

Answer to Anonymous Referee #1

by Julia Thomas & Annika Oertel on behalf of all Co-Authors

01 Oct 2026

We thank Anonymous Referee #1 to accept the review and appreciate their feedback. In the following, we provide a brief initial response to the two key issues raised by the reviewer, namely (1) the data access and (2) the structure of the manuscript.

1 Reply and explanations to comments about the data

We appreciate the reviewer's efforts to verify data accessibility and their constructive suggestions for improving access to the datasets. Below, we clarify the comments on data access.

Data volume

The original volume of the dataset - 1.1 TB per experiment - requires consideration of efficient options to compress the analysis files. We therefore compressed the GRIB files using Zstandard (zstd) and bundled them into monthly tar archives. Monthly directories were chosen to simplify the download for potential users, aiming to directly download the complete dataset for example for statistical analyses. Unfortunately, we did not sufficiently account for users who wish to access individual files or shorter time periods. To simplify the review process and access to data for reviewers, we had prepared a repository containing two days of example analysis files and provided this link to the ESSD editors. We apologize that the respective link has not been forwarded to you. Please find 48 example files of the Swabian MOSES 2023 campaign reanalysis here: <https://bwsyncandshare.kit.edu/s/6xPpfan4Ex4LaHo>.

Data Format

Data are available in GRIB format. GRIB is a standardized, highly compressed binary file format used to store meteorological data. It is maintained by the World Meteorological Organization (WMO; <https://www.dwd.de/EN/aboutus/it/functions/Teasergroup/grib.html>) and therefore follows their conventions as well as the conventions of operational weather forecasting centers such as the German weather service (DWD).

We apologize that this procedure caused you so much effort and that you were unable to open the decompressed GRIB files. The compressed dataset should be accessible in the following way:

(1) File extraction using

```
for file in *.zst; do
    zstd -dk "$file" -o "../path2data/${file%.zst}"
done
```

(2) And subsequently opening the individual analysis GRIB files as:

```
import xarray as xr
ds = xr.open_dataset(grib_file, engine="cfrib")
```

Please note that this requires the installation of the python cfrib package (<https://anaconda.org/conda-forge/cfrib>), in addition to xarray. We have independently tested the download, data extraction, and reading of the data.

Planned Revisions

We will optimize access to the data by restructuring the compressed folders. Specifically, we will provide several data repositories, one for each month and experiment, containing the original analysis files in GRIB format. That way, each user can decide whether to only download individual files or obtain multiple files in bulk as a tar archive.

Incomplete CSV-file

The csv files contain observation space statistics: As described in the manuscript, there are no additional campaign observations (SWM) assimilated in the control experiment (CTRL), and thus it is natural and correct that the respective columns in `ObsSpaceStatistics_20230_CTRL_BF_domain_1level.csv` have no entries. We will add a specific note of caution in the repository.

2 Reply and explanations to comments about the manuscript

We appreciate feedback on manuscript structure, which gives us the chance to reflect on the text and the articulation of the key messages. The key comments address (1) the length of the manuscript, (2) an unclear manuscript structure related to unclear goals and overly information dense figures. Below, we clarify the comments on the manuscript.

Manuscript length

- We included a comprehensive introduction to match it to the scientific background of potential users, ranging from data scientists, modelers specialized in physics parameterizations to DA experts. Therefore, we tried to motivate the use cases of the dataset from multiple perspectives. Importantly, we put our campaign reanalysis into context of existing reanalyses with the goal to differentiate it from existing datasets and outline the novelty of our approach and dataset. We will critically revise this part and streamline the introduction.
- To reduce the overall length, we aim to streamline the manuscript and focus on emphasizing the implications of the shown results.

Manuscript structure and content

- The overarching goal (i.e., the presentation of our reanalysis datasets) is stated in the introduction (see Subsection ‘A new reanalysis dataset based on the Swabian MOSES 2023 campaign’). We agree that the remaining manuscript outline, specifically Sections 4 and 5, has not been explained and motivated well. In particular, for readers that are not familiar with the concept of observation space and model space evaluation, Sections 4 and 5 might not sound very intuitive. We will re-think and clarify the structure and motivate clearly, why both perspectives are needed to thoroughly characterize the dataset.
- The information rich figures in Section 4 comprise typical ‘observation space statistics’ considered in data assimilation studies. These indeed show the reviewer’s specific suggestion 2: They illustrate that (i) the assimilation of additional observations works as it is supposed to, (ii) their assimilation does not degrade the analysis at the location of the observation, (iii) and that the agreement with the other operational observations is not degraded either. Such diagnostics are commonly applied in data assimilation, see e.g. Vural et al. (2024), Fourrié et al. (2019), Lange and Janjić (2016), Weissmann and Cardinali (2007), and Keil and Cardinali (2004). During revision of the manuscript, we will focus strongly on discussing the implications of the presented results as suggested by the reviewer.
- Complex observation error statistics are important to confirm that the a-priori choice of the set observation error is justified. Observation number statistics are important to assess the robustness of the statistics. We will streamline this Section and discuss the implications.
- As the reanalysis assimilates as many observations as possible, there is no remaining independent observation to validate the CMPG reanalysis (see specific aspect 3 raised by the reviewer). However, observation space statistics in Section 4 illustrate that on average for *all* observations the analysis is closer to the observations than it is for the model first-guess to observation. Evaluation of its added value for weather prediction can only be performed with reforecasts, i.e., by initializing forecasts from both CMPG and CTRL and subsequently validating the forecast quality. This is current work in progress and beyond the scope of this manuscript, which describes the reanalysis. Systematic reforecast experiments are listed as potential application of this dataset (see specific aspect 5 in the Review).

Additional comments

- We apologize for not providing a clear statement about which analysis files are included in the published dataset in the manuscript. The data contain the deterministic analysis files which are used in Section 5. Additional data, such as the 40 member ensemble, the ensemble mean and spread, are not published but can be made available upon request. We will clarify this in the manuscript.
- Specific aspect 5: Potential applications are mentioned in the concluding Section in lines 639 - 653. We can further elaborate on the listed applications.
- Specific aspect 6: We strongly agree to this aspect. In a revised version of the manuscript, we will conclude with a clear statement for what the dataset is explicitly useful.

References

- Fourrié, Nadia et al. (July 2019). “The AROME-WMED reanalyses of the first special observation period of the Hydrological cycle in the Mediterranean experiment (HyMeX)”. en. In: *Geoscientific Model Development* 12.7, pp. 2657–2678. ISSN: 1991-9603. DOI: 10.5194/gmd-12-2657-2019. URL: <https://gmd.copernicus.org/articles/12/2657/2019/> (visited on 06/18/2025).
- Keil, Christian and Carla Cardinali (Oct. 2004). “The ECMWF reanalysis of the MAP Special Observing Period”. en. In: *Quarterly Journal of the Royal Meteorological Society* 130.603, pp. 2827–2849. ISSN: 0035-9009, 1477-870X. DOI: 10.1256/qj.03.89. URL: <https://rmets.onlinelibrary.wiley.com/doi/10.1256/qj.03.89> (visited on 07/01/2025).
- Lange, Heiner and Tijana Janjić (May 2016). “Assimilation of Mode-S EHS Aircraft Observations in COSMO-KENDA”. en. In: *Monthly Weather Review* 144.5, pp. 1697–1711. ISSN: 0027-0644, 1520-0493. DOI: 10.1175/MWR-D-15-0112.1. URL: <https://journals.ametsoc.org/doi/10.1175/MWR-D-15-0112.1> (visited on 06/01/2026).
- Vural, Jasmin et al. (Jan. 2024). “Improving the representation of the atmospheric boundary layer by direct assimilation of ground-based microwave radiometer observations”. en. In: *Quarterly Journal of the Royal Meteorological Society* 150.759, pp. 1012–1028. ISSN: 0035-9009, 1477-870X. DOI: 10.1002/qj.4634. URL: <https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.4634> (visited on 05/13/2025).
- Weissmann, Martin and Carla Cardinali (Jan. 2007). “Impact of airborne Doppler lidar observations on ECMWF forecasts”. en. In: *Quarterly Journal of the Royal Meteorological Society* 133.622, pp. 107–116. ISSN: 0035-9009, 1477-870X. DOI: 10.1002/qj.16. URL: <https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.16> (visited on 08/05/2025).