Early evaluation of the VIIRS calibration, cloud mask and surface reflectance Earth data records

Eric Vermote a,⁎, Chris Justice b, Ivan Csiszar c

a NASA Goddard Space Flight Center, Terrestrial Information Systems Laboratory, Mail Code 619, Bldg 32, 5036H, Greenbelt, MD 20771, USA
b University of Maryland, Dept. of Geographical Sciences, USA
c NOAA/NESDIS, Center for Satellite Applications and Research, USA

Abstract

Surface reflectance is one of the key products from VIIRS and as with MODIS, is used in developing several higher-order land products. The VIIRS Surface Reflectance (SR) Intermediate Product (IP) is based on the heritage MODIS Collection 5 product (Vermote, El Saleous, & Justice, 2002). The quality and character of surface reflectance depend on the accuracy of the VIIRS Cloud Mask (VCM), the aerosol algorithms and the adequate calibration of the sensor. The focus of this paper is the early evaluation of the VIIRS SR product in the context of the maturity of the operational processing system, the Interface Data Processing System (IDPS). After a brief introduction, the paper presents the calibration performance and the role of the surface reflectance in calibration monitoring. The analysis of the performance of the cloud mask with a focus on vegetation monitoring (no snow conditions) shows typical problems over bright surfaces and high elevation sites. Also discussed is the performance of the aerosol input used in the atmospheric correction and in particular the artifacts generated by the use of the Navy Aerosol Analysis and Prediction System. Early quantitative results of the performance of the SR product over the AERONET sites show that with the few adjustments recommended, the accuracy is within the threshold specifications. The analysis of the adequacy of the SR product (Land PEATE adjusted version) in applications of societal benefits is then presented. We conclude with a set of recommendations to ensure consistency and continuity of the JPPS mission with the MODIS Land Climate Data Record.

1. Introduction

The directional surface reflectance product is a critical input for generating Vegetation Indices (VI), Leaf Area Index (LAI), Fraction of Absorbed Photosynthetically Active Radiation (FPAR), Bidirectional Reflectance Distribution Function (BRDF), Albedo, and Land Cover. The directional surface reflectance is obtained after the correction of the various atmospheric effects that affect the top of the atmosphere signal. The scattering by molecules and aerosols in the atmosphere is clearly evident in the comparison of top of the atmosphere and corrected reflectance in the visible part of the Electromagnetic Spectrum, as seen in Fig. 1. The first requirement for accurate atmospheric correction is a proper absolute calibration of the instrument. Calibration errors propagate through the whole atmospheric correction chain, in particular through the aerosol inversion and impact most of the bands in the visible part of the Spectrum, and will impact subsequent downstream products. It is very important therefore for the Surface Reflectance product team to keep assessing the instrument performance and independently monitoring the calibration. The procedure we have been using in the past for AVHRR NOAA16 and MODIS Aqua/Terra has been implemented for VIIRS, so that a continuous near real-time monitoring of VIIRS/Aqua cross-calibration can be achieved (Section 2).

One of the aspects which are often underestimated in generating a satellite derived product is the correct flagging of data that are suboptimal to enable users to include data quality considerations in their analysis (Roy et al., 2002). For surface reflectance no reliable retrieval can be done in the presence of cloud, cloud shadow, high aerosol or at high solar zenith angles. Although the detection of high aerosol is in some way straightforward (although high aerosol could be confused with clouds), the validation of cloud or cloud shadow detection algorithm remains quite challenging, although the availability of high quality remote sensing derived cloud information from CALIPSO (Winker, Pelon, & McCormick, 2003) with global coverage (although limited to a swath of 5 km) makes the assessment less problematic. We will present in Section 3 our evaluation of the VIIRS Cloud Mask, focusing on snow-free land by comparison to both CALIPSO data and the Aqua Cloud Mask.

Quantitative information on aerosols is a critical input to the atmospheric correction and by far the one that introduces the most uncertainties in the SR product. Both aerosol optical thickness and the aerosol model (size distribution, multispectral real and imaginary refractive indices and non-sphericity) have to be determined to perform accurate atmospheric correction. There is also some limitation to the

⁎ Corresponding author. Tel.: +1 301 614 5413; fax: +1 301 614 5269.
E-mail address: Eric.F.Vermote@nasa.gov (E. Vermote).
aerosol retrieval algorithm (over snow and bright surfaces) that necessitates the use of default values for correction. For MODIS, those values are fixed (aerosol optical thickness of 0.05, urban clean aerosol model), but for the VIIRS operational production system, they are currently read from the output of an aerosol transport model, the Navy Aerosol Analysis and Prediction System (NAAPS), the accuracy, spatial and temporal resolution of which are shown in this study to be inadequate for the purpose of atmospheric correction over land and cause obvious artifacts in the current SR product (Section 4).

One critical task in developing a data product is providing an accuracy assessment. The directional surface reflectance from MODIS has been validated through the various stage defined by the MODIS land validation approach (up to Stage III) (Morisette, Privette, & Justice, 2002). In particular, careful attention has been paid to the accuracy of the 6S radiative transfer code (Kotchenova & Vermote, 2007; Kotchenova, Vermote, Levy, & Lyapustin, 2008; Kotchenova, Vermote, Matarrese, & Klemm, 2006), used to correct atmospheric effects by a comprehensive comparison to a Monte Carlo model (Bréon, 1992). Both the Aqua and Terra Collection 5 Surface Reflectance products have been extensively compared to a reference reflectance dataset obtained for the AERONET sites (Holben et al., 1998), which gives ancillary inputs for atmospheric correction such as water vapor, aerosol optical thickness, size distribution and refractive indices. A preliminary assessment for Terra using data from 2003 was published (Vermote & Kotchenova, 2008), showing that the accuracy of the Surface Reflectance product was satisfactory except for band 3 (470 nm) which is used for aerosol inversion. The analysis has been completed for both Terra and Aqua for the 2000 to 2009 period. In summary, the study used c. 50,000 matchups between AERONET and MODIS, spread out over 200 locations, which translates to analysis of more than a million individual pixels. Using this robust dataset, the performance of the MODIS Collection 5 product has been confirmed to be well within the theoretical error bars: ± (0.01 + 10%) for threshold and: ± (0.005 + 5%) for objective. Using this protocol, we have derived the performance metrics of the VIIRS Surface Reflectance and compared them to the equivalent MODIS metrics (Section 5).

The performance of the surface reflectance is an important metric for a broad variety of science users as well as for long term applications of societal benefits. In this context, the MODIS daily climate modeling

Fig. 1. The VIIRS first light image acquired on November 21, 2011 in the visible bands; top is the top of the atmosphere (RGB) image and bottom is the prototype VIIRS corrected surface reflectance product.
of US agriculture and global commodity prices. Such application, namely the readiness of the VIIRS product to assess the agricultural drought of July 2012, which occurred over most of the Northern Hemisphere and had an important impact on several sectors of US agriculture and global commodity prices.

In Section 7, we will focus on improvements needed to the VIIRS surface reflectance algorithm and upstream products that are critical to ensure MODIS Land Climate Data Record continuity. For more details on these issues we refer the reader to Justice et al. (2013). Some of these recommendations have already been implemented in the Land Product Evaluation and Analysis Tool Element’s (PEATE) system (http://landweb.nasa.gov/NPP_QA) (Read et al., 2007) and have demonstrated their usefulness but have yet to be implemented in NOAA’s near-real time Interface Data Processing Segment (IDPS), which provides data for archive and distribution to the NOAA’s Comprehensive Large Array-data Stewardship System (CLASS), the data portal available to VIIRS data users.

2. VIIRS calibration monitoring

Accurate radiometric calibration is a prerequisite to creating a science-quality, time-series of surface reflectance and consequently, higher order downstream products. Calibration errors can propagate directly into the surface reflectance and create artificial variations that can be misinterpreted as trends, especially if these variations are due to a slow decay in the calibration mechanism. Vicarious calibration provides an additional source of calibration information, to verify and evaluate on-board calibration. We are using the approach we developed for cross calibration of MODIS (Vermote & Saleous, 2006a) to monitor the calibration in the visible to shortwave infrared bands and to provide correction terms as needed. The approach relies on using the multi-year Terra MODIS data set to derive spectral and directional characterizations of stable desert sites that can be used as invariant targets. A candidate list of such targets is provided in Cosneau, Leroy, and Briottet (1996). Subsets of Terra MODIS data are collected and undergo a rigorous screening based on the quality flags (no cloud, cloud shadow, adjacent cloud, high aerosol or snow). The spectral characterization is obtained by averaging the surface reflectance for each site and spectral band over the acquisition period. The directional characterization is derived using the MODIS Bidirectional Reflectance Distribution Function (BRDF) algorithm that relies on a kernel-driven linear BRDF model defined as a weighted sum of an isotropic parameter, RossThick and LiSparseReciprocal kernels (Schaaf et al., 2002).

Using the site directional characterization, we compute a MODIS-like surface reflectance at the MODIS acquisition time. Atmospheric parameters (surface pressure, gaseous content, water vapor, aerosol optical thickness) obtained from assimilated data, MODIS data, MODIS-like and/or ground measurements are then used in conjunction with the 6S radiative transfer code (Vermote, Tanré, Deuze, Herman, & Morcrette, 1997) to determine the MODIS Top Of Atmosphere (TOA) reflectance. The computed reflectance is compared to the acquired reflectance to infer changes in the instrument calibration. To assess this approach, we applied it to transfer the Terra MODIS calibration to the Aqua MODIS instrument. When applied to a stable ground site in Niger, the results of this approach agreed to within 1% of the Aqua MODIS on-board solar diffuser (Vermote & Saleous, 2006b). We also have applied this method to assess the calibration of ASTER (which acquired data in the same geometry as Terra) and VIIRS. For ASTER, the method has enabled us to highlight a problem with the saturation of the red band found for the desert site (Fig. 2, left) as well as issues with the on-board calibration system (Fig. 2, right). By comparing Aqua to VIIRS (Fig. 3), also using this technique, we have confirmed that the rapid degradation of the mirror in the red and near-infrared wavelengths (due to a deficient cleaning process just before launch) for details see (Barrie et al., 2012) was well accounted for by the on-board calibration system. We are now continuously monitoring the VIIRS calibration at two “calibration” sites using this technique (Fig. 4, right side), one over Australia (shown in Fig. 3, left side) and one over the Saharan Tibesti region (shown in Fig. 4, left side).

3. VIIRS cloud mask performance over Land (snow free)

While the validation of surface reflectance is greatly facilitated by AERONET data, the validation of the cloud mask remains a significant challenge. The CALIPSO mission and in particular the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) provides a unique, independent opportunity to evaluate cloud mask products. Despite its relatively narrow footprint (330 m to 5 km depending on the altitude of the layer sensed), CALIOP is acquiring data about 2 min after Aqua MODIS which makes it ideal for cloud mask evaluation. The VIIRS matchup with CALIPSO is less favorable with up to 12 minute difference between the two observations. Using the fraction of cloud at 330 m, 1 km and 5 km in the CALIOP footprint, we classify the footprint as cloudy if any of these three levels has a non-zero fraction of cloud. Two examples of the matchup analysis are shown in Fig. 5 (top) and Table 1 for two different dates corresponding respectively to data acquired prior to an important algorithm update to VCM (5/10/2012) and after the update (11/10/2012).

![Fig. 2. ASTER/MODIS multi-date comparison over a calibration site showing a saturation problem in the ASTER dataset (left), and an anomaly with the on-board calibration procedure (right).](image-url)
The number of matchups in this analysis is approximately 35,000. In Table 1, we report the performance of VCM as the percent of missed clouds (leakage), and the percent of false detections (False Det.). One can see a slight improvement in the performance, a small decrease in leakage rate from 8.9% to 7.2% after the update with roughly an equivalent increase in false detection rate from 1.9% to 3.9%. This trade off is perfectly acceptable as the downstream land algorithms favor fewer data with less cloud, than more data with increased contamination.

A similar analysis is conducted for the MODIS surface reflectance Internal Cloud Mask (ICM) and also reported in Table 1. The maturity of the ICM is leading to much better results, with very good performance, well in line with Land requirements. By design the ICM is conservative and tuned to the no-snow condition. Furthermore, by using the cloud adjacent quality flag available in the ICM to exclude even more data (see last column of Table 1), the leakage drops to close to 2%. A direct comparison was also undertaken between VCM and ICM for 11/10/2012 and clearly shows two issues with the VCM. Fig. 6a shows typically the difficulty of cloud masking over bright surfaces (in this case the Sahel) where an untuned algorithm can have a significant number of false detections. Fig. 6b also shows a typical problem with cloud masking when using the cirrus band over high elevation, bright surfaces: the threshold used under “normal” condition is triggered by the fact the surface is bright, but the amount of water vapor (that absorbs the radiation under high altitude cloud) is not high enough and thus the surface signal is “leaking” in the cirrus band. One solution for the cirrus test is to adapt the threshold as a function of target altitude.

4. Impact of aerosol optical thickness (climatology) on the VIIRS surface reflectance product

The aerosol retrieval over land can fail for a variety of reason, one being the brightness of the surface (Castanho et al., 2007; Levy,
In such cases the MODIS surface reflectance algorithm assumes a default “climatological” value for AOT corresponding to a clear day (0.05 at 550 nm and an urban clean aerosol model). The VIIRS surface reflectance algorithm is not using a default value but instead uses the output of the Navy Aerosol Analysis and Prediction System (NAAPS). The NAAPS is a global aerosol and visibility forecast model, which generates operational six-day global-scale forecasts for sulfate, dust, sea salt and smoke. NAAPS is based on a modification to the model developed by Christensen (1997). The NAAPS model output is available as 1° × 1°, at 6-h intervals and 24 vertical levels reaching 100 millibars. The NAAPS presumably has skills at its designed resolution and time step (or on a daily average), but it is ineffective at predicting the instantaneous AOT at VIIRS spatial resolution (1 km) and providing enough reliable information for performing atmospheric correction. This perspective is confirmed by a comparison of true color images presented in Fig. 7.

### 5. Quantitative assessment of the performance of the VIIRS surface reflectance over the AERONET sites

We have performed a preliminary estimate of the performance of the VIIRS SR over the AERONET sites. The protocol used is described in detail by Vermote and Kotchenova (2008) and will not be re-described here. The results are presented in Fig. 8a, for the VIIRS bands M5 (red) and M7 (near infrared) for 2012. Fig. 8a shows the ACCURACY or mean bias (red line), Precision or repeatability (green line) and Uncertainty (i.e. the quadratic sum of Accuracy and Precision) (blue line) of the SR as well the uncertainty specification (magenta line), that was derived from the MODIS theoretical error budget. While it is still early to make any strong statement about the performance, given the fact that there have been several updates to the VCM and Aerosol IP algorithms during 2012, but in general it can be seen that there is a bias at low reflectance in both red and near infrared.

For comparison, we show the performance of the MODIS Aqua Surface Reflectance product for corresponding bands band 1 (red), band 2 (near infrared) in Fig. 8b for the 2002-2009 period. The product behaves very well and is well under the specification for the medium range of reflectance (0.0 to 0.25), which corresponds to vegetated to

### Table 1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Leakage</td>
<td>8.9%</td>
<td>7.5%</td>
<td>2.6%</td>
<td>2.1%</td>
</tr>
<tr>
<td>False det.</td>
<td>1.9%</td>
<td>3.9%</td>
<td>6.5%</td>
<td>6.5%</td>
</tr>
</tbody>
</table>

The locations of the CALIPSO match up for the cloud mask performance evaluation (TOP: VCM on 5/10/2012 right, 11/10/2012 left; Bottom MODIS Aqua on 3/2/2007). VCM statistics and matchup location courtesy of Tom Kopp, VCM co-lead.
Fig. 6. a: An issue with the VIIRS Cloud Mask (VCM) over Sahel on 11/10/2012, due to the seed file currently used. The top image is the VIIRS IDPS generated product (clouds are white filled values), the middle is the VIIRS Land PEATE adjusted product (which is the product generated by the Land PEATE using the NASA Land Science Team adjusted version of the IDPS software), where the cloud are not set to filled values, the bottom is the MODIS Aqua product. The red circle outlines the false cloud detection in VCM due to a problem in the seed file used; the blue circle outlines an area of aerosol over-correction due to the use of the Navy Aerosol climatology (see Section 5.). b: An issue with the VIIRS Cloud Mask (VCM) over the Tibetan Plateau on 11/10/2012 due to the cirrus reflective test. The top image is the IDPS generated product (clouds are white filled value); the middle image is the Land PEATE adjusted product, where clouds are not set to filled values, the bottom is the equivalent MODIS Aqua data. The red circle outlines false cloud detections in VCM, due to problem in the cirrus reflective test over high elevation bright surfaces.
Fig. 6 (continued).
Fig. 7. Example of the issue with the use of NAAPS output in the atmospheric correction over Australia (04/20/13). The top panel is MODIS Aqua, the middle panel is the VIIRS IDPS output (clouds are white filled values), and the bottom panel is the VIIRS Land PEATE adjusted (the same stretch is used for each image). The bright yellow tone observed in the middle of Australia in the middle panel is the result of the use the NAAPS output (the only change in the atmospheric correction procedure between the Land PEATE adjusted version and the IDPS).
semi-arid surfaces, for which the accuracy is about 0.005 across the whole medium range of reflectance. It should be noted that the error budget was derived for an early version of the MODIS SR algorithm (Vermote & Saleous, 2006a) and will be revised for the MODIS Collection 6 algorithm. In particular, this early version did not attempt to make any determination of the aerosol model and the understanding of the different models was limited to a few studies (pre-AERONET era).

One can get an idea of the improvements still needed to the VIIRS Surface Reflectance product, by comparing Fig. 8a and b; however, as indicated previously, several problems have impacted the algorithm and

Fig. 8. a: Estimates of VIIRS surface reflectance Accuracy Precision and Uncertainty performances for VIIRS for band M5 (red), Left panel and M7 (Near-Infrared), Right Panel. The surface reflectance ‘truth’ is computed using the 6S radiative transfer code and the AERONET data for 2012. b: Same as 8a but for Aqua and for 2002–2009.
the general complexity of the IDPS system, resulting in delays in making changes, makes it difficult to get a large enough sample of consistently calibrated and processed surface reflectance data. The Land PEATE is currently planning a complete reprocessing of the VIIRS products with updated calibration and adjusted algorithms that will be a crucial input to future assessment of the products.

6. Adequacy of VIIRS surface reflectance production in an application of societal benefit

With evidence of climate change and global warming, the acceleration of the hydrological cycle is a strong possibility, translating in an increased number of "extreme" weather events such as flooding or

Fig. 9. A: Assessment of the impact of the 2012 Northern Hemisphere Drought from the MODIS Climate Modeling Grid daily NDVI data. b: A VIIRS NDVI anomaly (prototype) computed for the same date (July, 30th 2012) as the MODIS NDVI anomaly shown in Fig. 9, generated from data produced at the Land PEATE.
drought occurrence. The need to monitor the Earth system under these circumstances is extremely important, in particular in the area of food production. In the global economy, it is necessary to provide global, timely and unbiased assessments and forecasts of global agriculture production, to avoid large variations in commodity prices and organize food aid to prevent civil unrest in at-risk countries. The GLAM (Global Agricultural Monitoring) system developed by NASA and USDA at the Department of Geographical Sciences at University of Maryland has provided such a monitoring capability for the past 10 years (Becker-Reshef, Vermote, et al., 2010). An example of the GLAM input to the GEO agricultural monitoring community of practice, is shown on Fig. 9a at the peak of the crisis generated by the exceptional drought of 2012 (Karl et al., 2012).

As a prototype for the transition of VIIRS data into the GLAM system, we are using the VIIRS adjusted Land PEATE processing stream, post processing filtering and a preliminary adaptation of the MODIS Surface Reflectance Internal Cloud Mask algorithm and NDVI (ratio between the difference in the near-infrared and the red and the sum) anomaly, computed from the VIIRS daily observations in bands 11 (red) and 12 (near-infrared) (closer to MODIS band 1 and 2 than M5 and M7) and corrected for BRDF effects. Those bands are extracted from the VIIRS/NPP Daily Surface Reflectance Level 3 Global Diagnostic Data Record (DDR) at climate modeling grid (CMG - 0.05 degree) resolution (NPP_SRF_L3_CMG) available through NASA’s Level 1 and Atmosphere Archive and Distribution System (LAADS): http://laadsweb.nascom.nasa.gov/data/search.html.

Fig. 9a and b shows an example of the GLAM product and its adaptation to VIIRS. The anomaly image in Fig. 9a shows the cropland NDVI departure from the average (2000–2011) on July 30th 2012, highlighting hotspots of crops under stress during the 2012 droughts that affected the United States and the Black Sea region. The time-series curves below compare the daily development of croplands in 2012 (red) to average (2000–2011) in 3 important crop growing regions: Illinois, USA; Orenburg Oblast, Russia; Kostanay Oblast, Kazakhstan. The crop development through the season depicted by the NDVI shows consistent negative anomalies with regard to a ten year average, with highest discrepancies during the crop peak development period. In 2012 crops in the US, southern Europe and the Black Sea region suffered from prolonged high temperatures and lack of moisture, which resulted in significantly reduced production. This information was available one month prior to harvest and several months before the release of official statistics. The VIIRS NDVI anomaly is computed for the same day as the one presented in Fig. 9a and is shown in Fig. 9b. The similar spatial patterns of NDVI anomalies are obvious.

7. Conclusions and recommendations

The S-NPP VIIRS instrument was planned to generate data that will provide continuity in land observations from EOS MODIS. This meets an important requirement from the science community that is concerned with studying global change and application communities requiring multi-year comparisons. This also bodes well, as S-NPP was designed as a bridging-mission to the JPSS suite of operational instruments, which will have the same capability. The VIIRS advance over the previous NOAA operational polar orbiter, the Advanced Very High Resolution Radiometer is truly significant, with many of the MODIS capabilities and some improvements added. If we are successful in the launch of JPSS-1 with the same VIIRS instrument, then the research to operational transition in terms of instrument capabilities can be claimed a significant success.

The observations being collected by the VIIRS can provide continuity in observations from MODIS and are of sufficient quality for both science and operational applications; however, significant challenges need to be addressed so that the generated products do cover the full suite of MODIS land products or take full advantage of the capabilities provided by the instrument (Justice et al., 2013).

A performance evaluation of the SR IP is described above. After this preliminary product evaluation, we have made a number of suggested improvements to the product. However the operational system was not designed for extensive chain-testing of algorithm updates and as a result changes are slow to be made and can in turn create further problems. To develop the best possible SR product for science use, we recommend that rather than trying to transition an operational system (the IDPS) to research use, a purpose for which it was clearly not intended, a separate science research production system should be established with a dedicated testing system to allow improvements to be developed and tested in a Beta-Provisional, Stage 1 validation period with periodic reprocessing from Level 1 (calibration, geolocation) data, as was the case for MODIS. The MODIS LANCE (Land Atmosphere Near real-time Capability for EOS) experience has shown that science quality algorithms can be easily adapted for near real-time production (Murphy, Justice, Lowe, Maiden, & Davies, 2012). The VIIRS operational system and products could also benefit from new or improved algorithms developed and tested by the science community.

One critical consideration is that the IDPS was not designed to reprocess the data record but instead is designed for near real-time delivery. The need for a consistent product record from the start of operations is important for global change research but as shown above is also necessary for operational applications that are dependent on in multyear comparison or establishing trends. The development of climate data records (CDRs) from VIIRS will require a dedicated system outside of the IDPS. We therefore recommend that a CDR processing capability be developed for VIIRS with a science team driving the system requirements.

In summary, the VIIRS instrument is providing data from which a science quality SR product can be developed. The SR product currently running in the IDPS will we believe meet the needs of the operational community for which it was intended but a more flexible system is needed for implementing and testing the necessary improvements to the product.

This paper presents the basic tools and methods to fully evaluate SR-IP and its upstream dependencies. The evaluation will need to be reiterated as improvements are implemented and reprocessing is completed. After reprocessing, the scope of the cloud mask evaluation needs also to be broadened and a more statistically robust analysis needs to take place, to mimic the way the performance metrics are derived for SR-IP (Stage III validation), at a minimum a year of global coincident matches with CALIPSO data needs to be analyzed to fully cover the temporal/spatial variability in cloud and land systems.

References


