Response to Reviewer 2

Comment 1: The study investigates two methodologies (i.e, Interpolation and Machine Learning (ML)) for regionalizing Intensity-Duration-Frequency (IDF) curves across various time scales. The findings indicate that Machine Learning outperforms Interpolation, both in terms of accuracy and data requirements. Given its methodological rigour and practical relevance, the study holds significant importance for publication in HESS. However, the following questions could enhance the depth of the manuscript.

Response: We sincerely thank you for your insightful feedback and for recognizing the value of our work. We have addressed the points you raised in the detailed point-by-point response below and hope that our explanations resolve your queries.

Comment 2: With the possibility of extreme events predicted to increase in the future, why hasn't the study considered duration less than 1 hour (i.e. 30 min)?

Response: We agree that with the projected increase in extreme events, and rainfall intensities at sub-hourly durations is of critical importance. However, given that the station-derived IDF curves were used as foundational ground-truth data for both training and validating our machine learning models and for performing the spatial interpolations, any consideration of durations less than one hour would first require the development of station IDF curves at sub-hourly scales. To produce such curves fundamentally requires precipitation data recorded at a sub-hourly resolution. Throughout our study area of mainland China, such high-resolution station precipitation data is not widely available. Given the high demand for a large and spatially extensive sample to effectively train robust machine learning models and avoid overfitting, we decided to set the minimum duration for the IDF curves in our study to one hour. This decision allowed us to ensure the accuracy of the observed station IDF curves we used, and to guarantee enough stations to build a representative model at a national scale. Nevertheless, we acknowledge the need for shorter durations and see it as a valuable direction for future research. An investigation into these shorter durations would be a logical and important extension of this work if sub-hourly precipitation data become more readily accessible in the future. We believe our study provides a methodological framework for such future analyses.

Comment 3: The study mentions using widely adopted ML methods from previous research. Given LSTM's proven effectiveness in IDF study, why hasn't it been considered?

Response: We agree that Long Short-Term Memory (LSTM) models have demonstrated effectiveness in hydrological studies, particularly for tasks involving time-sequence data. We are aware that previous studies have utilized LSTM within the IDF framework, such as the temporal downscaling of precipitation data to enrich sample availability or the projection of non-stationary IDF curves in the future(Ameen et al., 2025; Bhattacharyya & Saha, 2023; Jahangiri et al., 2025). These applications capitalize on LSTM's strength in modeling temporal dependencies.

However, the primary objective of our study is different. Our work focuses on the spatial regionalization of IDF curves across mainland China, framed as a comparative analysis between traditional interpolation techniques and machine learning. To achieve this, our machine learning approach was designed as a spatial regression task. For each grid cell, the model predicts IDF curves based on a set of static features, including geographic coordinates, elevation, and various statistical metrics derived from annual and daily precipitation records. Since our input data do not constitute a time series, the architecture of an LSTM network is not directly suited for this specific regression problem. Furthermore, our investigation aimed to determine if machine learning methods, even when restricted to coarser daily gridded data, could achieve accuracy comparable to interpolation methods that rely on hourly observations. Our results confirmed this hypothesis, with methods like Gradient Boosting demonstrating robust performance and high accuracy. The success of the selected models in fulfilling our research objectives affirmed their suitability for the spatial regression task. Therefore, the inclusion of LSTM was not considered necessary for our comparative analysis. However, we acknowledge the potential for extending our work using LSTM. For instance, it could be applied to temporal downscaling in regions with sparse high-resolution data to generate more samples, or to investigate the non-stationarity of the regionalized IDF curves, which are promising directions for future research.

Comment 4: What are the limitations of the study?

Response: We thank the reviewer for this important question. We acknowledge that our study has several limitations, which can be broadly categorized into three areas: the data, the methodology, and the application.

First, regarding the data, the observational network of stations presents a degree of spatial heterogeneity. The sparse distribution of stations, particularly in the western regions of mainland China, limits both the predictive capability of our models and our ability to robustly validate the results in these specific areas. Furthermore, our analysis is based on record lengths ranging from 10 to 70 years, while our estimations extend to return periods of up to and exceeding 100 years. The long-range extrapolation based on the GEV distribution inevitably introduces a degree of uncertainty. Another data-related constraint is that our analysis was limited to a minimum rainfall duration of one hour, which means that more extreme, shorter-duration precipitation events were not considered.

In terms of methodology, a limitation is our estimation of stationary IDF curves. In the context of a rapidly changing climate, the potential non-stationarity of IDF curves is an important factor whose influence on our results warrants further investigation. Additionally, our selection of climate variables for prediction was focused on precipitation. The inclusion of other variables, such as humidity and temperature, was not explored in this study but could potentially improve the accuracy of our methods, especially in regions characterized by more complex climatic conditions.

As for application, while the regionalized products we provide at 0.1° and 0.5° spatial resolutions are an important advancement in mainland China, this level of granularity can still inevitably diminish the magnitude of precipitation extremes at specific locations. These spatial resolutions may not be sufficient for certain station-level requirements. This limitation could be more apparent in areas with high heterogeneity. Additionally, as we noted in the manuscript, the predictive accuracy is reduced in parts of the NW and SW regions due to the limited number of stations and the complex terrain. We have highlighted this uncertainty and advised that the results for these regions should be used with caution.

However, we believe these limitations do not detract from the successful achievement of our study's primary objectives or its overall significance. Instead, these limitations provide a clear and constructive direction for future work. These identified limitations create valuable opportunities for improvement through future research, including incorporation of higher-temporal-resolution precipitation data, the integration of a richer set of climate variables, retraining of models for complex terrain, and the extension of this framework to account for non-stationary IDF relationships.

Comment 5: Is it possible to transfer the outcomes to other geographically similar regions? If so, what considerations or adaptations would be necessary?

Response: From our perspective, the methodological framework and workflow developed in our study are indeed transferable to other geographically similar regions. However, the specific models developed and the selection of model types, as well as the resulting accuracy, should not be directly applied without training with local data. The successful application of our approach in a new region would be contingent upon a necessary process of data localization, feature engineering, and an uncertainty assessment.

To elaborate, our study compares two primary regionalization methods, one based on spatial interpolation of hourly station data and the other on machine learning with daily gridded precipitation. Both approaches demonstrated robust performance within our study area. When considering interpolation methods for a new region, a critical prerequisite is ensuring that the observational network has sufficient density and resolution to meet the demands of the chosen technique. Furthermore, our finding that mean annual precipitation was the optimal external drift variable for Kriging, while elevation was not always beneficial, highlights that the spatial structural relationship between the target variable and the covariates also influences the interpolation results.

This implies that the most suitable auxiliary variables for Kriging must be re-evaluated based on the characteristics of the target area.

In contrast, the machine learning approach, which relies on daily gridded precipitation data, is less dependent on the specific locations of gauging stations. This makes it a particularly viable option in regions where observational records with high temporal resolution are scarce. Nevertheless, our model was trained on data from mainland China. Its performance for a specific new region might be suboptimal without refinement. To improve predictive accuracy, we strongly recommend retraining the model with the inclusion of available local station data to better capture distinct regional precipitation patterns. Additionally, the quality and potential representativeness errors of the gridded precipitation dataset, which serves as the primary independent variable, must be carefully evaluated in the new context. It is also essential to validate which machine learning models and input features are the most effective for that specific region.

Additionally, the choice of regionalization method should be guided by regional characteristics. For instance, interpolation techniques may be more suitable for areas with stronger spatial autocorrelation in rainfall extremes, whereas machine learning might offer an advantage in regions where these spatial relationships are weaker or more complex. In short, our research confirms the applicability of traditional interpolation methods in mainland China and demonstrates the potential of machine learning for IDF curves regionalization. Yet, it also underscores a fundamental principle, echoing the “No Free Lunch” theorem, that no single method is always superior. In situations where the underlying assumptions of a method are not met, its performance will likely be poor. Therefore, when transferring our framework, we encourage future studies to incorporate region-specific features and conduct rigorous evaluations to enhance the credibility of their results. We believe our study provides a clear and detailed roadmap, offers new perspectives for developing high-precision regionalized IDF products, and serves as a valuable methodological and metric-based reference for further regional accuracy assessments in other similar regions.

Reference

Ameen, S. M., Aziz, S. Q., Dawood, A. H., Sabir, A. T., & Hawez, D. M. (2025). Utilizing machine learning and deep learning for precise intensity–duration–frequency (IDF) curve predictions. *Polytechnic Journal, 15*(1), 27–38. https://doi.org/10.59341/2707-7799.1848

Bhattacharyya, D., & Saha, U. (2023). Deep learning application for disaggregation of rainfall with emphasis on preservation of extreme rainfall characteristics for Indian monsoon conditions. *Stochastic Environmental Research and Risk Assessment, 37*(3), 1021–1038. https://doi.org/10.1007/s00477-022-02331-x

Jahangiri, M., Asghari, M., Niksokhan, M. H., & Nikoo, M. R. (2025). BiLSTM–Kalman framework for precipitation downscaling under multiple climate change scenarios. *Scientific Reports, 15*, 24354. https://doi.org/10.1038/s41598-025-08264-z