# Response to the Reviews

May 2, 2024

Arndt Kaps
Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR)
Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany
Email: arndt.kaps@dlr.de

Earth System Science Data
Manuscript Number: essd-2023-424

Dear Editor:
We have taken the reviewer's comments into account and agree that application of ML models to out-of-distribution data such as a different sensor is generally very difficult and great care has to be taken. As will be elaborated below, our approach is justified, however, due to both the similarities of the datasets (intercalibration, heritage channels) and the fact that we found good validation results against the 2B-CLDCLASS-LIDAR (CC-L) ground truth.
The main points of this discussion have been added to the revised manuscript.
Yours sincerely,

*Arndt Kaps on behalf of the author team*

## Response to Referees

### Earth System Science Data

| | |
|---|---|
| Manuscript: | essd-2023-424 |
| Title: | Characterizing clouds with the CCClim dataset, a machine learning cloud class climatology |
| Authors: | Arndt Kaps, Axel Lauer, Rèmi Kazeroni, Martin Stengel, Veronika Eyring |
| Date: | May 2, 2024 |

## Referee: 3

We thank the anonymous referee for his/her comments and for supporting publication of our manuscript. The original referee's comments are given in **blue**, our answers in **black**. All line numbers refer to the "track changes" version of the revised manuscript.

Overall I'm satisfied with your responses and revisions, but point #4 involves a fundamental issue for transfer learning that cannot be overlooked no matter viewing this issue from ML angle or from data rigor angle. Unfortunately this issue relates to the fundamental approach you use so I have to suggest major revision. Having said that, I do see and appreciate the discussion on scientific applications of your dataset. So I wholeheartedly support a final publication of your paper and dataset. ML is a great tool and should be embraced by our community, but we can't lose the rigor (which is also strongly suggested by numerous AI/ML research). In your case, it can be fixed using a collocated AVHRR-CC-L/CALIPSO dataset for training.

The second stage of the ML algorithm (random forest regression) is needed to be able to apply the algorithm to coarse-resolution data (having global climate models in mind). We share the concern of the reviewer that the ML algorithm might not be applicable to other datasets than MODIS. For this reason, we verified that the ML algorithm can indeed be applied to the out-of-distribution ESACCI dataset by comparing predictions of the ML algorithm applied to ESACCI data to the distribution of the original CC-L cloud class labels. As documented in Kaps et al. (2023), we found reasonable reproduction of the geographical distribution of all cloud types, which suggests that the method is robust enough to be applied to different input datasets if the data represent similar basic physical properties. In many settings, ML models have to be applied out-of-distribution an Kuma et al. (2023), Wang (2019), and Yuval and O'Gorman (2020) are some examples in climate science. When doing so, it is important to quantify the uncertainties induced by the domain shift, which for our method is documented in (Kaps et al., 2023).

Therefore, we disagree with the reviewer that in the case of deriving cloud class labels from physical cloud properties, the ML algorithm trained on MODIS data cannot be applied to AVHRR data. As an example, we reproduce Table VI from Kaps et al. (2023) quantifying the differences and correlations between the mean predicted fractions and CC-L per $2^\circ \times 2^\circ$ grid cell, thus demonstrating reasonable agreement 1 and allowing estimation of the uncertainty when applied to ESACCI.

We are using ESA's Cloud_cci dataset as it provides a consistent, long-term time series of cloud properties obtained from harmonized, reprocessed products from different satellite instruments. Especially relevant in

Table 1: Mean fraction of the predicted classes compared with the relative amounts of the classes in CC-L. $c_P$ is the Pearson-correlation between the geographical distributions. The last row shows the mean difference for pixels with predictions in the 90th percentile $\Delta_{90}$ relative to the mean $\mu_{90}$ of these predictions. Predictions are taken from a model trained on $(100\text{ km})^2$ from CUMULO and applied on 100×100 pixel Cloud_cci grid cells.

| | Ci | As | Ac | St | Sc | Cu | Ns | Dc |
|---|---|---|---|---|---|---|---|---|
| Predict. | 0.13 | 0.14 | 0.19 | 0.01 | 0.31 | 0.10 | 0.10 | 0.02 |
| CloudSat | 0.20 | 0.13 | 0.11 | 0.05 | 0.27 | 0.12 | 0.09 | 0.03 |
| $c_P$ | 0.87 | 0.80 | 0.60 | 0.18 | 0.88 | 0.84 | 0.83 | 0.36 |
| $\Delta_{90}/\mu_{90}$ | -29% | 49% | 49% | 1% | 18% | 14% | 30% | 39% |

that case is that two of the five AVHRR channels are intercalibrated with MODIS Aqua (Heidinger et al., 2010; Stengel et al., 2017) and that MODIS uses AVHRR heritage channels for its cloud property retrievals (Platnick et al., 2017). Version 3 of Cloud_cci has been shown to provide good global quality scores for cloud detection, cloud phase and ice water path based on validation results against A-Train sensors (Stengel et al., 2020). Collocating Cloud_cci data with CC-L is, however, non-trivial and would only provide a small number of usable samples for the few instances in which orbits overlap. Therefore, this collocation is in our opinion beyond the scope of this study.

Following the point of the reviewer, we added more discussion on the out-of-distribution application of our ML algorithm to the revised manuscript (lines 93-100):
"Application of an ML model to a different dataset (out-of-distribution) can be problematic and great care has to be taken to ensure no unexpected errors occur. Out-of-distribution application of ML models has been done in climate science before(e.g Kuma et al., 2023; Wang, 2019; Yuval and O'Gorman, 2020). Here, it is important to estimate the uncertainties induced by the domain shift. In our case, the regression model trained on MODIS cloud properties is applied to similar AVHRR retrievals. Applicability could be demonstrated by the reasonable reproduction of the geographical distribution of all cloud types when applying the model to ESACCI data (AVHRR), also providing an uncertainty estimate as documented in Kaps et al. (2023). This suggests that the method is robust enough to be applied to different datasets if they represent similar basic physical properties."

# References

Heidinger, A. K., W. C. Straka, C. C. Molling, J. T. Sullivan, and X. Wu (2010). "Deriving an inter-sensor consistent calibration for the AVHRR solar reflectance data record". In: *International Journal of Remote Sensing* 31.24, pp. 6493–6517. ISSN: 1366-5901. DOI: 10.1080/01431161.2010.496472.

Kaps, A., A. Lauer, G. Camps-Valls, P. Gentine, L. Gomez-Chova, and V. Eyring (2023). "Machine-Learned Cloud Classes From Satellite Data for Process-Oriented Climate Model Evaluation". In: *IEEE Transactions on Geoscience and Remote Sensing* 61, pp. 1–15. DOI: 10.1109/TGRS.2023.3237008.

Kuma, P., F. A.-M. Bender, A. Schuddeboom, A. J. McDonald, and Ø. Seland (2023). "Machine learning of cloud types in satellite observations and climate models". In: *Atmospheric Chemistry and Physics* 23.1, pp. 523–549. DOI: 10.5194/acp-23-523-2023.

Platnick, S. et al. (2017). "The MODIS Cloud Optical and Microphysical Products: Collection 6 Updates and Examples From Terra and Aqua". In: *IEEE Transactions on Geoscience and Remote Sensing* 55.1, pp. 502–525. DOI: 10.1109/tgrs.2016.2610522.

Stengel, M. et al. (2017). "Cloud property datasets retrieved from AVHRR, MODIS, AATSR and MERIS in the framework of the Cloud_cci project". In: *Earth System Science Data* 9.2, pp. 881–904. DOI: 10.5194/essd-9-881-2017.

Stengel, M. et al. (2020). "Cloud_cci Advanced Very High Resolution Radiometer post meridiem (AVHRR-PM) dataset version 3: 35-year climatology of global cloud and radiation properties". In: *Earth System Science Data* 12.1, pp. 41–60. DOI: 10.5194/essd-12-41-2020.

Wang, Z. (2019). *CloudSat 2B-CLDCLASS-LIDAR Product ProcessDescription and Interface Control Document*. Version p1_R05. URL: https://www.cloudsat.cira.colostate.edu/data-products/2b-cldclass-lidar.

Yuval, J. and P. A. O'Gorman (2020). "Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions". In: *Nature Communications* 11.1. ISSN: 2041-1723. DOI: 10.1038/s41467-020-17142-3.