Atmospheric
Measurement
Techniques

Open Access

EGU

Discussions

Interactive comment on "Cloud identification and classification from high spectral resolution data in the far and mid infrared" *by* Tiziano Maestri et al.

Anonymous Referee #2

Received and published: 30 April 2019

General Comments

The authors present the new cloud identification and classification algorithm CIC for far- and mid-infrared radiance measurements. The method is specifically designed for analysis of the ESA Earth Explorer candidate mission FORUM. Overall, I found this to be a scientifically sound study, which should be of interest for readers of AMT. The manuscript could be a bit more concise and some English language editing may be needed, but it is mostly clear. I would recommend the paper for publication in AMT.

Specific Comments

page 1, line 2-3: Saying the method uses 'a univariate distribution and a threshold' without additional information may not be clear to the reader. A distribution of which

variable or a threshold of which quantity?

page 1, line 1-10: I would suggest adding 1-2 sentences and numbers providing information on the accuracy and performance of the new CIC algorithm to the abstract.

page 3, line 3: An univariate distribution of which quantity?

page 3, line 29-32: It may not be too important to mention here that ERA-Interim data can be retrieved via a web interface. Saying 'surface height' is computed from temperature, pressure, and geopotential heights is a bit confusing, as surface height (or surface geopotential) is already a parameter in the ERA-Interim database. Do you mean 'geometric height' or 'height above the surface'?

page 4, Fig. 1: The RTX acronym/code was not introduced at this point. Also, the reader will not know what 'OD_deflt' means.

page 5, line 14-19: Is noise to be expected the leading error of the FORUM measurements? Is there a reference (e.g. an ESA report) available, providing more detailed information on the FORUM instrument? Instead of 'ideal measurement noise' perhaps say 'nominal measurement noise'?

page 6, Table 2: Can you provide a rough estimate on the percentage error of the noise estimates? The percentage errors might be rather large for the FIR part of the spectrum because measured radiances are quite low?

page 6, Table 3: 'Polar regions' covers both, polar winter and polar summer. Perhaps this should be split into the corresponding cases?

page 11, line 17-19: This is providing a specific definition of the similarity index used in this study. Did you consider any other potential measures of similarity? Was there a specific rationale to chose this definition?

page 13, Fig. 7: Make plots the same size.

page 15, line 9: Perhaps say that 'cardinality' means 'number of elements of the set',

which is simpler for non-mathematicians.

page 16, line 24 to p17, line 12: If the POSCO is not needed or evaluated during the rest of the manuscript, it may not have to be introduced here at all?

page 17, line 23: If I understood correctly, only a subset of all FORUM channels is considered. Is there any exploitable information content in the unused parts of the FORUM spectrum?

page 18, line 1: At this point, it may not be clear why a correlation between the CoI and DP can be found? From Fig. 9 this correlation does not become very evident, as these plots basically seem to show point clouds with some outliers? What is the correlation coefficient between CoI and DP?

page 19, line 9-14: What is the rationale for creating different TraNCs? Is this meant to reflect a priori knowledge on real cloud distributions in the atmosphere?

page 19, line 1: Was the tropical test case selected for presentation because it is the most difficult or most simple case?

page 19, line 1-2: I was wondering if it is sufficient to use only one test set (per class) for validation of the classification methods. The classification method might be tuned to work best only for the specific test set and may show different results for another test set. Did you consider to rerun the analysis with a different choice of test sets?

page 20, Fig. 10: Do the results presented here change if another test set is used?

page 21, Table 4: Perhaps choose more even bins for OD, i.e., 0.1 to 0.3 and 0.3 to 1 rather than 0.1 to 0.5 and 0.5 to 1.0?

page 21, line 13: Which CO2 band is meant?

page 22, line 10-12: Can you explain the initial decrease in performance when the first FIR channels are added to the classification?

page 23, Fig. 11: Perhaps it could help to add a few more data points to this figure, to better understand what is happening when the first FIR channels are added to the classification?

page 23, line 12-13: I am afraid I do not understand the sentence 'Note that the DP value is the minimum...' in this context. Can you explain it a bit better?

page 24, Fig. 12: I have some difficulty identifying any clear correlation between OD and CSID from these figures.

page 27, line 1-4: This sounds as if almost every linear algebra problem can be solved with an $O(n^2 \log n)$ algorithm, which is too general. I would rephrase this a bit and just refer to the algorithms used in this study.

page 29, line 21-22: Can you quantify this? How much higher are the scores?

page 30, line 2-3: Can you quantify this? How much better was the detection of thin cirrus?

All Figures: Please check and enlarge the font size of the labels to make them better readable.

Technical Corrections

page 1, line 8: change 'i.e' to 'i.e.'

page 2, line 10: introduce REFIR-PAD acronym

page 2, line 19-20: check that acronyms are properly introduced

page 2, line 31: change 'mostly widely' to 'most widely'

page 2, line 32: change 'Feature' to 'feature'

page 2, line 33: change 'Spectral Fitting' to 'spectral fitting'

page 3, line 12: change 'profiles' to 'conditions'

page 3, line 14: rephrase to 'the CIC algorithm'

page 3, line 32: change 'Era-Interim' to 'ERA-Interim'

page 4, line 8: change 'fulfil this information' to 'add information'

page 4, line 11: change 'of the spectrum representatives of ' to 'representative of'

page 4, line 15: change 'Scattnlay' to 'ScattNLay'

page 4, Table 1: change 'Cloud properties' to 'Cloud property'

page 6, line 12: change 'a presence' to 'the presence'

page 7, line 3: change '(CIC' to '(CIC)' (or delete)

page 11, line 4: change 'line' to either 'row' or 'column' (as applicable)

page 11, line 19 and page 12, line 15: remove extra brackets () for ETREM term

page 16, line 2: perhaps replace 'an algorithm' by 'a cloud classification algorithm' to be more specific?

page 16, line 19: change 'i,:' to 'i:'

page 17, line 15: change 'are evaluated' to 'is evaluated'

page 17, line 24: change 'but a small' to 'except for a small'

page 19, line 19: fix 'The the'

page 23, line 1: rephrase to 'a function'

page 23, line 8: rephrase to 'The left plot'

page 25, line 11: DP values are ... than 0.7 *and* for CoI

page 25, line 27: rephrase to 'When the elementary'

page 25, line 33: change 'intel' to 'Intel'

page 27, line 2: delete 'it is found'

page 27, line 12: change 'a multiple number' to 'different numbers'

page 28, Fig. 15: change 'not linear' to 'non-linear'

page 28, line 1: change 'every approximately' to 'about every'

page 28, line 5: rephrase to 'cloud spectra detection and classification'

page 29, line 6: change 'defines' to 'evaluates'

page 29, line 7: change 'provide' to 'provides'

page 29, line 11: change 'do' to 'does'

page 29, line 19: change 'somehow interpretable as' to 'related to'

page 29, line 23: rephrase 'computed to simulate'

page 29, line 33: change 'point out' to 'assess'

page 30, line 1: rephrase to '238-545 cmˆ-1 wavenumber range is improving the'

page 30, line 5-6: change 'noted as' to 'noted that'