**Manuscript number:** EGUSPHERE-2023-2885
**Title:** Algorithm for continual monitoring of fog based on geostationary satellite imagery
**Submitted to:** Atmospheric Measurement Techniques (AMT)

Dear editor and anonymous reviewers,

Once again, on behalf of all authors, we would like to kindly thank you for your very constructive comments and suggestions, and also for spending your precious time on reviewing our manuscript.

Please find below our answers to your inquiries.

Note that

a) list of references mentioned in this document is given on the last page.
b) due to an update in the reference manager used, all in-text and the bibliography fields included in the manuscript had to be updated. These changes are marked in the manuscript with track changes.

Sincerely yours,
Babak Jahani, Jan Cermak

## Referee #1

### Comments and suggestions

***C#1.1***

I appreciate the very detailed responses by the authors.

The revised manuscript is much better written and organized.

I agree that the POD and the FAR are sensitive to the fraction of fog/low stratus cases in the dataset, which is only 3% in this study. So a 30% of FAR means 1% of misidentified fog/low stratus cases in the whole dataset. Is 1% a big number? In terms of the absolute number, 1% relative to the whole dataset is small. But in terms of the predictive power, for every 4 reported fog cases, 1 of them is a false alarm, which is not small relative to reported fog/low stratus cases. In addition, a 80-90% POD means that 0.3-0.6% of the whole data which are fog/low stratus cases cannot be detected. In the absolute sense, 0.3-0.6% (relative to the whole dataset) is small, but10-20% (relative to the fog/low stratus cases) is not small. In this sense, the authors' response by introducing ACC and PFD does not address my comment about the high FAR directly.

In any case, it was not my purpose to debate how the verification statistics should be interpreted. Actually, with the current results, I think it is fine for the authors report their detection algorithm, but they should say a few words on how the accuracy of their detection method can be further improved (e.g. how can the FAR be reduced?) in the future.

The authors have added some discussions (Lines 190-200 and Lines 455-477) about the limitations of ground station data, such as the potential errors due to the spatiotemporal matching with the satellite pixels and the big difference between the spatial resolutions of 3x3 km2 satellite pixel versus a point measurement at the ground, which may have contributed to a part of the FAR. But again, I don't think these discussions would directly address the issue of the high FAR. Such limitations of ground station data are not unique to fog/low stratus detection; the same limitations are applicable to any comparison between satellite measurements and ground measurements (e.g. of greenhouse gases, surface temperature, etc). It is more important that the authors discuss about the potential deficiencies of the current detection algorithm besides the potential problems of the validation procedure.

***A#1.1***

Thank you for the clarification and the opportunity to account for this. In line with your suggestion, we have included some discussion on this topic in Sections 4 and 5. In summary, we have analyzed False Positive (FP; see Appendix B) and False Negative (FN; see Appendix B) predictions of the ML algorithm across the five datasets to gain insights on the deficits of the algorithm, and proposed ideas on how these deficit can be accounted for in the future.

In particular we have added the following in Section 4 (Results and discussion) at lines 395-415 of the revised manuscript:

“

Figure 3 provides detailed insights into the false-negative and false-positive (see Appendix B) predictions of the FLS ML technique across the five datasets. From Figure 3(a) it can be seen that the majority (~80-90%) of pixels falsely classified as FLS by the ML algorithm (i.e., false alarms) were actually clear-sky. Similarly, Figure 3(b) shows that ~83-92% of false-negative pixels (i.e., undetected FLS pixels) were predicted as clear-sky by the ML algorithm. These results suggest that the algorithm has room for improvement in distinguishing between clear-sky and FLS. Indeed, finding it challenging for the algorithm to separate clear-sky from FLS (especially in the complex geo- and atmospheric conditions of Europe) is expected, because FLS emits at temperatures close to the Earth's surface, resulting in only small differences between the LIR spectral signatures of clear-sky and FLS conditions. Although the algorithm shows some limitations in distinguishing between clear-sky and FLS, its performance is acceptable for meeting the objectives of the study. Especially, given that it is a lightweight approach relying solely on calibrated MSG-SEVIRI LIR data. Nevertheless, its accuracy can be further enhanced by incorporating auxiliary datasets. For example, integrating digital elevation maps (e.g., as in Egli et al., 2018), and surface or atmospheric temperature and humidity information from external sources (e.g. from reanalysis as in NWC SAF, 2019) could improve *POD* and *FAR*. Additionally, applying pixel- or spatial-based pre-/post-processing steps, such as those used in Pauli et al., (2022) and Andersen and Cermak, (2018) may help better identify FLS events and correct potential misclassifications, leading to a decreased *FAR*. These enhancements, however, fall outside the scope of the present study and are left for future work.

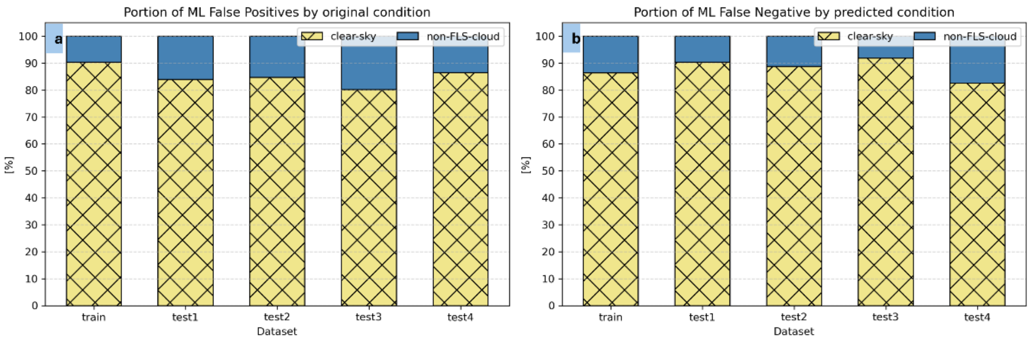


Figure 3. Detailed information on false-positive and false-negative (see Appendix B) predictions of the FLS ML technique across training, test1, test2, test3, test4 datasets. Panel (a) shows the percentage of false-positive instances that were clear-sky and non-FLS-cloud contaminated according to the ground truth dataset (i.e., MFL). Panel (b) shows the percentage of false-negative instances that were categorized as clear-sky and non-FLS-cloud by the ML algorithm.

“

And the following in Section 5 (Summary, Conclusions and outlook) at lines 480-485:

“

While the algorithm achieves the accuracy required for this study, a closer examination of its false-positive and false-negative (see Appendix B) predictions across the five datasets suggests that it could be improved in distinguishing between clear-sky and FLS. Its accuracy can be enhanced by incorporating auxiliary datasets (e.g. digital elevation maps and surface/atmospheric temperature and humidity information from external sources) and pixel- or spatial-based pre-/post-processing steps.

”

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

# Referee #2:

## Comments and suggestions

### *C#2.1*

In A#2.1 You agree that your study can benefit from comparison with external FLS products. However, your demonstrate that other products such as SAFNWC output, msg cpp algorithm or APOLLO-NG are not comparable to your results. Thank you for this long and documented answers.

However, you still did not mention any of these products and continue to show SOFOS as “the well-established and well-validated”.

I have no doubt on the quality of SOFOS, but such article cannot ignore the state of the art. SOFOS is not the only attempt to derive cloud low cloud data from Meteosat Second Generation. Moreover, the mention “well-establised” is not so appropriate. In which community is it “well established” ? Is there any operational weather service using SOFOS ?

I ask to the authors to cite more works on similar or related low cloud property retrieval from MSG satellite to inform the reader that author works and its associated reference SOFOS are valuable works among others.

Other comments have been well answered, modifications on the text have been successfully done.

### *A#2.1*

We describe SOFOS as an “established” algorithm in that it has been used as the basis for numerous scientific studies (about 10-20) since its first publication. It has been implemented by Deutscher Wetterdienst (DWD) staff into the satpy framework as fogpy as a basis for nowcasting applications. However, we do not insist on the label “well-established” and have replaced it by “existing” and discarded the term “state-of-the-art” where appropriate in the revised manuscript. Please see the Manuscript with track changes at lines 22, 88, 125, 315, 337, and 474.
Inline with your suggestion, the following references are now mentioned in the revised manuscript (please see lines 45, 60, and 410):

- NWC/GEO (NWC SAF, 2019),
- APOLLO-NG (Klüser et al., 2015),
- (Fuchs et al., 2022),
- MSG-CPP (we were not able to locate a reference for it) , and
- (Pauli et al., 2024)

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

## References

Andersen, H. and Cermak, J.: First fully diurnal fog and low cloud satellite detection reveals life cycle in the Namib, Atmos. Meas. Tech., 11, 5461–5470, https://doi.org/10.5194/amt-11-5461-2018, 2018.

Egli, S., Thies, S., and Bendix, J.: A hybrid approach for fog retrieval based on a combination of satellite and ground truth data, Remote Sens., 10, https://doi.org/10.3390/rs10040628, 2018.

Fuchs, J., Andersen, H., Cermak, J., Pauli, E., and Roebeling, R.: High-resolution satellite-based cloud detection for the analysis of land surface effects on boundary layer clouds, Atmos. Meas. Tech., 15, 4257–4270, https://doi.org/https://doi.org/10.5194/amt-15-4257-2022, 2022.

Klüser, L., Killius, N., and Gesell, G.: APOLLO-NG – A probabilistic interpretation of the APOLLO legacy for AVHRR heritage channels, Atmos. Meas. Tech., 8, 4155–4170, https://doi.org/10.5194/amt-8-4155-2015, 2015.

NWC SAF: Algorithm Theoretical Basis Document for Cloud Product Processors of the NWC/GEO (v2.1), https://doi.org/https://www.nwcsaf.org/Downloads/GEO/2018/Documents/Scientific_Docs/NWC-CDOP2-GEO-MFL-SCI-ATBD-Cloud_v2.1.pdf, 2019.

Pauli, E., Cermak, J., and Andersen, H.: A satellite-based climatology of fog and low stratus formation and dissipation times in central Europe, Q. J. R. Meteorol. Soc., 148, 1439–1454, https://doi.org/10.1002/qj.4272, 2022.

Pauli, E., Cermak, J., Bendix, J., and Stier, P.: Synoptic Scale Controls and Aerosol Effects on Fog and Low Stratus Life Cycle Processes in the Po Valley, Italy, Geophys. Res. Lett., 51, https://doi.org/10.1029/2024GL111490, 2024.