

**Ref:** essd-2022-334\_review

**Title:** EURADCLIM: The European climatological high-resolution gauge-adjusted radar precipitation dataset

**Authors:** Aart Overeem\*, Else van den Besselaar, Gerard van der Schrier, Jan Fokke Meirink, Emiel van der Plas, and Hidde Leijnse

### **General comments**

The data set is unique in the sense that the radar-rain gauge merging (hourly) is done at a European scale (over OPERA radar coverage) and is easily accessible. Sources of errors and limitations are discussed, and despite some limitations of the EURADCLIM dataset, this dataset can be useful in providing high-resolution (2 km, 1 hour) precipitation information.

However, the manuscript (and dataset) lacks clarity regarding its usage for further applications (e.g., missing indicators for erroneous EURADCLIM data, mainly literature-based ideas for potential applications, unclear definition of extreme cases). The authors created 1-hour and 24-hour rainfall accumulations with separate references. From the 1-h dataset, the 24-h dataset is generated in an hourly moving window and is 3 times larger in size than the 1-h dataset. The manuscript does not explain any special processing for the generation of this 24-h dataset nor its benefit. It is not clear why 24-h dataset should be produced separately with a separate reference.

Overall, the topic fits the scope of the journal's special issue, but the manuscript needs to be better clarified for its publication. Please see the comments below. Line numbers are indicated with "L"

Recommendation: Major

1. Section 4.3 and Fig.8: Two examples are presented to show the limitation of the EURADCLIM data (e.g., failing of clutter filtering effect). However, those radar echoes seem clearly a failure of radar processing for certain time steps (0045-0630 UTC which affected hourly EURADCLIM data from 0100-0700 that are more than 4 hours - please see also comment #11-, as briefly mentioned in L340 as well), which is unlikely "extreme precipitation" as stated in Figs. 8a and 8b. i) Have the authors cross-validated those events with other sources confirming the rainfall was extreme over the affected areas? ii) If the EURADCLIM aims to construct quality-checked past data (not in real-time adjustment) and those erroneous outputs are caused by at least a processing artefact (among other artefacts) that cannot be corrected even with the gauge adjustment for the limitation they pointed out, then the output analyses are expected how to deal with those time steps in their hourly merging and 24-hour accumulations or further applications. Instead, those results were explained in the context of the added value of the gauge adjustment (L347), which lacks discussion of the "limitation" of the EURADCLIM data. Although in L365-L372, the authors briefly mentioned a couple of ways to deal with those outlier cases, these do not seem to be applied to the current version of the EURADCLIM dataset. If an end-user is aware of such an outlier, the current version of the dataset will be unlikely used. Can the authors better propose or discuss what can help the users to use their climatological dataset?

2. Section 4.4: This section lacks a supporting explanation. How do they identify extreme events out of 8 years dataset? only by the radar 24 h accumulations? or a couple of flooding cases? For the presented extreme case analyses, were the outliers visually checked (as mentioned in section 4.2)? Over the flood-affected areas, have the authors checked that there were no rainfall estimates available produced (reanalyzed after the event) by regional and national products (that can justify L389-390)?
3. Unclear conclusion: (L423-L455). I am puzzled by their “recommendations” that seemed to target those who may replicate EURADCLIM. The improvement of EURADCLIM data will be made most likely by the authors, no? On the other hand, if the authors tried to promote their datasets to be used, I expect the recommendations in the context of the user perspective with some technical tips or examples, which perhaps they intended to address some with “strategic value” (L460-472). However, those “strategic values” are mostly speculated by addressing examples from the literature. One or two examples with detailed guidelines on the usage of the EURADCLIM data will be more useful (e.g., by improving section 4.4).

### Minor comments

1. L43-L44: The method of Park et al. (2019) is rather based on a systematic bias adjustment, instead of merging. Hence, the following sentence can be revised (with bold) as “an operational **gauge-adjusted** OPERA-based radar rainfall product for the European Rainfall-Induced Hazard Assessment (ERICHA) system. ~~employing data from ~1500 rain gauges.~~ This is used to compute flash flood hazard for Europe for the next 5-6 hours for the European Flood Awareness System (EFAS)”. Here, the number of rain gauges can be misleading because the daily bias map has been obtained from valid radar-rain gauge pairs, which may vary from different gauge sources, quality of both radar and gauge estimates, and days of rainfall over available gauge points. In any case, the algorithm can adapt to the use of more gauges.
2. L45-L46: “The use of this dataset is restricted to EFAS”, This is not correct (e.g., the dataset and the adjustment algorithms have continuously been used in the framework of several research projects since 2017) and do not add any relevant information; so, please remove it.
3. L50-L51: Please remove “and data from far more rain gauges are available (~7700)”. Perhaps the authors can mention that their dataset is built with different gauge sources from Park et al. (2019), but it is not necessary to point out the number of gauges unless adding an interesting research argument. The more gauges available, the better both methods would perform. In fact, the authors describe the gauge network used in the presented methods (in L104-L120), addressing the number of gauges deployed (~7700), which fits better here and shows the distribution clearly as well.
4. L56-L57: “Since it could also be applied in (near) real-time, its evaluation is also relevant for the existing gridded OPERA products”. Can this be better clarified?
5. L76: “It is based on the raw single site radar data, which have undergone Doppler clutter filtering.” Can this part be better explained?
6. L90-L95: Are the values of “nodata” and “undetected” directly from the OPERA 15-minute rain rate product? Or when calculating the 1-h precipitation accumulation, those values are assigned to the EURADCLIM data cell if not satisfying the full availability (similarly done as the OPERA data)? Here, can this full availability be better explained?

7. Fig. 2b, and L107: What does “Combined radar-gauge data availability (in %)” mean?
8. L103: Is it necessary to state “mid-June 2022” here? It is actually mentioned better in L115.
9. L121-L130: The authors comment that there are some uncertainties on the gauge aggregation time for the given ECA&D 24-h accumulation dataset and explain that “The end times of the observations display a large variability” in L127. Then, is such variability considered in the disaggregation of the 1-hour dataset applied in L199-L206?
10. Fig.3: For disaggregation, are the 24-h radar precipitation composites used?
11. L205: “at most 4 hours per 24-h interval”, does this threshold of 4 hours come from the observed results? I am a bit puzzled, there were days with missing hours of more than 4 hours (as in the case of Fig. 8.a), and if this is the case, how would the correction(merging) be applied?
12. Fig. 5 a-d: Maps of mean hourly precipitation over 2013-2020. The mean values are so low, which is not so evident to see the effects over some areas. Is it meaningful to present in terms of mean 1-h precipitation to highlight the effect of clutter filters?
13. L271-L273: Although the authors state that there has been a strong reduction in the coastal area of Norway (5a-5d), it is not clear to see. Can it be explained better?
14. Table 1: Why the values of mean daily precipitation corresponding to 4 different filters are the same?
15. L289: I guess “rho” here is the correlation coefficient. (Please add the symbol in the table or explicitly state “correlation” next to the symbol in L289).
16. L297-L300: “This confirms the effectiveness of the algorithms to further remove non-meteorological echoes”. Can this be better clarified? For no-rain only (0 mm), the effect of the static mask seems to be none based on daily accumulations on average.
17. L305: Regarding the results of an overestimation of 11%, please indicate explicitly EURADCLIM in the text.
18. L311: How do the authors separate between summer and winter cases?
19. Fig.6b-6d: Why there are some systematic “0” or near 0 values of EURADCLIM for all values of gauges?
20. L334-335: “After all, the gauges that...in the final EURADCLIM dataset”. Can this be better clarified? Does it mean that those gauges are used in the evaluation of the EURADCLIM dataset? Or if these are a part of EURADCLIM, how does this justify (link to) the previous statement?
21. L337: “For other regions, radar beam-blockage could play a role.” – Can this be explained better related to Fig.5 and Fig. 7 (e.g., Is it possible to identify those areas which indeed indicate the presented filter that does not seem to work based on the 8-year dataset? If so, in the comparison with gauges, those areas could have been excluded as well).
22. Fig.9: Why not use the same colour scale as those presented in Fig. 5? Can this be done in terms of daily accumulation?
23. L349-L350: Could it be better explained where those signatures are shown (in Fig. 9a or in Fig 9)? As mentioned in #22, can the results be better explained with respect to Fig. 5?
24. L359: “It is difficult to tell to what extent non-meteorological echoes play a role”. It is possible that invalid(strange) radar data (similar to the presented case in Spain) are included in the 8-year statistics over Croatia, Kosovo, Slovakia and Corsica. Have the authors checked those areas in particular?

25. L407-L409: This conclusion needs better supporting materials or references; Because the presented method is based on neither sub-daily gauge data nor a real-time feed environment, it is not clear yet what aspects can be improved. For production in real-time, better improvements in L365- L375 (Major comment 1) are expected.
26. Code availability: It is good that the authors provide some routines, but this code will not be sufficient to “reproduce” the results and the end-users required a good knowledge of the OPERA data structure. So, rephrase or remove “This helps end users to reproduce results from this study and to further explore and analyze the EURADCLIM dataset.”
27. Data availability: Hourly dataset is useful, however, if the 24-h precipitation datasets are simply generated by summing up hourly data in running windows, is there any specific reason to produce such big-size outputs with a separate reference?