Response to editor

Re: Manuscript #hess-2025-4469, entitled “Capturing the extremes: a quasi-comonotonicity-based algorithm for disaggregating daily to hourly rainfall”

Dear Editor,

Thank you for the opportunity to revise our manuscript. We appreciate the constructive comments provided by the four reviewers.

We have addressed all comments and revised the manuscript accordingly. A detailed point-by-point response is provided below.

Carlos Correa

ccorreag@aemet.es

Response to Anonymous Referee #1

This study presents the new algorithm Q-CODA for disaggregating precipitation data from daily to sub-daily values. The method leverages the relationship between daily precipitation totals and sub-daily maximum values, which is assumed and shown to be quasi-comonotonic, meaning that high daily precipitation totals are strongly associated with high sub-daily maximum values. The quasi-comonotonicity assumption is incorporated into the methodology via an upper bound copula. The algorithm is applied to 91 observational stations in Spain, and multiple other methodologies are included in the study as benchmarks.

A wide variety of evaluation metrics are applied, shedding light on which aspects of sub-daily precipitation the various methodologies handle well and which they struggle with (e.g., temporal consistency, the fraction of dry days, representation of extreme rainfall events).

The paper is generally well written and, as far as I can tell, methodologically sound. It appears that both Q-CODA and the benchmark methods are implemented according to best practices, which ensures a fair comparison.

I recommend that the paper be accepted with minor revisions, but have some suggestions that could strengthen the existing analysis.

We sincerely thank the reviewer for their careful reading of the manuscript and for their constructive and encouraging feedback. We appreciate the positive assessment of the study, particularly regarding the methodological soundness of Q-CODA, the fair implementation of the benchmark methods, and the comprehensiveness of the evaluation framework. We are also grateful for the recommendation of acceptance with minor revisions.

We have carefully addressed all the reviewer’s comments and believe they have helped improve the clarity and completeness of the manuscript. In the revised version, we have corrected formatting issues, added missing references, and clarified several aspects of the methodology. We have also improved the presentation of the quasi-comonotonicity assumption and its implications, and revised Section 3 to enhance its accessibility.

In addition, we have expanded the discussion of data quality control, evaluation metrics, performance variability across stations, and computational aspects of the algorithm. Following the reviewer's suggestions, we have introduced a dedicated Discussion section, where we further examine the quasi-comonotonicity assumption and discuss the potential applicability of Q-CODA to regions with different precipitation regimes, together with other aspects raised by the reviewers.

Below, we provide a detailed, point-by-point response to the reviewer's specific comments and indicate the corresponding changes made in the manuscript.

Specific Comments

Page 4, lines 120-121

There are some references missing from the list at the back. I found two, Lee et al. (2022) and Chodhury et al. (2025), but there may be more throughout the document. I recommend that the authors check all citations and ensure they are included in the reference list.

We thank the reviewer for pointing this out. The missing references (Lee et al., 2022; Chowdhury et al., 2025) have been added to the reference list in the revised manuscript. We have also conducted a thorough check of all in-text citations to ensure consistency and completeness, and have updated the reference list accordingly, including any additional citations introduced during the revision process.

Lee, J., Kim, U., Kim, S., and Kim, J.: Development and application of a rainfall temporal disaggregation method to project design rainfalls, Water, 14, 1401, https://doi.org/10.3390/w1409140, 2022.

Chowdhury, D. P., Roy, D., and Saha, U.: Incorporation of weather parameters in MMRC-K model for rainfall disaggregation, Stoch. Environ. Res. Risk Assess., 39, 289–308, https://doi.org/10.1007/s00477-024-02863-4, 2025.

Page 4, lines 126-127

Could you please describe the quality control that was applied to the observational data? It doesn't have to be a full and detailed account, but a brief mention of what type of homogeneity testing was applied or a reference to the dataset would be helpful.

We thank the reviewer for this helpful suggestion. We agree that a brief description of the quality control and homogeneity assessment improves transparency.

In the revised manuscript, we have expanded this section to include a concise description of the procedures applied. In addition to the completeness filtering (≥90% data availability), we assessed the homogeneity of the annual precipitation series derived from the hourly records using three standard tests: the Pettitt test, the Standard Normal Homogeneity Test (SNHT), and the Buishand range test (significance level of 0.05). To further verify these cases, we performed a spatial consistency check by comparing detected change points with those of neighboring stations within a 50 km radius. This analysis indicated that 86 out of the 91 AEMET stations showed no evidence of inhomogeneities. The small subset of stations flagged by the tests were geographically clustered in northern Spain and exhibited coincident change points associated with particularly wet years, rather than documented station relocations or instrumental changes, suggesting false positives linked to natural variability. Importantly, none of these stations correspond to the outliers identified in the evaluation metrics of the disaggregation methods, indicating that these potential inhomogeneities do not affect the interpretation of the performance results. This information has been incorporated into the data section (Sect. 2) of the manuscript.

Page 5-6, Section 3

The methodology could benefit from clearer presentation in certain sections. For example, on line 150-151, you write about "the upper bound copula C+, representing perfect positive dependence (comonotonicity)..." which is then presented in Equation (2). After the equation, you write: "This copula implies that X and Y increase together almost surely, which aligns with the quasi-comonotonic behaviour observed between daily precipitation totals and sub-daily maxima." The transition from perfect dependence to quasi-comonotonic behaviour is not entirely clear.

Additionally, while the benchmarks are described clearly, the description of the Q-CODA methodology is more difficult to follow, perhaps due to the level of detail. The flowchart certainly helps, and after reading Section 3 a couple of times it does make sense, but it could be more accessible. Could you add a simplified example showing how the comonotonic upper bound is translated into target maxima and used during the iterative adjustment process? This is key to understanding the method.

We thank the reviewer for this thoughtful and constructive comment. We appreciate the reviewer's careful reading of Section 3 and agree that the transition from the theoretical concept of perfect comonotonicity to the empirically observed quasi-comonotonic behaviour deserves further clarification, and that the methodological description could be made more accessible.

In the revised manuscript, we have clarified this point by explicitly stating that the Fréchet–Hoeffding upper bound copula is not assumed to represent the exact dependence structure in the data, but rather to define an idealized upper envelope corresponding to perfect positive dependence. We now explain more clearly that the empirical relationship between daily totals and sub-daily maxima is only quasi-comonotonic (i.e., very strong but not perfect), and that the upper bound copula is used as a practical way to translate this near-monotonic relationship into conditional target maxima that act as physically consistent constraints during disaggregation. In Section 3 we have added the following paragraph:

"In this study, the Fréchet–Hoeffding upper bound copula is not assumed to represent the exact dependence structure between daily precipitation totals and sub-daily maxima. Rather, it is used to define an idealized limiting case corresponding to perfect positive dependence (comonotonicity), which provides a useful theoretical reference for constructing physically consistent constraints during the disaggregation process. Empirical analyses (see Fig. 2) show that daily totals and sub-daily maxima exhibit a very strong but not perfectly monotonic relationship. We therefore refer to this behaviour as quasi-comonotonic: large daily totals are consistently associated with large sub-daily maxima, although variability remains due to event structure and timing. Within Q-CODA, the upper bound copula is used as a practical mechanism to translate this near-monotonic relationship into conditional target values for sub-daily extremes. These targets do not impose a deterministic mapping but instead provide an upper-envelope constraint that guides the reconstruction toward solutions that are statistically consistent with the observed dependence structure."

We have deleted "*almost*" and changed "*aligns*" to "*aligns approximately*" in the cited sentence:

"This copula implies that X and Y increase together ~~almost~~ surely, which aligns approximately with the quasi-comonotonic behaviour observed between daily precipitation totals and sub-daily maxima."

We have also added the following clarification introducing Sect. 3.4:

"Conceptually, the method operates in percentile space. For each wet day, the target daily precipitation total is first located within the empirical distribution of daily totals derived from the training data. Under the assumption of comonotonic behaviour, the same percentile is then used to infer a corresponding target value for the sub-daily maximum precipitation from its empirical cumulative distribution function. This step is equivalent to applying the Fréchet–Hoeffding upper bound, which links the marginal distributions through a shared rank structure. The resulting target maximum is subsequently used as a constraint within the iterative adjustment procedure to ensure that the reconstructed hourly sequence is consistent with both the daily total and the expected magnitude of sub-daily maxima."

To improve accessibility, we have also expanded the methodological description to better guide the reader through the logic of Q-CODA. In particular, we have added a simplified illustrative example in Appendix A showing how a given daily total is mapped to percentile space, how this percentile is used to derive target sub-daily maxima from the training distribution using the comonotonic assumption, and how these targets are then used within the iterative adjustment process to constrain the hourly sequence while preserving the daily total and the temporal structure. This addition complements the flowchart and helps clarify how the theoretical dependence structure is operationalized in practice. The revised manuscript now includes Appendix A ("Detailed step-by-step example of the Q-CODA process for a single day"), which was explicitly provided in our previous response to the reviewer's comment (available at: https://doi.org/10.5194/egusphere-2025-4469-AC2). Overall, these revisions aim to make the connection between the copula framework, the quasi-comonotonic empirical evidence, and the iterative reconstruction procedure more transparent and easier to follow.

Page 11, line 290

The parenthesis in the cited paper should be after the name when it is in running text, i.e., "The ELU activation function was chosen by Bhattacharyya et al. (2024)..." rather than "(Bhattacharyya et al. 2024)".

This issue recurs later on lines 302, 303, and 344, but I may have missed other instances. While this is a minor formatting issue, I recommend that the authors carefully check all citations to ensure they follow the correct format.

We thank the reviewer for noting this formatting issue. The citation style has been corrected in the indicated instances so that, when authors are mentioned in running text, the year appears in parentheses following the name (e.g., "Bhattacharyya et al. (2024)").

We have also reviewed the entire manuscript to ensure that all in-text citations consistently follow the appropriate formatting conventions.

General Comments

Evaluation Metrics

The evaluation metrics are described in the main text, but references or equations are not provided. While this makes the main text more readable, it would be useful to include an appendix with a fuller description of the metrics, including equations and references (e.g., for the Nash-Sutcliffe Efficiency and Wasserstein distance). This would improve the reproducibility of the study. I understand the code will also be made available, but it would still be good if the paper included all information needed to interpret and recreate the results.

We thank the reviewer for this valuable suggestion. We agree that providing a more complete description of the evaluation metrics improves reproducibility.

In response, we have added a new appendix (Appendix B) to the revised manuscript that includes a detailed description of all evaluation metrics used in the study, together with their mathematical formulations and key references (e.g., for the Nash–Sutcliffe Efficiency and the Wasserstein distance). The appendix clearly defines how each metric is computed for both daily maximum 1-hour precipitation and the hourly time series, including error measures, distribution-based metrics, bias indicators, intermittency diagnostics, event-duration statistics, and autocorrelation-based measures. While the main text retains a concise overview to preserve readability, this additional material ensures that all necessary information is now contained within the paper itself, allowing readers to fully interpret and independently reproduce the evaluation framework without relying solely on the accompanying code. The revised manuscript now includes Appendix B ("Evaluation metrics"), which was explicitly provided in our previous response to the reviewer's comment (available at: https://doi.org/10.5194/egusphere-2025-4469-AC2).

Assumption of Quasi-Comonotonicity

The assumption of quasi-comonotonicity underpins the methodology, and the authors argue that the heterogeneity of precipitation regimes in Spain makes the results broadly applicable. While this argument is pretty convincing, I can think of precipitation regimes where the quasi-comonotonicity assumption might not hold as well. For example:

- In regions like northwestern Europe or the Pacific Northwest of the US and Canada, storm tracks bring sustained moderate precipitation over long periods of time, which may weaken the relationship between daily totals and sub-daily maxima.
- In monsoon regions, short-lived, strong convective systems could contribute to very high sub-daily maxima without corresponding high daily totals. This could weaken or break the dependence between daily totals and sub-daily maxima.

Could the authors expand on the discussion of quasi-comonotonicity and the applicability of Q-CODA in other regions? This might fit best in the Discussion section.

How sensitive is Q-CODA to deviations from the quasi-comonotonicity assumption? For example, could the authors test the algorithm on artificial datasets where the correlation between daily totals and sub-daily maxima is systematically reduced? This would help determine how far the assumption can deviate before the method's performance degrades.

We thank the reviewer for this thoughtful and constructive comment. We agree that empirical quasi-comonotonicity is a key assumption underlying Q-CODA, and that its validity and implications may vary across different precipitation regimes.

In the revised manuscript, we have introduced a new Sect. 7 ("Discussion") to explicitly address these issues in a structured way. In Sect. 7.2 ("7.2 Applicability to other climatic regimes"), we extend the evaluation to independent observational stations from the NCEI/NOAA network, focusing on contrasting climatic regimes such as the Pacific Northwest and Florida, in order to explore the applicability of the method under markedly different precipitation characteristics. In addition, Sect. 7.3 ("7.3 Sensitivity to deviations from quasi-comonotonicity") presents a sensitivity analysis using artifitial datasets to assess the performance of Q-CODA under controlled deviations from the quasi-comonotonicity assumption.

Further methodological details and the initial justification of these analyses are provided in our previous detailed response to the reviewer's comment (available at: https://doi.org/10.5194/egusphere-2025-4469-AC2). The only modification introduced in the revised manuscript with respect to that response is the increase in the number of artifitial series from 6 to 15 in the sensitivity analysis.

To improve readability and avoid overloading the main text with figures, we have also added maps showing the location of the NCEI/NOAA stations in a new Appendix E ("Maps of Other Climatic Regimes").

Together, these additions clarify the role of the quasi-comonotonicity assumption, its expected variability across climates, and the sensitivity of the method to departures from it. These analyses provide additional insight into the robustness and climatic generality of the quasi-comonotonicity assumption.

Performance Variability Across Stations

The Q-CODA algorithm performs well across most evaluation metrics, but some outliers appear (in Figure 6b and 6f for example), suggesting that the method struggles a bit at certain stations. This is to be expected, but it would be interesting to investigate these cases further. For example:

- Do these outliers correspond to stations with lower correlation values in Figure 2, potentially indicating weaker quasi-comonotonic dependence?
- Are there locations where all methods perform poorly, potentially pointing to data quality issues or precipitation processes that are difficult to capture with disaggregation methods?

We thank the reviewer for this insightful comment. We agree that the presence of some outliers in Figs. 6b and 6f warrants further investigation, particularly in relation to station-specific characteristics and the strength of quasi-comonotonic dependence.

In the revised manuscript, we have added a new subsection in the Discussion (Sect. 7.1, “Performance variability across stations”) to address this issue. This section clarifies that Q-CODA generally performs consistently across AEMET stations. The outliers are not associated with systematic station-specific data quality problems, nor do they consistently correspond to weaker quasi-comonotonic dependence.

To avoid overloading the Discussion section with additional figures and extended diagnostics, and to maintain a clearer separation between the main discussion and supporting analyses, the detailed diagnostic study has been moved to Appendix D (“Performance variability across stations”). This decision also reflects our view that the analyses of alternative climatic regimes (new Sect. 7.2), sensitivity to quasi-comonotonicity (new Sect. 7.3), and spatial considerations (this latter raised by Reviewer #3, and added in new Sect. 7.4) are more central to the revised manuscript once station-specific issues and data-quality concerns have been ruled out. Appendix D was explicitly developed in response to this comment (available at: https://doi.org/10.5194/egusphere-2025-4469-AC2).

Computational Feasibility

The paper does not discuss the computational costs of Q-CODA relative to the benchmarks. Given the iterative nature of the algorithm, this could be an important consideration, particularly for larger datasets or real-time applications. Please add a discussion of the computational efficiency of Q-CODA as compared to the benchmark methods. Even better, you could perform a comparative test of the computational costs. I am thinking of runtime, where parallelization might be an option for some of these methods, but also computational costs and energy usage (which can be checked with tools like CodeCarbon in python).

We thank the reviewer for highlighting this important aspect. We agree that the computational efficiency of Q-CODA is a relevant consideration, particularly given its iterative nature and potential application to large datasets. In the revised manuscript, we have added a dedicated assessment of computational cost comparing Q-CODA with the state-of-the-art disaggregation methods. Following the reviewer’s suggestion, we used the CodeCarbon tool to quantify runtime, energy consumption, and estimated CO_2 emissions under identical computational conditions (same reference station and hardware setup). This provides a consistent basis for comparison across methods. The results have been included at the end of Sect. 6 (“Results”) in the revised manuscript, where we briefly discuss the relative computational performance of Q-CODA with respect to the state-of-the-art approaches. This addition confirms that, while Q-CODA is computationally more demanding than some simpler methods, its runtime remains fully feasible for offline and large-scale applications.

Response to Anonymous Referee #2

The manuscript presents Q-CODA, a novel copula-based framework for temporal disaggregation of daily precipitation into hourly time scales. The primary methodological contribution lies in exploiting the quasi-comonotonic dependence between daily total and sub-daily maxima though the Frechet-Hoeffding upper bound copula, combined with KNN based seeding and an iterative adjustment scheme. The approach is supported by a comprehensive multi-metric validation framework that takes into account accuracy, distributional fidelity, structural realism and hydrological relevance.

The manuscript is well presented, methodologically sophisticated and addresses a significant hydrological problem. The key strength of the paper lies in the explicit use of quasi-comonotonicity as a structural constraint, which poses as an advancement over the prevalent temporal disaggregation procedures. The comparative evaluation against multiple temporal disaggregation models under a 5-fold cross-validation scheme enhances the robustness and credibility of the results. The benchmark disaggregation methodologies have been implemented carefully and appear to be applied in a technically sound manner. The inclusion of precipitation dataset from 91 meteorological stations lying across diverse climatic regimes of Spain is commendable. The evaluation framework appropriately considers both upper-tail behaviour and key statistical and temporal properties of the hourly series, thereby ensuring that model performance is assessed rigorously.

Overall, the manuscript appears to have the potential to make an impactful contribution to the field of stochastic rainfall disaggregation.

My primary concerns largely overlap with those raised by Referee #1, specifically with respect to 1) the robustness and general applicability of the quasi-comonotonicity assumption 2) the need for improved clarity in the explanation of the proposed methodological framework 3) inclusion of a more detailed description of the evaluation metrics employed in the present study 4) the issue of missing references

Based on the author's response, in my assessment, these four concerns appear to have been adequately addressed. Therefore, I recommend acceptance of the manuscript, provided that the revisions described are fully reflected in the final version of the manuscript.

We sincerely thank Referee #2 for the careful reading of our manuscript and for the positive and encouraging assessment of our work. We are grateful for the recognition of the methodological contribution of Q-CODA, the robustness of the validation framework, and the potential relevance of the study for stochastic rainfall disaggregation. We also appreciate the referee's synthesis of the main concerns, which largely coincide with those raised by Referee #1: (1) the robustness and general applicability of the quasi-comonotonicity assumption, (2) the clarity of the methodological description, (3) the completeness of the evaluation metrics, and (4) the issue of missing references. The convergence of both referees on these aspects was particularly helpful in guiding the revision process.

In response, and as detailed in our reply to Referee #1 (https://doi.org/10.5194/egusphere-2025-4469-AC2), we have thoroughly revised the manuscript to address all these points. Specifically, we have improved the clarity of the methodological framework in Sect. 3, including the addition of a step-by-step illustrative example of the Q-CODA procedure in Appendix A ("Detailed step-by-step example of the Q-CODA process for a single day."). We have also included a comprehensive description of evaluation metrics in Appendix B ("Evaluation metrics"), completed and verified all references, and substantially strengthened the discussion on the quasi-comonotonicity assumption in new Sect. 7 (“Discussion”), including analyses in additional climatic regions (Sect. 7.2) and sensitivity experiments (Sect. 7.3).

We confirm that all revisions have been incorporated into the updated manuscript. We thank the referee again for the positive evaluation and constructive feedback, which have helped improve the clarity and robustness of the study.

Response to Anonymous Referee #3

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 the positive and constructive assessment of our manuscript. We greatly appreciate the recognition of the novelty of the proposed Q-CODA approach, its strong performance relative to the selected benchmark methods, and the clarity of the presentation. We also appreciate the reviewer’s comment regarding the conciseness of the original manuscript. While we have aimed to preserve this characteristic throughout the revision process, the feedback from four independent reviewers has led to the inclusion of several targeted improvements and clarifications. In particular, new material such as the new dedicated Discussion (Sect. 7) has been added to address key methodological aspects raised during the review process. We believe that these additions enhance the completeness and interpretability of the study without compromising its overall focus and structure.

We have carefully considered all the reviewer’s comments 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 helpful suggestion. We agree that a more explicit description of the KNN implementation improves clarity and reproducibility.

In the revised manuscript, Sect. 3.3 (“Initial Pattern Seeding via KNN”) has been expanded to include a formal mathematical description of the method. Specifically, we now detail the distance metric used to identify nearest neighbours (based on daily precipitation totals), the selection of the k closest candidates, and the random sampling procedure within this set. We also clarify how the selected hourly pattern is rescaled to match the target daily total. The full formulation of this procedure was originally provided in our detailed response to the reviewer (available at: https://doi.org/10.5194/egusphere-2025-4469-AC4) and has now been incorporated into the manuscript.

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 the original non-parametric formulation of Q-CODA limits its direct applicability to ungauged or unobserved locations.

In response, the revised manuscript introduces a semi-parametric regionalised extension of the method, referred to as Q-CODA-R (Q-CODA-Regionalised). This framework replaces the purely empirical distributions used in the comonotonic transformation with parametric (Bernoulli–Gamma) distributions, whose parameters are spatially interpolated using universal kriging. This allows the method to be applied at locations without local hourly observations while preserving the dependence structure between daily totals and sub-daily maxima. The spatial applicability of this approach is evaluated through a leave-one-station-out (LOSO) experiment over the 91 AEMET stations, ensuring that no information from the target station is used in the reconstruction. The results show that Q-CODA-R performs only slightly below the station-calibrated version, while still outperforming several benchmark station-calibrated methods across a range of metrics.

These developments are presented in a new subsection of the Discussion (Sect. 7.4, "Spatial component"), where both the capabilities and limitations of the regionalised framework are discussed. The Conclusions section has also been revised accordingly. To avoid overloading the Discussion section with a large number of additional figures, we have included complementary results in a new Appendix F ("Regionalization network map and expanded evaluation boxplots for the spatial component analysis"), which provides further visual support and metrics to the analysis presented in Sect. 7.4.

Further methodological details and the full formulation of the regionalised approach are provided in our previous detailed response to the reviewer (available at: https://doi.org/10.5194/egusphere-2025-4469-AC4).

5. 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 limited the applicability of Q-CODA to point-scale disaggregation and therefore constrained its direct use in mesoscale hydrological applications.

As described in our response to the previous comment, 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 using a leave-one-station-out framework. This represents a substantial step forward in terms of spatial transferability beyond the calibration station. However, we fully agree with the reviewer that this does not yet constitute a complete mesoscale hydrological framework. While Q-CODA-R allows pointwise spatial extrapolation of sub-daily precipitation characteristics, it does not explicitly model or enforce spatial coherence between locations. In particular, the current formulation does not address key aspects such as the preservation of spatial dependence structures (e.g. the synchronisation of extreme events across neighbouring sites). This limitation is now explicitly discussed in the revised manuscript. In particular, Sect. 7.4 ("Spatial component") provides a structured discussion of both the capabilities and the current limitations of the regionalised framework, and the Conclusions section has been substantially revised to clearly state that Q-CODA-R operates at the level of independent pointwise disaggregations rather than fully spatially consistent reconstructions. At the same time, the revised Conclusions also outline a practical pathway towards mesoscale applications. In particular, Q-CODA-R enables the generation of sub-daily series at a much larger number of locations using dense rain-gauge networks with only daily observations. These reconstructed series can then be used to derive local IDF curves, whose parameters may subsequently be spatially interpolated to obtain gridded products, following approaches such as Vicente-Serrano et al., (2025); note that in their study, the temporal resolution of the IDF durations was daily rather than hourly. In this sense, the proposed framework can be understood as an intermediate step towards enabling mesoscale hydrological applications at sub-daily resolution in data-scarce regions.

Response to Anonymous Referee #4

"Capturing the Extremes: A Quasi-Comonotonicity-Based Algorithm for Disaggregating Daily to Hourly Rainfall" is an interesting and novel study, which I recommend for publication after revisions to the manuscript.

We thank the reviewer for the positive assessment of our work and for recommending the manuscript for publication after revisions. We appreciate the recognition of the novelty and relevance of the proposed approach, and we address the specific comments below.

Please see my feedback and comments below:

a) The abstract should include quantifiable values of the main results.

We thank the reviewer for this helpful suggestion. We agree that including quantitative performance indicators in the abstract improves clarity and allows readers to better assess the magnitude of the methodological improvements.

In the revised manuscript, we have updated the abstract to incorporate representative quantitative results summarising the performance of Q-CODA relative to the benchmark methods. Specifically, we now report median values across stations for key evaluation metrics, including the 1-D Wasserstein distance, bias in the 99.9th percentile, event duration bias, lag-1 autocorrelation bias, and RMSE of IDF curves. These additions provide a concise yet informative overview of the main results. The revised abstract also reflects the additional analyses introduced in the new Discussion section (Sect. 7), including the evaluation across contrasting climatic regimes and the semi-parametric regionalised extension (Q-CODA-R), ensuring consistency between the summary and the full scope of the revised manuscript.

b) Figure 5 – Remove $P^{max}{}_{1h}$ from subplot titleand rescale (h). Recommended to add median boxplots.

We thank the reviewer for the suggestion. We have remove $P^{max}{}_{1h}$ from the subplot titles and rescale subplot (h) accordingly in Fig. 5. Regarding the inclusion of a median plot, we believe this could be unnecessary in this context. The current boxplots already capture the mean, variance, and extreme values, providing a comprehensive summary of the data distribution. In addition, the Q-Q plots requested in comment d), that has been added (Fig. 9) in new Sect. 7.2 of the revised manuscript, complement the boxplots by directly illustrating the alignment between simulated and observed values across the probability distribution. In our opinion, including a median panel would therefore be redundant, while the current combination of boxplots and Q-Q plots conveys both central tendency and variability in the disaggregated rainfall.

c) Figure 7 – Subplot titles should include only "T = XX years."

We thank the reviewer for this suggestion. We have updated the subplot titles in Figure 7 so that they include only "T = XX years".

d) Include a Q–Q plot comparing simulated and observed rainfall datasets for both daily hourly maxima and full hourly scales for selected stations. Selection could be based on elevation or regional grouping.

We thank the reviewer for this helpful suggestion. We agree that Q–Q plots provide a valuable complement to summary metrics by enabling a direct visual assessment of the agreement between observed and simulated rainfall distributions.

In response, we generated Q–Q plots for groups of stations based on a regional classification (Mediterranean and southern, inland, and northern/northwestern Spain), and evaluated both daily maximum 1-hour precipitation and the full hourly rainfall distribution for all disaggregation methods. After analysing these results, we found that both sets of Q–Q plots showed very similar comparative patterns across methods. To avoid redundancy and maintain a concise presentation, we therefore selected the Q–Q plots corresponding to daily maximum 1-hour precipitation, which are more directly aligned with the main focus of the study on sub-daily extremes. These Q–Q plots have now been incorporated at the beginning of Sect. 7.2 ("Applicability to other climatic regimes") in the revised manuscript. Their placement in this section allows us to first compare performance across different geographic regions within Spain, and then extend the analysis to contrasting climatic regimes (Pacific Northwest and Florida, USA), providing a coherent and structured evaluation of the method's behaviour under varying hydroclimatic conditions. The full set of Q–Q plots and additional details were previously provided in our previous detailed response to the reviewer's comments (available at: https://doi.org/10.5194/egusphere-2025-4469-AC5).

e) It would be much better if the author(s) could add a separate section for the discussion.

We thank the reviewer for this valuable suggestion. We agree that including a dedicated Discussion section improves the clarity, structure, and interpretability of the manuscript.

In the revised manuscript, we have introduced a new Sect. 7 (“Discussion”), which provides a structured synthesis of the key methodological and interpretative aspects raised during the review process. This section is organised into four subsections: (7.1) Performance variability across stations, (7.2) Applicability to other climatic regimes, (7.3) Sensitivity to deviations from quasi-comonotonicity, and (7.4) Spatial component. Together, these additions consolidate and expand upon the points raised by all referees, as outlined in our previous response to Referee #4 (available at: https://doi.org/10.5194/egusphere-2025-4469-AC5), providing a more comprehensive discussion of the general applicability, limitations, and potential extensions of the proposed Q-CODA framework.

More Comments:

1. L115 – How consistent is the data on a yearly basis? Although a station may have >90% overall data availability, some years could still have significant missing data.

We thank the reviewer for raising this important point regarding annual data consistency. We performed a detailed assessment of yearly data completeness across all 91 AEMET stations, including a stripe plot visualisation. The results show that, while some stations present missing data in specific years, particularly in the early part of the record (e.g. 1996, corresponding to the first year of the first cross-validation fold), the overall temporal coverage across all folds remains sufficient to reliably train and evaluate the disaggregation methods. To make this information accessible to the reader, we have included a reference in the revised manuscript (line 105 in Sect. 2) indicating that annual data completeness can be consulted in Fig. C1 (Appendix C), where the full stripe plot is provided. This is the same stripe plot that was previously presented (available at: https://doi.org/10.5194/egusphere-2025-4469-AC5).

2. General – Inconsistent use of “section” and “sect.”

We thank the reviewer for pointing out the inconsistent use of “section” and “Sect.” throughout the manuscript. We have carefully revised the entire text to ensure a uniform and consistent usage of section references in accordance with the journal’s formatting guidelines. In particular, we have aligned the notation with the conventions specified by Hydrology and Earth System Sciences (HESS) submission standards (https://www.hydrology-and-earth-system-sciences.net/submission.html), ensuring consistency throughout the manuscript.

3. L91–95 – Remove or move to another section. The same applies to L96–104 and L120–125.

We thank the reviewer for this suggestion. In response, we have removed the paragraphs corresponding to L91–95 and L96–104, as these were not essential for the flow of the introduction and were better integrated elsewhere in the manuscript. Regarding L120–125, we have decided to retain it in Sect. 2 (“Data and Study Area”), as it provides an important contextualisation of the study design by highlighting the comparatively large number of stations used in this work relative to many existing studies. We consider this relevant at this stage of the manuscript, as it helps to frame the representativeness and spatial coverage of the dataset before the methodological description is introduced.

4. 158 – Specify which previous methods are being referred to.

We thank the reviewer for this comment. In the revised manuscript, we have clarified this point by explicitly specifying the methods being referred to. The sentence has been reformulated to explicitly contrast Q-CODA with commonly used rainfall disaggregation approaches, including KNN, ANN-K, MMRC, Poisson-cluster models, and other copula-based frameworks. These methods typically represent the dependence between daily totals and sub-daily maxima in a purely statistical manner, without imposing an explicit structural constraint. In contrast, our approach exploits their quasi-comonotonic relationship as a key structural feature, as now clearly stated in the text.

5. L206 – Why was a 5-fold used? Also, divide the sentence for clarity.

We thank the reviewer for this comment. In the revised manuscript, we have clarified both the rationale for using a 5-fold cross-validation scheme and improved the readability of the corresponding sentence by restructuring it into a more explicit description. Specifically, a 5-fold cross-validation was selected as it provides a widely used compromise between bias (arising from reduced training data within each fold) and variance (sensitivity to the particular partitioning of the dataset), and is commonly adopted in scientific studies (Gutiérrez et al., 2019; Hernanz et al., 2022; Gong et al., 2026). This approach also ensures that each fold contains sufficient data for both training and validation while maintaining reasonable computational cost. To improve clarity, we have revised the first paragraph of Sect. 5 to explicitly incorporate this justification, together with the full description of the cross-validation setup and the definition of the temporal folds used in the experiments. The revised text also includes the corresponding references supporting this choice of cross-validation strategy.

6. L243 – Explain why 0.04 mm was selected.

We thank the reviewer for this comment. The tolerance value of 0.04 mm was selected to ensure that any adjustments remain below the rounding precision of one decimal place used in the precipitation dataset. This guarantees that the final daily totals are not affected while preserving internal consistency in the reconstructed hourly series.

This clarification has been added directly to the corresponding sentence in the revised manuscript.

7. L293 – Check the use of the word “by.”

We thank the reviewer for pointing this out. The sentence has been revised in the manuscript, both correcting the issue identified (i.e., the incorrect use of “by”) and improving the overall wording for clarity.

8. L331 – Rewrite this sentence for clarity.

We thank the reviewer for this comment. In line with Comment 3 from Referee #3 (https://doi.org/10.5194/egusphere-2025-4469-RC3), the paragraph containing previous line 331 has been removed, as it referred to the Neyman-Scott Rectangular Pulse (NSRP) model, which is a continuous rainfall simulation approach rather than a disaggregation method. This revision improves the clarity and focus of the section by ensuring that only relevant disaggregation methods are discussed.

9. 342 – Use the full form of MELODIST in the previous sentence.

The full form of MELODIST has been added in the preceding sentence in the revised manuscript, as suggested.

10. L375 – Clarify whether this analysis was done for each station individually or for all stations together.

The analysis was conducted individually for each station, and the boxplots shown in the manuscript summarise the distribution of results across all 91 stations to provide an aggregated view of the station-wise variability. This clarification has been added in the revised manuscript at the end of the first paragraph of Sect. 5.

11. L408, L416 – Explain how these are benchmark methods; if not justified, consider removing the term.

We thank the reviewer for this comment. The term “benchmark disaggregation methods” refers to a set of well-established and widely used approaches in the rainfall disaggregation literature, namely KNN, ANN-K, MMRC, and Poisson-cluster models. These methods were selected because they represent a representative set of state-of-the-art alternatives against which the performance of Q-CODA can be meaningfully assessed. This clarification has been added in the revised manuscript in the first paragraph of Sect. 4. However, if the editor or reviewers consider that the use of the term “benchmark” is not appropriate in this context, we would be happy to remove it or replace it accordingly in the final version.

12. L413–415 – Rewrite to avoid redundancy with the figure reference. Check the term “daily maximum”- it appears to refer to a daily maximum 1-hour value.

We thank the reviewer for this observation. The term “daily” in the original sentence was incorrect, as Fig. 5h refers to the maximum of the sub-daily 1-hour precipitation maxima rather than a daily maximum. This has been deleted in the revised manuscript and the sentence has been rephrased to reduce redundancy with the figure reference and improve overall clarity, as suggested by the reviewer.

13. L417 – Replace with “high-temporal-resolution.”

We thank the reviewer for this suggestion. The term has been replaced with “high-temporal-resolution” in the revised manuscript, as indicated.

14. L446, L473 – The use of “robust” needs more justification.

We thank the reviewer for this comment. In response, the term “robust” has been removed from lines L446 and L473 and replaced with more precise wording, as reflected in the revised manuscript. The updated sentences will read:

"Overall, the results in Fig. 5 demonstrate the superior ability of Q-CODA to reproduce both the magnitude and statistical characteristics of peak sub-daily precipitation. Its consistent performance across multiple metrics and spatial locations highlights its value as a disaggregation tool that reliably reproduces observed rainfall features."

"Similarly, the metrics in Fig. 6 show that Q-CODA achieves the best or near-best performance across all statistical and temporal features of the hourly rainfall series, providing a consistent advantage over state-of-the-art methods."

These revisions clarify the intended meaning while avoiding unsupported terms. In addition, we have reviewed the manuscript to avoid an overuse of the term "robust" more generally, ensuring that descriptions remain grounded in the quantitative results presented.

15. L450–453, L465 – Rewrite these lines as suggested above for clarity and conciseness.

We thank the reviewer for this comment. Lines 450–453 and 465 in the original manuscript have been changed in the revised manuscript in order to improve clarity and conciseness, as suggested. In particular, the description of the evaluated metrics (including bias in the temporal mean, temporal variance, zero proportion, mean event duration, and autocorrelations at multiple lags) has been streamlined, and the corresponding interpretation of Q-CODA's performance has been reformulated to avoid redundancy with the figure captions. These modifications have been incorporated into the revised manuscript and contribute to a more concise and clearer presentation of the results shown in Figs. 5 and 6.