Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2016-169-RC2, 2016

Atmospheric Measurement Techniques Discussions

Interactive comment on "Cloud Detection in All-Sky Images via Multi-scale Neighborhood Features and Multiple Supervised Learning Techniques" *by* H.-Y. Cheng and C.-L. Lin

J. Kleiss (Referee)

jkleiss@lclark.edu

Received and published: 12 November 2016

Dear Drs. Cheng and Lin,

Thank you for allowing us to view your revised transcript, especially with the clearly marked red font indicating changes. This is a high quality paper, and I appreciate the chance to read and provide comments to this work prior to publication.

I have one substantive comment, two requests for further clarification, and a number of very small comments about the text itself. All page numbers and line numbers refer to the "Revised Manuscript" submitted on Oct 27th.

1) Substantive comment. On Page 13, (lines 280-283) you discuss the choice of Thr_{upper} and Thr_{lower} when performing the RBR classification. It seems that any pixels that have RBR > Thr_{upper} are automatically marked as a non-cloud pixel. If so, this is incorrect. The pixels with RBR > Thr_{upper} should be removed from the analysis. This explains the results you show in Figure 4. If you decrease Thr_{upper}, then FN (the number of cloudy pixels incorrectly classified as non-cloud) will increase. This explains the dramatic reduction in $Recall$ that is depicted in Figure 4 as the Thr_{upper} decreases from 1 to 0.9.
It also is apparent in Figure 9c, where the RBR appears to capture the clouds quite well, except for the very bright regions where it fails, due to Thr_{upper}.
Finally, upon reading more closely, it appears that you are treating the sun region and vertical lines in a similar fashion. On Line 127, you declare that pixels in the Hough line region or sun region "are determined as non-cloud pixels." Considering the sample image in Figure 2d, this also seems erroneous: these pixels should not be set to non-cloud. They should simply be removed from the analysis as contaminated pixels.
This substantive comment certainly needs to be addressed, and may affect the results of the figures, since it should substantively change the $Recall$ metric of the RBR procedure.

1b) Note that point (r) below is also a substantive comment. What is the minimum spatial extent of a cloud that your method can resolve (in pixels). How does this affect your analysis?

2) First clarification request. I am not able to understand some aspects of Figure 3, and the text that describes it in lines 227-239. I do not understand what is the 3x3x3 neighborhood around a pixel p at level i. At first I wondered if these were the three color channels, so that Figure 3a shows the red, green, and blue channels stacked on top of each other for a 3x3 pixel region. However, if this is the case then

the three levels in Figure 3b, showing a 5x5 region, 7x7 region, and 9x9 region don't make sense to me. Therefore, in line 233, I don't understand why a pixel would have 27 x 4 votes. If I try to understand this, it seems that at level i, with $L_i = 7$, we actually consider the 21 x 21 pixel region around our pixel of interest: a 3x3 grid, where each grid has 7x7 pixels in it. Then all four of your classifiers are run on each of those 7x7 regions, and is determined as cloud or non-cloud for each classifier. Then it would be possible to add votes. I think the wording that threw me off was "3x3 neighborhood" and "3x3x3 neighborhood," because these aren't really neighborhoods. I would suggest using "neighboring image patches" instead. Possibly adding light grey grid lines in the 3x3x3 squares in Figure 3a would help the reader understand that you are considering the image patch and all neighboring image patches.

3) Second clarification request. I also struggle to understand the rationale and math in lines 241-252, regarding $P(x) \in cloud \mid Num_{i=1 \sim l}(x_{level_i} \in cloud)$ If the Latex formatter is not working, this is meant to be the left hand side of Eq. (4). I think lines 241-252 can be made much more clear. If we are using 5 levels of classification, then presumably each pixel in the image will have 5 resulting decisions if it is a cloud or non-cloud. (If I understand Figure 3 correctly, then level i and level l don't make sense to me, because you can't have $i - 1$ and $i + 1$ in these two cases, respectively.) There must be something simple you do to obtain your final determination if this pixels is a cloud, given the results of the 5 levels of determination.

Finally, here are some minor comments on the text.
a) line 39: please refrain from the use of "etc." in manuscripts.

b) line 50: "cloud" should be "could."

c) When citing the authors of a paper in a sentence, simply use a format like "The work by Long et al. (2006) suggested that ..." This recurs throughout the

manuscript on lines 57, 66, 70, 73, and 201.

d) The review of literature of past attempts to detecting, classifying, and tracking clouds is quite nice, thank you. It seems that you should also consider the work by Zhuo et al. (2014), Isosalo et al. (2007), and Liu et al. (2015), all of whom considered multiple spatial interrogations or segmented spatial regions of images as well.

e) Figure 1 is a helpful schematic. It seems that the training and manual labelling of the 250 images should factor into your system framework somewhere. Could you possibly add this? I would imagine a separate box that points into the "Classification via Multiple Classifiers" dashed box.

f) On Line 149 you say that including multiple color spaces "provides the classifier more information that is beneficial to performing classification." Do you have any evidence to support this?

g) Line 154, change "the feature" to "each feature" for more clarity.

h) Line 171 change to $R_{Eigenvalue}$ for consistency with Eq. (1)

i) Line 179 change "And the tree is built recursively." To "The tree is then built recursively."

j) In Eq. (3), what is the p in the denominator of the equation?

k) Line 234, it's not clear to me why we are referring to the feature vector, x_{Level_i}. Haven't the classifiers already independently determined cloud or non-cloud status for the feature vectors? I may be missing something here. I believe elsewhere you refer to this as pixel p at level i, can you just refer to the number of votes for this pixel as

$V_{cloud}(p_i)$?

l) Line 241, what is x (the second character in the line, in the $P(x \in cloud \ldots$ equation)?

m) Line 264. Please indicate the geographical area where the images were collected.

n) Line 275. RGR should be RBR.

o) In your discussion of Figure 4, can you please give a reason for the trends that we see? For example, Why does Accuracy increase as Thr_{lower} increases? Why does Recall decrease as Thr_{lower} increases?

p) Line 299. You state that $Thr_{lower} = 0.8$ and $Thr_{upper} = 0.9$ have the lowest detection accuracy (note the misspelled "detecction" Please use spell-check.) I presume you are only looking at the blue bar line. RBR's detection accuracy is about 0.78, while the 3 other methods appear to be 0.8 to 0.83. Is this really significant? Is there anything worth mentioning about the Precision and Recall? If not, why are they in the figure? A simple sentence explaining why you choose Accuracy as your conclusive metric against the others would be appreciated.

q) Line 306. Note that the notation is not consistent with Line 152. I'm OK with this, although it's a bit sloppy.

r) Why don't you use $L_1 = 1$ as the smallest level? It appears that you would have a small-scale limit on the size of clouds that you could resolve. What is that limit? Could you comment? You compare these results to other single-pixel results (e.g. Fig. 10). It seems to me that we can't quite compare them, since this method always

operates on at least a 5x5 grid region. (And, in fact, the 8 neighboring 5x5 regions as well? So the small scale limitation could be quite substantial!)

s) Line 308. Change "using feature" to "using the feature"

t) Line 309. This sentence didn't make sense to me. I suggest, "The value of Thr_{PCA} is typically between 90u) Line 314. Why do you say "cross-validated"? What does that mean?

v) In the caption to Fig. 6, refer the reader to Eq. 5.

w) Line 374. You say "fixed and dynamic RBR thresholding." I don't see that you did anything other than fixed thresholding, using 2 fixed threshold values of 0.8 and 0.9.

Many thanks for this fine manuscript, and your hard work. It is a contribution to our field.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2016-169, 2016.