Atmos. Meas. Tech. Discuss., 7, C4607–C4610, 2015
www.atmos-meas-tech-discuss.net/7/C4607/2015/


Atmospheric Measurement Techniques
Discussions

# *Interactive comment on* "Block based cloud classification with statistical features and distribution of local texture features" *by* H.-Y. Cheng and C.-C. Yu

**Anonymous Referee #2**

Received and published: 19 January 2015

The authors compare a bouquet of methodologies for cloud type classification. They also propose a new method and this is the basic novelty of their study. However, from the validation and the physics point of view, there are significant drawbacks.

Page 11782, line 15: in the process of block division, the cloud coverage percentage is lost. It is a significant drawback of the method that the pixels at each block are not classified as cloudy or clear. This means that the calculated features for each block are dependent from cloud coverage, i.e. the percentage of pixels that are cloudy compared with the total number of pixels. As a result, the detection of a specific cloud type e.g. cumulus could be wrong when the cloud covers just a fraction of the block.



The authors propose a threshold value to change the classification from stratus to cumulus, based on experiments. However, there is no information about the experiment types and the corresponding analysis, as well as there is no information about other misclassifications that may occur.

Page 11784, line 3: the clouds 3 and 4 are not dominant at all. From the number of the classified blocks, 7, 6 and 7 blocks correspond to cloud classes 3, 4 and 1 respectively. In addition, it is evident from the relevant figure (Fig. 6) that all blocks that were classified as 1 (cirrus), they are actually small cumulus or cumulus edges. The same classification error would be also possible for some non-classified blocks. The authors state that the not full blocks are not classified by their method. However, it is evident that this is not the case of this figure. There are not fully blocks that are classified as well. Overall, it seems to be possible that more blocks will be wrongly classified with this method. It is also highly questionable if the cirrus cases will be indentified and be classified as cirrus or as clear sky cases.

Figure 7 and relevant text: the samples provided for the training blocks with clear sky conditions seem not be the proper ones. In this figure, two examples are malformed by the sunlight and should be treated as clouds by an objective training method.

Figure 8 and relevant text: two examples of images are presented with 2 ground truth labels. The left one is classified as cirrus-and-cumulus and the right one as cirrus-and –cumulus. However, the first one is definitely cirrostratus-and-cumulus.

An all-sky camera was used with a lens that covers a field of view of 185o. However, it seems that in all figures a part of the sky is missing. There is no view of the surroundings. Even with a perfect horizon, the surroundings should be depicted with the specific field of view. In addition, the low clouds seem to be cut by the camera. The only possible explanation for theses effects is that the camera is located at a very high place. If this is not the case, it should be explained how these images were taken and if the authors masked the images and used just a part of the sky.

Figure 10 and relevant text: The validation procedure is highly questionable. The authors provide the average accuracy of the methods and the scores per cloud type. In addition, there is no information about the selection of the training and the testing datasets. Did they use enough images (blocks) for each cloud type? Did they use a diverse set of images including cases with several solar zenith angles? Did they use enough cases with the Sun blocked/semi-blocked by the clouds or be fully visible? What about cases that are difficult to classify (e.g. cumulus low in the horizon that seem to be like cirrus, thin cirrus)?. The authors compare their methods with one proposed by Heinle et all (2010). However, it has been revealed that this original method can provide better results if some new features are added/replaced (e.g the visible faction of the Sun, not blocked by clouds, the existence of raindrops) and if the images of the training and testing datasets are divided in subclasses based on the solar zenith angle (Kazantzidis et al., Atmospheric Research, 2012). Could the proposed method outperform the latter one?

Page 11774, line 25: The Principal Component Analysis (PCA) method was used to reduce the dimensionality of the extracted feature vectors. PCA is being used in several studies dealing with dimensional reduction of vectors. But, is this method reasonable to be used from a physics point of view? According to the standard PCA method, by subtracting the mean of each column from all the observations per column, the origin of the data in the initial space is translated to the cartesian origin. Therefore, the original unsigned data are probably assigned to negative values, which is contradictious to a physical interpretation of the data in the current application. What would be the results when using the uncentered PCA method?

Page 11776, line 20: scattered cumulus are not middle latitude clouds.

Page 11776, lines 5-15: it is correctly stated that under conditions with mixed cloud types, it is not appropriate to use the features of the entire image and classify the whole image as a certain cloud type. To solve this problem, the authors proposed the division and classification of the image in block. It is also stated that this method could



be very helpful, since the clouds in regions closer to the sun have higher impact on the irradiance changes. However, such cases are not presented in detail or analyzed in depth. From the sample images, it is not clear if the clouds close to the sun cannot be detected correctly when the sun is fully or partly visible.

Interactive comment on Atmos. Meas. Tech. Discuss., 7, 11771, 2014.