



HAL
open science

Validation of the LaSRC Cloud Detection Algorithm for Landsat 8 Images

Sergii Skakun, Eric Vermote, Jean-Claude Roger, Christopher Justice, Jeffrey Masek

► **To cite this version:**

Sergii Skakun, Eric Vermote, Jean-Claude Roger, Christopher Justice, Jeffrey Masek. Validation of the LaSRC Cloud Detection Algorithm for Landsat 8 Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2019, 12 (7), pp.2439-2446. 10.1109/JS-TARS.2019.2894553 . hal-03469854

HAL Id: hal-03469854

<https://uca.hal.science/hal-03469854>

Submitted on 26 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Validation of the LaSRC cloud detection algorithm for Landsat 8 images

Sergii Skakun, Eric F. Vermote, Jean-Claude Roger, Christopher O. Justice, and Jeffrey G. Masek

Abstract—This study aims at validating the cloud mask produced by the Land Surface Reflectance Code (LaSRC) for Landsat 8 data. To detect clouds in optical satellite imagery, LaSRC uses quality assurance (QA) layers, which are produced during the atmospheric correction process. The QA layers include a “Cloud mask”, which is based on the estimation of a residual metric showing the quality of aerosol inversion, and “High aerosol”, which shows the impact of aerosols on the derived surface reflectance. Validation is performed using the “L8 Biome” cloud validation dataset, which is produced by the US Geological Survey (USGS), and consists of 96 Landsat 8 scenes distributed globally over 12 different biomes. We show that the LaSRC cloud detection algorithm reliably identifies thick clouds with commission and omission errors less than 4%. Large cloud over-detection errors occur for thin clouds, which is due to the subjectivity of defining and extracting thin clouds in the reference dataset. We conclude this paper with recommendations on using the LaSRC QA layers, and give suggestions on reducing subjectivity, when generating cloud validation datasets.

Index Terms—cloud detection, validation, Landsat 8, LaSRC.

I. INTRODUCTION

CLOUD detection in optical satellite imagery is a prerequisite to producing and delivering high quality higher-order products. Cloud mis-detection (cloud omission) may lead to the retrieval of an erroneous signal from a satellite sensor (e.g. surface reflectance or vegetation indices), and these errors can propagate into higher level thematic products used, for example in agricultural, environmental or natural hazards monitoring applications. On the other hand, over-detection of clouds (cloud commission) may reduce the amount of valuable cloud-free data which, in this case, would be incorrectly identified as clouds. With the availability of high temporal resolution images (5–10 days) acquired by moderate spatial resolution (10–30 m) sensors aboard Landsat 8 and Sentinel-2 satellites, it is of great importance to develop and validate cloud detection algorithms, which can be applied globally for multiple conditions within an operational context. Development of cloud detection algorithms for Landsat-like remote sensing satellites has always been an area of active research. However, only a few algorithms

are used in operational context. This is due, in part, to the complexity of validating cloud detection algorithms and accounting for various cloud types and atmospheric conditions. Collecting reference (‘ground truth’) cloud data is a time and resource consuming task, and is usually performed through photo-interpretation of satellite images by analyst(s). However, in such a case, reference data have a certain degree of subjectivity, which should be considered, when reporting validation results. For example, Scaramuzza et al. (2012) [1] cross-compared results of visual photo-interpretation of clouds on Landsat 5 and Landsat 7 images by three different analysts and found an average 7% error due to subjectivity, with one scene having 25% difference among experts due to the substantial presence of thin clouds and fog.

The Automated Cloud Cover Algorithm (ACCA) [2], [3], [4] has historically served as the main algorithm for cloud screening in Landsat 4, Landsat 5 and Landsat 7 images. Additional adjustments were made to make ACCA work for Landsat 8 images [1]. The ACCA algorithm is based on top-of-atmosphere (TOA) reflectance data and/or thermal bands, and uses a set of empirical rules and several passes per pixel to determine, whether the pixel is cloudy or clear. The Fmask algorithm [5] builds upon ACCA and other algorithms and utilizes a set of physical based rules to identify cloudy pixels in Landsat images. The potentially cloudy pixels are further refined based on segmentation and geometrical properties of identified objects. The Fmask algorithm produces cloud and shadow layers, as well as a snow layer, which is produced using the Normalized Difference Snow Index (NDSI) and brightness temperature. The Fmask algorithm is constantly improving, and several further modifications have recently been proposed, including to be applied to other satellites such as Sentinel-2 [6], [7], [8]. The C programming language version of the FMask algorithm (called CFmask) is used by the US Geological Survey (USGS) in a production environment to deliver the quality assessment band for Landsat 8 products [9]. Other approaches try to incorporate experience gained with detecting clouds using MODerate Resolution Imaging Spectroradiometer (MODIS) images and adapt those rules to Landsat 8 images [10]. In fact, some rules for detecting cirrus clouds and snow in Landsat-8 imagery were derived from those used for MODIS [11], [12].

Manuscript received September 12, 2018. This work was supported by the NASA grant “Support for the HLS (Harmonized Landsat-Sentinel-2) Project” (no. NNX16AN88G).

S. Skakun, J.-C. Roger and C. O. Justice are with the Department of Geographical Sciences, University of Maryland, College Park, MD 20742 USA (e-mail: skakun@umd.edu; roger63@umd.edu; cjustice@umd.edu).

E. F. Vermote is with the Terrestrial Information Systems Laboratory (Code 619), NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA (e-mail: eric.f.vermote@nasa.gov).

J. G. Masek is with the Biospheric Sciences Laboratory (Code 618), NASA Goddard Space Flight Center, Greenbelt, MD 20771 USA (e-mail: jeffrey.g.masek@nasa.gov).

The above mentioned algorithms identify clouds using a single date satellite image. Other methods, such as those proposed in [8], [13], [14], [15], take advantage of multi-temporal satellite images, when a reference cloud-free image is first selected, and an image in question is compared to the reference image to identify clouds. While providing good performance, these methods have intrinsic drawbacks, such as complexity of automatically selecting the reference image or images, and the number of reference images required to accurately identify cloudy pixels.

A large body of work is dedicated to applying supervised machine learning algorithms for cloud detection in satellite imagery. In particular, decision tree (See5) [1], [9], neural networks [16], and support vector machine (SVM) [17] algorithms were trained to identify clouds in satellite imagery using mostly TOA reflectance values. There are two main problems with such approaches. First, supervised machine learning approaches require reference data for algorithms training and testing, collection of which, as mentioned above, is not a trivial task. Second, the use of TOA reflectance values without incorporating data on atmospheric state (e.g. aerosol optical thickness) may run into the problem of data normalization, which cannot be necessarily handled by the algorithms [18], especially when such data are absent from the training set.

In order to cross-compare operationally ready cloud detection algorithms to be used for Landsat quality assurance (QA) data products, Foga *et al.* (2017) [9] performed a large-scale analysis using USGS cloud reference datasets. These validation datasets included “L7 Irish” (for Landsat 7) [4], [1], “L8 SPARCS” [16] and “L8 Biome” [9] for Landsat 8. Among the algorithms tested were various versions of Fmask, ACCA, See5, and the Land Surface Reflectance Code (LaSRC) cloud mask. The LaSRC cloud mask is one of the products generated within the atmospheric correction process [19]. Foga *et al.* (2017) [9] reported a 73.07% overall accuracy (OA) for the LaSRC cloud mask with omission and commission errors 4.70% and 23.90%, respectively. However, that study did not provide algorithm description, did not describe criteria used, did not investigate the source of discrepancy between the LaSRC produced cloud masks and reference cloud masks, and results were not reported for the different types of clouds (thin and thick) available in the reference datasets and different types of biomes. The present study aims to fill this gap and provide a more comprehensive description and analysis of the LaSRC cloud detection algorithm for the “L8 Biome” validation data. In particular, we show that there are two major factors leading to relatively large LaSRC’s commission errors reported in [9]: (1) snow cover, for which LaSRC should not be directly applied; (2) thin (cirrus) clouds, which have a high degree of subjectivity in the reference dataset. We also show some inconsistencies in the “L8 Biome” reference cloud mask, which can be attributed to the subjectivity of manual cloud classification.

II. METHODOLOGY, DATA DESCRIPTION, AND PERFORMANCE METRICS

A. The LaSRC algorithm for cloud detection

LaSRC is a generic atmospheric correction algorithm for estimating land surface reflectance, taking into account absorption by atmospheric gases and scattering by molecules and aerosols [19], [20], and is based on the inversion of the 6SV radiative transfer code [21], [22]. LaSRC was first developed for MODIS [23] and further adopted for other remote sensing sensors such as Visible Infrared Imaging Radiometer Suite (VIIRS) [24]-[25], Operational Land Imager (OLI) aboard Landsat 8 satellite [19], and Multi-Spectral Instrument aboard Sentinel-2 satellites [20]. One of the major steps within this process is the retrieval of aerosol optical thickness (AOT), which is based on the relationship between red (band 4 in Landsat 8), and blue (1 and 2) and shortwave infrared (SWIR, band 6) spectral bands (Fig. 1).

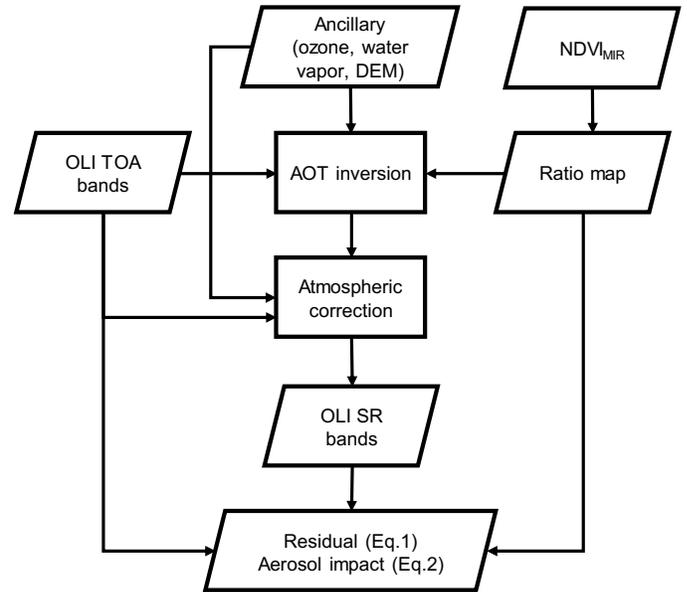


Fig. 1. A general flowchart of LaSRC algorithm.

First, the reference relationship is calculated at the global scale, at a coarse spatial resolution (0.05°) from multi-year MODIS/Terra and Multi-angle Imaging Spectroradiometer (MISR) data, and parameterized as a function of the Normalized Difference Vegetation Index (NDVI_{MIR}), computed using the Mid-IR (2.1 μm) band instead of red [26]: $NDVI_{MIR} = (NIR - MidIR/2)/(NIR + MidIR/2)$. Such parametrization allows accounting for seasonal variability in the red/blue ratio, and downscaling from coarse spatial resolution to the 30 m Landsat 8 pixel scale. In addition to retrieving the AOT map at 30 m resolution, LaSRC algorithm estimates the inversion “residual”, which is the metric of the goodness of AOT inversion:

$$\text{Residual} = \sqrt{\frac{1}{3} \left((\rho_S^1 - r_{1,4} \rho_S^4)^2 + (\rho_S^2 - r_{2,4} \rho_S^4)^2 + (\rho_S^6 - r_{6,4} \rho_S^4)^2 \right)}, \quad (1)$$

where ρ_S^1 , ρ_S^2 , ρ_S^4 and ρ_S^6 are the surface reflectance values in

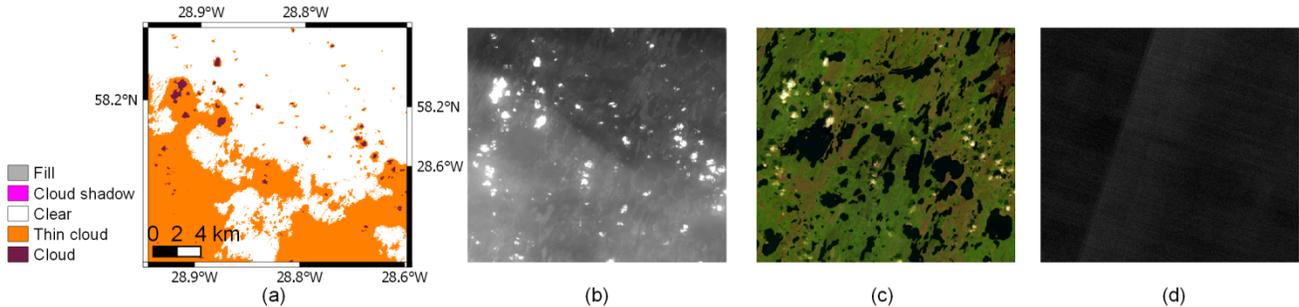


Fig. 2. A subset of the LC80350192014190LGN00 scene from the “L8Biome” cloud validation dataset, showing inconsistencies in detecting thin cloud: (a) classified validation data; (b) blue band 1 reflectance stretched from 0.15 to 0.22 (reflectance units); (c) false color composition of the LaSRC derived surface reflectance values from Landsat 8 bands 6-5-4 (SWIR1-NIR-Red, which correspond to 1.566-1.651 μm , 0.851-0.879 μm and 0.636-0.673 μm , respectively) stretched from 0 to 0.55; (d) cirrus band 9 reflectance (corresponds to 1.363-1.384 μm) stretched from 0.0005 to 0.006.

Landsat 8 bands 1, 2, 4 and 6, respectively, derived using AOT inverted from the red (band 4) and blue (band 1), and $r_{1,4}$, $r_{2,4}$ and $r_{6,4}$ are ratios between red and blue and SWIR bands derived from MODIS and MISR and downscaled at Landsat 8 spatial resolution.

This residual metric (Eq. 1) is the main criterion for detecting thick clouds, since the latter will either prevent the AOT inversion process from convergence, or will drive the residual metric to high values. It should be noted that, unlike most cloud detection algorithms which use TOA reflectance values, the LaSRC cloud detection algorithm uses surface reflectance residual values and incorporates data on atmospheric conditions. These differentiate the proposed approach from the previous ones. A specific threshold should be selected to determine cloudy and clear pixels in Landsat 8 imagery. This threshold was set to 0.05. Pixels with residual values exceeding this threshold are marked as cloudy ones. Pixels adjacent to clouds within 5 pixels (150 m) are separately masked as being ‘adjacent to clouds’. The rationality for marking adjacent cloudy pixels is that cloud boundaries are, as a rule, ambiguous and very difficult to detect precisely at moderate spatial resolution, and, therefore, such masking ensures that unreliable pixels are removed from further processing. The LaSRC algorithm also estimates a metric, which shows the impact of aerosols on the derived surface reflectance values:

$$\text{Aerosol impact} = \text{abs}(\rho_S^1 - \rho_T^1), \quad (2)$$

where ρ_S^1 and ρ_T^1 are surface reflectance and reflectance not corrected for aerosol for band 1 in Landsat 8, respectively. Pixels with aerosol impact values (Eq. 2) exceeding 0.03 (in reflectance units) are marked as ‘high aerosol’ values (low quality pixels) and should be excluded from further processing along with pixels detected as cloudy ones. Table I summarizes QA layers produced by the LaSRC algorithm. The thresholds (Table I) for detecting cloud and high aerosol were determined empirically from visual inspection and 6SV simulations.

B. Data description

The “L8 Biome” cloud validation dataset [27] consists of 96 Landsat 8 scenes, which were selected using a semi-random sampling by biomes [9]. These biomes included barren, forest, grass/crops, shrubland, snow/ice, urban, water and wetlands. For each biome, 12 Landsat 8 scenes were selected, and each

TABLE I
QA LAYERS PRODUCED BY THE LASRC ALGORITHM

QA data layer	Criterion	Threshold
Cloud mask	Residual (Eq. 1)	0.05
Adjacent to clouds	Neighborhood pixels identified as cloudy	5 pixels
High aerosol	Aerosol impact (Eq. 2)	0.03

scene was manually classified into the following classes: clear, thin cloud, cloud, and cloud shadow. It should be noted that no specific threshold was used to detect thin clouds, which were primarily determined by the analyst. Also, the cloud shadow class in the validation dataset was not provided for all the Landsat 8 scenes.

Out of 96 reference Landsat 8 scenes in the “L8 Biome” validation dataset, only 79 scenes were included for evaluating the LaSRC cloud detection algorithm. Among the 17 scenes not considered, 16 scenes were fully covered with snow/ice (12 from the snow/ice biome and 4 from other biomes), and 1 scene had inconsistencies in depicting thin clouds (Fig. 2). Fig. 2 shows a subset of the Landsat 8 scene (LC80350192014190LGN00) with inconsistent detection of thin clouds. While the reference image shows the existence of the boundary between thin clouds and clear pixels, no clear boundary exists in the satellite imagery, including when using true/false color combinations or the cirrus band. As per our analysis, this was the only scene with such major inconsistencies in the validation dataset. The reason for removing snow/ice scenes is that snow/ice pixels are detected as cloudy or high aerosol by the LaSRC algorithm, are considered low quality pixels and should not be used for further processing. This was the main reason for the OA=73.07% reported by Foga *et al.* (2017) [9], when major over-detection (CE=23.90%) was due to the presence of ice/snow.

C. Performance metrics

We used a set of performance metrics, similar to those used in [9], to validate the LaSRC cloud detection algorithm against the “L8 Biome” dataset. These metrics included overall

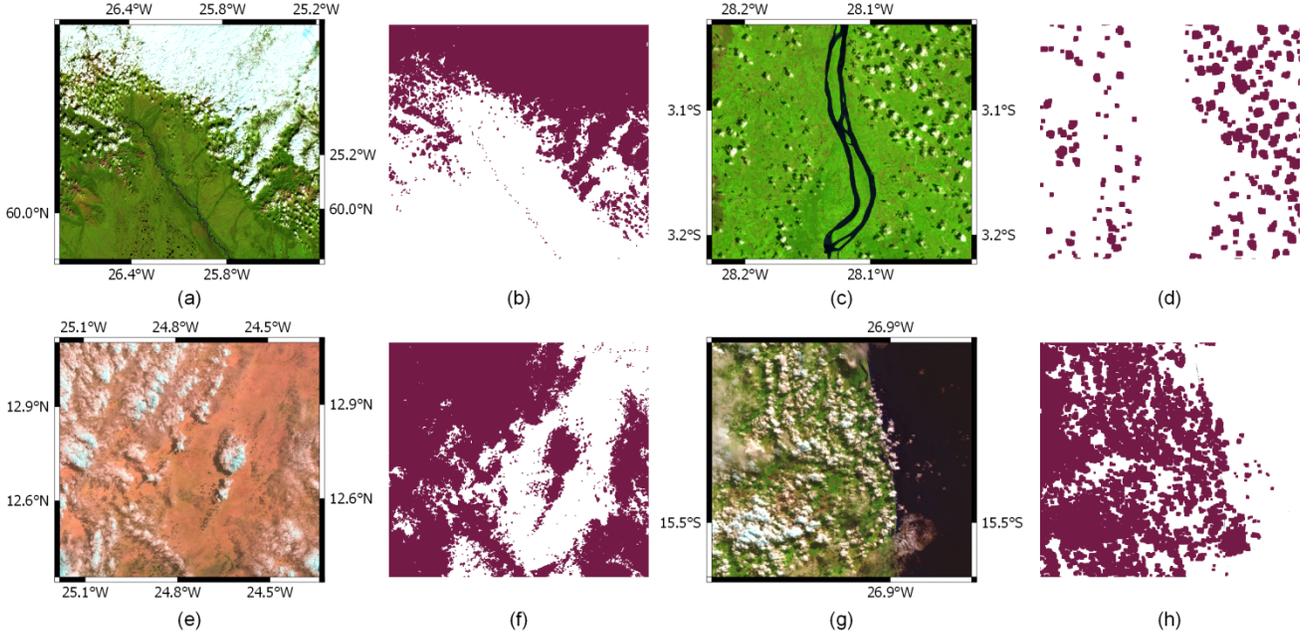


Fig. 3. Examples of cloud masks derived from LaSRC for scenes from different biomes along with the indicated overall accuracy: LC81080182014238LGN00 (wetlands biome, 93.72%) (a-b), LC81750622013304LGN00 (forest, 81.32%) (c-d), LC81750512013208LGN00 (grass/crops, 85.38%) (e-f), and LC82150712013152LGN00 (water, 83.15%) (g-h). Subplots (a), (c), (e), (g) show false color composites of Landsat 8 images (SWIR1-NIR-red), while corresponding cloud masks are show in (b), (d), (f), (h).

TABLE II

LASRC PERFORMANCE AGAINST “L8 BIOME” VALIDATION DATA IN TERMS OF OVERALL ACCURACY (OA, %) EQ. (3), CLOUD OMISSION ERROR (OE, %) EQ. (4), AND COMMISSION ERROR (CE, %) EQ. (5).

QA mask	OA	Thin cloud		Thick cloud		Cloud	
		OE	CE	OE	CE	OE	CE
Cloud	90.30	42.86	14.91	3.86	3.16	13.17	2.68
Cloud & adjacent cloud	89.16	23.13	37.74	2.19	13.02	7.19	10.72
Cloud & high aerosol	83.09	4.94	55.86	0.63	27.58	1.67	22.64
Cloud, adjacent & high aerosol	81.85	1.75	58.00	0.26	29.97	0.62	24.62

III. RESULTS

Fig. 3 shows examples of LaSRC cloud detection performance in some of the Landsat 8 images from the reference dataset.

Fig. 4 shows a histogram of the residual metric values (Eq. 1) for clear and cloudy pixels from all scenes of the reference datasets. We selected a rather conservative threshold value of 0.05 (Table I) and Fig. 4 validates the selection of this threshold, so clear and cloudy pixels can be discriminated.

Table II shows overall performance of the LaSRC cloud detection algorithm for Landsat 8 images in terms of OA, cloud OE and CE, depending on the QA layers applied.

Thick clouds were reliably detected by the LaSRC algorithm with OE=3.86% and CE=3.16%. Adding the adjacent cloud layer to the cloud mask decreased OE from 3.86% to 2.19%, while increased CE from 3.16% to 13.02%. This finding relates to the ambiguity of cloud boundary detection both in the validation dataset and LaSRC derived cloud mask. Adding the high aerosol layer to the cloud mask decreased OE from 3.86% to 0.63%, while substantially increased CE from 3.16% to 27.58%. This shows that the LaSRC cloud and high aerosol masks basically captured all the thick cloud pixels in the validation dataset.

accuracy (OA), cloud omission error (OE) and commission error (CE), and were calculated on a per pixel basis using the following equations:

$$OA = 100 \times \frac{n_{clear_as_clear} + n_{thin_as_cloud} + n_{thick_as_cloud}}{n_{total}}, \quad (3)$$

$$OE = 100 \times \frac{n_{total_cloud} - n_{thin_as_cloud} - n_{thick_as_cloud}}{n_{total_cloud}}, \quad (4)$$

$$CE = 100 \times \left(\frac{n_{clear_as_cloud} + n_{shadow_as_cloud}}{n_{thin_as_cloud} + n_{thick_as_cloud} + n_{clear_as_cloud} + n_{shadow_as_cloud}} \right), \quad (5)$$

where n_{total} is the total number of valid pixels in the scene; $n_{clear_as_clear}$, $n_{thin_as_cloud}$, $n_{thick_as_cloud}$ are the numbers of correctly classified pixels by the LaSRC algorithm; n_{total_cloud} is the total number of cloud pixels (both thin and thick) in the “L8 Biome” dataset; $n_{clear_as_cloud}$ and $n_{shadow_as_cloud}$ are the numbers of pixels incorrectly classified as cloud by the LaSRC algorithm. OE and CE errors were estimated for all clouds as well as for thin and thick clouds separately.

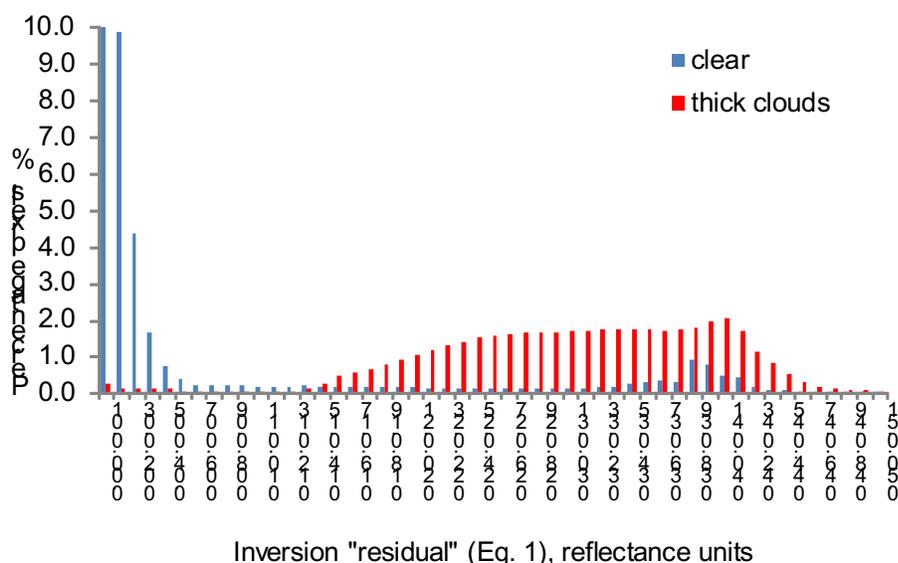


Fig. 4. A histogram of inversion residual metric (Eq. 1) for clear and cloudy pixels from the USGS reference dataset “L8 Biome. Total number of pixels is 2.8×10^9 .

TABLE III
LASRC PERFORMANCE DEPENDING ON THE BIOME TYPE. LASRC QA LAYERS CLOUD AND ADJACENT CLOUD WERE COMBINED.

Category	OA	Thin cloud		Thick cloud		Cloud	
		OE	CE	OE	CE	OE	CE
Barren	90.43	8.51	27.70	4.62	14.13	5.83	10.32
Forest	84.28	40.27	51.58	3.02	14.22	10.54	12.55
Grass/Crops	92.67	25.63	45.05	0.19	9.49	3.92	8.51
Shrubland	87.01	23.51	35.54	0.75	20.70	9.41	15.05
Urban	89.10	7.40	48.22	5.73	13.57	5.98	11.85
Water	90.23	24.97	40.32	0.42	10.66	5.08	9.21
Wetlands	90.38	30.50	24.50	0.12	9.51	9.77	7.36
Mean	89.16	22.97	38.99	2.12	13.19	7.22	10.69
Standard deviation	2.74	11.68	10.26	2.33	3.92	2.62	2.65

TABLE IV
LASRC PERFORMANCE DEPENDING ON THE BIOME TYPE. LASRC QA LAYERS CLOUD, ADJACENT CLOUD AND HIGH AEROSOL WERE COMBINED.

Category	OA	Thin cloud		Thick cloud		Cloud	
		OE	CE	OE	CE	OE	CE
Barren	80.82	1.82	53.74	0.22	33.85	0.72	26.21
Forest	81.98	2.39	58.34	0.48	25.80	0.86	21.79
Grass/Crops	76.61	0.62	76.78	0.02	36.09	0.11	32.54
Shrubland	84.47	0.83	44.94	0.04	33.24	0.34	23.62
Urban	84.22	0.75	65.48	0.78	24.57	0.78	21.75
Water	78.24	3.33	65.24	0.11	29.95	0.73	25.83
Wetlands	87.16	2.31	41.90	0.04	24.77	0.76	18.44
Mean	81.93	1.72	58.06	0.24	29.75	0.61	24.31
Standard deviation	3.70	1.03	12.31	0.29	4.77	0.28	4.50

As expected, thin clouds from the validation dataset were in disagreement with the LaSRC cloud layer only yielding OE=42.86% and CE=14.91%. Adding the high aerosol layer substantially improved detection of thin clouds with OE improving from 42.86% to 4.94%, however at the expense of overdetection, i.e. with CE increasing from 14.91% to 55.86%. Adding the adjacent cloud layer also improved detection of thin clouds: CE decreased from 42.86% to 23.13%. This means that a portion of thin clouds in the validation dataset were detected as the boundary of thick clouds. Further visual inspection confirms this explanation. Overall, performance of the LaSRC algorithm for thin clouds suggests that the LaSRC high aerosol layer can be used to detect thin clouds; however, large discrepancies with the “L8 Biome” validation dataset exist in terms of overdetection, which corresponds to the statement of subjectivity of detecting thin clouds in the reference dataset [9]. Moreover, when analyzing the distribution of the Landsat 8 cirrus band TOA reflectance values for thin clouds from the “L8 Biome” dataset, approximately 47% pixels identified as thin cloud are uniformly distributed in the range 0.0005 to 0.01 (reflectance units) (Fig. 5). This means that to detect such thin

clouds, a low threshold should be selected for the cirrus band. These thin clouds will be misidentified by most existing algorithms, which use the following thresholds (reflectance units): 0.0113 [10], 0.02 and 0.04 [5], [6]. The other 53% of pixels identified as thin cloud in the “L8 Biome” dataset have reflectance values in the cirrus band of more than 0.01.

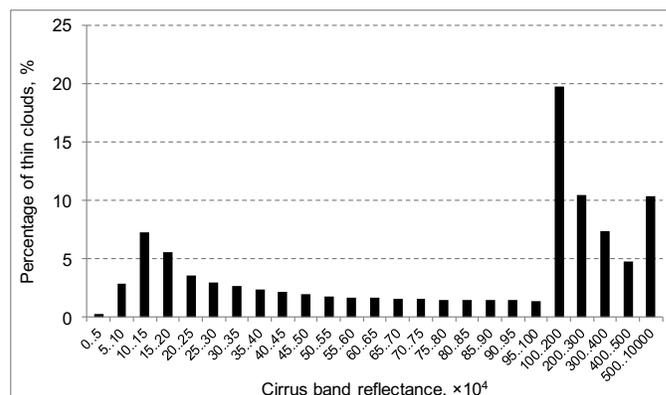


Fig. 5. Distribution of the Landsat 8 cirrus band TOA reflectance for thin cloud from the “L8 Biome” cloud validation dataset.

When thin and thick cloud masks from the validation dataset are combined together, the LaSRC cloud detection algorithm yielded the following performance (Table II): OE=13.17% and CE=2.68% for the cloud layer only, and OE=0.62% and CE=24.62% for the combined cloud, adjacent cloud and high aerosol layers. Again, these numbers are dependent on the subjective criteria used to define the cloud boundary and thin cloud.

Table III and IV show results of the LaSRC categorized by biomes (for two combinations of QA layers: cloud & adjacent cloud, and cloud, adjacent & high aerosol). Results from Table III and IV, in general, follow the same numerical pattern as Table II, and show similar performance across different biomes. For cloud and adjacent cloud layers (Table III), OE=7.22%±2.62% and CE=10.69%±2.65%, compared to OE=7.19% and CE=10.72% from Table II. For all three QA layers combined (Table IV), OE=0.61%±0.28% and CE=24.31%±4.50%, compared to OE=0.62% and CE=24.62% from Table II. These suggest that the LaSRC cloud detection algorithm performs reliably across different land cover conditions.

IV. RECOMMENDATIONS AND OUTLOOK

When using the LaSRC algorithm for atmospheric correction and cloud detection, users are recommended to mask out QA cloud and adjacent cloud pixels. Masking out the LaSRC derived high aerosol pixels depends on users' requirements and applications. Users requiring high quality land surface reflectance values with the lowest uncertainties should also mask out high aerosol pixels, in addition to cloud and adjacent cloud pixels, recognizing that doing so will decrease the number of observations available for analysis. Users with less strict requirements and intending to provide additional post-processing steps, such as time-series filtering/fitting or incorporating quality information into higher-level algorithms, may use surface reflectance values identified as high aerosol and mark them as lower quality compared to other pixels.

Users, working on snow/ice applications and requiring discrimination between clouds and snow/ice, are recommended to use other algorithms, for example Fmask [6], as LaSRC will mask out both snow/ice and clouds as unreliable retrievals.

Further research should be performed to provide a more objective validation of the LaSRC cloud mask, especially for thin clouds. A more formal and quantitative definition of thin cloud should be adopted, which will reduce subjectivity in producing 'ground truth' datasets for cloud validation. This, in our opinion, should combine, in addition to photointerpretation, ground based photos of the sky [28], e.g. from the Atmospheric Radiation Measurement (ARM) Climate Research Facility [29], and data on the atmospheric properties measured by NASA's Aerosol Robotic Network (AERONET) [30] of sites, which are distributed globally. Although ground based photos provide a limited field of view, which makes it difficult to cover the full satellite scene swath (e.g. from Landsat 8), they can be used as a 'seed point' to direct an analyst to manually produce a cloud mask or as training data to classify clouds for the full scene using machine learning approaches. Also, ground based photos together with AOT retrievals from the Aeronet stations can provide a cloud opacity classification [28] and, thus, can provide a quantitative approach for thin clouds definition. An example of a geo-referenced ground photo of clouds made synchronously with Landsat 8 acquisitions over NASA's Goddard Space Flight Center (GSFC) is shown in Fig. 6.

V. CONCLUSION

In this study, we validated the LaSRC derived cloud mask for snow/ice-free Landsat 8 images using the "L8 Biome" cloud validation dataset, generated by USGS. Within the atmospheric correction process, the LaSRC algorithm generates several QA layers related to clouds, which, unlike many other cloud detection algorithms, are based on land surface reflectance and incorporate information on atmospheric properties. The analysis was performed using 79 Landsat 8 scenes, which are distributed globally over multiple biomes. The LaSRC cloud detection algorithm was reliable in identifying thick clouds

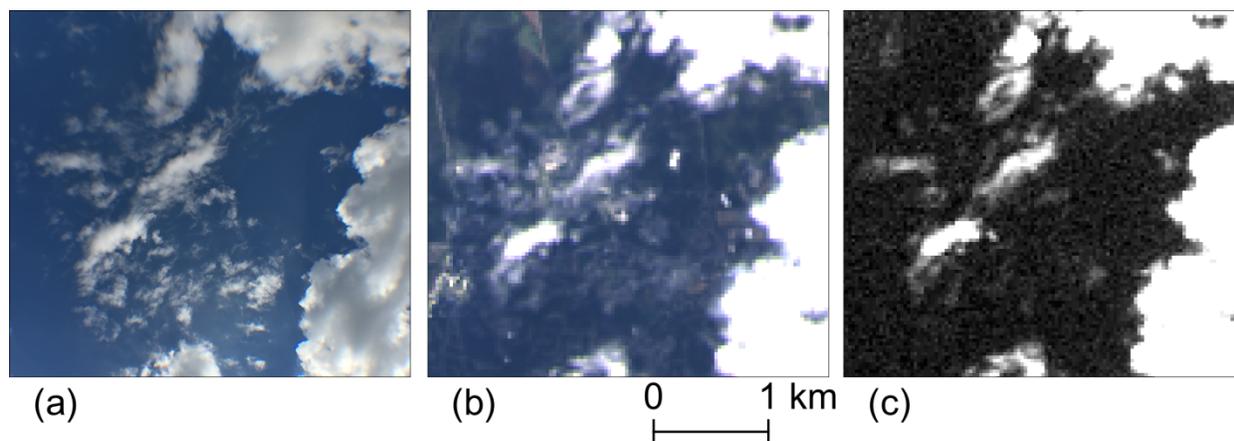


Fig. 6. A geo-referenced ground photo (a) and corresponding Landsat 8 images, namely a true-color composition of bands 4-3-2 (Red-Green-Blue) stretched from 0 to 0.25 (reflectance units) (b) and cirrus band stretched from 0.0005 to 0.006 (c), acquired on October 25, 2017 at 11:46 local time. The AOT at 500 nm was 0.1, according to AERONET measurements over NASA Goddard Spec Flight Center (GSFC) (the center of images is 76.84109°W, 38.99532°N).

from the reference data with commission and omission errors less than 4%. Larger discrepancies were observed for thin clouds, which can be related to the subjectivity in identifying thin cloud in the reference data. We have provided recommendations on using QA layers produced by LaSRC, when users are advised to mask out cloud and adjacent cloud pixels, as well as high aerosol pixels, depending on the applications and requirements. We have outlined further research to potentially reduce subjectivity in producing reference cloud masks in optical satellite imagery. In our opinion, these results are promising, since the same generic LaSRC algorithm can be relatively easily extended to other satellites, for example Sentinel-2 [31], especially taking into account the absence of a thermal band.

REFERENCES

- [1] P. L. Scaramuzza, M. A. Bouchard, and J. L. Dwyer, "Development of the Landsat data continuity mission cloud-cover assessment algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 4, pp. 1140–1154, 2012.
- [2] B. Hollingsworth, L. Chen, S. Reichenbach, and R. Irish, "Automated cloud cover assessment for Landsat TM images," in *Proc. SPIE Conf.*, vol. 2819, Imaging Spectrometry II, 1996, pp. 170–179.
- [3] R. R. Irish, "Landsat 7 automatic cloud cover assessment," in *Proc. SPIE Int. Soc. for Optical Eng.*, Apr. 2000, pp. 348–355.
- [4] R. R. Irish, J. L. Barker, S. N. Goward, and T. Arvidson, "Characterization of the Landsat-7 ETM+ automated cloud-cover assessment (ACCA) algorithm," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 10, pp. 1179–1188, Oct. 2006.
- [5] Z. Zhu, and C. E. Woodcock, "Object-based cloud and cloud shadow detection in Landsat imagery," *Remote Sens. Environ.*, vol. 118, pp. 83–94, 2012.
- [6] Z. Zhu, S. Wang, and C. E. Woodcock, "Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images," *Remote Sens. Environ.*, vol. 159, pp. 269–277, 2015.
- [7] S. Qiu, B. He, Z. Zhu, Z. Liao, and X. Quan, "Improving Fmask cloud and cloud shadow detection in mountainous area for Landsats 4–8 images," *Remote Sens. Environ.*, vol. 199, pp. 107–119, 2017.
- [8] Z. Zhu, and C. E. Woodcock, "Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change," *Remote Sens. Environ.*, vol. 152, pp. 217–234, 2014.
- [9] S. Foga, *et al.*, "Cloud detection algorithm comparison and validation for operational Landsat data products," *Remote Sens. Environ.*, vol. 194, pp. 379–390, 2017.
- [10] M. J. Wilson, and L. Oreopoulos, "Enhancing a simple MODIS cloud mask algorithm for the Landsat data continuity mission," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 723–731, 2013.
- [11] B. A. Baum, *et al.*, "MODIS cloud-top property refinements for Collection 6," *J. Appl. Meteorol. Climat.*, vol. 51, no. 6, pp. 1145–1163, 2012.
- [12] A. Frei, M. Tedesco, S. Lee, J. Foster, D. K. Hall, R. Kelly, and D. A. Robinson, "A review of global satellite-derived snow products," *Adv. Space Res.*, vol. 50, no. 8, pp. 1007–1029, 2012.
- [13] O. Hagolle, M. Huc, D. V. Pascual, and G. Dedieu, "A multi-temporal method for cloud detection, applied to FORMOSAT-2, VENUS, LANDSAT and SENTINEL-2 images," *Remote Sens. Environ.*, vol. 114, no. 8, pp. 1747–1755, 2010.
- [14] B. Wang, A. Ono, K. Muramatsu, and N. Fujiwara, "Automated detection and removal of clouds and their shadows from Landsat TM images," *IEICE Trans. Inf. Sys.*, vol. 82, no. 2, pp. 453–460, 1999.
- [15] D. Frantz, A. Röder, T. Udelhoven, and M. Schmidt, "Enhancing the detectability of clouds and their shadows in multitemporal dryland Landsat imagery: Extending Fmask," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 6, pp. 1242–1246, 2015.
- [16] M. J. Hughes, and D. J. Hayes, "Automated detection of cloud and cloud shadow in single-date Landsat imagery using neural networks and spatial post-processing," *Remote Sens.*, vol. 6, no. 6, pp. 4907–4926, 2014.
- [17] Y. Yuan, and X. Hu, "Bag-of-words and object-based classification for cloud extraction from satellite imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 8, pp. 4197–4205, 2015.
- [18] S. A. Ackerman, R. E. Holz, R. Frey, E. W. Eloranta, B. C. Maddux, and M. McGill, "Cloud detection with MODIS. Part II: validation," *J. Atmosph. Ocean. Technol.*, vol. 25, no. 7, pp. 1073–1086, 2008.
- [19] E. Vermote, C. Justice, M. Claverie, and B. Franch, "Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product," *Remote Sens. Environ.*, vol. 185, pp. 46–56, 2016.
- [20] G. Doxani, E. Vermote, J.-C. Roger, F. Gascon, S. Adriaensens, D. Frantz, *et al.*, "Atmospheric correction inter-comparison exercise," *Remote Sens.*, vol. 10, no. 2, art. no. 352, 2018.
- [21] E. Vermote, D. Tanre, J. Deuze, M. Herman, and J. Morcette, "Second Simulation of the Satellite Signal in the Solar Spectrum, 6S: an overview," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 3, pp. 675–686, 1997.
- [22] S. Y. Kotchenova, E. F. Vermote, R. Matarrese, and F. J. Klemm Jr, "Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part I: Path radiance," *Appl. Optics*, vol. 45, no. 26, pp. 6762–6774, 2006.
- [23] E. F. Vermote, and S. Kotchenova, "Atmospheric correction for the monitoring of land surfaces," *Journal of Geophysical Research: Atmospheres*, vol. 113, art. no. D23, 2008.
- [24] E. Vermote, C. Justice, and I. Csizsar, "Early evaluation of the VIIRS calibration, cloud mask and surface reflectance Earth data records," *Remote Sensing of Environment*, vol. 148, pp. 134–145, 2014.
- [25] S. Skakun, C. O. Justice, E. Vermote, and J. C. Roger, "Transitioning from MODIS to VIIRS: an analysis of inter-consistency of NDVI data sets for agricultural monitoring," *International Journal of Remote Sensing*, vol. 39, no. 4, pp. 971–992, 2018.
- [26] R. C. Levy, L. A. Remer, S. Mattoo, E. F. Vermote, and Y. J. Kaufman, "Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance," *J. Geophys. Res.: Atmosph.*, vol. 112, art. no. D13211, 2007.
- [27] L8 Biome Cloud Validation Masks. (2016). U.S. Geological Survey, data release. [Online]. Available: <http://doi.org/10.5066/F7251GDH>
- [28] M. S. Ghonima, B. Urquhart, C. W. Chow, J. E. Shields, A. Cazorla, and J. Kleissl, "A method for cloud detection and opacity classification based on ground based sky imagery," *Atmosph. Meas. Techn.*, vol. 5, no. 11, pp. 2881–2892, 2012.
- [29] G. M. Stokes, and S. E. Schwartz, "The Atmospheric Radiation Measurement (ARM) Program: Programmatic background and design of the cloud and radiation test bed," *Bull. Am. Meteorol. Soc.*, vol. 75, no. 7, pp. 1201–1221, 1994.
- [30] B. N. Holben, *et al.*, "AERONET—A federated instrument network and data archive for aerosol characterization," *Remote Sens. Environ.*, vol. 66, no. 1, pp. 1–16, 1998.
- [31] S. Skakun, E. Vermote, J.-C. Roger, and C. Justice, "Multispectral Misregistration of Sentinel-2A Images: Analysis and Implications for Potential Applications," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2408–2412, 2017.



Sergii Skakun received the M.S. (Hons.) degree in applied mathematics from the Physics and Technology Institute, NTUU "Kyiv Polytechnic Institute," Kyiv, Ukraine, in 2004, and the Ph.D. degree in computer science from the National Academy of Sciences of Ukraine, in 2005.

He is currently an Associate Research Professor at Department of Geographical Sciences at University of Maryland, College Park, MD, USA, and Research Scientist at the Terrestrial Information Systems Laboratory at NASA Goddard Space Flight Center (GSFC), Greenbelt, MD, USA. His research interests are in advancing methods, models and emerging technologies in the area of data science for heterogeneous remote sensing data fusion, processing and analysis, and their applications to the areas of societal benefit.



Eric Vermote received the Ph.D. degree in atmospheric optics from the University of Lille, Lille, France, in 1990.

Until 2012, he was a Research Professor with the Department of Geographical Sciences, University of Maryland, College Park, MD, USA. He is currently a Research Physical Scientist with the Terrestrial Information Systems Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD. He is a member of the Moderate Resolution Imaging Spectroradiometer Science (MODIS) Team, the Soumi-NPP VIIRS Science Team, and the Landsat Science Team, and is responsible for the atmospheric correction over land surfaces in the visible to middle infrared. His research interests are in radiative transfer modeling, vicarious calibration, atmospheric correction, aerosol retrieval, and the generation of climate data record for terrestrial studies.



Jean-Claude Roger received the Ph.D. degree in atmospheric physics from the University of Lille 1, Villeneuve-d'Ascq, France, in 1991 for a thesis on "Spatial Studies in Polarized Light - Preparation of the POLDER instrument".

He became a Full-Professor in atmospheric physics in 2007 at the University Blaise Pascal, Clermont-Ferrand, France. He is currently a Research Professor with the Department of Geographical Sciences, University of Maryland, College Park, MD, USA. He is also a Research Scientist with the Terrestrial Information Systems Laboratory (Code 619), NASA Goddard Space Flight Center, Greenbelt, MD, USA. He does research for land and atmosphere applications (remote sensing, radiative transfer, polarization, field experiments, etc.). He is particularly involved in atmospheric correction, aerosols characterization for climatic impact, land product with applications to agriculture, satellite Cal/Val exercises and radiative transfer code development such as 6S. He is a Co-PI of the CEOS-WGCV ACIX and CEOS-WGCV ACIX II – CMIX activities.



Chris Justice received the Ph.D. degree from the University of Reading, Reading, U.K., in 1977.

He became a Professor and Research Director of Geography with the University of Maryland, in 2001, and in 2010 became the Department Chair to the Department of Geographical Sciences. He is a Program Scientist for NASA's Land Cover Land Use Change (LCLUC) Program. He is the Land Discipline Leader for the NASA Moderate Imaging Spectroradiometer (MODIS) and the Soumi-NPP VIIRS Science Team and is responsible for the MODIS Fire Product. He is the Co-Chair of the NASA LANCE User Working Group.

He is a member of the international GOCF/GOLD-Fire Implementation Team. He is the Co-Chair of the GEOGLAM Initiative. His current research is on land cover and land use change, land observations and data products, global agricultural monitoring, and their associated information technology and decision support systems.



Jeffrey G. Masek received the B.A. degree in geology from Haverford College, Haverford, PA, USA, in 1989 and the Ph.D. degree in geological sciences from Cornell University, Ithaca, NY, USA, in 1994.

He is currently a Chief of the Biospheric Sciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA. He also serves as the NASA Landsat Project Scientist, and has previously held positions at University of Maryland, College Park, MD, USA; Hughes Information Systems; and Cornell University. His research interests include mapping land cover change in temperate environments, application of advanced computing to remote sensing, and satellite remote sensing techniques.