

Dear Referee:

Thank you for your comments concerning our manuscript entitled “Improved cloud mask algorithm for FY-3A/VIRR data over the northwest region of China”. Those comments are all valuable and very helpful for revising and improving our paper, as well as the important guiding significance to our research. We have made great efforts to make corresponding changes according to the comments. The main corrections in the paper and the responses to the referee’s comments are as following; the original comments are in italic and our answers are in normal font.

#### *General Comments*

*The authors present a method for cloud detection using FY-3A/VIRR observations over the varied surface types found in northwestern China. These include desert and mountainous regions with snow cover during the winter that are generally difficult scenes in which to perform cloud vs. clear sky discrimination. Bright and/or cold surfaces often mimic spectral cloud signatures. The algorithm described is a variation of the previously-published CLAUDIA (CLOUD and Aerosol Unbiased Decision Intellectual Algorithm) cloud detection scheme that is itself a variation on the MODIS cloud mask (MOD35). The authors develop three statistically-based, CLAUDIA-like reflective spectral cloud test thresholds for northwestern China that are augmented by a visibleinfrared bi-spectral test, NDSI, and NDVI indices to further sharpen cloud vs. clear boundaries. The cloud tests, NDVI, and NDSI thresholds development is sound, being based on previous work that is thoroughly documented in the text. The algorithm itself is reasonable and shows good potential for improved cloud detection in this area; however, the paper is weak on results and validation. The authors contrast their results to the operational VIRR and MOD35 cloud masks and ground-based human observed cloud amounts, but report only two VIRR/MOD35 comparisons (scenes). The human observer comparison is better but shows only ten reports for January, eight in February, and six in March, 2011, but the algorithm was developed using data from four months in 2010 and is intended to be used in all seasons. I would highly recommend adding more validation, at least including some data points from the warmer seasons.*

*The paper is generally well written and organized but there are many instances of incorrect tenses and awkward English construction that should be addressed (there are too many to list individually). The paper should have another good editing by a native English writer.*

Thank you for the suggestions about the results and validations. For the revised work, we have added overall 47 scenes collocated with the MODIS cloud mask products and 96 individual matchups between VIRR and the ground-based observations from two weather stations in 2011 for the improvements of the validation. The quantitative validation covered all months of 2011 including the warmer seasons that the referee recommended. The corresponding descriptions have been complemented in the revised paper (see Sect. 4).

The English in the paper has also been checked by a scientific and technical editing service website. (<http://www.textcheck.com/>)

#### *Specific Comments*

1. *Lines 159-160: The text mentions “diamonds” that are apparently seen on a computer screen. This is not germane to the topic and should be removed.*

The description of the data from MICAPS has been rewritten (see P.7 L.14-16). The word “diamonds” has been removed.

2. *Lines 178-181: Mention should be made here of Figure 4. Also, the word “settings” in line 178 implies values of thresholds when I think the authors really mean the forms of the thresholds as seen in Figure 4. I think the word “confidence” should be used instead of “overlapping” in line 180. A word like “delineation” should be used instead of “research” in line 181.*

Figure 4 has been renumbered as Figure 1. Also, the corresponding description of Figure 1 has been complemented in the revised paper (see P.8 L.5-6). The word “settings” has been replaced by “definitions” (see P.8 L.6). For the low and high limits, the former sentence may be confusing. It meant that these two parameters were determined by the minimum and maximum values of the overlapping intervals between the two categories according to the statistical data. The description of these two parameters has been rewritten in the revised paper (see P.8 L.7-10). The word “research” has been replaced by “delineation” (see P.8 L.12).

3. *Lines 184-185: What is “real total statistical pixels”? Do the authors mean “total numbers of pixels”? No mention of  $T$  should be made in this sentence since it does not appear in the equation in line 183.*

The former “real total statistical pixels” has been replaced by “total statistical numbers of pixels” (see P.8 L.15). The value of  $T$  was the threshold to determine which type the pixel belonged to. As the value of  $T$  changed, the loss function  $f$  varied. When the loss function  $f$  achieved a minimum such that the sum of the incorrect classified ratio reached the lowest value, the threshold value would be defined as the final  $T$ . The corresponding description of  $T$  has been complemented in the revised paper (see P.8 L.16-21).

4. *Line 199: “intensive” should be replaced by “narrow”; also, “clear” should be removed.*

The word “intensive” has been replaced by “narrow” (see P.9 L.6). The former “clear reflectance ranges” has been replaced by “the ranges of reflectance” (see P.9 L.6).

5. *Line 203: There is still some overlap between classes shown in Figure 1, so I think the word “mostly” should be inserted, as in “surface types can be mostly differentiated”.*

The word “mostly” has been inserted (see P.9 L.9).

6. *Line 206: The meaning of the lines in Figure 1b is unclear. I would say something like “Figure 1b also shows where the three specific channel 1 thresholds are located for clear vs. cloud/snow”.*

Per the referee’s comment, the description has been rewritten as “...are located for cloud types (clouds and snow) vs. clear types” (see P.9 L.13-14).

7. *Line 217: I would say something like “To separate snow and water into distinct categories. . .”*

The sentence has been changed as the referee’s suggestion (see P.9 L.24).

8. *Line 224: I would write “developed for the discrimination between snow and clouds/surface . . .”*

The sentence has been changed as the referee’s suggestion (see P.10 L.2-3).

9. *Line 261: “tend” should replace “prefer”.*

The word “prefer” has been replaced by “tend” (see P.11 L.14).

10. *Lines 293-295: This sentence confuses me. What selection is being referred to? Rewrite this sentence to make it clear or else remove it.*

The sentence has been removed.

11. *Line 310: Refer to these images as “gray scale” rather than simply “gray”.*

The word “gray” has been replaced by “gray-scale” as the referee’s suggestion.

12. *Line 319: What does “specific distributions of clouds” mean? Please rewrite to make clear.*

The former sentence has been changed as “... and cannot identify the clear-sky pixels between the clouds” (see P.14 L.12-13).

13. *Line 326: Please define H, M, and F in the text.*

For the revised work, we have added overall 47 scenes collocated with the MODIS cloud mask products, and three indices for the validation including the probability of detection (POD), the false-alarm ratio (FAR), and the hit rate (HR). The definitions of the validation scores were replaced by the Equations (10)-(14) in the revised paper. Also, the notations used for the definitions of the validation scores were shown in Table 5. The corresponding description has been added in the revised paper (see P.15 L.15-23 and P.16 L.1-3).

14. *Lines 327-331: I’m confused here. What does “Focused on the cloudy pixels” mean? In the results that follow, do you consider only cloudy pixels? But then how do you compute FAR? And what determines “truth”, MOD35 or something else? Please make this clear.*

To provide quantitative validation, the MODIS cloud mask product (MOD35), as a high-quality cloud mask product, was taken as “truth” for the evaluation of the proposed cloud mask algorithm (see P.15 L.2-5).

In order to make a clear description, the validation scores in this study were defined as following:

$$POD_{clear} = d/(c + d) ,$$

$$FAR_{clear} = b/(b + d) ,$$

$$POD_{cloudy} = a/(a + b) ,$$

$$FAR_{cloudy} = c/(a + c) ,$$

$$HR = (a + d)/(a + b + c + d) ,$$

where  $a$  represented the number of pixels identified as cloudy by both VIRR and MODIS.  $d$  represented the number of pixels identified as cloud-free by both VIRR and MODIS.  $b$  and  $c$  were the numbers of pixels showing different classifications between VIRR and MODIS.

The notations used in the equations were shown as Table 5 in the revised paper (shown as Table 1 here):

Table 1. The notations used for the definitions of POD, FAR, and HR scores.

Scenario	VIRR cloudy	VIRR clear
MODIS cloudy	$a$	$b$
MODIS clear	$c$	$d$

The probability of detection (POD) and the false-alarm ratio (FAR) were calculated for both cloudy and cloud-free conditions. The subscripts in the equations represented the cloudy and cloud-free conditions respectively.

The description of the validation scores has been complemented in the revised paper (see P.15 L.15-23, P.16 L.1-3, and Table 5).

15. *Lines 345-348: Same question as above but reversed. If considering only clear pixels, how is POD calculated? And what is “truth”?*

Refer to the reply to comment 14.

16. *Line 359: It is Table 5, not Table 4.*

Thank you for your attention. The Figures and Tables have been renumbered in the revision. We have revised the manuscript to minimize typographical, grammatical and bibliographical errors.

17. *Lines 370-374: It is stated that the cold temperatures and high reflectances of both cirrus clouds and surface make cirrus detection difficult but what about the 1.38  $\mu\text{m}$  channel? It would be nice to see a sentence describing why 1.38  $\mu\text{m}$  data cannot do the job here, since it is stated earlier that it is included in the algorithm specifically for this purpose.*

For the northwest region of China, the atmosphere is well known for its dryness especially over desert regions. The lack in water vapor in this area may cause difficulties for the identifications of

the thin cirrus clouds by 1.38 $\mu$ m observations (Ben-Dor, 1994). The corresponding description has been added in the revision (see P.18 L.18-20).

18. *Conclusion section: If you want to publish the paper more or less as is, I think some words should be included that list caveats to your results. There was no validation in the paper concerning results in different seasons (I'm including March as a winter month) and this is a major part of the threshold development. How does the algorithm handle transitions between seasons? This is part of the larger problem of having too little validation. I think there is good potential here but without more validation we can't know how well developed this algorithm really is. This should be stated as "future work" with regards to the algorithm.*

As the Referee's suggestion, more validations have been added to evaluate the performance of the proposed algorithm, which included 47 scenes collocated with the MODIS cloud mask products and 96 individual matchups between VIRR and the ground-based observations from two weather stations in 2011. It is true that the cases showing poor correlations with ground-based observations were mostly from the transitional months between seasons. The thresholds during transitional seasons without modifications could produce errors for the cloud detection results. The determination and validation of the thresholds during the transitional seasons are significantly potential areas to improve the cloud detection results. Considering the Referee's suggestion, we have rewritten the conclusion part with more information about the proposed algorithm and our "future work" (see Sect. 5).

19. *References: Section title should be "References".*

We are very sorry for our incorrect writing. The error has been corrected (see P.22 L.1).

20. *References should list all authors, rather than simply et al.*

The references have listed all authors (see References).

21. *Line 477: "Garder" should be "Gardner" (also line 44).*

The author's name is "Leonid C. Garder" (refer to doi: [http://dx.doi.org/10.1175/1520-0442\(1993\)006<2341:CDUSMO>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(1993)006<2341:CDUSMO>2.0.CO;2)).

22. *Figures: Figure 1: I would substitute the word "shown" for "labeled"; I would also write ". . .*

*the low and high confidence limits for clear skies vs. clouds and snow shown as black lines. . .”*

As the referee’s suggestion, the word “labeled” has been replaced by “shown”. The description “...for clear skies vs. clouds and snow...” has been complemented for Fig. 2 (former Fig. 1).

23. *Figure 3: Could the x’s in the color key at the top be made boxes or some other filled shape? It is very difficult to make out the colors.*

The color key in Fig. 4 (former Fig. 3) has been changed. I hope it is easier to make out the colors this way.

24. *Figure 5: For clarity, I would put the ranges of final confidences of clear skies on the figure near “Cloud Category” and “Clear Category”; also place “(clear)” under “Snow Pixels”.*

As the referee’s suggestion, we have added the corresponding details in Fig. 5.

25. *Figures 6, 8, 9: Use “gray scale” instead of “gray”.*

The word “gray” has been replaced by “gray-scale”.

#### Reference

Ben-Dor, E.: A precaution regarding cirrus cloud detection from airborne imaging spectrometer data using the 1.38  $\mu\text{m}$  water vapor band, *Remote Sensing of Environment*, 50, 346-350, 1994.