

# A convective-scale reanalysis for the ‘Swabian MOSES 2023’ field campaign by Thomas et al 2026

Review by Anonymous

September 28, 2026

The submitted manuscript is about a reanalysis data set created using the quasi-operational data assimilation framework of the DWD. By including a dense observation network from the Swabian MOSES 2023 field campaign, the dataset promises to be the most accurate possible reanalysis for this specific region over three months. In addition to those interested in the exact state of the atmosphere at that place in time, I expect the dataset to be of interest to researchers interested in convective-scale data assimilation, regional NWP, and field campaign organizers. I believe the dataset matches the goals of the journal well. (Disclaimer, this is my first time reviewing for ESSD).

However, the submitted manuscript has two critical flaws. The first is that I failed to actually access the data. The second is that I can not follow the structure of the manuscript. For these reasons I recommend the paper be rejected, but highly encourage the authors to rethink how the data should be accessed, and revise the manuscript for resubmission. I’ll begin by explaining my issues with the data, followed by my comments on the manuscript.

## The data

To look at the data I downloaded the SWM data for June. The first issue, it is only possible to download the full month. This is annoying, but I downloaded the full 250GB over night. Now the next annoyance. The downloaded tar file contains a single folder, unhelpfully name Unnamed entity, which in turn contains a single file, which is also a tar file. So I then needed to extract the tar file to get at the new tar file, which requires about 500GB of free storage. From the CMPG\_202306.tar file I could then extract single zst files. These I then tried to open using the code snippet provided (`import xarray as xr, ds = xr.open_dataset(name,engine='cfgrib')`), with two different python and xarray versions, which both ended in errors complaining about needing an ".idx" file. At this point I gave up.

I did manage to download and open the csv files of the observation space statistics, but at first glance these are missing data. For example in `ObsSpaceStatistics_20230_CTRL_BF_domain_1level.csv` the swm columns all have no entries.

## The manuscript

I do not understand how this manuscript is structured or what’s its goals are. The introduction is extremely long, but it does not explain what is happening in manuscript. What are these goals? In a regular journal I’d be asking about scientific questions, but here the focus should be on the data set. The dataset itself isn’t described much. For example, from the description I do not

understand if the dataset contains all 40 members, or just one, or maybe just the mean. A large part of chapter 4 and 5 are descriptions of very information dense plots. Why are you showing me all this? And why describe the plots, I can see those. What I need is help understanding the implications of what is shown. I'd highly recommend the authors completely revise their draft, turning it into a straight forward and clearly structured argument of what the data is and is not. Here a simplified possible example:

1. The reanalysis contains many more observations, and more observations is always better in theory, but in practice A, B, or C can make things worse. A shouldn't be a problem, but as many of these additional obs are not routinely assimilated, B and C could be an issue.
2. To make sure B and C are not an issue, we look at this, that, and this other thing, which shows that everything is working as it should.
3. Now that we have convinced you these observations are assimilated correctly, let's look at their impact by comparing both of the reanalysis against something independent in some sensible way.
4. Since the observations have the following characteristics, and the model has these characteristics, we expect the biggest improvements to be X, Y, and Z.
  - Oh wow, X is even better than we expected.
  - Y is not as good as hoped.
  - Z is better, but not in the way we expected.
5. And lastly, here are some cool examples of things you can do with our data.
6. In conclusion, this is a great dataset for these things, especially if you are interested in X and Z.