Interactive comment on "Cloud Detection over Snow and Ice with Oxygen A- and B-band Observations from the Earth Polychromatic Imaging Camera (EPIC)" *by* Yaping Zhou et al.

Anonymous Referee #2

Received and published: 22 November 2019

Zhou et al. described a cloud detection algorithm over snow and ice with oxygen A and B band. They have demonstrated that the new cloud mask algorithm is an improvement compared to the current EPIC cloud mask algorithm. They derived an analytic relationship between the double logarithm of the O2 band ratios and the surface elevation and the zenith angles. They also showed the limit of the algorithm for optically thin clouds and low elevations. I think the paper is fit the topic of AMT.

Thank you for reviewing the paper and providing thoughtful comments.

Specific comments

Page 8 lines 9-16. The authors described briefly the radiative transfer simulator for EPIC. Does the simulator have sphericity correction at the solar and viewing zenith angles larger than 80 degree?

The sphericity is not considered in the model. This is not a problem for EPIC, as the standard Level 2 cloud products are only generated for view zenith angle < 76^0

Page 8 lines 25-28.

Why the surface height in the simulations is from 0 to 15 km with 2.5 km increment? The surface height larger 9 km is not useful and the increment of 2.5 km is too large. The increment of 0.5 km would be a better option.

We removed simulations beyond surface height above 7.5km. It would be better to use increment of 0.5 km in height, but since the function with height is close to linear, we didn't redo the clear sky simulation.

Page 9, lines 26 – 30 It is not clear how the coefficients were derived. Could you explain it in detail?

We used a multivariate linear regression to do the fitting. The regression takes surface elevation (Z) and ln (m) as two independent variables and db ln (R_{abs}/R_{ref}) as dependent variable. The derived coefficients are used to prediction expected db ln (R_{abs}/R_{ref}) and then R_{abs}/R_{ref} . More details are added in the text.

Page 10 lines 24-26 How did you select the snow/ice surfaces?

Initially we used surface albedo of 0.8 to represent snow and ice surfaces in the model. Additional simulations are performed for surface albedo at 0.6 and 1.0 to cover the range of albedos over snow, sea ice surfaces. In the observational data, we used the surface type information included in the Langley GEO/LEO composite data, which is based on the IGBP surface type dataset and the Near-real-time Ice and Snow Extent (NISE) data set from the National Snow & Ice Data Center (NSIDC). We have added detailed references in the text.

Page 12 lines 4-9 Please explain more details about the regression. How did you design the model to predict the median . . .?

The same multivariate linear regression is applied to the observational data. The nature of regression is to provide a function that minimizes the total squared error which will approximately pass cross the median of each sample bins. But our cloud mask threshold is to find the upper bound of clear sky value so that all clear sky pixels will be under that curve idealistically. In reality, because there are many overlaps between clear and cloudy pixels as shown in Figure 5c and 5d, we could only move the divider up slightly to balance the clear and cloudy detection.

Page 12 lines 20 -21 What 'non-negligible uncertainties' do you mean here? Fig. 1 I,j

The reference cloud mask we used is based on GEO/LEO retrievals, which has its own uncertainties. Cloud contamination is one of the main causes of scatter in the clear sky regression. Other causes may include uncertainties in geolocation, surface elevation, atmospheric profile etc..

The 'Fitted threshold' is not easy to understand. Do you mean the fitted A-band and B-band ratio? Do you use the simulated A-band ratio, m, z, to derive the coefficients in Table1, then calculate the 'Fitted threshold' using these coefficients? Fig.1i,j shows that the fit is almost linear. Will it cause scatter if the coefficients are applied to other data not in the simulations? If the surface albedo is 0.6 or 0.9, could you get the same coefficients?

You are right. The fitted threshold refers to A-band and B-band ratios computed with regression coefficients. Ideally, everything being equal, the ratios for cloud sky should be larger than that of a clear sky. As mentioned earlier, we are trying to find the upper bound of the clear sky ratios. The regression is derived with simulations using surface albedo of 0.8. To test if these coefficients work for other surface albedos, we conducted new clear sky sensitivities with surface albedo of 0.6 and 1.0 and results are shown in Fig. 2. In majority of the cases, the clear sky A-band and B-band ratios are not sensitive to surface albedo, the fitting is problematic at large zenith angles (>76°) that EPIC does not retrieve.

Fig.2 Since the algorithm also detects clouds over snow/ice on top of mountains, could you make a similar plot for surface height of 2.5 km or 5 km?

We added cloudy sky sensitivities for surface elevation of 2.5 km (Figure 4). Compared to surface at sea level, the cloud detection algorithm is less sensitive over high mountains; more thin and low clouds will be undetected. Discussions are added to the text.

Fig. 3 How do you explain the scatter in the clear-sky plots?

The reference cloud mask used here is based on multi-sensor, including those from both geosynchronous orbit (GEO) or low Earth orbits (LEO). In general, the sensors used have better cloud detection capabilities than EPIC, but misclassifications still exist. The scatter we see can come from multiple sources, include cloud contamination, surface elevation uncertainty, cross-sensor consistency, geolocation error, atmospheric profile uncertainties, etc.

Fig. 4 It seems that you have to use more digits in the colorbar for (a,b). For (d) please use integer in the colorbar.

Done.