General comments:

The authors provide a new algorithm for disaggregation daily into hourly precipitation with special focus on preserving sub-daily maxima. The main innovation is to use a Fréchet–Hoeffding upper bound copula to condition sub-daily maximum precipitation values on daily totals. The method outperforms four well selected benchmark methods i.e. a nearest-neighbour resampling method, a Poisson cluster-based rainfall generator, a multiplicative random cascades and a deep learning technique. The manuscript is short, well written to the point concise avoiding unnecessary text burden. The presented approach is novel and shows excellent performance. I have only a few comments for improvement of the manuscript. Publication is recommended after minor revisions.

We sincerely thank the reviewer for this positive and constructive assessment of our manuscript. We appreciate the recognition of the novelty of the proposed Q-CODA approach, as well as its performance relative to the selected benchmark methods. We are also grateful for the reviewer's encouraging comments on the clarity and conciseness of the manuscript. We have carefully considered all of the reviewer's suggestions and have revised the manuscript accordingly. Below, we provide a detailed point-by-point response to each comment.

Comments for improvement:

1. Section 3.3: Please provide some more information about the implementation of the k-nn method, e.g. how is a neighbour sampled (equation) and how many neighbours are considered (equation).

We thank the reviewer for this suggestion. Following this comment, Section 3.3 will be expanded to provide a more explicit mathematical description of the K-Nearest Neighbours (KNN) implementation, including both the distance metric and the neighbour selection procedure. Specifically, since only one predictor (daily total precipitation) is used, the Euclidean distance (between the target daily total P_d^{target} and each candidate training value $P_{d,i}^{\text{train}}$ within the same season) reduces to the absolute difference:

$$d_i = \left|P_d^{target} - P_{d,i}^{train}\right|$$

The set of neighbours is then defined as the indices corresponding to the k smallest distances:

$$\mathcal{N}_k = argsort(d)_{i:k}$$

where $k = 10$. From this set, a single neighbour $j \in N_k$ is randomly selected with equal probability.

The associated hourly precipitation pattern $\boldsymbol{h} = (h_1, h_2, \dots, h_{24})$ of the selected neighbour j is scaled proportionally to match the target total:

$$\boldsymbol{h} \leftarrow \boldsymbol{h} \times \frac{P_{\text{d}}^{\text{target}}}{\sum_{\text{i}=1}^{24} \text{h}_{\text{i}}},$$

2. Section 4.4: This refinement of the hourly precipitation pattern in the 3rd step of the algorithm is quite empirical and may contradict the model concepts of step 1 and 2. Please comment.

We thank the reviewer for this insightful comment and for pointing out a lack of clarity in the description of Sect. 3.4 (referred to as Sect. 4.4 in the comment).

We would like to clarify that step 3 of the methodology consists of two conceptually distinct components.

The first component is a deterministic iterative adjustment procedure that modifies the initial KNN-based hourly pattern (the KNN seed obtained in step 2) so that it simultaneously satisfies the prescribed daily total and the sub-daily maxima obtained in step 1 via the comonotonic transformation. This component is fully consistent with the modelling framework, as it enforces the dependence structure and constraints defined in the previous steps.

The second component is a refinement process targeting the temporal dependence structure of the hourly series. Its primary purpose is to address a systematic limitation of KNN, namely the underestimation of temporal autocorrelation, as reported by Li et al. (2018) and also observed in our evaluation (Fig. 6 in the manuscript). KNN does not explicitly preserve the persistence properties of the original time series, often resulting in reduced autocorrelation. The refinement process in Q-CODA is therefore designed to correct this bias by locally redistributing precipitation within short temporal windows in a way that increases autocorrelation. Importantly, this refinement operates strictly within the feasible solution space defined by the first component of step 3: all candidate modifications are only accepted if they preserve the daily total and the sub-daily maxima (within a small tolerance). Therefore, the refinement does not contradict or relax the constraints imposed before, but rather complements them by improving a statistical property (autocorrelation) that is not explicitly controlled in the earlier stages.

3. Lines 320ff: The discussion of the NRSP is not relevant here, since it is not a disaggregation approach but a continuous simulation model. This section could be shortened or removed.

We thank the reviewer for this comment and agree that the discussion of the Neyman-Scott Rectangular Pulse (NSRP) model may be a digression from the main focus of this section, as it is a continuous rainfall simulation model rather than a disaggregation approach.

Following the reviewer's suggestion, we have removed the paragraph referring to NSRP (second paragraph in Sect. 4.3) in order to improve the focus and conciseness of the manuscript. The section now concentrates exclusively on the description of the Poisson-Cluster Disaggregation Method (PCDM) as a benchmark disaggregation approach.

4. The model is mainly non-parametric or relying on empirical probability distributions. This makes its regionalisation i.e. the application for unobserved locations almost impossible. This limitation should be discussed in the conclusions.

We thank the reviewer for this important observation. We agree that, in its original formulation, Q-CODA was primarily based on empirical distributions and therefore had limited direct applicability to ungauged or unobserved locations.

We have now explicitly addressed this limitation by extending the method into a semi-parametric regionalised framework, hereafter referred to as Q-CODA-R (Q-CODA-Regionalised). More specifically, in the revised manuscript, in particular in the discussion to address the spatial issue, we will replace the purely empirical cumulative distribution functions (CDFs) used in the comonotonic transformation with seasonal Bernoulli-Gamma parametric CDFs fitted for each station, each season, and each accumulation duration (1, 2, 6, 12 and 24 h). The parameters of these distributions were then spatially regionalised using universal kriging spherical variogram models, with longitude, latitude, and elevation as external predictors. Formally, for each station, season, and accumulation duration τ (with $\tau = 1, 2, 6, 12, 24$ h), precipitation maxima is modelled using a mixed Bernoulli-Gamma distribution. Let $P_\tau^{\max}$ denote the precipitation amount associated with duration τ for a given season. Its probability distribution is represented as

$$Pr(P_\tau^{max} = 0) = p_{0,\tau},$$

$$Pr(P_\tau^{max} > 0) = 1 - p_{0,\tau},$$

and, conditional on $P_\tau^{\max} > 0$,

$$P_\tau^{max} \mid P_\tau^{max} > 0 \sim Gamma(\alpha_\tau, \theta_\tau)$$

Where $p_{0,\tau}$ is the probability mass at zero precipitation, α_τ is the Gamma shape parameter, and θ_τ is the Gamma scale parameter. The parameters are estimated independently for each station and season from the observed daily maxima series for each accumulation duration. Specifically, $p_{0,\tau}$ is estimated as the empirical proportion of zero values,

$$p_{0,\tau} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}\,(P_{\tau,i}^{max} = 0)$$

Where n is the number of available observations and $\mathbf{1}(\cdot)$ is the indicator function. The Gamma parameters $(\alpha_\tau, \theta_\tau)$ are then fitted by maximum likelihood using only the strictly positive observations $\{P_{\tau,i}^{max} : P_{\tau,i}^{max} > 0\}$, with the location parameter fixed at zero.

The resulting CDF for duration τ is therefore

$$F_{P_\tau^{max}}(x) = \begin{cases} 0, & x < 0 \\ p_{0,\tau}, & x = 0 \\ p_{0,\tau} + \left(1 - p_{0,\tau}\right)G(x; \alpha_\tau, \theta_\tau), & x > 0 \end{cases}$$

Where $G(x; \alpha_\tau, \theta_\tau)$ denotes the Gamma CDF with shape α_τ and scale θ_τ.

$$G(x; \alpha_\tau, \theta_\tau) = \frac{1}{\Gamma(\alpha_\tau)} \cdot \frac{1}{\theta_\tau^{\alpha_\tau}} \int_0^x u^{\alpha_\tau - 1}\, e^{-u/\theta_\tau} du$$

$$\Gamma(\alpha_\tau) = \int_0^\infty t^{\alpha_\tau - 1} e^{-t} dt$$

The Gamma CDF parameters are estimated numerically using SciPy (scipy.stats.gamma.fit, with floc = 0) library developed by Virtanen et al. (2020).

After fitting these parameters at station level, the spatial fields of parameters $p_{0,\tau}$, α_τ, and θ_τ are regionalised by universal kriging using PyKrige (doi: 10.5281/zenodo.3738604). For any ungauged target location s, this yields interpolated parameter estimates $\hat{p}_{0,\tau}(s)$, $\hat{\alpha}_\tau(s)$, and $\hat{\theta}_\tau(s)$, from which a location-specific parametric CDF can be reconstructed as

$$\hat{F}_{P_\tau^{max}}(x; s) = \begin{cases} 0, & x < 0 \\ \hat{p}_{0,\tau}(s), & x = 0 \\ \hat{p}_{0,\tau}(s) + \left(1 - \hat{p}_{0,\tau}(s)\right) G\left(x; \hat{\alpha}_\tau(s), \hat{\theta}_\tau(s)\right), & x > 0 \end{cases}$$

This reconstructed seasonal and duration-specific CDF is the one subsequently used within the comonotonic transformation step of Q-CODA-R. In practice, for a given daily precipitation total P_d, the corresponding non-exceedance probability is first obtained from the kriged 24 h distribution,

$$q = \hat{F}_{P_d}(P_d; s) \equiv \hat{F}_{P_{24}}(P_{24}; s)$$

For $\tau < 24$, the target sub-daily maxima are obtained by applying the same non-exceedance probability q to the corresponding mixed Bernoulli-Gamma distribution. If $q < \hat{p}_{0,\tau}(s)$, then $\hat{P}_\tau^{max} = 0$; otherwise, the conditional Gamma quantile is computed as:

$$q_\tau^* = \frac{q - \hat{p}_{0,\tau}(s)}{1 - \hat{p}_{0,\tau}(s)}$$

and this same probability level is then comonotonically transferred to the shorter durations through the inverse CDFs using SciPy (scipy.stats.gamma.ppf):

$$\hat{P}_\tau^{max} = G^{-1}\left(q_\tau^*; \hat{\alpha}_\tau(s), \hat{\theta}_\tau(s)\right), \qquad \tau \in \{1,2,6,12\}$$

It is equivalent to:

$$\hat{P}_\tau^{max} = \hat{F}_{P_\tau^{max}}^{-1}(q; s), \qquad \tau \in \{1,2,6,12\}$$

thus preserving the comonotonic dependence structure between the daily total and the sub-daily maxima while allowing the transformation to be applied at locations without local hourly training data.

To evaluate whether this extension truly enables application at unobserved locations, we designed a strict leave-one-station-out (LOSO) experiment over the 91 AEMET stations used in the benchmark comparison (all with > 90% completeness). Importantly, for each target station, the regionalised method was applied without using any data from that station itself. Instead: (1) the parametric CDFs were reconstructed from kriged Bernoulli-Gamma parameters, and (2) the KNN seed day used to reconstruct the intradaily temporal structure was taken from the geographically nearest station, rather than from the target station. This second component is precisely what makes Q-CODA-R a semi-parametric, rather than a fully parametric, regionalised framework, since the seed-day still relies on an empirical K-nearest-neighbour analogue rather than on a fully parameterised stochastic model.

To improve the spatial support of the regionalisation step, the kriging models were fitted using a larger set of stations with at least 35% completeness over 1996-2024 (approximately ≥ 10 years of hourly observations). This broader network was used only to estimate the spatial structure of the distribution parameters, while the actual performance evaluation remained restricted to the independent set of 91 high-quality evaluation stations (>90% completeness). See Images 1 (map) and 2 (note that Image 2 is a stripe plot analogous to the one presented in the response to Referee #4, but instead of showing stations with more than 90% data completeness, it displays stations with more than 35% completeness). Importantly, the inclusion of stations with lower completeness in the regionalisation stage should be regarded as an additional challenge for the regionalised semi-parametric framework, since the fitted Bernoulli-Gamma parameters are estimated from less complete and therefore potentially noisier hourly records. Consequently, the reported performance of the semi-parametric Q-CODA-R should be interpreted in light of this stricter and more demanding regionalisation setting.

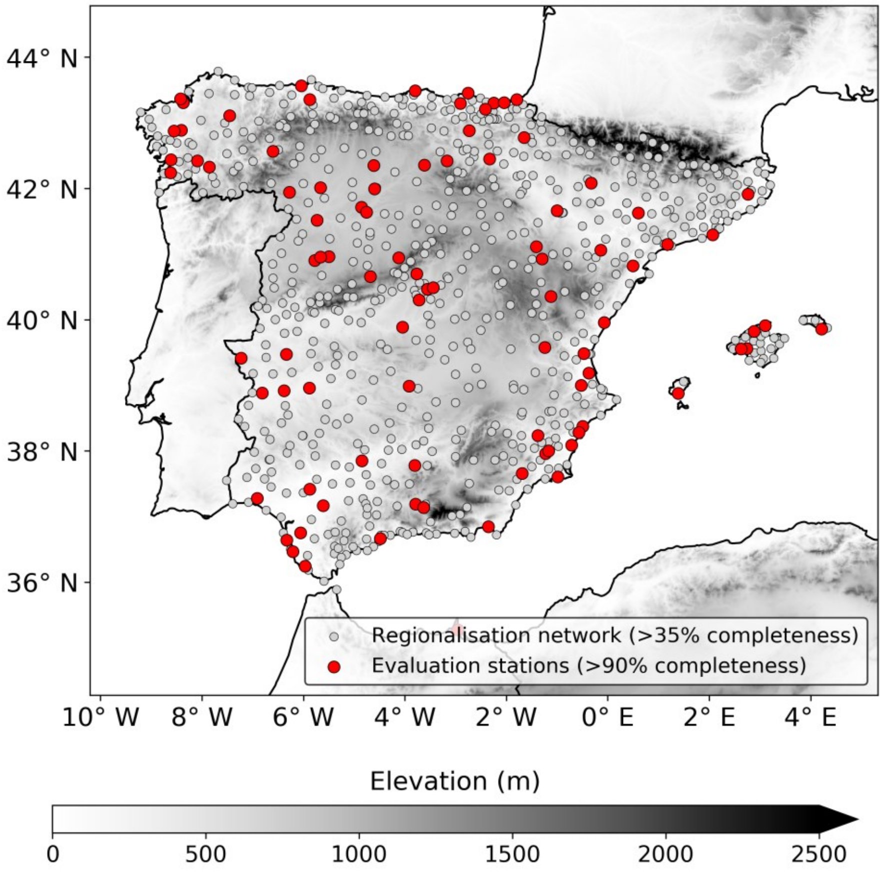

Image 1: Spatial distribution of the stations used in the regionalised semi-parametric Q-CODA-R framework. The 91 evaluation stations (>90% hourly data completeness over 1996-2024) were used exclusively for independent leave-one-station-out validation of

Q-CODA-R, whereas the regionalisation network also includes additional stations with at least 35% completeness over the same period.

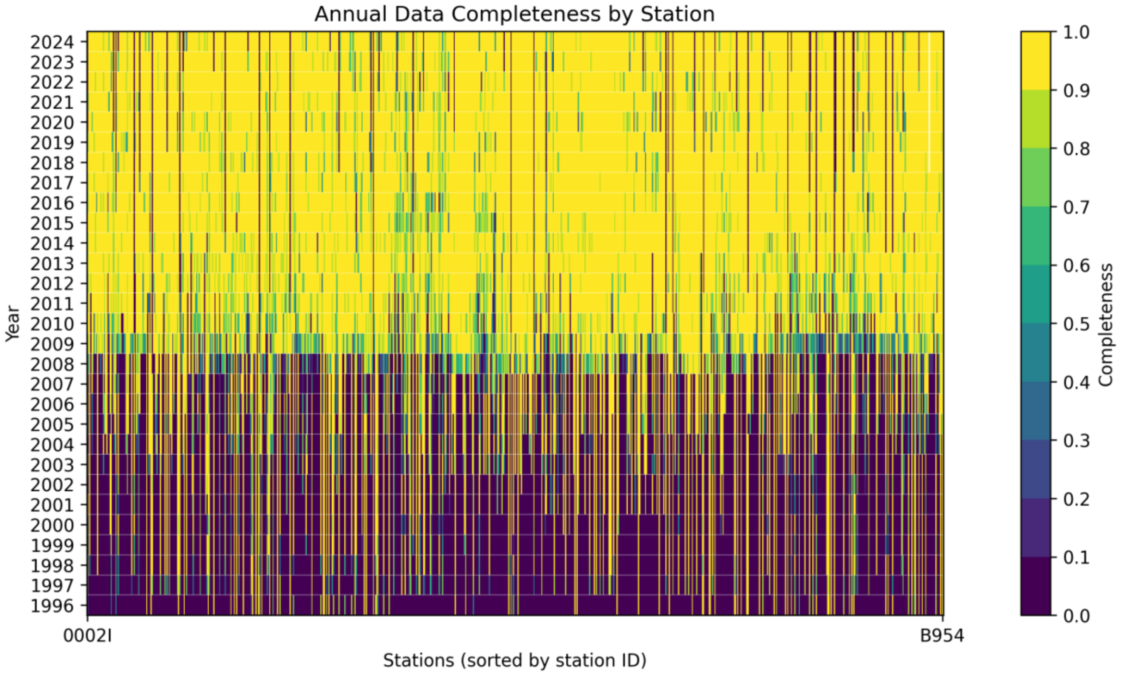

Image 2: Annual hourly data completeness of the stations used in the Q-CODA-R regionalisation framework over 1996-2024. Each column represents a station and each row a year, with colours showing the proportion of available hourly precipitation data. Stations were included in the regionalisation network if their total completeness exceeded 35% over the full period (approximately ≥ 10 years of hourly observations). The stripe shows that the universal kriging of the seasonal Bernoulli-Gamma parameters was supported by a spatially broader but temporally less complete network than the high-quality evaluation subset, making the regionalisation experiment more demanding. It can also be observed that, from 2009 onwards, most of the network exhibits generally high hourly data completeness.

The results (Images 3, 4, and 5) show that the regionalised semi-parametric version, Q-CODA-R, performs only slightly worse than the original station-calibrated Q-CODA, which has direct access to the target station's own historical record. Nevertheless, Q-CODA-R still outperforms several state-of-the-art disaggregation methods (KNN, ANN-K, PCDM, and MMRC) across a broad range of evaluation metrics (previously presented in Figures 5, 6, and 8 of the manuscript), despite the fact that these competing methods were calibrated using observations from the target station itself. In particular, Q-CODA-R shows strong skill in reproducing sub-daily P_{1h}^{max}-related statistics, including MAE, NSE, upper percentiles, and absolute maxima (see Image 3), as well as key properties of the full hourly precipitation series, such as mean event duration and autocorrelation (see Image 4). It also yields lower errors in the reconstruction of IDF curves than several of the benchmark methods (see Image 5).

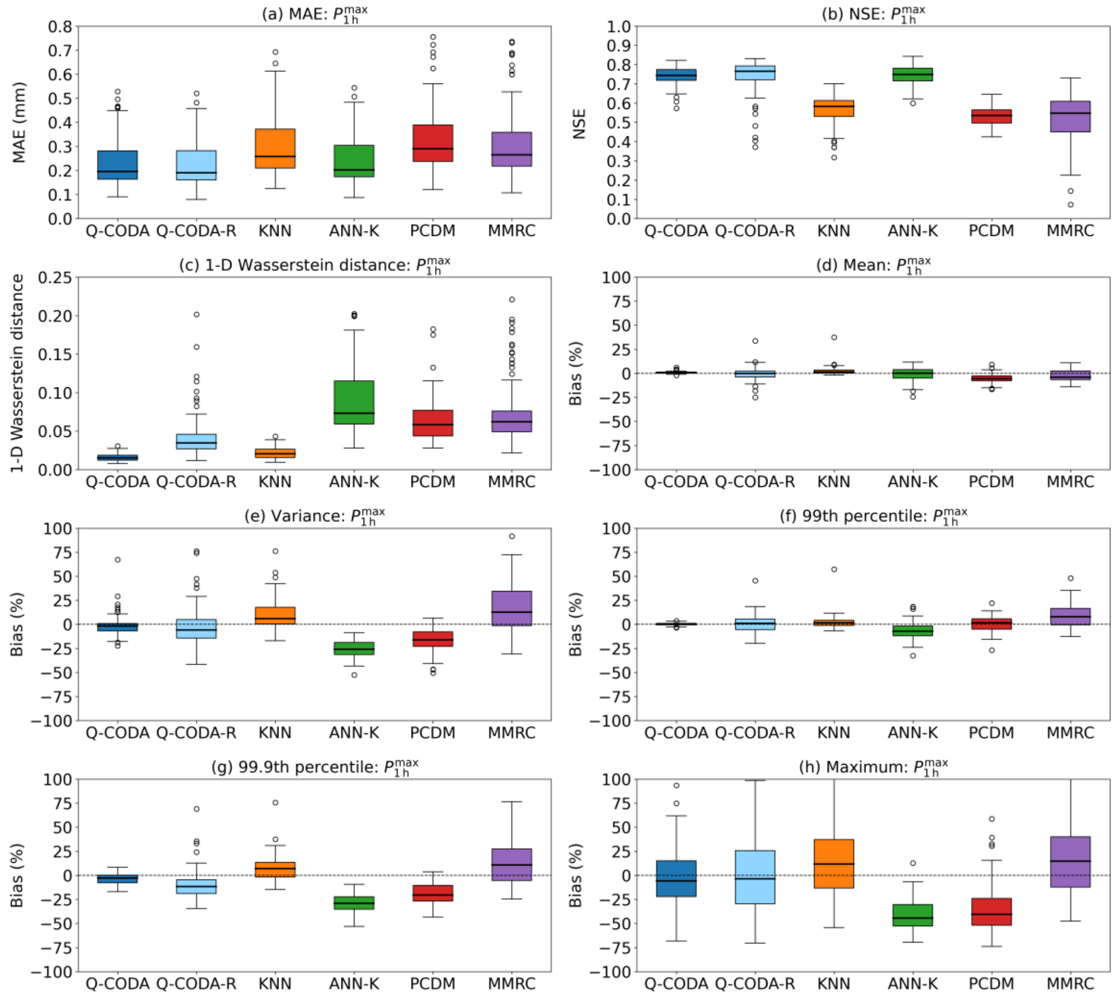

Image 3: Comparative evaluation of disaggregation methods for reconstructing daily maximum 1-hour precipitation (P_{1h}^{max}) across 91 meteorological stations. Each panel shows boxplots summarizing the spatial distribution of key performance metrics: (a) Mean Absolute Error (MAE), (b) Nash-Sutcliffe Efficiency (NSE), (c) 1-D Wasserstein distance, (d) bias in the mean, (e) bias in the variance, (f) bias in the 99th percentile, (g) bias in the 99.9th percentile, and (h) bias in the maximum. All methods except Q-CODA-R are evaluated using 5-fold cross-validation, whereas Q-CODA-R is evaluated in a leave-one-station-out (LOSO) framework, ensuring that no data from the target station are used in training Q-CODA-R.

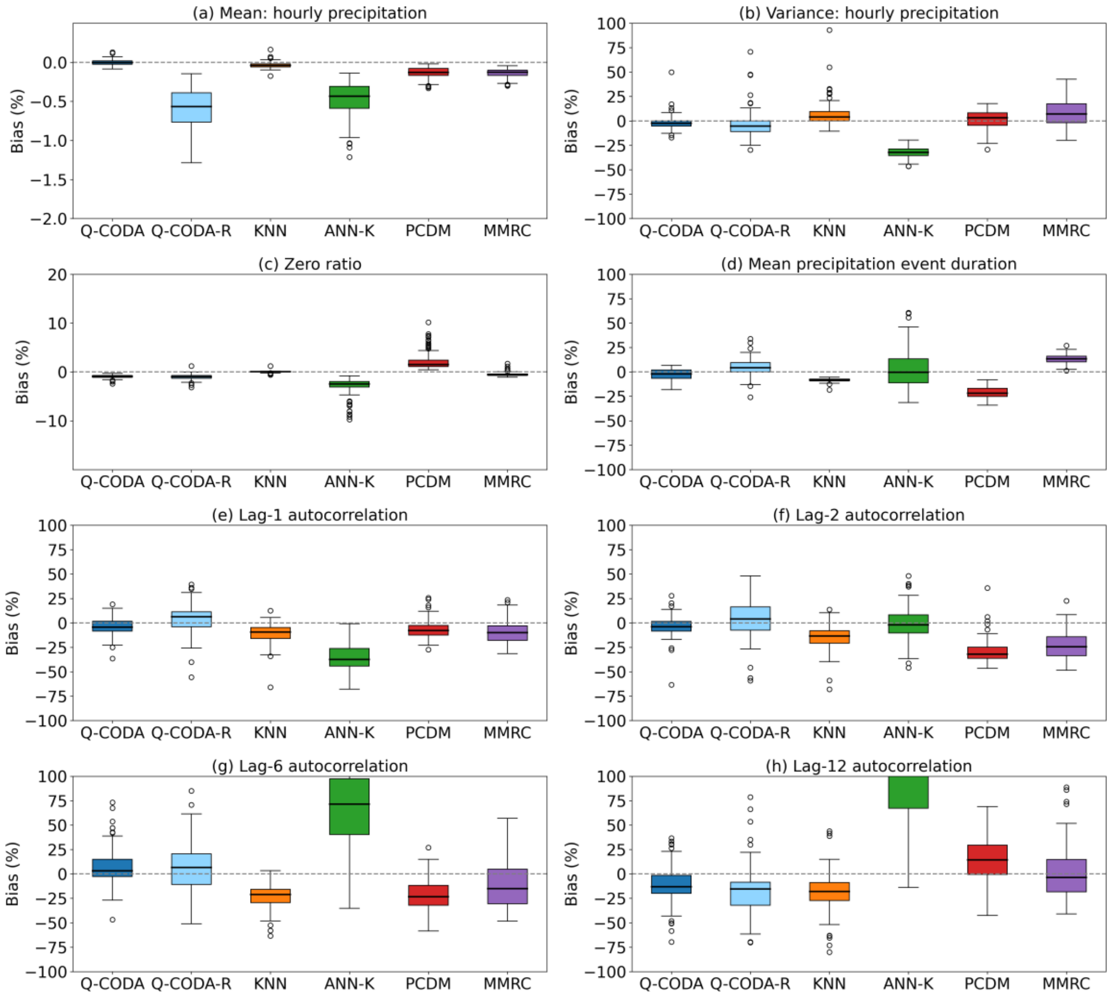

Image 4: Comparative evaluation of disaggregation methods for reconstructing complete hourly precipitation time series across 91 meteorological stations. Each panel presents boxplots summarizing the spatial distribution of key performance metrics: (a) mean hourly precipitation, (b) hourly precipitation variance, (c) bias in zero proportion, (d) bias in mean precipitation event duration, (e) bias in lag-1 autocorrelation, (f) bias in lag-2 autocorrelation, (g) bias in lag-6 autocorrelation, and (h) bias in lag-12 autocorrelation. All methods except Q-CODA-R are evaluated using 5-fold cross-validation, whereas Q-CODA-R is evaluated in a leave-one-station-out (LOSO) framework, ensuring that no data from the target station are used in training Q-CODA-R.

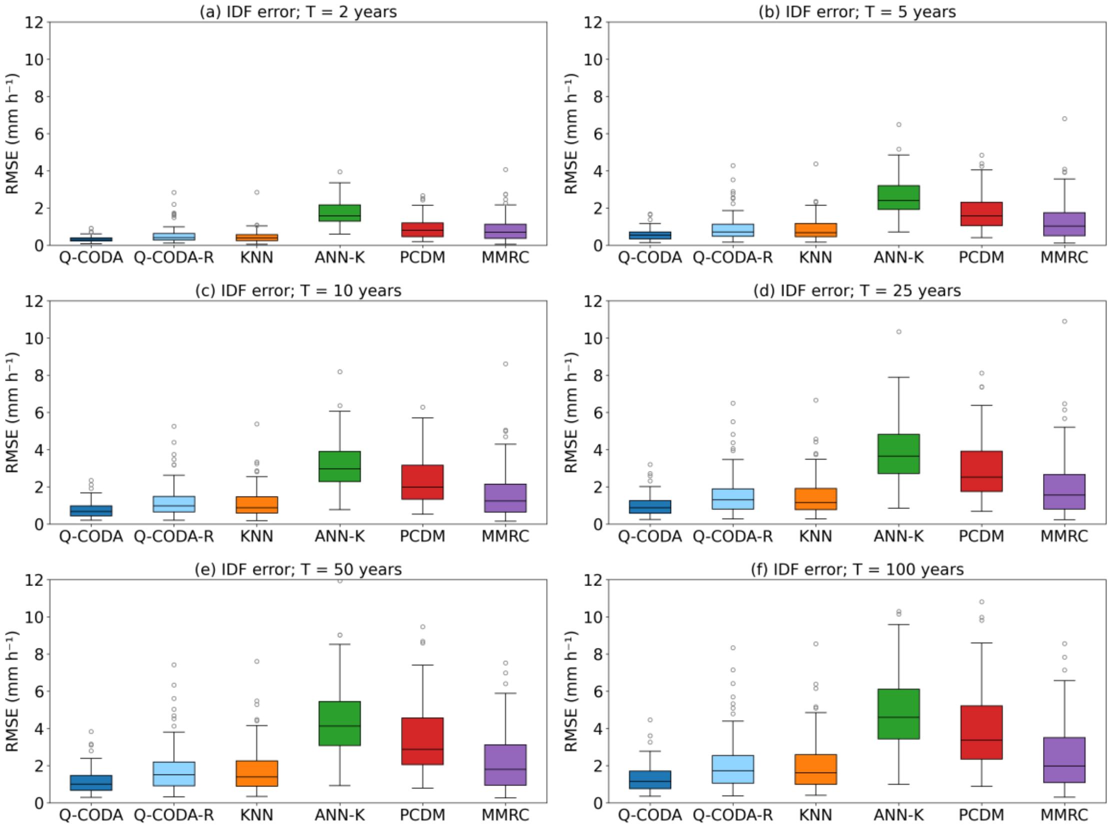

Image 5: Spatial distribution of root mean squared error (RMSE) between observed and simulated IDF (Intensity-Duration-Frequency) curves across 91 meteorological stations, grouped by disaggregation method. All methods except Q-CODA-R are evaluated using 5-fold cross-validation, whereas Q-CODA-R is evaluated in a leave-one-station-out (LOSO) framework, ensuring that no data from the target station are used in training Q-CODA-R.

We therefore agree with the reviewer that regionalisation was a key limitation of the original non-parametric formulation. However, the revised manuscript will demonstrate that this limitation can be substantially overcome by introducing a semi-parametric formulation and spatial interpolation of the parameters, Q-CODA-R. To make this point explicit in the manuscript, we will incorporate this new methodology and its validation results into the revised version, and we will substantially expand the new Discussion section proposed by referees. In particular, we will add a new dedicated subsection specifically addressing the spatial applicability of Q-CODA-R, including both the opportunities and limitations of the proposed semi-parametric framework. The Conclusions section will be revised accordingly. Adittionaly, we will update the versions of pyqcoda on GitHub and PyPI to include the additional option to run the method in semi-parametric mode.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... and SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nat. Methods, 17(3), 261-272, doi:10.1038/s41592-019-0686-2, 2020.

The missing spatial component does currently not allow mesoscale hydrological applications. This should also be discussed as limitation more in detail in the conclusions.

We thank the reviewer for this important comment. We agree that, in its original formulation, the lack of an explicit spatial component substantially limited the applicability of Q-CODA beyond point-scale station disaggregation, and therefore restricted its direct use for mesoscale hydrological applications.

As described in our response to Comment 4, this limitation has now been partially addressed through the introduction of a semi-parametric regionalised version of the method, Q-CODA-R, which enables application at ungauged locations and has been explicitly validated under a leave-one-station-out framework. This constitutes a substantial step forward with respect to the original non-parametric formulation and shows that the method can be meaningfully extended beyond the calibration station itself. However, we agree with the reviewer that this does not yet amount to a full demonstration of mesoscale hydrological applicability in the strict sense. The current regionalised framework allows pointwise spatial extrapolation of sub-daily precipitation characteristics, but a true mesoscale application would additionally require assessing the spatial coherence and synchronicity of the reconstructed sub-daily precipitation fields across multiple locations simultaneously. This includes, for example, the preservation of spatial dependence structures, the consistency of jointly reconstructed extremes, and the reliability of spatially continuous derived products such as gridded IDF surfaces. These aspects are beyond the scope of the present manuscript, which is primarily focused on methodological development and station-based validation. Nevertheless, we agree that this remaining limitation should be made explicit. Accordingly, in the revised manuscript we will expand both the Discussion and the Conclusions to clarify that, although Q-CODA-R represents a substantial advance in terms of spatial transferability, its current validation remains focused on independent pointwise application, rather than on fully spatially distributed hydrometeorological reconstruction. To make this clearer, we will also include a dedicated subsection in the new Discussion section specifically addressing the spatial dimension of the method, including its current regionalisation capability, its remaining limitations, and its implications for future mesoscale hydrological applications.

Finally, we now identify as a natural next step the application of Q-CODA-R to a much denser network of rain gauges for which only daily precipitation records are available. The purpose of this would be to generate sub-daily precipitation series at a much larger number of locations than is currently possible using only stations with observed hourly data. Once these sub-daily series are obtained, IDF curves could be estimated at each of those stations, and their corresponding parameters could then be spatially interpolated to produce spatially continuous IDF products over a regular grid. This is consistent with the general approach followed by Vicente-Serrano et al. (2025), who derived gridded IDF information by interpolating the parameters of pointwise IDF curves; note that in their study, the temporal resolution of the IDF durations was daily rather than hourly. In this sense, the regionalised disaggregation framework that we propose can be understood as a necessary methodological step toward enabling this type of mesoscale hydrological application at sub-daily resolution in regions where hourly precipitation observations are relatively sparse.

Vicente-Serrano, S. M., Beguería, S., Reig, F., Royo, A., Arretxea, M., Gil, M., ... & Nieto, R.: Developing science-informed maps and climate services for extreme rainfall in Spain, Nat. Hazards, 1-22, doi:10.1007/s11069-025-07731-0, 2025.