2014

The VIIRS Cloud Mask: Progress in the first year of S-NPP toward a common cloud detection scheme

Thomas Kopp
Aerospace Corporation, Thomas.J.Kopp@aero.org

William Thomas
The MITRE Corporation

Andrew Heidinger
NOAA/NESDIS Center for Satellite Applications and Research

Richard Frey
University of Wisconsin-Madison

Keith Hutchson
Redondo Beach, California

See next page for additional authors

Authors
Thomas Kopp, William Thomas, Andrew Heidinger, Richard Frey, Keith Hutchson, Barbara Lisager, Kurt Brueske, and Bonnie Reed

AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2013JD020458

Special Section:
Suomi NPP Calibration and Validation Scientific Results

Key Points:
- The VIIRS Cloud Mask is meeting most of its objectives after 1 year
- The VIIRS Cloud Mask serves the needs of a wide variety of VIIRS applications
- The VIIRS Cloud Mask is comparable to cloud masks from NOAA and NASA

Correspondence to:
T. J. Kopp,
Thomas.J.Kopp@aero.org

Citation:
Kopp, T. J., W. Thomas, A. K. Heidinger, D. Botambekov, R. A. Frey, K. D. Hutchison, B. D. Iisager, K. Brueske, and B. Reed (2014), The VIIRS Cloud Mask: Progress in the first year of S-NPP toward a common cloud detection scheme, *J. Geophys. Res. Atmos.*, *119*, 2441–2456, doi:10.1002/2013JD020458.

Received 28 JUN 2013
Accepted 14 DEC 2013
Accepted article online 5 JAN 2014
Published online 5 MAR 2014

The VIIRS Cloud Mask: Progress in the first year of S-NPP toward a common cloud detection scheme

Thomas J. Kopp[1], William Thomas[2], Andrew K. Heidinger[3], Denis Botambekov[4], Richard A. Frey[4], Keith D. Hutchison[5], Barbara D. Iisager[5], Kurt Brueske[6], and Bonnie Reed[7]

[1]The Aerospace Corporation, El Segundo, California, USA, [2]The MITRE Corporation, McLean, Virginia, USA, [3]NOAA/NESDIS Center for Satellite Applications and Research, Madison, Wisconsin, USA, [4]CIMSS, University of Wisconsin-Madison, Madison, Wisconsin, USA, [5]Northrop Grumman, Redondo Beach, California, USA, [6]Raytheon, Omaha, Nebraska, USA, [7]Global Science and Technology Corporation, Inc., Greenbelt, Maryland, USA

Abstract The Visible Infrared Imager Radiometer Suite (VIIRS) Cloud Mask (VCM) determines, on a pixel-by-pixel basis, whether or not a given location contains cloud. The VCM serves as an intermediate product (IP) between the production of VIIRS sensor data records and 22 downstream Environmental Data Records that each depends upon the VCM output. As such, the validation of the VCM IP is critical to the success of the Suomi National Polar-orbiting Partnership (S-NPP) product suite. The methods used to validate the VCM and the current results are presented in this paper. Detailed analyses of golden granules along with tools providing deep insights into granule performance, and specific cloud detection tests reveal the details behind a given granule's performance. Matchup results with CALIPSO, in turn, indicate the large-scale performance of the VCM and whether or not it is meeting its specifications. Comparisons with other cloud masks indicate comparable performance for the determination of clear pixels. As of September 2013 the VCM is either meeting or within 2% of all of its documented requirements.

1. Introduction

The Environmental Data Records (EDRs) produced by the Joint Polar Satellite System (JPSS) program, other than Imagery, depend upon accurate knowledge of the presence, or absence, of cloud. Whether it is a cloud-based product that needs to know where cloud exists, or an ocean/land/snow/aerosol product that needs to know where cloud is not, noting where cloud is present is a necessary precondition for that EDR to be produced. The design of the JPSS program dictated that all products would be produced on the Interface Data Processing Segment (IDPS), with attention paid not just to algorithm accuracy and precision but with attention to latency as well. Therefore, it was decided to produce a single cloud mask adequate for operational algorithm needs. The Visible Infrared Imager Radiometer Suite Cloud Mask (VCM) algorithm was designed with these requirements in mind.

The VCM evolved significantly during the period before the launch of Suomi National Polar-orbiting Partnership (S-NPP). A small team led by Northrop Grumman Aerospace Systems identified a number of improvements during this prelaunch time frame, assisted by lessons learned from the cloud mask derived from the Moderate Resolution Imaging Spectroradiometer (MODIS). Additional features beyond cloud identification, to include cloud phase, identification of multilayered clouds, differentiation between clouds and aerosols, and cloud shadows, were added or improved to address downstream EDR needs. The science now incorporated into the VCM combines features and lessons learned from prior programs, as well as newer algorithms that exploit the unique data collected by the Visible Infrared Imager Radiometer Suite (VIIRS) sensor, in order to maximize the quality of the VCM.

2. The VCM Algorithm

The VCM algorithm has been discussed extensively in other publications [*Hutchison and Jackson*, 2003; *Hutchison et al.*, 2005, 2008, 2009] and the VCM Algorithm Theoretical Basis Document (ATBD) [*National Aeronautics and Space Administration (NASA)*, 2013a], so only a high-level summary is presented here. The VCM cloud detection tests are grouped by surface type and solar illumination conditions, as illustrated in

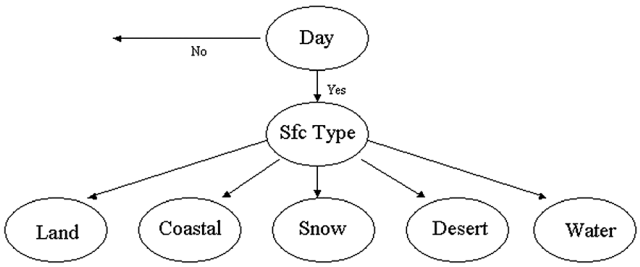

Figure 1. VCM branching logic.

Figure 1. Since each of these backgrounds contains its own set of cloud detection tests and related thresholds, one can work on a particular scene or background and any changes do not impact other types. This ability to isolate VCM performance to any given background benefits not just the validation itself but also allows the VCM validation team to isolate areas of concern and resolve them with minimal risk to areas under different conditions. Each cloud detection test employs three thresholds: a high cloud-free confidence, a low cloud-free confidence, and a midpoint threshold. Thus, the final set of VCM thresholds are based upon hundreds of tunable parameters, many of which vary with Sun-Earth-sensor viewing angle and atmospheric path moisture, as well as surface conditions. Cloud phase is also determined, using a method described in *Pavolonis et al*. [2005]. Clouds detected may be liquid water, mixed, opaque ice, cirrus, or overlap cloud phases, where mixed indicates a single cloud layer with a combination of ice and water particles, while overlap means separate ice and water cloud layers are present in the same pixel.

The specific VCM cloud detection tests applied during daytime conditions are shown in Table 1. The channels of VIIRS used to produce the VCM are found in the first column of Table 1, where "M" denotes VIIRS moderate resolution spectral bands and "I" denotes imagery resolution bands [*Hutchison and Cracknell*, 2006]. M bands have a spatial resolution of 750 m at nadir which increases to about 1.5 km at edge-of-scan, while I bands have resolutions of 375 m and 750 m, respectively. The VCM takes advantage of many of the bands available, with each band exploited for a specific purpose. Each individual cloud detection test is primarily designed to focus on a particular type of cloud, with the aggregation of the results from all tests combined to create the final cloud confidence analysis. Assuming three tests are applied to a given pixel, the overall cloud confidence is based on the cube root of the product of the probabilities for these three tests [*Hutchison et al*., 2011]. Based on this overall cloud-free probability, the VCM categorizes a pixel as confidently cloudy, probably cloudy, probably clear, or confidently clear in M band pixel space. In general the VCM is "clear-sky conservative," that is, if even just one cloud detection test identifies the pixel as cloud covered with high confidence, a cloud is

Table 1. Cloud Tests Used in the VIIRS Daytime Cloud Mask Algorithm as a Function of Surface Type

Cloud Tests	Water	Land	Desert	Coast	Snow
M9 (1.38 μm) reflectance test	X	X (if total path water vapor $>$ 0.50 cm)	X (if total path water vapor (tpwv) $>$ 0.50)	X (if tpwv $>$ 0.20)	X
M15–M16 (10.75–12.01 μm) brightness temperature difference (BTD)	X	X	X	X	
Tri-spectral M14, M15, M16 (8.55, 1076, 1201 μm) BTD test	X				
M15–M12 (10.76–3.70 μm) BTD test	X (if no sun glint)	X if TOC NDVI $>$ 0.27	X (if Lat $\geq$ 60° or $\leq -60°$)	X (if no sun glint) and TOC_NDVI $>$ 0.27	X
M12–M13 (3.70–4.05 μm) BTD test	X (if $-90° <$ Lat $< 90°$)	X (if $-90° <$ Lat $< 90°$) and TOC NDVI $>$ 0.27	X	X	X (if $-90° <$ Lat $< 90°$)
M5 (0.672 μm) reflectance test		X (M1 used when TOC NDVI $<$ 0.2)		X (M1 used when TOC NDVI $<$ 0.2)	
M7 (0.865 μm) reflectance test	X				
M7/M5 (0.865/0.672 μm) reflectance ratio test	X	X if (M5 $\geq$ 0.1)			
M1 (0.412 μm) reflectance test			X (if $-60° <$ Lat $< 60°$)		
I5 (11.45 μm) spatial test	X				
I2 (0.865 μm) reflectance test	X				

Table 2. VCM Performance Requirements

EDR Attribute	Specified Value
Probability of Correct Typing	
Ocean, day, COT > 1.0 tau	94%
Land, day, COT > 1.0 tau	90%
Ocean, night, COT > 1.0 tau	85%
Land, night, COT > 1.0 tau	88%
Cloud Leakage Rate	
Ocean, day, COT > 1.0 tau, no sun glint	1%
Land, day, COT > 1.0 tau	3%
Ocean and land, night, COT > 1.0 tau	5%
False Alarm Rate	
Ocean, day, COT > 1.0 tau	5%
Land and desert, day, COT > 1.0 tau	7%
Top-of-canopy NDVI 0.2 < NDVI < 0.4	
Ocean and land, night, COT > 1.0 tau	8%

considered to exist, i.e., the cloud-free probability is zero. To reduce cloud impacts on ocean color and sea surface temperature analyses, I bands are used toward the end of the VCM algorithm logic to test for the presence of cloud edges in pixels found over the ocean that are initially classified as confidently clear. However, the results of these I band-based tests can only change the VCM cloud confidence from confidently clear to probably clear or probably cloudy.

The higher-level requirements for the VCM are shown in Table 2. There are three fundamental parameters the VCM must consider: the probability of correct typing (PCT), leakage, and false alarms. The mathematical definitions of these performance metrics are shown in *Hutchison et al.* [2011]. PCT quantifies the raw number/percentage of clear/cloudy pixels as correct. Leakage occurs when a pixel identified as confidently clear in reality contains cloud. False alarms are the opposite of leakage and will identify a pixel as confidently cloudy when it should be clear. Note that in all of the requirements thin clouds with cloud optical thickness (COT) values of less than 1.0 tau are excluded from the requirement.

3. VCM Validation Methods

The JPSS program has established five categories of validation; in sequential order they are Beta, Provisional, and Validation stages 1, 2, and 3. For each product, Beta maturity indicates an early release product with known problems. Frequent updates to the algorithm are expected. When the data start to show stability in product performance and improvement in performance the product elevates to Provisional maturity, at which point incremental improvements may still occur. A product achieves Validated maturity once its performance is well defined over a range of representative conditions; each stage (1, 2, and 3) achieved as further quantitative results are obtained that validate the product. The VCM achieved the Provisional stage of validation on 20 February 2013. Further improvements are ongoing as the VCM algorithm moves through the different validation phases.

Validation of the VCM is performed on many different scales, since different types of errors are revealed with each approach. Qualitative analysis occurs on a large (hundreds) number of specific granules to isolate cloud detection performance on a test-by-test basis. Qualitative analysis in this case involves looking at various cloud mask bits or quality flags compared to various combinations of images using the moderate band sensor data records. Quantitative analysis on individual granules is done via creating "golden granules" (GG), in which a subject matter expert manually analyzes a chosen granule on a pixel-by-pixel basis resulting in a truth value for every pixel within that granule. On the larger scale regional and global quantitative analyses are determined via matchups with the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO). Finally, other VIIRS Calibration and Validation (Cal/Val) teams provide feedback of both a qualitative and quantitative manner on how the VCM impacts their particular EDR of interest.

In order to determine the root cause of VCM errors, a tool was designed that allows a scientist to display both the overall results of the VCM and the results of each individual cloud detection test. This tool is a Linux-based tool that operates at multiple locations. An example is shown in Figure 2. The display is capable of creating a red/green/blue image from any of the 16 moderate bands available from VIIRS. The user may toggle on and off any of the items in the three columns underneath the image on the right side, and each of those items represents an output from the VCM. The first column contains noncloud information relevant to the analysis, the second column has the overall and individual cloud test results, and the third column features the cloud phase results. The user may click on any of the associated boxes and display any number of desired items as needed to evaluate the scene. A specific example is discussed in section 4, paragraph 18.

Golden granules (GGs) are selected to assess VCM performance over three-contiguous VIIRS granules where a deficiency or challenging scene condition has been identified. Thus, the use of GGs is critical for

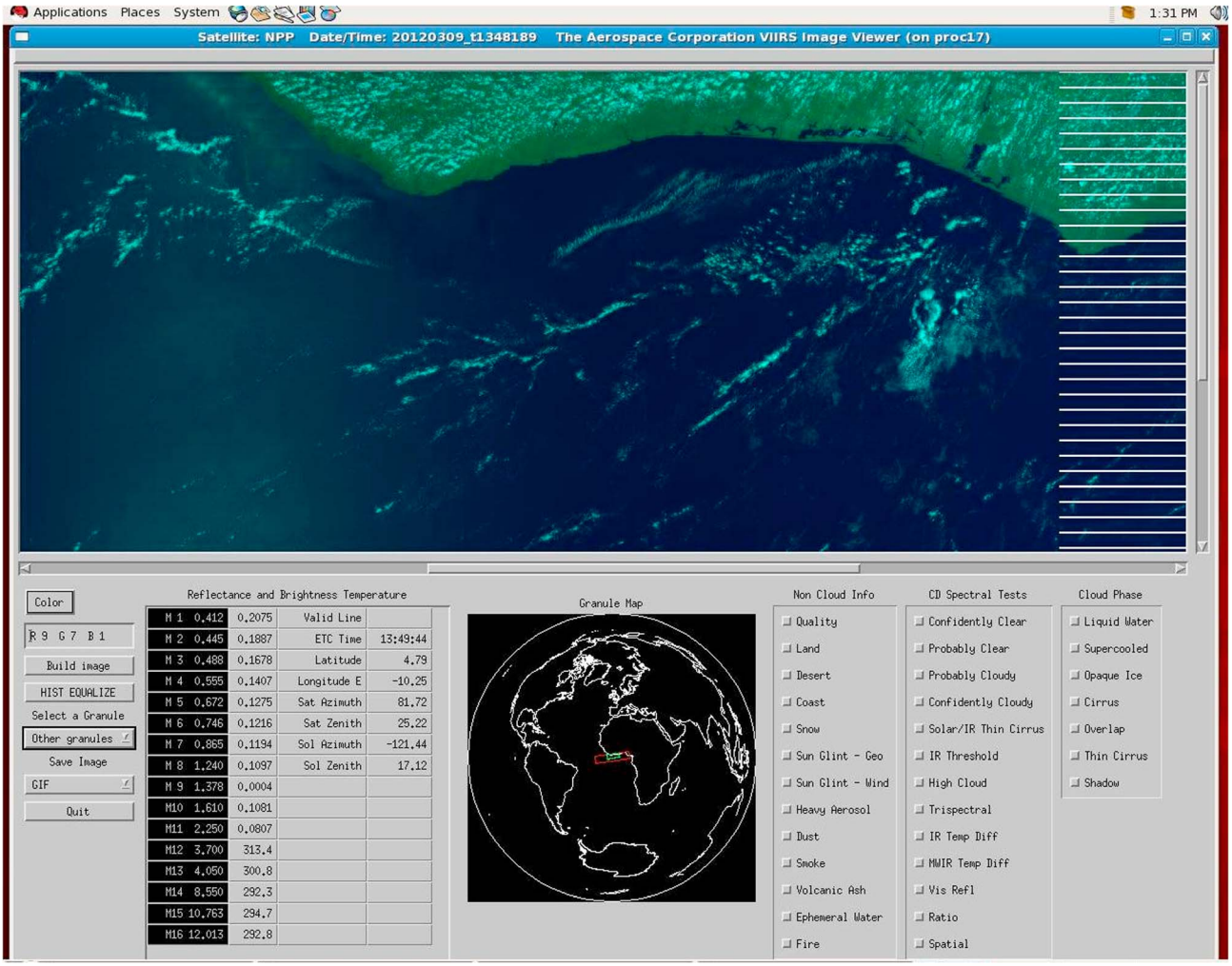

Figure 2. Example of the tool used to evaluate the VCM for a specific granule. The bottom left contains the distribution of the red, green, and blue color guns, in this case M bands 9, 7, and 3, respectively. The next column to the right contains the reflectance and brightness temperature values from the VIIRS M bands. The following column to the right contains characteristics of the scene geometry. In the bottom right are various flags which may be turned on and off in the display.

testing proposed software or threshold updates. Most of the 25 golden granules chosen to date were selected to address a specific problem found with the VCM results by the VCM, Cloud, Aerosol, and Snow/Ice Cal/Val teams, with the consequence that the summary of all of the golden granules covers a multitude of seasons, surface conditions, and cloud types. On one occasion, several golden granules were selected specifically to address concerns raised by the S-NPP Ozone Mapping and Profiler Suite Cal/Val team. Consequently, the final set of VCM golden granules cover a global range of geophysical conditions and represent the most challenging conditions expected to be encountered by this operational system. As an example, the northern and middle granules of GG#1 are shown in Figure 3a. This GG was selected to address a problem of misclassifying low-level clouds with glaciated tops as sea ice in the VCM snow-ice preprocessor as described in the next section. The problem has only recently been fully resolved [*Hutchison et al*., 2013]. A manually generated cloud mask is created for every Golden Granule, and it is assessed by at least three Subject Matter Experts (SMEs). One SME performs the initial analysis, the others then review that output (with no initial communication from the SME), and only after these reviews is the analysis considered completed.

Referencing Figure 3a contains a color composite of a VIIRS moderate resolution image where the VIIRS M5 band is assigned the red gun, the M10 band the green gun, and the M9 band the blue gun of a color display. In this color image, lower level water clouds appear yellow due to strong reflectances of these clouds in the M5 and M10 bands, but lower reflectances in the M9 band. Higher-level water clouds, on the other hand, appear white to

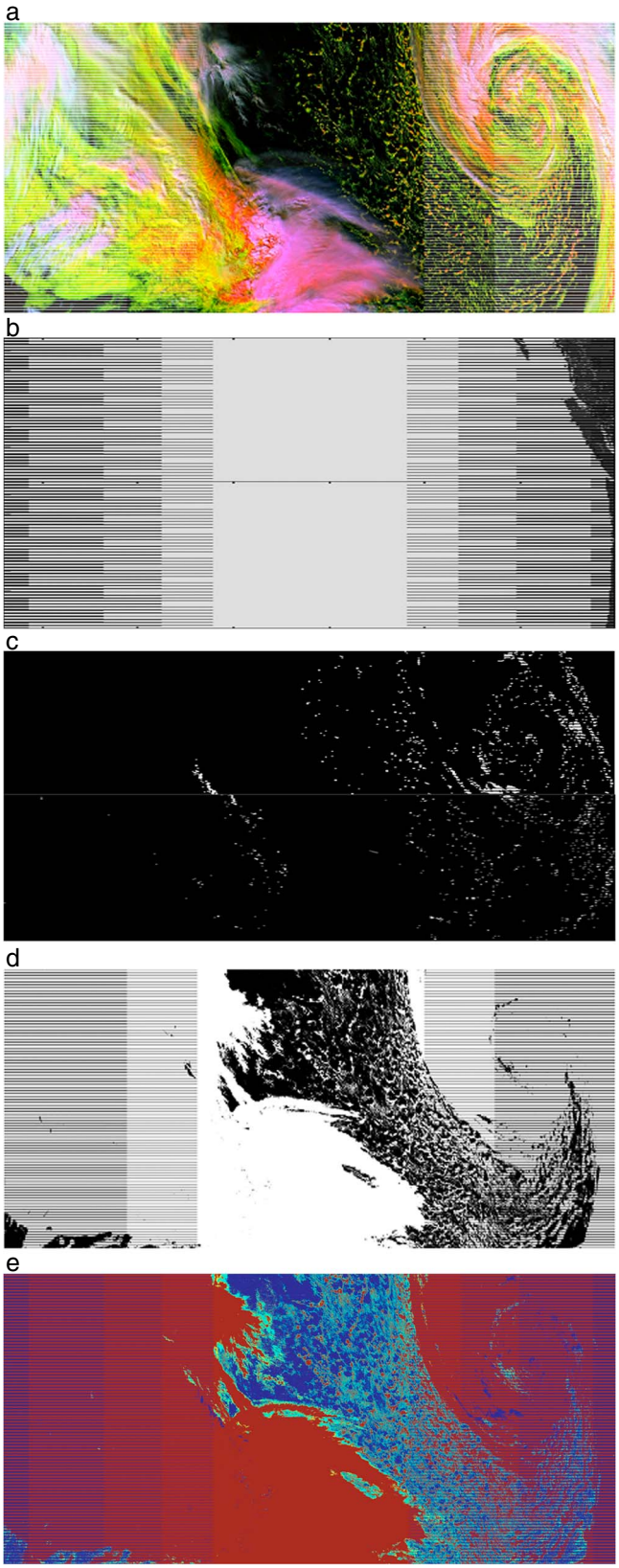

Figure 3

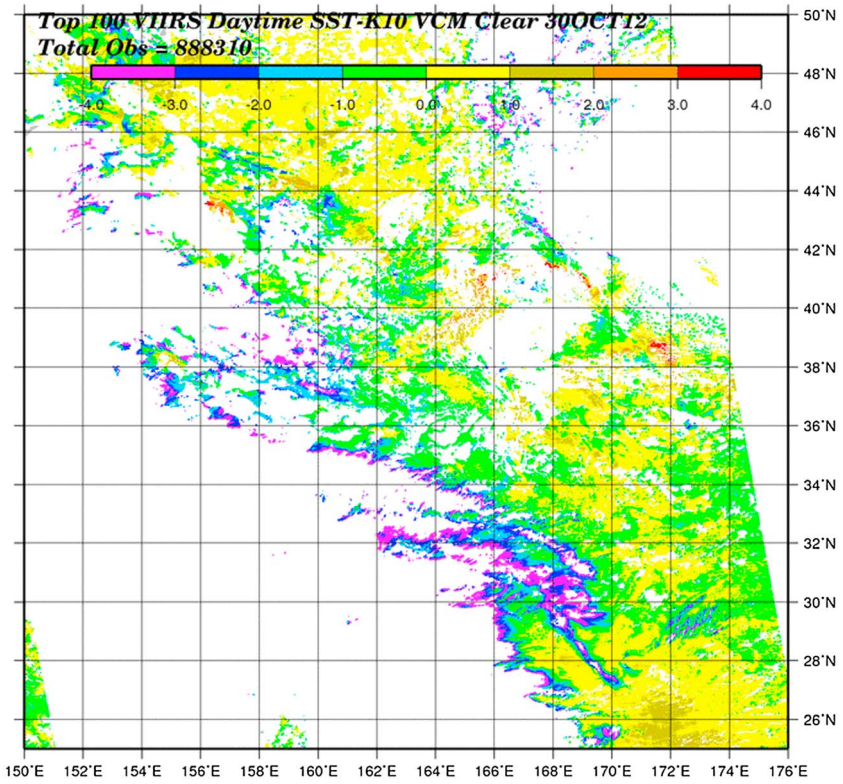

Figure 4. Plot showing sea surface temperature biases over the Pacific on 30 October 2012.

gray since their reflectances are similar in all three bands. Relatively lower level clouds with glaciated tops and higher-level ice clouds appear pink due to reduced reflectances in the M10 band compared to M5 and M9. Thin higher-level ice clouds appear blue since the reflectance is higher in M9 than the other bands. The ocean appears dark since reflectances are low in all three bands. Figure 3b shows the VCM land-water mask for these granules, note the west coastlines of the U.S. and Canada on the far right; while Figure 3c contains the internally generated VCM snow/ice mask. Figure 3d shows the manually generated cloud mask created by the subject matter experts. Figure 3e shows the VCM cloud confidence results, where dark red is confidently cloudy, dark blue is confidently clear, light blue is probably clear, and yellow is probably cloudy. Once a GG has been created, it serves as correlative data for establishing the current performance of the VCM on that granule and as a testing tool for evaluating the updated VCM performance for each potential upgrade. Actual results obtained with one of the many GGs created from VIIRS imagery are discussed in section 4, paragraph 14. To date approximately 25 GGs have been constructed to evaluate the VCM daytime performance while another 5–15 are planned to focus on nighttime performance.

To prove the requirements shown in Table 2, validation must include quantitative analyses over large scales. This is completed via the use of matchups with well-calibrated platforms. The primary platform used is CALIPSO, due to its relatively close proximity to the S-NPP orbit. The matchup methodology is based on finding the closest Aqua: MODIS or S-NPP: VIIRS pixel to CALIPSO 1km cloud layer pixel within ±12 min of time difference. This procedure is described in more detail by *Heidinger et al.* [2012]. Results may be summarized globally or for specific conditions such as day/night, ocean/land, or even latitude bands. This is necessary as some requirements are broken down in that manner. Matchups are also used to verify cloud phase. These are discussed in section 4.

Figure 3. (a) Daytime granules over the Gulf of Alaska on 17 February 2012. False color image; red = M5, green = M10, blue = M9. Pink shading is opaque ice clouds, blue is thinner cirrus, yellow indicates lower level water clouds, and gray represents high-level water clouds. (b) VCM land (dark) water (light) mask for daytime granules shows coastal regions of U.S. states of Washington and Oregon along with British Columbia, Canada. (c) VCM internally generated snow/ice (white) versus nonsnow/ice (black) analysis shows clouds over the ocean are misclassified as snow/ice. (d) Manually generated cloud mask analysis. (e) VCM cloud confidence analysis shows areas that contain clouds are misclassified as cloud free (blue), especially in the head of the comma-shaped cloud.

Table 3. Thresholds for Cloud Detection Tests Over Snow/Ice Backgrounds in Daytime Scenes[a]

Parameter Name	Type	Description
SD_M12_M13_Hi	Float	Confident clear threshold used in the snow/day M12–M13 emission difference test
SD_M12_M13_Mid	Float	Clear/cloudy threshold used in the snow/day M12–M13 emission difference test
SD_M12_M13_Lo	Float	Confident cloudy threshold used in the snow/day M12–M13 emission difference test
SD_M12_M15_Hi	Float	Confident clear threshold used in the snow/day M15–M12 emission difference test when terrain height is less or equal to high elevation threshold, HiElevThresh
SD_M12_M15_Mid	Float	Clear/cloudy threshold used in the snow/day M15–M12 emission difference test when terrain height is less or equal to high elevation threshold, HiElevThresh
SD_M12_M15_Lo	Float	Confident cloudy threshold used in the snow/day M15–M12 emission difference test when terrain height is less or equal to high elevation threshold, HiElevThresh
SD_M12_M15_HiHiElev	Float	Confident clear threshold used in the snow/day M15 – M12 emission difference test when terrain height is greater than high elevation threshold, HiElevThresh
SD_M12_M15_MidHiElev	Float	Clear/cloudy threshold used in the snow/day M15–M12 emission difference test when terrain height is greater than high elevation threshold, HiElevThresh
SD_M12_M15_LoHiElev	Float	Confident cloudy threshold used in the snow/day M15–M12 emission difference test when terrain height is greater than high elevation threshold, HiElevThresh
SD_M9_PTPW_INFLECTION	32-bit floating point	Total path integrated water vapor value at desert/day M9 versus path total precipitable water (TPW) inflection pt;
SD_M9_HI_ZERO_TPW_REFLECTANCE	64-bit floating point	M9 high clear-sky confidence reflectance at 0 cm total precipitable water for the snow/day M9 cloud confidence reflectance test; value percent reflectance
SD_M9_MID_ZERO_TPW_REFLECTANCE	64-bit floating point	M9 cloud/no cloud reflectance at 0 cm total precipitable water for the snow/day M9 cloud confidence reflectance test; value percent reflectance
SD_M9_LO_ZERO_TPW_REFLECTANCE	64-bit floating point	M9 low clear-sky confidence reflectance at 0 cm total precipitable water for the snow/day M9 cloud confidence reflectance test; value percent reflectance
SD_M9_HI_POLY_COEFS	64-bit floating point	Zeroth- to first-order polynomial coeffs on path TPW used in the confidently clear threshold calculation for the snow/day M9 cloud confidence reflectance and thin cirrus tests; calculation yields percent reflectance 1-D Array Size of Dimension(s): 2
SD_M9_MID_POLY_COEFS	64-bit floating point	Zeroth- to first-order polynomial coeffs on path TPW used in the Clear/Cloudy threshold calculation for the snow/day M9 cloud confidence reflectance and thin cirrus tests; calculation yields percent reflectance 1-D Array Size of Dimension(s): 2
SD_M9_LO_POLY_COEFS	64-bit floating point	Zeroth- to first-order polynomial coeffs on path TPW used in the confidently cloudy threshold calculation for the snow/day M9 cloud confidence reflectance and thin cirrus tests; calculation yields percent reflectance 1-D Array Size of Dimension(s): 2

[a]The applicable bands are identified by M?#, where # is the number of the band employed.

The final component of validating the VCM comes via coordination with the other VIIRS EDR Cal/Val teams. Such coordination can come from a number of forms. One example is shown in Figure 4. In this case the sea surface temperature team has generated an image of biases which may indicate cloud leakage. The larger negative biases were investigated by the VCM team to determine if a cloud should have been identified, and if so what test should have detected that particular cloud. Both the tool noted in paragraph 8 and GGs are used to determine if these kinds of clouds represent a wider-scale problem and guide the VCM Cal/Val team toward the optimal solution. Similar feedback is received from the other VIIRS EDR Cal/Val teams.

The on-orbit tuning of VCM thresholds is a critical process that is linked to both the methods of validation and the results obtained with the VCM algorithm. The idea of using thresholds that may be adjusted for optimizing cloud mask output dates back early automated cloud mask efforts within the Real Time Nephanalysis used by the Air Force Weather Agency [*Hamill et al.*, 1992]. As described in paragraph 4, the VCM is comprised of hundreds of thresholds. A sample of thresholds and their descriptions for the case of cloud detection over snow cover during the day is shown in Table 3. These thresholds were implemented in the VCM as tunable parameters, as defined in the VCM Operational Algorithm Description (OAD) [*National Aeronautics and Space Administration (NASA)*, 2013b], to ensure they can be easily adjusted during testing without recompiling the code hosted on the operational IDPS system. The S-NPP algorithm change process allows for a VCM threshold update to move efficiently through the algorithm review process and into operations at the IDPS on the order of a few weeks.

Table 4. Initial VCM Performance for the VIIRS Granule Collected on 17 February 2012 at 2218 UTC and Interim Performance as of 1 Year Later

Category	Number of Pixels	PCT	False Alarms	Leakage	PCPCs
Ocean day, outside glint	1,932,290	98.0%	1.8%	0.0%	10.8%
	(1,954.070)	(99.7%)	(0.3%)	(0.0%)	(11/8%)
Coast	763	100.0%	0.0%	0.0%	0.0%
	(760)	(100.0%)	(0.0%)	(0.0%)	(0.0%)
Inland water	58	100.0%	0.0%	0.0%	0.0%
	(58)	(100.0%)	(0.0%)	(0.0)	(0.0%)
Land, nondegraded	9383	100.0%	0.0%	0.0%	0.0%
	(9,354)	(100.0%)	(0.0%)	(0.0%)	(0.0%)
Snow	22,779	35.8%	0.0%	55.6%	13.4%
	(1,737)	(97.5%)	(2.5%)	(2.3%)	(7.8%)
Desert	8	100.0%	0.0%	0.0%	0.0%
	(6)	(100.0%)	(0.0%)	(0.0%)	(0.0%)

4. Results

This section will cover examples and results from the various methods of validation as discussed in section 3 Starting with GGs, and returning to the GG in Figure 3, this data set was selected to investigate the cause of VCM cloud leakages over some ocean surface conditions. Figure 3a shows low-level clouds having glaciated tops, indicated by the pink color of some open-celled cumulus (donut-shaped) cloud fields. These clouds extend from the lower right corner through the middle upper section of the scene. These glaciated clouds were misclassified as sea ice, as shown in Figure 3c, by the at-launch VCM snow/ice preprocessor. As a result, cloud tests listed in Table 1 that are normally used to detect these low-level clouds over the ocean were not applied because the background was typed as snow/ice, allowing these clouds to go undetected, as shown in Figures 3d and 3e. The updated logic used to resolve these misclassifications in the VCM snow/ice preprocessor has recently been implemented in the IDPS as described and demonstrated in the literature [*Hutchison et al.*, 2013]. Table 4 shows the improvements realized from the algorithm tuning and code modifications that occurred between February 2012 (unmodified/untuned IDPS at-launch software) and February 2013. The current performance values are shown in parentheses. Rather than being displayed as a simple overall cloud confidence for the granule, the results are stratified by background type, consistent with the surface types in Figure 1 and the VCM performance requirements listed in Table 2. In this case, the focused improvement is shown in the results for snow and ice. Significant leakage in the at-launch algorithm of 55.6% was sharply reduced to 2.3% because misclassifications of snow/ice over the ocean were eliminated, which in turn allowed a more full set of cloud detection tests to be applied over the corrected background. The overall PCT

Table 5. A Comparison of VCM and Collection 6 MOD35 Cloud Detection Agreements to the CALIOP Product for the Period From September 2013[a]

	VIIRS Cloud Mask (VCM)			MOD35 Collection 6		
	Sample Size $> 1.7 \times 10^6$			Sample Size $> 8.6 \times 10^6$		
	Hit Rate (%)	False Alarms (%)	Missed Clouds (%)	Hit Rate (%)	False Alarms (%)	Missed Clouds (%)
Global, all	89.5	3.9	5.3	94.3	2.2	2.9
60°S–60°N	94.9	3.0	1.5	96.8	2.3	0.5
Global day	92.7	3.5	3.0	96.5	1.7	1.5
60°S–60°N day	94.8	3.5	1.1	97.5	1.8	0.6
Global night	86.4	4.4	7.3	92.2	2.7	4.1
60°S–60°N night	95.0	2.4	2.0	96.1	2.8	0.5
60°S–60°N water day	95.3	3.5	0.6	97.4	2.1	0.3
60°S–60°N water night	95.2	2.8	1.3	95.1	3.8	0.4
60°S–60°N land day	93.9	3.6	2.2	97.7	0.9	1.1
60°S–60°N land night	94.5	1.5	3.7	98.8	0.3	0.7
Desert day	96.0	1.2	2.8	97.8	0.5	1.5
Desert night	94.0	0.9	4.9	98.6	0.3	0.9

[a]Comparisons were made for only those clouds with CALIOP COT $\geq$ 1.0. Only confident clear and confident cloudy designations from the VIIRS and MODIS cloud masks were considered.

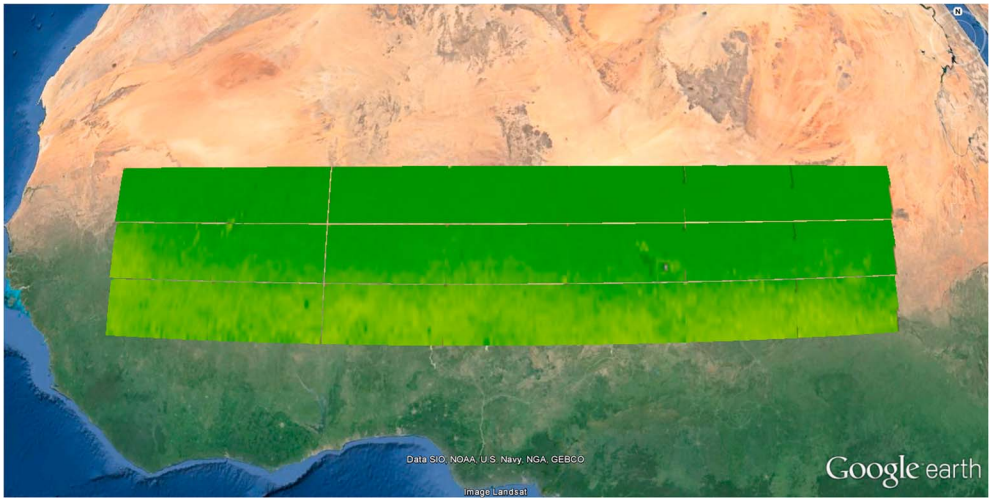

Figure 5. Representation of NDVI changes over Australia as of January 2013. Darker green shades indicate higher NDVI increases.

over the same background type improved from 35.8% to 97.5%. Such quantitative results are the strength of employing a GG strategy as part of the overall VCM Cal/Val process.

The matchups with CALIPSO reveal the large-scale capabilities and shortcomings from the VCM. The operational VCM as of September 2013 and the latest version of Aqua MOD35 (Collection 6) cloud detection data were compared to collocated Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) data for September 2013. These results are shown in Table 5. Performance results (i.e., PCT, false alarm, and leakage) are computed for the binary cloud mask, which is defined as pixels classified as confidently clear and cloudy pixels [*Hutchison et al.*, 2011]. Only clouds with optical depths greater than or equal to 1.0 are considered. VCM and MOD35 pixels were collocated separately with CALIOP so that the numbers of comparisons are different; however, the MODIS-AQUA and VIIRS orbits are both polar and Sun synchronous with equator crossing time near 1:30 P.M. local time and orbit altitude differences of ~120 km. Comparisons are divided into several basic scene types seen in the first column. The other columns show agreement (PCT), false clear (false alarms), and missed clouds (leakage) in percent for the VCM (left) and MOD35 (right). VCM agreements with CALIOP are greater than 90% overall between 60°S and 60°N latitude and lag MOD35 by about 2%. Performance is better over water than land, with VCM agreements roughly 4% less than MOD35 for land scenes and about 2% less for daytime water but about the same for nighttime. The larger differences seen between VCM and MODIS in the global category are due to lesser VCM agreements in polar regions. The lower performance of the VCM algorithm is not unexpected at this phase of program maturity, as these more difficult areas take some time and effort to fine tune in cloud detection algorithms. Furthermore, the VCM depends upon ancillary data (e.g., snow/ice) whose algorithms are also undergoing validation, and they are not produced routinely on the operational IDPS. Comparison of the VCM performance in Table 5 to the requirements in Table 2 helps to determine the extent that the VCM achieves its specifications. The results indicate the VCM is near or has achieved the requirements for PCT and false alarms for all but polar locations; however, leakage values are still higher than desired. Several reasons for the higher leakage in the polar regions have recently been identified, and significant improvement is expected in 2014 as planned algorithm and tuning updates are implemented.

In addition to evaluating the VCM cloud confidence, CALIPSO may also be used to determine the performance of the cloud phase algorithm. Cloud phase has significant impacts on the downstream cloud algorithms, which use cloud phase as a branching tool to decide which algorithm (i.e., water/ice) to employ. Initial results are shown in Table 5. These are generally consistent with MODIS results. Until cloud optical thicknesses (COT) exceed approximately 0.5, the values increase with COT, which is consistent with expectations. The VCM algorithm is flexible enough that certain cloud phase algorithm elements may also be tuned to improve performance.

Interactions with other Cal/Val teams, as discussed previously, have resulted in concrete improvements to the VCM. For example, the VCM has a dependence on a top-of-canopy (TOC) Normalized Difference Vegetation Index (NDVI) database; however, the determination of TOC NDVI itself is the responsibility of the Land

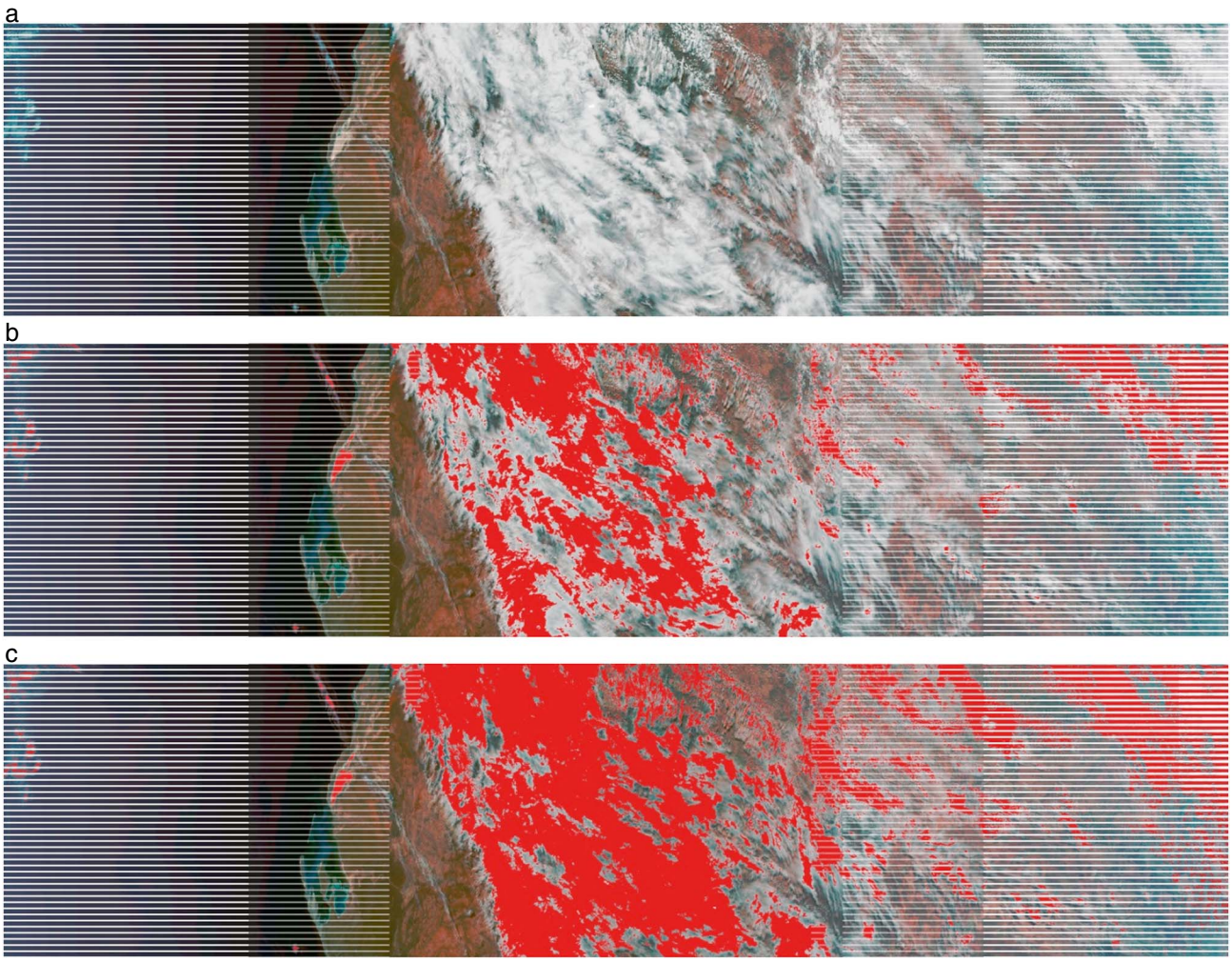

Figure 6. Example of improvements from tuning the VCM. (a) Original true color image covering Western Australia, 17 July 2013. (b) Initial results from the visual brightness cloud detection test. (c) Final results from the same test after tuning was completed.

Product team. However, any impacts to the VCM and any necessary adjustments in the internal algorithms are the responsibility of the VCM team. Figure 5 shows NDVI changes from January to April 2013. These changes led to leakage in the VCM. The VCM team then tuned the appropriate thresholds, and the resulting differences are shown in Figure 6. This granule is from Western Australia on 17 July 2013. In this case, the visual brightness cloud detection test, which depends in part upon TOC NDVI, was not identifying cloud sufficiently. Obvious cloud features were not being identified as cloudy, and no other available cloud detection test was able to identify these clouds. Using the TOC NDVI inputs from the Land team and the sensor data record (SDR) values available from the qualitative analysis tool, the related thresholds were adjusted to account for the changes induced by the underlying TOC NDVI values. As seen in Figure 6c, the induced leakage has been significantly reduced. This is only one of many examples where the VIIRS EDR teams work collaboratively to improve the VCM and hence the output of the downstream EDRs as well.

The larger results of Table 5 are more pessimistic than those presented in Table 4 due, primarily, to the nature and timing of the calculations. The GG results in Table 4 reflect all algorithm modifications and threshold updates implemented or planned through late 2013. Matchup data require extensive granule collection and computation and are only done on archived VIIRS data. The collocated results, as shown in Table 5, were calculated using the VCM as it stood operationally in October 2012. Also, thresholds are necessarily tuned by investigation of a relatively small number of VIIRS granules, one of them usually a particular GG. In fact, GGs are usually chosen because they display a particular scene type where a problem in VCM results has been

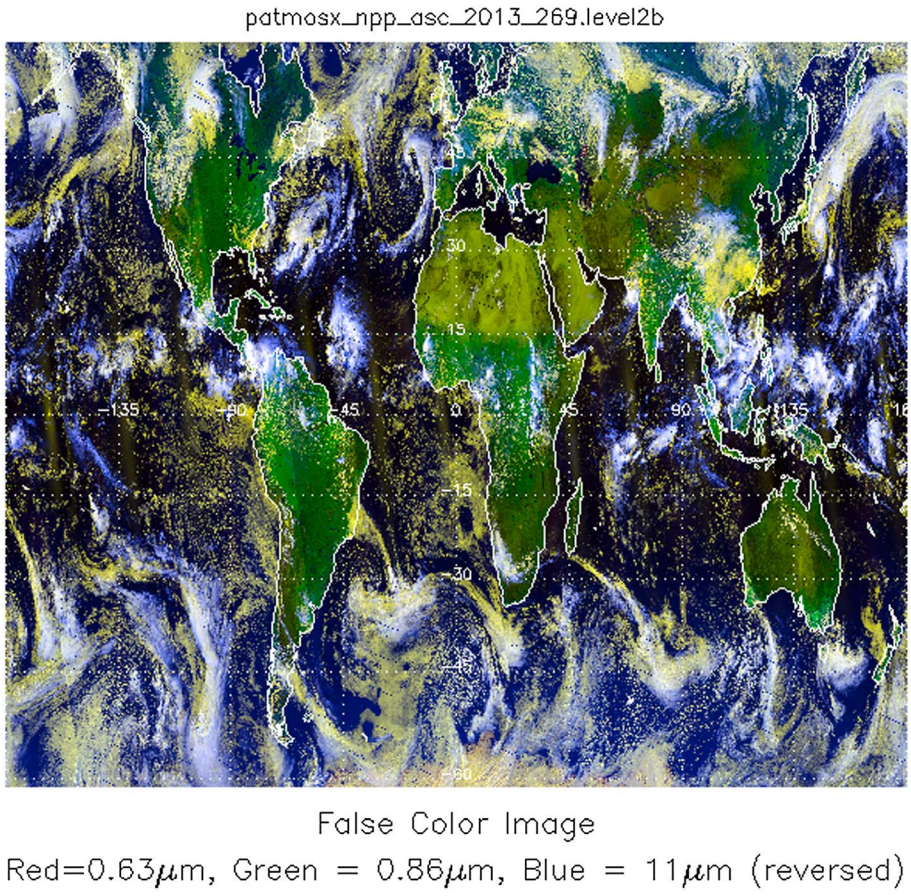

Figure 7. False color image of S-NPP data taken during 26 September 2013 daytime.

identified. In contrast to this process, collocated CALIOP comparisons cover multiple days or seasons with varying atmospheric conditions, cloud properties, solar illumination, and ancillary data quality. Each of the primary measurement parameters of the VCM (PCT, false alarms, and leakage) will improve as additional enhancements and corrections are implemented. The VCM continues to evolve, so these results herein will only get better as the validation of the VCM progresses over the next 12 to 18 months. There is high confidence the VCM will ultimately attain all of the specifications levied upon it.

5. Global Comparisons With Other Cloud Masks

The sections above describe the quantitative validation and comparisons done with the VCM. These comparisons were able to show how the VCM performs relative to the specifications placed on it. However, it is also instructive to compare the VCM directly against other masks to understand its specific characteristics and how they compare to other cloud mask products. One such product is the cloud mask generated by NOAA with the Clouds from advanced very high resolution radiometer Extended Processing System (CLAVR-x). CLAVR-x is run operationally on the advanced very high resolution radiometer (AVHRR) which is the predecessor sensor to VIIRS. The CLAVR-x system has been modified to process VIIRS data, and the CLAVR-x cloud mask [*Heidinger et al*., 2012] has been updated to include additional VIIRS channels. The CLAVR-x mask is also a four-level mask but employs a naïve Bayesian methodology. In the following analysis and discussion, the CLAVR-x mask will be referred to as the NOAA mask. The NOAA cloud mask is currently used by the NOAA National Centers for Environmental Prediction for verifying the cloudiness in its forecast models and approximates the NOAA mask expected for the Geostationary Operational Environmental Satellite, R-series (GOES-R) Advanced Baseline Imager. Therefore, comparisons of the characteristics of these masks are relevant. Another mask that is relevant to the VCM is the NASA MODIS cloud mask [*Ackerman et al*., 1998]. An

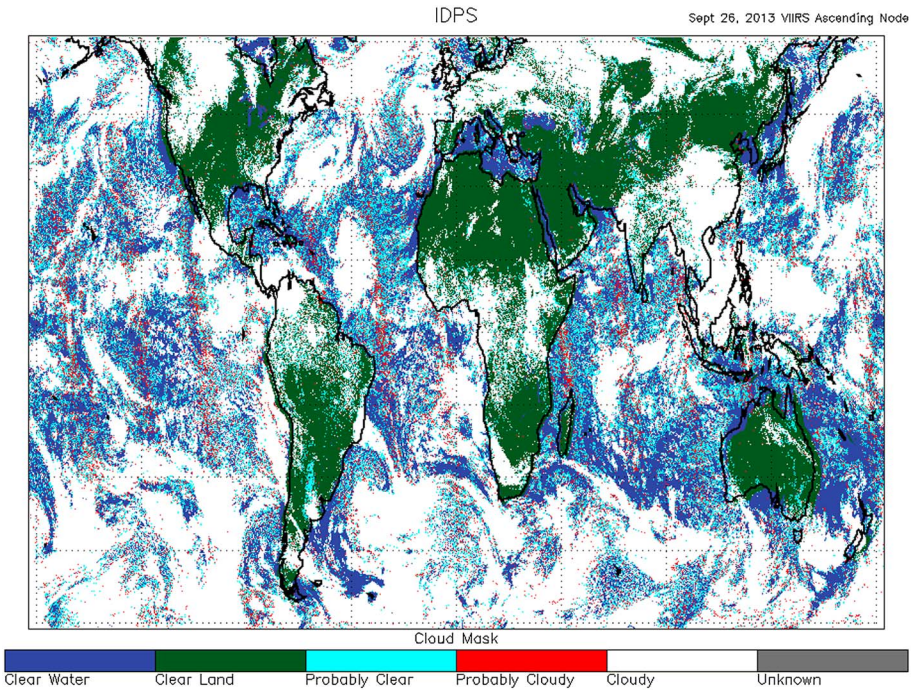

Figure 8. VIIRS IDPS Cloud Mask for daytime 26 September 2013.

early version of the MODIS mask was used in developing the VCM. Significant development has occurred in both masks, and communication between the VCM and MODIS teams has been ongoing for several years.

In this section, we compare one recent day of the VCM, NOAA, and the NASA mask to explore their overall consistency. These comparisons are for daytime data collected on 26 September 2013. The IDPS VCM and NOAA mask were generated on S-NPP data, while the NASA mask was generated on AQUA/MODIS data. The

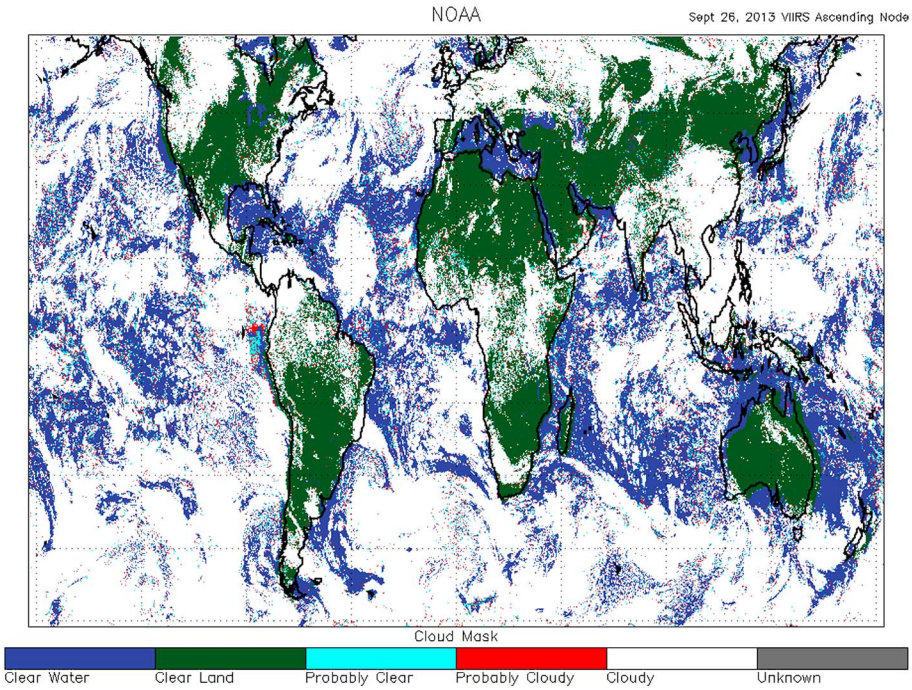

Figure 9. NOAA VIIRS Cloud Mask for daytime 26 September 2013.

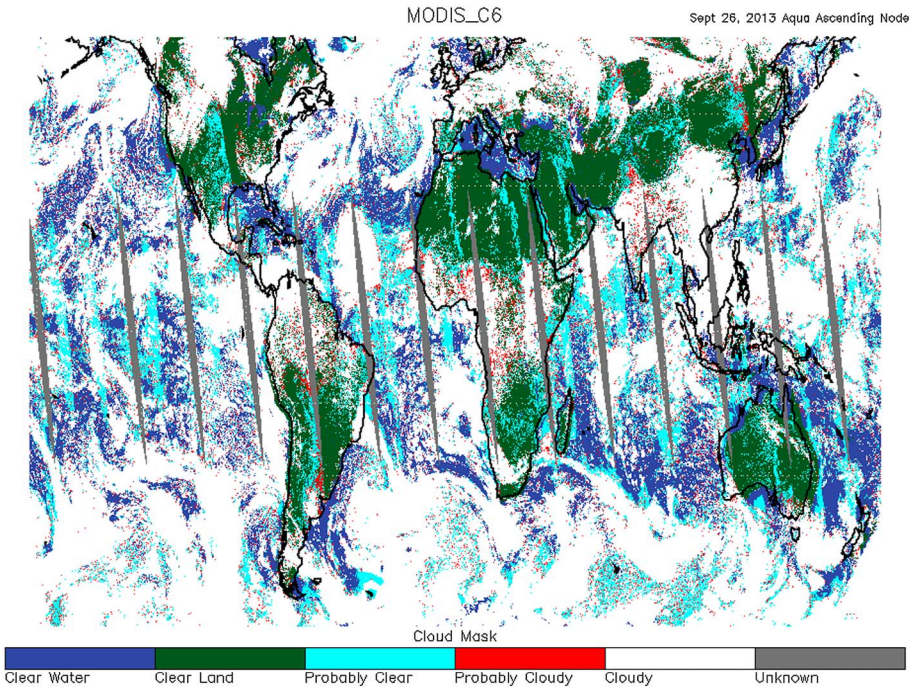

Figure 10. NASA MODIS (MYD35) Collection 6 Cloud Mask for daytime 26 September 2013.

global data were generated by mapping the granule data on a 0.1° latitude/longitude grid. Nearest neighbor sampling was used with the exception that bowtie deleted pixels were excluded. Figure 7 shows a false color image generated using the M5, M7, and M15 observations. In this color combination, high opaque clouds appear white, transparent high clouds are bluish white, low clouds are yellowish white, clear ocean is dark blue, and vegetated land surfaces appear green. Figure 8 shows the VCM, Figure 9 shows the NOAA mask, and Figure 10 shows the NASA mask (run on MODIS). In these figures, white pixels are confidently cloudy, red pixels are probably cloudy, cyan pixels are probably clear, and confidently clear pixels are colored blue for water surfaces or green for land surfaces. One obvious difference between the masks is the distribution of the probably clear pixels. The NOAA mask has relatively few probably clear pixels except over desert or snow-covered land. The VCM has many more probably clear pixels especially over the ocean in regions of small-scale cloudiness. This results from VIIRS imagery band (375 m) spatial tests that are used in the VCM algorithm to detect partially cloud-filled pixels in the VIIRS radiometric bands (750 m). The NASA mask also has a significant number of probably clear pixels with a concentration in regions of oceanic glint. These differences are consistent with the design philosophy of each mask. The VCM and NASA masks (MOD35) share a core methodology; however, the MOD35 is a more clear-sky conservative algorithm. The VCM design attempts to strike a more even balance between false alarms and leakage. This is seen in the choice of clear-sky confidence thresholds, where the various cloud categories (i.e., confident clear, probably clear, probably cloudy, and cloudy) are indicated at higher confidence values in the MOD35 algorithm, reducing the range of values indicating probably clear and probably cloudy. In addition, MOD35 performs "clear-sky" tests that in some cases allow an initial probably clear decision to be increased to confident clear when certain spectral criteria permit.

Because the VCM and NOAA masks were run on the same data, it is useful to construct a difference image for these results. Figure 11 shows the difference of the binary (clear or cloudy) masks derived from the VCM and NOAA masks. The binary masks were generated by reclassifying probably clear and confidently clear as clear and probably cloudy and confidently cloudy as cloudy. Pixels where the binary mask agreed on clear are colored blue for water surfaces and green for land surfaces, while pixels where the binary cloud masks agreed on cloudy are colored white. Cyan pixels are those where the VCM was cloudy but NOAA reported clear. Red pixels are where NOAA reported cloudy but the VCM reported clear.

Figure 11 reveals that the two masks agree very well on the distribution of binary clear and cloudy classifications. The results in Figure 11 show that roughly 99% of confidently cloudy VCM pixels are also labeled

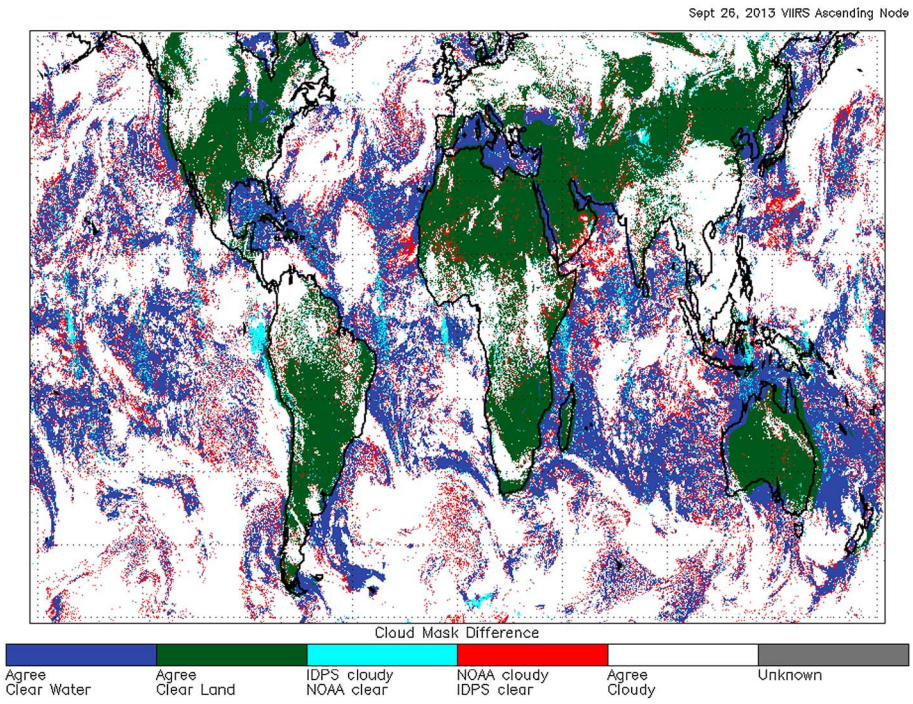

Figure 11. Difference in the binary (clear/cloudy) IDPS and NOAA masks.

confidently cloudy by the NOAA mask. The value over land is 96%. The agreement for pixels labeled confidently clear in the VCM is less than that for confidently cloudy, but the numbers are still high with the agreement over land being 89% and over water being 86%. However, Figure 11 does indicate some differences in characteristics of the two masks. It is important to note that there is no objective truth in this analysis, and these differences do not imply faults in either mask. Nonetheless, the VCM appears to generate higher numbers of probably clear than the NOAA mask and imagery bands are not used in the VCM algorithm over land surfaces. Roughly 15% of the ocean is classified as probably clear in the VCM, while that number is below 5% for the NOAA mask. This higher number of probably clear classifications over the ocean in the VCM may potentially be addressed by additional tuning, although the program is considering an algorithm adjustment as well, recall the VCM is only at its Provisional stage of validation. The NOAA mask also tends to classify many

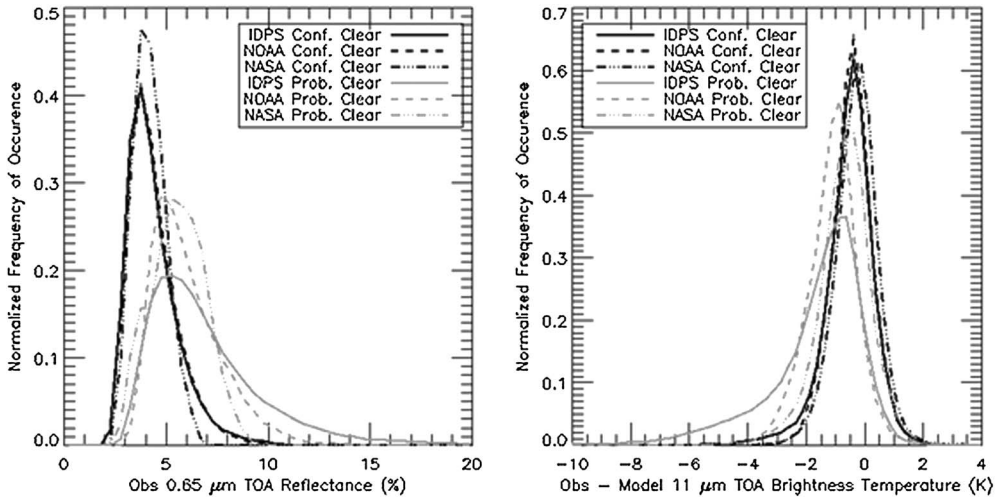

Figure 12. Distributions of the (left) clear and probably clear observed 0.65 μm reflectance observations and (right) observed minus model 11 μm brightness temperature. Data observed on 26 September 2013 for daytime ice-free and glint-free ocean pixels between 60°S and 60°N. IDPS and NOAA are generated using S-NPP data. NASA generated using AQUA-MODIS data.

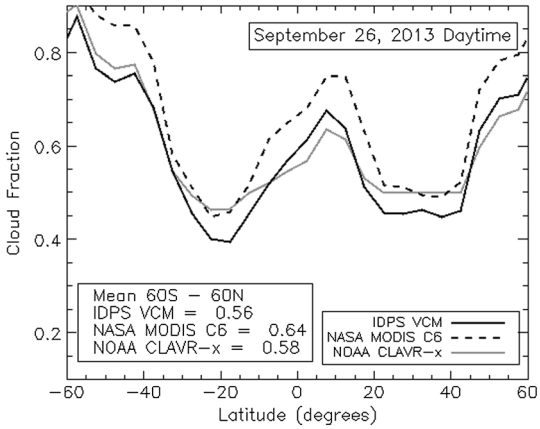

Figure 13. Zonal Cloud Fraction Comparison for 26 September 2013. MODIS results are generated from the MYD35 Collection 6 data. NOAA CLAVR-x was processed on the IDPS VIIRS SDRs using the NOAA VIIRS Cloud Mask. Cloud fraction was computed from the four-level masks. Clear pixels were assumed to be cloud free. Cloudy pixels were assumed to be 100% cloudy. Probably clear pixels were assumed to be 33% cloudy and probably cloudy pixels were assumed to be 66% cloudy. Mean values were computed using area weighting.

pixels as confidently cloudy where the VCM calls them probably clear, and from Figure 7, this appears to occur on the edges of cloud systems.

Figures 8–11 show the global characteristics of the VCM compared to the NOAA and NASA masks. The biggest differences occur in the partition of pixels between the confident clear and probably clear categories. Figure 12 shows the radiometric impact of this partition for each of the masks using 0.65 and 11 μm measurements in order to capture the impacts of different types of cloud contamination. Figure 12 (left) shows the probability distribution functions of the observed top-of-atmosphere 0.65 μm reflectance of glint-free pixels for the confidently clear (black) and probably clear (grey) pixels of each mask. The curves are normalized so that the area under the curve is the same; differences due to the number of pixels are removed. The absence of glint was determined from the geometric glint flag of the VCM, which is set to true if the glint zenith angle is less than 40°. The IDPS results are plotted with solid lines, the NOAA results are plotted with dashed lines, and the NASA results are plotted with dash-dotted lines. The confident clear distributions are expected to peak near low values of reflectance with a minimal tail toward higher values of reflectance. Higher values of reflectance are assumed to indicate cloud contamination. The probably clear distributions should include values with higher reflectances. This pattern is generally seen in Figure 12 (left). The MODIS confidently clear distribution appears to have the smallest tail toward higher reflectance. The NOAA and IDPS confidently clear distributions are similar to each other. The most obvious difference is the probably clear distribution where the IDPS shows a much larger tail toward higher reflectance than the NOAA or NASA results. This indicates the potential for encountering cloud contamination is higher in the IDPS probably clear results than the NOAA or NASA probably clear. This behavior is consistent with the definition and the larger number of probably clear pixels in the VCM results.

Figure 12 (right) shows the same analysis applied to the difference of observed and modeled top-of-atmosphere 11 μm brightness temperatures. The computation of the modeled clear-sky 11 μm brightness temperature is taken from the Pathfinder Atmospheres Processing System (A. K. Heidinger et al., The AVHRR Pathfinder Atmospheres Climate Data Set, submitted to *Bulletin of the American Meteorological Society*, 2013) applied to S-NPP and AQUA/MODIS. The clear distributions should peak near zero, and values significantly less than zero are indicative of cloud contamination. Figure 12 (right) shows the IDPS, NOAA, and NASA distribution do peak near zero. The IDPS results show the largest tail toward negative values. The offset of the mode of the MODIS results from the mode of the IDPS and NOAA results is likely due to differences in the VIIRS and MODIS infrared radiative transfer used in Pathfinder Atmospheres - Extended (PATMOS-x). The probably clear IDPS results show the largest tail toward negative values which are expected for the reasons given above. In summary, both panels in Figure 12 indicate the IDPS confidently clear pixels provide data that is comparable to that from NOAA and NASA though the indications of cloud contamination are slightly larger for the IDPS in both distributions than for NASA or NOAA.

It is also instructive to compare the VCM's performance in terms of its distribution of cloud fraction since cloud fraction is often used to compare different cloud data sets [*Stubenrauch et al.*, 2013] and to verify numerical weather or climate prediction models. Figure 13 shows the zonal cloud fractions computed from the same VCM, NOAA and NASA mask data used in Figures 8–10. Cloud fraction was computed from the four-level masks with confidently clear pixels assumed to be cloud free, confidently cloudy pixels assumed to be 100% cloudy, probably clear pixels assumed to be 33% cloudy, and probably cloudy pixels assumed to be 66% cloudy. The zonal distributions were computed for 5° latitude zones. All pixels were included in this analysis and no attempt was made to discard data from the multiple views at higher latitudes. In addition,

area-weighted mean values of cloud fraction for the entire region were computed and shown in the figure legend. The VCM tends to offer a similar zonal cloud fraction for most zones relative to the NOAA results. The MODIS cloud fraction is the highest for all zones reflecting the more clear-sky conservative nature of that mask. The difference between the VCM and MODIS appears quasi-constant for all zones; this indicates a systematic difference in cloud sensitivity across all cloud types between the two masks. In terms of the mean cloud fractions, the VCM is 9% below MODIS but 1% below NOAA. As demonstrated by the recent Global Energy and Water Cycle Experiment Cloud Assessment Report [*Stubenrauch et al.*, 2013], these differences are not unexpected based on the range of cloud fractions observed from the contributing data sets. The good correlation in the shape of these distributions and level of agreement in the values indicates that the VCM performance is consistent in all regions for this daytime nonpolar analysis for one day.

6. Conclusions

The VCM is being validated across many scales since the launch of S-NPP. Tools that allow detailed analyses of specific granules reveal any shortcomings of particular cloud detection tests. The development of GGs permits detailed quantitative analyses for specific scenes and their related cloud detection tests and often serves as an effective independent data set. Extensive matchup comparisons reveal the performance of the VCM across space and time on larger scales. This multipronged effort has led to significant improvements in the VCM since the launch of S-NPP and gives the VCM Cal/Val team a path for continued VCM validation. With the added software updates already scheduled for implementation into the IDPS, the VCM is expected to achieve all of its requirements. Comparison of its global distribution cloud fraction also indicates the VCM is working well compared to other established methods.

Acknowledgments

This work was funded by the JPSS Program Office Data Products and Algorithms Division (DPA) led by Heather Kilcoyne. Much of the data used here was provided by the Atmospheric PEATE at the University of Wisconsin SSEC. The views, opinions, and findings contained in this report should not be construed as an official National Oceanic and Atmospheric Administration or U.S. Government position, policy, or decision. All trademarks, service marks, and trade names are the property of their respective owners.

References

Ackerman, S. A., K. I. Strabala, W. P. Menzel, R. A. Frey, C. C. Moeller, and L. E. Gumley (1998), Discriminating clear sky from clouds with MODIS, *J. Geophys. Res.*, *103*, 32,141–32,157.

Hamill, T. M., R. P. d'Entremont, and J. T. Bunting (1992), A description of the Air Force real-time nephanalysis model, *Weather Forecast.*, *7*, 288–306.

Heidinger, A. K., A. T. Evan, M. J. Foster, and A. Walther (2012), A naive Bayesian cloud-detection scheme derived from CALIPSO and applied within PATMOS-x, *J. Appl. Meteorol. Climatol.*, *51*(6), 1129–1144.

Hutchison, K. D., and J. M. Jackson (2003), Cloud detection over desert regions using the 412 nanometer MODIS channel, *Geophys. Res. Lett.*, *30*(23), 2187, doi:10.1029/2003GL018446.

Hutchison, K. D., and A. P. Cracknell (2006), *VIIRS—A New Operational Cloud Imager*, vol. 218, CRC Press of Taylor and Francis Ltd, London.

Hutchison, K. D., J. K. Roskovensky, J. M. Jackson, A. K. Heidinger, T. J. Kopp, M. J. Pavolonis, and R. Frey (2005), Automated cloud detection and typing of data collected by the Visible Infrared Imager Radiometer Suite (VIIRS), *Int. J. Remote Sens.*, *20*, 4681–4706.

Hutchison, K. D., B. Iisager, T. Kopp, and J. M. Jackson (2008), Discriminating between clouds and aerosols in the VIIRS Cloud Mask algorithms, *J. Atmos. Oceanic Technol.*, *25*, 501–518.

Hutchison, K. D., R. L. Mahoney, E. F. Vermote, T. J. Kopp, J. M. Jackson, A. Sei, and B. D. Iisager (2009), A geometry-based approach for identifying cloud shadows in the VIIRS Cloud Mask algorithm for NPOESS, *J. Atmos. Oceanic Technol.*, *26*, 1388–1397.

Hutchison, K. D., B. D. Iisager, and B. I. Hauss (2011), The use of global synthetic data for pre-launch tuning of the VIIRS Cloud Mask algorithm, *Int. J. Remote Sens.*, *33*, 1400–1423, doi:10.1080/01431161.2011.571299.

Hutchison, K. D., B. D. Iisager, and R. L. Mahoney (2013), Enhanced snow and ice identification with the VIIRS Cloud Mask algorithm, *Remote Sens. Lett.*, *4*, 929–936, doi:10.1080/2150704X.2013.815381.

National Aeronautics and Space Administration (NASA) (2013a), Joint Polar Satellite System (JPSS) VIIRS Cloud Mask (VCM) Algorithm Theoretical Basis Document (ATBD), Version B. [Available at http://npp.gsfc.nasa.gov/science/sciencedocuments/474-00033_ATBD-VIIRS-Cloud_Mask_B.pdf].

National Aeronautics and Space Administration (NASA) (2013b), Joint Polar Satellite System (JPSS) VIIRS Cloud Mask (VCM) Operational Algorithm Document (OAD), Version D. [Available at http://npp.gsfc.nasa.gov/science/sciencedocuments/2013-07/474-00062_OAD-VIIRS-Cloud-Mask_D.pdf].

Pavolonis, M. J., A. K. Heidinger, and T. Uttal (2005), Daytime cloud typing from AVHRR and VIIRS: Algorithm description, validation, and computation, *J. Appl. Meteorol.*, *44*, 804–826.

Stubenrauch, C. J., et al. (2013), Assessment of global cloud datasets from satellites: Project and database initiated by the GEWEX radiation panel, *Bull. Am. Meteorol. Soc.*, *94*(7), 1031–1049, doi:10.1175/BAMS-D-12-00117.