

Referee 1: Georgia Papacharalampous

DOI: 10.5194/egusphere-2026-4282-RC1

Date of comment: 21.08.26

No.	Comment	Response
1	<i>Lines 73–74: Was any additional quality control applied to the dataset beyond those originally performed by the data provider?</i>	We did not perform any additional quality control. However, as selection criteria for choosing our two case study stations, we had set that they (a) be from divergent Austrian precipitation regions (as defined by Yaqub et al., 2011) and (b) have sufficiently complete time series. We have added this information in a more explicit form to the relevant line.

Changes in the manuscript.

(preprint L73–74) “No additional quality checks were performed on the data beyond routine quality control by the provider. However, one criterion for the selection of the two stations for the case study was data completeness (with values of 98.1% and 93.0% at the selected stations).”

2	<i>Lines 93–94: Maybe, something like “weights equal to the cluster share of 10-min above-threshold (wet) observations” instead of “weights equal to the cluster share of 10-min observations”?</i>	Thank you for the careful reading – this comment showed us that our language on thresholds for filtering the precipitation time series was imprecise in several places. We did not adopt the specific wording “above-threshold (wet)”, because the fits use all positive 10-min values within an identified event. We now consistently refer to “positive within-event 10-minute observations”.
---	---	---

Changes in the manuscript.

(preprint L93–95) “4. We obtain an overall precipitation distribution as a weighted mixture of the cluster-wise distributions, with weights equal to the cluster’s share of the positive within-event 10-minute observations. Finally, we compare our aggregated, cluster-wise mixture to a baseline Gamma distribution fitted to all positive within-event observations irrespective of clusters.”

(preprint L180–182) “We fitted Gamma distributions to strictly positive precipitation intensities, focusing on wet-interval (wet-day) distributions. Specifically, we excluded zeros and sub-threshold values (at or below the wet-interval threshold used in event definition) and all positive 10-min precipitation intensities within the identified events. Then we estimated parameters by L-moments – once using all positive precipitation observations to obtain a baseline distribution, and once separately for each cluster.”

(preprint L191) “Here, w_k is the proportion of positive within-event precipitation observations assigned to cluster k , n_k is the number of positive within-event observations in cluster k , [...].”

3	<i>Lines 80–81: Since the number of flashes per hour were available on a $1\text{ km} \times 1\text{ km}$ grid, maybe something like this is closer to the approach followed: “For each station, an event was flagged as convective if at least one flash occurred during the event window within a discrete raster buffer comprising all grid cells whose centers fall within a 7 km radius from the station.” (or “all grid cells whose boundaries encompass...”)</i>	We thank the referee for the valuable reminder about the structure of the lightning data and have adapted the manuscript accordingly.
---	--	---

Changes in the manuscript.

(preprint L80–81) “For each station, an event was flagged as convective if at least one flash occurred during the event window within a discrete raster buffer comprising all grid cells whose centers fall within a 7 km radius from the station.”

continued on next page

No.	Comment	Response
4	<i>The severity, duration and mean intensity features are all included in the k-means clustering, even though mean intensity is severity divided by duration. This could introduce some redundancy or uneven weighting of information. Were alternative combinations of driving features explored and what motivated the final choice? What are the benefits of this specific combination?</i>	We share this concern, which we also considered during the analysis. While these three parameters of course stem from the classical drought event definitions provided by Yevjevich et al. (1967), they do exhibit a degree of redundancy that leads to high correlation between them - which is after all well known for its potential to destabilise clustering solutions (see e.g. Dolnicar et al., 2018, p. 50), as well as affect variable weighting. In our analysed stations, this correlation is especially high between severity and duration (0.67 and 0.77, respectively). For this reason, we conducted the partitioning clustering both on the complete, theory-driven variable set and on a reduced set consisting only of magnitude, duration, time_to_peak, and flash. The results of these analyses were virtually identical: In both stations, 95% and 96% of events respectively were assigned to the identical cluster. For our goodness-of-fit measures, the percentages of improvement over the simple Gamma model were identical for 10 out of 14 of our metric-station combinations (with changes of around 1% for the other four). As the choice of variable sets did not seem to have an impact on the results, our final choice was an editorial one: We decided that reporting on the complete “classical variable set” would improve the readability of the paper without affecting the results. We have not yet made any changes to the manuscript that would reflect this discussion point, but we are happy to provide the alternative analysis on the reduced variable set for the partitioning clustering case in the Supplementary Materials if so advised.
5	<i>Out of curiosity, was the clustering tested without the lightning feature in the two examined locations? As noted already in the paper, the proposed framework is feature-set agnostic. Removing the dependency on lightning data in future implementations could broaden the practical applicability of the framework in regions where such data are unavailable.</i>	In preliminary analyses, removing lightning tended to critically destabilise clustering solutions. We have now repeated the full workflow on our two stations without lightning: For clusterwise regression, this destabilised solutions and reduced final fit in both cases, while in the partitioning clustering, comparable results could be reached at least for station Dornbirn. We have added a brief discussion of these results, together with a recommendation for settings without lightning data, to the Discussion.

Changes in the manuscript.

(inserted after preprint L333) “Conversely, the question may arise whether the present approach can be applied in settings or world regions where lightning data are not available. In our analyses, removing lightning from the variable set substantially reduced both clustering stability and final fit in the clusterwise regression case, for both evaluated stations. In the partitioning case, it reduced stability at Graz, whereas results at Dornbirn were comparable to those obtained with lightning. Our recommendation for cases where lightning data are not available is thus to substitute other sources of information on convection, such as CAPE, and to select variables based on cluster stability and final goodness of fit, as is done in this study.”

continued on next page

No.	Comment	Response
6	<i>Maybe another brief discussion on the extent to which the proposed framework can be applied region-wise as opposed to strictly location-wise (station-specific) could be included. Similarly, it could be relevant to discuss how useful the framework is expected to be at other temporal resolutions, and what specific adaptations would be required to handle them.</i>	Thank you for the suggestion; we have now added our ideas on both topics to the Discussion. For regionalisation, we outline two options: combining the framework with classical regional frequency analysis (Hosking and Wallis, 1997), with station-specific weights incorporated into the clustering, or fitting a single pooled spatio-temporal model with static site covariates (e.g. elevation, longitude, latitude) to extrapolate to ungauged locations. Regarding temporal resolution, we discuss which of our event characteristics become less informative at coarser resolutions and suggest alternative variables.

Changes in the manuscript.

(inserted after preprint L333) “Regional or hierarchical mixtures that share information across stations could stabilise parameter estimates for data-sparse clusters. Additionally, the framework could be extended to estimate precipitation distributions at ungauged locations. One option is to combine the framework with established regional frequency analysis methods (Hosking and Wallis, 1997, pp. 9–11), with station-specific weights incorporated directly into the clustering procedure. Alternatively, a single spatio-temporal model could be fitted by pooling events from all stations and including static site-specific covariates, such as elevation, longitude, and latitude. This approach would allow the model to represent spatial variation and extrapolate precipitation distributions to ungauged locations.”

(inserted after preprint L333) “The proposed framework was optimised on sub-hourly data, where process heterogeneity has been well documented (Dallan et al., 2022; Marra et al., 2026). At coarser temporal resolutions, several of the event characteristics used here become less informative. For example, time to peak becomes uninformative when the accumulation interval approaches the event duration. Similarly, a binary lightning indicator has reduced discriminatory power, making convective events more difficult to identify. Moreover, reliable lightning observations have been available in Austria only from 2010 onward. Potential alternatives for identifying convective precipitation include convective available potential energy (CAPE) or weather-type classifications. Although some variables may need to be replaced when applying the framework at longer temporal scales (e.g., daily precipitation measurements), the general methodological framework remains applicable.”

7	<i>In relation to the specific questions posed and explored in the paper, it appears that the answers, particularly for question (ii), may be somewhat sensitive to the choice of case study. Under this view, maybe the paper could briefly discuss a broader, large-scale comparison of clustering approaches as a future direction towards a generalized implementation of the proposed framework.</i>	Thank you for pointing this out – while we had selected the two stations specifically to showcase the approach under diverse precipitation conditions, (1) this aspect was not discussed sufficiently in the submitted version, and (2), you are of course right, two stations from one single country remain an important limitation. We have now revised several passages in the manuscript to again strengthen the regional comparison aspect, and we have extended the discussion to critically evaluate our research questions in light of the limited number of stations in our case study.
---	---	---

continued on next page

No.	Comment	Response
Changes in the manuscript.		
	(preprint L252) “[...] In absolute terms, this improvement is more pronounced in Graz than in Dornbirn, [...]”	
	(inserted after preprint L275) “Taken together, the results indicate that the tendencies seen in this approach are the same for both stations from divergent Austrian precipitation regions: clusterwise mixed Gamma distributions improve overall fit, while this gain is larger for clusters derived by clusterwise regression than for partitioning clustering. At Dornbirn, where rainfall events are often longer and less intense than in Graz, the baseline fit of the simple Gamma distribution was better, but also for but even here the cluster-aggregated approach yielded a relative improvement of 73% in D_{KL} (comparable to the 66% encountered in Graz). This tendency held for both stations despite the contrasting precipitation regimes and different numbers of clusters K selected.”	
	(preprint L294–295) “[...] with larger absolute improvements at Graz and still visible gains at Dornbirn where the baseline already performs comparatively well; in relative terms, the reduction in D_{KL} was in fact somewhat larger in Dornbirn.”	
	(preprint L319) “Two Three limitations merit discussion.”	
	(inserted after preprint L333) “Third, the approach has so far been tested on only two stations from a single country. We selected these two stations specifically to test the behaviour of the approach in divergent precipitation regimes (see Yaqub et al., 2011), and regional influence seems to affect our three research questions differently:	
	Regarding (i), improving overall precipitation distribution fit by aggregating finite mixtures of distributions: We found this to be the case for both of our investigated stations and both clustering approaches applied.	
	Regarding (ii), as stated, the tendency for our stations is clear: Clusterwise regression notably outperforms partitioning clustering in improving the final distribution fit, while producing clusters that – at first glance – are harder to interpret as event types. However, this claim should be tested on a wider range of stations/regions and especially clustering algorithms.	
	Regarding (iii), location of improvements: For our two stations, relative improvement was generally higher in upper quantiles (see $MAE_{[Q_{60}, Q_{100}]}$, with again larger relative reductions for clusterwise regression than for partitioning clustering). This behaviour is especially pronounced for Dornbirn, the longer-duration, lower-intensity station, while for Graz, improvements are also quite marked in quantiles 20–40. This aspect appears the most station- and possibly region-specific of the three. Thus, a broader, large-scale comparison of clustering approaches across many stations and climatic regions is the natural next step towards a generalised implementation of this framework.”	
8	<i>Given that practical applications are highlighted as key motivations (such as stochastic weather generators), it would be relevant to also provide a brief conceptual outline on how the mixtures can be utilized in practice. For instance, in stochastic weather generators, this could involve simulating conditioned states and subsequently drawing precipitation values from cluster-specific distributions.</i>	Thank you for this helpful suggestion. We agree that the practical use of the proposed mixture framework should be clarified. We have added a short conceptual outline to the Discussion. In an event-based weather generator, the occurrence model would first generate event start times or inter-arrival times, possibly jointly with event type. The event-type probabilities may depend on season, location, circulation type, or other process covariates. Given the event type, event-level characteristics such as duration, severity, magnitude, and time-to-peak should be generated from their type-specific joint distribution. Finally, the within-event rainfall profile can be simulated conditional on both the event type and these event characteristics. We also clarify that preserving the multivariate dependence among event characteristics is essential. This could be achieved by resampling observed event vectors or normalised hyetographs within each type, by fitting type-specific multivariate or copula models, or by using conditional disaggregation models that generate a temporal rainfall profile consistent with simulated event properties. Thus, the proposed mixture should be viewed as a process-aware distributional/intensity component that can be embedded within a broader event-based weather generator, rather than as a complete generator by itself.

continued on next page

No.	Comment	Response
Changes in the manuscript.		
	<i>(inserted after preprint L341)</i> “For stochastic weather generation, the proposed framework is most naturally embedded in an event-based model. Event start times or inter-arrival times would first be generated, possibly jointly with event type. The event type may depend on season, location, circulation type or other process covariates. Conditional on this type, event characteristics such as duration, severity, magnitude and time-to-peak should be generated from their type-specific joint distribution, after which the within-event 10-min rainfall profile can be simulated from a type-specific distributional or disaggregation model. Such an implementation must preserve the multivariate dependence among event characteristics. This can be done by resampling observed event vectors or normalised hyetographs within each type, by fitting type-specific multivariate or copula models, or by using conditional disaggregation models. The mixture proposed here is therefore best viewed as a process-aware intensity component for a broader event-based weather generator, not as a complete generator by itself.”	
9	<i>Since an important application of the proposed framework is in stochastic weather generators, the manuscript could benefit from including a few additional key references in this domain to further strengthen the context.</i>	Thank you for this suggestion. We agree that the stochastic weather-generator context should be strengthened. We have added references covering both classical daily weather generators and event-based or sub-daily stochastic rainfall models. Specifically, we added Richardson (1981) as an early stochastic weather-generator reference, Kilsby et al. (2007) and Wilks (2010) for climate-impact and downscaling applications, and Rodriguez-Iturbe et al. (1987), Onof et al. (2000), and Burton et al. (2008) to acknowledge the event-based and sub-daily stochastic rainfall modelling tradition. This better positions our framework as a process-aware distributional component that could be embedded within existing weather-generator or event-based rainfall-generator structures.
Changes in the manuscript.		
	<i>(inserted after preprint L21)</i> “Stochastic weather generators range from early daily occurrence–intensity models (Richardson, 1981) to multi-site and climate-impact applications (Kilsby et al., 2007; Wilks, 2010). At sub-daily scales, event-based and point-process rainfall generators provide a closely related modelling tradition, representing rainfall as clustered storm or cell processes (Rodriguez-Iturbe et al., 1987; Onof et al., 2000; Burton et al., 2008). In both settings, the representation of wet-interval intensities remains central, motivating improved but still operationally tractable distributional formulations.”	
10	<i>A few typos and notation inconsistencies exist throughout the paper and should be addressed. For instance, the symbol “k” is used both for a cluster’s serial index and for the total number of clusters. Additionally, “Dimension 3” in line 205 should ideally be styled as “Dim. 3” (similarly for “Dimension 4” in line 206).</i>	We apologize for the typos and thank you for pointing them out. We have corrected them in the affected manuscript lines as well as figure labels/headings.
Changes in the manuscript.		
	<i>(throughout)</i> The total number of clusters is now denoted by K , the cluster index by k (preprint L148, 169, 171–177, 339; panel titles of Figs. 2–5; axis labels and captions of Figs. B1–B3). <i>(preprint L205–206)</i> “Dimension 3” and “Dimension 4” changed to “Dim. 3” and “Dim. 4”. <i>(preprint L322)</i> “increasing” corrected to “increasing”. In addition, we corrected the order of the station maxima (<i>preprint L72</i>) (“28.1 and 25.7 mm (10 min) ^{−1} ”, matching the station order Graz Universität, Dornbirn) and changed “events in Dornbirn” to “events at Dornbirn” (<i>preprint L111</i>).	

References

- Burton, A., Kilsby, C. G., Fowler, H. J., Cowpertwait, P. S. P., and O’Connell, P. E.: RainSim: A spatial–temporal stochastic rainfall modelling system, *Environmental Modelling & Software*, 23, 1356–1369, doi:10.1016/j.envsoft.2008.04.003, 2008.
- Dallan, E., Borga, M., Zaramella, M., and Marra, F.: Enhanced summer convection explains observed trends in extreme subdaily precipitation in the Eastern Italian Alps, *Geophysical Research Letters*, 49, e2021GL096727, doi:10.1029/2021GL096727, 2022.
- Dolnicar, S., Grün, B., and Leisch, F.: Market segmentation analysis, in: *Market segmentation analysis: Understanding it, doing it, and making it useful*, pp. 11–22, Springer, doi:10.1007/978-981-10-8818-6_2, 2018.
- Hosking, J. R. M. and Wallis, J. R.: *Regional frequency analysis*, Cambridge University Press, 1997.
- Kilsby, C. G., Jones, P., Burton, A., Ford, A., Fowler, H. J., Harpham, C., James, P., Smith, A., and Wilby, R.: A daily weather generator for use in climate change studies, *Environmental Modelling & Software*, 22, 1705–1719, doi:10.1016/j.envsoft.2007.02.005, 2007.
- Marra, F., Dallan, E., Canale, A., Prosdocimi, I., Papacharalampous, G., Borga, M., and Papalex-iou, S. M.: Apparent heavy tails of sub-daily precipitation explained by the coexistence of lighter-tailed processes, *Geophysical Research Letters*, 53, e2025GL119705, doi:10.1029/2025GL119705, 2026.
- Onof, C., Chandler, R. E., Kakou, A., Northrop, P., Wheeler, H. S., and Isham, V.: Rainfall modelling using Poisson-cluster processes: a review of developments, *Stochastic Environmental Research and Risk Assessment*, 14, 384–411, doi:10.1007/s004770000043, 2000.
- Richardson, C. W.: Stochastic simulation of daily precipitation, temperature, and solar radiation, *Water Resources Research*, 17, 182–190, doi:10.1029/WR017i001p00182, 1981.
- Rodriguez-Iturbe, I., Cox, D. R., and Isham, V.: Some models for rainfall based on stochastic point processes, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 410, 269–288, doi:10.1098/rspa.1987.0039, 1987.
- Wilks, D. S.: Use of stochastic weather generators for precipitation downscaling, *WIREs Climate Change*, 1, 898–907, doi:10.1002/wcc.85, 2010.
- Yaqub, A., Seibert, P., and Formayer, H.: Diurnal precipitation cycle in Austria, *Theoretical and Applied Climatology*, 103, 109–118, doi:10.1007/s00704-010-0281-z, 2011.
- Yevjevich, V. M. et al.: An objective approach to definitions and investigations of continental hydrologic droughts, vol. 23, Colorado State University Fort Collins, CO, USA, <http://hdl.handle.net/10217/61303>, 1967.