

Review of the paper “Bayesian Cloud Top Phase Determination for Meteosat Second Generation” by Mayer et al.

The manuscript is accurate, well-structured and well written. I recommend for it to be accepted after the authors have addressed the few minor points I list below.

General comments

One thing that is mentioned in the abstract and in the introduction is the low/inadequate accuracy of traditional cloud typing methods in detecting cloud top phase. Perhaps a comparison with at least one another product available online for cloud type retrieval (could be from geostationary or low earth orbit instruments) could be interesting. It could be briefly shown as direct comparison or discussed. It could be for example the NWC-SAF Cloud Type product (even though it does not explicitly report the cloud phase, though the cloud types can be broadly related to cloud phase).

In the discussion in chapter 8.1 it is mentioned that a difficult case is represented, perhaps unsurprisingly, by overlapping cloud layers with a high thin ice cloud over a low liquid cloud. Does the measure of certainty provide an extra information that can be used to isolate these cases? More in general, it would be interesting also to show an example of how the certainty measure looks like for a typical example of the product as in Fig. 7. This is very useful especially when the first phase choice is only marginally more certain than the second choice (as it is mentioned in chapter 8.2 in relation to the POD of MP and SC types). This discussion could be added to chapter 8.3 which currently is fairly limited in content.

A clarification regarding the certainty measure: how should it be interpreted when the probabilities of the two most probable states are both about 0.5 (one slightly higher than the other), while for all the other states probabilities are ~ 0 ? In this case has the certainty measure the same meaning as for a case where for example $P(q^* | M, A) = 0.5$ and the remaining 5 states have all $P = 0.1$?

Minor and technical comments

Abstract, line3: “mainly distinguished between” -> mainly identified/detected

Section 3.1, line 124: the sentence “These probabilities...” is repeated twice.

Section 4.1: not clear where the information content for the prior is shown in Fig 3. I imagine is the first panel, but from the figure caption is not clear.

Section 4.3, Line 229: this applies also to fractional cloud cover?

Section 4.3: given the strong dependency on the surface emissivity at wavelength around $8.7 \mu\text{m}$, should an emissivity map also be taken into account, or the surface type is enough to account for the effect?

Section 4.5: The surface type can be used as a proxy for surface albedo, but does this also include the spectral variation of the albedo? Are significant changes in surface albedo between $0.6 \mu\text{m}$ and $1.6 \mu\text{m}$ (e.g. over snowy surfaces) important in this context or the surface type is enough?

Section 4.6: why the mutual information between $RR_{1.6/0.6}$ and q seems to benefit by the inclusion of the surface type or the $R_{1.6}$ but not from the use of both together?

Section 8.1, line 388: “by nature/DARDAR to it” not very clear, please rephrase. Also, missing bracket in “(see Fig 7”.

Section 8.1, line 418: “spatially shifted against the DARDAR clouds” please clarify.

Section 8.2, line 438: “less sensitive to optically thin clouds than Lidar”

Section 8.2, line 480: As discussed in the text, is it clear that SC and MP types are often difficult to distinguish, and that the certainty computed by ProPS is often marginally higher for either of the two types. Also, what is the confidence in the DARDAR supercooled water classification?

Section 9: It would be interesting to briefly discuss how a new sensor (e.g. MTG FCI) would impact such an algorithm, both in terms of the information contained in the new channels and in the surface resolution.

Figure 3: the caption could be clearer. Also, at first I did not understand that the first column of each panel (apart from the third) represents the starting point of each information content before the introduction of each new condition.

Figure 6: perhaps adding another RGB composite helps a better comparison with the categorization as many of the high thin cirrus are lost in the RGB shown in the current figure.