Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2019-345-RC1, 2019

Atmospheric Measurement Techniques Discussions

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 #1

Received and published: 29 October 2019

This study presents an updated DSCOVR/EPIC cloud detection over snow and ice surfaces. It improves the current scheme by better accounting for changes in surface altitude and the solar or viewing zenith angles.

The topic is appropriate, the method physically sound, the general structure sensible, and the improvements in EPIC's cloud flagging look good. However, some of the radiative transfer choices seem unphysical, details needed for replication are missing, and discussion of several important issues is absent. Unfortunately, the revisions I propose may mean re-running the radiative transfer and recalculating the thresholds so I request major revisions.

I expect that the main conclusions of the paper to be solid and that the authors should have little trouble in dealing with my comments. With revisions I would judge the science and presentation to be of higher quality and would support its publication.

1. Specific comments:

1.1 General flow and clarity The order is sensible but important details sometimes appear late in the paper in a way that confused me. For example, I don't see an explicit statement that the training & validation is versus the GEO/LEO dataset until P11. This should be in the introduction and mentioned when talking about performance (e.g. P4L18). It's also not immediately clear what is new. So the old algorithm doesn't account for surface height (P16L11, 16 pages in!), but what else exactly? Please explain in the introduction, and see the line-by-line technical comments.

1.2 Incomplete information regarding methodology How did you get the regression statistics (e.g. P9L26–30)? I first assumed simultaneous multi-variate least squares, but P12L7 makes me think not. In Table 1 do you have error bars? I also think your equation is complex (see section below), how do you handle the imaginary part?

I don't see your snow & ice surface definition until P11L23–25 which I think says you're using GEO/LEO data, and only "permanent" snow, i.e. not seasonal? But then why are there so many samples over N America and Eurasia in Figure 7(a) but not in Figure 8(a)? If you develop using permanent snow then this needs to be said in the introduction and potential issues with e.g. snow-covered forests with lower albedo need to be discussed.

On P12L16–24 "when applied to a different dataset". What is this different dataset? Did you subsample the full dataset? Does this different set have the same distribution of SZA, time etc?

1.3 Mathematical issues P6L25 Eq. (7) and (8). Could you expand on the switch to c_0? I work it out as complex:

$\ln(R_{abs}/R_{ref}) = mce^{-z/H}$
$-\ln(R_{abs}/R_{ref}) = -mce^{-z/H}$
$\ln(-ln(R_{abs}/R_{ref})) = \ln(-mce^{-z/H}) = \ln(-1) + \ln(mce^{-z/H})$
$\ln(-\ln R_{abs}/R_{ref})) = i\pi + \ln(c) + \ln(m) - z/H$
$c_0 = i\pi + \ln(K_a w_1 \rho_0 H)$
Please explain my error or comment on how this affects your regression.

Minor points: dln(x) is widely used in calculus, how about something without another standard meaning, like dbln(x) or ln"(x). I also think you lost an (R_{abs}/R_{ref}) in Eq.(9) on P7L5.

1.4 Radiative transfer (RT) description & choices

I think that the general approach is sensible but some details aren't clear and several RT inputs are physically unrealistic. These are my biggest technical issue with the paper and are the primary reason I propose major revisions.

You simulate liquid clouds above 2.5 km, consistently 1 km thick, and over frozen surfaces up to 15 km in altitude with albedo of 0.8.

Firstly, CALIOP sees liquid in Arctic clouds <2.5 km (Cesana et al. 2012, doi: 10.1029/2012GL053385) and the ARM site in Alaska also sees lots of these lower clouds (e.g. Zhao & Wang 2010, doi: 10.1029/2010JD014285). Your higher clouds should typically be ice, which may affect both $R_{ref}(\tau)$ and the in-cloud path lengths.

I would like to see sub-2.5 km clouds included in your RT. You might need to exclude them from your threshold calculation to prevent too many false positives, but these low clouds are particularly difficult for LEO/GEO-based infrared detection. The implications of this for your testing & validation should be discussed and even if you find you can't reliably test for these low clouds, then you should be explicit about this limitation.

The fixed geometric thickness also might affect your thresholds somewhat. A 1 km thick liquid cloud with $\tau = 3$ is very low N_d and should have unrealistically large within-cloud

path lengths. This might contribute to the discussion on P10L15–19. I'd propose a thickness that varies realistically with τ based on number concentration or a published relationship (e.g. for liquid clouds Eq. 2 from Chiu et al. 2014, doi: 10.5194/acp-14-8389-2014).

Your 0.8 albedo for both bands needs support. At the very least, you need to consider what this means for e.g. snow covered forests where the albedo is substantially lower, and may vary between the bands. Perhaps some simple physical argument with discussion of the limitations might be enough.

Finally, I don't know surfaces on Earth >10 km altitude. Why include your 15 km surfaces? Your observation sample should lack such cases, does this affect the regression statistics for the RT sample in Table 1?

I'd also appreciate some other details. The paragraph P8L9–16 is a good place to explicitly state that within-cloud absorption by O_2 is included in your RT (it is, right?). I also assume the EPIC ILS are broad enough that line broadening barely matters but would like a comment on this plus a reference to your spectroscopic database.

Summary: I believe your RT should include clouds <2.5 km (which may be liquid, supported by appropriate references), but higher clouds should contain ice and geometric thicknesses should vary realistically in tau. Then just recalculate the thresholds and statistics. I suspect your results will be robust to these choices, but some details could differ and your results would be more physically defensible. I think you should also at least discuss the very high surface altitudes and whether your regressions are affected.

1.5 Discussion of LEO/GEO limitations

It's fair enough that you test versus GEO/LEO, but you should explain how their limitations are relevant to your analysis. Examples of the sorts of references that should be included in the discussion are Wang et al. (2016, doi: 10.1002/2016JD025239) for MODIS collection 6, Karlsson & Håkansson (2018, doi: 10.5194/amt-11-633-2018) for

AVHRR and Shang et al. (2018, doi: 10.1038/s41598-018-19431-w) for Himawari-8.

Technical comments

2.1 General: Please check for missing articles or pluralisation where you currently treat normal nouns as proper nouns. Example insertions in square brackets:

P4L2 “the height of [the] effective reflective layer”
P4L4: “for use over [the] land surface” or “for use over land surface[s]”
P4L12: “but [a] large discrepancy is found”
P4L17 “over snow/ice surface” → “over snow- or ice-covered surfaces” (the hyphens here are an opitional style choice)

There are others. Plus the Andes, which takes the definite article despite its capitalisation:

P15L30: “. . .and the southern tip of [the] Andes”
P16L12–13: “. . .southern tip of [the] Andes could. . .”

2.2 Line-by-line

P1L24–26: When talking about performance statistics, please mention against what you are comparing. Some form of: “against a product based on multiple other passive sensors” or similar.

P3L1–7 para: “reference channel”, change to “weakly absorbing reference channel” to help those unfamiliar with the approach.

P1L27: “Less significant” – I don’t see significance tests, perhaps “less substantial”?

P2L7–8: not sure what long haul means.

P2L12: “narrow” is not an absolute. On P3L3 you mention the channel centres, could you add typical FWHM or another statistic that describes the spectral width there?

P2L19: “winter”, “summer” please specify (I assume) “boreal”

P4L10: “are performing reasonably well” – this value judgment depends on assumptions about the performance of other cloud flags. I would prefer “show good agreement” with some performance statistic(s) in brackets.

P4L11–12: comment that “accuracy rate” and “correct cloud detection rate” will be defined later, or describe here.

P5L2: “based on well-known and well-mixed atmospheric O_2 gaseous absorption” – this looks to me like the adjectives both refer to “absorption”, but it isn’t exactly “well-mixed absorption” you mean. How about something like “well-known gaseous absorption of well-mixed atmospheric O_2”.

P5L3: “..gaseous absorption, therefore, changes in observed radiance in the expected O_2 band” – I find “therefore” a weird link here, I don’t think it’s the O_2 band that’s “expected”. How about “. . .gaseous absorption. Changes in observed radiance in the O_2 band are expected to contain. . .”

P8L14 “convoluted with” → “convolved with” (I believe this is the verb for mathematical convolution, please check).

P8L21–22: a reference or pointer to the atmosphere definitions would be handy.

P9L1: “duplicate the quantitative relationship”. . . maybe “simplified relationship”? The RT model uses quantitative relationships too.

P14L23–28: this is a stylistic preference, but why pick a, b, c, d? In my opinion the standard notation (TP/TN/FP/FN for True/False Positive/Negative) is more easily understood and would make help me to interpret Equations (11)–(13) on sight.

P16L2: “Comparison show that. . .” → “Comparison shows that” (typo missing “s”)

P16L7–8 – “indicates high cloud fraction (>80 %) over. . .”, I’d just say “indicates cloud fraction > 80 % over. . .” because “high cloud fraction (>80 %)” could also be >80 % coverage of high-altitude clouds. This also appears on P14L17, where “high cloud fraction”

is >95 %. I'd be tempted to change "low cloud fraction (<5 %) and high cloud fractions (>95 %) categories" to "cloud fraction < 5 % and cloud fraction > 95 % categories".

P16L17: "achieved high accuracy. . ." → "has improved accuracy" ("high" again seems too subjective to me).

P17L1–3 this explanation seems physical but I don't think it's accessible to a non-specialist. How about: "This method is based on the fact that photons reflected by clouds above the surface will travel, on average, a shorter distance through the atmosphere and so experience less absorption by O_2" or similar?

P17L16: "these performance matrices". I would prefer "metrics" because you haven't explicitly introduced results as a matrix previously, and also these values are not the matrix itself, but derived from it (e.g. accuracy score = trace of normalised confusion matrix).

P17L11: "Model derived algorithm is chosen because of its stable performance". Do you mean that you chose the model algorithm because it performs better for the sample that was not used in training the obs-based dataset? If so, please change sentence to say this and, as requested earlier, describe how the datasets differ.

P26L9: ". . .on the right side of black lines will be identified as clear sky. . .": this implies that you use the black lines as a threshold, but I think you prefer the red dashed lines. Please rephrase to be clear that the black lines are a possible selection but you don't use them (if this is true).

P27L5: This is very nitpicky, but the (d) colour bar makes it look like you have continuous cloud mask values. I'd personally change the colour bar tick mark locations to be the actual flag values (1, 2, 3, 4 instead of 1.0, 1.6,. . .)

P28L1 : Figure 5, could you add a legend or some text indicator on one of the panels for the colours? This isn't vital given it's in the caption, but it would be nicer.

P32L5: Figure 9 caption: "matrix" → "metrics" as above.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2019-345, 2019.