Landsat 8 remote sensing reflectance (R$_{rs}$) products: Evaluations, intercomparisons, and enhancements

Nima Pahlevan [a,b,*], John R. Schott [c], Bryan A. Franz [a], Giuseppe Zibordi [d], Brian Markham [a], Sean Bailey [a], Crystal B. Schaaf [e], Michael Ondrusek [f], Steven Greb [g], Christopher M. Strait [h]

[a] NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, MD 20771, USA
[b] Science Systems and Applications, Inc., 10210 Greenbelt Road, Lanham, MD 20706, USA
[c] Rochester Institute of Technology, 54 Lomb Memorial Dr, Rochester, NY 14623, USA
[d] European Commission, Joint Research Center, 21027 Ispra, Italy
[e] University of Massachusetts Boston, 100 Morrissey Blvd., Boston, MA 02125, USA
[f] NOAA, NESDIS, STAR, SOCD, College Park, MD 20740, USA
[g] Wisconsin Department of Natural Resources, 2801 Progress Rd., Madison, WI 53716, USA
[h] Upstate Freshwater Institute, 224 Midler Park Dr, Syracuse, NY 13206, USA

ARTICLE INFO

ABSTRACT

The Operational Land Imager (OLI) onboard Landsat-8 is generating high-quality aquatic science products, the most critical of which is the remote sensing reflectance (R$_{rs}$), defined as the ratio of water-leaving radiance to the total downwelling irradiance just above water. The quality of the R$_{rs}$ products has not, however, been extensively assessed. This manuscript provides a comprehensive evaluation of Level-1B, i.e., top of atmosphere reflectance, and R$_{rs}$ products available from OLI imagery under near-ideal atmospheric conditions in moderately turbid waters. The procedure includes a) evaluations of the R$_{rs}$ products at sites included in the Ocean Color component of the Aerosol Robotic Network (AERONET-OC), b) intercomparisons and cross-calibrations against other ocean color products, and c) optimizations of vicarious calibration gains across the entire OLI observing swath. Results indicate that the near-infrared and shortwave infrared (NIR-SWIR) band combinations yield the most robust and stable R$_{rs}$ retrievals in moderately turbid waters. Intercomparisons against products derived from the Visible Infrared Imaging Radiometer Suite (VIIRS) and the Moderate Resolution Imaging Spectroradiometer onboard the Aqua platform (MODISA) indicate slight across-track non-uniformities (<1%) associated with OLI scenes in the blue bands. In both product domains (TOA and R$_{rs}$), on average, the OLI products were found larger in radiometric responses in the blue channels. Following the implementation of updated vicarious calibration gains and accounting for across-track non-uniformities, matchup analyses using independent in-situ validation data confirmed improvements in R$_{rs}$ products. These findings further support high-fidelity OLI-derived aquatic science products in terms of both demonstrating a robust atmospheric correction method and providing consistent products across OLI's imaging swath.

1. Introduction

The Operational Land Imager (OLI) onboard Landsat-8 outperforms the heritage Landsat optical sensors both radiometrically and spectrally. The OLI is a 12 bit pushbroom sensor, which allows for significantly higher signal-to-noise ratio (SNR) than that of the 8 bit whiskbroom Thematic Mapper (TM) and the Enhanced Thematic Mapper Plus (ETM+) (Pahlevan et al., 2014; Pahlevan and Schott, 2013). In addition, the new coastal/aerosol band centered at 443 nm allows for determination of the inherent optical properties (IOPs) of the upper water column

at the chlorophyll *a* peak absorption. The major advantage of the OLI over existing ocean color missions like the Moderate Resolution Imaging Spectroradiometer (MODIS) is its relatively fine spatial sampling, providing the capability to map optically active components (OACs) of upper water column in inland (e.g., lakes, ponds, rivers, wetlands) and near-shore waters. The Landsat-8 revisit time, however, is 16 days, thus only providing periodic snapshots of the near-surface water condition. The OLI data when combined with data acquired with the Multi-Spectral Imager (MSI) aboard Sentinel-2A (and Sentinel-2B in the near future) will allow for more frequent observations at these moderate spatial resolutions. Although this satellite constellation is intended primarily for monitoring landuse/landcover, within the next 10–15 years, the triplet of these sensors will provide a robust set of data products enabling limnologists and ecologists as well as water resource

managers to monitor aquatic systems with minimal costs and at more frequent revisit times (on the order of three days over mid-latitude areas).

A number of studies have started to show the quality of the Landsat-8 data is sufficient for water-related studies. Using the near simultaneous overpasses (n-SNOs) with the MODIS onboard Aqua (hereafter referred to as MODISA) and the Visible Infrared Imaging Radiometer Suite (VIIRS) onboard the Suomi National Polar-Orbiting Partnership (SNPP), it was shown that the OLI radiometric observations at the top of atmosphere (TOA) are well in agreement (on average, <2%) with those of MODISA and VIIRS in reflectance domain (Pahlevan et al., 2014). Although that analysis used a limited set of intercomparisons in 2013, the method was validated against a vicarious calibration approach utilizing in-situ radiometric observations made at the Aerosol Robotic Network Ocean Color (AERONET-OC) sites (Zibordi et al., 2009). The results led to a recommendation on using the TOA reflectance (ρ_t) calibration over the TOA radiance calibration for the retrieval of high-quality aquatic science products. Note that OLI calibration has been conducted in reflectance domain using regular observations of solar diffuser panels (Markham et al., 2014). The performance of the atmospheric correction (ACO) of OLI imagery has recently been demonstrated by both Franz et al. (2015) through an implementation in the SeaWiFS Data Analysis System (SeaDAS) and Vanhellemont and Ruddick (2014) through an implementation in a software suite termed ACOLITE. The goal of the atmospheric correction is to retrieve the spectral remote sensing reflectance ($R_{rs}(\lambda)$ defined as the ratio of water-leaving radiance to the total downwelling irradiance just above the water surface). Note that the wavelength-dependency is, hereafter, dropped for brevity. This quantity (R_{rs}) is critical for deriving the properties of OACs of water, including the IOPs, concentrations of chlorophyll a (Chl), and suspended particulate matter (SPM). Since the launch of Landsat-8 in early 2013, several studies addressing water quality in inland and coastal waters have been conducted. Lymburner et al. (2016) used historic and current Landsat-class instruments to map total suspended sediments (TSS) in the Eastern lakes of Australia. With limited in-situ validations for TSS, they used a land-based ACO method (Vermote et al., 1997) to derive surface reflectance (ρ), which includes sky glint contributions (Mobley, 1999). While the relative shapes of ρ were found reasonable, no in-situ validation was attempted to investigate the quality of products. It is expected that such a land-based ACO may give seemingly reasonable radiometric properties over small inland bodies of water. This is in part due to the fact that the land-based ACO spatially interpolates over-land estimated aerosol types/load over these small bodies of water and retrieve surface reflectance. Therefore, as recommended in Lymburner et al. (2016) an over-water ACO should be considered for high-fidelity retrievals of water constituents (e.g. Chl, TSS, and the colored dissolved organic matter; CDOM) or IOPs. Similarly, care must be taken when standard surface reflectance products (such as those currently provided by the US Geological Survey) are used for aquatic science applications. Furthermore, Brando et al. (2015) utilized OLI-derived R_{rs} to estimate turbidity maps in the northern Adriatic Sea. Gradients were found to qualitatively agree well with modeled circulation patterns and sea surface salinity maps. Barnes et al. (2015) examined OLI-derived R_{rs} to identify dredging activities in the Port of Miami using a thresholding criterion. Lee et al. (2016) developed a semi-analytical approach for the retrieval of Secchi-disk depth from OLI imagery and implemented the method for one OLI scene over the mouth of Jiulongjiang River in southeast China.

The overall goal of this study is to evaluate and enhance OLI-derived R_{rs} products in moderately turbid waters through a comprehensive analysis of the OLI radiometric performance for the R_{rs} and TOA (ρ_t) products for the 2013–2015 timeframe. The analysis begins with the evaluation of OLI-derived R_{rs} products at the AERONET-OC sites. The OLI products (ρ_t and R_{rs}) are then compared against MODISA and VIIRS ocean color products at n-SNOs. This allows for the characterization of OLI's product quality across its 185 km observing swath. Further,

time series of OLI-derived R_{rs} over selected oligotrophic/mesotrophic/eutrophic lake waters are also analyzed against those derived from VIIRS data to independently examine their consistency. The results of the intercomparisons are utilized to extend the vicarious calibration gains (Franz et al., 2007; Gordon, 1998) computed using the Marine Optical Buoy (MOBY; Clark et al. (2003)) observations across the OLI image swath. The validity of the refined, across-track calibration gains is then evaluated against independent in-situ observations collected from various research cruises. The manuscript, thus, develops through sections addressing a) the performance of the atmospheric correction and R_{rs} products, b) intercomparisons of OLI and ocean color products at n-SNOs and over lake waters, and c) the enhancement of OLI's scene-wide across-track performance. Each section provides the associated methodology and results. An additional section provides a discussion on the implications of results, sources of uncertainties, and directions of future research followed by the summary and conclusions.

2. Evaluation of the atmospheric correction (ACO)

The goal here is to evaluate the performance of atmospheric correction implemented in SeaDAS (Franz et al., 2015). However, for completeness, a few of major differences in the implementation as compared to ACOLITE is provided below.

The operational ACO implemented for the removal of aerosol contribution uses a band ratio (epsilon value) of Rayleigh-corrected radiance/reflectance within the near-infrared (NIR) and the shortwave infrared (SWIR) regions, where water-leaving radiance can be assumed negligible to determine the dominant aerosol type and load (Gordon, 1997; Gordon and Wang, 1994). The algorithm utilizes aerosol look-up-tables (LUTs) comprised of epsilon values for 10 aerosol models specified for 8 different humidity levels (Ahmad et al., 2010). However, the implementation of aerosol correction in ACOLITE is based upon exponential extrapolations of the ratio of multiply scattered SWIR aerosol reflectances into the visible bands (Vanhellemont and Ruddick, 2015). Thus, there is no use of aerosol LUTs.

Currently, the SeaDAS implementation incorporates per-pixel geometry information, i.e., solar zenith angle (SZA), solar azimuth angles (SAZ), viewing zenith angle (VZA) and viewing azimuth angle (VAZ), into the processing whereas ACOLITE uses a scene center SZA and a single VAZ across a scene. The incorporation of per-pixel geometry information can be fundamental in high-latitude regions where SZA varies significantly across OLI scenes (185×185 km^2) as well as in lower latitude regions where sunglint (from either the sea surface or high aerosol load) will begin to degrade product quality (an accurate geometry allows for a better estimation of atmospheric scattering and the effects of sea surface reflectance in the ACO process).

The ACOLITE (Vanhellemont and Ruddick, 2014) processing assumes that the aerosol does not significantly spatially vary and automatically applies an aerosol model estimated over a pre-specified window size. A median Rayleigh-corrected reflectance computed per image block is applied in contrast to a per-pixel aerosol removal. This strategy has been shown to significantly improve the quality of the R_{rs} products by minimizing the influence of NIR/SWIR instrument noise. Note that OLI's NIR and SWIR bands were not specifically designed for the typical radiances encountered over water bodies (Pahlevan et al., 2014). In this respect, SeaDAS also provides the capability to define a rectangular window in which the median or mean values can be computed and applied to reduce the influence of noise on the derived products.

2.1. Approach

Three different aerosol correction methods already built into SeaDAS were applied to OLI imagery acquired over AERONET-OC sites. These methods include the use of band combinations of a) 865–1609 nm, b) 865–2201 nm, and c) 1609–2201 nm with the NIR iterations enabled

(Bailey et al., 2010; Stumpf et al., 2004). The aerosol correction was implemented across 7×7 data-elements to minimize the per-pixel noisy epsilon determination. Note that *no* TOA vicarious gains were applied to OLI images prior to this process. The R_{rs} spectra were derived from OLI images by taking the median values over 7×7 data-elements centered at the AERONET-OC sites. Note that 3×3 data-elements (at the center) were discarded to avoid potential adjacency effects from the sites' superstructures.

Since AERONET-OC measurements are performed with 10-nm wide spectral bands (Zibordi et al., 2009) that differ from the OLI band centers, a look-up-table (LUT) created using Hydrolight (Mobley and Sundman, 2008) was applied to derive a hyperspectral model of the in-situ measured R_{rs} spectrum, from which spectral band-pass corrections were derived. A full description of the LUT creation is given in Appendix A. The determination of the best-fit spectral model was carried out via spectral fitting of the modeled spectra and the in-situ multispectral spectrum. The modeled spectrum that satisfies the minimum root-mean squared difference (RMSD) is assumed to represent in-situ hyperspectral R_{rs}:

$$RMSD_j = \sqrt{\sum_{i=1}^{N} \left(R_{rs}^{i} - R_{rs}^{i,j} \right)^2 / N} \qquad (1)$$

where R_{rs}^{i} is the in-situ measurement for channel i, $R_{rs}^{i,j}$ is the modeled spectra resampled to the spectral bands of AERONET-OC for channel i and jth record of the LUT, N is the number of the visible channels (N = 6 for most AERONET-OC sites). Note that, here, modeled hyperspectral R_{rs} were spectrally resampled to 10-nm wide square channels. The corresponding hyperspectral R_{rs} associated with the best fit are further resampled to the OLI relative spectral response (RSR) functions (Pahlevan et al., 2014). The ratio of the two resampled spectra is applied to the OLI-derived R_{rs} to adjust the retrievals to the measured spectra. To perform the assessment, in-situ measurements collected near simultaneously (<1.5 h) with only cloud- and glint-free OLI images were considered. As a result, >80 cloud-free scenes with relatively low aerosol loads, i.e., AOT (869) < 0.1, were identified.

2.2. Results

The fidelity of the OLI-derived R_{rs} associated with the different aerosol removal techniques implemented in SeaDAS is shown in the form of scatterplots in Fig. 1 for n = 94 matchups and four OLI visible bands (443, 482, 561, and 655 nm). At a first glance, it is evident that the performance of ACO using the NIR-SWIR band combinations is considerably better than that of the SWIR bands. The SWIR-only band combination seems to significantly underestimate the R_{rs} in moderately turbid waters typically found at the AERONET-OC sites. Here, we define

moderately turbid waters for the sites where $R_{rs}(655) < 0.6e^{-3}$ [1/sr]. The main reason for the noisy retrievals associated with the SWIR method (1609–2201 nm) is likely explained by the low sensitivity of the corresponding band ratio to various aerosol models, coupled with the low SNR of the SWIR bands. The NIR-SWIR band combination, overall, yields more precise retrievals. However, during peak melt/rain seasons when major river discharges dominate the optical properties of the upper water column, the results of the SWIR band combination tend to converge to the NIR-SWIR outcomes. A handful of such cases were observed at the Wave_CIS (Mississippi Delta), Zeebrugg (Belgium coastal waters) sites, and for a few inland water bodies (not shown here). In other instances where significantly large upwelling NIR radiances are observed, the NIR iteration technique (Bailey et al., 2010) may fail due to the improper bio-optical modeling of the NIR signal. In such cases, the SWIR-only approach seemed to yield better agreement with in-situ (although this observation will have to be extensively tested). The statistics associated with the performance of the various ACO approaches are tabulated in Table 1.

The metrics include the root mean squared errors (similar to Eq. (1)) and the median percentage difference (MPD). The percentage difference at each site was computed as follows

$$PD_i = \left[\left(R_{rs}^{i} - R_{rs}^{i,AER} \right) / R_{rs}^{i,AER} \right] \times 100 \qquad (2)$$

where R_{rs}^{i} is the OLI-derived value in band i and $R_{rs}^{i,AER}$ is the corresponding in-situ measurements. The mean difference (Mean Diff. [1/sr]) refers to the histogram mean of the differences. While the qualitative matchup analysis does not allude to any appreciable differences in the performances of the 865–1609 nm versus the 865–2201 nm method, the statistics suggests that the 865–1609 nm method yields more precise $R_{rs}(443)$. This technique also leads to a slightly better performance at the other three bands. This is potentially due to the differences in the absolute calibration of the two SWIR bands and the errors in the extrapolation of the estimated aerosols (Gordon and Wang, 1994) into the visible bands, i.e., larger errors are expected when aerosols are estimated using two distant channels in the spectral space. It should further be noted that these retrievals were made across different parts of the OLI image swath.

3. OLI intercomparisons with ocean color products

Following the evaluation of OLI-derived R_{rs} products, the OLI radiometric performance across its observing swath was examined using ocean color products to support high-fidelity aquatic science applications. Each OLI image is formed of 14 physically separate linear arrays of 494 detectors (pixels) aligned such that they yield the ~185 km
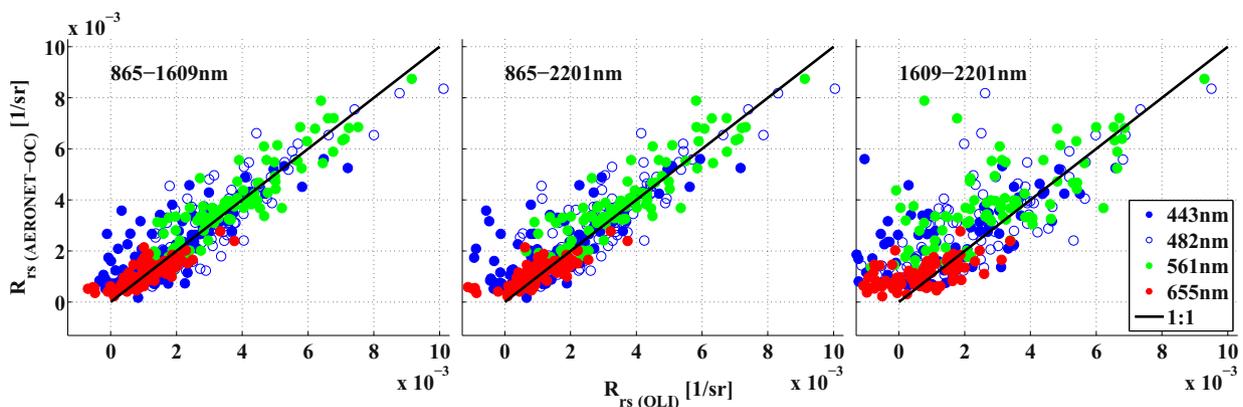

Fig. 1. The matchup analysis (n = 94) carried out at the AERONET-OC sites to evaluate the performance of the atmospheric correction (implemented in SeaDAS) using three different band combinations. The 1609–2201 nm band combination yields the highest uncertainties in the retrievals (right). The 865–1609 nm band combination provides improved retrievals in the 443 nm channel than the 865–2201 nm method. See Table 1 for detailed statistics.

Table 1
The performance statistics associated with the precision of the retrievals using the three band combination methods, based on match-ups with AERONET-OC.

	Band [nm]	RMSE [1/sr]	MPD [%]	Mean Diff. [1/sr]	Slope [ ]	Intercept [ ]
865–1609 nm	443	0.000986	−13.56	−4.3341E−04	0.71	0.000969
	482	0.000839	1.42	−3.6202E−05	0.86	0.000479
	561	0.000739	−1.41	−4.5550E−05	0.83	0.000713
	655	0.000426	−0.59	−4.6916E−05	0.56	0.000484
865–2201 nm	443	0.001173	−22.91	−6.2904E−04	0.66	0.001189
	482	0.000944	−2.16	−2.0278E−04	0.82	0.000735
	561	0.000802	−2.26	−1.6336E−04	0.81	0.000890
	655	0.000492	−4.66	−1.2102E−04	0.51	0.000573
1609–2201 nm	443	0.005006	−69.37	−2.6365E−03	0.07	0.002299
	482	0.004376	−30.25	−1.9935E−03	0.13	0.003051
	561	0.003496	−22.80	−1.5870E−03	0.22	0.003401
	655	0.002754	−53.65	−1.2632E−03	0.07	0.001060

swath. Fig. 2 illustrates the layout of the 14 focal plane modules (FPMs). The across-track nonuniformity of OLI products has not yet been quantified over water bodies. Typically, as part of routine calibration exercises, an across-track nonuniformity analysis is determined from observations of solar diffusers (100% reflector) and the shutter (0% reflector) covering the full dynamic range of the signals (Markham et al., 2014). Here, we leverage the direct intercomparisons of radiometric observations of OLI and concurrent ocean color products from MODISA and VIIRS to essentially cross-calibrate OLI for its across-track nonuniformity artifacts.

3.1. Approach

To evaluate the consistency in the radiometric responses associated with the FPMs, an updated intercomparison approach originally described in Pahlevan et al. (2014) is applied. The near-simultaneous nadir overpasses (n-SNOs) of Landsat-8 and Aqua as well as of Landsat-8 and SNPP are utilized to investigate the consistency of the TOA radiometric measurements as well as the R_{rs} products for pairs of OLI-MODISA and OLI-VIIRS images. The identification of matchup (cross-calibration) sites across the many OLI scenes since launch allows to determine adjustments minimizing these artifacts across various FPMs and thus provide consistent spatial products over inland and coastal waters. The TOA comparisons are made for the visible (443, 482, 561, and 655 nm) and the NIR (865 nm) bands, but only for the visible bands for the R_{rs} performance analysis.

The procedure is implemented as follows:

1. Identify potential OLI-MODISA and OLI-VIIRS matching scenes. The time-lapse (Δt) between the overpasses is restricted to <20 min. While the Landsat-8 and Aqua orbit crossings occur at ~72°N within 14–15 min, the Landsat-8 and SNPP lag time ranges from 3 to 20 min.

Due to ice formation in winter at high latitudes in the Northern hemisphere, the analysis is only carried out for scenes acquired from mid-April to early October.

2. The ocean color (OC) images (MODISA and VIIRS) provided by the NASA Ocean Biology Processing Group (OBPG) are processed using SeaDAS to obtain ρ_t and R_{rs} products using the standard processing flags to mask out potential cloud or sea-ice contaminations. These products are vicariously calibrated with MOBY measurements.

3. The OLI ρ_t data along with the quality-assurance (QA) band were obtained from the US Geological Survey web portal. For a consistent computation of scene-wide ρ_t, per-pixel SZA was applied.

4. Identify potential matchup sites from all of the image pairs that fall within the −8°<VZA<8° range of MODISA and VIIRS observation angles. The analyses were further constrained to the matchups for which the differences in the VZAs were <5°.

5. Filter the matchup sites using the following criteria: a) applying a threshold (T) to the ratio of ρ_t in the NIR bands (0.96<T<1.04) to detect potential changes in the atmosphere within a $\Delta t = 20$min, b) masking out pixels with high aerosol loads (i.e., AOT(865) > 0.05), and c) examining the spatial homogeneity of the OC pixels within 3 km×3 km regions by evaluating the corresponding coefficient of variations (CV). More specifically, the OC pixels whose CV$_{\rho_t(865)}$>0.2 and OLI pixels whose CV$_{R_{rs}(561)}$>0.6 are ignored. These optimal thresholds were determined experimentally by evaluating the standard deviation of the matchups versus the number of data points and are expected to lead to the identification of most ideal matchup sites across each OLI scene. These potential matchup sites are then sorted based upon their spatial variability and (to reduce the computational burden) only top 60 matchups are selected as final matchups per OLI-MODISA and OLI-VIIRS scene pairs.

6. Account for differences in the spatial sampling by using pre-launch point-spread functions (PSFs) of MODISA and VIIRS (Pahlevan et al., 2016). These were applied to OLI imagery at each matchup site to calculate the average of ρ_t over 3 km×3 km regions. Note that only the near-nadir VIIRS PSF is applicable to this study. The support (width) and the orientation of the PSFs were adjusted given the MODISA and/or VIIRS per-site geolocation data (e.g., VZA, VAZ). The unit-area PSFs designed at a very fine grid size are downsampled to 30-m grid cells given a viewing geometry to enable implementation to OLI images.

7. Account for differences in the relative spectral response (RSR) functions at TOA through the application of forward radiative transfer (RT) modeling. For this purpose, the MODerate resolution atmospheric TRANsmission (MODTRAN) was utilized (Berk et al., 2006). To do so, given the sun-sensor geometry, per-site ancillary data, including ozone content, wind speed, water vapor, NO_2, and surface pressure (extracted from the OC ancillary data) were supplied to MODTRAN. In addition, the hyperspectral R_{rs} required for these simulations were
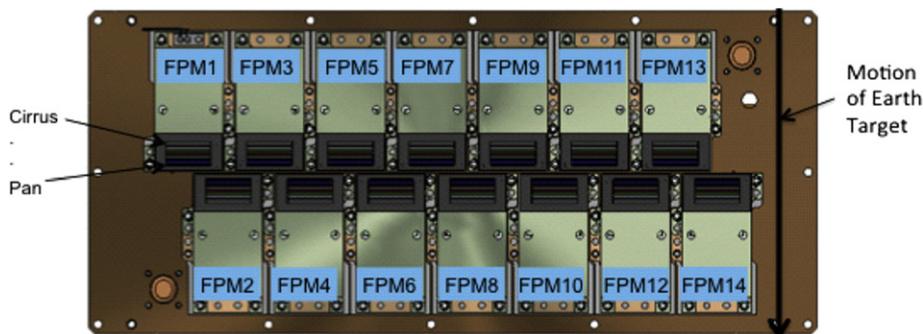
Fig. 2. The layout of the OLI focal plane (Markham et al., 2014). The 185 km swath of an OLI scene is imaged by the detectors distributed across 14 different focal plane modules (FPMs). The odd and even FPMs, which contain nine VNIR + SWIR bands, view each single location on the ground at slightly different viewing angles along track. The panchromatic (Pan) and the Cirrus bands are the closest and farthest bands relative to the satellite nadir vector. Note that FPM01 is on the western side and FPM14 is on the eastern of the OLI scene.

determined from MODISA and/or VIIRS R_{rs} products through spectral fitting as described in Section 2.1. This was performed at each match-up site assuming the in-water and atmospheric optical properties are known. However, recognizing that the aerosol optical properties derived from OC data processing are affected by large uncertainties, the derived AOT(865) was scaled up to ± 80% to model a variety of TOA reflectance spectra. Amongst the 80 different simulated TOA cases (2 aerosol models × 40 AOTs), the curve that best fits the MODISA and/or VIIRS ρ_t was identified. The hyperspectral ρ_t corresponding to the fitted curve was then convolved with the spectral responses of the sensors to compute the corresponding differences. Adjustments were then applied to OLI observations on a band-by-band basis.

8. The matching OLI images were also atmospherically corrected using the 865–2201 nm band combination (currently the default setting in SeaDAS) to obtain R_{rs} for its four relatively broad visible bands (Pahlevan et al., 2014). To allow for consistent intercomparisons of R_{rs}, a similar procedure described in Section 2.1 was followed to account for the differences in the OLI-MODISA and OLI-VIIRS RSRs. It should also be noted that the Rayleigh-corrected reflectances in the NIR/SWIR and the visible bands were spatially averaged with 7×7-element and 3×3-element windows, respectively, to reduce noisy R_{rs} retrievals.

3.2. Results

3.2.1. TOA products

As discussed, the near-nadir observations of MODISA and VIIRS permit the characterization of OLI's across-track nonuniformity over bodies of water. Fig. 3 illustrates temporally average (median) TOA intercomparisons over the years 2013–2015 for all the visible bands. Note that, no calibration adjustments have been applied to the OLI ρ_t at this stage, i.e., standard data obtained from the USGS has been utilized. On the other hand, the MODISA and VIIRS ocean color data underwent

standard NASA OBPG processing (e.g., vicarious calibration, detector destriping). The error bars indicate one standard deviation (σ) derived from all the filtered data points for each FPM during the 2013–2015 timeframe. The OLI across-track nonuniformity is found to be <1.5% for all the bands except in the red band, where high uncertainties hamper the fidelity of intercomparisons. Note that this uncertainty increases with a decrease in the signal. In addition, OLI's SNR associated with this channel (over water bodies) is well below those of MODISA and VIIRS (Pahlevan et al., 2014). On average, OLI exhibits larger radiometric responses in the blue bands (i.e., <2%) and in the green band (i.e., <0.5%) than those of MODISA and VIIRS. On average, 200 matchups (data points) were found for each FPM. Taking uncertainties in the proposed methodology into considerations, the OLI ρ_t, on average, agrees with MODISA and VIIRS data within ± 2%. Note, however, that the analysis refers to the median values calculated for the 3-year period as opposed to the analysis presented in Pahlevan et al. (2014) wherein image pairs in 2013 were only utilized for gain calculations. If the uncertainty in the methodology is represented by $\sigma/\sqrt{n}$ (where n is the number of matchups per FPM), the average uncertainty is estimated to be ~5e^{-4}, which is much smaller than the variations observed between FPMs. Therefore, these across-track nonuniformities are considered significant for aquatic science applications. Although the two independent intercomparisons primarily exhibit similar patterns across track, there are slight differences in the patterns of the blue bands for the FPM11 to FPM14. Overall, the uncertainties in the intercomparisons stem primarily from residual differences in RSRs (caused by either modeling of R_{rs} or ρ_t) and changes in the atmospheric conditions during Δt (Pahlevan and Schott, 2012). It should be noted that the uncertainty in estimation of the median gains was different for different FPMs according to the number of n-SNO events. For example, <100 matchups were found for FPM01 and FPM09 for OLI-MODISA and for FPM13 and FPM14 for OLI-VIIRS, respectively. Hence, larger uncertainties in final gain computation (see Section 5) are anticipated for these FPMs.
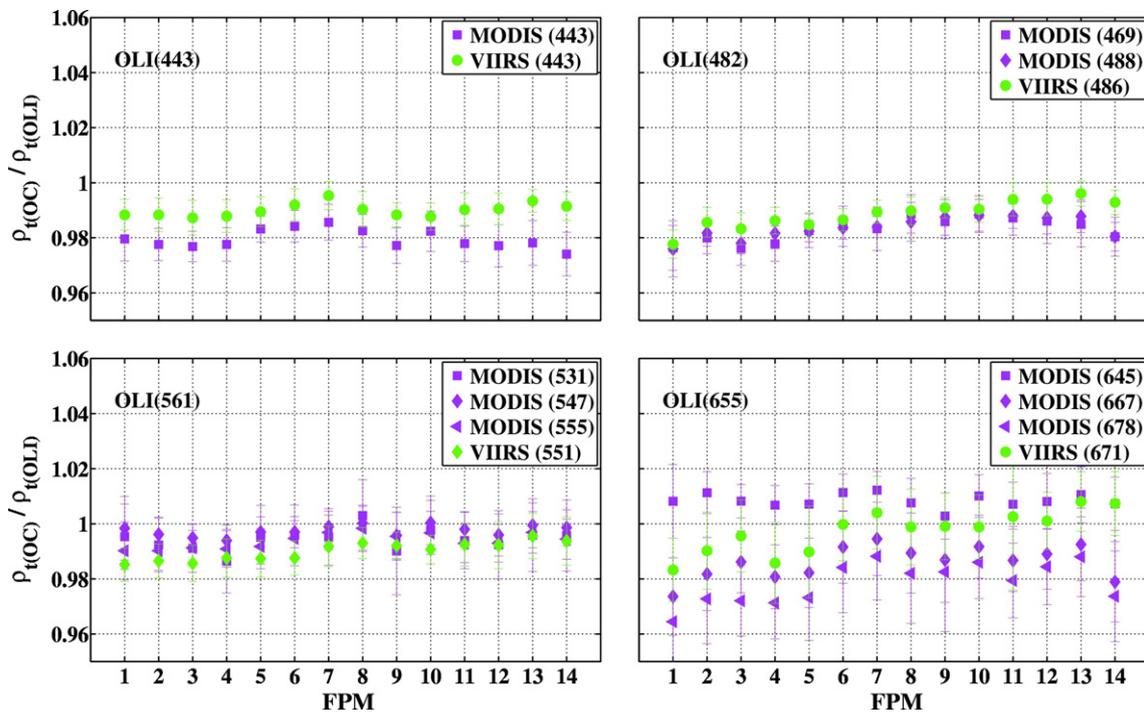
Fig. 3. The OLI across-track non-uniformity shown for the four OLI visible bands. Each data point indicates median values of the ratios of TOA reflectances where OC represents MODISA and/or VIIRS. The OLI aross-track nonuniformity is estimated to be <1% for the blue and green bands. Although high uncertainties are associated with the analysis of the red band, the variability across the swath is ~1.5%. The OLI TOA signals are, on average, larger than those of VIIRS and MODISA. Note also that discrepancies between the VIIRS and MODISA responses are observed. Error bars represent one standard deviation and the average standard error (uncertainty) is estimated to be ~5e^{-4} suggesting that the across-track variations are significant.

The OLI responses may also be used as a reference to compare consistency in the MODISA and VIIRS TOA observations. The intercomparisons imply that MODISA and VIIRS are consistent (<0.5% difference) in all the visible bands, except in the 443 nm where ~1% difference is found. There are, however, noticeable inconsistencies found in the radiometric responses of MODISA red bands where the corresponding land channel (645 nm) exhibit relatively larger responses than those of (non-saturated) ocean bands (667 and 678 nm). This discrepancy may be explained by the minimal overlap between the land and ocean channels where there are differences in the shape of the TOA and water-leaving radiance spectra. In addition, although the intercomparisons have been made in typically dry atmospheres (<60% humidity), there are still water vapor absorption features (~657 nm) that can contribute to the OLI and MODISA (land) TOA signals. Such features are absent in 667 nm and 678 nm ocean channels. Further visual inspection of the intercomparisons in the red bands indicates that the OLI and VIIRS red channels are generally in agreement for FPM > 5 (i.e., eastern side of the OLI scene). It should also be noted that the 645 nm channel of MODISA is designed for the higher radiances experienced over land, and thus operates at a much lower SNR than do the 667 and 678 nm channels over oceans (Franz et al., 2006).

The intercomparisons in the NIR bands are illustrated in Fig. 4. Although it has been shown (through simulations) that ~5% errors in the absolute radiometric responses of the NIR bands would not have appreciable impacts on the magnitude or shape of the R_{rs} spectra (Franz et al., 2007; Wang and Gordon, 2002), in practice, the absolute radiometric responses of these bands may affect the performance of the atmospheric corrections under circumstances that had not been fully investigated (e.g., during high aerosol loads or in coastal areas). It is notable that while the OLI NIR channel (865 nm) is relatively in agreement with MODISA 869 nm (ocean band), there is, on average, an ~3.5% difference in OLI and VIIRS (or MODISA 859 nm band) responses. This does suggest a possible discrepancy between MODISA and VIIRS calibration in the 865 nm–869 nm channels that warrants further investigation.

This discrepancy is consistent across the OLI focal plane (various FPMs) for various OLI-VIIRS and OLI-MODISA intercomparisons. This observation also implies that MODISA and VIIRS are not in full agreement in the 869 nm ocean channel whereas MODISA land band and the VIIRS channels are consistent. The TOA intercomparisons in the SWIR bands are not shown here for brevity. Briefly, results indicate very similar patterns with different magnitudes across the focal plane as in the NIR band (Fig. 4) for both the OLI-VIIRS and OLI-MODISA intercomparisons. The mean across-track and temporal biases for the 1609 and 2201 nm channels were found to be ~1.106 and ~0.993 for OLI-MODISA SWIR bands and ~0.812 and ~0.781 for the OLI-VIIRS SWIR

bands, respectively. Note that standard NASA ocean color processing does not make use of these VIIRS SWIR bands for ACO, due to known calibration issues (Eplee et al., 2015).

3.2.2. R_{rs} products

Similarly, the OLI-derived R_{rs} products can be compared with the OC-derived products at each matchup site. The atmospheric conditions are near ideal (i.e., low aerosol load) and almost identical at each site. The temporally average (median) R_{rs} intercomparisons are illustrated in Fig. 5. The R_{rs} intercomparisons, to a great extent, are consistent with the results in Section 3.2.1 showing that the OLI tends to overestimate the signal (or water-leaving radiances) when compared to the OC products. Although the overall observed across-track nonuniformity is similar to those seen in Fig. 3, some inconsistencies are observed. For instance, for OLI-VIIRS, <1% difference in $\rho_t(443)$ is expected to translate to >10% difference in $R_{rs}(443)$ in mesotrophic waters (Gordon, 1998). However, this difference was found to be ~15–20% in $R_{rs}(443)$. This may be attributed to the differences in the ACO. These discrepancies do not appear in the other three visible channels.

It is also found that the OLI-derived $R_{rs}(443)$ in the blue bands are slightly larger (~10%, on average) towards the edges of the OLI scene than at the scene center. In fact, the OLI products are, in general, found more consistent with those of OC towards the scene center. The average uncertainty (standard error) in these intercomaprisons is estimated to be ~0.007 derived from the standard deviations shown as error bars in the plots. The variability across FPMs in the green channel amounts to less than <10%. The analysis in the red band is largely affected by the intercomparison uncertainties (mean std ~15%) in the red part of the spectrum due to the very low backscattering from water. However, the data points (temporally median values) for all the red bands show similar variability (but different magnitudes) across the OLI focal plane. Although the differences in RSRs have largely been compensated, there remain unaccounted differences. For example, the MODISA 678 nm and VIIRS 671 nm bands partially overlap with the chlorophyll-a fluorescence peak that is outside of the spectral coverage of the corresponding OLI channel (see RSRs in Pahlevan et al. (2014)). In addition, the modeled R_{rs} spectra built into the LUT in Section 2.1 to account for differences in RSRs were based upon an average chlorophyll-a absorption spectrum (see Appendix) that may not represent all water types encountered at matchup sites. For these reasons, these two channels exhibit relatively similar discrepancies against the OLI red channel. On the other hand, the OLI-MODISA intercomparisons in the 645 and 667 nm agree quite well (<5% difference, on average). Further visual inspection of the plots indicates that the VIIRS-derived R_{rs} is only slightly lower in the green and red region than that of MODISA. These discrepancies, however, are within the uncertainty of the methodology presented here.

To further give insights into the consistency of the OLI-VIIRS and OLI-MODISA within the full dynamic range of R_{rs}, scatterplots are shown in Fig. 6. These scatterplots indicate that the products are more consistent for low intensity levels than for the mid-range values and additionally that the OLI is generally in reasonable agreement with MODISA and VIIRS in the green and the red channels.

Also, the root-mean squared differences (RMSD; 1/sr) indicate that OLI and VIIRS are in a better agreement than OLI and MODISA in the $R_{rs}(443)$. These signal-dependent differences may be attributed to a) calibration differences and b) the differences in the removal of aerosols using the NIR-SWIR method versus the operational NIR method (for MODISA and VIIRS). It should also be noted that the scatterplots of OLI-MODISA TOA data (not shown here) also showed signal-dependent variations in the 443 nm whereas the OLI-VIIRS TOA data showed an almost a 1:1 slope for all the channels. This may be related to the known degradation in MODISA radiometric stability after 2012 (Franz et al., 2016). The comparison of the OLI-VIIRS scatterplots for ρ_t and R_{rs} supports the propagation of errors in the ACO through the OLI-derived R_{rs} products. Note that the data points include all the products available within the 2013–2015 timeframe, during which more matchups were
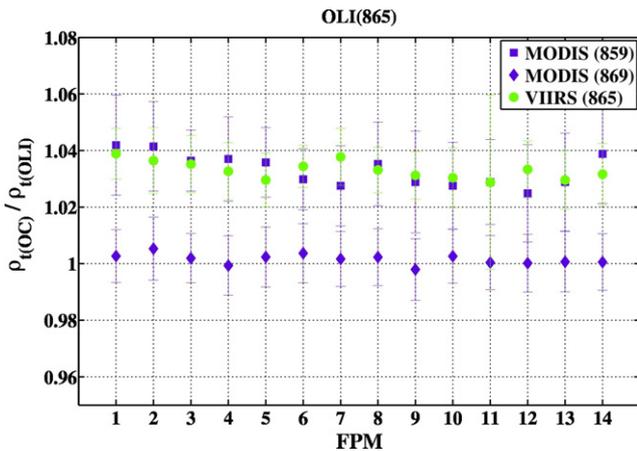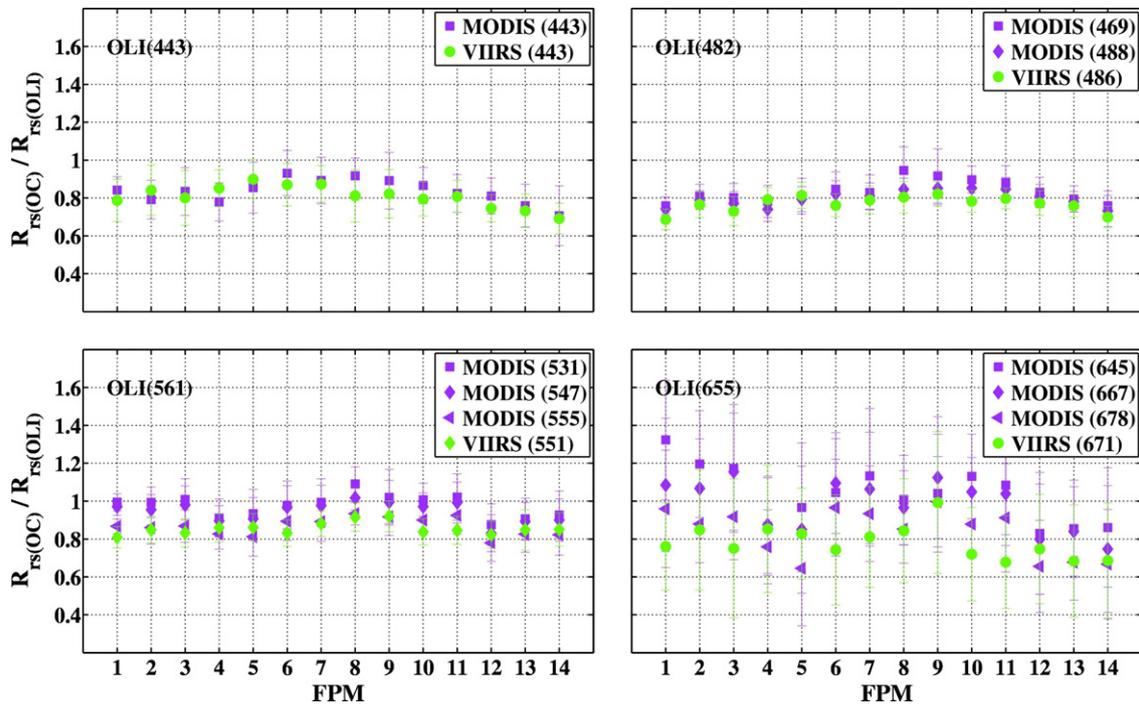
Fig. 4. The intercomparisons in the NIR region. While OLI and MODISA ocean channel agree well across the focal plane, there is ~3.5% difference in the OLI and VIIRS radiometric responses. Note that the across-track nonuniformity of the OLI-VIIRS and OLI-MODISA (land band) are very consistent.

Fig. 5. The OLI-MODISA and OLI-VIIRS intercomparisons of R$_{rs}$ products. The overall trends are similar to the ones in Fig. 3 where TOA intercomparisons are demonstrated, i.e., the OLI products exhibit larger signals and are more consistent with those of VIIRS and MODISA towards the scene center. The two intercomparisons in the blue channels are in good agreement (<5% difference, on average). The larger uncertainties in the red band are expected due to the magnitude of the signal. Error bars represent one standard deviation and the average standard error (uncertainty) is estimated to be ~0.007 (<1%).

found for the OLI-MODISA intercomparisons than those found for the OLI-VIIRS matchups.

The analysis provided in this section does not capture the full seasonal radiometric performance of OLI; in fact, it only includes high-latitude summer scenes. Therefore, to further assess the quality of the time series of OLI products at a more frequent rate (every 16 days), the OLI-derived products over selected inland waters for 2013–2015 are investigated.
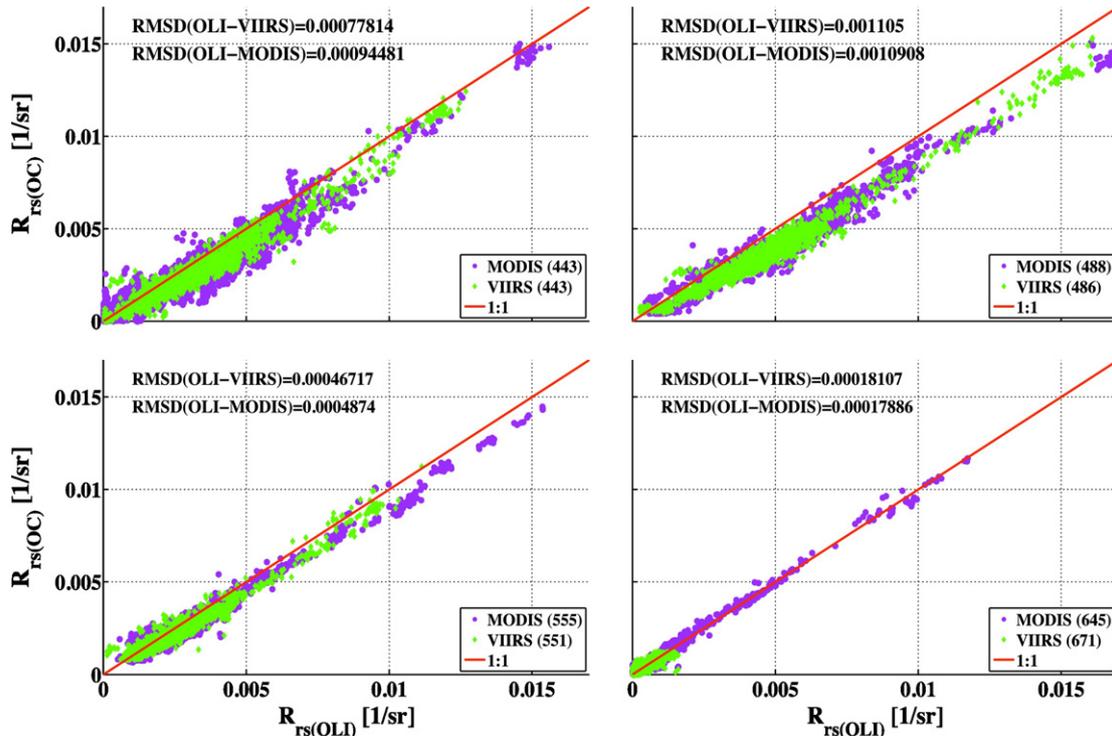
Fig. 6. The scatterplots of OLI- and OC-derived R$_{rs}$ products. From the RMSDs, it is inferred that OLI and VIIRS products are, on average, in better agreement than that of OLI-MODISA (although the maximum difference is found in the 443 nm channel). Further, a signal-dependent difference is clearly observed in the blue bands.

4. Site-specific temporal stability in R_{rs}

To examine the temporal (i.e., 16-day) stability of OLI-derived R_{rs} products, we utilized VIIRS-derived products over various lake waters. This analysis also serves as an independent cross-validation effort for the relative accuracy of OLI-derived R_{rs} products. The lake systems considered here span over a range of optical water types according to the seasonally average VIIRS-derived chlorophyll-a (Chl) products. These lakes include Crater Lake (OR, USA); Chl ≈ 0.2 mg/m³, Lake Tahoe (CA, USA) and Lake General (Argentina); Chl ≈ 0.6 mg/m³, and Lake Garda (Italy) and Lake Dengizkul (Uzbekistan) with 2 < Chl < 10 mg/m³. The OLI images were processed as described in Section 3.1. The VIIRS L1B data were also processed to produce R_{rs} using an equivalent operational processing scheme (746–865 nm band combination for aerosol removal) except that the straylight window size was reduced to 7 × 3-array window to allow for retrievals over inland waters. Furthermore, the median value of a 3 × 3-element window was chosen to represent VIIRS-derived R_{rs} per scene. It should also be noted that the VIIRS retrievals were restricted to VZA < 20° to minimize potential contaminations by adjacent land surfaces and clouds (even though this does not imply nil adjacency effects). Also, as part of the standard SeaDAS processing, a digital elevation model was utilized for the Rayleigh correction. This ensures consistency in the R_{rs} retrievals for both OLI and VIIRS.

4.1. Results

The R_{rs} products derived over Crater Lake (left column) and Lake Tahoe (right column) are shown in Fig. 7 for the three visible bands. The optical regime of these lakes is primarily driven by absorptions of phytoplankton and pure water, which facilitates a relatively robust atmospheric correction, i.e., near-zero water-leaving radiance in the NIR and SWIR regions. The data are shown for a time series since the beginning of the year 2013.

The VIIRS data (squares) represent 16-day median values to enable temporally consistent comparisons. It is evident that the OLI reproduces the same seasonal patterns measured by the VIIRS in the blue bands. For these relatively less productive lakes, the OLI-derived time series becomes noisier than that of VIIRS in the green channel. This observed noise is more pronounced for Crater Lake than Lake Tahoe because of its lower productivity. This is perhaps because a) the OLI prodcuts are used "as is" whereas VIIRS products are median filtered, b) of the noisy nature of OLI's NIR-SWIR detectors (even after a region-based aerosol determination), and c) of the better performance of the VIIRS ACO using two NIR bands. The resulting noisy $R_{rs}(555)$ values will have implications for the retrieval of Chl using band ratio techniques for Crater Lake. Further differences may be explained by the differences in the overpass time/dates and the lakes' temporal dynamics.

The median values of the time series shown on the plots can also provide insights into the overall differences in the products. It is found that OLI-derived products are consistently higher than that of VIIRS except in the $R_{rs}(443)$ for Lake Tahoe where the OLI and VIIRS values agree well (within 5%). Note that the two lakes are positioned over two different (adjacent) FPMs, i.e., FPM07 and FPM08, as indicated on the plots. This overestimation amounts to ~10–40% for Crater Lake in all the visible bands. For Lake Tahoe, the OLI-derived products are only 9% larger in magnitude than that of VIIRS in the 482 and 561 nm channels. The time series for the $R_{rs}(655)$ products (not shown here) was dominated primarily by negative values ($-0.0005 < R_{rs}(655) < 0$) for both OLI and VIIRS with OLI exhibiting larger variability. Negative retrievals (in the red band) were also found for other lake systems except for Lake Dengizkul where the optical regime is governed by relatively high Chl (~6 mg/m³, on average) and suspended sediment loads. Note that the trends in Chl are not shown here (beyond the scope of this research).

5. Scene-wide calibration

While prior analyses in Section 3 demonstrated how consistent the OLI-VIIRS and OLI-MODISA products are, they do not provide insights into the absolute radiometric performance of OLI. The goal here is to update the MOBY-derived TOA calibration gains and extend that to the entire OLI observing swath using findings from Section 3.

High temporal revisit times of missions like MODISA or VIIRS permits statistically robust derivation of vicarious gains, i.e., $G(\lambda)$, over the MOBY site. However, since the Landsat-8 launch, there have been only a handful of suitable OLI scenes (i.e., cloud-free with low aerosol loading) over the MOBY site. In addition, this site is always located within the OLI's FPM10, which limits the ability to derive scene-wide gains to minimize its across-track nonuniformity (Section 3.2.1). It is, thus, imperative to utilize MOBY observations in conjunction with intercomparison (cross-calibration) results (Section 3.2.1) to ensure highest-quality products across the OLI image swaths.

5.1. Approach

A similar procedure described in Franz et al. (2007) and implemented in SeaDAS was followed using the MOBY R_{rs} data to compute the gains from the available scenes for FPM10. These scenes were acquired in wintertime and are characterized by very low aerosol loading (i.e., AOT(865) < 0.1) with a minimal spatial variability in R_{rs} products in the proximity of the MOBY. Note that the gains initially reported in Franz et al. (2015) were based on the TOA radiance products, which are updated here to use the ρ_t product (Markham et al., 2014; Pahlevan et al., 2014). The vicarious gains for FPM10 were averaged using both 865–2201 nm and 865–1609 nm band combinations. The average gains were computed as G(443) = 1.011, G(482) = 1.0101, G(561) = 1.0072, G(655) = 1.0098, G(865) = 1.00, G(1609) = 0.956, and G(2201) = 1.01. Note that the NIR gain is fixed at unity in both methods and new SWIR gains are computed. The calibration gains for the rest of the FPMs are derived relative to FPM10 gains:

$$G_j = \frac{G_{w,j}}{G_{w,10}} G \qquad (3)$$

where $G_{w,j}$ is the weighted average of the two intercomparisons illustrated in (Fig. 3), $G_{w,10}$ is the weighted average gain for FPM10, G is the vicarious calibration gain for FPM10, and G_j is the final computed gain for the jth FPM. Note that the spectral dependency has been dropped. The final per-FPM gains are shown in Fig. 8. As expected, the across-track nonuniformity resemble those shown in Fig.3. After applying the gains, the R_{rs} retrievals are re-evaluated against AERONET-OC observations and validated with independent in-situ measurements.

5.2. Results and validations

The overall retrieval errors after applying updated vicarious gains (computed for all the AERONET-OC sites) are provided in Table 2. Here, the statistics are compared to those tabulated in Table 1, which denote retrieval errors associated with OLI standard products (i.e., prior to calibration updates). It is inferred that the overall retrieval noise (as determined by the RMSE) has been reduced, on average, by 10%. Slight improvements in the slope parameter are also observed. The most significant improvements are found in the 1609–2201 nm method for which the mean differences, i.e., histogram means of the differences, are diminished by an order of magnitude. While the bias in retrievals can mostly be attributed to biases in the the SWIR bands, the noise is likely associated with the FPM-to-FPM variability in the SWIR bands. Following the gain updates, the 865–1609 nm method still marginally outperforms the 865–2201 nm method. These findings indicate that the ACO can be sensitive to small calibration uncertainties in the
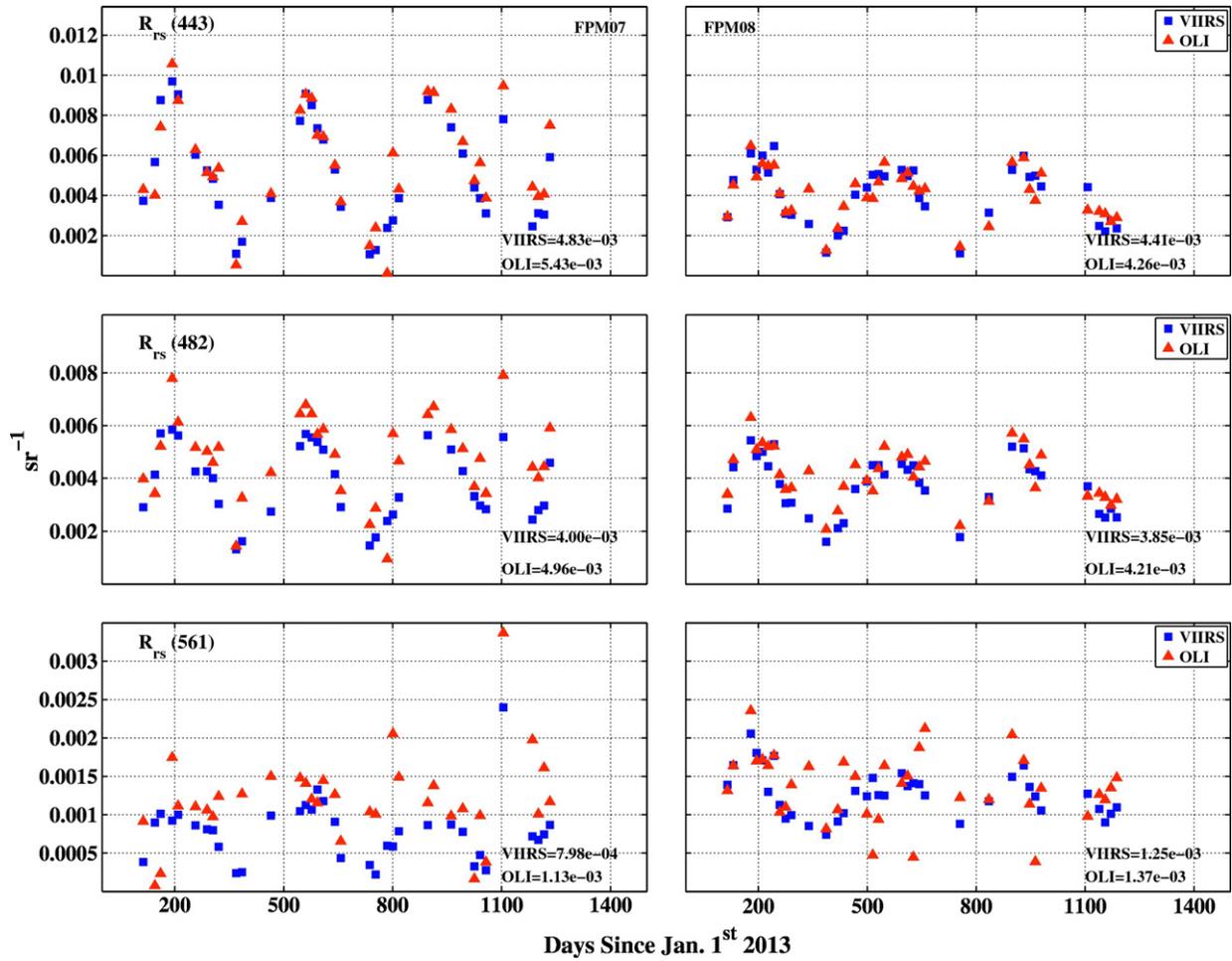
Fig. 7. The time series of $R_{rs}(443)$, $R_{rs}(482)$, and $R_{rs}(561)$ derived from OLI (triangles) and VIIRS (squares) over the Crater Lake (left) and Lake Tahoe (right). The OLI is found sufficiently sensitive to changes in seasonal cycles in water column. The temporally median values (annotated) indicate that OLI products are brighter than those of VIIRS. For these water types characterized by low productivity, OLI-derived $R_{rs}(561)$ is found noisier than that of VIIRS.

SWIR bands. It is believed that the SWIR across-track nonuniformity can further be reduced via further analyses of numerous OLI scenes.

It is also important to evaluate the level of improvements for each FPM. In the following, we provide retrieval errors in R_{rs} as % differences for updated gains (G_{update}) and the unity gains (G_{unity}). Fig. 9 shows average retrieval errors over AERONET-OC sites for four individual

FPMs (FPM04, FPM06, FPM07, FPM09) as examples. The errors are given only for the two NIR-SWIR band combinations. It is clearly seen that the gain adjustments have diminished the retrieval errors for these FPMs. The enhancements, on average, have the largest impact on the 443 nm channel. Note, however, that, the errors are minimized, on average, for each FPM.
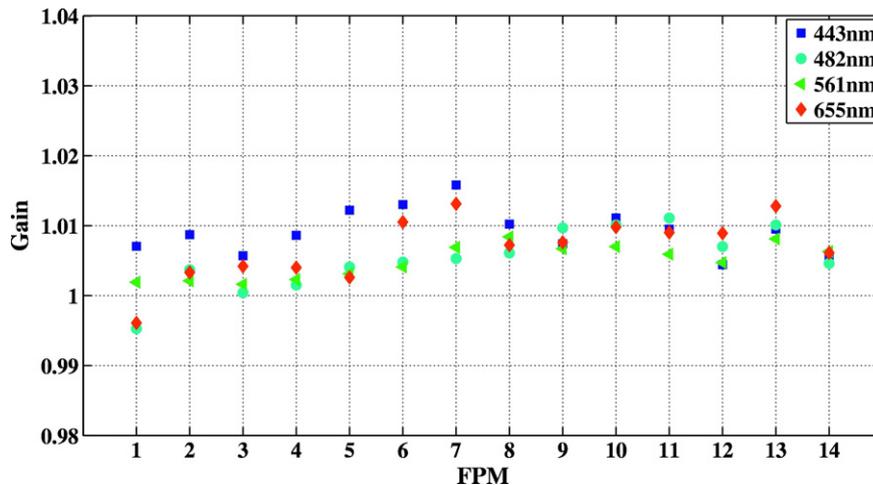

Fig. 8. The per-FPM calibration adjustments computed using a combination of MOBY observations and cross-calibration against ocean color missions (Eq. (3)). The derived gains will be updated as more matchups become available.

Table 2
The overall statistics associated with retrieval errors after calibration enhancements.

|	Band [nm]	RMSE [1/sr]	MPD [%]	Mean Diff. [1/sr]	Slope [ ]	Intercept [ ]
865–1609 nm	443	0.000910	4.63	−2.7398E−05	0.71	0.000695
	482	0.000779	1.35	−2.1924E−05	0.88	0.000408
	561	0.000697	−5.73	−6.1709E−05	0.88	0.000655
	655	0.000418	1.61	−1.5007E−05	0.57	0.000454
865–2201 nm	443	0.000961	12.29	8.4391E−05	0.68	0.000671
	482	0.000832	7.33	7.3005E−05	0.85	0.000409
	561	0.000745	−3.83	−1.5523E−04	0.84	0.000751
	655	0.000469	7.84	2.2126E−05	0.53	0.000484
1609–2201 nm	443	0.003169	48.89	5.4605E−04	0.15	0.001864
	482	0.002663	30.64	4.7937E−04	0.29	0.002170
	561	0.002140	13.15	1.6676E−04	0.39	0.002318
	655	0.001726	50.83	2.9759E−04	0.13	0.000868

The improvements in the retrievals have been verified with a set of independent in-situ R_{rs} measurements. These validation datasets (30 matchups in total) come from robust measurement techniques (e.g., Lee et al., 2013; Mueller et al., 2003). The measurements were carried out in the Massachusetts Bay, Chesapeake Bay, Lake Michigan, Green Bay, Boston Harbor, and coasts of South Carolina. The percentage errors (Eq. (2)) averaged for each FPM are tabulated in Table 3. These validations further strongly suggest that the updated vicarious gains reduce the retrieval errors (in particular in the 443 nm channel). The most pronounced enhancements are associated with the blue channels at the edge of the swath, i.e., FPM01. Note that these retrievals were performed under near-ideal atmospheric conditions, i.e., AOT(865) < 0.1. Nevertheless, imperfect atmospheric correction results in somewhat large errors as indicated by % differences.

To further examine the impacts of improvements in OLI-derived R_{rs} products, several cloud-free OLI scenes over open oceans (nearby islands) were processed and analyzed before and after the implementation of gains shown in Fig. 8. The nonuniformity was tested by evaluating the standard deviation computed across track for subsamples of rows, which satisfied minimal in-water variability amongst all the pixel rows. By comparing standard deviations before and after gain implementations, on average, ~10% reduction in across-track variability

was found. The improvement ranged from 0 to 20% depending on the band. Fig. 10 illustrates four full-swath R_{rs}(443) (a & b) and R_{rs}(655) (c &d) products derived from ρ_t before (upper panels) and after implementing the gains. These plots qualitatively show how the application of across-track gains has reduced the FPM-to-FPM variability. At first glance, one infers that the overall brightness of the products have increased due to applying gains > 1 for most FPMs. However, residual variability is still pronounced. These variabilities may be attributed to a) uncertainties in gain computations, b) non-uniformity within each FPM, and c) large across-track (FPM-to-FPM) nonuniformities in the NIR and SWIR bands. Analyzing many more OLI-derived R_{rs} products (not shown here) indicated slight signal-dependent FPM-to-FPM variability, i.e., higher variability in clear waters and less for highly turbid waters. Larger FPM-to-FPM differences were also observed for imagery acquired at high solar zenith angles and under challenging atmospheric conditions (e.g., high aerosol loads or when sun glint is present).

6. Discussions and future directions

Throughout this manuscript, we have examined the agreement of OLI products with in-situ data (R_{rs}) as well as with those obtained from MODISA and VIIRS products (i.e., ρ_t and R_{rs}). While it was demonstrated that the performance of the NIR-SWIR band combination well exceeds that of the 1609–2201 nm (SWIR) method, the latter method showed promising retrievals in instances where large outgoing water-leaving radiances are found in the NIR region (as also recently demonstrated in Vanhellemont and Ruddick, 2015). In such circumstances (i.e., extremely turbid waters), the NIR iteration method (Bailey et al., 2010), which uses a generic bio-optical model to estimate water-leaving radiance, is expected to exhibit degraded performance. On the other hand, the poor sensitivity of the band ratio of the SWIR bands to aerosols together with the noisy SWIR detectors adversely impacts the R_{rs} retrievals in moderately turbid waters. Thus, a similar switching option available for processing MODIS data (Shi and Wang, 2009) built into SeaDAS is required for the OLI processing to allow the algorithm to choose either NIR-SWIR or SWIR band combinations given the turbidity level. The majority of the retrievals presented here were carried out in relatively near-ideal atmospheric conditions. In general, as one
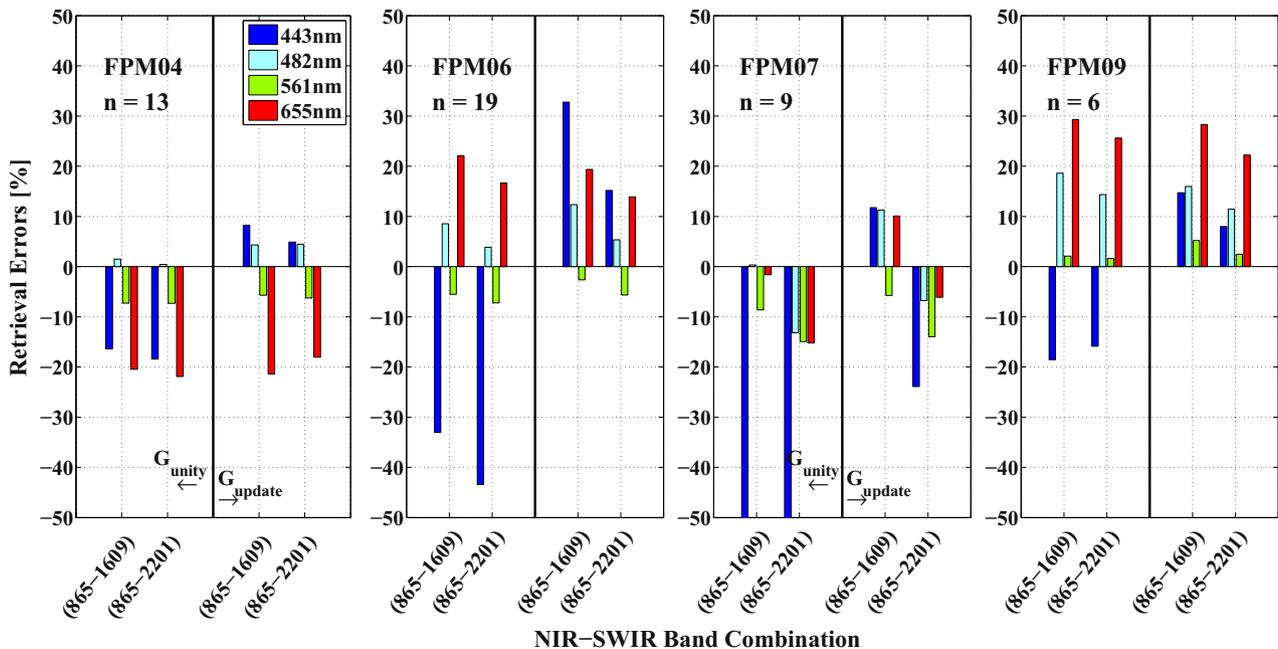
Fig. 9. The retrieval errors shown for two different band combinations and gain settings (i.e., unity and update). The retrieval errors are reduced by extending the MOBY-derived gains across the OLI focal plane and further minimizing for each FPM using AERONET-OC observations.

Table 3
The average per-FPM retrieval errors (%) found for in-situ matchups. The retrievals are carried out using the 865–1609 nm method. The integers in round brackets refer to the number of samples.

Band	FPM01 (4)		FPM04 (4)		FPM05 (6)		FPM06 (3)		FPM07 (7)		FPM08 (4)		FPM09 (1)		FPM10 (1)	
	G_{unity}	G_{update}	G_{unity}	G_{update}	G_{unity}	G_{update}	G_{unity}	G_{update}	G_{unity}	G_{update}	G_{unity}	G_{update}	G_{unity}	G_{update}	G_{unity}	G_{update}
443	36.11	17.74	−61.54	−34.7	−70.44	−46.96	−98.7	−49.7	−49.84	−22.6	−32.07	−18.12	−62.91	28.86	−12.93	−12.92
482	49.17	14.15	−16.11	−14.11	−19.54	−15.54	−42.67	−31.08	−23.8	−19.23	−16.67	−11.18	−33.39	4.23	−11.33	−11.72
561	13.88	9.56	−6.03	−5.32	−14.73	−13.23	−20.85	−16.32	−19.45	−16.6	−10.09	−8.18	−22.11	−11.32	11.06	−12.08
655	9.67	10.54	−15.23	−12.36	−30.87	−26.87	−42.42	−32.05	−23.12	−11.96	−8.87	−4.08	−21.39	−2.67	−4.6	−3.92

would expect (Gordon et al., 1997), the quality of products degrades with the increase in the aerosol loading and at higher solar zenith angles, which together would result in larger errors or masking of pixels. Detailed analysis of such issues is beyond the scope of this study. However, it is well worth scrutinizing the performance of the atmospheric corrections for various atmospheric conditions in the future. Note further that we did not make any attempt here to evaluate the performance of atmospheric correction over extremely eutrophic, turbid or CDOM-rich waters. These conditions occur primarily in the proximity of human settlements. Although per-pixel geometry is accounted for in the SeaDAS ACO processing, residual discrepancies for adjacent FPMs (i.e., banding) has been observed at high AOTs/solar zenith angles. This artifact is believed to be due to the differences in the viewing angles of odd and even FPM detectors. Therefore, further research needs to be dedicated to the retrievals at large solar angles and high AOTs to maximize the utility of OLI (or MSI) in the future for water quality mapping. Adopting new aerosol models specifically representing land/coastal aerosols will also potentially enhance retrievals from OLI (or MSI).

The major sources of errors contributing to the uncertainties in the intercomparison analysis presented here were a) the time difference in the overpass times of satellites and b) inaccurate characterization of the differences in the spectral response functions. Both of these factors may contribute to noise in the time series analyses and can be reduced by increasing sample size. Moreover, the lack of knowledge on a) the hyperspectral distribution of water-leaving radiances and b) the amount of total water vapor in the atmosphere further impacts these intercomparisons. For

instance, the quality of the matchup analysis degraded for calibration sites located in high-latitude eutrophic waters due to inaccurate modeling of hyperspectral water-leaving radiances. This is expected, as aquatic systems with higher trophic levels are rich in spectral information that may not be captured in our modeled spectra (Section 2.1 & Appendix A). The final vicarious gains (shown in Fig. 8) were also subject to uncertainties associated with the MOBY-derived gains, the cross-calibration approach, and corresponding image artifacts (e.g., striping) attributed to OLI, MODISA, and VIIRS. The standard error (i.e., $\sigma/\sqrt{n}$) of the MOBY-derived gains and the cross-calibration method in the blue bands were ~0.008 and ~$5e^{-4}$, respectively, which indicates higher uncertainties in MOBY-derived gains. Therefore, the vicarious gains need to get updated as more MOBY matchups become available. Although the standard error in cross-calibration is relatively low, attempts will also be made to further update per-FPM gains in the future. Currently, similar analyses are being conducted for the MSI onboard Sentinel-2A to evaluate the consistency of the corresponding R_{rs} products with those derived from OLI to provide seamless moderate resolution aquatic science products in coastal/inland waters.

While the across-track nonuniformity of the OLI at TOA was found < 1.5% for most of the FPMs, our results suggest more rigorous requirements for pre-launch characterization of OLI-2 onboard Landsat-9 currently planned for launch in 2021. The OLI's across-track and along-track non-uniformity may originate from a) variations in the spectral response between detectors and focal plane modules, b) uncorrected non-linearity at low signal levels, and c) data quantization effects. The OLI-2
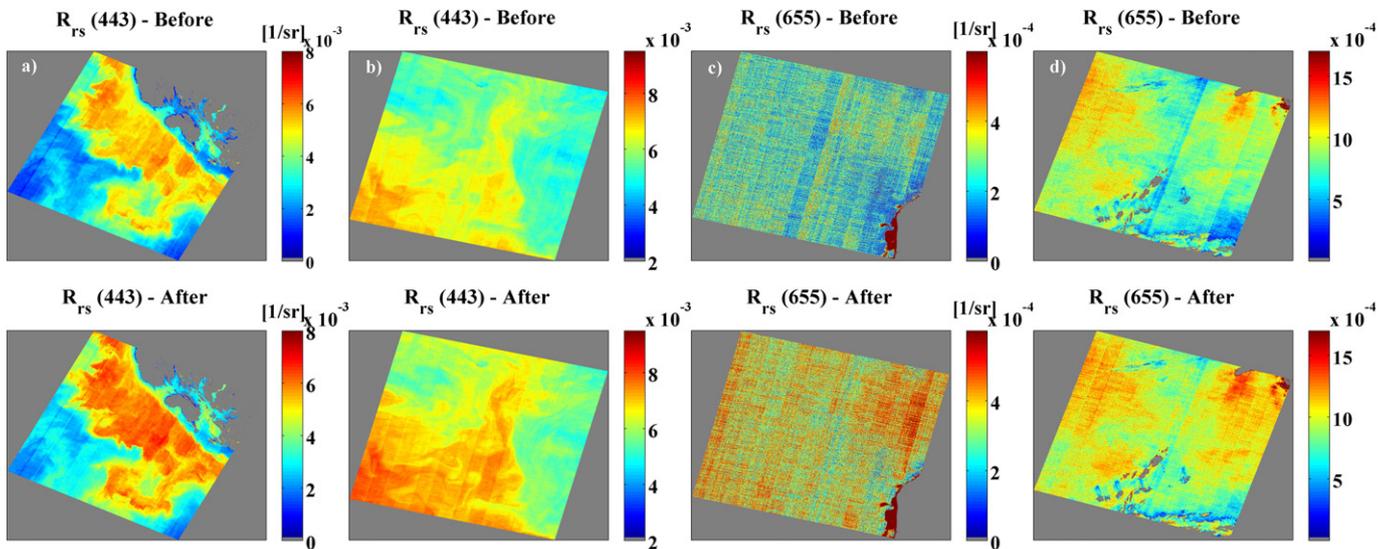
Fig. 10. The Rrs(443) & Rrs(655) products derived from four OLI scenes before (top) and after (bottom) the implementation of the updated gains. In general, the overall signal has increased and the across-track nonuniformity is reduced. However, there are still variabilities due to a) uncertainties in gain computations, b) non-uniformity across each FPM, and c) large across-track (FPM-to-FPM) nonuniformity in the NIR and SWIR bands. The corresponding scene IDs are LC81780102013213LGN00 (a), LC80410382015043LGN00 (b), LC80750882013131LGN01 (c), and LC80230182015253LGN00 (d).

will be fully spectrally characterized per detector at the integrated instrument level. In addition, all 14-bits from the OLI will be retained and transmitted to the ground, so that the data will not be quantization-noise limited at low signal levels. Further, more detailed non-linearity tests at low radiance levels are planned. Together, these characterizations will yield enhanced product fidelity for aquatic science/applications.

7. Summary and conclusion

This manuscript provides a comprehensive analysis of the OLI-derived R_{rs} products for aquatic science/applications. The following major findings are summarized below:

- In near-ideal atmospheric conditions, the NIR-SWIR band combination provides much more stable R_{rs} retrievals than those from the SWIR-based method in moderately turbid waters. The 865–1609 nm band combination marginally outperforms the 865–2201 nm method.
- The OLI across-track nonuniformity over bodies of waters is <1.5% for most of the focal plane modules (FPMs) and spectral bands. The satellite inter-comparisons at near simultaneous nadir overpasses (n-SNOs) allowed for a high-fidelity characterization of the across-track nonuniformity in the 443, 482, and 561 nm bands.
- The OLI-derived R_{rs} products are found brighter (~10%, on average) in the blue bands than those of VIIRS and MODISA. The products are found most consistent in the green band.
- The retrievals of R_{rs} from OLI and VIIRS over five different lake systems demonstrated that the OLI is sufficiently radiometrically sensitive to seasonal variations in different aquatic ecosystems.
- Analyzing the scatterplots of intercomparisons (OLI-VIIRS) showed a linear signal-dependent trend for the blue and green bands. While the OLI products tend to be higher in magnitude within larger signal ranges (e.g., R_{rs} (443) > 0.003), they tend to become darker (or agree more) at lower signal levels. This trend is believed to be due to calibration uncertainties in the SWIR bands as well as the differences in the aerosol removal techniques.
- The updated calibration gains were demonstrated to reduce the retrieval errors. The refined (adjusted) gains were validated against independent in-situ observations, which indicated improvements primarily in the R_{rs} (443).

Designed for global landuse/landcover change monitoring, the OLI has demonstrated sufficient radiometric performance to be of potential use to the aquatic science community. This effort was dedicated to a) diagnosing potential shortcomings of the OLI for aquatic science applications, b) fully evaluating best-practice methodologies for atmospheric correction, and c) providing adjustments to further enhance data products. The results will help maximize the utility of the OLI aquatic science products for studying inland and coastal waters and assist users in better understanding potential inconsistencies with ocean color products. This analysis serves as a first step to provide seamless aquatic products from inland and nearshore waters into shelf waters and open oceans by merging moderate-resolution and ocean color products. Over the next decade, the combined OLI-MSI products have the potential to support operational monitoring of water quality at temporal and spatial scales (10 to 60 m) that have never been possible before. It is, thus, critical for the science community to devise new algorithms and methodologies explicitly suited for Landsat-class satellites to help informed decision-making and mitigate public health issues relevant to water quality.

Appendix A. Appendix

To create the LUT of various R_{rs} spectra utilized in Sections 2 & 3, the widely used Hydrolight software suite was used (Mobley and Sundman, 2008). The specific IOPs shown in Fig. A1 were used and remained unchanged for all the simulations. The spectra shown are normalized to 440 nm. The mineral specific absorption and scattering coefficients correspond to calcareous sand (available through Hydrolight package).

The IOPs were scaled by varying concentrations of Chl (mg/m^3) and TSS (g/m^3) as well as CDOM absorption at 443 nm, i.e., a_{cdom}(443). The Chl and TSS were varied up to 18 units at fine increments to cover a broad range of water types. The a_{cdom}(443) was similarly varied up to 3 m^{-1}. Note that only physically realistic combinations of constituents were supplied (e.g., a_{cdom}(443) = 2 m^{-1} was not used with Chl = TSS < 0.5). These variations resulted in >68000 R_{rs} spectra within the visible domain. The LUT yielded a normal distribution for a(443) ranging from 0.05 to 4.5 m^{-1} with a mean and standard deviation of 2.1 and 0.86 m^{-1}, respectively. The backscattering coefficients at 655 nm, i.e., b_b(655), ranged from 0.001 to 0.3 m^{-1} with the median value of 0.065 m^{-1}. The other parameters, including the Fournier-Forand phase functions (Fournier and Forand, 1994) associated with backscattering ratio of 2%, infinite water-column depth, wind speed of 5 m/s, and solar zenith angle of 45°, remained constant for all the simulations.

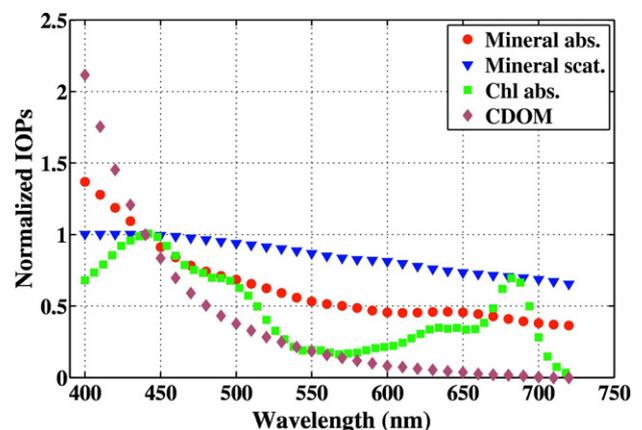
Fig. A1. The normalized IOPs utilized for the developments of the LUT applied in Sections 2 & 3.

References

Ahmad, Z., Franz, B.A., McClain, C.R., Kwiatkowska, E.J., Werdell, J., Shettle, E.P., Holben, B.N., 2010. 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. 49, 5545–5560.

Bailey, S.W., Franz, B.A., Werdell, P.J., 2010. Estimation of near-infrared water-leaving reflectance for satellite ocean color data processing. Opt. Express 18, 7521–7527.

Barnes, B.B., Hu, C., Kovach, C., Silverstein, R.N., 2015. Sediment plumes induced by the Port of Miami dredging: analysis and interpretation using Landsat and MODIS data. Remote Sens. Environ. 170, 328–339.

Berk, A., Anderson, G.P., Acharya, P.K., Bernstein, L.S., Muratov, L., Lee, J., Fox, M., Adler-Golden, S.M., Chetwynd, J.J.H., Hoke, M.L., Lockwood, R.B., Gardner, J.A., Cooley, T.W., Borel, C.C., Lewis, P.E., Shettle, E.P., 2006. MODTRAN5: 2006 Update, 62331F-62331F.

Brando, V., Braga, F., Zaggia, L., Giardino, C., Bresciani, M., Matta, E., Bellafiore, D., Ferrarin, C., Maicu, F., Benetazzo, A., 2015. High-resolution satellite turbidity and sea surface temperature observations of river plume interactions during a significant flood event. Ocean Sci. 11, 909.

Clark, D.K., Yarbrough, M.A., Feinholz, M., Flora, S., Broenkow, W., Kim, Y.S., Johnson, B.C., Brown, S.W., Yuen, M., Mueller, J.L., 2003. MOBY, a radiometric buoy for performance monitoring and vicarious calibration of satellite ocean color sensors: measurement and data analysis protocols. Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 4, 3–34.

Eplee, R.E., Turpie, K.R., Meister, G., Patt, F.S., Franz, B.A., 2015. Updates to the on-orbit calibration of SNPP VIIRS for ocean color applications. *SPIE Optical Engineering + Applications*. International Society for Optics and Photonics (pp. 96071P-96071P-96018).

Fournier, G.R., Forand, J.L., 1994. Analytic Phase Function for Ocean Water. pp. 194–201.

Franz, B.A., Werdell, P.J., Meister, G., Kwiatkowska, E.J., Bailey, S.W., Ahmad, Z., McClain, C.R., 2006. MODIS land bands for ocean remote sensing applications. Proc. Ocean Optics XVIII, Montreal, Canada.

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

Franz, B.A., Bailey, S.W., Kuring, N., Werdell, P.J., 2015. Ocean color measurements with the Operational Land Imager on Landsat-8: implementation and evaluation in SeaDAS. J. Appl. Remote. Sens. 9, 096070.

Franz, B.A., Siegel, D.S., Behrenfeld, M.J., Signorini, S.R., 2016. Global ocean phytoplankton [in State of the Climate in 2015]. Bulletin of the American Meteorological Society, pp. S87–S89.

Gordon, H.R., 1997. Atmospheric correction of ocean color imagery in the earth observing system era. Journal of Geophysical Research: Atmospheres (1984–2012) 102, 17081–17106.

Gordon, H.R., 1998. In-orbit calibration strategy for ocean color sensors. Remote Sens. Environ. 63, 265–278.

Gordon, H.R., Wang, M., 1994. Retrieval of water-leaving radiance and aerosol optical thickness over the oceans with SeaWiFS: a preliminary algorithm. Appl. Opt. 33, 443–452.

Gordon, H.R., Du, T., Zhang, T., 1997. Atmospheric correction of ocean color sensors: analysis of the effects of residual instrument polarization sensitivity. Appl. Opt. 36, 6938–6948.

Lee, Z., Pahlevan, N., Ahn, Y.-H., Greb, S., O'Donnell, D., 2013. Robust approach to directly measuring water-leaving radiance in the field. Appl. Opt. 52, 1693–1701.

Lee, Z., Shang, S., Qi, L., Yan, J., Lin, G., 2016. A semi-analytical scheme to estimate Secchi-disk depth from Landsat-8 measurements. Remote Sens. Environ. 177, 101–106.

Lymburner, L., Botha, E., Hestir, E., Anstee, J., Sagar, S., Dekker, A., Malthus, T., 2016. Landsat 8: providing continuity and increased precision for measuring multi-decadal time series of total suspended matter. Remote Sensing of Environment 185, 108–118.

Markham, B., Barsi, J., Kvaran, G., Ong, L., Kaita, E., Biggar, S., Czapla-Myers, J., Mishra, N., Helder, D., 2014. Landsat-8 operational land imager radiometric calibration and stability. Remote Sens. 6, 12275–12308.

Mobley, C.D., 1999. Estimation of the remote-sensing reflectance from above-surface measurements. Appl. Opt. 38, 7442–7455.

Mobley, C.D., Sundman, L.K., 2008. Hydrolight 5, Ecolight 5 User Guide. Sequoia Scientific, Inc, Bellevue, p. 97.

Mueller, J.L., Fargion, G.S., McClain, C.R., Pegau, S., Zanefeld, J., Mitchell, B.G., Kahru, M., Wieland, J., Stramska, M., 2003. Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 4, Volume IV: Inherent Optical Properties: Instruments, Characterizations, Field Measurements and Data Analysis Protocols.

Pahlevan, N., Schott, J., 2013. Leveraging EO-1 to evaluate capability of new generation of Landsat sensors for coastal/inland water studies. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 6, 360–374.

Pahlevan, N., Schott, J.R., 2012. Characterizing the relative calibration of Landsat-7 (ETM+) visible bands with Terra (MODIS) over clear waters: The implications for monitoring water resources. Remote Sens. Environ. 125, 167–180.

Pahlevan, N., Lee, Z., Wei, J., Schaff, C., Schott, J., Berk, A., 2014. On-orbit radiometric characterization of OLI (Landsat-8) for applications in aquatic remote sensing. Remote Sens. Environ. 154, 272–284.

Pahlevan, N., Sarkar, S., Franz, B.A., 2016. Uncertainties in coastal ocean color products: impacts of spatial sampling. Remote Sens. Environ. 181, 14–26.

Shi, W., Wang, M., 2009. An assessment of the black ocean pixel assumption for MODIS SWIR bands. Remote Sens. Environ. 113, 1587–1597.

Stumpf, R.P., Arnone, R.A., Gould, J.R., Martinolich, P.M., Ransibrahmanakul, V., 2004. A partially coupled ocean-atmosphere model for retrieval of water-leaving radiance from SeaWiFS in coastal waters. In: Patt, F.S., et al. (Eds.), 2003: Algorithm Updates for the Fourth SeaWiFS Data Reprocessing. National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, MD.

Vanhellemont, Q., Ruddick, K., 2014. Turbid wakes associated with offshore wind turbines observed with Landsat 8. Remote Sens. Environ. 145, 105–115.

Vanhellemont, Q., Ruddick, K., 2015. Advantages of high quality SWIR bands for ocean colour processing: examples from Landsat-8. Remote Sens. Environ. 161, 89–106.

Vermote, E.F., El Saleous, N., Justice, C.O., Kaufman, Y.J., Privette, J.L., Remer, L., Roger, J.C., Tanr'e, D., 1997. Atmospheric correction of visible to middle-infrared EOS-MODIS data over land surfaces: background, operational algorithm and validation. Journal of Geophysical Research: Atmospheres 102, 17131–17141.

Wang, M., Gordon, H.R., 2002. Calibration of ocean color scanners: how much error is acceptable in the near infrared? Remote Sens. Environ. 82, 497–504.

Zibordi, G., Berthon, J.-F., Mélin, F., D'Alimonte, D., Kaitala, S., 2009. 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. 113, 2574–2591.