Geoscience and Remote Sensing Letters



Implementation on Landsat Data of a Simple Cloud Mask Algorithm Developed for MODIS Land Bands

Journal:	Geoscience and Remote Sensing Letters
Manuscript ID:	Draft
Manuscript Type:	Letters
Date Submitted by the Author:	n/a
Complete List of Authors:	Oreopoulos, Lazaros; NASA-GSFC, Laboratory for Atmospheres Wilson, Michael; UMBC, GEST Varnai, Tamas; University of Maryland, Baltimore County, JCET
Key Words:	Clouds, Satellites

SCHOLARONE™ Manuscripts

Implementation on Landsat Data of a Simple Cloud Mask Algorithm Developed for MODIS Land Bands

by

Lazaros Oreopoulos¹, Michael J. Wilson^{2,1} and Tamás Várnai^{3,1}

- 1. Laboratory for Atmospheres, NASA-Goddard Space Flight Center, Code 613.2, Greenbelt, MD 20771, Lazaros. Oreopoulos@nasa.gov
- 2. Goddard Earth Sciences and Technology Center, University of Maryland Baltimore County, Baltimore, MD 21250, Michael.J.Wilson-1@nasa.gov
- 3. Joint Center for Earth Systems Technology, University of Maryland Baltimore County, Baltimore, MD 21250, Tamas. Varnai@nasa.gov

Submitted to

IEEE Geoscience and Remote Sensing Letters

August 2010

Abstract

This letter assesses the performance on Landsat-7 images of a modified version of a cloud masking algorithm originally developed for clear-sky compositing of Moderate Resolution Imaging Spectroradiometer (MODIS) images at northern mid-latitudes. While data from recent Landsat missions include measurements at thermal wavelengths, and such measurements are also planned for the next mission, thermal tests are not included in the suggested algorithm in its present form to maintain greater versatility and ease of use. To evaluate the masking algorithm we take advantage of the availability of manual (visual) cloud masks developed at USGS for the collection of Landsat scenes used here. As part of our evaluation we also include the Automated Cloud Cover Assessment (ACCA) algorithm that includes thermal tests and is used operationally by the Landsat-7 mission to provide scene cloud fractions, but no cloud masks. We show that the suggested algorithm can perform about as well as ACCA both in terms of scene cloud fraction and pixel-level cloud identification. Specifically, we find that the algorithm gives an error of 1.3% for the scene cloud fraction of 156 scenes, and a root mean square error of 7.2%, while it agrees with the manual mask for 93% of the pixels, figures very similar to those from ACCA (1.2%, 7.1%, 93.7%).

Index Terms

Clouds, Enhanced Thematic Mapper, Landsat, Moderate Resolution Imaging Spectroradiometer (MODIS), masking, Operational Land Imager (OLI), remote sensing, satellite, Thermal InfraRed Sensor (TIRS)

I. INTRODUCTION

The presence of clouds in images acquired by the Landsat program is usually an undesirable, but generally unavoidable fact. With the emphasis of the program being on land imaging, suspended liquid/ice particles fully or partially obscure the desired observational target. Knowledge of cloud amount in a Landsat scene and the location of clouds is therefore valuable information that facilitates proper scene selection by Landsat data users, scene compositing from multiple scenes, and scheduling of future acquisitions [1]. Presently, Landsat-7 images come with metadata that provide the total cloud fraction of the scene (the fraction of cloudy over the total number of pixels) as well as the cloud fraction in each of the four scene quadrants. These cloud "scores" are generated by the Automated Cloud Cover Assessment (ACCA) algorithm [2]. Unfortunately, a classification of individual pixels as either cloudy or cloud-free (i.e., a "cloud mask") is not provided, forcing data users to perform their own cloud screening whenever their application requires it. This will change for the next Landsat mission, the Landsat Data Continuity Mission (LDCM), for which a cloud mask product is planned [3].

The purpose of this letter is to revisit a simple clear pixel detection algorithm developed for MODIS 250/500 m land bands [4], unassisted by thermal data, and examine whether it can provide pixel-level clear-cloudy sky discrimination for Landsat scenes at very small computational cost. While we apply the algorithm in this paper only to Enhanced Thematic Mapper Plus (ETM+) Landsat-7 data, it should be also applicable to historic Landsat-4 and Landsat-5 data from the Thematic Mapper instrument, as well as data from the Operational Land Imager (OLI) sensor of the upcoming LDCM. The cloud-shadow detection component, of the original algorithm has not yet been fully validated in our Landsat implementation and will not be further discussed in this paper.

II. THE CLOUD MASKING ALGORITHM

The clear/cloud mask scheme introduced in [4] (hereafter "LTK scheme") is a simple threshold scheme that uses only four MODIS 250/500 m resolution bands, specifically bands 1, 2, 3, and 6. These bands have approximate spectral equivalents in the ETM+ instrument aboard Landsat-7 (Table 1). The LTK scheme threshold selection for surface type classification and cloud detection is based on typical spectral signatures of five major pixel classes: non-vegetated land, vegetated land, water, ice/snow and cloudy pixels, as depicted in Fig. 6 of [4]. The scheme successively applies threshold tests to first classify non-vegetated pixels, followed by the classification of ice/snow, water, and cloudy pixels. Any pixels not classified to any of the above classes are assigned to the vegetated class. A flow chart presenting our modified LTK scheme is provided in Fig. 1. After exhaustive testing of a variety of plausible adjustments to the LTK thresholds to improve its performance, we settled on two threshold modifications in the last step of the algorithm that separates cloud and vegetated pixel classes. These new thresholds resulted in substantially better agreement between the cloud/clear masks from LTK and those from a manual "truth" mask (discussed below) for a large collection of ETM+ scenes. Both the original and modified LTK scheme threshold values are provided in the last box of the Fig. 1 flow chart, and the performance of both variants of the scheme are contrasted in the next section. The decrease of the ETM+ band 1 threshold (MODIS band 3) is consistent with the values of the spectral reflectance plot for vegetated land shown in Fig. 6 of [4], which do not seem to exceed 0.1. However, the decrease of the band 5 threshold seems somewhat inconsistent with the observed values of MODIS band 6 reflectances in the same plot, which seem to range between 0.1 and 0.18. The fact that a lower value appears to work better for Landsat may be due to the difference in spectral range and central wavelength location of the MODIS and Landsat bands.

Numerous other threshold modifications also improved upon the original scheme, but none worked as well as the two modifications that were eventually adopted. While we realize that a cloud masking scheme developed for an instrument with similar spectral characteristics, but with bands of different spectral widths, different spatial resolution, and off-nadir viewing capabilities, should not necessarily translate perfectly to ETM+, we found nevertheless that in practice the LTK scheme carries over quite well from MODIS to Landsat observations.

III. ALGORITHM IMPLEMENTATION ON LANDSAT-7 SCENES

The modified LTK scheme is applied to a collection of 156 Landsat scenes, a subset of the 212 scenes used by [2] to evaluate the performance of ACCA scene-averaged cloud fractions. The criteria used to select the original dataset of 212 scenes is provided in [2]. These scenes are approximately evenly apportioned among 9 latitude zones covering the entire globe. The present subset of 156 consists of the scenes for which it was determined by USGS-EROS personnel that a reliable cloud mask can be obtained. The manual mask was developed via a visual assessment procedure [5]: Three experienced USGS imagery analysts performed manual assessment of the scenes in [2]. 11 scenes were examined by all three in order to obtain the approximate error of the procedure, which was found to be about 7% on average [5]. The process involved opening each full resolution scene in Adobe Photoshop in a variety of RGB combinations, including overlays of the (resampled) thermal band when necessary. The analysts then used appropriate Photoshop image processing functions to isolate clouds. Two classes of clouds were identified: thick and thin. Cloud pixels were labeled as thin if they were transparent but still visually identifiable as clouds. For the purposes of this paper, no distinction is made between thin and thick clouds in the quantitative metrics of the LTK scheme performance, but only when

remote sensing group at Boston University (BU) [6], which recommended that the collection of 156 scenes be further reduced by 14. Here we provide results for both the 156 and 142 scene sets. As will be shown, both the ACCA and the LTK schemes agree better with the manual mask for the smaller subset, a fact that seems to give further credence to the BU evaluation. In addition to the USGS manual mask, we also have for these scenes the manually-determined cloud fraction used as "truth" in [2]. The set of 156 contains 35 scenes from midlatitudes (30°-45°N or 30°-45°S), 41 scenes from the subtropics (15°-30°N or 15°-30°S), 21 scenes from the tropics (15°S – 15°N), 33 scenes from austral or boreal latitudes (45°-60°S or 45°-60°N), and 26 scenes from the polar regions (60°-90°N or 60°-90°S). The BU group flagged as unreliable 7 polar visual masks (4 from south and 3 from north), 1 austral mask, 2 tropical masks, 1 midlatitude south mask, and 3 midlatitude north masks. The fact that the original polar group of 44 scenes was reduced by the combined manual mask screening re-evaluation of USGS and BU to 19 scenes should come as no surprise, since cloud/ice/snow discrimination is very difficult even in visual image analysis.

Fig. 2 shows the outcome resulting from the original LTK scheme for a sample scene with clear vegetated and non-vegetated land pixels, water pixels, and a fair amount of cloudy pixels. The scheme appears to perform a reasonably good pixel classification and the clear/cloud mask, although slightly worse than ACCA, makes the correct distinction between clear and cloudy pixels more than 90% of the time.

The overall performance of the LTK scheme in terms of the "cloud score" (the cloud fraction of the entire scene) can be seen in Fig. 3 for the 156 scene (top) and the 142 scene (bottom) sets. The left panels correspond to the original LTK scheme and the right panels to our modified version. ACCA results are included for comparison. The legends in each plot also

contain summary metrics such as the overall bias in scene cloud fraction, the root mean square error of the scene cloud fraction, and the number of "bad" scenes, defined here as scenes with LTK or ACCA cloud fraction absolute differences from the manual mask ("cloud fraction errors") that exceed 10% (=0.1 when cloud fraction is measured in a scale from 0 to 1). Modifying the LTK scheme results in noticeable improvements which bring it on par, according to our performance metrics, with the more complex ACCA scheme which includes thermal tests.

Using the same panel arrangement as in Fig. 3, Fig. 4 shows the performance of the original and modified LTK schemes in terms of the percentage of pixels for which the algorithms agree that a pixel is clear or cloudy ("mask agreement"). Again, the ACCA results are included for comparison, allowing us to create a scatterplot of this metric where each scene is represented by a point. Lines at the 80% agreement level are meant to isolate the poorer performers, discussed further below. The benefits of modifying the LTK scheme are evident, as it more closely approaches ACCA levels of performance. Note that that only 9 scenes have LTK mask agreements below 80% (4 for the set of 142), but still always above 65%.

If good cloud masking capabilities is the objective, then the mask agreement of Fig. 4 is a better evaluator of the scheme's skill. If only the scene-average cloud fraction (score) is of interest, and cancelling pixel misclassifications are tolerable, then the results of Fig. 3 are more relevant. An obvious question is whether our collection of scenes includes cases with small scene-average cloud fraction errors, but low mask agreement. Fig. 5 is a scatterplot of mask agreement vs. cloud fraction error. As expected, there is a strong anticorrelation between the two quantities. Scenes with small cloud fraction error usually exhibit high values of mask agreement. With an arbitrary choice of 5% cloud fraction error and 80% mask agreement, only one scene falls in the quadrant that indicates good cloud fraction estimates due to cancelling errors.

- We now examine why our modified LTK algorithm performs poorly for certain scenes, either in terms of cloud fraction errors or mask agreements. We identified scenes belonging to both of these categories of poor performance based on 10% (17 out of 156 scenes) and 80% (9 out of 156 scenes) thresholds, respectively, in order to investigate this question. 7 of the 9 scenes that do not pass the 80% mask agreement threshold also belong to the subset of 17 scenes that do not satisfy the 10% cloud fraction error criterion, so the number of unique "bad" scenes is 19. These 19 scenes have the following characteritics:
- (a) 5 belong to the south pole latitude zone where cloud discrimination from ice and snow isnotoriously difficult.
- 125 (b) 7 exhibit greater than 10% cloud fraction error also between the manual USGS cloud scores 126 and the manual cloud scores of [2], which suggests that these scenes pose cloud identification 127 challenges even when visually inspecting RGB composites.
- 128 (c) 7 exhibit also poor ACCA performance (greater than 10% scene cloud fraction error); only
 129 three of these belong to the 7 of category (b), yielding a total of 11 "difficult" scenes.
- 130 (d) For these 11 scenes, 6 have less than 80% pixel-level agreement for the ACCA algorithm as

 131 well, and 7 were deemed to have unreliable USGS visual masks by the BU team (i.e. they

 132 belong to the set of 156, but not to the set of 142).
 - (e) 12 of the 19 scenes have high amounts of thin clouds, specifically a ratio of thin cloud pixels to total number of cloudy pixels higher than the median value of 0.31 (derived from the 134 out of 156 scenes with non-zero cloudiness). 3 scenes have actually a ratio greater than 0.9 while only 2 have a ratio smaller than 0.1. Thin cloud is very difficult to identify in land-dominated scenes with a simple threshold algorithm relying only on solar bands. The LTK scheme should therefore be used with caution for cloud masking when visual image

inspection or other evidence (e.g., thermal band signatures) indicates the presence of thin clouds.

IV. SUMMARY AND CONCLUSIONS

We have revisited a cloud/clear masking algorithm initially developed for MODIS clear-image compositing and applied two threshold modifications that significantly improve its performance when applied to a set of 156 Landsat scenes selected to cover the full range of Earth geographical zones. The algorithm uses four Landsat solar bands that roughly correspond to the MODIS bands of the original algorithm. We have found that despite its simplicity the algorithm works quite well, giving a bias error of 1.3% for the scene cloud fraction of the 156 scenes, and a root mean square error of 7.2%. The algorithm agrees with the pixel classification (clear/cloudy) of a manual (visual) mask for 93% of the pixels, on average. These performance metrics (1.3%, 7.2%, 93.0%) are very close to those (1.2%, 7.1%, 93.7%) of the more sophisticated Landsat-7 operation algorithm (ACCA), which also incorporates thermal band tests.

Two motivations for bypassing thermal tests are simplicity and speed. The modified LTK scheme of this paper can be coded much easier by a non-expert than ACCA with its involved "pass two" portion which, while helpful for re-classifying ambiguous pixels, has the drawback of added complexity and greater execution time. Another reason to consider the scheme of this paper, that may become relevant for future missions such as LDCM, is that cloud masking can continue to operate even if no thermal data are available. This possibility is certainly not remote for LDCM given the fact that solar and thermal sensing capabilities will be partitioned between two instruments, OLI and TIRS, the latter of which has a shorter design life. While this modified LTK scheme can be applied to historical Landsat data, its availability for future Landsat

acquisitions is also important since it provides an extra cloud masking assessment opportunity for whatever operational cloud mask algorithm is eventually adopted. The algorithm can also be part of consensus cloud masks or masks with confidence level flags based on the degree of agreement between an ensemble of distinct masking schemes.

Weaknesses of the modified LTK scheme that we have exposed in this work include its limited ability to identify thin clouds and clouds over snowy or icy surfaces. Further work can conceivably be undertaken to add a thermal component to the LTK algorithm and/or to add threshold tests for the 1.38 µm "thin cirrus" band of LDCM's OLI instrument. Both of these elements have the potential to improve the LTK-based scheme significantly. Finally, it is recognized that the testing of a masking algorithm on a collection of 156 scenes is not exhaustive or conclusive, even if the scenes were selected to encompass most of the surface, solar geometry and cloud type diversity encountered around the globe. Unfortunately, scheme evaluations that involve manually generated masks cannot by nature be very extensive because the laborious nature of visual pixel classification.

ACKNOWLEDGMENT

We would to thank Richard Irish (formerly of SSAI/NASA-GSFC) for providing the 212 Landsat-7 scenes used in [2] and with his manual cloud scores; Pat Scaramuzza (USGS/SGT) for providing the visual cloud masks for the same dataset; and Curtis Woodcock (Boston University) for his team's assessment of the USGS masks. Finally, we would like to acknowledge funding by USGS under contract 06CRCN0022.

183 References

- 184 [1] T. Arvidson, J. Gasch, and S. N. Goward, "Landsat 7's long term acquisition plan An
- innovative approach to building a global archive", Remote Sensing of Environment, vol. 78, no.
- 186 1-2, pp. 13-26, Oct. 2001.
- 187 [2] R. R. Irish, J. L. Barker, S. N. Goward, and T. Arvidson, "Characterization of the Landsat-7
- 188 ETM Automated Cloud-Cover Assessment (ACCA) Algorithm", Photogrammetric Engineering
- 489 & Remote Sensing, vol. 72, no 10, pp. 1179-1188, Oct. 2006.
- 190 [3] T. Loveland, United States Geological Survey Earth Resources Observation Systems
- 191 (EROS) Data Center (EDC), Sioux Falls, SD, private communication, November 2009.
- 192 [4] Y. Luo, A. P. Trishchenko, and K. V. Khlopenkov, "Developing clear-sky, cloud and cloud
- shadow mask for producing clear-sky composites at 250-meter spatial resolution for the seven
- MODIS land bands over Canada and North America," Remote Sensing of Environment, vol. 112,
- 195 no. 12, pp. 4167-4185, Dec. 2008.
- 196 [5] P. Scaramuzza, United States Geological Survey Earth Resources Observation Systems
- 197 (EROS) Data Center (EDC), Sioux Falls, SD, private communication, November 2009.
- 198 [6] C. Woodcock, Boston University Boston, MA, private communication, April 2010.

List of Tables

200 Table I

CORRESPONDENCE BETWEEN MODIS AND ETM+ BANDS

MODIS	ETM+
Band 1 (620-670 nm)	Band 3 (630-690 nm)
Band 2 (841-876 nm)	Band 4 (750-900 nm)
Band 3 (459-479 nm)	Band 1 (450-515 nm)
Band 6 (1628-1652 nm)	Band 5 (1550-1750 nm)

Correspondence between the MODIS bands used in the original LTK scheme and the ETM+

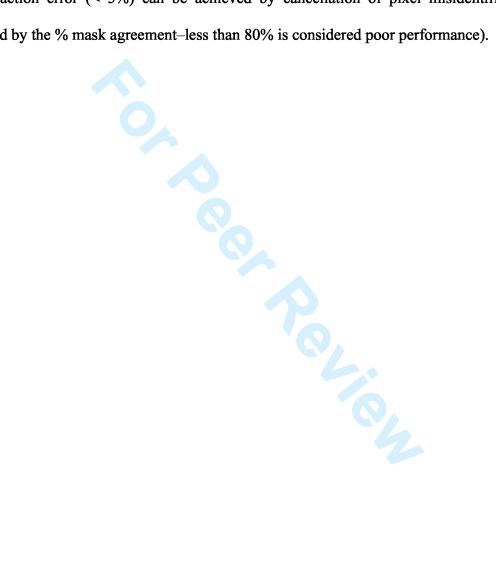
bands used in this study for the modified LTK scheme.

List of Figures

- Figure 1. Flow chart depicting the LTK clear/cloud masking part of the LTK scheme. The numbers in parentheses in red are the original LTK thresholds for the reflectances of the equivalent MODIS bands.
- Figure 2. An example of applying the original LTK scheme on a Landsat-7 scene acquired on April 22, 2001. (top) true color RGB image; (bottom) LTK pixel classification.
 - Figure 3. Comparison between manually-determined and cloud mask algorithm scene cloud fractions (ACCA or LTK). Top row corresponds to the set of 156 Landsat scenes, and bottom row to the reduced set of 142 scenes (see text for details). The left plots are for the original LTK scheme, while the right plots show results after our modification. "Bad" in the legends refers to a scene for which the cloud fraction error is greater than 10% (=0.1 when cloud fraction is measured in a scale from 0 to 1).
 - Figure 4. Comparison between LTK and ACCA mask agreement (in %) for the original LTK scheme (left panels) and the modified LTK scheme (right panels). The top row is for the set of

156 Landsat scenes while the bottom row is for the reduced set of 142 Landsat-7 scenes. "Bad" in the legends refers to a scene for which the mask agreement is less than 80%.

Figure 5. Scatterplot of the mask agreement of the modified LTK scheme against cloud fraction error. The left lower quadrant identifies the number of scenes (one in this case) where a low cloud fraction error (< 5%) can be achieved by cancellation of pixel misidentifications (as measured by the % mask agreement—less than 80% is considered poor performance).



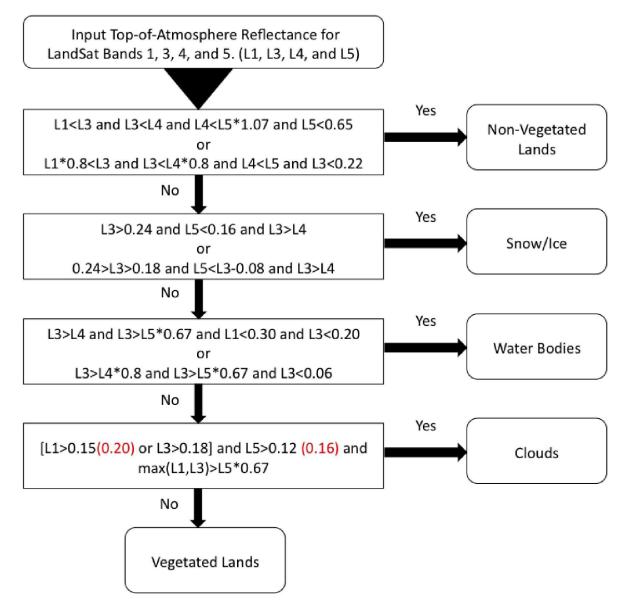


Figure 1. Flow chart depicting the LTK clear/cloud masking part of the LTK scheme. The numbers in parentheses in red are the original LTK thresholds for the reflectances of the equivalent MODIS bands.

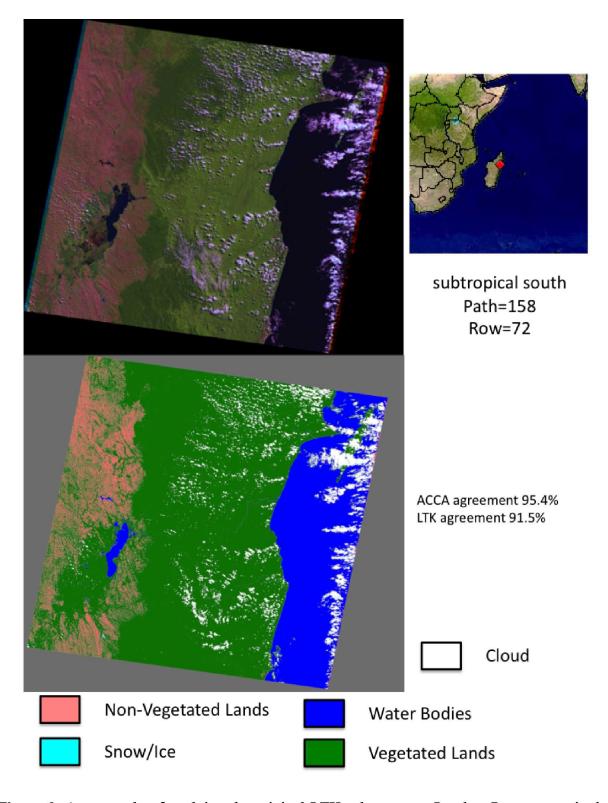


Figure 2. An example of applying the original LTK scheme on a Landsat-7 scene acquired on April 22, 2001. (top) true color RGB image; (bottom) LTK pixel classification.

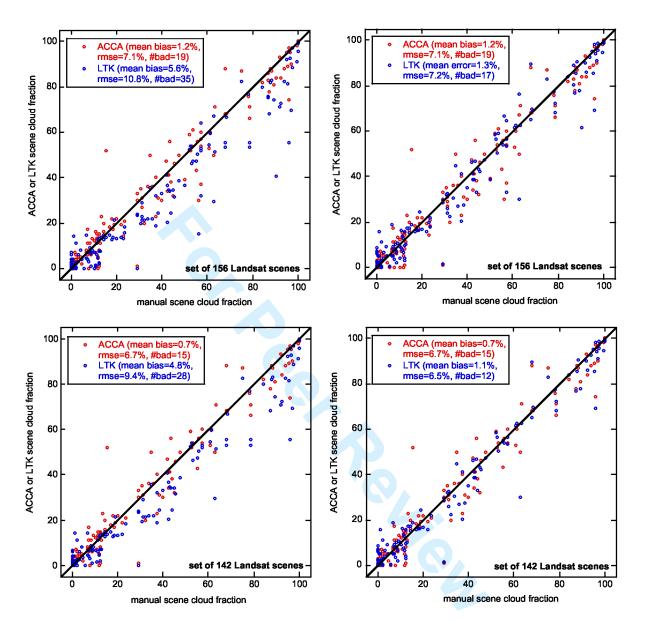


Figure 3. Comparison between manually-determined and cloud mask algorithm scene cloud fractions (ACCA or LTK). Top row corresponds to the set of 156 Landsat scenes, and bottom row to the reduced set of 142 scenes (see text for details). The left plots are for the original LTK scheme, while the right plots show results after our modification. "Bad" in the legends refers to a scene for which the cloud fraction error is greater than 10% (=0.1 when cloud fraction is measured in a scale from 0 to 1).

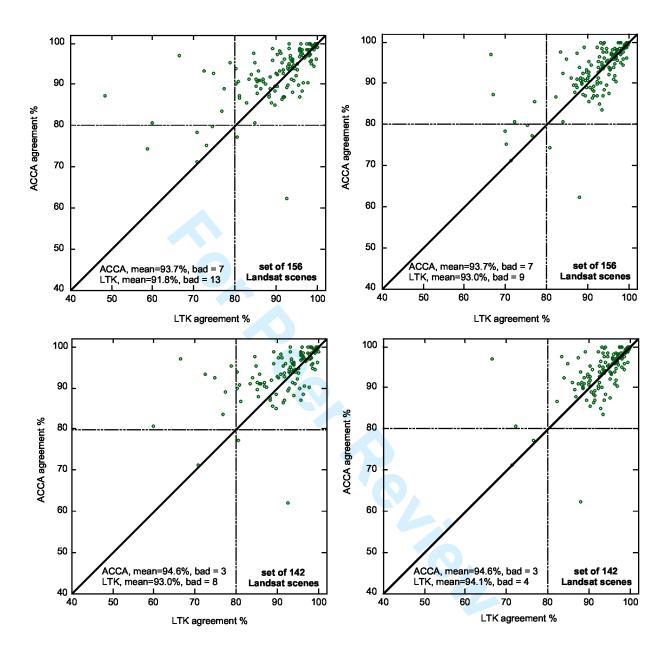
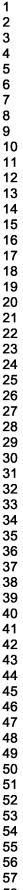


Figure 4. Comparison between LTK and ACCA mask agreement (in %) for the original LTK scheme (left panels) and the modified LTK scheme (right panels). The top row is for the set of 156 Landsat scenes while the bottom row is for the reduced set of 142 Landsat-7 scenes. "Bad" in the legends refers to a scene for which the mask agreement is less than 80%.



set of 156 Landsat scenes LTK agreement % O O absolute scene cloud fraction error

Figure 5. Scatterplot of mask agreement of the modified LTK scheme against cloud fraction error. The left lower quadrant identifies the number of scenes (one in this case) where a low cloud fraction error (< 5%) can be achieved by cancellation of pixel misidentifications (as measured by the % mask agreement—less than 80% is considered poor performance).