Suomi NPP VIIRS Ocean Color Data Product Early Mission Assessment

Kevin R. Turpie*[a], Wayne D. Robinson[a], Bryan A. Franz[b], Robert E. Eplee, Jr.[a]
Gerhard Meister[b], Gwyn F. Fireman[a], Frederick S. Patt[a], Robert A. Barnes[a], Charles R. McClain[b]

[a] Science Applications International Corporation, Beltsville, Maryland 20705, USA
[b] NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

ABSTRACT

Following the launch of the Visible Infrared Imaging Radiometer Suite (VIIRS) aboard the Suomi National Polar-orbiting Partnership (NPP) spacecraft, the NASA NPP VIIRS Ocean Science Team (VOST) began an evaluation of ocean color data products to determine whether they could continue the existing NASA ocean color climate data record (CDR). The VOST developed an independent evaluation product based on NASA algorithms with a reprocessing capability. Here we present a preliminary assessment of both the operational ocean color data products and the NASA evaluation data products regarding their applicability to NASA science objectives.

Keywords: NPP, VIIRS, Ocean Color, satellite remote sensing, climate data record.

1. INTRODUCTION

On 28 October 2011, NASA launched the Visible Infrared Imaging Radiometer Suite (VIIRS) aboard the Suomi National Polar-orbiting Partnership (NPP) spacecraft. VIIRS is being used by NOAA to routinely generate measurements of the Earth's surface and atmosphere, which are referred to as Environment Data Records (EDR). The ocean color EDR includes normalized water-leaving radiance, inherent optical properties (absorption and phytoplankton backscatter coefficients, a and b_{bp}, at five wavelengths) based on an algorithm developed by Carder et al.[1], and chlorophyll *a* concentration using the three channel version of the empirical algorithm developed by O'Reilly et al.[2] In the months that followed the launch, the NASA/Goddard Space Flight Center (GSFC) NPP VIIRS Ocean Science Team (VOST), began evaluating ocean color data products. Their task was to determine whether the NOAA ocean color EDR products would meet NASA science objectives, including continuing the existing climate data record (CDR) established with earlier NASA missions. Before the NPP launch, the VOST argued that the NOAA EDR products could not support continuity of the existing NASA ocean color CDR because they were based on different algorithms and because of the lack of support for reprocessing. Both situations could possibly lead to inconsistencies in the continued CDR. Thus to determine whether VIIRS measurements would facilitate NASA CDR continuity, the VOST faced the challenge of developing an independent evaluation data production capability that both used NASA selected algorithms and supported reprocessing. The effort included an independent calibration strategy and leveraged existing processing capabilities that supported heritage ocean color data production. NOAA operational ocean color products were compared to products generated by this independent research processing system and data from other sensors. A preliminary VOST evaluation of these VIIRS ocean color data products is herein presented.

2. BACKGROUND

Although the launch was picture perfect, the prelaunch preparation of Suomi NPP VIIRS met many challenges and delays. Thanks to the dedicated work of many engineers and analysts, most of the challenges with the sensor were eventually met. By the end of 2009, extensive prelaunch sensor testing indicated that no known significant problems remained that would prevent VIIRS from potentially carrying on the ocean color legacy established with the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) and the MODerate resolution Imaging Spectroradiometer (MODIS)[3].

---

* kevin.r.turpie@nasa.gov; phone 1 301 286 9996; fax 1 301 286 0268; oceancolor.gsfc.nasa.gov

Extensive characterization testing of the VIIRS instrument was carried out in the last half of 2009 and extensive observatory-level testing was performed by NIST early in 2010. An assessment of potential risks to ocean color quality was done based on this prelaunch knowledge of the instrument and an examination of ground system capabilities[3]. A predominant instrument issue at the time was the presence of optical communication between bands via the spectral filter array, a phenomenon also known as *optical crosstalk*. What was learned from the characterization data, and later reported, was that the effects of optical crosstalk were markedly smaller than expected from earlier experiments with the engineering design unit and subsequent bench tests with the flight unit integrated filter array (2005-2008). It was therefore left to determine whether the remaining effects of optical crosstalk were significant to ocean color and whether the uncertainty in the characterization was sufficient to form a confident assessment.

Review of the Raytheon Performance Verification Reports (PVR) for the spectral characterization of the VIIRS instrument[4,5] indicated the uncertainty determination was sufficiently thorough and robust, and that the predicted uncertainty was sufficiently small to perform credible assessments of the effects to ocean color EDR quality. To assess those potential ramifications, characterization data was incorporated into the VIIRS Data Simulator[6], which was enhanced to simulate various instrument effects. Over most of the ocean, the results showed that small effects would be present in water-leaving radiances and that the expected effect would be minimal to chlorophyll *a* concentration, with coastal waters being a possible exception if the effects were left uncorrected[6]. Furthermore, sending bands were generally adjacent on the focal plane to bands receiving crosstalk, and thus the spatial effects of optical crosstalk werefound to be usually only one to three samples distant. This was confirmed after launch through examination of lunar images taken with VIIRS, which showed no noticeable image ghosting that would result from crosstalk spatial effects.

As apparently a tradeoff with optical crosstalk, large out-of-band light leaks were observed in the instrument spectral response during testing. Most notable was a 1% bias in the radiometric response for the 410 nm band stemming from NIR wavelengths when observing a solar spectrum and about a 4% bias for the 551 nm band stemming from the blue side of the corresponding band pass. The magnitude of these effects did not exceed that of light leaks associated with the SeaWiFS and MODIS sensors. Although these effects were manifested differently in VIIRS, the belief was that the same correction methodology used for those instruments would be as effective.

Thus, the strategy implemented by NOAA and adopted by NASA to address both the remaining optical crosstalk spectral effects and out-of-band light leaks was to incorporate the full band response of the instrument for each band into the atmospheric correction tables. This approach was originally developed by Gordon[7] to address out-of-band light leaks in the SeaWiFS instrument. That response characteristic was represented by the relative spectral response (RSR) curves derived from instrument characterization testing. Early in 2011, the NPP/JPSS project made RSR curves available that were derived from both the thermal/vacuum test data and NIST observatory-level test data using the Spectral Irradiance and radiance Responsivity Calibrations using Uniform Sources (SIRCUS)[8] system. Because this test involved flooding the entire focal plane with monochromatic light, the results included the spectral effects of crosstalk. These hybrid RSR curves were used to generate new ocean color atmospheric correction tables later in 2011, and are now being used in the NASA evaluation data processing software. The same RSR curves are applied in the operational processing of the EDR products.

After launch, it was clear from calibration system measurements that the instrument's response to light was decreasing more rapidly than expected in the 672 nm, 748 nm, and 865 nm channels. NASA assembled a special engineering team, which carefully evaluated the instrument behavior and traced the problem to tungsten oxide contamination of four of the telescope mirrors during their manufacture. It was further determined that this contamination also affected the 1240nm, 1380nm, and 1610nm channels, which had not yet been activated. Tungsten oxide is a photoactive substance that when exposed to UV radiation causes VIIRS to become increasingly less sensitive in the six aforementioned VisNIR and SWIR channels.

It was further found that the Solar Diffuser Stability Monitor (SDSM) detectors were also degrading at similar wavelengths, but project engineers noted that this was an independent phenomenon. In addition, close examination of the Solar Diffuser reflectance using data from the SDSM revealed that the reference panel was yellowing at a unprecedented rate; probably because the calibrator aperture was designed without a cover and open in the ram direction of the spacecraft. To further understand these effects, considerable data are being analyzed and their overall affect on ocean color data products will be monitored.

Fortunately, the instrument response degradation is trackable and expected to eventually level off so that useful measurements can continue to be made. If the response were to continue to decrease in the NIR channels, alternative

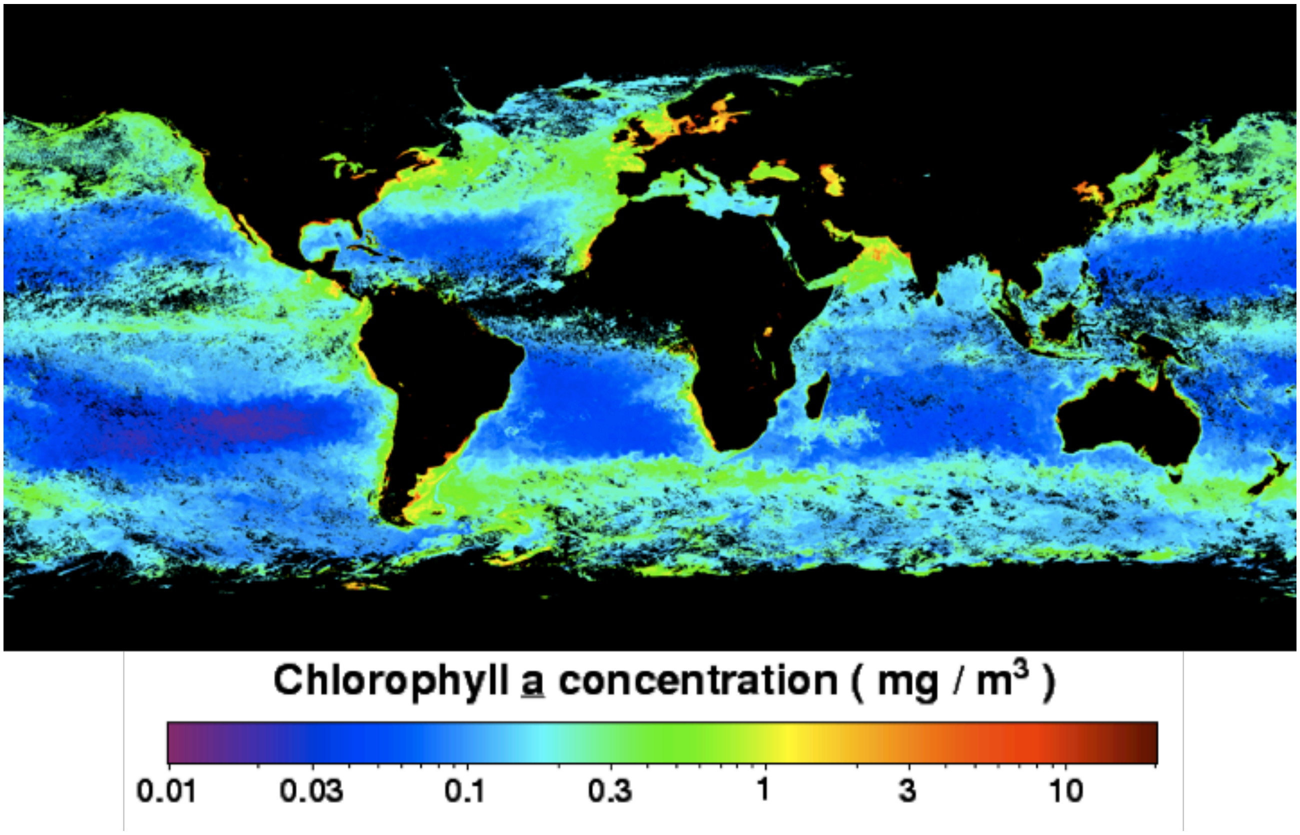

Figure 1. Global composite of chlorophyll a concentration – This level-3 composite of chlorophyll a concentration is based on data taken from the VIIRS sensor for the entire month of March 2012. NASA L2 algorithms and calibration were employed in this case for evaluation purposes. Level-3 products are currently not operationally produced by the Suomi NPP ground system. (Courtesy of NASA/GSFC Ocean Biology Processing Group).

methods can be employed that would boost the signal. In the mean time, the red and NIR channel responses are being monitored using the on-board solar calibration system and lunar calibration. The SWIR channels, conversely, had degraded significantly before they had been turned on, hence making it difficult to relate their current response to prelaunch calibration. Further monitoring and evaluation of the SDSM data will be needed to determine whether its degradation poses a risk to future calibration of VIIRS.

Despite the potential resolution of instrument performance risks, concerns regarding support for ocean color research remained with the data production. These included the lack of support for reprocessing, the exclusion of some data products that were originally generated with SeaWiFS and MODIS, and the use of atmospheric correction algorithms that were not entirely consistent with those employed in the climate data records established with SeaWiFS and MODIS. Also, some doubt lingered regarding whether sufficient support would be provided for ocean color validation[3]. Finally, the operational system does not produce Level 3 products (i.e., global composites of data products) as shown in Figure 1, which are a key tool in understanding large-scale spatial and temporal trends in the data.

3. EVALUATION

The evaluation of the VIIRS ocean color data involved an assessment of both the instrument calibration and flight data. The operational calibration of the VIIRS instrument was reviewed in two aspects. First, the calibration algorithm theoretical basis document, the calibration plan, and the operational code used to perform the count-to-radiance conversion were carefully reviewed. The second aspect of the evaluation of VIIRS ocean color capabilities was the development of an independent calibration capability. This new capability provided an opportunity to independently validate the operational calibration process using measurements from the solar calibration system and calibration measurements taken from the Moon, as described by Eplee et al.[9] More importantly, the independent calibration, along with a time-dependent processing of raw data records (RDR) from the sensor, facilitated a reprocessing capability that is not available in the operational processing stream and critical to producing a consistent data record.

Using the existing computational infrastructure currently supporting the production of ocean color data products for MODIS and other heritage missions, a full suite of derived products could be produced using algorithms consistent with the established ocean color CDR. These products, provided in geolocated satellite swath format, are referred to as NASA Level-2 (L2) data products and are analogous to the NOAA EDR data products. The EDR ocean color product suite, as previously defined in Section 1, are essentially a small subset of the standard NASA L2 data products generated for the science community for other missions. The NASA suite of products also includes heritage products not being continued by the NOAA EDR (e.g., Particulate Inorganic Carbon). Also, the NASA L2 production capability can also easily include experimental products, which cannot be generated by the operational NOAA system.

Furthermore, with some adjustment to accommodate the NOAA EDR data products, the same NASA processing system provides the ability to generate global composites of NASA L2 and EDR data in Platte Carre projection, yielding what are referred to as Level-3 (L3) Standard Mapped Images (SMI), and in Integerized Sinusoidal Equal Area Grid (ISEAG) projection, creating what are called L3 binned data products. Both of these are crucial to understanding geographic and temporal trends of the data on synoptic and global scales. L3 production however required a careful look at data flags and masks to determine what data from the EDR or NASA L2 were to be aggregated into the L3 data products.

Once the existing NASA capabilities were augmented to support VIIRS, evaluation of the VIIRS flight data could begin. In the early mission assessment, the data was compared to mission averages from the Sea-viewing WIde Field-of-view Sensor (SeaWiFS) and to contemporaneous data from the MODIS aboard the Aqua spacecraft. Comparisons were done using L2 data to explore small-scale artifacts, beginning with striping patterns along the satellite track direction. In addition, L3 data was evaluated against SeaWiFS and MODIS to compare large-scale trending.

3.1 Masking and Flagging

The general state of the flags and masks used by the NOAA EDR data products and the NASA L2 evaluation products was examined. This analysis was used to : 1) identify which EDR flags are comparable to NASA's L2 flag set that would be necessary screen EDR data for inclusion into L3 products, and 2) determine which of those EDR flags performed their job properly. Analysis of EDR flags and masking was limited by the early state of the instrument calibration. In addition, the VIIRS cloud mask algorithm is subject to ongoing improvements to its performance during the early stage of the mission, which also limits the conclusions that can be drawn. Therefore, the observed performance of flags and masks must be continued throughout the first year of the mission, if not beyond, and any results presented here regarding masks should be considered preliminary.

Each pixel in an EDR or L2 product is associated with a group of digital bits called flags, each denoting the state of certain conditions that pertain to the quality of the pixel’s data value. There are 32 possible flags for the L2 product, 25 of which are directly ocean related, and there are 47 flags for the EDR, some with up to 8 states. A mapping of the EDR flags to the L2 flags based on their functionality and how well they currently operate is shown in Table A.1 and A.2 in the Appendix A. Instead of the geophysical quantity, each pixel also can contain a special numerical value indicating that it has been masked. These special numerical values, or fill values, indicate that a quantity could not be computed from the instrument measurements. A single fill value is used in the NASA L2 files, while there are 8 different numbers used in the NOAA EDR, each indicating some additional information about the failure to compute a value.

The analysis involved examination of flags and masks for several granules over open ocean and coastal regions. All granules showed similarities in flag performance. The L2 flags created by NASA algorithms all behave mostly as expected. Whereas this was often true for the EDR flags, many of the EDR flags were found to be extremely aggressive, appearing to extraneously mark more data as poor quality than expected. In fact, a significant portion of randomly scattered pixels were marked as poor quality for no apparent reason.

Most of the sporadic losses of data in the EDR data products appeared to be associated with false positive and negative indications of the presence of clouds, and probably reflect the currently immature status of associated flags. Some of the masked data also have flags set indicating confidently clear sky or occur over relatively clear water. In the clear water cases, many masked pixels can have coccolithophore or high DOM flags set. The density of the random data loss also appears to rise toward the end of scan with abrupt increases coincident with the 3:1 to 2:1 aggregation zone transition and likewise with the 2:2 to 1:1 transition. It was also observed that pixels corresponding to inland water are masked in the EDR data products. Thus, the EDR data products currently contain no geophysical values for areas like the Great Lakes of the USA. In addition, whenever any normalized water-leaving radiance becomes less than zero in a pixel, all EDR values associated with that pixel are masked. Therefore, in regions like the gyres, where the red band (672 nm) can

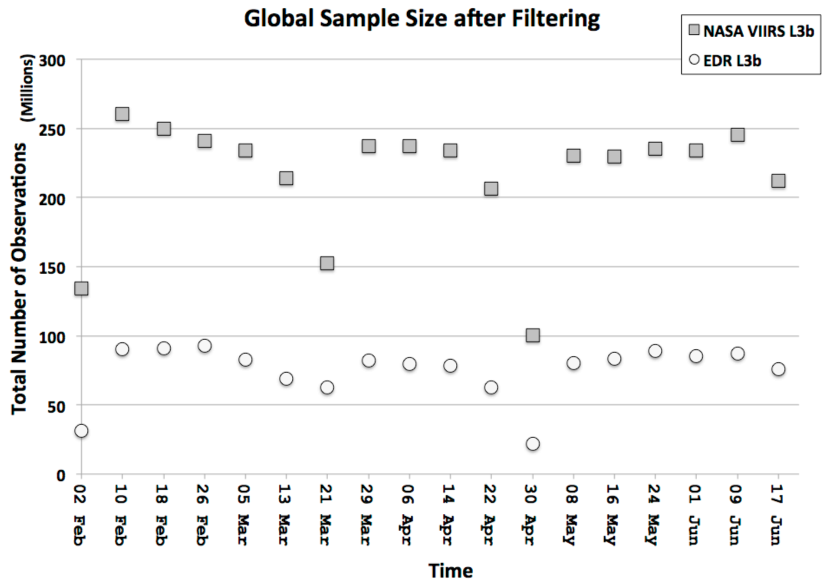

Figure 2. **Comparison of data sampling.** The number of observations within each eight-day L3 ISEAG bin are totaled over all bins common to both the EDR and NASA L2 products. The resulting time series covers the viable period for the EDR products, which were first based on a proper calibration on the 6th of February 2012.

become negative, valid chlorophyll *a* concentration and values in the other bands for normalized water-leaving radiances and IOP coefficients are discarded. This stems from a historical feature of the EDR algorithm. An earlier version of the algorithm required all bands to compute the IOP values and the chlorophyll *a* concentration using the Carder semi-analytic algorithm. NOAA has recognized both cases of extraneous data loss and requests to change this masking behavior in the EDR data products have been documented internally.

The EDR stray light flag is a significant contributor to the reduction in data in the EDR L3 products. A stray light flag appears in both the EDR and L2 products and is necessary to cull regions around clouds, land, and other bright targets when selecting data for aggregation into L3 products. NOAA EDR and NASA L2 algorithms set these flags differently. The NASA L2 stray light flag algorithm used for VIIRS is analogous to the one used for SeaWiFS and MODIS and appears to be performing as expected. Unfortunately, the EDR stray light flag removes a significantly larger amount of data in the clear regions around bright targets and many pixels spuriously are marked as contaminated with stray light. However, if the EDR version of the flag is not used, then too much contaminated data gets included in the corresponding L3 products. A similar situation arises with other EDR flags, but not as pronouncedly. For instance, the flagging of pixels with sensor zenith angle greater than 53° causes a significantly larger data cull than the 60° threshold used for the NASA L2 data products. It was found that the EDR flag that corresponds to the NASA L2 flag indicating a poor chlorophyll *a* data value (CHLWARN) was too conservative. In particular, chlorophyll *a* concentration can frequently fall below the EDR flag criterion of 0.05 mg m^{-3} in the gyres, thus it was not used in selecting EDR data for L3 data product generation. The combined effect of the more aggressive flagging criteria and masking results causes L3 products that are created from the NOAA EDR data to include a much smaller measurement sample compared to the L3 products based on NASA L2 data products.

There are many other EDR flags that were considered that do not have NASA L3 selection flag analogs. There are general EDR flags that correspond to the ocean color quality in the visible bands (410, 443, 486, 551, 672 nm). These all appear to flag significant amounts of the retrieved normalized water-leaving radiance in the EDR. The absorption (IOP_A) quality flags are relatively benign, but still can sporadically mark large regions as poor quality. The scattering (IOP_S) flag currently culls almost all the valid retrieved data in the EDR. Hence, both of these are not used in the generation EDR L3 products.

The land / water flag in the EDR was examined. The EDR land / water flag has four states: open sea, coastal water, inland water and land. The coastal water state affects a relatively small area, supposedly flagging waters immediately against coastlines. However, the marking of coastal water shows anomalies along the western coasts of land masses (e.g., the Californian coast). In such cases, the coastal water state frequently appears inside regions marked as land, making a landward progression of open water, land, coastal water, and then land again. Thus, the coastal water state appears to be offset eastward along scan from it's proper location. Other than this behavior, the land and other water conditions appear to be accurate.

In creating L3 data products from NASA L2 data, it was found that the standard NASA flag for indicating a problem in the atmospheric correction algorithm (ATMWARN) was not performing optimally and so was not used in forming L3 products. The NASA ATMWARN flag was tending to be triggered by the NIR radiance ratios falling out of the expected range (i.e., 0.85 – 1.35) was because of early difficulties with instrument calibration. The corresponding calibration issues were resolved, but this flag continued to be omitted to be consistent with the EDR product. Once the calibration improves, the NASA L3 data products can be reprocessed with this flag in use. The corresponding EDR flag, for similar reasons, was also not performing well, and was also omitted from L3 product generation. However, the EDR data products will never be reprocessed according to NOAA mission requirements.

Thus, in generating EDR L3 data products, only EDR flags corresponding to the standard suite of NASA L3 flags used with the exception of the ATMWARN and CHLORWARN. The current translation of the EDR flags to NASA L3 selection flags analogs appears to work reasonably. However, the number of observations included in the EDR L3 binned data was greatly reduced compared to the corresponding NASA L3 data (see Figure 2). Figure 2 also shows significant drops in data during outages that occurred in late March and late April 2012. Furthermore, the aggressive nature of the EDR flags and masks caused large areas to be excluded from global spatial coverage of the EDR data (as evident by the more widespread green regions in Figure 3b).

3.2 Striping

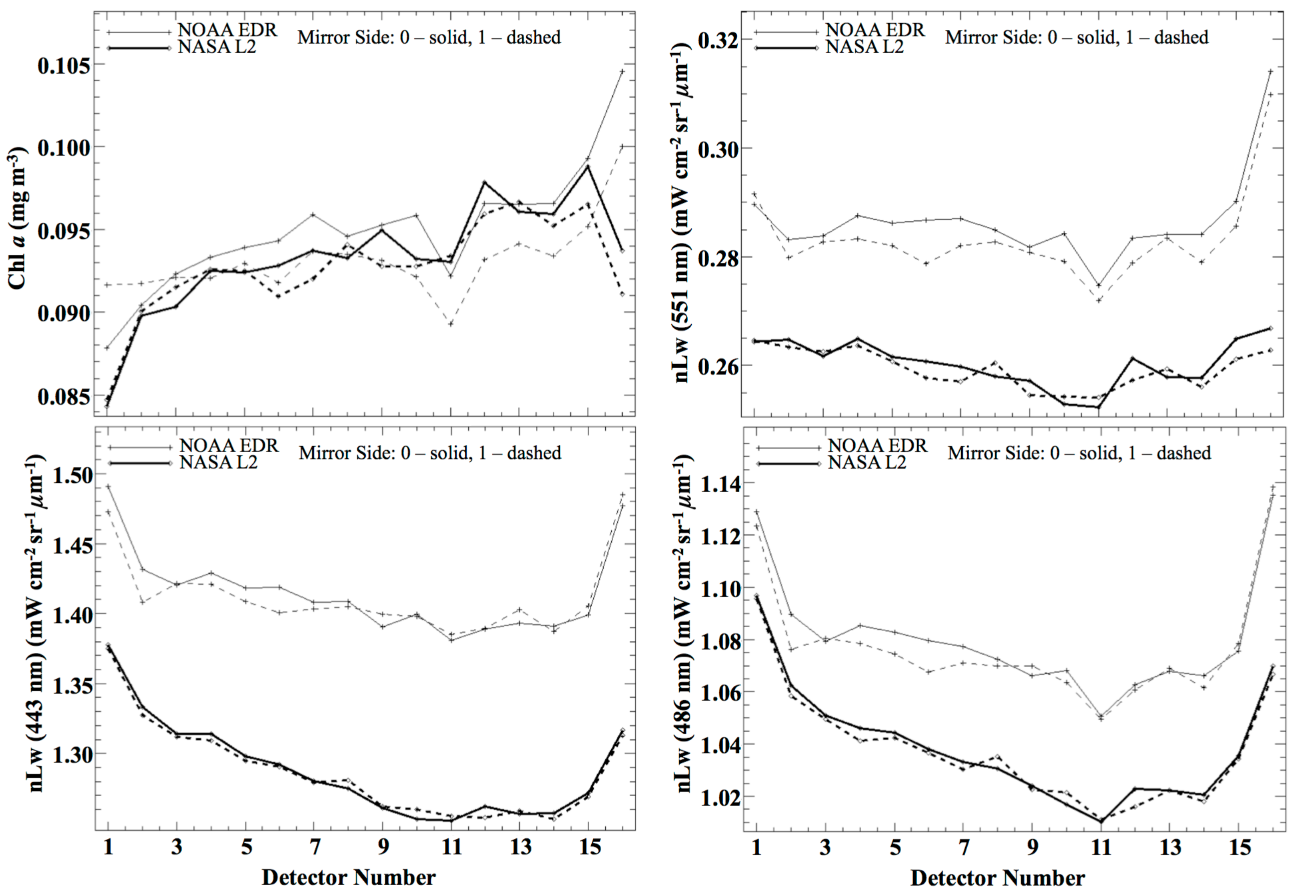

Figure 3. **Striping pattern for the EDR and NASA Level-2 data products.** These plots were based on data taken from a clear subregion of a satellite swath over the open ocean. The results are representative of several other similar cases taken from other granules at different times and location around the globe. Generally, the striping pattern is similar between the EDR and NASA L2, although there are still disagreement in the surface radiometry. The pronounced effects associated with the detectors near the edge of the scan is remarkable and unique to viewing the Earth.

Early evaluation also included looking at the quality of EDR L2 data products at the granule level. An area of interest for a push broom scanner like VIIRS is any systematic variation that can lead to scan-parallel striped patterns in the satellite image. VIIRS has sixteen detectors along the track direction of each scan in a satellite swath. Measurement biases at each detector and for each side of the instrument half-angle mirror leads to stripe patterns. The striping, when severe, can degrade the image quality and affect the comparison of satellite data to *in situ* data. Much of the striping stemming from biases in the prelaunch counts-to-radiance calibration is removed using the solar diffuser by both NASA[9] and NOAA. Furthermore, some striping in the top-of-the-atmosphere radiometry is expected, due to real variation in illumination geometry with each detector that is removed by the atmospheric correction algorithm. However, residual striping was observed in the derived products (e.g., chlorophyll *a* concentration). That these residual effects were not observed when viewing the solar calibration source, and occur primarily in the derived products, suggests other instrument effects may be the cause. Evaluation of potential candidates sources, such as variation of spectral or polarization response across detectors, will require further analysis.

Figure 3 shows statistical averages by detector of four derived products over a region of open ocean for a single, representative granule. These averages were taken within the 3:1 aggregation zones only to avoid bow-tie deletion effects that are characteristic of the VIIRS instrument design. Although the EDR and NASA L2 products do not entirely agree in radiometry due to the nascent calibration, they have similar amounts and patterns of striping in chlorophyll *a* concentration, which in turn comes from striping in the ratios derived from the normalized water-leaving radiance for the bands shown. The prominence of the effect at the edge detectors is remarkable, but the cause is not clear. The residual striping at this stage of the mission is still large compared to MODIS Aqua data and will need to be addressed to bring the VIIRS data product quality in line with heritage missions.

3.3 Level-3 Analysis

L3 binned data products for chlorophyll *a* concentration data taken over eight-day intervals from 2nd of February to the 25th of June 2012 were collected for large-scale trending analysis. The chlorophyll a concentration data quality is dependent on the quality of the retrieved water-leaving radiometry at 443, 486, and 551 nm. Separate L3 products derived from NOAA EDR products and NASA L2 products were processed into 9 km resolution ISEAG projections. For each eight-day period, a corresponding eight-day SeaWiFS L3 dataset was acquired at the same spatial resolution. The SeaWiFS L3 datasets contained twelve-year composites of all data for each eight-day period. A time series of these SeaWiFS L3 mission averages were used to determine roughly whether the EDR and NASA L2 contain reasonable values for their corresponding eight-day period and to facilitate a time series comparison that accounts for the seasonal signal. Figure 4 plots the common-bin difference between VIIRS L3 and SeaWiFS $\log_{10}$(chl *a*) verses the SeaWiFS values for one eight-day period, but is fairly representative of plots for other periods. The vertical dispersion in the Figure 4 plots shows the spread of differences as a function of SeaWiFS mission mean chlorophyll value. Generally, the differences seen in Figures 4a and 4b (EDR and NASA L2, respectively) are well within a factor of two, which is the accepted uncertainty of the chlorophyll *a* concentration algorithm[10]. However, in addition to estimation errors, much of the vertical spread in these scatter plots stems from real spatial and temporal variation with respect to the long-term average. Figures 4c shows the same spread in the comparison of MODIS Aqua chlorophyll *a* concentration to the corresponding SeaWiFS mission means. Difference scatter plots between VIIRS and MODIS Aqua, for the same eight-day period, show a much tighter spread because more small-scaled features are matched. Interestingly, for the period shown in Figure 4, the EDR and NASA L2 show little bias, except towards the extremes, but as shown later, this can change with time. Although some of the evident overestimate of low values and underestimate of high values could come from errors in the maturing calibration, it is also expected that the SeaWiFS mission mean would also wash out some of the extremes that would be expected in an eight-day composite. Finally, although both NOAA and NASA VIIRS datasets have low biases with respect to the SeaWiFS mission mean, it is notable that the MODIS Aqua chlorophyll *a* concentration is globally lower than SeaWiFS or VIIRS. This may indicate errors in the MODIS Aqua, for which the time series is know to suffer increased uncertainties in the late mission period.

Figure 5 shows the spatial distribution of the relative chlorophyll *a* concentration differences between the VIIRS L3 SMI datasets and the MODIS Aqua datasets. Several observations can be made regarding this plot. Generally, over most of the ocean, the differences are relatively small, with slightly higher values for VIIRS as suggested by the Figure 4 plots. In both cases, the largest positive differences (VIIRS > MODIS) are spread sporadically across the northern Pacific Ocean and the largest negative difference (VIIRS < MODIS) are most notable in the northeastern Atlantic Ocean, North Sea, and Baltic Sea. The cause of these localized differences is not yet clear, but it is consistent with the tendency to underestimate high values and overestimate low values. It is also remarkable that the greatest differences are predominantly in the northern hemisphere, while vast areas of the Southern Ocean have very low differences. Finally,

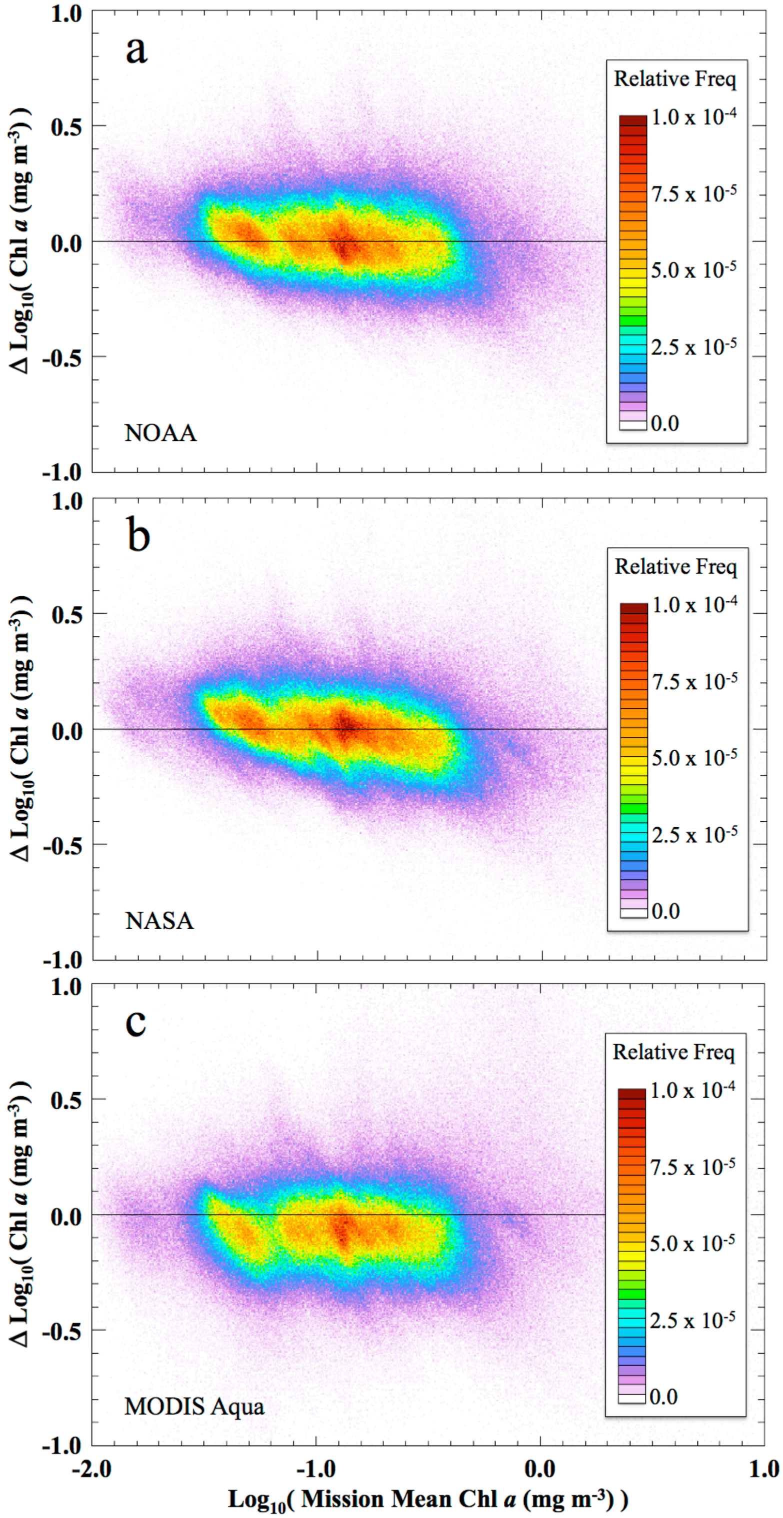

Figure 4. **Difference from SeaWifS mission mean chlorophyll *a***. Differences from the SeaWiFS L3 mission mean for the corresponding eight-day periods were taken for L3 ISEAG bins common to L3 products derived from the a) NOAA EDR, b) NASA L2, and c) MODIS Aqua. In this plot, the eight-day period beginning with the 21st of March 2012 is shown.

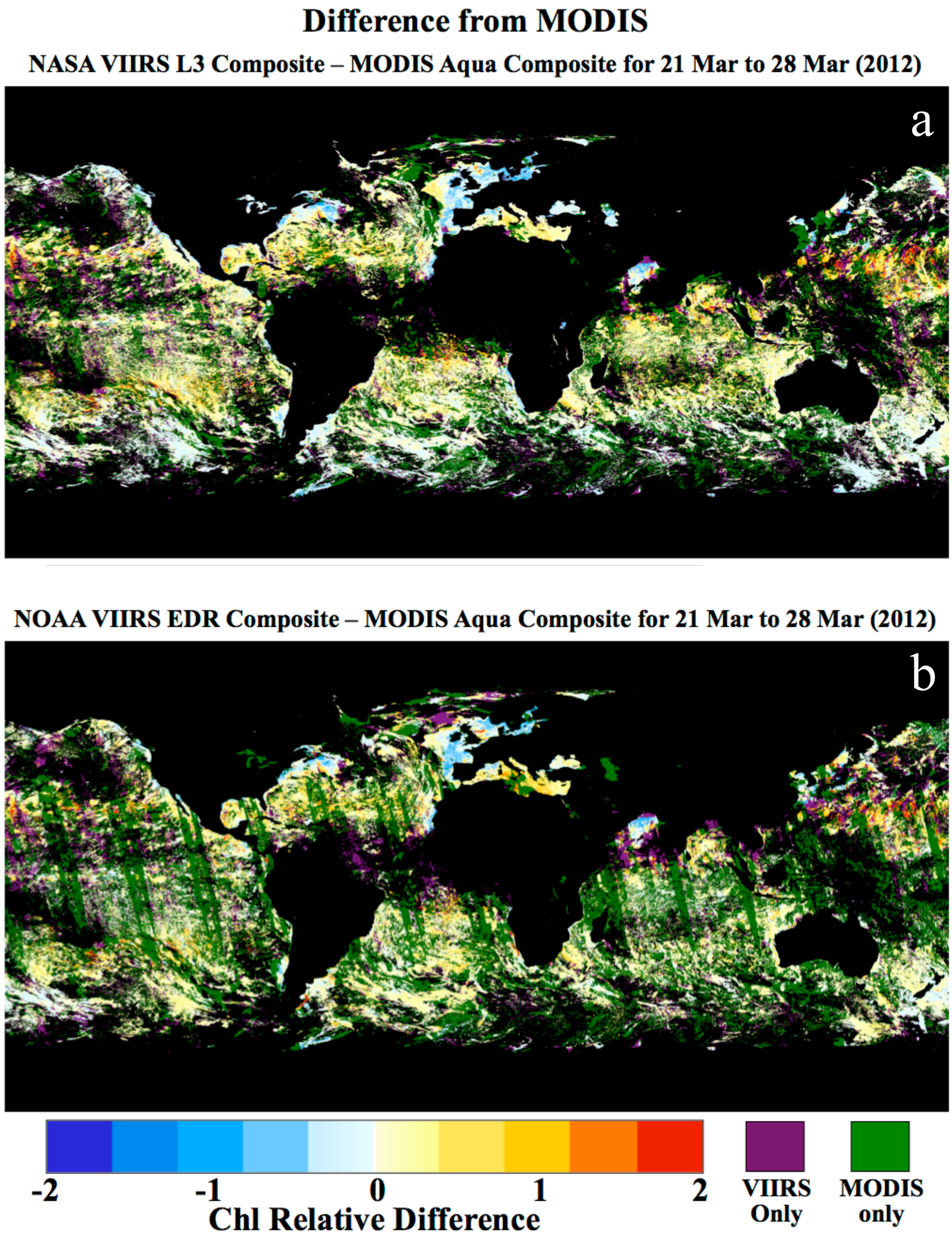

Figure 5. **Maps of relative difference.** These maps are derived from composites of data projected into L3 SMI products for an eight-day composite starting on the 21st of March 2012. The results show the distribution of the relative difference between chlorophyll *a* concentration as estimated using VIIRS and MODIS Aqua data. The VIIRS data was based on a) NASA L2 data and b) based on NOAA EDR data. The magenta areas indicate where there was VIIRS values present, but not MODIS data. Green areas indicate where the was no VIIRS data, but MODIS data was present. Black indicates the absence of chlorophyll *a* concentration data.

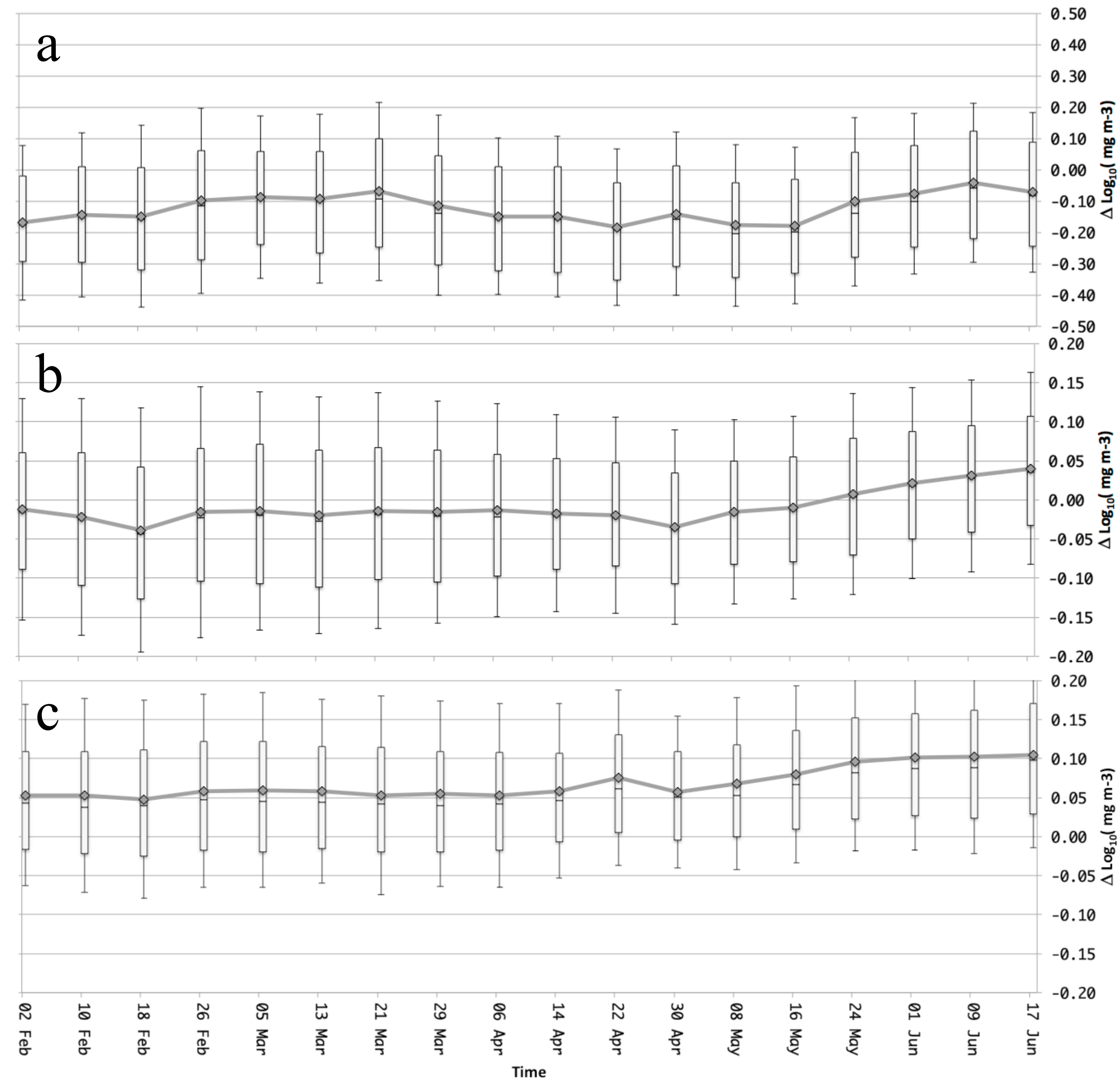

Figure 6. **Time series of the NASA L3 difference from SeaWiFS mission mean.** The time axis gives the start date for each eight-day period. The broken-line gives the mean bin difference of $\log_{10}$(Chl *a*) between the NASA L3 and SeaWiFS L3 mission mean for common bins that correspond the SeaWiFS mission mean at a) 0.05, b) 0.2, and c) 1.0 mg m^{-3}. The boxes give the first quartile, median, and third quartile, and the whiskers plot the standard deviation across all common bin differences.

Figure 5 also shows the masking applied to the EDR greatly affects its spatial sampling of the globe, especially at lower latitudes. Emphasis of higher latitudes can cause the variation in global scale averages to be less stable.

The differences from the SeaWiFS mission means shown in Figure 4b were subdivided into intervals of 0.1 log units and plotted as times series. Figure 6 shows plots for SeaWiFS chlorophyll *a* value intervals around 0.05, 0.2, and 1.0 mg m^{-3}. As before, the high values tend to be underestimated while low values are overestimated. The trends are fairly steady except for some slight variation near data outages (e.g., late April) and a gradual upward trend from late April through late June that may be a real geophysical difference between the 2012 seasonal cycle and the SeaWiFS climatological signal. The effect of the outages emphasizes the importance of data spatial coverage and sampling to large-scale temporal statistics. Figure 7 gives the same kind of plot for the EDR differences, but the traces appear slightly more erratic. At least two possible explanations can be considered for further analysis. First, the lack of

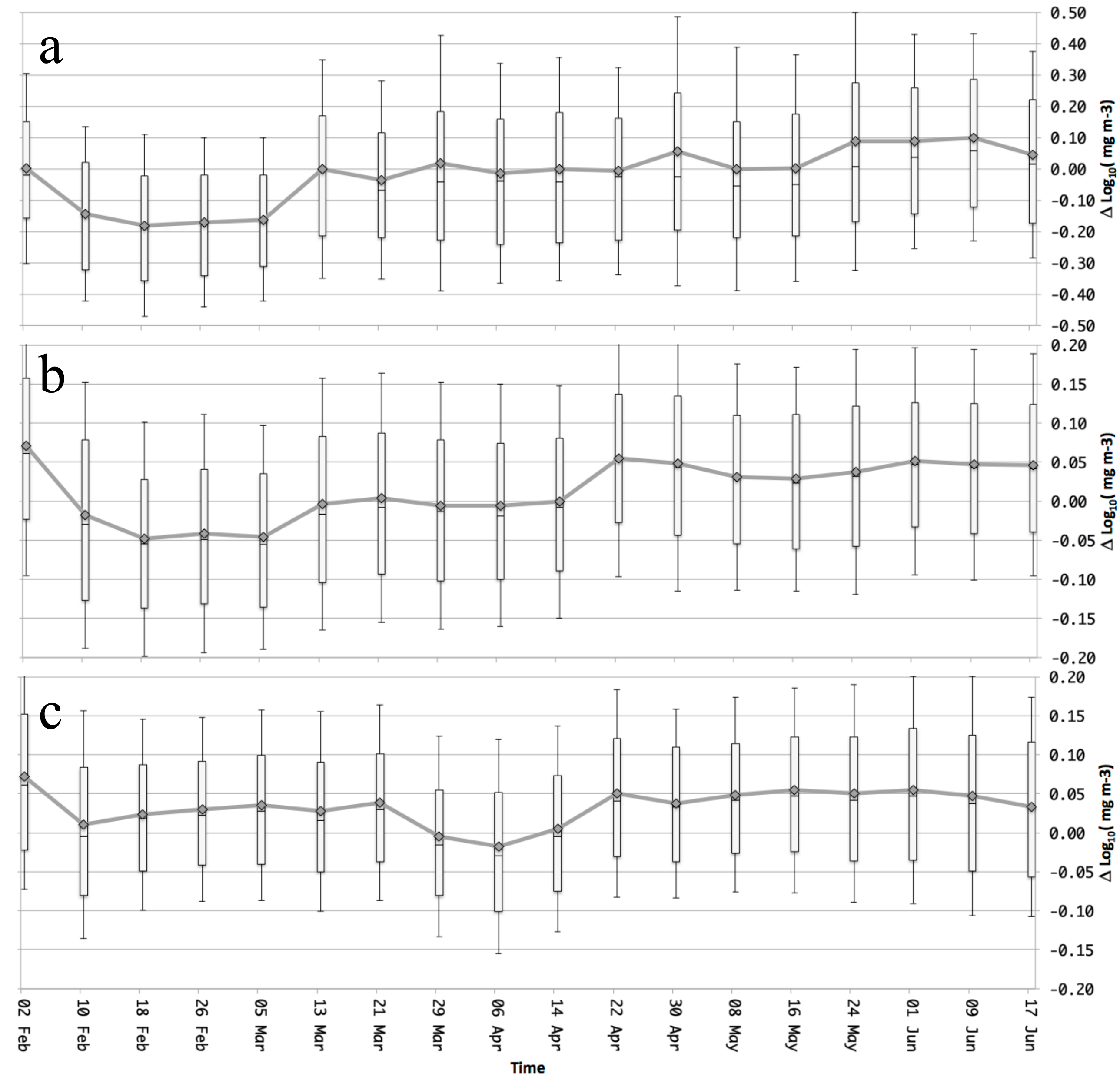

Figure 7. **Time series of the EDR L3 difference from SeaWiFS mission mean.** This is the same as Figure 6, but pertains to the EDR L3 data products. This plot illustrates the larger temporal variation, which may stem from variations in the NOAA calibration or in the sparser sampling within bins of the EDR caused by aggressive masking and flags.

reprocessing causes any artifact stemming from the developing NOAA calibration to become a permanent feature in the EDR series, unlike the NASA L2 and L3 evaluation data products. Another possible explanation is that the larger variation stems from a poorer, more variable sampling within bins, which is a byproduct of the aggressive masking associated with the EDR. These explanations are not mutually exclusive and both situations could be possible contributors to the observed difference in temporal variation. More importantly, both situations would lead to permanent artifacts in the EDR time series. This further supports the argument that the NOAA EDR will not be suitable as a good climate record because of the lack of reprocessing. A more direct comparison between the NASA evaluation and the NOAA EDR products is illustrated in Figure 8. The plot shows chlorophyll *a* concentration averages over the deep-water region of the open ocean (i.e., waters with depth > 1000 m) for both the EDR and NASA evaluation data for each 8-day L3 composite. The SeaWiFS mission means for the corresponding 8-day time periods are also shown and the range of the SeaWiFS 8-Day composition mean over the entire mission (about 0.15 to 0.22 mg m^{-3}) are indicated by the gray region. Clearly, the EDR mean is greatly out of family with the record established over a dozen years with the

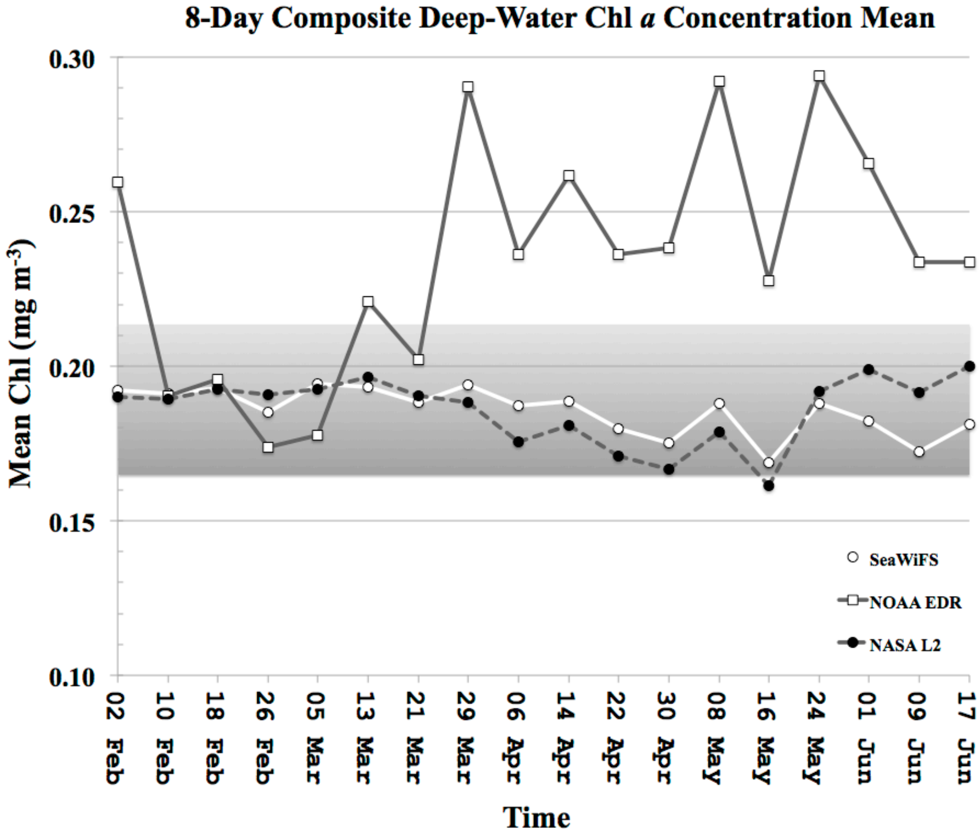

Figure 8. **Chlorophyll Mean Time Series.** Mean chlorophyll was averaged for each 8-day composite over waters with depth > 1000 m. The top and bottom of the gray region on the plot give the mission min and max of this statistic over the entire SeaWiFS mission. The SeaWiFS curve corresponds to the mission mean spatially averaged ove the deep water region.

SeaWiFS mission and with the current the NASA VIIRS evaluation product. There is no plan to improve the EDR through reprocessing, and so it is likely to remain as shown.

4. SUMMARY

An early mission assessment of the Suomi NPP VIIRS ocean color data products was carried out by the VOST. The instrument does seem to be performing well and that is reflected in the quality of data being retrieved. Although the instrument response and solar calibration system are degrading at a phenomenal rate at some wavelengths, most of this behavior is trackable and appears correctable. To incorporate this knowledge into a data production system that uses heritage algorithms and could reprocess the data, the VOST set up an independent calibration capability that feeds directly into the existing NASA ocean color processing computational infrastructure. Using the existing system with heritage NASA algorithms facilitates creation of VIIRS data products that are consistent with the CDR that was established with heritage NASA missions. Moreover, this will provide an opportunity to examine data products of interest to the science community, particularly those that were generated for heritage missions, but are not being carried by the current NOAA EDR ocean color product suite, or new data products that cannot currently be produced by the NOAA operational data processing system. Finally, minor augmentations were made to extend the existing capability to also generate L3 data products from NOAA EDR products. Generation of the L3 products is a key tool to perform large-scale spatial or temporal trending of satellite data products and is a capability missing in the NOAA operation data processing system.

To create L3 products for the EDR data products required use of EDR flags and was subject to EDR masking. The over aggressive nature of these two data filters appears to have substantially reduced the data sample available for analysis or inclusion into L3 products. Too many data points are marked as poor quality or simply masked out of existence when the EDR data products are made. The former case leads to the conundrum of whether to include all data marked by a flag as poor quality, adding noise to the L3 data product, or to exclude the data causing a substantial drop in data sample and coverage. The VOST develop a suite of flags to create L3 products from NOAA EDR products that appear to be reasonable for the early mission. However, this subsequent data loss may have affected the stability of regional and global averages for the EDR L3 products.

Examination of the derived products indicate that both the NOAA EDR and NASA L2 data products continue to mature. Some products show greater striping than seen in MODIS Aqua data. Also, the striping patterns between the NOAA EDR and NASA L2 are very similar. This artifact will need to be addressed to bring the VIIRS ocean color data quality fully in line with the heritage. Chlorophyll *a* concentration L3 data for both the NOAA EDR operational products and the NASA L2 evaluation products appear comparable in quantity, and roughly in range, to the corresponding SeaWiFS mission means. They also closely match MODIS Aqua in eight-day composites spatially, especially in the southern hemisphere. However, much of the high concentrations tend to be underestimated and low concentrations tend to be overestimated. These likely stem for biases in the early calibration, which can be refined with time. For instance, the refinement of the current vicarious calibration will likely be achieved through collection of more data within the next one to three years.

Monitoring of the L3 time series will continue through the first year of the Suomi NPP mission and may be expanded to include NASA evaluation products not found in the NOAA EDR, but are of interest to the NASA science community (e.g., Particulate Inorganic Carbon, Particulate Organic Carbon, and the Attenuation Coefficient at 490 nm). The VOST will continue to work to refine the independent calibration of VIIRS. Further evaluation of VIIRS instrument behavior may involve looking at second order instrument effects, including monitoring of the signal-to-noise ratio, especially in the NIR and SWIR bands, which are rapidly losing responsivity. In addition, sensor behavior leading to residual striping and scan effects, which could stem from errors in the characterization of spectral, response-verses-scan (RVS), and polarization responses, will be examined as more flight data is collected and calibration of the instrument improves.

APPENDIX A. Matching NASA L2 and NOAA EDR flags

This Appendix provides details regarding how the NOAA EDR data product flags can be related to the NASA L2 product flags. Identification of this relationship is a necessary step in creating L3 data products from NOAA EDR products that are analogous to the NASA L3 data products. The Table A.1 provide a map that matches the two sets of flags. In Table A.1, unless noted otherwise, a value of 0 for any flag bit indicates a condition that is necessary, but perhaps not sufficient, for good quality (e.g., a datum is within a valid range or a calculation succeeded). Triangles at the top of the NASA L2 columns indicate flags used for L2 masking and L3 data selection. If a flag indicated by a black triangle is set, then the a fill value is using for the corresponding NASA L2 pixel. If any of the flags indicated by triangles are set, then the data is not used in the L3 product. The EDR flag corresponding to the NASA L2 CHLWARN flag is not currently used to produce EDR L3 products because the EDR algorithm has too high of a lower bound for its valid range criterion. This ATMWARN flag and its EDR counterpart are not used because the current state of the calibration is pushing the NIR band ratios out of range. For additional information, see corresponding notes in Table A.2 as indicated by the column below labeled NOTES.

Table A.1 **Map of NOAA EDR data product flags to NASA L2 flags.** Gray columns indicate the standard set of flags used by NASA to mask or cull data in L2 or L3. See text and Table A.2 for further details.

NASA Standard L2 Flags

NOAA EDR Flags

EDR flag name	Bit #	Description	NOTES	ATMFAIL	LAND	PRODWARN	HIGLINT	HILT	HISATZEN	COASTZ	SPARE_8	STRAYLIGHT	CLDICE	COCCOLITH	TURBIDW	HISOLZEN	SPARE_14	LOWLW	CHLFAIL	NAVWARN	ABSAER	SPARE_19	MAXAERITER	MODGLINT	CHLWARN	ATMWARN	SPARE_24	SEAICE	NAVFAIL	FILTER	SSTWARN	SSTFAIL	HIPOL	PRODFAIL	SPARE_32
				▼	▼		▽	▼	▽			▽	▼	▽		▽		▽	▽	▽			▽		▽				▽	▽					
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
QF1_VIIRSOCCEDR	0	412 nm OC quality	A																																
	1	445 nm OC quality	A																																
	2	488 nm OC quality	A																																
	3	555 nm OC quality	A																																
	4	672 nm OC quality	A																																
	5	CHL-a quality	B																																
	6	IOP_a 412 quality	A																																
	7	IOP_s 412 quality	A																																
QF2_VIIRSOCCEDR	0	IOP_a 445 quality	A																																
	1	IOP_s 445 quality	A																																
	2	IOP_a 488 quality	A																																
	3	IOP_s 488 quality	A																																
	4	IOP_a 555 quality	A																																
	5	IOP_s 555 quality	A																																
	6	IOP_a 672 quality	A																																
	7	IOP_s 672 quality	A																																
QF3_VIIRSOCCEDR	0	SDR M1-7 quality																																	
	1	O_3 quality																																	
	2	Wind speed > 8 m/s	G																																
	3	Epsilon out of range	C																						X									
	4	Atmospheric correction success (0 = all good), otherwise, step that failed		X																															
	5																																		
	6																																		
	7	Spare																																	
QF4_VIIRSOCCEDR	0	Land / water (0-2 = water, 3 = land)	Q		X																														
	1																																		
	2	Snow / ice											X																						
	3	Solz > 70 deg														X																			
	4	Sun glint					X																												
	5	Sensor zenith angle > 53° or pixel too wide	D						X																										
	6	Shallow water (< 50m)								X																									
	7	Spare																																	
QF5_VIIRSOCCEDR	0	Cloud confidence (0 = conf. Clear)	E										X																						
	1																																		
	2	Adjacent pix not clear	F									X																							
	3	Cirrus	I																																
	4	Cloud shadow	J															X																	
	5	Heavy aerosol	K																																
	6	Absorbing aerosol	L																		X														
	7	AOT > 0.3																																	
QF6_VIIRSOCCEDR	0	Turbid water	M												X																				
	1	Coccolithophores	N											X																					
	2	High DOM	O																																
	3	Chlorophyll *a* range (0 = none retrieved)	P																X																
	4																																		
	5																																		
	6	Carder algorithm branching.																																	
	7																																		
QF7_VIIRSOCCEDR	0	Normalize water-leaving radiance is out of range																																	
	1	Chlorophyll *a* concentration out of range	R																						X										
	2	IOP_a in one or more bands is out of range																																	
	3	IOP_s in one or more bands is out of range																																	
	4	input SST quality	H																																
	5	Bright target										X																							
	6	Spare																																	
	7	Spare																																	
Other		Bad lat, lon																												X						

Table A.2 **Notes corresponding to Table A.1.** Letters in the 1st column correspond to items so marked in the 2nd column of Table A.

	NOTES
A	The ocean color and IOP quality are poor if : - turbid water is present, - normalize water-leaving radiance for that band is out-of-range, or - coccolithophorids are present.
B	Chlorophyll *a* concentration poor quality flag is set if : - the computed value is out-of-range, - any IOP is out-of-range (N/A?), or - coccolithophores are present.
C	Data is flagged as poor quality if the NIR band ratio (epsilon) is outside of the range of 0.85 – 1.35.
D	This flag is also set to 1 if the horizontal cell size > 1.3 km.
E	Cloud states are : - 0–confidently clear - 1–probably clear - 2–probably cloudy -3–confidently cloudy.
F	If any adjacent pixel is anything less than confidently clear, this flag is set to 1.
G	Note that the NASA atmospheric correction algorithm sets the whitecap radiance to a constant when windspeed exceeds 12 km/sec, but no corresponding flag is set by that algorithm in the NASA L2 product.
H	Fraction of light into or out of a pixel is above threshold.
I	The cirrus flag is determined by tests with the SWIR band M9 and with TIR bands M15 and M16
J	The cloud shadow flag is geometry-based. The TIR band, M15, and atmospheric profile are used to get a cloud top height at edge. This, solar and spacecraft geometry are used to find shadow pixels. The edge is expanded by 1 pixel beyond this.
K	Heavy aerosol is determined over ocean with tests on M1+5, M11+1 (spectral tests) and on M5 (spatial test)
L	The presence of absorbing aerosol, which can undermine the atmospheric correction quality, is indicated when the single scattering albedo in the 551 nm band (M4) < 0.7.
M	Turbid water is indicated when remote sensing reflectance at 551 nm is greater than 0.012.
N	The presence of coccolithophores are indicated when all of the follow criteria are true : - nLw(443 nm) $\geq$ 1.1 W m^{-2} sr^{-1} nm^{-1}, - nLw(551 nm) $\geq$ 0.81 W m^{-2} sr^{-1} nm^{-1}, - La(748 nm) < 1.1 W m^{-2} sr^{-1} nm^{-1}, and - 0.6 $\leq$ nLw(443 nm)/nLw(551 nm) $\leq$ 1.1.
O	High concentration of Dissolved Organic Matter is indicated when IOP_a(410 nm) > 2 m^{-1}.
P	Chlorophyll *a* concentration states: - 0 – no Chl *a* concentration computed - 1 – Chl *a* < 1. mg m^{-3} - 2 – 1. < Chl *a* < 10 - 3 – Chl *a* > 10 mg m^{-3}.
Q	Water mask states: - 0 – sea water - 1 – coastal - 2 – inland - 3 – land.
R	Good chlorophyll *a* concentration range defined by the range 0.05 → 50 mg m^{-3}.

REFERENCES

[1] Carder, K.L., Chen, F.R., Lee, Z.P., Hawes, S.K., "Semianalytic Moderate- Resolution Imaging Spectrometer algorithms for chlorophyll *a* and absorption with bio- optical domains based on nitrate-depletion temperatures," Journal of Geophysical Research 104(C3), pp 5403–5421 (1999).

[2] O'Reilly, J.E., Maritorena, S., Mitchell, B.G., Siegel, D.A., Carder, K.L., Garver, S.A., Kahru, M., McClain, C.R., "Ocean color algorithms for SeaWiFS," Journal of Geophysical Research 103(C11), pp 24937-24953 (1998).

[3] Turpie, K.R., Meister, G., Eplee, G., Barnes, R.A., Franz, B., Patt, F.S., Robinson, W.D., McClain, C.R., "Assessment of NPP VIIRS ocean color data products: Hope and risk," Proc. of SPIE Vol. 8153 81530M0-1, Earth Observing Systems XVI, Butler, J.J., Xiong, X., Gu, X., eds., San Diego, CA. 9 pp (2011).

[4] Raytheon, (2009a). "Performance verification report – VIIRS FU1 Relative spectral response. Performance Verification Plan," PVP154640-101, Section 4.6. 204 pp.

[5] Raytheon, (2009b). "Performance verification report – VIIRS FU1 Crosstalk. Performance Verification Plan," PVP154640-101, Section 4.7.2. 162 pp.

[6] Robinson, W.D., Patt, F.S., Franz, B.A., Turpie, K.R., McClain, C.R., (2009). "A VIIRS ocean data simulator," *Proceedings of the SPIE*, 7452(3-5), 3-5 Aug, San Diego. DOI: 10.1117/12.827058.

[7] Gordon, H.R., (1995). "Remote sensing of ocean color: a methodology for dealing with broad spectral bands and significant out-of-band response," Applied Optics, 34(36), pp 8363-8374.

[8] Brown, S.W., Eppeldauer, G.P., and Lykke, K.R., (2006). "Facility for spectral irradiance and radiance responsivity calibrations using uniform sources," Applied Optics, 45(32), pp 8218-8237.

[9] Eplee, R.E., Turpie, K.R., Fireman, G.,F., Meister, G., Stone,T.C., Patt, F.S., Franz, B.A., Bailey, S.W., Robinson, W.D., McClain, C.R., "VIIRS On-Orbit Calibration for Ocean Color Data Processing," in Earth Observing Systems XVII J.J. Butler, X Xiong, and X. Gu, eds., Proc. SPIE 8510, 85101G, (2012).

[10] Campbell, J., Antoine, D., Armstrong, R., Arrigo, K., Balch, W., Barber, R., Behrenfeld, M, Bidigare, R., Bishop, J., Carr, M-E., Esaias, W., Falkowski, P., Hoepffner, N., Iverson, R., Kiefer, D., Lohrenz, S., Marra, J., Morel, A., Ryan, J., Vedemikov, V., Water, K., Yentsch, C., Yoder, J., (2002). "Comparison of algorithms for estimating ocean primary production from surface chlorophyll, temperature, and irradiance," Global Biogeochemical Cycles, 16(3). pp 9.1-9.12.