta Rrs^{LCT}(\lambda) \rangle \quad (1)$$

where $\langle \rangle$ denotes the ensemble mean.

The rationale of the SIASS method mainly exploits an adaptive bias correction scheme to remove aforementioned systematic bias and location-dependent bias among sensors simultaneously via the proper integration of spectral feature extraction, quantile-quantile (Q-Q) adjustment, empirical mode decomposition (EMD), and extreme learning machine (ELM). The systematic bias can be processed through spectral feature extraction and Q-Q adjustment in sequence, and the location-dependent bias may be removed by using EMD and ELM as a whole. Within the systematic bias correction, spectral feature extraction is designed to process a handful of pixels with no cloud contamination and retrieve the major spectral information generating two referenced data sets for both baseline and

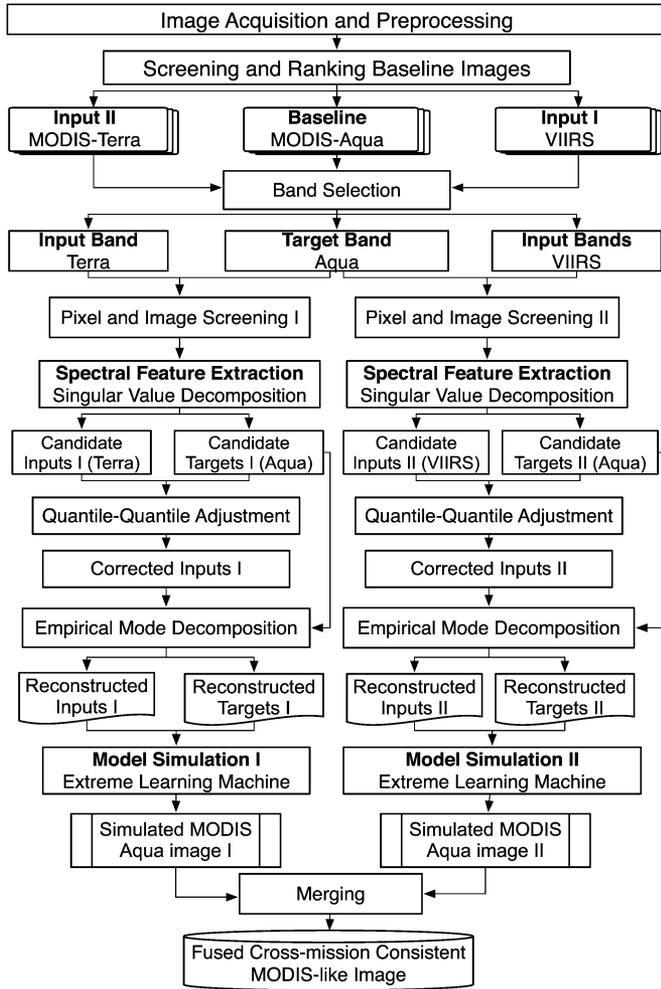


Fig. 2. Schematic flowchart of the cross-mission ocean color reflectance merging scheme with SIASS.

complementary satellite sensors, respectively, from which the bias correction relationship may be signified. Given any observed data set, the systematic bias can be eliminated by the Q–Q adjustment globally based on the two relevant referenced data sets. Because the decomposition of the EMD is based on the local characteristic time scale of the data, it can be applied to analyze some nonlinear and nonstationary processes in hydrodynamic environments and help for addressing the location-dependent bias. After the systematic bias correction, EMD is applied to remove white noise-like high-frequency fluctuations embedded in the two pairwise time series systematic bias-corrected reference data sets and to reconstruct them into new data sets for location-dependent bias correction. Finally, ELM is applied to build relationships between the pairwise reconstructed data sets to eliminate the location-dependent bias pixel by pixel.

Here, a schematic flowchart is presented to demonstrate the SIASS procedure applied to merge ocean color reflectance from three cross-mission satellite sensors: MODIS-Terra, MODIS-Aqua, and VIIRS (see Fig. 2). For a set of three sensors, one of them must be chosen as the baseline, while the others are complementary. In other words, observations collected by the remaining sensors must be projected to be consistent

with the observations up to the baseline sensor level. The baseline sensor can be screened by using three predetermined criteria. The purpose is to maximize the spatial coverage of clear pixels (i.e., pixels having valid observation values) in the lake before merging cross-mission ocean color reflectance from three sensor observations. First, any single sensor that can show the highest number of clear pixels should have a comparative advantage. Second, any pairwise sensors (MODIS-Aqua and MODIS-Terra, abbreviated as AT hereafter; MODIS-Aqua and VIIRS, abbreviated as AV hereafter; MODIS-Terra and VIIRS, abbreviated as TV hereafter) that may contribute to more clear pixels with a higher coverage ratio collectively may gain more comparative advantage. Finally, with such three pairwise settings (i.e., AT, AV, and TV), the smaller the difference between ocean color reflectance observations at the common band (i.e., the same wavelength and 443 nm in this study) among sensors (i.e., lower relative biases between sensors), the more the comparative advantage.

A graphical sketch of the SIASS method for bias correction is presented in Fig. 3. According to the predetermined criteria as described above, MODIS-Aqua is selected as the baseline sensor, whereas MODIS-Terra and VIIRS are selected as the complementary satellite sensors. For those mismatched bands between baseline and complementary satellite sensors, two bands in the same neighborhood of complementary sensors may be selected as an alternative to simulate and generate the same band information. As an example, two VIIRS bands may be projected to a relevant band of MODIS-Aqua for bias correction. For the purpose of demonstration, two bands associated with VIIRS and one band associated with MODIS-Aqua were used in Fig. 3 (see Fig. 3, top) to illustrate how the new band information can be generated by removing associated bias. The four key methods in SIASS can be delineated in a greater detail as follows.

1) *Spectral Feature Extraction*: In the image screening step for both the baseline and complementary satellite sensors, each clear pixel within the overlapped time period may be screened out for the construction of a reference time-series database. The reference time-series database for each pixel may have a good memory effect of the hidden patterns of systematic bias so that the cumulative distribution function (cdf) can be computed for holistic consideration of bias correction. By picking up the candidate pixels for adjustment one by one over the entire study region, the individual cdf associated with candidate pixel values over the study period can also be generated. When addressing each pixel on a rolling basis, the bias correction for the entire study region may be achieved stepwise by using the Q–Q adjustment to eliminate systematic bias.

Before the generation of each cdf curve, spectral feature extraction must be performed using singular value decomposition (i.e., a signal processing scheme) to create a reference time series for subsequent characterization of bias among sensors (see Fig. 3, bottom). The reference time series is the first principal component (PC) of all matchups in the overlapped study time period and regions. Following (4)–(12), each pixel time series can be fully corrected to eliminate the systematic bias through the Q–Q adjustment (see Fig. 3, middle).

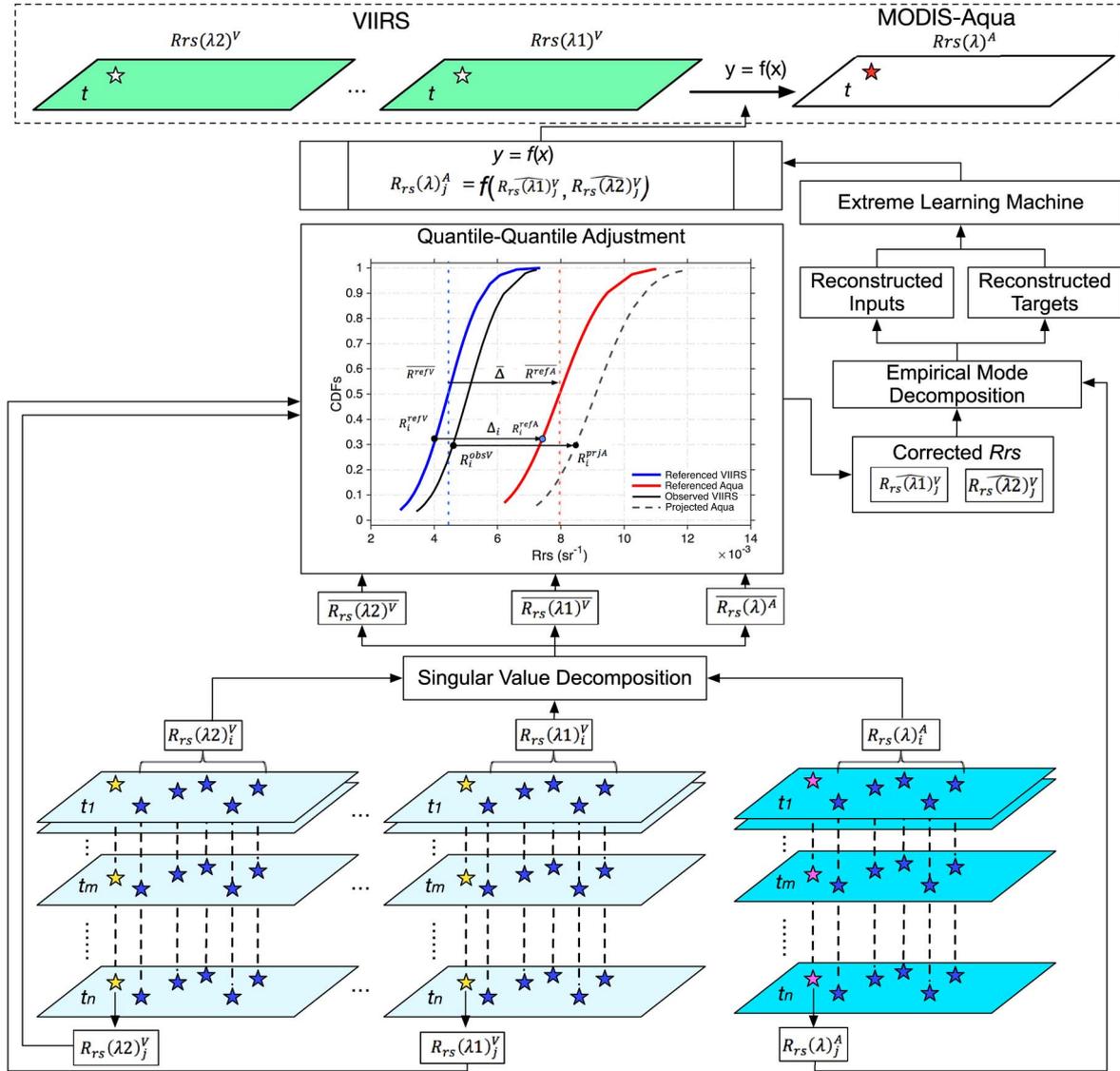


Fig. 3. Graphical sketch of the core operation of the SIASS method for bias correction.

2) *Q-Q Adjustment*: As aforementioned, systematic bias (i.e., $\Delta Rrs^{\text{SAT}}(\lambda)$) results mainly from instrumental and algorithmic differences among sensors. Ideally, this bias is nearly consistent and thus can be quantified through the sensor inter-comparisons. The simple way to characterize this systematic bias is to calculate the average differences through a matchup analysis, indicated as

$$\Delta Rrs^{\text{SAT}}(\lambda) = \frac{1}{N} \sum_{i=1}^N (Rrs_i^{\text{SAT1}}(\lambda) - Rrs_i^{\text{SAT2}}(\lambda)) \quad (2)$$

$$\langle \Delta Rrs^{\text{SAT}}(\lambda) \rangle = \frac{1}{T} \sum_{t=1}^T (Rrs_t^{\text{SAT}}(\lambda)) \quad (3)$$

where i denotes the number of matchups among sensors on the same date, and N is the total number of i . t denotes the time (day number), and T is the total number of t .

In practice, due to different instrumental response functions and the degradation of sensors along with the time, this bias cannot be always consistent. Therefore, simply adding the average differences cannot eliminate discrepancies among

cross-mission sensors; on the contrary, new biases could be introduced. To overcome this weakness, we adopted the Q-Q adjustment method, an adaptive method originally proposed to calibrate the projected regional climate model outputs to the observed local scales [23]. This method has proven effective in calibrating model projections of climate parameters, such as air temperature [24] and precipitation [25].

Similarly, this theory can be adopted to calibrate cross-mission sensor observations to reproduce observations from one sensor (i.e., complementary satellite sensor) based on the given data from the other sensor (i.e., the baseline satellite sensor). Here, the Q-Q adjustment is utilized to characterize the sensor-dependent systematic bias. Based on the Q-Q adjustment theory (see Fig. 4), one observation from VIIRS (i.e., Rrs_i^{obsV}) can be calibrated to the MODIS-Aqua level by adding the associated cross-mission bias [ΔRrs , formulated in (1)] between these two sets of sensor observations pixel by pixel, which can be modeled as

$$Rrs_i^{\text{prjA}} = Rrs_i^{\text{obsV}} + \Delta Rrs_i \quad (4)$$

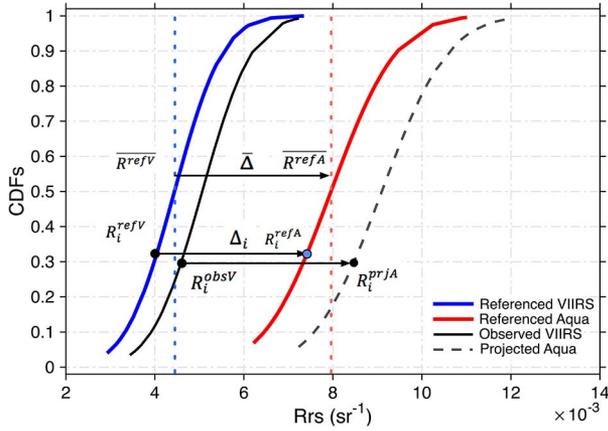


Fig. 4. Illustrative example of the Q-Q adjustment method for calibrating cross-mission sensor observations. Observations from one sensor (solid black contour) can be calibrated to the other sensor level (dashed black contour) by removing associated biases, which can be characterized from common differences between cross-mission sensor observations during the overlapped time period (blue and red contours).

where Rrs_i^{prjA} is the projected observation after bias correction, and i is the i th percentiles in cdf's of Rrs_i^{obsV} . In Fig. 4, referenced VIIRS time-series data stand for common observations of VIIRS after the treatment of spectral feature extraction during the overlapped time period with MODIS-Aqua. Similarly, referenced Aqua time series stand for common observations of MODIS-Aqua after the treatment of spectral feature extraction during the overlapped time period with VIIRS. Observed VIIRS time-series data stand for raw VIIRS observations with no further treatment. Projected Aqua time-series data stand for projected data at the Aqua level from the observed VIIRS observations.

In ΔRrs_i , the associated systematic bias at the i th percentiles (ΔRrs_i^{SAT}) can be expressed as the sum of the mean radiometric shift ($\bar{\Delta}$) plus the corresponding deviation Δ'_i , i.e.,

$$\Delta Rrs_i^{SAT} = g\bar{\Delta} + f\Delta'_i \quad (5)$$

where

$$\Delta_i = Rrs_i^{refA} - Rrs_i^{refV} \quad (6)$$

$$\bar{\Delta} = \frac{1}{N} \sum_{i=1}^N \Delta_i = \overline{Rrs^{refA}} - \overline{Rrs^{refV}} \quad (7)$$

$$\Delta'_i = \Delta_i - \bar{\Delta} \quad (8)$$

$$g = \frac{\left(\sum_{i=1}^N Rrs_i^{obsV} \right) / N}{\left(\sum_{i=1}^N Rrs_i^{refV} \right) / N} = \frac{\overline{Rrs^{obsV}}}{\overline{Rrs^{refV}}} \quad (9)$$

$$f = \frac{IQR_{Rrs^{obsV}}}{IQR_{Rrs^{refV}}} \quad (10)$$

$$IQR_{Rrs^{obsV}} = Rrs^{obsV}|_{p=75\%} - Rrs^{obsV}|_{p=25\%} \quad (11)$$

$$IQR_{Rrs^{refV}} = Rrs^{refV}|_{p=75\%} - Rrs^{refV}|_{p=25\%}. \quad (12)$$

Terms $IQR_{Rrs^{obsV}}$ and $IQR_{Rrs^{refV}}$ in (10)–(12) are the interquartile ranges of the observed and referenced observations, re-

spectively, calculated as the differences between the 75th ($p = 75\%$) and 25th ($p = 25\%$) percentiles.

In this paper, Rrs^{refA} and Rrs^{refV} were derived from the matchups of common observations of both sensor observations after the spectral feature extraction. The basic underlying theory is that induced discrepancies due to instrumental and algorithmic differences may be memorized along with the time, which can be characterized from the long-term historical observations (see Fig. 3, bottom panel). From (5)–(12), it is observed that ΔRrs_i^{SAT} is mainly modulated by g and f ; if $g = f = 1$, ΔRrs_i^{SAT} would be a special case in which only Δ_i is added to Rrs_i^{obsV} without any further adjustment. To avoid uncertainty that might result from outliers in observations, the median values of Rrs^{refA} and Rrs^{refV} instead of the ensemble mean were used to calculate the average radiometric shift $\bar{\Delta}$ in (7). Similar values were applied in (9) to calculate the factor g as well.

By adopting the Q-Q adjustment method, systematic bias ΔRrs^{SAT} at all common bands among sensors can be greatly mitigated. Because this adjustment method is based totally on observations, it can thus be utilized to synthesize spectral information for mismatched bands, such as the band at 531-nm wavelengths between MODIS and VIIRS. To achieve this goal, for instance, observations from two neighboring bands of VIIRS (i.e., 486 and 551 nm) can be employed as the baseline information for possible reconstruction.

3) *EMD*: Due to the highly dynamic nature of aquatic environments, time series of ocean color reflectance from different sensors at one particular pixel on the same date might not agree well with each other. The temporal differences among sensors could result in changes of the spectral characteristics in an aquatic environment because of bio-optical and biochemical processes, as well as water mass advection. The work in [26] revealed semidiurnal to diurnal fluctuations of concentrations of suspended sediment and phytoplankton in Tampa Bay, Florida, through coastal buoys and satellite observations. In other words, these time series will be too chaotic to build reliable relationships, even with advanced machine learning tools. Although these fluctuations do exist in those ocean color reflectance observations, to some extent, these high-frequency fluctuations might be referred to as white noise that could be removed to better characterize long-term relationships.

Toward the end to removing the white noise-like fluctuations, EMD was applied to reconstruct the projected time series in this study. Unlike the well-known Fourier transform, which requires linear and stationary data to avoid energy spreading in energy–frequency domain, EMD can decompose any nonlinear and nonstationary time series into a finite and often small number of intrinsic mode functions (IMFs) that admit well-behaved Hilbert transforms [27]. The decomposition is based on the direct extraction of the energy associated with various intrinsic time scales, and the IMFs are extracted level by level from the highest frequency local oscillations riding on the corresponding lower frequency part of the data until no complete oscillation can be extracted in the residual. In other words, the first IMF is one time series with the highest frequency oscillations included and thus can usually be considered the white noise, which must be removed. Therefore, the reconstructed time-series data

set can be produced by adding all remnant IMFs together, except the first one. The reconstructed time series instead of the projected time-series data set is used to characterize the location-dependent relationships among sensors with the aid of ELM.

4) *ELM*: After removing the systematic bias, location-dependent bias $\Delta Rrs^{LCT}(\lambda)$ must also be addressed. As previously described, $\Delta Rrs^{LCT}(\lambda)$ at one geographic grid could be different from the others; therefore, relationships should be analyzed individually for each particular pixel. Traditional methods such as linear regression could be the simplest approach; however, the significant dynamic nature of an aquatic environment prevents linear regression from building reliable relationships with sufficient accuracy. In this paper, a fast machine learning tool, i.e., ELM, was utilized to establish complex relationships among sensors that can be used to estimate the location-dependent bias at each geographical grid.

ELM was proposed to improve learning speed and accuracy via the single-hidden feedforward neural networks (SLFNs) [28]. Unlike traditional machine learning algorithms that require adjusting and tuning all parameters in SLFNs, the input weights and hidden layer biases in ELM can be randomly assigned if the activation is infinitely differentiable with the unchanged hidden layer output matrix [28]. The randomly chosen input weights and bias concepts enable ELM to directly respond to the inputs and rapidly find a solution without the iterative adjustment required in gradient descent-based algorithms. ELMs can provide universal generalization and classification performance, not only with the smallest training error but also with the smallest norm of weights at extreme learning speeds that can be thousands of times faster than traditional feedforward network learning algorithms, such as backward propagation algorithms in artificial neural network modeling analysis. Consequently, ELMs can be easily implemented in various applications. In this paper, ELM was utilized to explore the complex relationships between the systematic bias-corrected time series and associated observed Aqua time series. Once a reliable relationship is characterized, it can be used to correct the location-dependent bias.

III. RESULTS

A. Screening and Ranking of Baseline Sensors

Before the implementation of any cross-mission sensor observations merging scheme, baseline sensors must first be selected through screening and ranking steps. To achieve this, one sensor must be selected as the target, whereas the other sensors remain complementary, providing observations to be projected onto this target sensor level successively. In this paper, ocean color reflectance derived from three satellite sensors (i.e., MODIS-Terra, MODIS-Aqua, and VIIRS) were employed for possible data merging purposes. To determine the baseline sensor among these three sensors, two statistics addressing the three predetermined criteria in above were calculated among sensors as quantified indicators (i.e., average monthly coverage percentage and average monthly mean absolute bias; see Figs. 5 and 6, respectively). Average monthly coverage

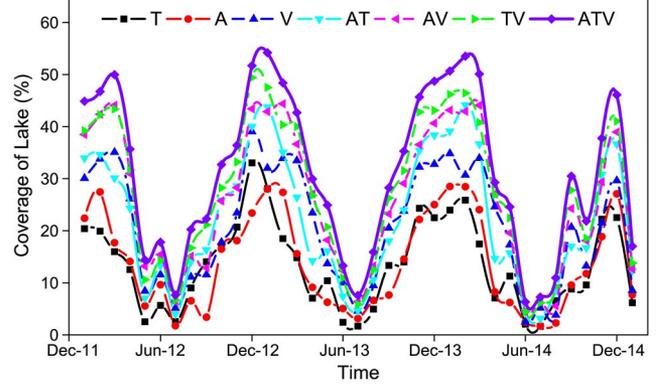


Fig. 5. Average monthly coverage (POC) during the 2012–2015 time period for the four possible combinations of these sensors. The combinations of sensors are identified by the letters associated with each individual sensor (AT: Aqua+Terra, AV: Aqua+VIIRS, and ATV: Aqua+Terra+VIIRS).

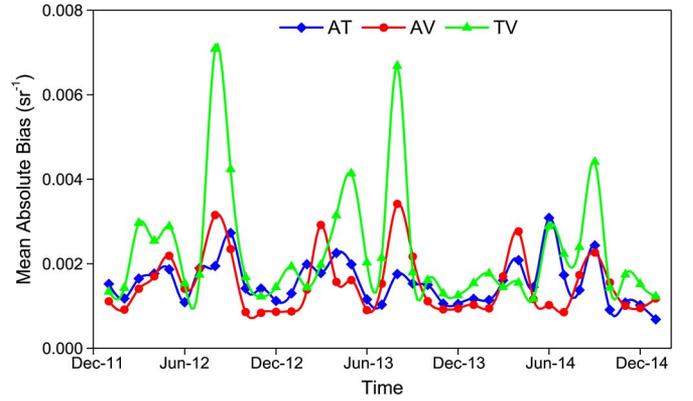


Fig. 6. Monthly mean absolute bias (MAB) of observed ocean color reflectance at 443 nm among sensors. Combinations of sensors are identified by the letter associated with each individual sensor: Terra (T), Aqua (A), and VIIRS (V).

percentage (i.e., \overline{POC}) was calculated as the ensemble mean of daily clear pixels coverage percentage (POC) over the lake each month, i.e.,

$$POC = 100 * \frac{N_{clear}}{N_{total}} \quad (13)$$

$$\overline{POC} = \frac{1}{T} \sum_{t=1}^T POC_t \quad (14)$$

where N_{clear} is the number of clear water pixels (i.e., having data value), and N_{total} is the total number of water pixels over the lake. Similarly, the average monthly mean absolute bias (\overline{MAB}) was calculated as the average of the mean absolute bias (MAB) between clear pixel matchups among sensors, i.e.,

$$MAB = 100 * \frac{1}{N} \sum_{i=1}^N |Rrs_i^{SAT1} - Rrs_i^{SAT2}| \quad (15)$$

$$\overline{MAB} = \frac{1}{T} \sum_{t=1}^T MAB_t. \quad (16)$$

The benefits of merged cross-mission sensor observations are intuitive (see Fig. 5). Compared with any single-sensor observations, the \overline{POC} of merged products over Lake Nicaragua

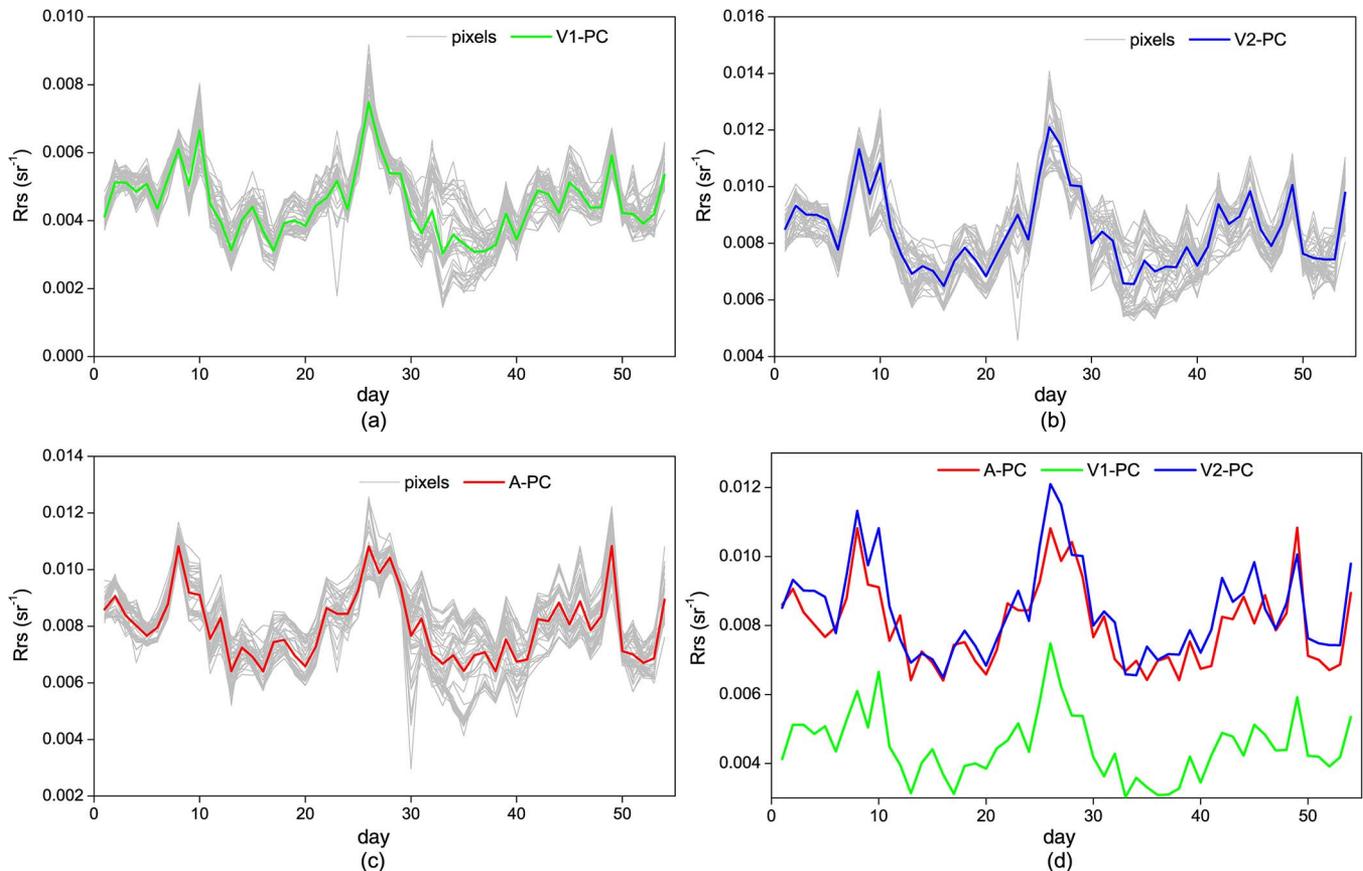


Fig. 7. First PC (PC1) and associated ocean color reflectance time series of MODIS-Aqua (531 nm) and VIIRS (488 and 551 nm). A-PC = PC1 of MODIS-Aqua; V1-PC = PC1 of VIIRS at 488 nm; V2-PC = PC1 of VIIRS at 551 nm. (a) VIIRS at 488 nm. (b) VIIRS at 551 nm. (c) Aqua at 531 nm. (d) PCs.

is nearly twice that of any single-sensor observations. Regarding the \overline{POC} of a single-sensor mission, VIIRS has a slightly larger \overline{POC} than that of MODIS, particularly during dry seasons (November–April). Meanwhile, the \overline{POC} of merged products between any two sensors indicates that merging VIIRS with any MODIS observations to gain a larger coverage ratio would be beneficial. This can be inferred from the larger \overline{POC} of AV and TV compared with that of AT. Within this context, VIIRS has priority than MODIS to be selected as the baseline sensor. However, comparisons of \overline{MAB} between any two sensors indicate that better agreement can be achieved by choosing MODIS-Aqua as the baseline sensor to merge with other two sensors. This selection can be deduced from smaller \overline{MAB} of AV and AT when comparing with that of TV. By taking MODIS-Aqua as the baseline sensor, the other two sensor observations are all needed to be projected onto the MODIS-Aqua level, which avoids addressing extreme biases between MODIS-Terra and VIIRS (e.g., large biases of TV; see Fig. 6). Therefore, considering the spectral (i.e., number of available bands) and overpassing time differences (see Table I), as well as the statistics of coverage ratio and relative biases between sensors, observations derived from MODIS-Aqua were selected as the baseline information for merging with observations from VIIRS and MODIS-Terra successively over Lake Nicaragua. In the merging scheme, VIIRS observations were first merged with those of MODIS-Aqua, and the merged products were

further merged with observations from MODIS-Terra. In addition, observations from MODIS-Aqua were reserved without any correction or computation. Only those grids having no information (i.e., no data value) in MODIS-Aqua (but with values in other sensors) were projected from other two sensors observations and then merged with those of MODIS-Aqua.

B. Systematic Bias Correction

In this paper, the Q–Q adjustment method was adopted to remove systematic bias among sensors. As described in Section II-B, historical time series of ocean color reflectance Rrs^{SAT} were utilized as references for systematic bias ΔRrs_i^{SAT} characterization. Here, all available observations of each sensor over Lake Nicaragua during the 2012–2015 overlapped time period were used to derive the relevant reference time series. To merge observations at MODIS 531-nm wavelengths, observations from two VIIRS bands at 486- and 551-nm wavelengths were used to generate the associated spectral information of 531 nm at the MODIS-Aqua level. Matchups between MODIS-Aqua and VIIRS bands were first extracted from both historical time series, respectively (denoted by pixels in Fig. 7). For this purpose, all historical images were ranked from high to low based on the number of matchups between both sensor observations for each date; as more observations recorded at different days were included, the number of matchups decreased. To

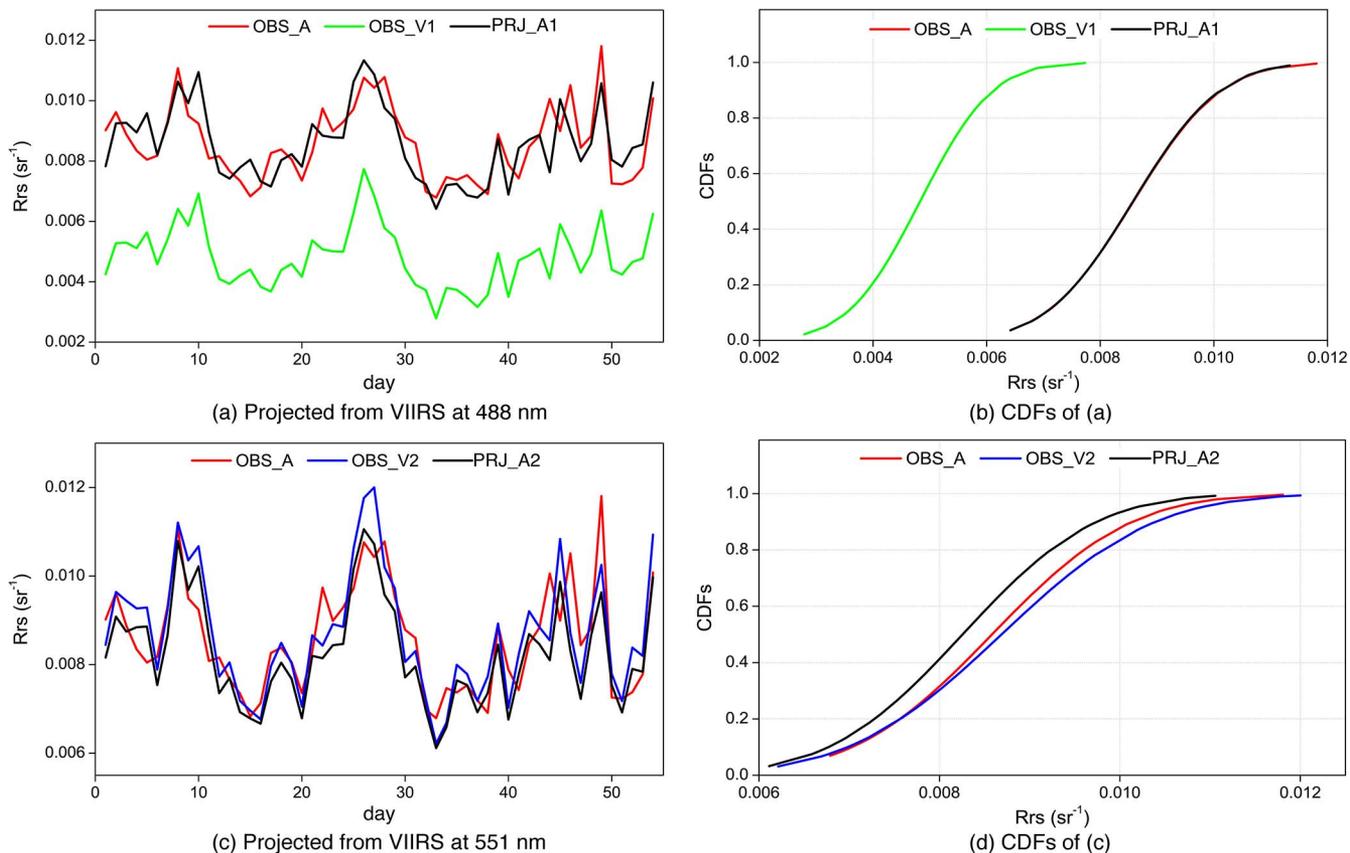


Fig. 8. Comparisons of ocean color reflectance before and after systematic bias correction. (a) Time series of observed $Rrs^{VIIRS(488)}$ (OBS-V1), $Rrs^{AQUA(531)}$ (OBS-A), and projected time series (PRJ_A1) from OBS-V1. (b) CDFs of each time series in (a). (c) Same as in (a) but from $Rrs^{VIIRS(551)}$ (OBS-V2). (d) CDFs of each time series in (c).

better characterize the systematic bias among sensors, sufficient matchups should be sampled. Here, one criterion was created as time was increased one by one until the number of matchups among sensors in the time series was $< 1\%$ of the maximum number of matchups (i.e., the number of pixels on the first date in the ranked time series).

To simplify the work and avoid unnecessary uncertainties, PCs of these observations were computed via singular value decomposition, and PCs were then utilized as reference time series for the Q-Q adjustment rather than using the chaotic raw time series. Statistics indicate that the PC1 of these observations can explain almost 95% of the total variance; therefore, PC1 of each sensor observation should be applied as the historical reference in the Q-Q adjustment for systematic bias characterization. Because the correction scheme is totally data dependent, large outliers in the reference time series would thus introduce uncertainties and new biases. To obtain a quality-assured correction, a quality control approach was applied. After calculating the PC1 of each observation (e.g., V1-PC, V2-PC, and A-PC in Fig. 7), the absolute bias between V-PC and A-PC was computed accordingly, i.e.,

$$\text{Bias}_i = 100 * \frac{|V-PC_i - A-PC_i|}{A-PC_i} \quad (17)$$

where i is the number of days in the extracted discrete time series. For quality control purposes, data points with values larger than the total average plus one standard deviation

(i.e., $> 95\%$ in cdf's of the PC bias) of the whole bias time series may be first screened out; associated data points in the PC time series in those corresponding days with large bias may then be removed.

Following the aforementioned screening steps and quality control assessment (see Fig. 7), only 54 days of observations were reserved for further systematic bias characterization, suggesting an urgent need for data merging over this region to improve the relevant spatial and temporal coverage of observations in turn. Comparisons of the derived PC of each sensor suggest that although the fluctuations of each sensor observation are different, the long-term variability (i.e., trends) is still similar to each other. This process is the foundation for conducting the cross-mission sensor bias correction.

Once the referenced time series (i.e., PC time series) data are created, the Q-Q adjustment method can be adopted for systematic bias correction among sensors. Following (5)–(12), observations collected by one sensor can be calibrated to the other based on the referenced time-series data. For instance, a comparison of $Rrs^{VIIRS(488)}$ and $Rrs^{VIIRS(551)}$ before and after being projected to $Rrs^{AQUA(531)}$ indicates that the systematic bias among sensors can be largely removed by adopting the Q-Q adjustment method (see Fig. 8). This can be evidenced by the fact that the cdf of projected time series (PRJ_A1) is almost overlapped with that of $Rrs^{AQUA(531)}$ after bias correction [see Fig. 8(b)]. In addition, it is indicative that the projected time-series data are not a simple transition

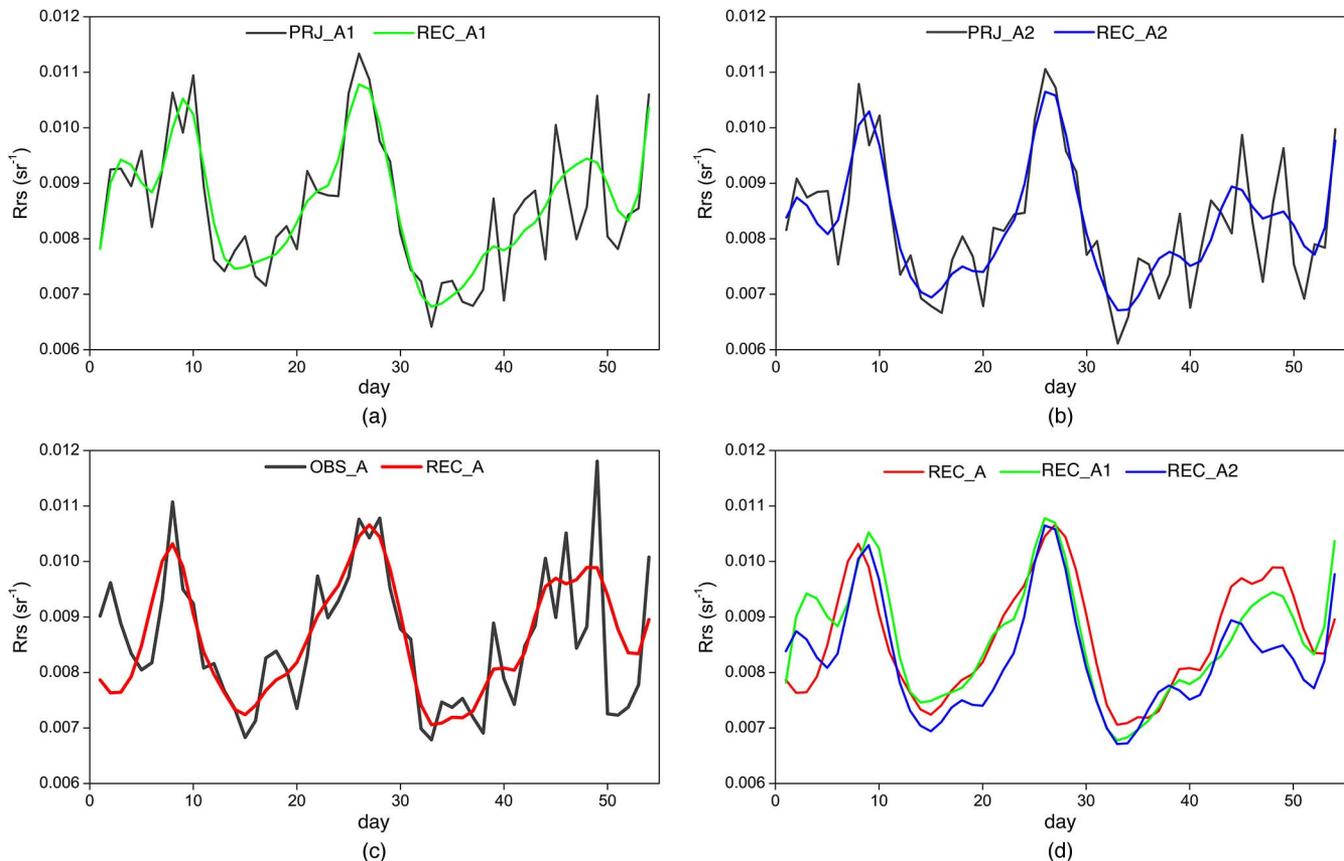


Fig. 9. Comparisons of the time series before and after reconstruction. Here, PRJ denotes projected time series in Fig. 8 (A1 is associated with VIIR 488 nm, and A2 is associated with VIIRS 551 nm), whereas REC denotes reconstructed time series. (a) Reconstructed from the projected time series I (PRJ_A1). (b) Reconstructed from the projected time series II (PRJ_A2). (c) Reconstructed from Aqua 531 nm (OBS_A). (d) Reconstructed time series.

from the original raw data by adding the average differences between two referenced time series [e.g., the differences are not consistent between time series before and after systematic bias correction in Fig. 8(c)]. During the Q–Q adjustment, minor corrections were made for observations with small values, whereas large corrections were made for large values at band 551 nm of VIIRS, reflecting a reasonable correction logic. This is due to the different instrumental responses or algorithmic differences in dealing with extreme values associated with different sensors. Large biases occur mainly for those peak values between VIIRS and MODIS-Aqua observations [see Fig. 8(c)]; therefore, it is an advantage of using such an adaptive method such as Q–Q adjustment for correcting systematic biases among cross-mission sensor observations.

C. Location-Dependent Bias Correction

Once the systematic bias can be removed from each particular pixel, relationships can be established between the projected time series and the associated target sensor time series for possible removal of the embedded location-dependent bias at this grid. Due to the highly dynamic nature of aquatic environments, the original reflectance time series are always nonlinear and nonstationary, with fluctuations randomly distributed (e.g., similar to white noise). As described in Section II-B, building a robust model for prediction with those chaotic time series is difficult, regardless of which method is chosen. With the aid

of EMD, these fluctuations are screened out in the first IMF, which has the highest frequency along with the lowest energy. Thus, removing these high-frequency signals may not affect the long-term variability of the original signals; rather, it would improve the stability and efficiency in model generalization. Comparisons of the time-series before and after reconstruction via EMD clearly show that the reconstructed time-series data are much smoother after removing those high-frequency fluctuations than the original projected time series; however, the reconstructed time series still maintain their original long-term variability (see Fig. 9).

With the reconstructed time series, ELM is used to establish relationships for removing the possible location-dependent bias ΔRrs^{LCT} . In this paper, reconstructed MODIS-Aqua time-series data were defined as targets, whereas reconstructed time series of the projected observations (from VIIRS or MODIS-Terra) time-series data were used as inputs for machine learning purposes. We randomly screened 70% of these reconstructed time-series data for ELM training, whereas the remnant was used for testing purposes (i.e., model validation). The stopping criterion for training was created based on the coefficient of determination (i.e., R^2) between the predicted value for the remaining 30% of inputs and the associated targets. The training process continues until the R^2 value reaches 0.8. Once a robust model is generated, it may then be used for prediction based on the given new observations. To avoid random simulation and improve robustness, multiple models were simulated.

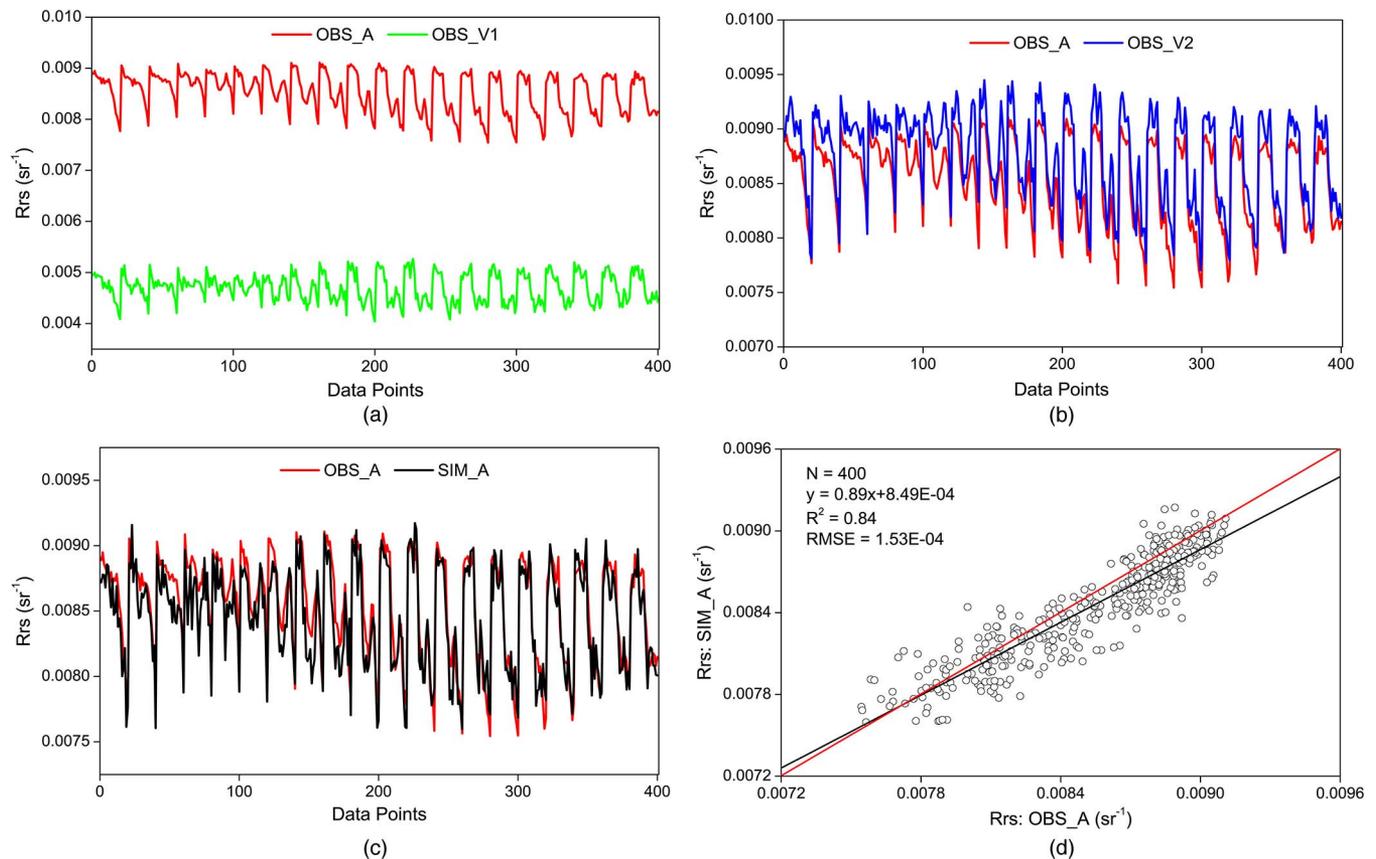


Fig. 10. Comparisons between the observed and simulated data for 400 pixels at 531 nm on January 10, 2014. OBS_V1 denotes the observations from VIIRS at 488 nm and OBS_V2 for 551 nm; OBS_A are the observations from MODIS-Aqua at 531 nm. (a) Aqua 531 nm vs. VIIRS 488 nm. (b) Aqua 531 nm vs. VIIRS 551 nm. (c) Aqua 531 nm vs. synthesized. (d) Scatter plot of (c).

In this paper, 30 trials were simulated for each particular pixel.

Normally, ELM performs at an extremely fast speed. In some extreme situations, however, such as cases for pixels with limited observations in the time series, ELM might fail to generalize a robust model due to data deficiency; consequently, no prediction can be generated. In this situation, a systematic bias-corrected data value may be assigned directly, which means no location-dependent bias correction is conducted for this geographical grid. In addition, after 30 trials, the mean of this predicted value is calculated to represent the final prediction performance at each geographical grid.

D. Experimental Results

1) *Spectral Information Synthesis*: To verify the efficacy of SIASS, experimental analyses were conducted for spectral information synthesis (such as VIIRS 488 and 551 nm) at mismatched bands, as well as spectral adjustment at the common bands. In this paper, the validation scheme was performed by applying SIASS to reconstruct 400 observed MODIS-Aqua ocean color reflectance in one image from associated VIIRS and MODIS-Terra observations on January 10, 2014. These 400 data points (i.e., pixel values) were extracted from a 20×20 clear scene in each sensor observation. Spectral information of two adjacent projected VIIRS bands with wavelengths of

488 and 551 nm were synthesized to generate observations at MODIS-Aqua 531-nm wavelengths (see Fig. 10). Before applying SIASS, distinctive biases were found between VIIRS and MODIS-Aqua observations. MODIS-Aqua at 531 nm overestimated by almost twice that of VIIRS at 488 nm, whereas the VIIRS at 551 nm slightly overestimated that of MODIS-Aqua at 531 nm. Inconsistent biases (e.g., small bias for minimum; large bias for maximum) between observations from MODIS-Aqua 531 nm and VIIRS 551 nm also indicate the complexity of bias correction among sensors, arising mainly from the nonlinear and nonstationary nature of biases among sensors. These inconsistencies might result from instrumental responses, algorithmic differences, and calibration uncertainties, as well as aquatic dynamics due to local overpassing time differences of satellite sensors. However, the synthesized time-series data from VIIRS 488 and 551 nm largely avoid relevant issues as mentioned earlier.

Fair agreement can be observed between sensor measurements and reconstructed spectral information at 531 nm [see Fig. 10(c)], which strongly supports the efficacy of the SIASS method and indicates that the SIASS is capable of removing the prominent nonlinear and nonstationary biases. The synthesized spectral information from neighboring bands preserve the unique spectral characteristics of these mismatched bands, which are valuable for better monitoring and understanding the changing aquatic environment. This advantage distinguishes

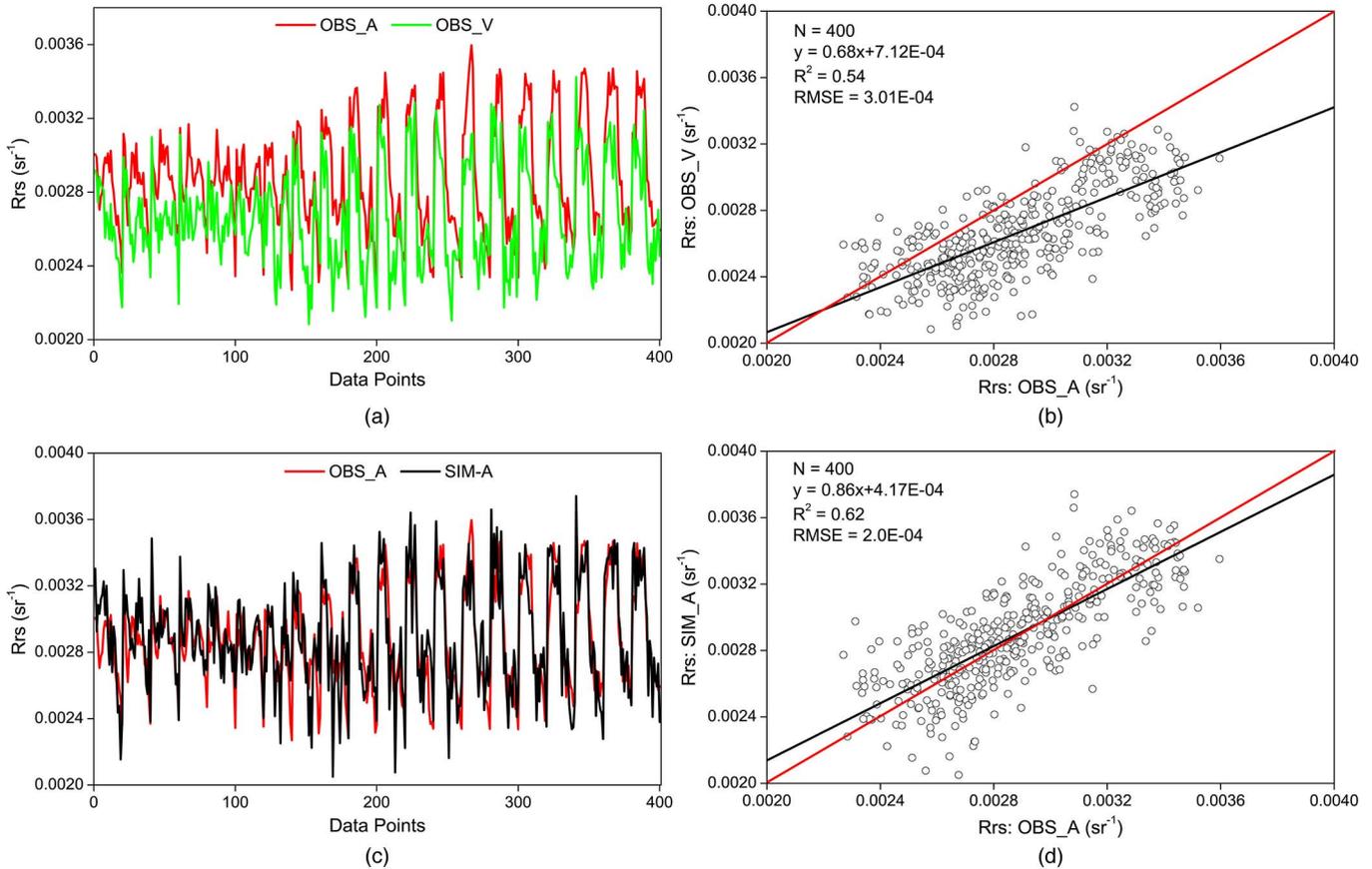


Fig. 11. Comparisons between the observed and simulated data for 400 pixels at 443 nm on January 10, 2014. Observations from VIIRS at 443 nm were used to simulate the associated spectral information at MODIS-Aqua 443 nm. (a) Aqua 443 nm vs. VIIRS 443 nm. (b) Scatter plot of (a). (c) Aqua 443 nm vs. simulated. (d) Scatter plot of (c).

SIASS from other similar bias correction schemes that may lose some valuable spectral information because they only work for the common bands among sensors. Prior to the development of SIASS, this weakness could largely be avoided by using complex bio-optical models or band-shifting processes [10], [11], [15], [16].

2) *Spectral Information Adaptation*: For the common bands among sensors, observations are calibrated from one sensor to the other through similar processes such as reconstructing spectral information for those wavelengths in mismatched bands; this process is thus termed as spectral information adaptation. The only difference between adaptation and synthesis schemes is that only one-band observation instead of two is used as input. Comparisons between the observed and projected spectral information from different sensor observations at different wavelengths (see Figs. 11–13) suggest that the consistency between sensors is significantly improved after correction, with the largest effects at 443 nm.

Comparisons of observations from different sensors at the same wavelength (e.g., 443 nm) indicate that the biases resulting from the temporal differences might be most significant when compared with other biases resulting from instrumental and algorithmic differences. In addition to the uncertainties in absolute radiometric calibration processes between sensors (see Table I), the only difference between MODIS-Terra and

MODIS-Aqua is nearly 3 h local overpassing time differences; other factors (e.g., instrument design, wavelengths, and algorithm) are similar. Nevertheless, the RMSE between MODIS-Aqua and Terra [see Fig. 13(b)] is much larger than that between MODIS-Aqua and VIIRS (instrumental, algorithmic, and small temporal differences existed). The possible reason for this might result from the highly dynamic nature of aquatic environments, which is sensitive to temporal differences with significant short-term temporal variability [26].

Similar effects can also be observed at 531 nm. Synthesized data from two VIIRS observations (see Fig. 10) agree even more closely than those of projections based directly on MODIS-Terra observations (see Fig. 12). However, because of inherent synthesis variability, bias-corrected data still follow the original observation variability. This effect was also evident in comparisons of R^2 before and after calibration. Values of R^2 were not greatly improved even after calibration, and in some cases, they were even reduced. Although R^2 could not provide meaningful insights, to some extent, the results suggest that the corrected data are still dependent on the original observations, particularly the variability.

To confirm the consistent improvement of ocean color reflectance before and after the spectral information adaptation with SIASS, the mean relative difference (MRD) was calculated pairwise for ocean color reflectance observations between

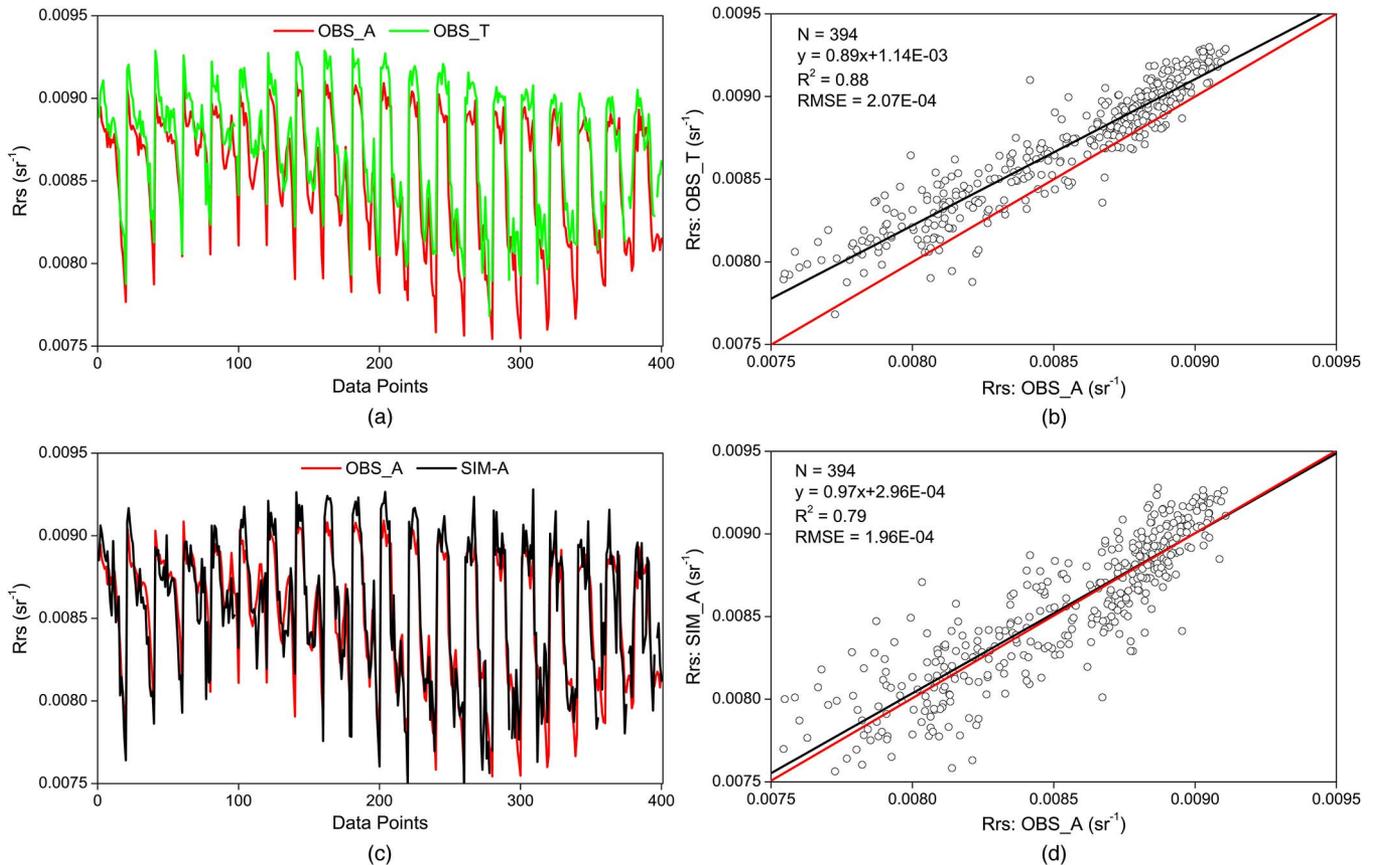


Fig. 12. Comparisons between the observed and simulated data for 400 pixels at 531 nm on January 10, 2014. Observations from VIIRS at 443 nm were used to simulate the associated spectral information at MODIS-Aqua 531 nm. (a) Aqua 531 nm vs. Terra 531 nm. (b) Scatter plot of (a). (c) Aqua 531 nm vs. simulated. (d) Scatter plot of (c).

MODIS-Aqua and either of the two other satellites, as well as the simulated data, i.e.,

$$\text{MRD} = 100 \times \frac{1}{N} \sum_{i=1}^N \frac{\text{Rrs}_i^{\text{SAT1}} - \text{Rrs}_i^{\text{SAT2}}}{\text{Rrs}_i^{\text{SAT2}}} \quad (18)$$

where $\text{Rrs}_i^{\text{SAT2}}$ denotes the ocean color reflectance observations from MODIS-Aqua; $\text{Rrs}_i^{\text{SAT1}}$ denotes observations from MODIS-Terra or VIIRS or simulated ocean color reflectance; and N is the number of samples in each satellite observation. The associated MRD values derived from the results (see Figs. 10–13) were comparatively summarized (see Table II).

Before applying SIASS, apparent discrepancies (or inconsistencies) were observed between MODIS-Aqua and the other two satellites, particularly at 443 nm (see Table II). VIIRS significantly underestimated ocean color reflectance observations, whereas MODIS-Terra significantly overestimated these values relative to the MODIS-Aqua. Therefore, correcting these large cross-mission biases is essential before merging observations associated with these three platforms. The largest correction effect was confirmed at 443 nm, with an $\text{MRD} < 1\%$ after applying SIASS.

3) *Application of Mapping Chlor-*a* Concentrations*: A real-world application of SIASS was performed to map the chlor-*a* concentrations in Lake Nicaragua using NASA’s operational

chlor-*a* algorithm of MODIS (i.e., OC3M), which is an empirical fourth-order polynomial relationship derived from *in situ* measurements of chlor-*a* and blue-to-green band ratios of ocean color reflectance [29], [30], i.e.,

$$C_{\text{chlor-}a} = 10^{(a_0 + \sum_{i=1}^4 a_i * R^i)} \quad (19)$$

$$R = \log_{10} \left(\frac{\text{Rrs}_{443} > \text{Rrs}_{488}}{\text{Rrs}_{551}} \right) \quad (20)$$

where the sensor-unique coefficients of a_0 – a_4 are 0.2424, -2.7423 , 1.8017, 0.0015, and -1.2280 , respectively. The numerator in (20) specifies retrieval of the greatest value (i.e., maximum) of the band ratios between Rrs_{443} and Rrs_{488} .

To compute chlor-*a* concentrations, observations of Rrs_{443} , Rrs_{488} , and Rrs_{551} over Lake Nicaragua were retrieved from MODIS-Aqua, MODIS-Terra, and VIIRS, respectively, on December 5, 2013. By applying SIASS, merged ocean color reflectance products at these three distinctive wavelengths were generated (see Figs. 14–16, respectively). These maps show that the spatial coverage ratio of MODIS-Aqua observations over Lake Nicaragua was significantly improved by merging with VIIRS and MODIS-Terra successively. Before merging, MODIS-Aqua had a clear coverage ratio of 35.28% of the lake. The coverage percentage was improved to 67.97% after merging with VIIRS, nearly twice that before merging, and increased to 83.63% after coalescing with observations from

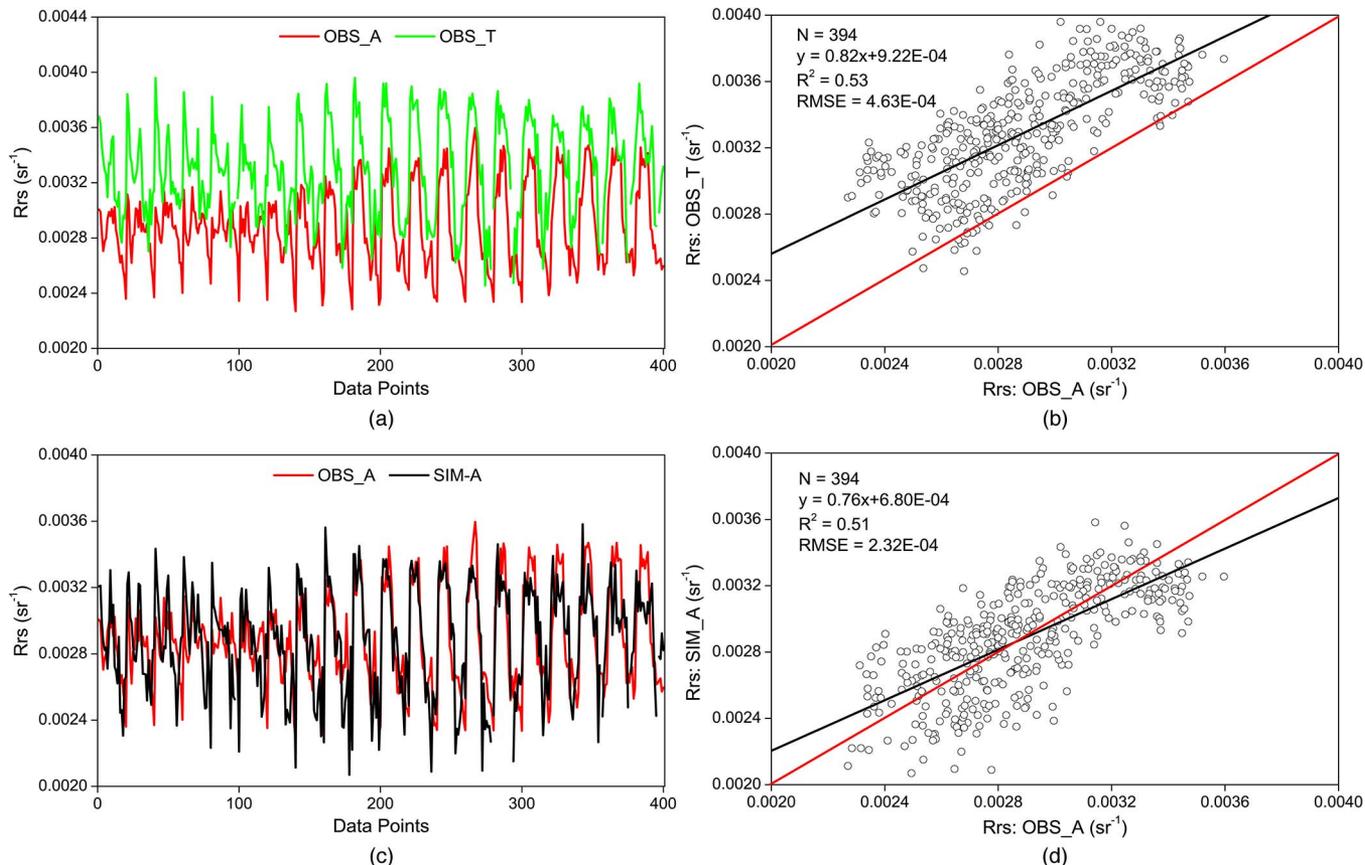


Fig. 13. Comparisons between the observed and simulated data for 400 pixels at 443 nm on January 10, 2014. Observations from MODIS-Terra at 443 nm were used to simulate the associated spectral information at MODIS-Aqua 443 nm. (a) Aqua 443 nm vs. Terra 443 nm. (b) Scatter plot of (a). (c) Aqua 443 nm vs. simulated. (d) Scatter plot of (c).

TABLE II
COMPARISONS OF MRD BETWEEN PAIRWISE OCEAN COLOR
REFLECTANCE FROM THE CROSS-MISSION OBSERVATIONS
BEFORE AND AFTER SPECTRAL ADAPTATION WITH SIASS

| Sensor Pair | AV | | AT | |
|----------------|---------------|--------------|---------------|--------------|
| | Before (%) | After (%) | Before (%) | After (%) |
| 531 nm | 44.96/2.7* | 0.98 | 1.88 | 0.23 |
| 443 nm | -7.31 | 0.53 | 14.21 | 0.02 |

* two bands of VIIRS (488 nm/551 nm) observations (Fig. 10).

MODIS-Terra finally. Some obvious uncertainties remain in the merged products, particularly after merging with observations from MODIS-Terra at blue band (443 nm). This effect might result from the quality-unassured observations (i.e., we did not filter the satellite observations with flags) and aquatic dynamics due to large temporal differences, as well as biases arising from ELM model simulations in characterizing the possible relationships to remove the corresponding location-dependent bias.

Following the empirical relationships in (19) and (20), chlor-*a* concentrations were derived and mapped (see Fig. 17), showing that high concentrations of chlor-*a* were mainly observed in the mid-west coast of the lake, whereas low concentrations were observed in the central lake. Over some geographic grids, the chlor-*a* concentrations were slightly lower than those at

adjacent grids, mainly due to overestimation of observations at the two blue bands in the correction scheme, particularly at 443 nm, where two distinctive patches with large values were observed in the southern areas of the lake. In addition, large outliers were observed over the northeast of the lake, which is classified as a large extreme value present in regions based on observations from MODIS-Terra. These outliers might result not only from the quality-unassured ocean color reflectance observations but also from the large temporal differences associated with aquatic dynamics, both of which possibly cause failure of chlor-*a* concentration retrieval. Despite these possible uncertainties, the merged products significantly improve the spatial and temporal coverage by combining observations from available cross-mission sensors. This merging scheme allows us to better monitor and understand the dynamics of aquatic environment by deriving biophysical parameters that can be utilized as water quality indicators.

IV. DISCUSSION

Discrepancies between cross-mission ocean color radiometric products can be largely explained by the differences resulting from instrumental to algorithmic and temporal aspects, such as sensor design, calibration, atmospheric correction, data processing, and short-term aquatic variability due to different local overpassing time of sensors. Significant biases arising

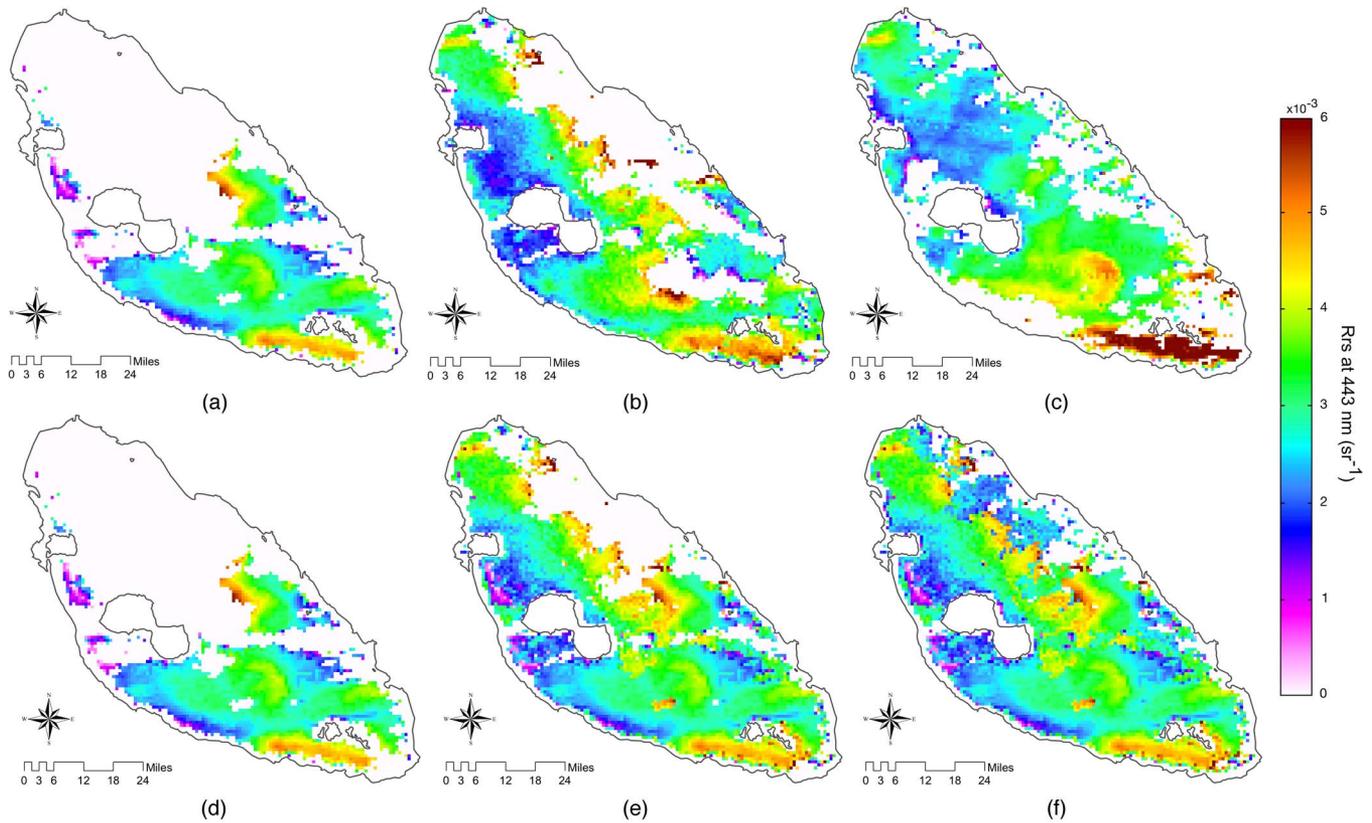


Fig. 14. Comparisons of MODIS-Aqua ocean color reflectance before and after merging with VIIRS and MODIS-Terra at 443 nm on December 5, 2013. The original observations of each sensor are shown in the upper panel, whereas the merged observations are shown in the bottom panel. (a) MODIS-Aqua. (b) VIIRS. (c) MODIS-Terra. (d) MODIS-Aqua. (e) MODIS-Aqua fused with VIIRS. (f) MODIS-Aqua fused with VIIRS and MODIS-Terra.

from these differences should be eliminated before merging these cross-mission sensor observations. Differing from methods relying on complex bio-optical models and scarce *in situ* measurements to eliminate these discrepancies, SIASS is proposed in this study to manage these biases for possible merging purposes.

Two kinds of biases, sensor-dependent systematic bias and location-dependent bias, were managed between cross-mission sensor observations. The sensor-dependent systematic bias, abbreviated as systematic bias, mainly results from the instrumental and algorithmic differences among sensors. An important assumption here is that the variability of this bias remains unchanged or has few changes over time. This bias can therefore be characterized based on the common observations recorded at the overlapping time window among sensors. SIASS uses this advantage to remove the systematic bias among sensors using an adaptive statistical adjustment method (i.e., Q–Q adjustment method). Results indicate that this adjustment can mitigate the systematic bias between radiometric observations from cross-mission sensors significantly. Nevertheless, some concerns must be highlighted. As aforementioned, we assume that the systematic bias remains unchanged or has few changes along the time horizon; however, sensor degradation (i.e., sensitivity drifting) could introduce new uncertainties to the adjustment process. Because the adjustment is highly dependent on the cdf's of sensor observations, different weights would be assigned to the same pixel for adjustment as the degraded

observations might alter the distribution of cdf's. This adaptive adjustment scheme is also dependent on the amount of common observations. The more sensor common observations, the better the accuracy. With a limited number of observations, the calculated cdf's of observations could have large intervals, which, in turn, might result in large uncertainties when quantifying those factors for bias correction in (5)–(12). Therefore, to achieve a quality-assured calibration, sufficient common observations should be guaranteed. In addition, biases are mainly characterized through distribution mapping between the observed and two reference time series. The Q–Q adjustment method requires that these time series have the same number of samples to guarantee fair intercomparisons. In our study case, the observed time series has one more data point (i.e., the one to be projected) than the two referenced time-series data. Although this does not seem to change the cdf's of the observed time series apparently, some uncertainties might be introduced. Within this context, the method should be slightly modified to avoid this issue in the future.

Unlike other linear or nonlinear models used for band-shifting purposes, spectral information at mismatched wavelengths in one sensor can be reconstructed based on observations from two neighboring wavelength observations in the other sensor through the Q–Q adjustment method. This spectral information synthesis scheme not only works for the band-shifting purposes for the same sensor but is also capable of generating associated spectral information from different sensors. Increased spectral

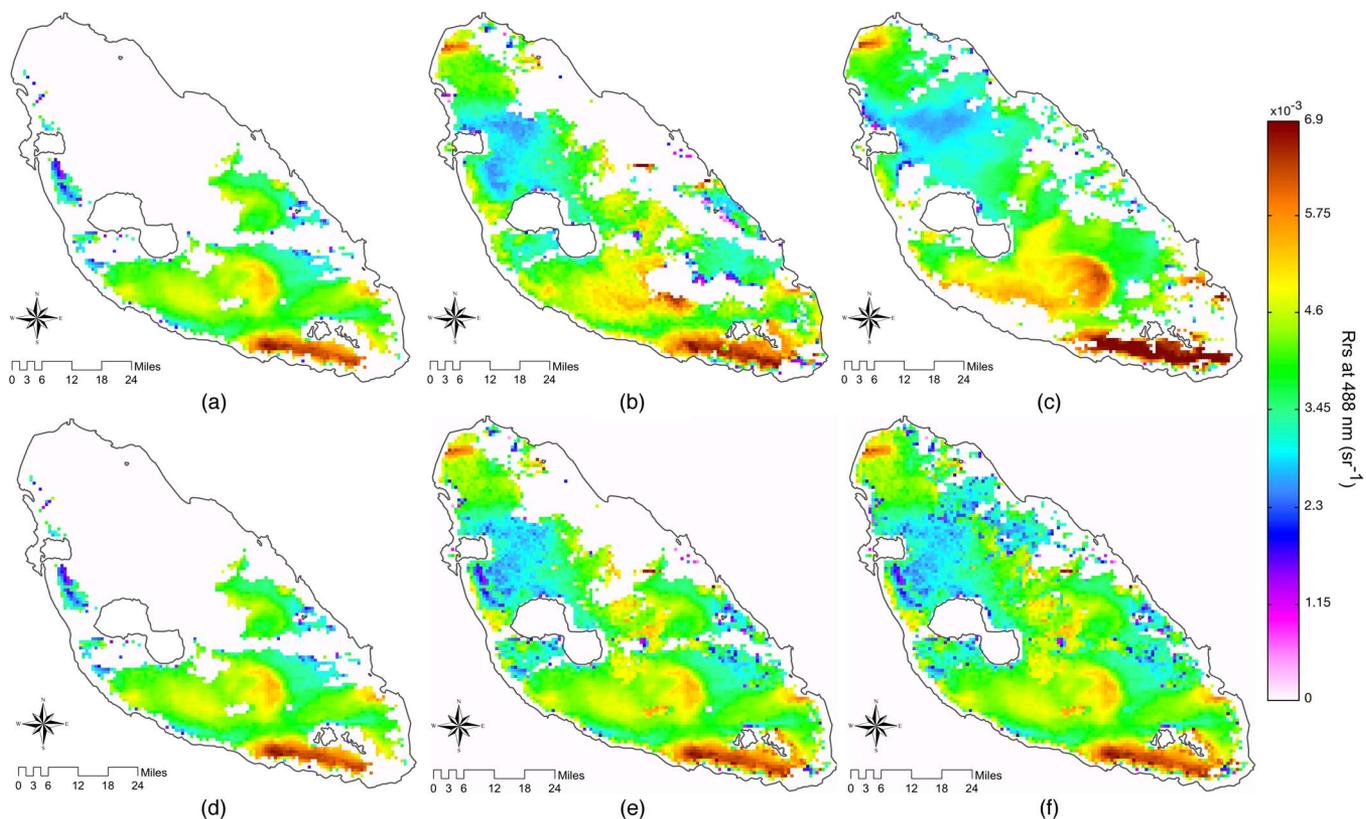


Fig. 15. Comparisons of MODIS-Aqua ocean color reflectance before and after merging with VIIRS and MODIS-Terra at 488 nm on December 5, 2013. The original observations of each sensor are shown in the upper panel, whereas the merged observations are shown in the bottom panel. (a) MODIS-Aqua. (b) VIIRS. (c) MODIS-Terra. (d) MODIS-Aqua. (e) MODIS-Aqua fused with VIIRS. (f) MODIS-Aqua fused with VIIRS and MODIS-Terra.

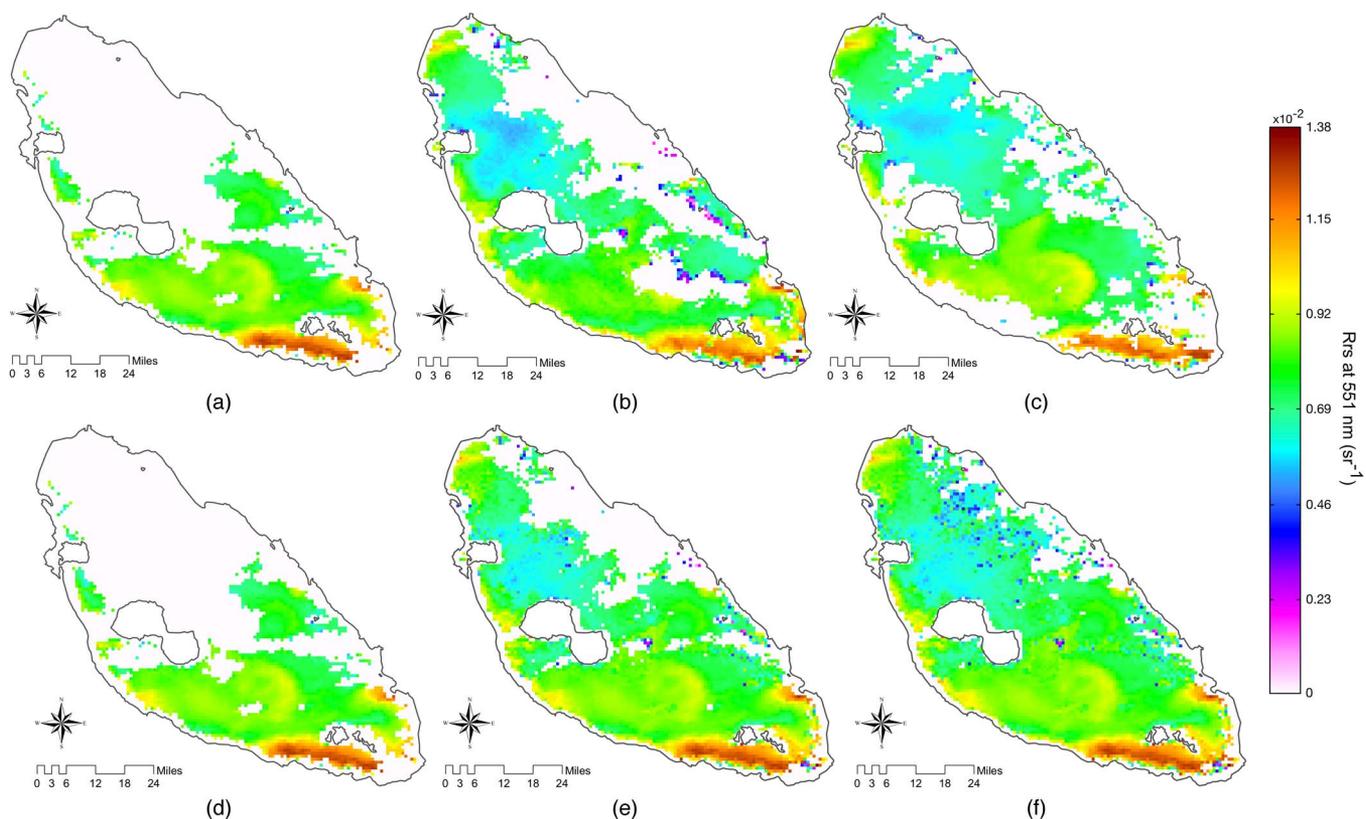


Fig. 16. Comparisons of MODIS-Aqua ocean color reflectance before and after merging with VIIRS and MODIS-Terra at 551 nm on December 5, 2013. The original observations of each sensor are shown in the upper panel, whereas the merged observations are shown in the bottom panel. (a) MODIS-Aqua. (b) VIIRS. (c) MODIS-Terra. (d) MODIS-Aqua. (e) MODIS-Aqua fused with VIIRS. (f) MODIS-Aqua fused with VIIRS and MODIS-Terra.

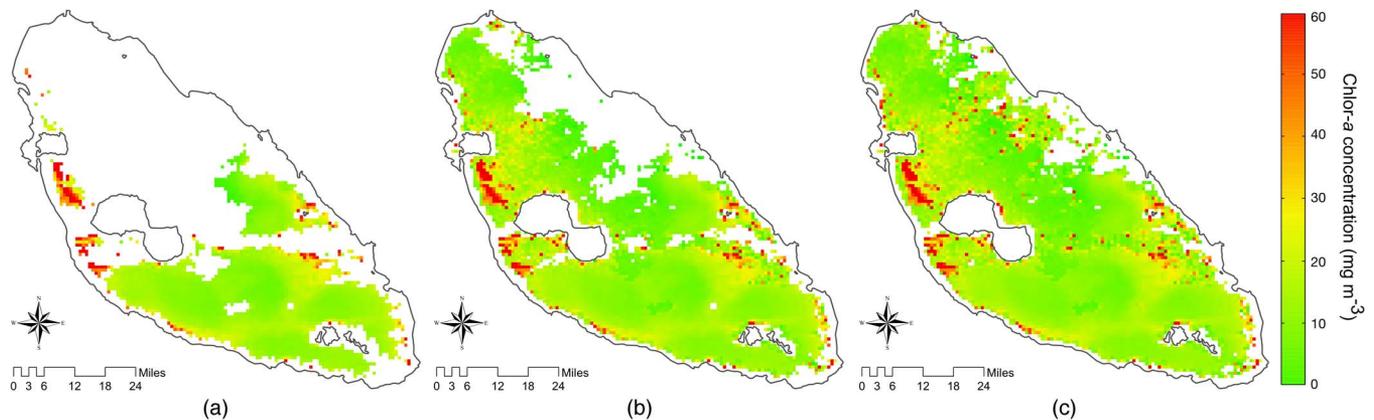


Fig. 17. Chlorophyll-*a* (chlor-*a*) concentrations derived from ocean color reflectance at 443, 488, and 551 nm on December 5, 2013 based on the NASA's operational chlor-*a* algorithm OC3M. (a) MODIS-Aqua. (b) MODIS-Aqua fused with VIIRS. (c) MODIS-Aqua fused with VIIRS and MODIS-Terra.

information may provide valuable data sources for subsequent data mining purposes, which can help us better monitor parameters related to the quality and status of aquatic environment, such as chlor-*a* concentration.

The location-dependent bias could be affected by the hydrodynamics in the highly complex aquatic environment due to the advection of water masses and the local overpassing time of the satellite sensors. Similarly, we assume that these aquatic dynamics have little variability in each particular pixel of a geographical grid, which means that dynamics varying with time can differ for different grids but should remain unchanged or have few changes for the same grid. Statistical relationships are estimated between sensor observations and are then employed to predict the bias-corrected value based on given new observations in this bias-correction scheme. Due to the dynamic nature of aquatic environments, time series of radiometric observations with chaotic fluctuations such as white noise are always present, making it difficult to generalize a robust relationship with fair accuracy, even with advanced machine learning tools. Within this context, a signal processing method, i.e., EMD, was applied here to remove the white noise-like fluctuations embedded in the time-series data for building relationships. This method, to some extent, improves the performance by quickly generalizing a robust model for location-dependent bias corrections.

Finally, ELM, a machine learning tool with an extremely fast speed in learning processes, was implemented to establish possible relationships that account for the location-dependent bias. For most cases, ELM performed well in generating a robust model to characterize the possible location-dependent bias among sensors, but it might fail in some cases, particularly when few matchups were available for training purposes. Consequently, for this situation, the established ELM model cannot be considered as a robust one for data prediction purposes. In some cases, the predicted value deviates far from the anticipated one if the given input falls out of the normal range of data utilized for model training, evidenced by experiencing a negative value or a value 10 times larger than the normal value. In this paper, quality control was applied to avoid such large outliers. Pixels with a predicted value > 1.5 times or < 0.5 times the given inputs must be masked as no

predicted value available at this grid and given a flag of *NaN* (i.e., Not a Number). Moreover, due to stochastic weights and bias assignment schemes in ELMs, different solutions can be achieved for the same problem. To some extent, these possible drawbacks reduce overall accuracy, which can be seen from the scattered matchups between the observed and corrected data (see Figs. 11–13). This issue also limits the applications of this ELM method to deal with some extreme cases (e.g., pixels with few observations in the historical time-series data). Meanwhile, multiple trials can be conducted by using the ensemble means of these robust predictions to represent the overall prediction accuracy. In addition, other possible methods, such as the second highest value [31] and the modified maximum average [32], also can be utilized to optimize the predicted multiple results for possible accuracy improvements.

Considering all possible causes of spectral discrepancies at similar wavelengths between sensors, the temporal differences might be more salient than others. This finding can be confirmed by the direct comparison among results given the fact that the associated biases between the MODIS-Aqua and Terra are even larger than those between MODIS-Aqua and VIIRS, since MODIS instruments aboard Terra and Aqua have the same design and data processing algorithms, and thus, the biases between these two MODIS sensor observations should ideally be smaller than those between MODIS-Aqua and VIIRS, which have different instrumental design and data processing algorithms. On the contrary, biases between MODIS-Aqua and VIIRS are even smaller than those between MODIS-Aqua and MODIS-Terra. Hence, this effect might be attributed to the temporal differences in local overpassing time among sensors, which is larger between MODIS-Aqua and Terra than that between Aqua and VIIRS. In addition to the temporal differences, discrepancies could result from uncertainties during calibration and retrieval processes as well. Aquatic environments, however, are highly dynamic compared with terrestrial environments. Inner dynamic factors, such as water mass advection as well as biochemical and biophysical processes, combined with external forcing factors such as wind, could synergistically alter the variability of an aquatic environment and affect the quality of remote sensing products. The challenge of remote sensing in an aquatic environment is not just limited to the temporal

differences that cause dynamics and a weak signal-to-noise ratio; it is also due to the chaotic characteristics in an aquatic environment, which make it difficult to statistically predict and handle these differences among sensors.

As SIASS mainly employs common observations among sensors to characterize and remove the systematic and location-dependent biases, it only addresses the long-term common biases between sensors. Consequently, uncertainties and biases might remain in the bias-corrected products, particularly for days when observations are sampled during highly dynamic periods of the aquatic environment. In these situations, SIASS could not eliminate all the biases based on the historical memories, and some biases are thus preserved in the simulated products.

Although the spatial and temporal coverage of merged products derived from biases-corrected observations via SIASS can be improved significantly, it is still difficult to achieve a full clear coverage of the study area in some cases. To restore that missed spectral information, information reconstruction methods such as SMart Information Reconstruction (SMIR) [17], Neighborhood Similar Pixel Interpolator (NSPI) [33], Geostatistical Neighborhood Similar Pixel Interpolator (GNSPI) [34], and Weighted Linear Regression (WLR) integrated with a regularization method [35] can be applied to further recover the missing information toward a full clear coverage of study regions.

In addition, back to the original data source used in this study, due to the limited valid satellite ocean color reflectance observations over the study area (resulting from severe clouds contamination), no flag filtering was applied to the original observations before further computation. Thus, uncertainties might exist in some of these observations. Although those uncertainties do not seem to affect our study objectives, biases might be introduced into the simulated ocean color reflectance products due to error propagation. We therefore advise excluding those quality-unassured observations through flag filtering in real-world operational applications.

V. CONCLUSION

By taking advantage of temporal overlaps between successive generations of satellite sensors, an adaptive statistical method, i.e., SIASS, was proposed in this study to remove the systematic and location-dependent biases between cross-mission ocean color sensors for observation merging purposes. With the aid of the Q-Q adjustment method and a machine-learning-based correction scheme, SIASS was able to remove the instrumental- and algorithmic-related systematic biases, as well as location-dependent bias. Compared with previous methods using complex bio-optical models and scarce *in situ* measurements for possible bias correction, SIASS is more adaptive as it relies on common observations from cross-mission sensors solely. This advantage makes the SIASS technique transferrable and applicable to include any other satellite sensor observations with similar features in various occasions.

Differing from previous merging schemes that only consider observations from the common band observations between

cross-mission sensors, SIASS can also synthesize spectral information for those wavelengths in mismatched bands at one sensor based on observations in two neighboring bands collected from the other sensor. This spectral information synthesis scheme helps preserve more spectral characteristics, allowing us to better monitor and understand the changing aquatic environment. Increased spectral information provides essential data sources for data mining purposes, which can be used to derive some distinctive indicators that are provided in the operational products.

Experimental results in this study suggest that SIASS is able to remove biases among cross-mission sensors significantly with accuracy. Final merged data products have better spatial and temporal coverage than any individual sensor. This advantage resolves the drawback of data scarcity in some regions such as the tropics, where dense cloud cover is frequent and hinders the successive monitoring of water quality in inland waters and coastal regions.

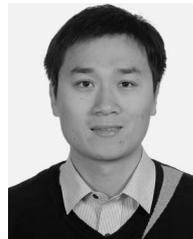
ACKNOWLEDGMENT

The authors would like to thank the three anonymous reviewers for their comments and constructive suggestions to improve this paper. They would also like to thank the NASA Ocean Color Science Team for providing the MODIS and VIIRS ocean color products, as well as the SeaDAS software for processing these data, and Nanyang Technological University for providing the ELM codes.

REFERENCES

- [1] 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, Jan. 1994.
- [2] M. Wang and S. W. Bailey, "Correction of sun glint contamination on the SeaWiFS ocean and atmosphere products," *Appl. Opt.*, vol. 40, no. 27, pp. 4790–4798, Sep. 2001.
- [3] K. D. Moore, K. J. Voss, and H. R. Gordon, "Spectral reflectance of whitecaps: Their contribution to water-leaving radiance," *J. Geophys. Res.*, vol. 105, no. C3, pp. 6493–6499, Mar. 2000.
- [4] W. W. Gregg, "Coverage opportunities for global ocean color in a multimission era," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 5, pp. 1620–1627, Sep. 1998.
- [5] W. W. Gregg and R. H. Woodward, "Improvements in coverage frequency of ocean color: Combining data from SeaWiFS and MODIS," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 4, pp. 1350–1353, 1998.
- [6] 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.
- [7] 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.
- [8] D. D'Alimonte, G. Zibordi, and F. Mélin, "A statistical method for generating cross-mission consistent normalized water-leaving radiances," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 12, pp. 4075–4093, Dec. 2008.
- [9] F. Mélin and G. Zibordi, "Optically based technique for producing merged spectra of water-leaving radiances from ocean color remote sensing," *Appl. Opt.*, vol. 46, no. 18, pp. 3856–3869, 2007.
- [10] S. Maritorena and D. A. Siegel, "Consistent merging of satellite ocean color data sets using a bio-optical model," *Remote Sens. Environ.*, vol. 94, no. 4, pp. 429–440, Feb. 2005.
- [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, 2010.

- [12] E. J. Kwiatkowska and G. S. Fargion, "Application of machine-learning techniques toward the creation of a consistent and calibrated global chlorophyll concentration baseline dataset using remotely sensed ocean color data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 12, pp. 2844–2860, Dec. 2003.
- [13] W. W. Gregg and N. W. Casey, "Improving the consistency of ocean color data: A step toward climate data records," *Geophys. Res. Lett.*, vol. 37, no. 4, 2010, Art. ID. L04605.
- [14] W. W. Gregg, N. W. Casey, J. E. O'Reilly, and W. E. Esaias, "An empirical approach to ocean color data: Reducing bias and the need for post-launch radiometric re-calibration," *Remote Sens. Environ.*, vol. 113, no. 8, pp. 1598–1612, Aug. 2009.
- [15] Z. Lee, S. Shang, C. Hu, and G. Zibordi, "Spectral interdependence of remote-sensing reflectance and its implications on the design of ocean color satellite sensors," *Appl. Opt.*, vol. 53, no. 15, pp. 3301–3310, May 2014.
- [16] F. Mélin and G. Sclep, "Band shifting for ocean color multi-spectral reflectance data," *Opt. Exp.*, vol. 23, no. 3, p. 2262–2279, Feb. 2015.
- [17] N.-B. Chang, K. Bai, and C.-F. Chen, "Smart information reconstruction via time-space-spectrum continuum for cloud removal in satellite images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 5, pp. 1898–1912, May 2015.
- [18] M. Wang *et al.*, "Impacts of VIIRS SDR performance on ocean color products," *J. Geophys. Res. Atmos.*, vol. 118, no. 18, pp. 10 347–10 360, 2013.
- [19] 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.
- [20] H. R. Gordon and D. K. Clark, "Clear water radiances for atmospheric correction of coastal zone color scanner imagery," *Appl. Opt.*, vol. 20, no. 24, pp. 4175–4180, Dec. 1981.
- [21] A. Morel, D. Antoine, and B. Gentili, "Bidirectional reflectance of oceanic waters: Accounting for Raman emission and varying particle scattering phase function," *Appl. Opt.*, vol. 41, no. 30, pp. 6289–6306, Oct. 2002.
- [22] S. L. Castro, G. A. Wick, D. L. Jackson, and W. J. Emery, "Error characterization of infrared and microwave satellite sea surface temperature products for merging and analysis," *J. Geophys. Res. Ocean.*, vol. 113, 2008, Art. ID. C03010.
- [23] A. Amengual, V. Homar, R. Romero, S. Alonso, and C. Ramis, "A statistical adjustment of regional climate model outputs to local scales: Application to Platja de Palma, Spain," *J. Clim.*, vol. 25, no. 3, pp. 939–957, Sep. 2011.
- [24] B. Gerelchuluun and J. B. Ahn, "Air temperature distribution over Mongolia using dynamical downscaling and statistical correction," *Int. J. Climatol.*, vol. 34, no. 7, pp. 2464–2476, Jun. 2014.
- [25] J. Osca, R. Romero, and S. Alonso, "Precipitation projections for Spain by means of a weather typing statistical method," *Global Planet. Change*, vol. 109, pp. 46–63, Oct. 2013.
- [26] Z. Chen, C. Hu, F. E. Muller-Karger, and M. E. Luther, "Short-term variability of suspended sediment and phytoplankton in Tampa Bay, Florida: Observations from a coastal oceanographic tower and ocean color satellites," *Estuarine Coastal Shelf Sci.*, vol. 89, no. 1, pp. 62–72, Sep. 2010.
- [27] N. E. Huang *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Royal Soc. A, Math. Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.
- [28] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, Dec. 2006.
- [29] P. J. Werdell and S. W. Bailey, "An improved in-situ bio-optical data set for ocean color algorithm development and satellite data product validation," *Remote Sens. Environ.*, vol. 98, no. 1, pp. 122–140, 2005.
- [30] J. E. O'Reilly *et al.*, "Ocean color chlorophyll algorithms for SeaWiFS," *J. Geophys. Res.*, vol. 103, no. C11, pp. 24 937–24 953, Oct. 1998.
- [31] B. Choudhury and C. Tucker, "Satellite observed seasonal and inter-annual variation of vegetation over the Kalahari, The Great Victoria Desert, and The Great Sandy Desert: 1979–1984," *Remote Sens. Environ.*, vol. 23, no. 2, pp. 233–241, Nov. 1987.
- [32] D. G. Long, Q. P. Remund, and D. L. Daum, "A cloud-removal algorithm for SSM/I data," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 1, pp. 54–62, Jan. 1999.
- [33] J. Chen, X. Zhu, J. E. Vogelmann, F. Gao, and S. Jin, "A simple and effective method for filling gaps in Landsat ETM+ SLC-off images," *Remote Sens. Environ.*, vol. 115, no. 4, pp. 1053–1064, Apr. 2011.
- [34] X. Zhu, D. Liu, and J. Chen, "A new geostatistical approach for filling gaps in Landsat ETM+ SLC-off images," *Remote Sens. Environ.*, vol. 124, pp. 49–60, Sep. 2012.
- [35] C. Zeng, H. Shen, and L. Zhang, "Recovering missing pixels for Landsat ETM+ SLC-off imagery using multi-temporal regression analysis and a regularization method," *Remote Sens. Environ.*, vol. 131, pp. 182–194, Apr. 2013.



Kaixu Bai received the B.S. degree in urban and rural planning from Northeast Agricultural University, Harbin, China, in 2009 and the M.S. degree in remote sensing and geographic information system from East China Normal University, Shanghai, China, in 2011. He is currently working toward the Ph.D. degree at East China Normal University.

He is currently a Visiting Scholar at the Department of Civil, Environmental, and Construction Engineering, University of Central Florida, Orlando, FL, USA. His research interests include environment

remote sensing and climate change detection and attribution.



Ni-Bin Chang (SM'10) received the B.Sc. degree in civil engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1983 and the M.Sc. and Ph.D. degrees in environmental systems engineering from Cornell University, Ithaca, NY, USA, in 1989 and 1991, respectively.

He is currently a Professor with the University of Central Florida, Orlando, FL, USA. His research interests include integrated data fusion and mining for environmental, ecological, and hydrological applications.



Chi-Farn Chen received the B.Sc. degree in atmospheric physics from National Central University, Jhongli, Taiwan, in 1974 and the M.S. degree in meteorology and the Ph.D. degree in civil and environmental engineering from the University of Wisconsin-Madison, Madison, WI, USA, in 1981 and 1989, respectively.

He is currently a Professor with the Center for Space and Remote Sensing Research, National Central University. His research interests include geographic information systems (GIS), digital image processing, and integration of remote sensing, and GIS for environmental applications.