

Uncertainties in Coastal Ocean Color Products: Impacts of Spatial Sampling

Nima Pahlevan^{a,b}, Sudipta Sarkar^{a,b}, Bryan A. Franz^a

^aNASA Goddard Space Flight Center, 8800 Greenbelt Rd, Greenbelt, MD, 20771

^bScience Systems and Applications, Inc., 10210 Greenbelt Rd, Suite 600,
Lanham, MD 20706

ABSTRACT

With increasing demands for ocean color (OC) products with improved accuracy and well characterized, per-retrieval uncertainty budgets, it is vital to decompose overall estimated errors into their primary components. Amongst various contributing elements (e.g., instrument calibration, atmospheric correction, inversion algorithms) in the uncertainty of an OC observation, less attention has been paid to uncertainties associated with spatial sampling. In this paper, we simulate MODIS and VIIRS OC products from 30m resolution OC products derived from the Operational Land Imager (OLI) aboard Landsat-8, to examine impacts of spatial sampling on both cross-sensor product intercomparisons and in-situ validations of R_{rs} products in coastal waters. The simulations were carried out for OLI scenes “scanned” for one full orbital-repeat cycle of each ocean color satellite. While some view-angle dependent differences in simulated Aqua-MODIS and VIIRS were observed, the average uncertainties (absolute) in product intercomparisons (due to differences in spatial sampling) at regional scales are found to be 1.8%, 1.9%, 2.4%, 4.3%, 2.7%, 1.8%, and 4% for the $R_{rs}(443)$, $R_{rs}(482)$, $R_{rs}(561)$, $R_{rs}(655)$, [Chl_a], $K_d(482)$, and $b_{bp}(655)$ products, respectively. It is also found that, depending on in-water spatial variability and the sensor’s footprint size, the errors for an

27 in-situ validation location in coastal areas can reach as high as $\pm 18\%$. We conclude that
28 a) expected biases induced by the spatial sampling in product intercomparisons are
29 mitigated when products are averaged over at least $7km \times 7km$ windows, b) VIIRS
30 observations, with improved consistency in cross-track spatial sampling yields more
31 precise calibration/validation results than MODIS, and c) use of a single pixel centered
32 on in-situ coastal sites provides an optimal sampling size for validation efforts. These
33 findings will have implications for enhancing our understanding of uncertainties in ocean
34 color retrievals and for planning of future calibration/validation exercises.

35 **1. INTRODUCTION**

36 Following four decades of research and development, ocean color (OC) products
37 from spaceborne remote sensing instruments now play a critical role in the management
38 and monitoring of coastal ecosystems, which are under increased stress due to human
39 population growth in coastal areas and associated anthropogenic impacts such as
40 pollution and agricultural run-off (McGranahan et al. 2007; Nixon 1995; Vitousek et al.
41 1997; Vörösmarty et al. 2000). Repeatable, timely, and reliable OC products provide a
42 viable and efficient tool for monitoring of these coastal ecosystems. Coastal OC products
43 provide a synoptic view of coastal ecosystems at an instance of time and when
44 assimilated with coupled hydrodynamic-ecosystem coastal models (Allen et al. 2008;
45 Gohin et al. 2005; IOCCG 1997; Natvik and Evensen 2003; Ouillon et al. 2004) allow for
46 nowcasting and forecasting of environmental conditions to aid in hazard mitigation
47 efforts (e.g., occurrence of harmful algal blooms). Within the suite of OC products
48 distributed by NASA, the remote sensing reflectance (R_{rs} ; defined as the ratio of water-

49 leaving radiance and the total downwelling irradiance just above the surface) plays a
50 central role in determining the optical and biogeochemical properties of coastal oceans.

51 These observations and products, however, come from various satellite sensors with
52 known, partially known, or unknown uncertainties in their radiometric observations or
53 derived geophysical products. These sensors and their processing approaches are
54 managed by various space agencies (<http://www.ioccg.org>), and despite international
55 efforts to coordinate satellite ocean color programs (e.g., <http://ceos.org>), the derived
56 products can be expected to differ due to uncertainties associated with sensor design and
57 calibration and geophysical retrieval algorithms. Examples of satellite ocean color
58 sensors are the Sea-viewing Wide Field-of-view Sensor onboard Orbview-2 (McClain et
59 al. 2004; O'Reilly et al. 1998), the Moderate Resolution Imaging Spectroradiometer
60 (MODIS) onboard both Aqua and Terra platforms (Esaias et al. 1998), and the recently
61 launched Visible Infrared Imaging Radiometer Suite (VIIRS) onboard the Suomi
62 National Preparatory Partnership (SNPP). To ensure climate-quality products and
63 consistent data record, the OC products are validated against in-situ observations, and
64 individual satellite records are compared over their common mission overlapping periods.
65 Each of these comparison methods carries unique uncertainties.

66 The in-situ R_{rs} measurements are commonly made during research cruises or at
67 stationary field stations (Antoine et al. 2008a; Zibordi et al. 2009a; Zibordi et al. 2009b).
68 The ocean color component of the AErosol RObotic NETwork (AERONET-OC) is a
69 good example of an internationally coordinated automated network of radiometric
70 observations in coastal areas (Zibordi et al. 2009b). The sources of uncertainties in such
71 field validation efforts are attributed to a) the instrument calibration (Antoine et al.

72 2008a; Clark et al. 1997; Mueller et al. 2004), b) the post-processing techniques (Zibordi
73 et al. 2004), c) differences in the spectral sampling of the in-situ instruments and that of
74 the OC sensor (Wang 1999), and d) the spatial representativeness of the in-situ
75 observation (Mélin and Franz 2014).

76 The uncertainty in intercomparisons of OC-derived R_{rs} products, however,
77 corresponds to a) the sensors' radiometric performance on-orbit and the strategies
78 employed for vicarious calibration (Bailey et al. 2008; Franz et al. 2007; Zibordi et al.
79 2015) b) the atmospheric correction (Gordon 1997; Müller et al. 2015), c) differences in
80 spectral/spatial sampling, and d) the observation characteristics, i.e., angular
81 dependencies, including bidirectional reflectance distribution function (BRDF) effects
82 (Meister et al. 2012; Morel et al. 2002; Morel and Gentili 1996). This is further
83 complicated if a higher-level product intercomparison (e.g., chlorophyll-a, inherent
84 optical properties) is desired (Lee et al. 2010; Wang et al. 2005). The OC products
85 obtained from different sensors are often compared either over open oceans (Franz et al.
86 2005; Hu et al. 2013) or coastal waters (Barnes and Hu 2015; Hu and Le 2014; Ladner et
87 al. 2014; Mélin et al. 2011) to examine consistency amongst satellite observations and
88 detect any abnormal trending. In addition, various techniques have been proposed to
89 merge (fuse) different OC products (Maritorena et al. 2010; Maritorena and Siegel 2005;
90 Mélin et al. 2011; Mélin et al. 2009) to build comprehensive datasets less prone to spatial
91 and temporal gaps.

92 Among the sources of uncertainties for in-situ OC product validations (or calibration)
93 and cross-sensor intercomparisons, less attention has been paid to the characterization of
94 the impact of inherent in-water spatial variability and its relevance to sensors' spatial

95 sampling. Instead, efforts have been made to avoid or minimize the impact of spatial
96 sampling through various statistical analyses (Bailey and Werdell 2006; Mélin et al.
97 2007; Moore et al. 2015). One possible reason for the lack of prior research on the
98 uncertainties induced by differences in spatial sampling is the inability to decompose the
99 bulk uncertainty/bias into its components noted above. Bailey and Werdell (2006)
100 suggested performing a statistical test within 5×5-pixel boxes centered on the in-situ
101 matchups, where the coefficient of variation (CV; the ratio of standard deviation and
102 mean) of valid pixels is evaluated. Some other researchers have reported use of 3×3-
103 pixel boxes to analyze spatial variability at coastal sites (Zibordi et al. 2009a). Zibordi et
104 al. (2006) asserted that the error in the representativeness of in-situ measurements could
105 generally be treated as a random error, which can be minimized by increasing the sample
106 numbers and averaging over time. The statistical variability observed in this manner is a
107 combination of spatial variability and temporal variability, where the latter may be
108 associated with the instrument calibration or algorithm errors. A common uncertainty
109 goal for ocean color missions is to retrieve remote sensing reflectance with accuracies
110 better than 5% in the blue channels over the open ocean (Hooker et al. 1992).
111 Nevertheless, the reported biases for different coastal systems range from $\pm 2\%$ to $\pm 35\%$
112 in R_{rs} products (Moore et al. 2015; Zibordi et al. 2009a; Zibordi et al. 2012). Moore et al.
113 (2015) reported bias values (for SeaWiFS and Aqua MODIS) larger than those published
114 in Zibordi et al. (2012) over optically complex waters. The appreciable discrepancies in
115 the results were due to the differences in the quality-control of the in-situ matchups.
116 Zibordi et al. (2012) showed that the differences in R_{rs} products (satellite-derived versus
117 in-situ) increase with the increase in the solar zenith angles and aerosol optical thickness.

118 Furthermore, the Group on Earth Observations (GEO), the Committee on Earth
119 Observation Systems (CEOS) and its working groups, together with the International
120 Ocean Colour Coordinating Group (IOCCG) are leading efforts in providing traceable
121 measurements with associated uncertainties to the user community. For a full treatment
122 of the error budget, it is desired that the uncertainties be specified at the component level,
123 i.e., instrument, atmospheric correction, spectral/spatial sampling error, inversion
124 algorithm errors, etc. For instance, for in-situ validations, both in-situ instruments
125 (Antoine et al. 2008a; Hooker and Maritorena 2000) and the OC sensor should have
126 traceable, and meaningful total uncertainty budget (Chander et al. 2013).

127 The present research provides a full and unique analysis of the impact of sensors'
128 spatial sampling (and their footprint sizes) on uncertainties corresponding to a) product
129 intercomparisons at regional scales and b) in-situ validation practices at coastal ocean
130 sites (local scales). In order to isolate the uncertainties to a sensor's spatial sampling, we
131 follow a simulation approach. Our focus is to model Terra MODIS (MODIST), Aqua
132 MODIS (MODISA), and VIIRS OC products at regional and local scales using the 30m
133 OC products derived from the Operational Land Imager (OLI) onboard Landsat-8 (Irons
134 et al. 2012). To assess the impact of spatial sampling on satellite product
135 intercomparisons, the simulations are performed on a daily basis for 30 (globally
136 distributed) OLI scenes for a 16-day orbital repeat cycle of Aqua, Terra, and SNPP. Then,
137 simulated MODISA products are independently compared against a) simulated VIIRS
138 and b) simulated MODIST products. To simulate the impact on in-situ
139 calibration/validation efforts, the OC products (MODISA, MODIST, and VIIRS) are
140 simulated with more than 100 OLI scenes over eight different known in-situ stations

141 (covering various environmental conditions) for a similar orbit cycle as above. The OLI
142 observations surrounding the stations represent the “truth” and are employed to examine
143 the matchup quality against the simulated coarse resolution OC products. With this
144 approach, the OLI-derived OC products and their spectral contents form the basis of the
145 simulations.

146 The following section (Section 2) elaborates on the necessary steps pursued to
147 simulate OC products from relatively fine-resolution OLI products. Section 3 contains
148 results associated with uncertainties in product intercomparisons and in-situ validations,
149 which is followed by the discussions in Section 4. In the conclusion section, we make
150 recommendations for improving uncertainty assessment of OC products and provide
151 guidelines for future developments.

152 **2. METHODS**

153 The VIIRS instrument aboard SNPP was launched in 2011 to pursue the heritage
154 global measurements made by MODIS on Aqua and Terra. Although VIIRS shares
155 similar operational and climate science requirements with MODIS, it has notable
156 differences in design, calibration, and onboard processing (Baker 2011; Cao et al. 2013).
157 MODIS is a rotating mirror, which sweeps one scan across track while projecting it onto
158 10 square detectors (1km ocean bands). In contrast, VIIRS is a rotating telescope with a
159 half-angle mirror, which projects one sweep of scan onto 16 rectangular detectors, i.e.,
160 750m Moderate (M)-bands. The Aqua and SNPP missions are both in ascending orbits
161 crossing the equator around 13:30 local time, while Terra (similar to Landsat-8) is in a
162 morning orbit with nominal 10:30 equatorial crossing time.

163 The projected area of a detector (square or a rectangle) on the ground represents an
164 ideal spatial sample. However, the extent within which an instrument views (and
165 samples) a spot on the ground is commonly larger than the projected detector area due to
166 imperfect optics, jitter, electronics, and post-processing (Holst 2008; Schott 2007). The
167 spatial sampling function is recognized as the point spread function (PSF). The overall
168 PSF of an imaging system is the convolution of the above-noted, component-specific
169 PSFs. The overall spatial performance is specified by the $PSF(x, y)$, which is the product
170 of the measured line spread functions along-track (y) and along-scan (x). For a pair of
171 MODISA-VIIRS or MODIST-MODISA daily observations/products, each instrument
172 views a target of interest from a different vantage point, which yields varying projected
173 sampling area, i.e., the ground sample distance (GSD). The GSD is commonly defined at
174 the Full-Width-Half-Maximum (FWHM) of the PSF (Schott 2007). The size and the
175 shape of a footprint (one spatial sample) of an optical imager is a function of the orbit, the
176 viewing geometry (line-of-sight), and the instrument effective spatial performance. The
177 spatial performance is commonly characterized prior to launch (Lin et al. 2013) and is
178 monitored commonly in the frequency domain throughout the mission either through
179 onboard measurements (Xiong et al. 2006) or through observations of reference targets
180 (moon, uniform edges, etc.) (Xiong et al. 2004).

181 When comparing two OC products at the swath level, the differences due to spatial
182 sampling largely depend on the spatial heterogeneity of the site, i.e., pixel/scene content.
183 Similarly, the fidelity of an in-situ matchup is highly affected by the spatial uniformity
184 around the site. In the following, we describe the simulation procedure implemented for
185 different OLI scenes representing distinct coastal water conditions to obtain reliable

186 statistics on the extent of the impact of spatial sampling on coastal OC products. A
187 thorough description of the procedure is first given for the regional product
188 intercomparisons followed by a short explanation of the simulations implemented at in-
189 situ sites.

190 **2.1. Regional Simulations**

191 **2.1.1. Retrieval of Rrs Products from OLI**

192 The OLI onboard Landsat 8 has been shown to considerably outperform previous
193 generation of Landsat sensors when studying aquatic systems (Gerace et al. 2013;
194 Pahlevan and Schott 2013). This is because of its improved SNR, the 12-bit radiometric
195 resolution, and the addition of the new 443-nm channel (Pahlevan et al. 2014;
196 Vanhellemont and Ruddick 2014). The OLI Level-1 data products (TOA reflectance)
197 were obtained from the USGS web portal (<http://glovis.usgs.gov/>). In order to provide
198 30m OC products, the OLI data were processed using the SeaDAS package (Franz et al.
199 2015). The calibration gains obtained via cross-calibration at the top-of-atmosphere
200 (TOA) reflectance domain were applied to the original OLI observations (Pahlevan et al.
201 2014). However, it should be emphasized that the absolute product accuracy of the OLI
202 products are not critical in our study. What is important is to obtain relative in-water
203 spatial features present in various environmental conditions. Owing to the non-zero
204 reflectance in the near-infrared (NIR) portion of the spectrum, we used two different
205 atmospheric corrections (ACO) and chose the OLI products with more valid and
206 smoother products based on analyzing the standard deviations for all valid pixels. The
207 ACO was implemented using a) the combination of the two short-wave infrared (SWIR)
208 bands (Vanhellemont and Ruddick 2015) and b) the combination of NIR-SWIR (865nm-
209 2201nm) with the iterations on the NIR channel to account for non-zero reflectance

210 (Bailey et al. 2010). In these techniques, the aerosol model selection is carried out using
211 Rayleigh-corrected SWIR or NIR-SWIR reflectances. Prior to the aerosol-model
212 selection, the input Rayleigh-corrected reflectances (within the NIR and SWIR channels)
213 were smoothed using a 5×5 averaging filter to increase effective SNR (Gerace et al.
214 2013). For the product intercomparison, 25 OLI cloud-free scenes (globally distributed
215 as shown in Fig. 1) were processed. The preliminary products over high-latitude areas
216 and turbid atmospheric conditions showed noisy R_{rs} retrievals. This is attributed to the
217 noisy nature of the Rayleigh-corrected reflectances (even after filtering) leading to error
218 in aerosol model selection (Franz et al. 2015). Therefore, the OLI-derived R_{rs} products
219 were further passed through a 3×3 median filter to remove residual noise. This, however,
220 reduces the inherent resolution of OLI data. At this point, the OLI-derived R_{rs} products
221 are supplied to the simulation process as described below. The retrieval of secondary
222 products, including chlorophyll-a ([Chla],(O'Reilly et al. 1998)), diffuse attenuation of
223 downwelling irradiance at 490nm ($K_d(490)$,(Mueller 2000)), and particulate
224 backscattering at 655nm ($b_{bp}(655)$, (Lee et al. 2012)) will be derived (Section 2.1.3) from
225 the simulated coarse-resolution R_{rs} products (Lee et al. 2012).

226 **2.1.2. Implementation**

227 To simulate MODIS (MODIST and MODISA) and VIIRS R_{rs} products using OLI-
228 derived R_{rs} products, the spatial sampling function (PSF) and the (per-observation) line-
229 of-sight information of the sensors are needed.

230 **2.1.2.1. PSF Modeling**

231 As described, the PSF is essentially a 2D distribution function that weighs the
 232 incoming surface-reflected (or emitted) radiance. For the purpose of this study, the Line
 233 Spread Functions (LSFs) measured during the pre-launch characterization of MODIS and
 234 VIIRS were employed (Barnes et al. 1998; Lin et al. 2013). Note that, due to inherent

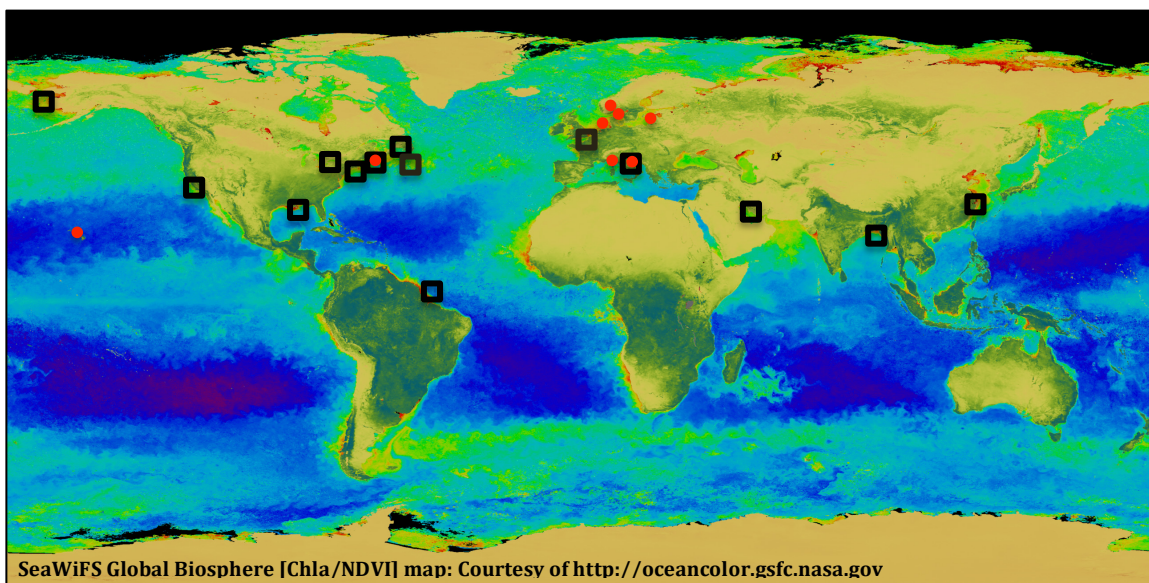


Fig. 1. The sites (indicated by boxes) where OLI scenes are processed to Level-2 OC products for intercomparisons of products at regional scales (Section 2.1). For these locations, the MODISA, MODIST, and VIIRS OC products were simulated according to their corresponding viewing geometries from DOY=1 to DOY=16 in 2015. The red dots denote the locations where simulated matchups were produced and analyzed (Section 2.2).

235

236

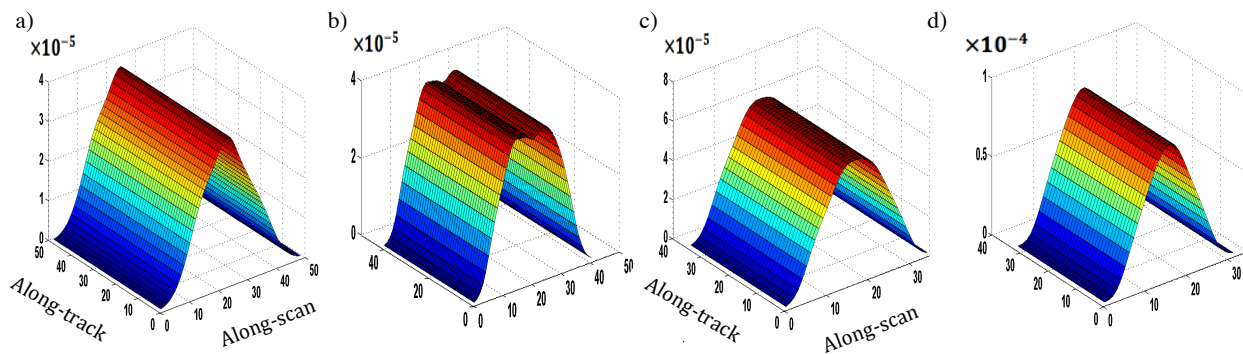


Fig. 2. The band/detector average point spread functions (PSFs) shown for a) MODIS, b) VIIRS near-nadir zone (PSF-V3g), c) VIIRS mid-range zone (PSF-V2g), and d) VIIRS edge-of-scan zone (PSF-V1g). Note that it is assumed that the along-track response is rectangular (ideal response). All the functions have unit-area. The coordinates are in arbitrary pixel units.

237 similarities in shape, only one set of LSF was utilized for MODISA and MODIST. The
238 LSFs had been characterized for each detector in each band. In general, the MODIS
239 along-scan LSFs, i.e., $LSF(x)$, can be approximated as triangular functions (Wolfe et al.
240 1998).

241 For VIIRS, however, the onboard aggregation yields different along-scan LSFs for each
242 aggregation zone (Cao et al. 2014). VIIRS has been designed with rectangular detectors
243 (with the smaller dimension along-scan) to allow for aggregating (averaging) multiple
244 samples. At near nadir scan angles ($\alpha < 31.72^\circ$) every three pixels are aggregated, in
245 mid-range angles ($31.72^\circ < \alpha < 44.86^\circ$) every two pixels are aggregated, and at the
246 edge of the scan ($\alpha > 44.86^\circ$) no averaging is implemented. This scheme provides near-
247 uniform sampling of the Earth surface in the along-scan (x) direction and reduces the
248 “pixel growth” at the edge of the scan to less than $2.2\times$ the linear dimension of the nadir
249 pixels (Fig. 8 in Cao et al. (2014)). The MODIS along-scan samples at the edge of the
250 scan, however, are nearly five times larger than those at nadir. On the other hand, at the
251 edge of the scan, the size of the MODIS and VIIRS along-track footprints are very
252 similar, i.e., \sim twice larger than the footprint sizes at nadir-viewing angles (Wolfe et al.
253 2013). For an accurate simulation of the effects of spatial sampling on OC products, the
254 four different $LSF(x)$ (Fig. 2) are modeled to represent the spatial sampling of MODIS
255 and VIIRS. The along-scan LSFs, as a result, are modeled as the summation of multiple
256 Gaussian functions, which provided the best fit to the measured LSFs:

257

$$258 \quad LSF(x) = \sum_{i=1}^5 a_i \exp\left(\frac{x-b_i}{c_i}\right) \quad (1)$$

259 where b and c are the mean and the standard deviation for each Gaussian function,
260 respectively, and $i = 1, \dots, 5$ represents indices attributed to the Gaussian functions. The
261 $PSF(x, y)$ is a separable function and can be constructed by assuming a rectangular LSF
262 along-track (y) as below (Lin et al. 2013):

263

$$264 \quad PSF(x, y) = LSF(x) \times LSF(y) \quad (2)$$

265

266 where $LSF(x)$ and $LSF(y)$ are the along-scan and along-track spatial responses of
267 MODIS and VIIRS, respectively. Fig. 2 shows the unit-area PSFs employed to simulate
268 MODIS and VIIRS scenes. Note that the discrepancies amongst band- or detector-
269 dependent LSFs are assumed negligible and average sampling functions of 1000m ocean
270 bands (MODIS) and 750m M-bands (VIIRS) are used. Also, in this study, the optical
271 scattering (leakage) outside of the instantaneous field-of-view of a detector has been
272 ignored (Meister and McClain 2010). The PSFs constructed using Eq.2 are further scaled
273 and rotated per observation (pixel) according to the line-of-sight geometries.

274 **2.1.2.2. Line-of-sight Geometry**

275 In addition to the PSF modeling, the corresponding line-of-sight information is
276 extracted from the geolocation products. The MODIS geolocation products (MYD03 and
277 MOD03 corresponding to MODISA and MODIST, respectively) were obtained from the
278 NASA's Level-1 and Atmosphere Archive and Distribution System (LAADS) while
279 those of VIIRS (GMODO) were downloaded from the NOAA's Comprehensive Large
280 Array-data Stewardship System (CLASS). The geolocation products allow for
281 reconstructing per-pixel (-observation) line-of-sight geometries. The line-of-sight

282 information includes per-pixel geographic latitude and longitude, the view zenith angle
 283 (VZA), the view azimuth angle (VAA) relative to the North, and the satellite range (R),
 284 which is defined as the distance between the sensor and ground (Baker 2011). While the
 285 standard MODIS geolocation products are available at 1km grid resolution, the inherent
 286 VIIRS geolocation grids are at 750m.

287 The 30m OLI R_{rs} products are sampled with MODIS and VIIRS PSFs and their
 288 respective line-of-sight geometries on a daily basis (for 16 days). The supports of the
 289 PSFs (Fig. 2) are essentially scaled and rotated in along-track and along-scan directions
 290 given per-pixel VZA, VAA, and R (Schowengerdt 1997). The PSFs are convolved with
 291 OLI R_{rs} products in a discrete form as follows:

292

$$293 \quad S_n^k = \sum_m \sum_i \sum_j PSF_{(VZA_m, VAA_m)}(i, j) * R_{rs}^k(i, j) \quad (3)$$

294

295 where S_n^k is the simulated VIIRS or MODIS sub-granule swath for the n th day
 296 ($1 < n < 16$), k is the channel number, $R_{rs}^k(i, j)$ indicates the remote sensing reflectance
 297 attributed to a pixel location (i, j) on the OLI grid for band k , and $PSF_{(VZA_m, VAA_m)}$ is the
 298 PSF implemented for a given VZA_m and VAA_m , where m is an observation (pixel) index
 299 associated with MODIS or VIIRS geolocation products (Wolfe et al. 2002). The four OLI
 300 visible channels (i.e., $k = 1, 2, 3, 4$ which represent 443, 482, 562, and 655nm) are
 301 processed during the simulation (Pahlevan and Schott 2013).

302 The MODIS footprint size was computed using range, VZA, and MODIS effective
 303 focal length (380mm) for the along-scan and along-track directions (Schowengerdt 1997).
 304 On the other hand, due to the onboard aggregation scheme of VIIRS, there is no such a

305 direct relationship between VZA and footprint size. Therefore, the footprint size of
306 VIIRS in along-scan and along-track was calculated by quadratic modeling of the scan-
307 angle-versus-foot-print size relationship (Cao et al. 2014). This was done by specifying
308 observational aggregation zone (for VIIRS) and recalculating the scan angle from a given
309 VZA as below

310

$$311 \quad \alpha = \sin^{-1}(\sin(180 - VZA) R_e / (R_e + H)) \quad (4)$$

312

313 where R_e is the Earth radius for a given latitude and H stands for the VIIRS orbit altitude
314 at nadir. To provide concrete examples, the along-scan size of a MODIS and VIIRS
315 observation cell at nadir is approximately 2200m and 1160m, respectively, which are
316 nearly 2.2 and 1.5 times larger than the projected size of the detectors on the ground (e.g.,
317 ~ 750 m for VIIRS). With this approach, simulated VIIRS or MODIS granules (5-min
318 MODIS granules in case of MODIS) are created. Since only the OLI effective area is
319 filled with valid R_{rs} values in the granules, we refer to the resulting swath as sub-granules.
320 Fig. 3 shows an example of an OLI scene and the simulated MODISA, MODIST, and
321 VIIRS. In this example, an OLI-derived chlorophyll-a product is “scanned” given
322 MODISA, MODIST, and VIIRS viewing geometries. Note the differences in the average
323 VZAs for the three simulated ocean color scenes.

324 **2.1.2.3. Time Period**

325 While the OLI R_{rs} products come from various coastal systems during different
326 times/seasons (see Appendix), simulations of MODIS and VIIRS granules are performed

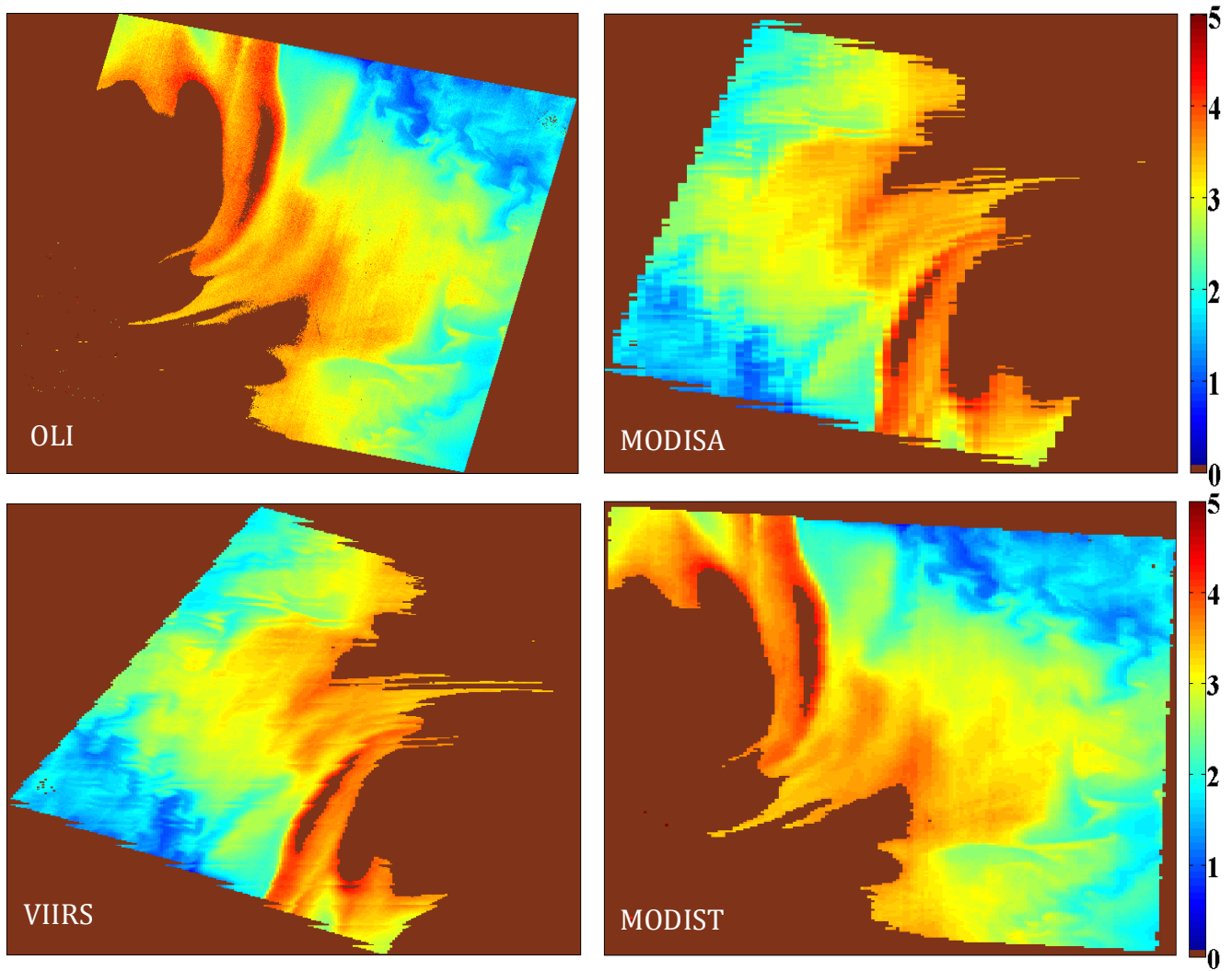


Fig. 3. Examples of simulated MODISA, VIIRS, and MODIST swaths for [Chla] (mg/m^3) fields derived from the OLI-derived [Chla] products over the Yangtze River mouth, East coast of China. The average view zenith angles (VZAs) are 58.18° , 51.5° , and 21.5° for MODISA, VIIRS, and MODIST swaths, respectively. The features are reproduced at different viewing conditions. Note that OLI and MODIST are in descending orbits. Also, the extremely turbid waters over the basin are masked.

327

328

329

330

331

332

333

334 for an arbitrary orbital repeat cycle of Aqua/Terra/SNPP starting from January 1st
335 (DOY=1) to January 16th (DOY=16) in 2015. This is to maintain the orbit geometries
336 similar for all the simulations, which allows for consistent product intercomparisons
337 across all simulated products. Note that only viewing geometries of ocean color sensors
338 are extracted over this period of time and the OLI scenes correspond to different days of
339 year and environmental conditions. It should be noted that daily intercomparisons are
340 carried out only for days that both granules fully cover an OLI scene across track. This is
341 to ensure that there is a one-to-one comparison between simulated MODIS and VIIRS R_{rs}
342 products. Note that for higher latitude OLI scenes (e.g., Alaska scenes), there are more
343 than 16 pairs of MODISA-VIIRS (or MODIST-MODISA) intercomparisons.

344 **2.1.3. Retrieval of Secondary OC Products**

345 After MODISA, MODIST, and VIIRS R_{rs} swath-level products are simulated, [Chla],
346 $K_d(490)$, and $b_{bp}(655)$ are computed using standard algorithms available in SeaDAS
347 (<http://seadas.gsfc.nasa.gov>). For [Chla], we used the OC3 algorithm (O'Reilly et al.
348 1998; O'Reilly et al. 2000), which was recently adapted for the spectral bands of OLI
349 (Franz et al. 2015). Similarly, for $K_d(490)$ we used the KD2 algorithm (Mueller 2000)
350 with the OLI-specific tuning provided in SeaDAS. And finally, for $b_{bp}(655)$ we used the
351 QAA inherent optical properties algorithm of Lee et al. (2002). It should be further
352 emphasized that the absolute accuracy of retrievals are not critical in our study. What is
353 important is capturing varying coastal-ocean spatial features (swirls, river plumes, eddies,
354 etc.) of different sizes through OLI observations and the derived products.

355 **2.1.4. Gridding**

356 The swath-level R_{rs} products are commonly distorted geographically (Fig. 3)
357 rendering it impossible to make pixel-to-pixel comparison. For a particular day of Terra
358 and Aqua's overpasses (e.g., DOY=4), consider arbitrary MODISA and MODIST
359 observations, which view an OLI scene at two significantly different viewing angles. The
360 per-observation inter-comparisons of the two swaths are hindered by significantly
361 different levels of distortions for the pair of observation (e.g., ~6km MODISA pixels
362 against ~1km MODIST pixels). For this reason, the simulated sub-granules (S_n^k) are
363 gridded to 1km sinusoidal grid cells (S_n^{kG}) (Wolfe et al. 1998) using the nearest-neighbor
364 technique (Campbell et al. 1996). The 1km grid was chosen to closely match the near-
365 nadir nominal resolution of MODIS and VIIRS over spatially variable coastal waters.
366 The gridding allows for a uniform cell-by-cell comparison.

367 **2.2. Local Simulations**

368 With the simulation routine in place, it is also possible to provide insights into how
369 well a discrete in-situ sample taken during a field campaign or at a field station represents
370 an observation made by MODISA, MODIST, or VIIRS (Mélin et al. 2007; Moore et al.
371 2015). To do so, OLI pixels were used to represent a sample measurement in the field or
372 at a site. These locations can well be arbitrary, however, it was decided to carry out this
373 analysis at known field sites where routine radiometric observations are made for OC
374 calibration/validation purposes. We chose AERONET-OC stations (Zibordi et al. 2009b)
375 located near Venice (Venise), Martha's Vineyard (MVCO), Gustav Dalén (Gustav),
376 Helsinki Light House (Helsinki), Zeebrugge, and Palgrunden (Fig. 1) to represent
377 spatially variable waters at the proximity of coastal waters (<http://aeronet.gsfc.nasa.gov>).
378 We further used the locations at the Marine Optical Buoy (MOBY) (Clark et al. 1997)
379 and the BOUSSOLE site (Antoine et al. 2008b) to assess spatial variability in clear

380 waters. The latter two sites are buoys where in-situ radiometric observations are made to
381 calibrate TOA OC observations. It is stressed that we do not incorporate the actual
382 radiometric measurements at these sites. All the available cloud-/glint-free OLI scenes
383 over the selected sites (totaling 120) were obtained from the USGS database. Similar to
384 the previous section, the field sites were observed (scanned) for one orbit cycle (DOY=1
385 through DOY=16) of Aqua, Terra, and SNPP. Therefore, the 16-day simulated products
386 (locally) are computed for each OLI sample scene (Section 2.1.2.3).

387 For processing the OLI TOA products, although best-practice atmospheric correction
388 was employed (similar to Section 2.1.1; a combination of NIR-SWIR), there remains a
389 small percentage of OLI pixels flagged for poor quality (due to ship wakes, significant
390 resuspension events, etc.). Like the previous section, the OLI-derived OC products were
391 processed with a 3×3 median filter prior to implementing the simulations (Section 2.1.2).
392 The simulation was performed for one pixel, i.e., 1×1 window, as well as a 3×3 window,
393 and a 5×5 window of MODISA, MODIST, and VIIRS pixels surrounding the site
394 location to scrutinize in-water spatial variability at observation scales. Following multiple
395 experiments, it was decided to discard simulation results (OC pixels) for which more than
396 5% of the OLI pixels (falling under one ocean color footprint) are flagged. The major
397 reasons for pixel-flagging in case of our cloud-free OLI scenes were failure of the
398 atmospheric correction and the presence of land boundaries or marine vessels). When this
399 criterion is passed, the remaining flagged pixels (< 5% of total) are ignored and not
400 included in further statistical analysis. Moreover, the simulation ocean color footprints
401 were discarded if CV calculated for a window exceeded a threshold of 0.15 (Bailey and
402 Werdell 2006). To further minimize impacts of OLI residual scene noise (i.e., striping),

403 the same criterion was tested for the OLI pixels falling under the window under
 404 investigation. To simulate the in-situ radiometric observation at the site, a window of size
 405 5×5 (centered at the site location) was used to extract OLI-derived pixels totaling 16.
 406 This is to , exclude the center 3×3 elements to avoid adjacency effects (Storey et al.
 407 2014). The median value computed from the 16 pixels was found to best represent the
 408 measurement at the site. To emulate calibration/validation activities at these sites, the
 409 gridding scheme (Section 2.1.4) was excluded in this section. Therefore, comparisons
 410 were made at swath-level observations.

411 **2.3. Metrics**

412 To gauge how two sets of products (derived from sensors A and B) compare, multiple
 413 metrics are defined. The per-cell (i) percent difference is defined as:

$$414 \quad RPD^i = (R_A^i - R_B^i) / [(R_A^i + R_B^i) / 2] \quad [\times 100] \quad (5)$$

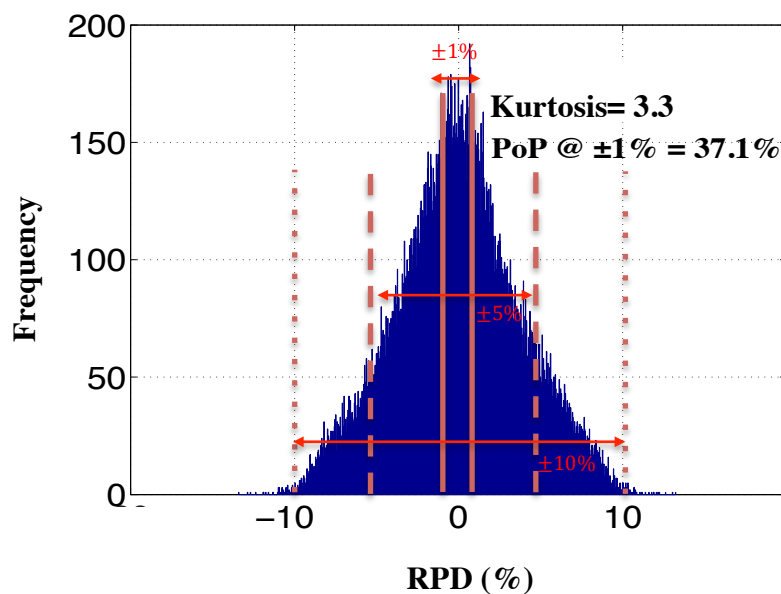
415
 416 where R_A^i and R_B^i stand for per-grid cell (i) simulated products for products derived from
 417 sensors A and B. In a slightly different manner, the absolute relative percent difference
 418 can be calculated

$$419 \quad ARPD^i = |R_A^i - R_B^i| / [(R_A^i + R_B^i) / 2] \quad [\times 100] \quad (6)$$

420 The median and mean values of RPD and ARPD over an entire grid can be used to
 421 specify the overall discrepancies for a product pair (Table 1). To further get insights into
 422 the differences in an absolute sense (product units), we also provide the root-mean-
 423 squared difference for N grid cells computed as below

$$424 \quad RMSD = \sqrt{\frac{\sum_{i=1}^N (R_A^i - R_B^i)^2}{N}} \quad (7)$$

426 Also, histograms of the RPD sub-granules (Eq. 5) provide valuable statistics on a per-
 427 grid cell basis. Based on a histogram, multiple discrepancy metrics can be defined. In
 428 particular, we use the 1st momentum of the histogram to refer to the mean bias. We also
 429 define the Percent of Pixels (PoP) within the $\pm z\%$ range of the histogram mean. The PoP
 430 (%) is a metric that specifies the percentage of pixels that fall within the $\pm z\%$ of the
 431 mean of the histogram. For instance, when $z = 1\%$, the percentage of grid cells
 432 exhibiting differences less than 1%, i.e., $-1\% < RPD^i < +1\%$, is expressed. This is
 433 represented by PoP @ $z = \pm 1\%$. Fig. 4 illustrates a sample RPD histogram computed
 434 from simulated VIIRS and MODISA sub-granule R_{rs} products with z ranging within
 435 $\pm 10\%$. For a given pair of simulated gridded products, the narrower the histogram the
 436 more homogenous the site is. To eliminate outliers from our analysis, we restrict RPD
 437 within the $\pm 60\%$ range. These outliers commonly occur at land-water interfaces, where



446 Fig. 4. A typical histogram derived from product intercomparisons.
 447 Amongst different histogram-derived parameters, the Percent of Pixels
 (PoP) is shown above for various levels of relative percentage
 difference, i.e., RPD= 1%, 5%, and 10%. Also, histogram kurtosis
 indicates how spread is the distribution around the histogram mean.
 The narrower the histogram, the more consistent a pair of products are.

448 discrepancies in the footprint sizes result in seemingly large differences. The other
 449 histogram-derived metric is kurtosis (K). Kurtosis is the fourth moment of the
 450 distributions, which is a descriptor of the shape of the distribution and explains how
 451 narrow a distribution is relative to that of a normal. For a normal distribution, the kurtosis
 452 value is ~ 3 . The narrower the shape of a histogram, the larger the kurtosis value is. For
 453 example, when $K = 10$, a product pair is found more consistent than if K were equal to
 454 3. The overall analysis of these metrics helps understand how the combination of
 455 imaging geometry and spatial sampling contribute to the differences in ocean color
 456 products at regional scales.

457 For the OC simulations over field stations, we used a slightly different metric termed
 458 the percent difference (PD) for the station (s) as below

$$459 \quad PD^s = (R_X^s - R_{OLI}^s)/R_{OLI}^s \quad [\times 100] \quad (8)$$

460 where R_X^s represents the median (within a 3×3 or 5×5 box) simulated observations (for
 461 the sensor X) and R_{OLI}^s is the median OLI OC products within a 5×5 box centered over
 462 the field station s . Note that we excluded the 3×3 pixels at the center, which effectively
 463 allowed 16 pixels for calculating the median value. The PD can be utilized to obtain an
 464 estimate of the temporal mean bias at a field station (e.g., MOBY). The absolute
 465 percentage difference can also be used to provide further insights into matchup analysis:

$$466 \quad APD^s = |R_X^s - R_{OLI}^s|/R_{OLI}^s \quad [\times 100] \quad (9)$$

467 The metric APD is used to explain how close a simulated in-situ R_{rs} is to that derived
 468 from satellite observations. Note that for simulating matchups, we only evaluate
 469 differences in R_{rs} .

470 3. RESULTS

471 The uncertainties induced by spatial sampling are presented for two data quality
472 assessment approaches: a) intercomparisons of products derived from two different OC
473 sensors and b) products compared against “in-situ measurements” at eight different sites.
474 While the former provides insights into the discrepancies in products at regional scales
475 (Landsat scene size; $185km \times 185km$), the latter is an assessment of spatial
476 representativeness at the selected field sites.

477 **3.1. Regional Assessment**

478 **3.1.1. Product Consistency**

479 In this section, the kurtosis value is used to describe the overall discrepancies between
480 product pairs. Fig. 5 illustrates the kurtosis values (K) as a function of differences in the
481 mean view zenith angles (\overline{VZA}) for the three selected sites. The mean view zenith angle is
482 computed by taking the average of VZAs with which the OC sensor “views” (scans) an
483 OLI scene. The differences in \overline{VZA} , i.e., $\Delta VZA = \overline{VZA}_V - \overline{VZA}_{MA}$ and $\Delta VZA =$
484 $\overline{VZA}_{MT} - \overline{VZA}_{MA}$, are binned into 10° intervals. The subscripts MA, MT, and V
485 correspond to MODISA, MODIST, and VIIRS, respectively. For each OLI scene, at least
486 12 pairs of simulations for MODISA-VIIRS and MODIST-MODISA are incorporated.
487 The error bars denote one-standard deviation for the data points (associated with mean K)
488 situated within a ΔVZA bin. The differences in angles (ΔVZA) range from -60° to $+60^\circ$
489 on the x-axes. Figs. 5a-c correspond to VIIRS-MODISA intercomparisons of R_{rs} , [Chla],
490 $K_d(482)$, and $b_{bp}(655)$ products. It is clearly seen that the interconsistency between all
491 products improves towards positive $\Delta VZAs$, i.e., narrower histograms on the right side of
492 the plots. On average, the products are more consistent by a factor of two when the edge-
493 of-the-scan VIIRS products are compared against near-nadir products of MODISA.

494 This is attributed to larger along-scan footprint sizes of VIIRS at the edge of the scan
 495 ($\sim 2.7\text{km}$) against the 2.2km along-scan MODISA footprint sizes at nadir-viewing
 496 geometries. On the other hand, the discrepancies between MODIST and MODISA
 497
 498

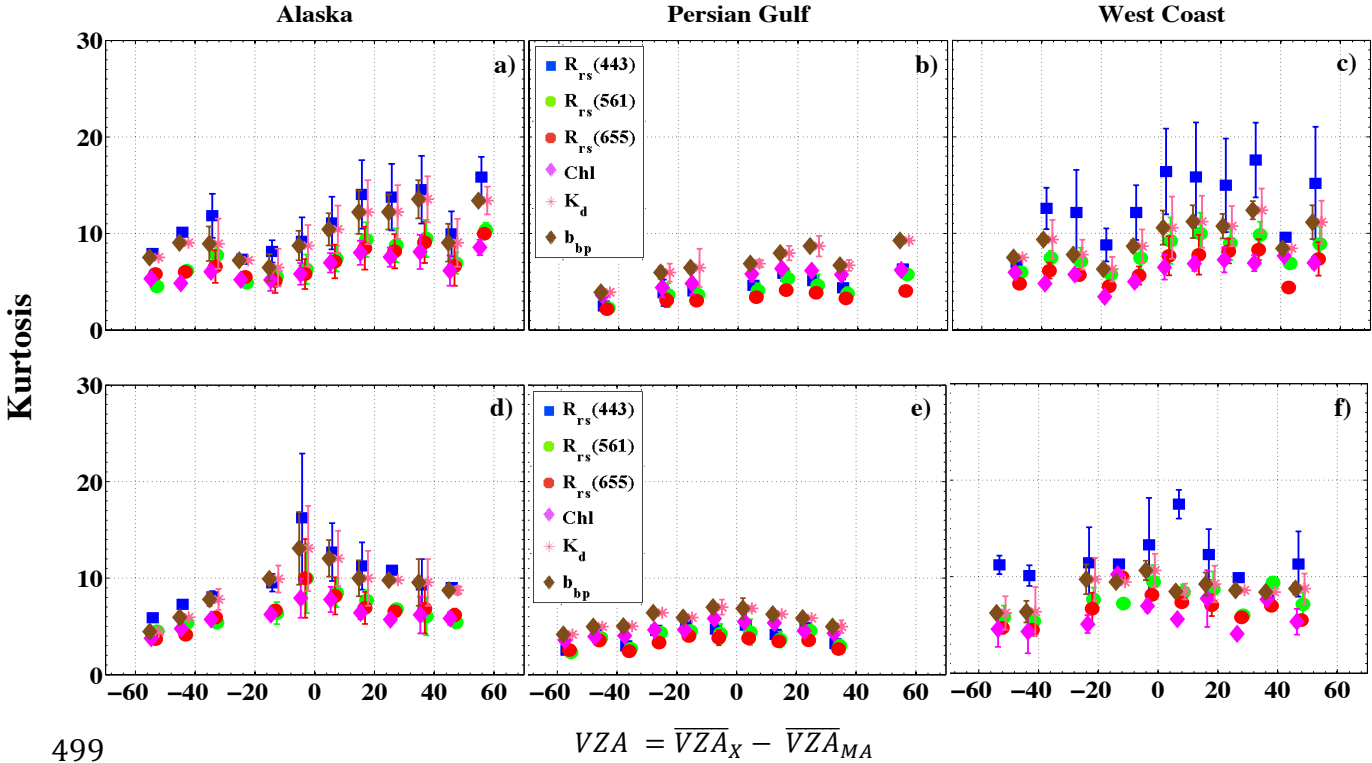


Fig. 5. The kurtosis values (derived from the RPD histograms) are shown as a function of differences in mean VZAs (\overline{VZA}). The subscript X denotes VIIRS (V) or MODIST (MT). The top row shows the asymmetric trends when MODISA and VIIRS products are compared. The x-axes denote $\Delta VZA = \overline{VZA}_V - \overline{VZA}_{MA}$. This trend implies that the products are more consistent when $\overline{VZA}_V > \overline{VZA}_{MA}$, which yield similar footprint sizes for the two observations. The bottom row corresponds to the intercomparisons of MODISA and MODIST products as a function of $\Delta VZA = \overline{VZA}_{MT} - \overline{VZA}_{MA}$. Although MODISA and MODIST are in afternoon and morning orbits, there is no particular trends found when comparing the associated products at different VZAs.

500
 501
 502
 503
 504

505 products exhibit symmetric trends around $\Delta VZA = 0$ (Fig. 5d-f). It is, thus, inferred that
506 although Aqua and Terra are in different morning and afternoon orbits, due to their
507 similarities in spatial sampling and footprint size, there are no distinct trends observed on
508 either side of the plots. The K values also provide clues on the in-scene spatial variability
509 at each site. For instance, the K values for the Persian Gulf site are, on average, smaller
510 than those for the Alaska and the Monterey Bay sites (indicating its larger inherent in-
511 water variability). While K values (and also PoP not shown here) show evidence for the
512 angular dependency of product interconsistency at regional scales, the mean RPD (same
513 plots as in Fig. 5 but for mean RPD not shown here for brevity) does not evidently exhibit
514 angular dependency. This implies that the use of representative statistics (e.g., mean RPD
515 and ARPD) computed at regional scales can minimize the overall product inconsistencies
516 (as opposed to pixel-by-pixel comparisons captured by K and PoP). Table 1 contains the
517 overall statistics calculated for all sites. As expected, for all the products compared in this
518 study, the mean RPD values converge to zero validating that overall intercomparisons at
519 regional scales ($185km \times 185km$ in this study) are immune from the effects of
520 differences in spatial sampling. Note that the RPD ranges within the $\pm 60\%$ with
521 histograms resembling the normal distribution, that is, the differences can reach up to
522 60%. The comparisons in the ARPD domain (absolute relative percent differences
523 computed for all scenes) show percent differences falling within the 1% to 4.5% range.
524 This varies depending on whether the mean or median metrics are considered. We refer
525 to the mean ARPD values (averaged for the two sets of intercomparisons), as the
526

527 Table 1. The overall product intercomparisons statistics, i.e., median (med), mean, and RMSE, averaged
 528 and tabulated for all sites.
 529

		$R_{rs}(443)$	$R_{rs}(482)$	$R_{rs}(561)$	$R_{rs}(655)$	[Chla]	$K_d(482)$	$B_{bp}(655)$
VIIRS Vs. MODISA	Med (RPD) [%]	0.0035	0.0029	0.0005	0.0004	-0.0045	-0.0008	-0.0031
	Mean (RPD)[%]	0.0133	0.0049	-0.0060	-0.0291	-0.0100	-0.0014	-0.0164
	Med (ARPD)[%]	0.7802	0.8114	1.0604	1.5287	1.5814	0.9795	1.9116
	Mean (ARPD) [%]	1.5870	1.6899	2.1482	4.0010	2.8950	1.9385	4.2637
	RMSE	0.0004	0.0004	0.0004	0.0002	0.1423	0.0124	0.0022
MODIST Vs. MODISA	Med (RPD) [%]	0.0009	0.0007	-0.0001	-0.0043	-0.0141	-0.0032	0.0000
	Mean (RPD)[%]	-0.0052	-0.0021	-0.0096	-0.0055	-0.0431	-0.0188	-0.0273
	Med (ARPD)[%]	0.8660	0.8848	1.2238	2.2307	1.3995	0.7839	1.4984
	Mean (ARPD) [%]	2.0921	2.0638	2.6193	4.6604	2.6351	1.6529	3.7957
	RMSE	0.0005	0.0005	0.0005	0.0003	0.1351	0.0123	0.0022

530

531 *intercomparison uncertainty* induced by differences in the spatial sampling that one
 532 should anticipate for similar practices in coastal waters. The values computed as 1.8%,
 533 1.9%, 2.4%, 4.3%, 2.7%, 1.8%, and 4% for the $R_{rs}(443)$, $R_{rs}(482)$, $R_{rs}(561)$, $R_{rs}(655)$,
 534 [Chla], $K_d(482)$, and $b_{bp}(655)$ products, respectively (Table 1). The RMSD values
 535 indicate that, overall, the differences in R_{rs} and [Chla] products are $< 0.0005 sr^{-1}$ and
 536 $< 0.15 mg/m^3$, respectively.

537 3.1.2. Uncertainty Thresholds

538 As noted, the histograms of the RPD grids can be used to interpret the inconsistencies
 539 between a product pair. Fig. 6 shows how the percentage of pixels (PoP) lying within a
 540 threshold ($\pm z\%$) can change. The results are shown for six sample sites. These curves are
 541 the mean values derived from all the daily intercomparisons within one MODISA-VIIRS
 542 orbit cycle (MODIST-MODISA comparisons are not shown here). Each marker indicates
 543 the percentage of pixels falling within a relative percent difference (RPD) shown on the
 544 x-axes as z . Note that the x-axes are shown from 1 to 10% (recall from Section 3.2 that

545 the RPDs can reach as high as 60%). The y-axis indicates the PoP (%). Taking the graph
546 associated with the Alaska site as an example, when a 1% RPD in products (R_{rs} , [Chla],
547 $K_d(490)$, $b_{bp}(655)$) is adopted as an accepted threshold (i.e., pixels are assumed
548 consistent), there is, on average, only 40-60% of the pixels that meet this threshold. As
549 one accepts higher uncertainty thresholds more pixels fall within the threshold. As it
550 appears from Figs. 6a & 6d, the product discrepancy is dependent on inherent spatial
551 variability of the coastal ecosystems under study. For the sites with less spatial variability
552 (inferred via visual analyses), over 90% of the products are in good agreement, when
553 uncertainty threshold is set at 2%. Since the shape of the histograms resemble normal
554 distributions, it is also possible to express the uncertainties in intercomparisons as
555 probability. These sites include East China and Saint Lawrence. Consider MODISA-
556 VIIRS discrepancy in [Chla] for the East China site (Fig. 6a), it can be inferred that there
557 is a 68% chance for any given pixel to exhibit differences below $\pm 1\%$. For the Saint
558 Lawrence site, the inconsistency curves quickly reach 100% for all the products at the ~
559 5% threshold. These results (including those not shown here) indicate that there is, on
560 average, 80% chance that product intercomparisons show inherent differences below 8%
561 (due to the combination of in-scene spatial variability and sensors' viewing geometries),
562 a threshold at which the product can be considered consistent. In other words, there is a
563 20% chance that products show more than ~8% discrepancies. Depending on the
564 application, one may choose different thresholds for such intercomparison analyses.
565 Visual inspections of RPD grids verify that high discrepancies occur at the proximity of
566 land-water interfaces, where the two instruments view such high-contrast regions from

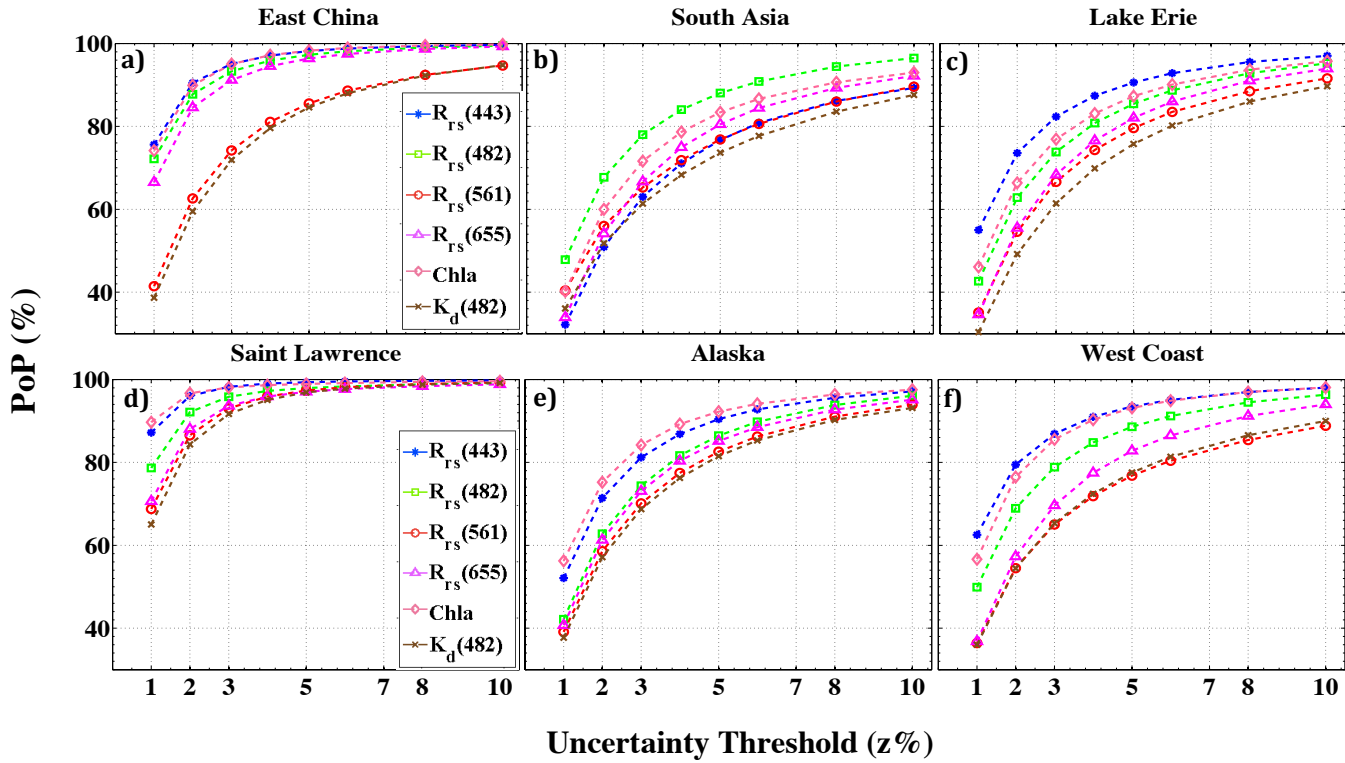


Fig. 6. The spatial interconsistency curves showing the percentage of pixels (PoP) found consistent at different $|RPD|$ thresholds (x-axes). As the absolute RPD increases more pixels (% PoP) are expected to fall within the threshold. On average, more than 80% of the pixels are considered consistent when $z=8\%$ difference in products are adopted as the threshold.

567

568

569

570

571

572

573

574

575

576

577 different vantage points. Note that in actual imaging conditions, adjacency effects
578 may further complicate the discrepancies (Meister and McClain 2010). Although the
579 original OLI products have undergone smoothing and de-noising, there are some signal-
580 dependent trends in the spatial consistency curves (as seen in Fig.6). The red channel and
581 the $b_{bp}(655)$ products (derived from MODISA and VIIRS) commonly show the largest
582 discrepancies, whereas (for sites with $[Chla] < 0.3 \text{ mg/m}^3$) R_{rs} (443) product is found to
583 be the most consistent product. This is consistent with the previously published works
584 (Mélin et al. 2007; Moore et al. 2015) and is attributed to the low signal levels
585 measured/retrieved. For all the sites and geometries, K_d (490) is, on average, amongst the
586 relatively consistent products while the $[Chla]$ products exhibit less consistency (slightly
587 better than those of R_{rs} (655) and b_{bp} (655)).

588 The graphs shown in Fig. 6 indicate the interconsistency between products derived
589 from MODISA and VIIRS. It is also important to obtain insights into how VIIRS-
590 MODISA cross-comparisons differ from those for MODIST-MODISA. This would
591 reveal how differences in the orbits and viewing geometries can, overall, contribute to
592 intercomparison in coastal ocean waters. Fig. 7 shows the median ratio of PoPs
593 associated with the MODISA-VIIRS and MODIST-MODISA intercomparisons as a
594 function of varying thresholds (τ). It is inferred that the MODISA- and VIIRS-derived
595 products are more consistent (12% at maximum) than MODIST-MODISA products.

596 For instance, there are 8-12% more pixels falling within $\pm 1\%$ threshold (τ) when
597 MODISA and VIIRS products are compared, i.e., VIIRS and MODISA are slightly in a
598 better agreement than MODIST and MODISA. This indicates the importance of satellite
599 orbits (ascending versus descending) on product interconsistency and long-term global

600

601

602

603

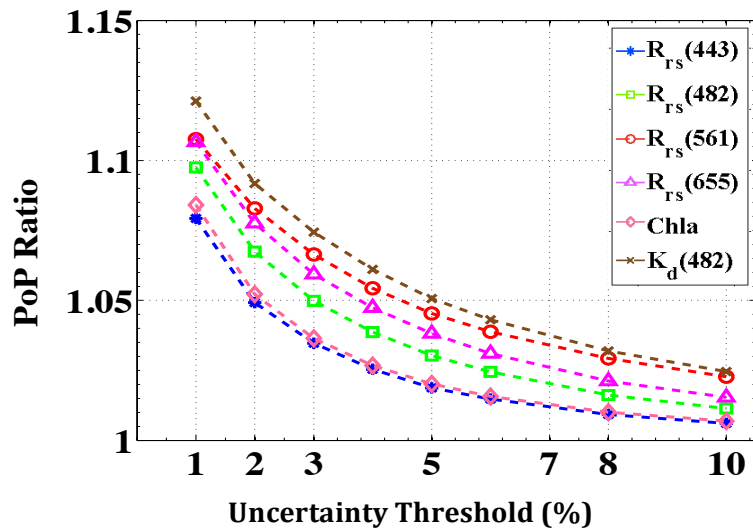
604

605

606

607

608



609

Fig. 7. The ratio of MODISA-VIIRS and MODIST-MODISA intercomparisons (based on PoP) as a function of the uncertainty threshold (absolute RPD). Overall, MODISA and VIIRS products are more consistent than when products of MODIST and MODISA are compared.

610

611

612

613

614

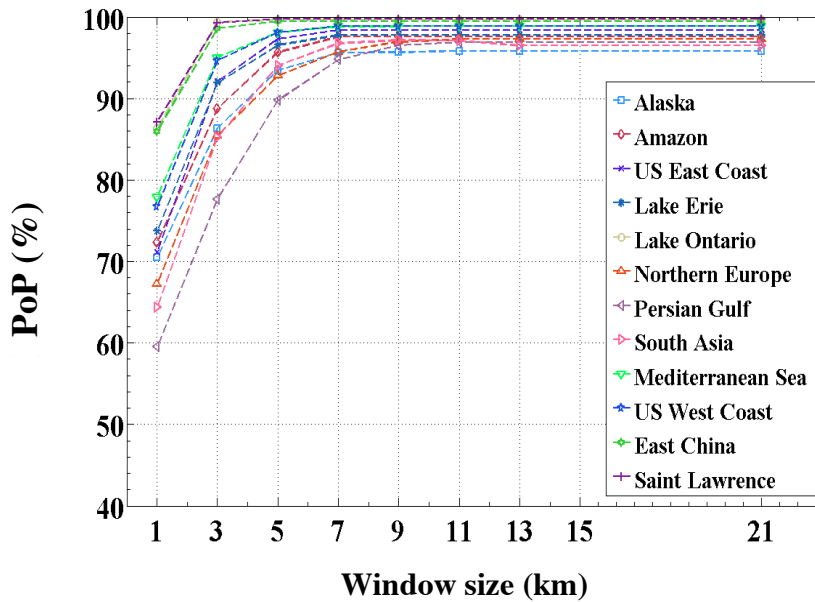
615

616

617

618

619



620

621

Fig. 8. The spatial consistency curves shown for $R_{rs}(443)$ as a function of scale, i.e., $1\text{km}\times 1\text{km}$, $3\text{km}\times 3\text{km}$, etc. The y-axis denotes PoP at $z=1\%$, which increases as the effective grid cell size increases. The optimal scale, on average, is found at 7km.

622 monitoring of coastal oceans. As larger thresholds are adopted, the intercomparisons are
623 expected to become very similar for the two sensor pairs (e.g., less than 3% difference in
624 PoP at $z = \pm 10\%$ threshold). Note that these intercomparisons encompass all
625 observations (pixels) across various OLI scenes, which permits to gauge consistency on a
626 per-grid cell basis between sensor (product) pairs.

627 **3.1.3. Scale-dependency**

628 The intercomparisons presented in the previous section showed product discrepancies
629 at the 1km grid cell size (Section 2.1.4). In this section, we will discuss the spatial scales
630 at which nearly all the pixels meet a $z = \pm 1\%$ difference threshold (RPD). For this
631 purpose, the 1km grid cell is filtered using varying window sizes (i.e., 3×3 , ..., 21×21).
632 The products (smoothed at different scales) are then compared (i.e., calculating RPD) to
633 quantify at what scale, i.e., window size, they can be regarded as consistent, i.e., the
634 inconsistencies reach a minimum. At such scales, the product intercomparisons can be
635 considered insensitive to the effects of spatial sampling. To exclude the artifacts due to
636 the averaging at the land-water interfaces, the grid cells 21km from these boundaries
637 were not incorporated in the analysis. Fig. 8 illustrates the spatial interconsistency curves
638 for the Rrs(443) products shown for the different sites studied here. The y-axis is the PoP
639 (%) specified for $z = \pm 1\%$ threshold and the x-axis denotes the window size used to
640 smooth the products. As expected the product inconsistencies are at minimum when
641 comparing relatively uniform coastal scenes. This is noticeable for the Saint Lawrence
642 and East China sites, for which the curves quickly peaks when products are smoothed
643 with 3×3 windows, i.e., effective grid size of $3km \times 3km$. However, for most sites, the
644 spatial interconsistency curves plateau (does not necessarily reach 100%) when the

645 effective grid cell size is $\sim 7\text{km}$. This indicates that for relatively reliable product
646 intercomparisons in offshore regions, one may need to perform a low-pass filtering with
647 window sizes of at least 7×7 to *minimize* the impact of high-frequency variations in
648 coastal ocean waters. In our case studies, the exception is the Persian Gulf site where the
649 corresponding curve does not reach a maximum until 9×9 windows are implemented. It
650 is, therefore, surmised that the intercomparison studies at $9\text{km} \times 9\text{km}$ standard Level-3
651 products are expected to be least impacted by the effects of spatial sampling, i.e.,
652 differences in the viewing geometries, orbits, footprint sizes, etc.). Note that depending
653 on the inherent spatial variability of the sites, there may be some residual differences in
654 the products as seen in Fig. 8 for spatial scales larger than 9km . It is stressed that the
655 result in this section applies to the offshore regions ($> 21\text{km}$ off the coastlines), where
656 RPD values, in general, remain below 10% ($|z| < 10\%$).

657 **3.2. Spatial representativeness at In-situ sites**

658 As described in Section 2.2, for the local spatial analysis, the OLI observations at the
659 proximity of the field sites are regarded as the “sea-truth” and compared against
660 simulated coarse OC products. Fig. 9 shows PDs for the available OLI-derived $R_{rs}(443)$
661 products at the MVCO, Venise, and Zeebrugge sites. Note that the OLI scenes available
662 from 2013 to 2015 are shown for one annual period and grouped according to the OLI
663 observation dates, i.e., OLI scenes. For these sites, MODISA, MODIST, and VIIRS
664 simulated observations are averaged over 3×3 boxes. Several conclusions can be drawn
665 from the graphs. First, although the $CV < 0.15$ threshold was implemented to discard

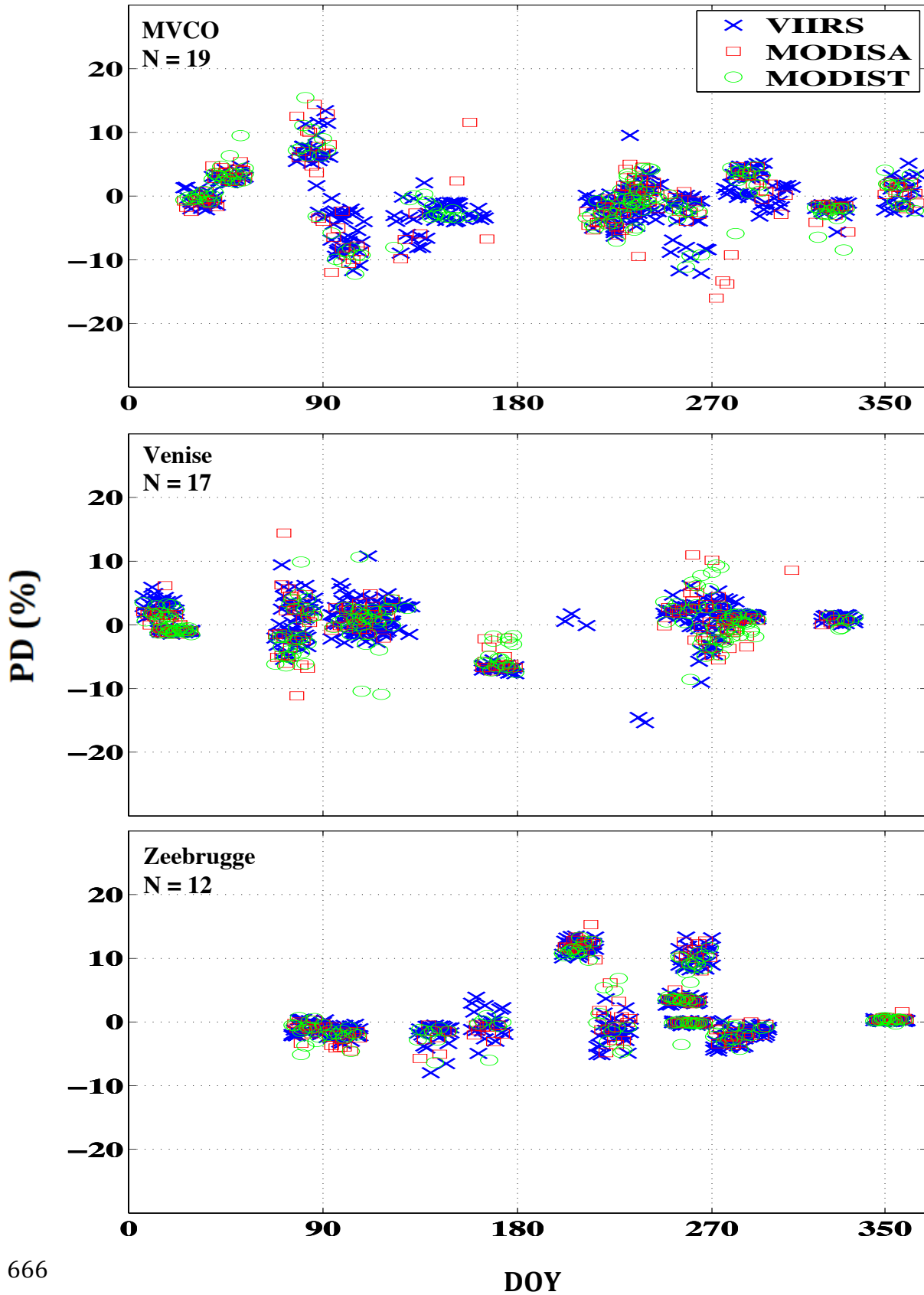
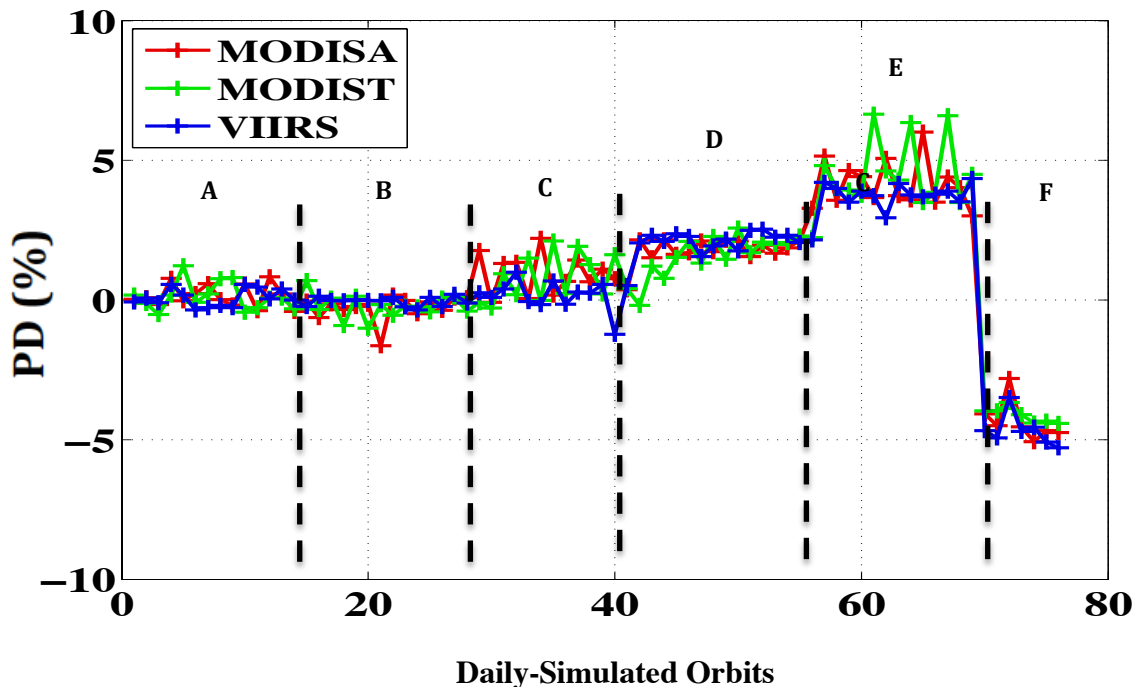


Fig. 9. The simulated percent differences (PD) for $R_{rs}(443)$ (Eq. 8) shown for the three AERONET-OC stations for all available OLI scenes (N) at the sites. The PDs are calculated for 3×3 simulated OC pixels and a representative OLI OC pixel. The PD ranges from -20% to +30% for these sites and is mainly dependent on the environmental conditions (in-water features) captured by the OLI scene. The MODISA and MODIST show more day-to-day variability with respect to VIIRS.

667 simulated outliers, the daily simulated PDs (Eq. 8) can reach as high as $\pm 18\%$. The
668 largest PDs may be attributed to influxes of terrestrial inputs and algal bloom events.
669 Second, MODISA and MODIST exhibit larger day-to-day variability in PD (vertical
670 axes) than that for VIIRS during an orbit cycle for most of the coastal conditions. The
671 per-orbit cycle variability of PD, on average, ranges from 1% to 15%. Third, spatial
672 variability around each site primarily drives the magnitude of the difference (note the
673 cluster of data points associated with each OLI scene). On the other hand, the variations
674 in viewing geometry result in a random variability (clutter) around the average PD. Note
675 that the PDs (or APDs) in the red channel are the largest (i.e., $< 25\%$) amongst all the
676 channels due to the relatively small signal (Eq. 9).

677 This random variability in the day-to-day simulations is lowest for the MOBY and
678 BOUSSOLE sites. Fig. 10 illustrates variability for different orbits for six (A to F) OLI
679 scenes at the MOBY site. To comply with the existing calibration protocols of ocean
680 color products at the MOBY site (Bailey and Werdell 2006), the median of simulated
681 ocean color observations are computed within 5×5 -element windows and compared
682 against representative OLI pixels (Section 2.3). For the scenes A, B, and C, the mean PD
683 is nearly zero with some random variability from orbit to orbit. The random variability
684 (around the mean trend) ranges from 0.5% to 1.5%. In general, VIIRS exhibits less
685 variability relative to MODISA and MODIST, in particular when spatial variability
686 around the site increases (e.g., scene E). The overall (mean) trend generally remains
687 within the $\pm 5\%$ range. The differences found for the BOUSSOLE site (not shown here)
688 range within $\pm 4\%$. It is thus worthwhile noting that the variations in the footprint size

689 increase uncertainties (random variability) and may introduce bias in
690 calibration/validation efforts at these spatially homogenous sites.



691

Fig. 10. The PD (%) shown for three OLI-derived R_{rs} (443) products (i.e., A, B, C, D, E, and F) for the MOBY site. The curves correspond to MODISA (red), MODIST (green), and VIIRS (blue) products. The solid lines denote PDs corresponding to 5×5 windows. The larger variability associated with MODIST and MODISA with respect to VIIRS is noticeable.

692

693

694

695

696

697

698

699

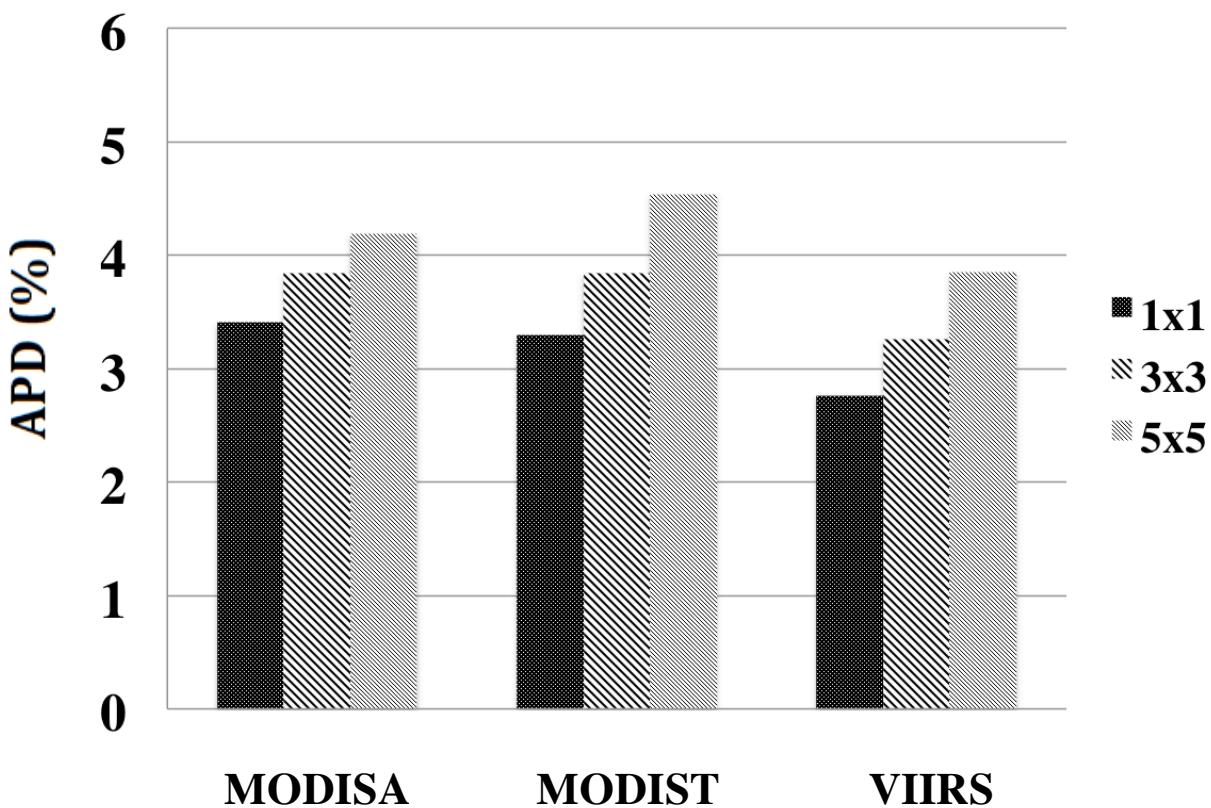
700

701

702

703 In the OC matchup analysis, after performing a multi-stage filtering (Bailey and
704 Werdell 2006), the uncertainties are commonly attributed to the sensor calibration (at a
705 reference wavelength), and radiative transfer modeling (atmospheric correction), as well
706 as atmospheric conditions. Temporal averaging, however, diminishes impacts of spatial
707 mismatch at validation sites (Bailey et al. 2008; Franz et al. 2007). To provide insights
708 into the overall expected biases for the coastal sites for matchup analysis, we present
709 average statistics derived from the metrics defined in Section 2.3. After discarding all the
710 outliers (i.e., flagged simulated OC observations), 1091, 1084, and 2043 valid
711 “matchups” were incorporated in this statistical analysis for MODISA, MODIST, and
712 VIIRS, respectively. Fig. 11 illustrates the average absolute differences (APDs) for the
713 443nm channel. It is inferred that, on average, the APD (Eq. 9) increases with the
714 increase in the window size at these coastal sites. This is expected as larger spatial
715 variability is introduced in the matchup analysis by incorporating more OC observations
716 in coastal waters. Compromise has to be made as increasing the window size lowers the
717 noise contribution. Furthermore, there is, on average, 10-15% less difference associated
718 with simulated matchups for VIIRS when compared to those of MODISA and MODIST.
719 Although use of a single (center) pixel gives rise to a minimum difference in our
720 simulated matchup analysis, in practice, residual detector striping/banding adds another
721 source of uncertainty in the analysis. Thus, depending on the instrument performance (or
722 the efficiency of the de-striping approach) and the environmental conditions (turbidity,
723 resuspension, river plumes, algal blooms, etc.) either center pixel (1×1) or the

724 median/mean of a 3×3 window is recommended for use. Currently, the use of a 3×3-
 725 element window is common practice for validating ocean color products in coastal waters
 726 (Zibordi et al. 2009a). Although an apparent random variability is expected for



727

Fig. 11. The overall mean absolute differences (APDs) computed for simulated MODISA (N=1091), MODIST (N=1084), and VIIRS (N=2043) for different window sizes shown for the OLI 443nm channel. The windows are centered over the AERONET-OC (costal) sites.

728

729

730

731

732

733

734

735

736 Table 2. The average, per-band statistics computed for all the AERONET-OC (coastal) sites.
737

	Band (nm)	1x1			3x3			5x5		
		APD (%)	RMSD (1/sr)	PD (%)	APD (%)	RMSD (1/sr)	PD (%)	APD (%)	RMSD (1/sr)	PD (%)
MODISA	443	3.4	0.00024	0.8	3.8	0.00031	1	4.2	0.00033	1.4
	482	3.3	0.00029	0.7	3.7	0.00039	0.7	4.1	0.00043	1
	561	3.4	0.00040	0.6	4	0.00049	0.7	4.5	0.00055	1
	655	6.8	0.00038	0.6	8.1	0.00042	1.1	8.3	0.00047	1.9
MODIST	443	3.3	0.00023	0.8	3.8	0.00028	0.9	4.5	0.00036	1.6
	482	3.2	0.00027	0.6	3.7	0.00035	0.7	4.4	0.00045	1
	561	3.4	0.00037	0.5	4.3	0.00046	0.6	5.1	0.00056	1
	655	6.2	0.00038	1	8.4	0.00042	1.6	9	0.00049	1.4
VIIRS	443	2.8	0.00020	0.5	3.3	0.00027	0.6	3.8	0.00033	0.9
	482	2.7	0.00024	0.4	3.2	0.00034	0.6	3.8	0.00042	0.7
	561	2.9	0.00032	0.4	3.6	0.00045	0.6	4.3	0.00054	0.8
	655	5.6	0.00032	0.9	7.4	0.00041	1.8	8.6	0.00047	2.3

738

739

740

741

742

743

744

745

746

747

748

749

750

751 observations made from different orbits (Fig. 9) over varying environmental conditions
 752 (turbidity, resuspension, river plumes, etc.), we found that there is a bias that may not be
 753 removed by temporal averaging. At the coastal sites, the mean bias is found to be
 754 $\sim 0.8\%$, which is the average of the mean difference, i.e., PD, derived for all the spectral
 755 bands and window sizes (Table 2). For the BOUSSOLE and MOBY sites, the temporally
 756 averaged PDs are found to fall within the $\pm 0.1\%$ range for the 443nm channel. The
 757 average PD, APD, and RMSD metrics are tabulated in Table 2. Note that these
 758 differences are attributed to matchup analysis at field stations where temporal averaging
 759 is possible. However, at in-situ sampling stations (taken during research cruises), the
 760 difference may be as high as $\pm 18\%$ (Fig. 9) depending on the footprint size and
 761 environmental conditions.

762 In addition to the overall expected differences, we also give insights into the mean
 763 variability, i.e., standard deviation around mean difference, shown along the y-axes of
 764 Figs. 9 for all the sites as a function of window size for the coastal sites. We found the
 765 uncertainty (variability) to be 3%, 3.5% and 3.9% for the 443nm channel, and 3%, 4.5%,

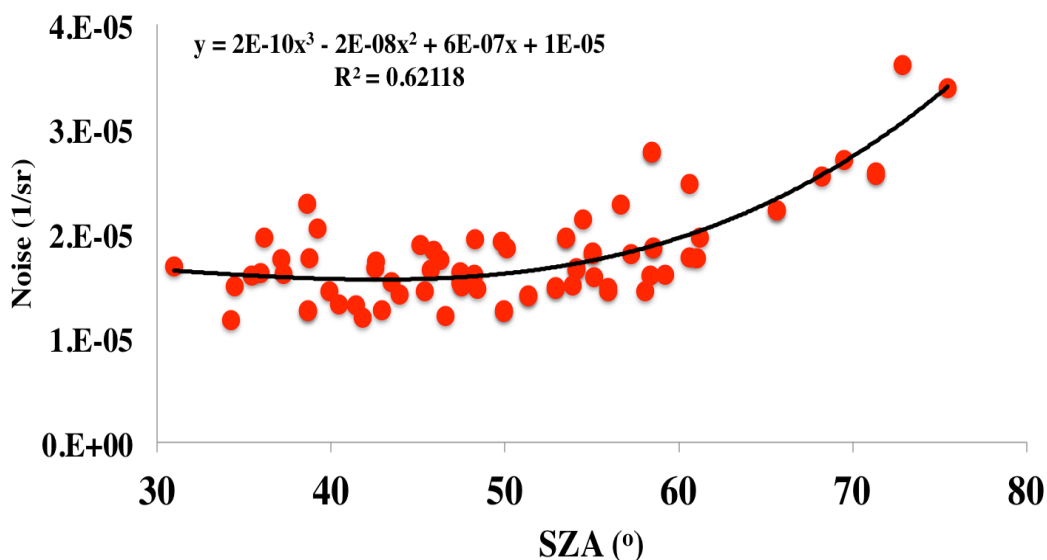


Fig. 12. The OLI-derived $R_n(443)$ noise (1/sr) is shown as a function of solar zenith angle (SZA). The standard deviation was calculated over 5×5 -element windows over uniform bodies of water from various OLI scenes (N=75). The product noise remains relatively stable for low to

766 and 7.8% for the 561nm channel when 1×1, 3×3, and 5×5 windows are used
767 respectively. In contrast, the variability in PD (or APD) at the BOUSSOLE and MOBY
768 sites (located in blue waters) was reduced by increasing the window size, i.e., the 5×5-
769 element window provided minimum variability.

770 **4. DISCUSSIONS**

771 The basis of the spatial analysis presented here was the OLI-derived ocean color
772 products. Although OLI offers dramatically improved SNR when compared with the
773 previous Landsat sensors, the striping and banding effects limit the quality of the ocean
774 color products at the inherent 30-m resolution. This is, in particular, noticeable at lower
775 signal levels when solar zenith angle (SZA) is high. Fig. 12 illustrates the $R_{rs}(443)$ noise
776 as a function of SZA computed over uniform bodies of water across various independent
777 OLI-derived products. In order to surmount this instrument-induced variability within
778 OLI-derived products, the high-frequency components were removed by running a 3×3-
779 element window size. To evaluate the impact of the filtering, we conducted large-area
780 simulations (Section 3.1) for a handful of OLI scenes and found that the results (Table 1)
781 changed only by a fraction of a percent. The results presented in Section 3.2, however,
782 showed a more appreciable difference (< 5% in Table 2) for some of the coastal sites
783 located at higher latitudes where the impact of striping in OLI products was severe. These
784 scenes (characterized with low SNR) were therefore reprocessed with an average
785 Ångström exponent (derived over a subscene) to minimize the effects of striping.

786 The OLI scenes applied in Section 2.1 represents a large range of productivity/
787 turbidity ($0 < [Chla] < 15 \text{ mg/m}^3$) from Amazon River basin, Mississippi River and
788 Bay of Bengal to the Great Lakes. Our assumptions (e.g., use of a single sampling

789 function for MODIS across all channels; Fig. 2) in the simulation process have no
790 significant effects on the results. This was examined using a slightly different PSF over
791 three OLI scenes. The average APDs presented in Section 3.1 provide average estimates
792 for an overall error budget analysis when comparing gridded ocean color products at 1km
793 resolution. However, as discussed in Section 3.1, the actual per-pixel discrepancies may
794 differ depending on in-water features and viewing geometries.

795 Although the differences (Section 3.2) were found to reach up to ~18% in the blue-
796 green channels, the spatial mismatch in different (OLI) channels tend to agree in their
797 signs (Table 2) implying that even at highly spatially variable sites the impact on the
798 ocean color products derived from band-ratio algorithms (e.g., [Chla]) are expected to be
799 minimal. Note, however, that this does not hold true for all cases studied here, i.e., 10-
800 20% of the simulated matchups showed contradictory trends. The average statistics
801 (Table 2) extracted from more than 100 OLI scenes over six different coastal sites
802 provide a set of robust estimates for error budget analysis in ocean color product
803 validation chain in coastal waters. While the revisit cycle of these polar orbiters (SNPP,
804 Aqua, and Terra) similarly repeats every 16 days, we conducted a sensitivity study on the
805 choice of the orbit cycle where DOY=5 through DOY=20 orbits of Aqua-MODIS was
806 utilized for simulations over the MVCO site. The overall APD at the site changed only on
807 the order of 0.1% indicating that the choice of the 16-day orbit cycle is insignificant.

808 **5. CONCLUSIONS**

809 This study presents a novel, comprehensive approach to characterize the uncertainties
810 associated with the product intercomparisons and in-situ validation efforts induced by
811 different spatial sampling effects of three OC imagers. The high quality, moderate-

812 resolution (~30m) OLI-derived OC products obtained over various regions/locations are
813 central to the results presented here. Here, we emphasize the critical conclusions of the
814 study.

815 The following conclusions can be drawn from the results pertaining to the product
816 intercomparisons at regional scales:

- 817 • Due to the differences in the spatial sampling schemes of MODISA and VIIRS, there
818 is a viewing-angle (footprint-size) dependency on the product intercomparisons.
819 Therefore, care must be taken when high-fidelity per-pixel intercomparison is
820 desired.
- 821 • The mean absolute percent differences (APD) in product intercomparisons due to the
822 differences in the spatial sampling are estimated to be 1.8%, 1.9%, 2.4%, 4.3%,
823 2.7%, 1.8%, and 4% for the $R_{rs}(443)$, $R_{rs}(482)$, $R_{rs}(561)$, $R_{rs}(655)$, [Chla], $K_d(490)$,
824 and $b_{bp}(655)$ products, respectively.
- 825 • The differences in the products in the offshore regions are mitigated if a pair of
826 products is smoothed over window sizes of at least $7km \times 7km$.

827 The radiometric differences at coastal in-situ sites always include some level of
828 uncertainty in the spatial representativeness of the site. With the strategies set forth here,
829 we provide the following conclusions to give insights into the magnitude of these
830 uncertainties:

- 831 • The ocean color observation centered on the stationary in-situ coastal sites provides
832 the minimum temporally averaged bias and variability. Assuming residual
833 striping/banding effects in the products, a 3×3 -element window is recommended.

834 On the other hand, 5×5-element boxes yield both lower biases and uncertainties
835 over clear open waters.

836 • The contribution of spatial sampling to differences between satellite and in-situ
837 measurements in near-shore areas can reach as large as 18%. This varies according
838 to the in-water spatial variability and the satellite orbit, which determines the
839 footprint size (sensor viewing angle).

840 • The APD in the spatial representativeness of a field site (for R_{rs} products) is found,
841 on average, to be 3.7%, 3.5%, 3.9%, 8.3% for the 443, 482, 561, and 655nm
842 channels, respectively.

843 • The daily observations of VIIRS show less random variability (induced by spatial
844 sampling) than those of MODISA and MODIST at the calibration sites (e.g.,
845 MOBY) as well as the coastal validation sites. In addition, the number of valid
846 simulated matchups for VIIRS was found twice larger than that of MODIS
847 suggesting that more robust statistical analysis is possible for the VIIRS products.

848 The results of this study, for the first time, allow for taking one step forward for a full
849 quantification of the overall error budget analysis of coastal ocean products by isolating
850 the errors associated with spatial sampling. Similar sensitivity analyses can be performed
851 to estimate uncertainties/biases in spectral sampling and geolocation errors to further
852 decompose the overall error budget. We further highlight the advantages of near-uniform
853 along-scan spatial sampling (assuming no loss in the radiometric performance) for future
854 ocean color missions like the Pre-Aerosol, Cloud, and ocean Ecosystem (PACE) and the
855 planned GEOstationary Coastal and Air Pollution Events (GEO-CAPE) to enhance our

856 ability in validating products and quantifying corresponding uncertainties for reliable
857 monitoring of the changing coastal waters.

858

859 **Acknowledgement**

860 Financial support by the Geo-CAPE program at the NASA headquarters and support
861 from Antonio Mannino with the NASA Ocean Biology Processing Group (OBPG) are
862 greatly appreciated. We are also grateful to Robert E. Wolfe and Gary Lin with NASA
863 GSFC's Terrestrial Information Systems Lab for the discussions of spatial performance
864 of VIIRS and MODIS instruments. The computing support at the Terrestrial Information
865 Systems Lab by Miguel O. Román and Ed Masuoka is acknowledged. We are also
866 grateful to the anonymous reviewers for their thoughtful comments that help improve this
867 manuscript.

868

869 **References**

870

871 Allen, J.I., Smyth, T.J., Siddorn, J.R., & Holt, M. (2008). How well can we forecast high biomass algal
872 bloom events in a eutrophic coastal sea? *Harmful Algae*, 8, 70-76

873 Antoine, D., d'Ortenzio, F., Hooker, S.B., Bécu, G., Gentili, B., Tailliez, D., & Scott, A.J. (2008a).
874 Assessment of uncertainty in the ocean reflectance determined by three satellite ocean color sensors
875 (MERIS, SeaWiFS and MODIS-A) at an offshore site in the Mediterranean Sea (BOUSSOLE project).
876 *Journal of Geophysical Research: Oceans*, 113

877 Antoine, D., Guevel, P., Deste, J.-F., Bécu, G., Louis, F., Scott, A.J., & Bardey, P. (2008b). The
878 "BOUSSOLE" buoy-a new transparent-to-swell taut mooring dedicated to marine optics: Design, tests, and
879 performance at sea. *Journal of Atmospheric and Oceanic Technology*, 25, 968-989

880 Bailey, S.W., Franz, B.A., & Werdell, P.J. (2010). Estimation of near-infrared water-leaving reflectance
881 for satellite ocean color data processing. *Optics express*, 18, 7521-7527

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

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

- 886 Baker, N. (2011). VIIRS Geolocation Algorithm Theoretical Basis Document (ATBD). In. Goddard
887 Space Flight Center, Greenbelt, MD: Joint Polar Satellite System (JPPS)
- 888 Barnes, B.B., & Hu, C. (2015). Cross-Sensor Continuity of Satellite-Derived Water Clarity in the Gulf
889 of Mexico: Insights Into Temporal Aliasing and Implications for Long-Term Water Clarity Assessment.
890 *Geoscience and Remote Sensing, IEEE Transactions on*, 53, 1761-1772
- 891 Barnes, W.L., Pagano, T.S., & Salomonson, V.V. (1998). Prelaunch characteristics of the moderate
892 resolution imaging spectroradiometer (MODIS) on EOS-AM1. *Geoscience and Remote Sensing, IEEE*
893 *Transactions on*, 36, 1088-1100
- 894 Campbell, J.W., Blaisdell, J.M., & Darzi, M. (1996). Level-3 Sea WiFS data products: spatial and
895 temporal binning algorithms. *Oceanographic Literature Review*, 9, 952
- 896 Cao, C., De Luccia, F.J., Xiong, X., Wolfe, R., & Weng, F. (2014). Early on-orbit performance of the
897 visible infrared imaging radiometer suite onboard the suomi national polar-orbiting Partnership (S-NPP)
898 satellite. *Geoscience and Remote Sensing, IEEE Transactions on*, 52, 1142-1156
- 899 Cao, C., Xiong, J., Blonski, S., Liu, Q., Uprety, S., Shao, X., Bai, Y., & Weng, F. (2013). Suomi NPP
900 VIIRS sensor data record verification, validation, and long-term performance monitoring. *Journal of*
901 *Geophysical Research: Atmospheres*, 118, 11,664-611,678
- 902 Chander, G., Hewison, T.J., Fox, N., Wu, X., Xiong, X., & Blackwell, W.J. (2013). Overview of
903 intercalibration of satellite instruments. *Ieee Transactions on Geoscience and Remote Sensing*, 51, 1056-
904 1080
- 905 Clark, D., Gordon, H., Voss, K., Ge, Y., Broenkow, W., & Trees, C. (1997). Validation of atmospheric
906 correction over the oceans. *Journal of Geophysical Research: Atmospheres (1984–2012)*, 102, 17209-
907 17217
- 908 Esaias, W.E., Abbott, M.R., Barton, I., Brown, O.B., Campbell, J.W., Carder, K.L., Clark, D.K., Evans,
909 R.H., Hoge, F.E., & Gordon, H.R. (1998). An overview of MODIS capabilities for ocean science
910 observations. *Geoscience and Remote Sensing, IEEE Transactions on*, 36, 1250-1265
- 911 Franz, B.A., Bailey, S.W., Kuring, N., & Werdell, P.J. (2015). Ocean color measurements with the
912 Operational Land Imager on Landsat-8: implementation and evaluation in SeaDAS. *Journal of Applied*
913 *Remote Sensing*, 9, 096070-096070
- 914 Franz, B.A., Bailey, S.W., Werdell, P.J., & McClain, C.R. (2007). Sensor-independent approach to the
915 vicarious calibration of satellite ocean color radiometry. *Applied Optics*, 46, 5068-5082
- 916 Franz, B.A., Werdell, P.J., Meister, G., Bailey, S.W., Eplee Jr, R.E., Feldman, G.C., Kwiatkowskaa, E.,
917 McClain, C.R., Patt, F.S., & Thomas, D. (2005). The continuity of ocean color measurements from
918 SeaWiFS to MODIS. In, *Optics & Photonics 2005* (pp. 58820W-58820W-58813): International Society for
919 Optics and Photonics
- 920 Gerace, A.D., Schott, J.R., & Nevins, R. (2013). Increased potential to monitor water quality in the
921 near-shore environment with Landsat's next-generation satellite. *Journal of Applied Remote Sensing*, 7,
922 073558-073558
- 923 Gohin, F., Loyer, S., Lunven, M., Labry, C., Froidefond, J.-M., Delmas, D., Huret, M., & Herbland, A.
924 (2005). Satellite-derived parameters for biological modelling in coastal waters: Illustration over the eastern
925 continental shelf of the Bay of Biscay. *Remote Sensing of Environment*, 95, 29-46

- 926 Gordon, H.R. (1997). Atmospheric correction of ocean color imagery in the Earth Observing System
927 era. *Journal of Geophysical Research: Atmospheres* (1984–2012), 102, 17081-17106
- 928 Holst, G.C. (2008). *Electro-optical Imaging System Performance*. (Fifth ed.). Bellingham JCD Buliding
929 and SPIE
- 930 Hooker, S., Esaias, W., Feldman, G., Gregg, W., & McClain, C. (1992). In SB Hooker, & ER Firestone.
931 *An overview of SeaWiFS and Ocean Color, NASA Tech. Memo, 104566, 24*
- 932 Hooker, S.B., & Maritorena, S. (2000). An evaluation of oceanographic radiometers and deployment
933 methodologies. *Journal of Atmospheric and Oceanic Technology*, 17, 811-830
- 934 Hu, C., Feng, L., & Lee, Z. (2013). Uncertainties of SeaWiFS and MODIS remote sensing reflectance:
935 Implications from clear water measurements. *Remote Sensing of Environment*, 133, 168-182
- 936 Hu, C., & Le, C. (2014). Ocean color continuity from VIIRS measurements over Tampa Bay.
937 *Geoscience and Remote Sensing Letters, IEEE, 11, 945-949*
- 938 IOCCG (1997). Ocean-colour data merging. In W.W.A. Gregg, J., Kwiatkowska, E., Maritorena, S.,
939 Mélin, F., Murakami, H., Pinnock, S., Pottier, C. (Ed.), *Reports of the International Ocean Colour*
940 *Coordinating Group* (p. 68). Dartmouth, Canada: IOCCG
- 941 Irons, J.R., Dwyer, J.L., & Barsi, J.A. (2012). The next Landsat satellite: The Landsat Data Continuity
942 Mission. *Remote Sensing of Environment*, 122, 11-21
- 943 Ladner, S., Arnone, R., Vandermeulen, R., Martinolich, P., Lawson, A., Bowers, J., Crout, R.,
944 Ondrusek, M., & Fargion, G. (2014). Inter-satellite comparison and evaluation of navy Suomi-NPP VIIRS
945 and MODIS-Aqua ocean color properties. In, *SPIE Sensing Technology+ Applications* (pp. 911107-
946 911107-911109): International Society for Optics and Photonics
- 947 Lee, Z., Arnone, R., Hu, C., Werdell, P.J., & Lubac, B. (2010). Uncertainties of optical parameters and
948 their propagations in an analytical ocean color inversion algorithm. *Applied Optics*, 49, 369-381
- 949 Lee, Z., Hu, C., Arnone, R., & Liu, Z. (2012). Impact of sub-pixel variations on ocean color remote
950 sensing products. *Optics express*, 20, 20844-20854
- 951 Lee, Z.P., Carder, K.L., & Arnone, R.A. (2002). Deriving inherent optical properties from water color:
952 a multiband quasi-analytical algorithm for optically deep waters. *Applied Optics*, 41, 5755-5772
- 953 Lin, G., Tilton, J.C., Wolfe, R.E., Tewari, K.P., & Nishihama, M. (2013). SNPP VIIRS spectral bands
954 co-registration and spatial response characterization. In J. J. Butler, X. Xiong, & X. Gu (Eds.), *Earth*
955 *Observing Systems XVIII* (pp. 88661G-88661G-88615). San Diego: SPIE
- 956 Maritorena, S., d'Andon, O.H.F., Mangin, A., & Siegel, D.A. (2010). Merged satellite ocean color data
957 products using a bio-optical model: Characteristics, benefits and issues. *Remote Sensing of Environment*,
958 114, 1791-1804
- 959 Maritorena, S., & Siegel, D.A. (2005). Consistent merging of satellite ocean color data sets using a bio-
960 optical model. *Remote Sensing of Environment*, 94, 429-440
- 961 McClain, C.R., Feldman, G.C., & Hooker, S.B. (2004). An overview of the SeaWiFS project and
962 strategies for producing a climate research quality global ocean bio-optical time series. *Deep Sea Research*
963 *Part II: Topical Studies in Oceanography*, 51, 5-42

- 964 McGranahan, G., Balk, D., & Anderson, B. (2007). The rising tide: assessing the risks of climate
965 change and human settlements in low elevation coastal zones. *Environment and urbanization*, 19, 17-37
- 966 Meister, G., Franz, B., Kwiatkowska, E.J., & McClain, C.R. (2012). Corrections to the calibration of
967 MODIS Aqua ocean color bands derived from SeaWiFS data. *Geoscience and Remote Sensing, IEEE*
968 *Transactions on*, 50, 310-319
- 969 Meister, G., & McClain, C.R. (2010). Point-spread function of the ocean color bands of the Moderate
970 Resolution Imaging Spectroradiometer on Aqua. *Applied Optics*, 49, 6276-6285
- 971 Mélin, F., & Franz, B.A. (2014). Assessment of Satellite Ocean Colour Radiometry and Derived
972 Geophysical Products. In C.D. G. Zibordi, A. Parr (Ed.), *Optical Radiometry for Ocean Climate*
973 *Measurements, Experimental Methods in Physical Sciences* (pp. 609-632): Elsevier Academic Press.
- 974 Mélin, F., Vantrepotte, V., Clerici, M., D'Alimonte, D., Zibordi, G., Berthon, J.-F., & Canuti, E. (2011).
975 Multi-sensor satellite time series of optical properties and chlorophyll-a concentration in the Adriatic Sea.
976 *Progress in Oceanography*, 91, 229-244
- 977 Mélin, F., Zibordi, G., & Berthon, J.-F. (2007). Assessment of satellite ocean color products at a coastal
978 site. *Remote Sensing of Environment*, 110, 192-215
- 979 Mélin, F., Zibordi, G., & Djavidnia, S. (2009). Merged series of normalized water leaving radiances
980 obtained from multiple satellite missions for the Mediterranean Sea. *Advances in Space Research*, 43, 423-
981 437
- 982 Moore, T.S., Campbell, J.W., & Feng, H. (2015). Characterizing the uncertainties in spectral remote
983 sensing reflectance for SeaWiFS and MODIS-Aqua based on global in situ matchup data sets. *Remote*
984 *Sensing of Environment*, 159, 14-27
- 985 Morel, A., Antoine, D., & Gentili, B. (2002). Bidirectional Reflectance of Oceanic Waters: Accounting
986 for Raman Emission and Varying Particle Scattering Phase Function. *Applied Optics*, 41, 6289-6306
- 987 Morel, A., & Gentili, B. (1996). Diffuse reflectance of oceanic waters. III. Implication of
988 bidirectionality for the remote-sensing problem. *Applied Optics*, 35, 4850-4862
- 989 Mueller, J.L. (2000). SeaWiFS algorithm for the diffuse attenuation coefficient, K (490), using water-
990 leaving radiances at 490 and 555 nm. *SeaWiFS postlaunch calibration and validation analyses, part*, 3, 24-
991 27
- 992 Mueller, J.L., Fargion, G.S., McClain, C.R., Mueller, J., Brown, S., Clark, D., Johnson, B., Yoon, H.,
993 Lykke, K., & Flora, S. (2004). Ocean Optics Protocols For Satellite Ocean Color Sensor Validation,
994 Revision 5, Volume VI: Special Topics in Ocean Optics Protocols, Part 2. In: NASA Report 211621, 2003,
995 1-36
- 996 Müller, D., Krasemann, H., Brewin, R.J., Brockmann, C., Deschamps, P.-Y., Doerffer, R., Fomferra,
997 N., Franz, B.A., Grant, M.G., & Groom, S.B. (2015). The Ocean Colour Climate Change Initiative: II.
998 Spatial and temporal homogeneity of satellite data retrieval due to systematic effects in atmospheric
999 correction processors. *Remote Sensing of Environment*, 162, 257-270
- 1000 Natvik, L.-J., & Evensen, G. (2003). Assimilation of ocean colour data into a biochemical model of the
1001 North Atlantic: Part 1. Data assimilation experiments. *Journal of Marine Systems*, 40, 127-153
- 1002 Nixon, S.W. (1995). Coastal marine eutrophication: a definition, social causes, and future concerns.
1003 *Ophelia*, 41, 199-219

- 1004 O'Reilly, J.E., Maritorena, S., Mitchell, B.G., Siegel, D.A., Carder, K.L., Garver, S.A., Kahru, M., &
1005 McClain, C. (1998). Ocean color chlorophyll algorithms for SeaWiFS. *Journal of Geophysical Research:*
1006 *Oceans (1978–2012)*, 103, 24937-24953
- 1007 O'Reilly, J.E., Maritorena, S., Siegel, D.A., O'Brien, M.C., Toole, D., Mitchell, B.G., Kahru, M.,
1008 Chavez, F.P., Strutton, P., & Cota, G.F. (2000). Ocean color chlorophyll a algorithms for SeaWiFS, OC2,
1009 and OC4: Version 4. *SeaWiFS postlaunch calibration and validation analyses, part, 3*, 9-23
- 1010 Ouillon, S., Douillet, P., & Andréfouët, S. (2004). Coupling satellite data with in situ measurements and
1011 numerical modeling to study fine suspended-sediment transport: a study for the lagoon of New Caledonia.
1012 *Coral Reefs*, 23, 109-122
- 1013 Pahlevan, N., Lee, Z., Wei, J., Schaff, C., Schott, J., & Berk, A. (2014). On-orbit radiometric
1014 characterization of OLI (Landsat-8) for applications in aquatic remote sensing. *Remote Sensing of*
1015 *Environment*, 154, 272–284
- 1016 Pahlevan, N., & Schott, J.R. (2013). Leveraging EO-1 to Evaluate Capability of New Generation of
1017 Landsat Sensors for Coastal/Inland Water Studies. *Selected Topics in Applied Earth Observations and*
1018 *Remote Sensing, IEEE Journal of*, 6, 360-374
- 1019 Schott, J.R. (2007). *Remote Sensing The Image Chain Approach*. (2nd ed.). New York: Oxford
1020 University Press
- 1021 Schowengerdt, R.A. (1997). *Remote Sensing: Models and Methods for Image Processing*. (2nd ed.).
1022 Chesnutt Hill, MA: Academic Press
- 1023 Storey, J., Choate, M., & Lee, K. (2014). Landsat 8 Operational Land Imager on-orbit geometric
1024 calibration and performance. *Remote Sensing*, 6, 11127-11152
- 1025 Vanhellemont, Q., & Ruddick, K. (2014). Turbid wakes associated with offshore wind turbines
1026 observed with Landsat 8. *Remote Sensing of Environment*, 145, 105-115
- 1027 Vanhellemont, Q., & Ruddick, K. (2015). Advantages of high quality SWIR bands for ocean colour
1028 processing: Examples from Landsat-8. *Remote Sensing of Environment*, 161, 89-106
- 1029 Vitousek, P.M., Mooney, H.A., Lubchenco, J., & Melillo, J.M. (1997). Human domination of Earth's
1030 ecosystems. *Science*, 277, 494-499
- 1031 Vörösmarty, C.J., Green, P., Salisbury, J., & Lammers, R.B. (2000). Global water resources:
1032 vulnerability from climate change and population growth. *Science*, 289, 284-288
- 1033 Wang, M. (1999). A sensitivity study of the SeaWiFS atmospheric correction algorithm: Effects of
1034 spectral band variations. *Remote Sensing of Environment*, 67, 348-359
- 1035 Wang, P., Boss, E.S., & Roesler, C. (2005). Uncertainties of inherent optical properties obtained from
1036 semianalytical inversions of ocean color. *Applied Optics*, 44, 4074-4085
- 1037 Wolfe, R.E., Lin, G., Nishihama, M., Tewari, K.P., Tilton, J.C., & Isaacman, A.R. (2013). Suomi NPP
1038 VIIRS prelaunch and on-orbit geometric calibration and characterization. *Journal of Geophysical*
1039 *Research: Atmospheres*, 118, 11,508-511,521
- 1040 Wolfe, R.E., Nishihama, M., Fleig, A.J., Kuyper, J.A., Roy, D.P., Storey, J.C., & Patt, F.S. (2002).
1041 Achieving sub-pixel geolocation accuracy in support of MODIS land science. *Remote Sensing of*
1042 *Environment*, 83, 31-49

1043 Wolfe, R.E., Roy, D.P., & Vermote, E. (1998). MODIS land data storage, gridding, and compositing
1044 methodology: Level 2 grid. *Geoscience and Remote Sensing, IEEE Transactions on*, 36, 1324-1338

1045 Xiong, X., Che, N., & Barnes, W. (2006). Terra MODIS on-orbit spectral characterization and
1046 performance. *Geoscience and Remote Sensing, IEEE Transactions on*, 44, 2198-2206

1047 Xiong, X., Sun, J., Xiong, S., & Barnes, W.L. (2004). Using the moon for MODIS on-orbit spatial
1048 characterization. In (pp. 480-487)

1049 Zibordi, G., Berthon, J.-F., Mélin, F., D'Alimonte, D., & Kaitala, S. (2009a). Validation of satellite
1050 ocean color primary products at optically complex coastal sites: Northern Adriatic Sea, Northern Baltic
1051 Proper and Gulf of Finland. *Remote Sensing of Environment*, 113, 2574-2591

1052 Zibordi, G., Mélin, F., & Berthon, J.-F. (2012). Intra-annual variations of biases in remote sensing
1053 primary ocean color products at a coastal site. *Remote Sensing of Environment*, 124, 627-636

1054 Zibordi, G., Mélin, F., Berthon, J.-F., Holben, B., Slutsker, I., Giles, D., D'Alimonte, D., Vandemark,
1055 D., Feng, H., & Schuster, G. (2009b). AERONET-OC: a network for the validation of ocean color primary
1056 products. *Journal of Atmospheric and Oceanic Technology*, 26, 1634-1651

1057 Zibordi, G., Mélin, F., & Berthon, J.F. (2006). Comparison of SeaWiFS, MODIS and MERIS
1058 radiometric products at a coastal site. *Geophysical Research Letters*, 33

1059 Zibordi, G., Mélin, F., Hooker, S.B., D'Alimonte, D., & Holben, B. (2004). An autonomous above-
1060 water system for the validation of ocean color radiance data. *Geoscience and Remote Sensing, IEEE*
1061 *Transactions on*, 42, 401-415

1062 Zibordi, G., Mélin, F., Voss, K.J., Johnson, B.C., Franz, B.A., Kwiatkowska, E., Huot, J.-P., Wang, M.,
1063 & Antoine, D. (2015). System vicarious calibration for ocean color climate change applications:
1064 Requirements for in situ data. *Remote Sensing of Environment*, 159, 361-369

1065

1066

1067

1068

1069

1070

1071

1072

1073

1074

1075

1076

1077

1078

1079

1080

1081

1082

1083

1084

1085

1086 **Appendix**

1087 Table A1. The OLI scenes processed for the product intercomparison study (Section 2.1) are tabulated

LC80090262014248LGN00	LC80440342014285LGN00	LC81370452015076LGN00
LC80110302013227LGN00	LC80440352015064LGN00	LC81640402014310LGN00
LC80110322014086LGN00	LC80450342014276LGN00	LC81650392015064LGN00
LC80130322014100LGN00	LC80790152014210LGN00	LC81910212014115LGN00
LC80140372014299LGN00	LC80790162014210LGN00	LC81910292013208LGN00
LC80170302014096LGN00	LC80840122015168LGN00	LC81910292013224LGN00
LC80190312014270LGN00	LC80840132015168LGN00	LC81980222014068LGN00
LC80190402014046LGN00	LC81180372015071LGN00	LC82240592013327LGN00
LC80210402015079LGN00	LC81180392013241LGN00	LC82250582014225LGN00
LC80440342014109LGN00	LC81360452015069LGN00	

- 1088
- 1089
- 1090
- 1091
- 1092
- 1093
- 1094
- 1095
- 1096
- 1097
- 1098
- 1099
- 1100
- 1101
- 1102
- 1103
- 1104
- 1105
- 1106
- 1107
- 1108
- 1109
- 1110
- 1111
- 1112
- 1113
- 1114
- 1115
- 1116
- 1117
- 1118
- 1119
- 1120

1121

List of Figures

1122

1123

1124

1125

Fig. 1. The sites (indicated by boxes) where OLI scenes are processed to Level-2 OC products for intercomparisons of products at regional scales (Section 2.1). For these locations, the MODISA, MODIST, and VIIRS OC products were simulated according to their corresponding viewing geometries from DOY=1 to DOY=16 in 2015. The red dots denote the locations where simulated matchups were produced (Section 2.2).

1126

1127

1128

1129

Fig. 2. The band/detector average point spread functions (PSFs) shown for a) MODIS, b) VIIRS near-nadir zone (PSF-V3g), c) VIIRS mid-range zone (PSF-V2g), and d) VIIRS edge-of-scan zone (PSF-V1g). Note that it is assumed that the along-track response is rectangular (ideal response). All the functions have unit-area. The coordinates are in arbitrary pixel units.

1130

1131

1132

1133

1134

Fig. 3. Examples of simulated MODISA, VIIRS, and MODIST swaths for [Chla] (mg/m^3) fields derived from the OLI-derived [Chla] products over the Yangtze River mouth, East coast of China. The average view zenith angles (VZAs) are 58.18° , 51.5° , and 21.5° for MODISA, VIIRS, and MODIST swaths, respectively. The features are reproduced at different viewing conditions. Note that OLI and MODIST are in descending orbits. Also, the extremely turbid waters over the basin are masked.

1135

1136

1137

1138

Fig. 4. A typical histogram derived from product intercomparisons. Amongst different histogram-derived parameters, the Percent of Pixels (PoP) is shown above for various levels of relative percentage difference, i.e., RPD= 1%, 5%, and 10%. Also, histogram kurtosis indicates how spread is the distribution around the histogram mean. The narrower the histogram, the more consistent a pair of products are.

1139

1140

1141

1142

1143

1144

1145

Fig. 5. The kurtosis values (derived from the RPD histograms) are shown as a function of differences in mean VZAs (\overline{VZA}). The subscript X denotes VIIRS (V) or MODIST (MT). The top row shows the asymmetric trends when MODISA and VIIRS products are compared. The x-axes denote $\Delta VZA = \overline{VZA}_V - \overline{VZA}_{MA}$. This trend implies that the products are more consistent when $\overline{VZA}_V > \overline{VZA}_{MA}$, which yield similar footprint sizes for the two observations. The bottom row corresponds to the intercomparisons of MODISA and MODIST products as a function of $\Delta VZA = \overline{VZA}_{MT} - \overline{VZA}_{MA}$. Although MODISA and MODIST are in afternoon and morning orbits, there is no particular trends found when comparing the associated products at different VZAs.

1146

1147

1148

1149

Fig. 6. The spatial interconsistency curves showing the percentage of pixels (PoP) found consistent at different [RPD] thresholds (x-axes). As the absolute RPD increases more pixels (% PoP) are expected to fall within the threshold. On average, more than 80% of the pixels are considered consistent when $z=8\%$ difference in products are adopted as the threshold.

1150

1151

1152

Fig. 7. The ratio of MODISA-VIIRS and MODIST-MODISA intercomparisons (based on PoP) as a function of the uncertainty threshold (absolute RPD). Overall, MODISA and VIIRS products are more consistent than when products of MODIST and MODISA are compared.

1153

1154

1155

Fig. 8. The spatial consistency curves shown for $R_{rs}(443)$ as a function of scale, i.e., $1km \times 1km$, $3km \times 3km$, etc. The y-axis denotes PoP at $z=1\%$, which increases as the effective grid cell size increases. The optimal scale, on average, is found at 7km.

1156

1157

1158

1159

1160

Fig. 9. The simulated percent differences (PD) for $R_{rs}(443)$ (Eq. 8) shown for the three AERONET-OC stations for all available OLI scenes (N) at the sites. The PDs are calculated for 3×3 simulated OC pixels and a representative OLI OC pixel. The PD ranges from -20% to +30% for these sites and is mainly dependent on the environmental conditions (in-water features) captured by the OLI scene. The MODISA and MODIST show more day-to-day variability with respect to VIIRS.

1161

1162

1163

1164

Fig. 10. The PD (%) shown for three OLI-derived $R_{rs}(443)$ products (i.e., A, B, C, D, E, and F) for the MOBY site. The curves correspond to MODISA (red), MODIST (green), and VIIRS (blue) products. The solid lines denote PDs corresponding to 5×5 -element windows. The larger variability associated with MODIST and MODISA with respect to VIIRS is noticeable.

1165

1166

1167

Fig. 11. The overall mean absolute differences (APDs) computed for simulated MODISA (N=1091), MODIST (N=1084), and VIIRS (N=2043) for different window sizes shown for the OLI 443nm channel. The windows are centered over the AERONET-OC (costal) sites.

1168

1169

Fig. 12. The OLI-derived $R_{rs}(443)$ noise (1/sr) is shown as a function of solar zenith angle (SZA). The standard deviation was calculated over 5×5 -element windows over uniform bodies of water from various OLI scenes

1170 (N=75). The product noise remains relatively stable for low to medium range SZA but increases beyond
1171 SZA=58°.