Journal of Hydrology 675 (2026) 135660

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Simulating sub-daily rainfall time series in the absence of sub-daily observations

F. Cappelli [a], E. Volpi [b], A. Langousis [c], R. Deidda [d], S.M. Papalexiou [e], A. Perdios [c], S. Grimaldi [a,*]

[a] DIBAF Department, Università degli Studi della Tuscia, Viterbo, Italy
[b] DICITA Department, Università degli Studi di Roma Tre, Rome, Italy
[c] Department of Civil Engineering, University of Patras, Patras, Greece
[d] Dipartimento di Ingegneria Civile, Ambientale edArchitettura, University of Cagliari, Cagliari, Italy
[e] Institute of Global Water Security, Hamburg University of Technology, Hamburg, Germany

ARTICLE INFO

This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Uwe Haberlandt, Associate Editor

Keywords:
Sub-daily rainfall simulation
Complete Stochastic Modelling System
Multifractal modelling
Canonical Multiplicative Random Cascade
Parsimonious rainfall model calibration

ABSTRACT

This paper presents a novel framework for simulating sub-daily rainfall time series in the absence of sub-daily observations, which are typically essential for calibrating conventional approaches. The proposed approach combines two classes of models: a daily rainfall model that generates long synthetic daily time series and a disaggregation model based on multifractal theory to refine the temporal resolution to sub-daily scales. The implemented procedure is parsimonious and relies solely on the observed daily rainfall time series and the power law exponent n of the intensity–duration–frequency curves, information widely available to practitioners.

The framework was tested on a challenging case study consisting of 70 rain gauges in the Arno River basin (Italy), each with 20 years of continuous 15-minute rainfall data. The performance was evaluated by comparing key statistical attributes estimated from the benchmark dataset and the simulated rainfall time series at 15-minute temporal resolution, such as the dry frequency, the autocorrelation at lags 1 and 10, and the dependence of rainfall intensity on the duration of spatial averaging and return period as embodied in the well-established notion of intensity–duration–frequency (IDF) curves.

The results show promising agreement, despite the limited sample size, which introduces some calibration challenges. The relative errors of the selected attributes fall within ± 15% for most of the analyzed time series, indicating that the framework offers a valuable alternative for hydrological studies, particularly in contexts where sub-daily observations are scarce or entirely absent.

1. Introduction

Accurate simulation of sub-daily rainfall time series is a critical requirement for a wide range of hydrological and climate applications, including flood frequency analysis, infrastructure design, and water resource management (Volpi et al., 2024; Fisher et al., 2025). These applications require long observed time series. As such data are often unavailable, synthetic series—potentially spanning several millennia—are used as an alternative to better capture the full range of variability inherent in rainfall processes. However, existing sub-daily rainfall simulation models commonly found in the literature require high-resolution calibration data (Pui et al., 2012; Beven, 2021; Grimaldi et al., 2022; Northrop, 2024). Alternative strategies have been proposed to bridge the scale gap between daily and sub-daily rainfall, including multifractal and universal multifractal formulations that exploit the scaling properties of rainfall processes (e.g., Gires et al., 2020), as well as point-process based models developed for sub-hourly rainfall generation (e.g., Park et al., 2021). Nevertheless, most of these approaches either rely on fine-scale observations for calibration or operate under assumptions that differ from the specific objective addressed here. Moreover, when high-resolution datasets are available, they are typically limited to 10–30 years of observations, which may not adequately capture the full physical variability inherent in the rainfall process.

In contrast, hydrologists generally have access to much longer

* Corresponding author at: Via San Camillo de Lellis, snc 01100, Viterbo, Italy.
E-mail addresses: Francesco.cappelli@unitus.it (F. Cappelli), Elena.volpi@uniroma3.it (E. Volpi), andlag@upatras.gr (A. Langousis), rdeidda@unica.it (R. Deidda), simon.papalexiou@tuhh.de (S.M. Papalexiou), tassosper13@hotmail.com (A. Perdios), Salvatore.grimaldi@unitus.it (S. Grimaldi).

https://doi.org/10.1016/j.jhydrol.2026.135660
Received 20 October 2025; Received in revised form 26 February 2026; Accepted 9 May 2026
Available online 10 May 2026

continuous daily rainfall datasets, spanning 50–100 years, as well as maximum annual rainfall, non-continuous, records over various durations (e.g., 1, 3, 6, 12, and 24 h). These datasets form the backbone of key hydrological tools, such as intensity–duration–frequency (IDF) curves, which are essential for flood analysis and infrastructure planning (Koutsoyiannis, 2022). Given their availability and informativeness, leveraging daily rainfall time series and IDF curves presents a promising avenue for simulating sub-daily rainfall processes.

This study introduces an innovative framework for simulating sub-daily rainfall time series without requiring continuous sub-daily observational data for calibration. The proposed approach combines a daily rainfall simulation model with a multifractal disaggregation scheme. Specifically, the daily simulation model is calibrated using observed daily rainfall data, while the disaggregation process employs a parsimonious procedure based on IDF curves. By eliminating the need of sub-daily continuous-time observations, this framework significantly broadens its applicability, particularly in regions where high-resolution rainfall data are scarce or unavailable. This approach not only addresses the limitations of existing models – due to limited calibration information – but also offers a versatile tool for hydrological and climate modeling across diverse geographical and climatic contexts.

For the simulation of daily rainfall time series, we adopt CoSMoS-2s (Papalexiou, 2022), a two-state intermittent model within the broader CoSMoS (Complete Stochastic Modelling System; Papalexiou, 2018). CoSMoS provides a general stochastic architecture for simulating hydroclimatic variables via explicit separation of marginal and dependence components, and has been extended to multisite and space–time random field settings (Papalexiou et al., 2021a, 2023). Open-source implementations are available in R (Papalexiou et al., 2021b) and Python (Cappelli et al., 2024). CoSMoS-2s reproduces rainfall intermittency while preserving the prescribed marginal distribution and autocorrelation structure at the daily scale. The model is selected for its flexibility and modular design, which allows transparent configuration of its components.

As for the disaggregation model we refer to the multiplicative cascade model based on multifractal theory. It relies on the assumption that the rainfall process shows scale-invariant behavior over a finite range of temporal scales relevant to hydrologic applications, utilizing a straightforward theoretical approximation of the IDF curves (Langousis et al. 2009). We opted for this model since the multifractal approach is centred on the scale-invariant behaviour of IDF curves.

In this contribution, we assess the proposed framework using a large dataset collected from 70 raingauges within the Arno River basin, located in Central Italy. Previous research (Cappelli et al., 2025a; Cappelli et al., n.d.) has evaluated the performance of the two individual models implemented and tested on the same dataset. In the following sections, we focus on assessing the performance of the entire framework, which combines the two individual models into a unified approach.

The manuscript is structured as follows. Section 2 introduces the concept of the proposed framework and provides a brief review of the theoretical background of the two models. Section 3 describes the testing strategy and presents the results obtained for a single case study. The section also summarizes the overall performance across the full set of 70 raingauges. Section 4 discusses the impact of sample size and highlights the advantages of simulating long daily time series, supported by an analysis of parameter behavior. Section 5 concludes with a summary of the key findings and outlines potential future research directions.

2. The framework for simulating sub-daily rainfall time series from daily observations

As stated in the introduction, currently available models for simulating sub-daily rainfall time series require continuous sub-daily observations for calibration. The proposed framework aims to overcome this crucial limitation by combining two approaches, i.e. a daily rainfall generator (CoSMoS-2s) and a multifractal disaggregation model (MRC), as illustrated in Fig. 1. The minimal required inputs for applying the framework (shown in blue in Fig. 1) are the observed daily time series and the power law exponent n of the IDF curves.

By calibrating the COSMOS-2s model using the available daily observations, it is possible to generate synthetic daily rainfall time series. Subsequently, using the power law exponent n of the intensity–duration–frequency curves, the multifractal disaggregation model is calibrated, enabling the transformation of the synthetic daily time series into synthetic sub-daily rainfall time series, which represent the final output of the framework.

In the following sub-sections, a brief overview of the two combined modeling procedures is provided, with particular attention to the parsimonious calibration of the multifractal disaggregation model, which was only recently introduced (Cappelli et al. 2025a).

2.1. Daily rainfall model simulation

In the univariate CoSMoS-2s implementation (Fig. 2), rainfall generation is represented through two coupled processes: a binary wet/dry process and a continuous rainfall intensity process, each modeled separately with Gaussian parent autoregressive formulations. In the present study, CoSMoS-2s is configured using sample autocorrelation coefficients to specify the dependence structure, as described below.

For the wet/dry process, the steps are as follows: (1) The rainfall record is transformed into a binary sequence, where wet days are assigned the value 1 and dry days 0. (2) Sample autocorrelations are computed separately for each seasonal period. (3) The sample binary autocorrelation coefficients $\rho_B(\tau)$, estimated from the observed time series, are transformed into their parent Gaussian equivalent $\rho_{ZB}(\tau)$ using the closed-form correlation transformation function for binary processes $\rho_{ZB}(\tau) = 1-(1-\rho_B(\tau))^{c_{p_0}}$ the parameter c_{p_0} is obtained analytically as a function of p_0 following Papalexiou (2022). (4) A Gaussian AR(2) model is then employed to reproduce reproduce explicitly the first two autocorrelation lags. (5) Daily probabilities of zero rainfall are estimated for each seasonal period. (6) The simulated Gaussian series is converted into binary wet/dry states by thresholding according to the seasonal probability of zero rainfall p_0; specifically, if $\Phi(Z_t) \leq p_0$ (where Φ is the standard normal CDF and Z_t the simulated Gaussian value), the day is classified as dry, otherwise as wet (section 6 in Papalexiou, 2022).

For the rainfall intensity process, a conditional sequence is applied: (1) means and standard deviations of nonzero daily rainfall are estimated on a seasonal period basis; (2) Gamma distributions parameterized by these seasonal patterns are fitted to represent the marginal behavior (any other distribution can be used); (3) sample conditional Spearman autocorrelation coefficients are estimated for each seasonal period considering sequences only of wet days (using the same rationale as in Eq. 22 in Papalexiou, 2022); (4) these sample conditional autocorrelation coefficients $\rho_S(\tau)$ are transformed into their parent Gaussian equivalents $\rho_{ZS}(\tau)$ via the closed-form correlation transformation function $\rho_{ZS}(\tau) = 2\sin(\pi/6\rho_S(\tau))$ (see also Fig. 2); (5) a Gaussian AR(2) model is then used to reproduce the temporal dependence of nonzero rainfall; and (6) the simulated Gaussian series are mapped back into rainfall intensities through the inverse of the fitted Gamma distributions.

Finally, the binary wet/dry simulations are combined with the rainfall intensity simulations to produce synthetic rainfall series that preserve the marginal properties and temporal autocorrelations observed in the data. Further methodological details are provided in Papalexiou 2018, 2022.

2.2. Sub-daily multifractal disaggregation cascade model

In the proposed framework we include the modelling approach developed in Cappelli et al. (2025a and 2025b), which relies on a simple

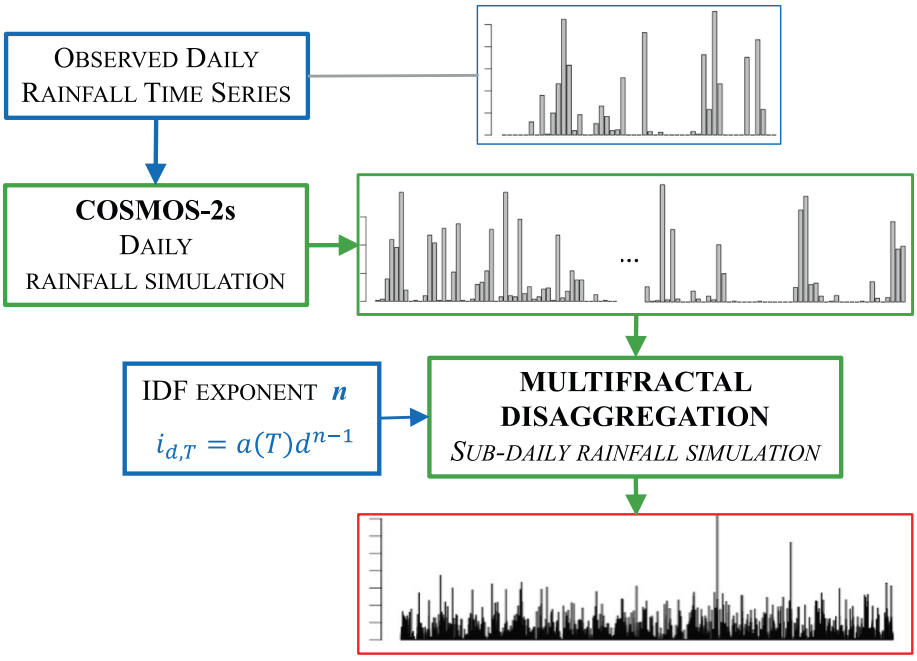


Fig. 1. The proposed framework scheme. In blue, the necessary input information; in green, the modelling steps; in red, the model output. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

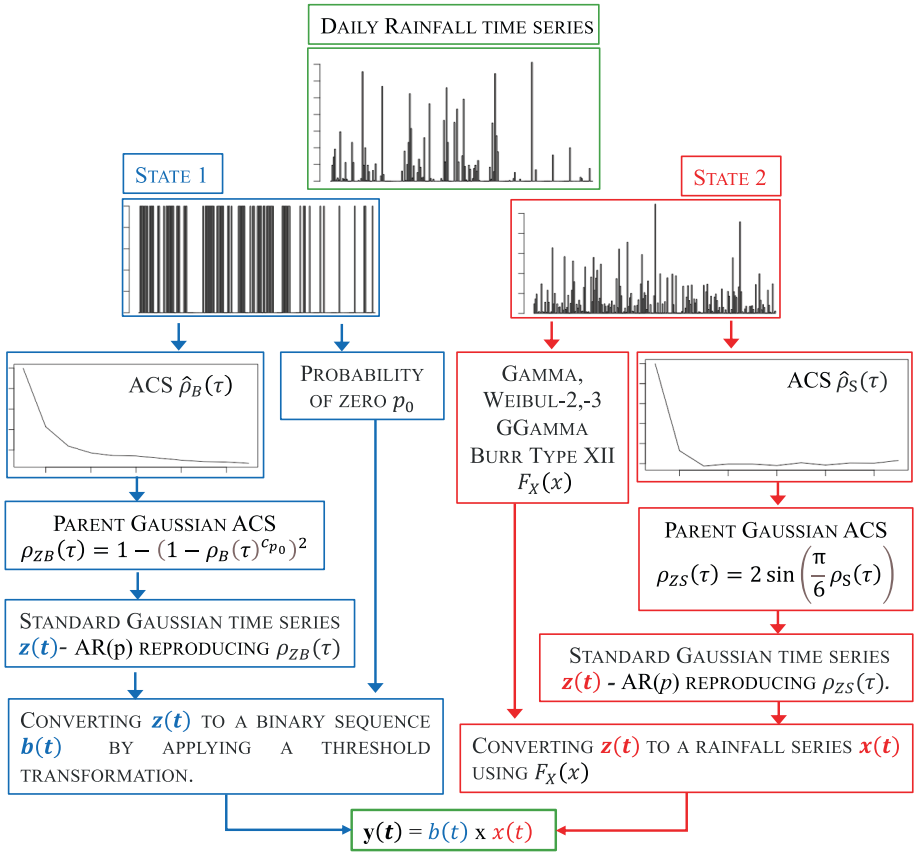


Fig. 2. The CoSMoS-2s framework for a single daily precipitation time series.

analytical approximation to the Intensity-Duration-Frequency (IDF) curves based on multifractal theory (see Langousis et al., 2009). According to Cappelli et al. (2025a and 2025b), the average rainfall intensity over duration d, I_d, is approximated by a stationary beta-lognormal multifractal process and subsequently modeled using a discrete binary (multiplicity $m = 2$) multiplicative random cascade (MRC) with beta-lognormal generator. The latter results as the product of two independent random factors. The first factor $A_{\beta,m=2}$ is equal to 0

or 1 with probability $1-2^{-C_\beta}$ and 2^{-C_β}, respectively, and is used to model the alternation of wet and dry intervals. The second factor, $A_{LN,m=2}$, which describes the intensity of the fluctuations inside wet intervals, follows a lognormal distribution with log-mean $m_{A_{LN,m=2}} = (C_\beta - C_{LN})\ln 2$, and log-variance $\sigma^2_{A_{LN,m=2}} = 2C_{LN}\ln 2$; i.e. $\ln A_{LN,m=2}\ N\left[(C_\beta - C_{LN})\ln 2, 2C_{LN}\ln 2\right]$. The parameters C_β and C_{LN} are non-negative and satisfy the condition $C_\beta + C_{LN} < 1$ (see, e.g., Langousis and Veneziano, 2007). Disaggregation by the MRC model requires the estimation of three parameters, i.e. C_β that controls the fraction of dry periods, C_{LN} that controls the amplitude of the multiplicative fluctuations when it rains, and D the maximum temporal scale up to which multifractal scale invariance holds. The aforementioned parameters specify the log–log slope and spacing of the multifractal IDF curves, and can be obtained by properly fitting (see Cappelli et al., 2025a, and Fig. 3 for an illustration and Section 3.1 below) the multifractal analytical approximation to empirically derived IDF curves in the form (see e.g. Veneziano et al., 2007; Tyralis and Langousis, 2019; Emmanouil et al., 2020):

$$i_{d,T} = a(T)d^{n-1} \tag{1}$$

where $i_{d,T}$ [mm/h] is the rainfall intensity over duration d [h] with return period T [yr], n [-] is a positive constant with $0 < n < 1$, and $a(T)$ [mm/h^n] is a function independent of d. After the parameters of the MRC model have been estimated, disaggregation of daily rainfall rates proceeds utilizing the cascade version described in Cappelli et al. (2025b) that preserves exactly the 0-level (i.e. daily) rainfall averages.

3. Illustrative example

In this section, we outline the strategy adopted to evaluate the performance of the proposed framework. To this aim, the analysis relies on a comprehensive rainfall database specifically assembled for testing and comparing different rainfall models. This dataset was built using rainfall observations from a dense raingauge network operating in the Arno river basin, located in Central Italy (see Fig. 4 for a geographical reference). Out of 123 available stations, 70 were selected, as a benchmark, based on their completeness and data quality, ensuring continuous and contemporaneous observations at a 15-minute resolution with minimal data gaps over a 20-year period. These time series were already used in previous studies to test the individual models included in this framework. Readers can refer to Cappelli et al. (2025b) for a detailed description of the selected dataset.

In the following Sections the performance evaluation of the coupled framework depicted in Fig. 1 is presented in two parts. In Section 3.1, we conduct a detailed analysis focusing on a single raingauge while in Section 3.2, we extend the analysis to all 70 selected time series, providing a broader view of the framework's performance across the entire dataset.

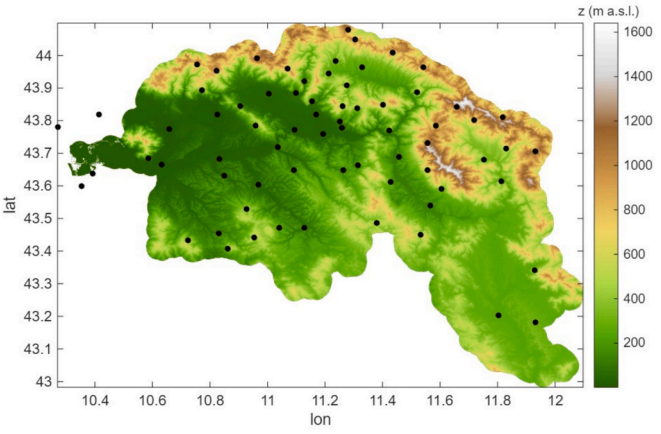


Fig. 4. Geographic distribution of the 70 selected rainfall time series (black dots) of the Arno River basin, Central Italy (
Adapted from Fig. 2 of Cappelli et al., 2025b).

3.1. Single case study results

The aim is to present the steps involved in applying the proposed framework to a single rain gauge time series, including detailed examples. Fig. 5 illustrates the key properties of the selected series (Fig. 5a): the seasonal empirical Spearman autocorrelation structure (Fig. 5b), and the empirical cumulative distribution function of rainfall depth for the four seasons (Fig. 5c).

In Fig. 5b we observe that Spring (green dots) and Winter (purple dots) exhibit higher short-term memory compared to Autumn (orange dots) and Summer. From the ECDF plot of rainfall depths (Fig. 5c), Summer exhibits low rainfall depths, whereas Winter, Autumn and Spring display higher exceedance probabilities of heavier rainfall.

The original 15-minute resolution data, used here as a benchmark for the case study, were aggregated to daily time series. These upscaled data were then used to estimate the parameters of both model components. Finally, the proposed two-step nested framework was applied: a synthetic daily time series was first generated using COSMOS-2s, then disaggregated back to 15-minute resolution to simulate 2,000 years of rainfall at the final temporal scale.

As illustrated in Fig. 1, the first step consists of calibrating the CoSMoS-2s model. To account for seasonality (see Section 2.1), the time series was divided into four periods—Winter, Spring, Summer, and Autumn. This strategy mitigates both sample size limitations and risks of over-parameterization by pooling data within each season.

Based on this seasonal grouping, several candidate distributions were fitted separately for each season to ensure consistency in parameterization. Among these, the Gamma distribution provided the best fit. The fitting was performed numerically by optimizing the distribution parameters to minimize the distance between the empirical and

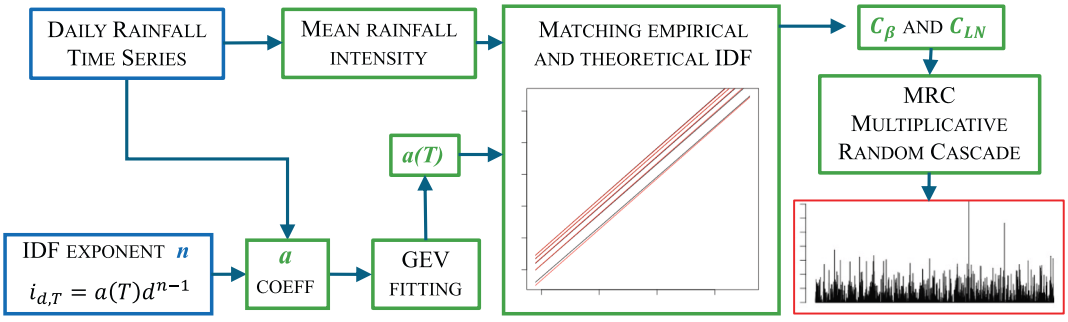


Fig. 3. Multifractal disaggregation model parsimonious calibration procedure. In blue, the required input information; in green, the modelling steps; in red, the model output. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

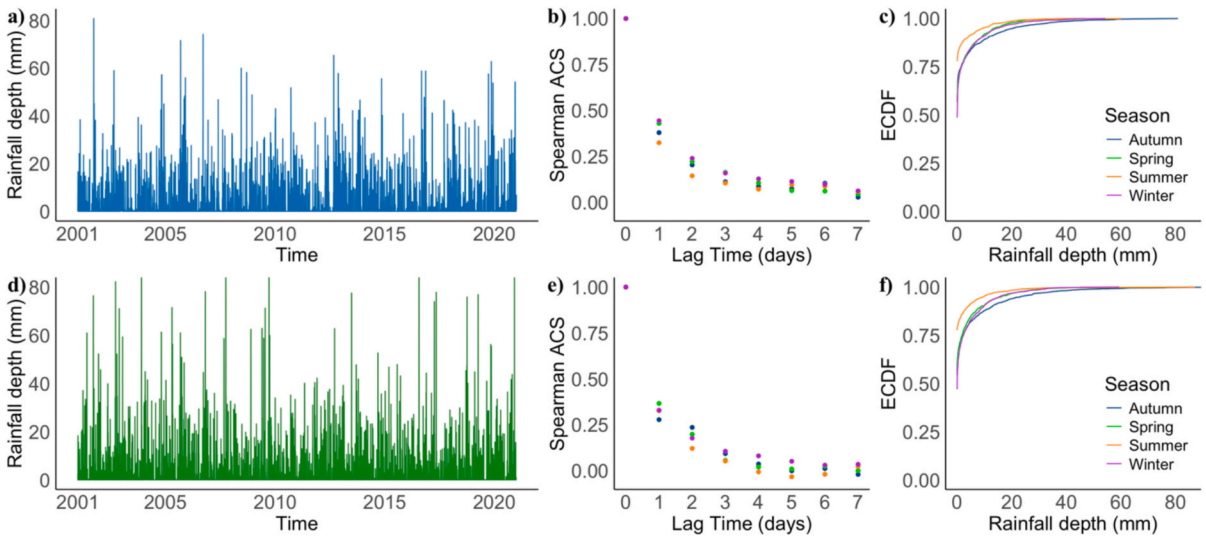


Fig. 5. Raingauge TOS01000861: (a) observed daily rainfall time series, and related (b) empirical Spearman autocorrelation structure (ACS) by season, (c) empirical cumulative distribution function (ECDF) by season, (d) 20-year daily simulated rainfall time series, and related (e) empirical Spearman ACS by season, and (f) empirical ECDF by season.

theoretical cumulative distribution functions of daily rainfall records. To further assess adequacy, we examined the behavior of the fitted distribution in the tails, as a good fit in the central range may conceal discrepancies at the extremes, which are particularly relevant in hydrological applications. Specifically, Fig. 6 compares the Gamma distribution fitted to non-zero daily rainfall depths with the corresponding empirical distribution for the four seasons in terms of exceedance probability. The fits are satisfactory across all seasons.

With respect to the autocorrelation structure, we adopted an empirical approach whereby the first three sample autocorrelation coefficients are directly estimated from the observed time series. No theoretical autocorrelation function is fitted. The AR(2) model is constructed to reproduce the first two lags explicitly, while the third lag is used as a consistency check rather than being imposed. Since the daily rainfall time series exhibits weak autocorrelation beyond short lags, setting $p = 2$ is sufficient to capture the dominant short-term dependence while avoiding unnecessary parameterization.

The parameters of the Gamma distribution and the autocorrelation structure of the binary and continuous time series (Table 1) reveal clear seasonal differences. Specifically (Table 1, left panel), Autumn exhibits a highly skewed distribution, likely due to the combination of many days with low precipitation and the occurrence of more intense events. Winter also shows a skewed distribution, though with moderate but less extreme rainfall compared to Autumn, while Spring and Summer display slightly less skewed distributions.

Regarding the autocorrelation structure, the ACS coefficients of the binary time series at lag 1 (Table 1, right panel) are relatively high across all seasons, particularly in Spring (0.40) and Winter (0.35), indicating that rainfall tends to recur on consecutive days. In Summer and Autumn, the values at lag 1 are slightly lower (~0.32), suggesting greater daily variability in rainfall occurrence. As expected, autocorrelations decrease progressively at longer lags in all seasons.

Based on the fitted parameters reported in Table 1, a 2,000-year long daily time series was stochastically simulated. The resulting series (Fig. 5d) exhibits more intense rainfall events compared to the observed data—an expected outcome, given that a 2,000-year simulation is more likely to encompass rare or extreme rainfall episodes. Both the simulated and observed series share a similar autocorrelation behavior: a steep decline at the first lag and values nearing zero after just a few lags. However, the observed series shows slightly stronger seasonal variability at lag 1. Additional details concerning the CoSMoS-2s implementation and performance evaluation are provided in Cappelli et al. (in

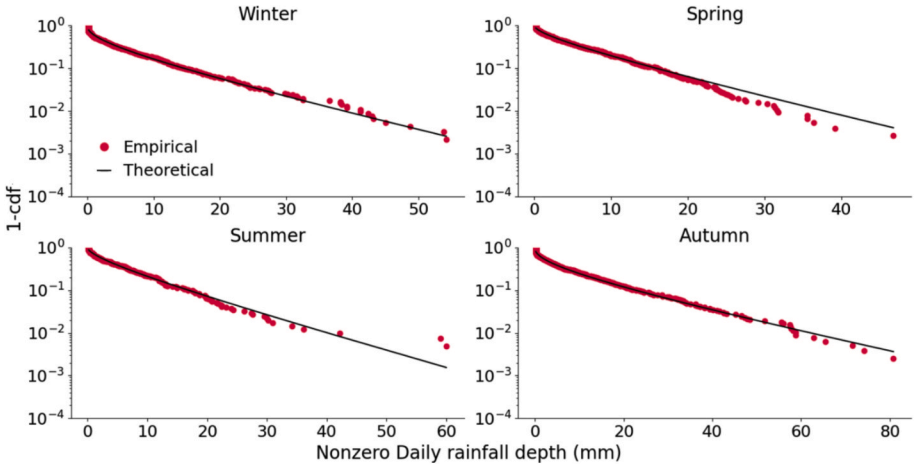


Fig. 6. The Gamma fitted distribution to nonzero daily rainfall depths compared with the empirical distribution for the four seasons in terms of exceedance probability (1-CDF).

Table 1
Gamma parameter estimates after calibration and the Spearman's ACS coefficients ($\rho_S(1)$, $\rho_S(2)$ and $\rho_S(3)$) of the Binary and Continuous time series.

Season	Gamma Parameters		Binary TS			Continuous TS		
	Shape	*Scale*	$\rho_S(1)$	$\rho_S(2)$	$\rho_S(3)$	$\rho_S(1)$	$\rho_S(2)$	$\rho_S(3)$
Winter	0.40	12.92	0.35	0.20	0.13	0.21	0.09	0.12
Spring	0.55	10.91	0.40	0.22	0.15	0.09	0.04	0.01
Summer	0.55	11.59	0.32	0.15	0.07	−0.07	−0.04	0.04
Autumn	0.37	21.44	0.32	0.22	0.13	0.12	0.14	0.07

review).

Based on the fitted parameters, the simulated 2,000-year daily time series is then used as input for the multifractal disaggregation model (i.e., the multiplicative random cascade described in Section 2.2), together with the exponent n estimated from the benchmark series at 15-minute resolution.

The first step of the disaggregation procedure (Fig. 3) involves estimation of the MRC parameters based on available data. For this purpose, the exponent n of the empirical IDF curves (assumed known a priori in practical applications) is here estimated by best fitting Equation (1) (see also Equation (7) in Cappelli et al., 2025a) to the annual maxima obtained after averaging the historical series to various durations of temporal averaging. Then, a Generalized Extreme Value (GEV) distribution model is fitted to the a values estimated on the daily rainfall maxima (inverting the Eq. (1) extracted from the historical rainfall series, and the estimated GEV parameters are used to obtain the $a(T)$ function; see Equation (1) herein and Equation (6) in Cappelli et al. (2025a). As a final step, the empirical (Equation (1) and theoretical (multifractal; Equation (1) in Cappelli et al., 2025b, and Equation (4) in Cappelli et al., 2025a) IDF curves for selected return periods (T) are matched (Fig. 7) to infer the parameters C_β and C_{LN} of the MRC model. In Fig. 7a, the match between theoretical (red) and empirical (black) IDF curves is shown for $T = 2$, 4, 6, 8, and 10 years—note that the calibration range is limited by the length of the observed record. Due to the statistical scaling nature of the cascade model, the agreement between observed and theoretical curves is excellent (see also the objective function values in Fig. 7b).

Once these parameters are calibrated (summarized in Fig. 7b), the multiplicative random cascade is applied to disaggregate the synthetic daily series down to the desired sub-daily resolution—set in this case to match that of the observed data.

To evaluate the 2,000-year synthetic 15-minute rainfall series, we compare its main statistical properties against the observed 15-minute benchmark data. The comparison focuses on five key metrics: the IDF curves, autocorrelation structure, marginal distribution of 15-minute rainfall depth (return period–rainfall amount relationship), the exponent n, and the dry-interval frequency (percentage of dry intervals).

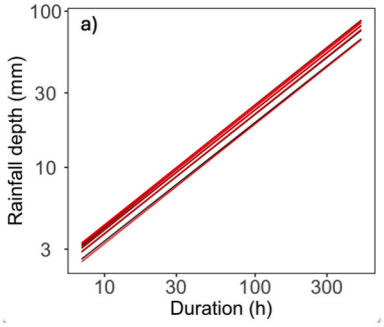
Fig. 8 presents the IDF comparison for four return periods (2, 5, 10, and 20 years). We emphasize lower return periods, as the benchmark dataset spans only 20 years. The figure contrasts the IDF curves derived from the benchmark series (20 years) with those from the synthetic series (2,000 years). For $T = 2$ and 10 years, the simulation replicates the theoretical behavior calibrated from the observed data. The remaining panels display results for return periods not used during calibration, and even there, the framework achieves good agreement with the benchmark. This agreement is critical, as IDF curves reflect rainfall characteristics closely tied to flood generation and the realism of simulated runoff time series.

Fig. 9a compares the autocorrelation structures. The disaggregation model effectively preserves the observed features, although it exhibits a slight general underestimation. Fig. 9b illustrates the 15-minute rainfall accumulation as a function of return period. Despite the limited sample size, the synthetic and observed time series display remarkable similarity, including in the upper tail of the distribution.

A comparison between the exponent n—which represents the slope of the intensity–duration–frequency (IDF) curves—and the dry-interval frequency (estimated on wet days) reveals a close agreement between observed and simulated data. The exponent derived from the 2,000-year synthetic series ($n = 0.246$) shows a slight overestimation compared with the value obtained from the 20-year observed record ($n = 0.235$). In contrast, the frequency of dry intervals is almost identical in the two datasets, with 89.6% for the observed series and 89.2% for the simulation. This strong correspondence confirms the model's ability to accurately reproduce the intermittency and the proportion of dry intervals characteristic of the observed rainfall time series.

In addition, Fig. 10 compares the dry-interval frequency (dry-frequency) at various temporal aggregation scales for both the observed and simulated series. The time scales analyzed include 15 min, 30 min, 1 h, 3 h, 6 h, 12 h, 1 d, 3 d, 5 d, and 7 d. As expected, dry-frequency naturally declines as the aggregation scale increases, since longer intervals are less likely to remain completely dry (i.e., they are more likely to include precipitation within them).

At finer scales (15 min to 1 h), the synthetic and observed series nearly overlap, indicating the model accurately reproduces the frequency of dry periods at high resolution. At coarser scales (3 h and above), the simulation slightly overestimates dry-frequency compared to observations, with the difference increasing to approximately 7–10% at daily to weekly scales. Nonetheless, this modest bias remains well within acceptable limits for hydrologic applications, underscoring the model's robust capability in reproducing dry interval intermittency and scaling behavior across temporal resolutions.



b) Input-Output calibration parameters	
GEV-location (mm)	26.16
GEV-scale (mm)	6.17
GEV-shape (-)	-0.19
Mean annual rainfall intensity (mm/h)	0.103
n-value (-)	0.235
C_b (-)	0.615
C_{Ln} (-)	0.019
D (h)	168.99
Objective function value (mm²/h²)	4.41

Fig. 7. Multiplicative random cascade model calibration. a) comparison among theoretical (Eq. (1)) and empirical (Eq. (2)) IDFs referring to Return Periods 2, 4, 6, 8, 10 years; b) Input parameter (GEV-location, GEV-scale, GEV-shape, Mean annual rainfall intensity, n value) and Output parameters (C_β, C_{LN}, D) with the objective function value resulting from the optimization.

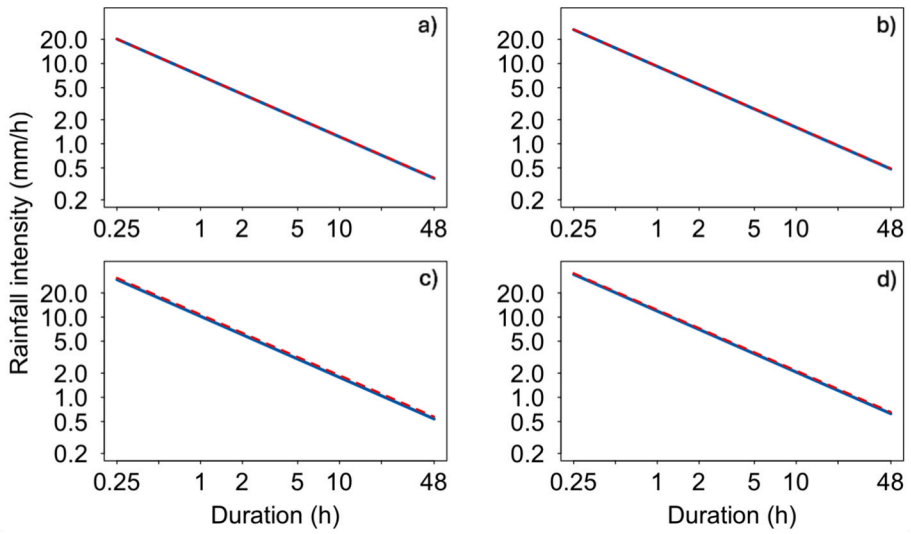


Fig. 8. Intensity-duration-frequency curves in log–log plots. (blue line) Benchmark observed rainfall time series; (red dashed line) proposed framework simulation. Each plot refers to a different return period: (a) $T = 2$ years; (b) $T = 5$ years; (c) $T = 10$ years; (d) $T = 20$ years. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

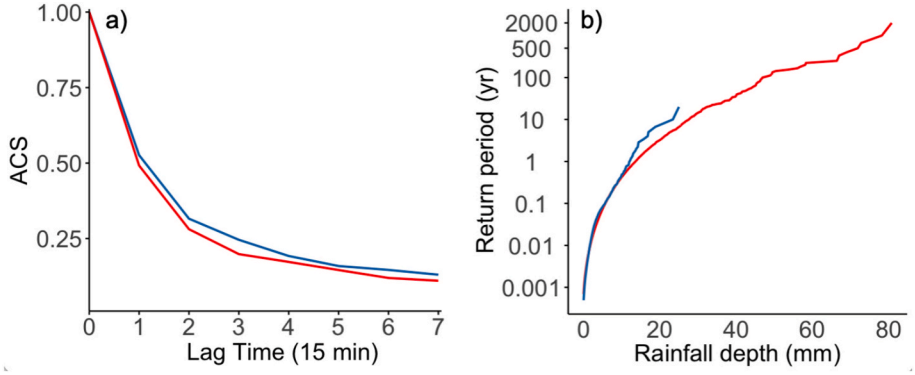


Fig. 9. Autocorrelation structure (ACS) (a) and 15-minute rainfall depth as a function of the empirical return period (b). (blue line) Benchmark observed rainfall time series; (red line) proposed framework simulation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

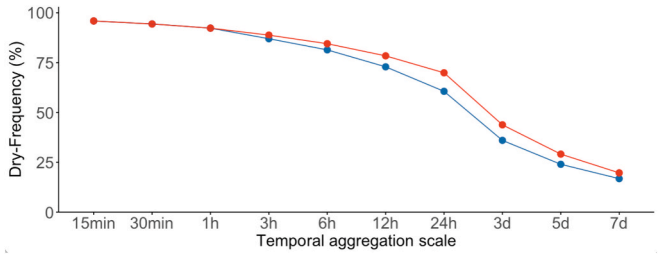


Fig. 10. Comparison of dry-frequency between the observed (blue) and simulated (red) rainfall time series calculated on different time aggregation scales, from 15-minute up to 7-day: 15-minute, 30-minute, 1-hour, 3- hour, 6-hour, 12-hour, 24-hour, 3-day, 5-day, and 7-day. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

A deeper comparison of wet and dry spell lengths, as well as rainfall intensity during events, is presented in Figs. 11 and 12. Rainfall events are defined here as continuous sequences of wet 15-min time steps; therefore, they represent uninterrupted rainfall segments rather than fully independent meteorological events. Overall, the results indicate a satisfactory reproduction of both properties, particularly for dry spell lengths and for rainfall intensity associated with shorter events. However, a slight underestimation of longer wet spell durations is evident in both figures. This discrepancy affects only a small fraction of the analysed events (approximately 1%) and is likely related to a modest overestimation of the parameter C_β.

As a general comment, although the performance evaluation is challenging due to the restricted sample size, the proposed framework delivers promising results despite being calibrated solely by constraining the exponent n. This further reinforces the importance of the recently introduced parsimonious calibration procedure.

3.2. Complete database analysis

To provide a comprehensive assessment of the proposed framework's performance, the entire procedure described in the previous section was repeated for all 70 rainfall time series available in the benchmark database. Regarding the application of CoSMoS-2s, we opted to use the Gamma distribution for all time series in order to automate the fitting and simulation procedure. This choice was based on the fact that the Gamma distribution proved to be optimal for approximately 75% of the stations in the database. As a result, the performance reported here should be considered indicative rather than optimal. A station-by-station inference would likely lead to improved results, but such a

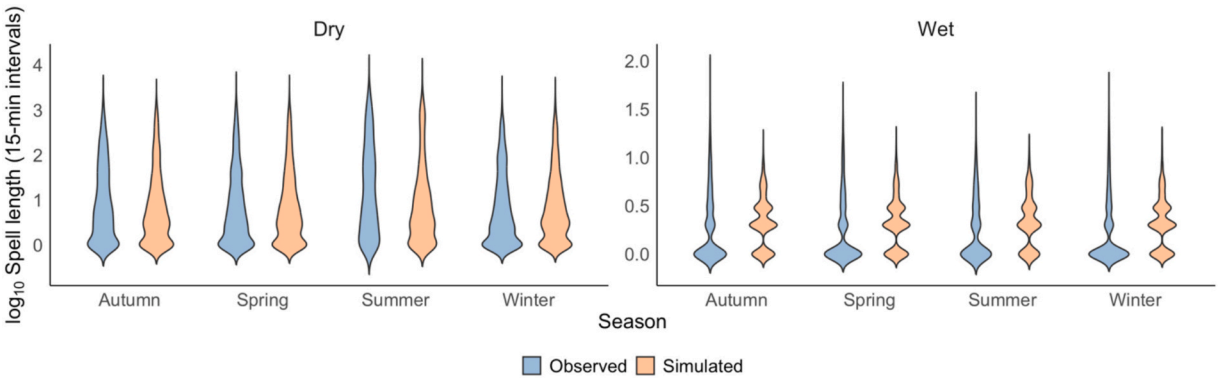


Fig. 11. Comparison of dry and wet spell lengths estimated from observed and simulated rainfall time series (20-year records) across the four seasons considered in the analysis.

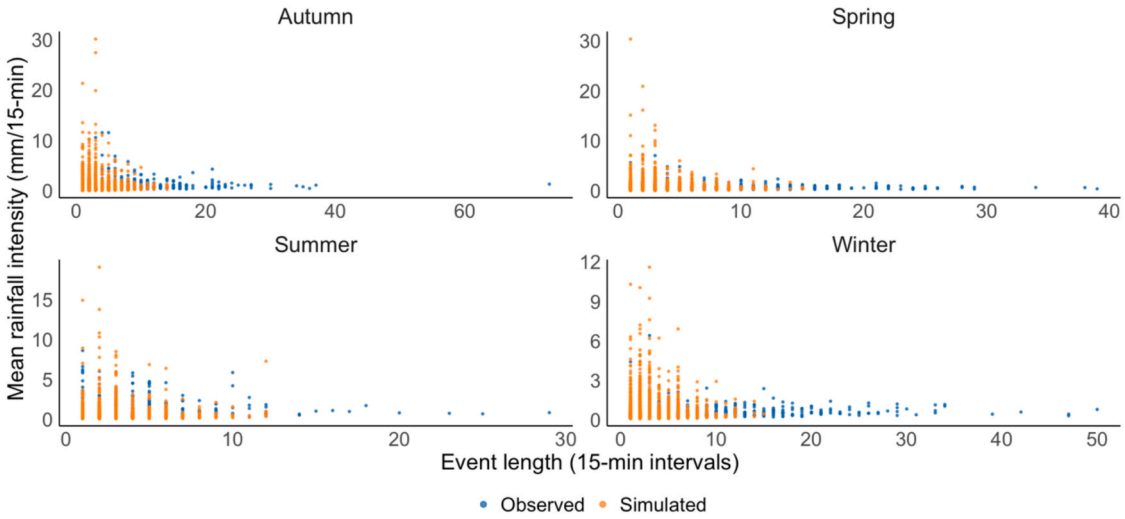


Fig. 12. Rainfall intensity as a function of event length, estimated from observed and simulated 20-year rainfall records across the four seasons considered in the analysis.

detailed analysis falls outside the scope of this paper, which focuses instead on demonstrating the applicability of the method at a broader scale.

Firstly, at each site, we evaluated the performance in reproducing relevant attributes (a) by the following relative error $RE(a)$

$$RE(a) = 1 - \frac{a_s}{a_b} \tag{2}$$

where a_s and a_b correspond to any relevant attribute (a) evaluated on the synthetic 2000-year (s) and benchmark (b) time series.

For the sake of clarity and to offer a more efficient overview of the results, Fig. 11 summarizes the comparison between the proposed framework simulations and the benchmark data with density plots based on 70 $RE(a)$ values for each relevant attribute. Specifically, results based on the reproduction of the following main attributes of the rainfall time series are shown in each panel of Fig. 13: the exponent n, the dry frequency at 15 min temporal resolution, ACS at lag 1 and lag 10, and 15-minute rainfall depth for three return periods (1, 5, and 10 years).

The results are encouraging, especially considering the limited sample size. For the exponent n (Fig. 13a), the mode of the empirical distribution is close to zero, and 73% of the raingauges show an error within ± 15%. This is also because the IDF curve slope n is used for disaggregation model calibration. Further, the entire dataset accurately captures the dry frequency attribute (Fig. 13b), which depicts an empirical distribution among the raingauges strongly concentrated around the zero error value. Note that, unlike n, the zero frequency at 15-minute resolution is not considered for the random cascade model calibration.

When examining the selected autocorrelation coefficients, there is a frequent but limited underestimation at lag 1 (Fig. 13c), with 97% of the database showing errors within the ± 15% range. However, a more evident underestimation is observed at lag 10 (Fig. 13d), where the mode corresponds to a 12% error, indicating a slight deviation from the expected value.

A similar pattern is observed in the three density plots representing 15-minute rainfall amounts for the three selected return periods 1, 5, and 10 years (Fig. 13e,f,g), where the bias reduces and the uncertainty increases with increasing return period (as expected). Although some variability is evident, the overall performance confirms the robustness of the proposed framework.

4. The role of synthetic daily rainfall time series as input

As introduced in Sections 1 and 2, the primary goal of the proposed procedure is to optimize the available information by leveraging longer daily time series. This advantage is crucial because traditional flood frequency analyses typically focus on large return periods (up to 1000 years, Read and Vogel, 2015), and the longer the observed sample, the more representative the information on rainfall fluctuations becomes.

In addition to this evident benefit, we identified another intrinsic

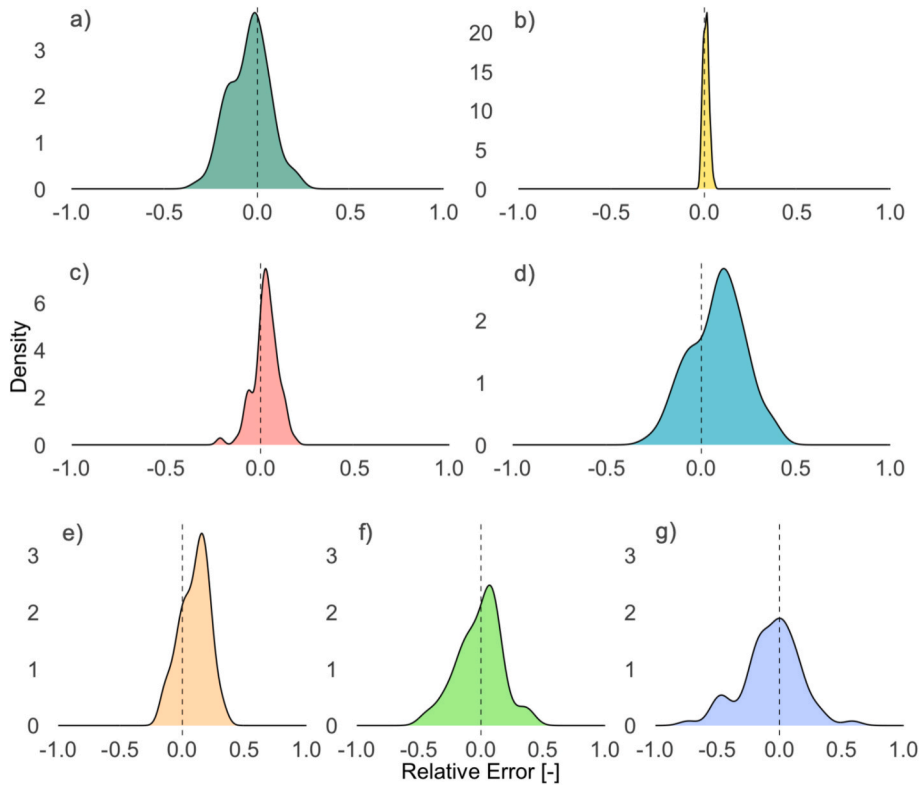


Fig. 13. Performance evaluation of the proposed framework on 70 rainfall time series. For each time series (simulated and observed) an attribute is quantified by the relative error *RE* in Eq. (2). The evaluated attributes include: (a) exponent *n*, (b) dry frequency, (c and d) ACS coefficient at lag 1 and 10, and (e,f, and g) the three-rainfall intensity (at 15-minute) related to 1, 5, and 10 of return period. Each plot represents the empirical density plot of the 70 *RE* values of a specific attribute.

added value of the framework related to the use of a daily simulation model and the subsequent application of multifractal disaggregation to the synthetic daily time series. Specifically, we observed greater stability in the calibration process.

When only 20 years of data are available, fitting a GEV distribution can lead to unrealistic parameter estimates, propagating this bias into the empirical IDF curves and, consequently, into the C_β and C_{LN} estimates used for the multiplicative random cascade. To support this observation, Fig. 14a shows the C_β and C_{LN} pairs derived from the observed 20-year time series (blue dots) for all 70 raingauges. The data exhibit significant dispersion, reaching both the upper and lower bounds of the parameter ranges. In contrast, the red dots (Fig. 14c)—representing the C_β and C_{LN} pairs obtained from the 2000-year synthetic time series—are well clustered and consistent with the limited spatial variability of the case study.

While one might argue that this result could be driven by the fact that the synthetic data were generated using a specific distribution function, however this assumption is only partially correct; indeed, the orange dots in Fig. 14b represent the C_β and C_{LN} pairs derived from the first 20 years of the 2000-year synthetic time series. The resulting cloud of points closely resembles that obtained from the observed data, confirming that sample size plays a critical role in model stability. Thus, using a long synthetic daily time series as input significantly mitigates this instability and enhances the robustness of the procedure.

5. Conclusion and future remarks

In this work, we proposed an original framework for simulating sub-daily rainfall time series without the need for sub-daily observations, which are currently mandatory for calibrating models available in the literature. The idea behind the framework is simple yet innovative: it combines two classes of rainfall models—a daily rainfall model to simulate daily rainfall time series and a disaggregation model to refine the time resolution of the simulated daily data. The framework is calibrated through a parsimonious procedure requiring only two inputs: the daily time series and the power law exponent *n* of the

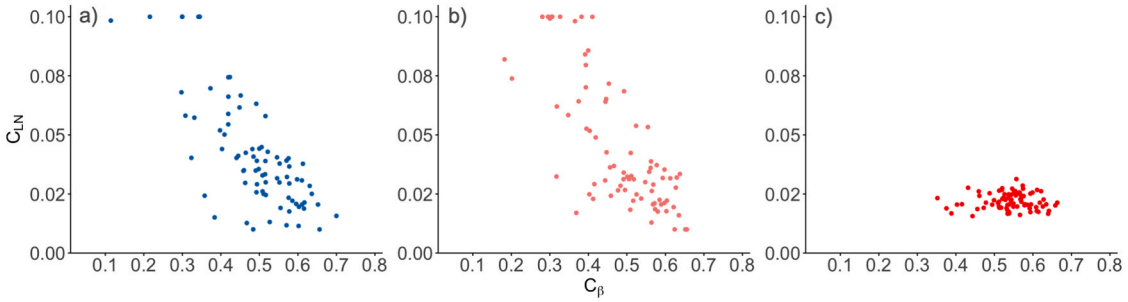


Fig. 14. Multifractal disaggregation model parameter C_β and C_{LN}. Each point identifies the pair values related to one raingauge. (a) benchmark time series; (b) the first 20-year of the proposed framework simulation; (c) the 2000-year proposed framework simulation.

intensity–duration–frequency (IDF) curves so using minimal information—precisely the information for which the longest hydrological records are typically available.

In this study, we employed the Complete Stochastic Modelling System (CoSMoS) using the two-state intermittent model (CoSMoS-2s) for daily rainfall simulation and a multiplicative cascade model based on multifractal theory for disaggregating the daily series to sub-daily scales.

We evaluated the performance of the framework using a challenging case study involving 70 rain gauges located in the Arno River basin. The dataset consists of high-quality 15-minute continuous and contemporaneous rainfall time series spanning 20 years. The comprehensive analysis supports promising conclusions regarding the performance of the proposed framework, specifically, several key rainfall attributes were compared between the benchmark data and the simulated time series, including the IDF exponent *n*, dry frequency, autocorrelation structure (ACS) at lags 1 and 10, and the return period–rainfall amount relationships for three return periods (1, 5, and 10 years). The agreement between the simulated and observed data was encouraging, with only a minor portion of the dataset being not well reproduced. This discrepancy is most likely due to the limited 20-year sample size, which may bias the calibration process.

Beyond evaluating the framework's reliability, we investigated the role of the long synthetic input daily rainfall time series. Our findings confirmed that a 20-year time series is too short for reliable parameter estimation in such models. However, using a 2000-year synthetic daily rainfall time series effectively mitigates biases in the calibration procedure, providing more robust results.

Several directions for future research can be pursued to improve and expand the proposed framework, both in terms of methodological refinement and broader applicability. First, an important step would be to test the framework in different climatic regions, applying it to diverse datasets. This would help assess its robustness and generalizability across varying hydrological conditions, particularly in regions characterized by different rainfall regimes, such as arid areas with sporadic but intense rainfall or tropical climates with highly seasonal precipitation patterns. Expanding the geographical scope of the analysis would offer valuable insights into the adaptability of the model and its potential limitations.

Another useful direction is the integration of the simulated sub-daily time series into continuous rainfall-runoff models. This would allow for a more comprehensive evaluation of the framework's impact on hydrological predictions, particularly for extreme flood events. Understanding how the synthetic sub-daily time series influences the simulation of peak flows could provide useful feedback for eventually improving the rainfall modelling.

The proposed framework also offers considerable flexibility, as both the daily rainfall model and the multiplicative cascade scheme can be substituted with alternative models. Exploring different combinations of these models could help identify optimal configurations that improve performance under specific conditions. For instance, testing more complex daily rainfall models or adopting alternative disaggregation techniques might enhance the framework's accuracy in reproducing key rainfall features. Such experimentation would help determine whether there are more effective combinations that better capture the temporal variability and extremes of rainfall.

Furthermore, the framework was designed with a long-term goal of extending to multisite applications, where daily rainfall models are computationally efficient and relatively simple to implement. This multisite approach could significantly enhance the framework's applicability for regional studies, where preserving the spatial coherence of rainfall patterns is critical. An interesting avenue for future research would be to evaluate how much of this spatial correlation is retained at the sub-daily scale after disaggregation and its effects in distributed hydrological modeling.

In conclusion, the proposed framework presents a promising and practical solution for generating sub-daily rainfall time series from minimal input data. Its flexibility and scalability make it a valuable tool for hydrological studies and flood risk assessments, especially in contexts where sub-daily observations are limited or unavailable.

CRediT authorship contribution statement

F. Cappelli: Writing – review & editing, Methodology, Data curation, Conceptualization. **E. Volpi:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization. **A. Langousis:** Writing – review & editing, Methodology, Conceptualization. **R. Deidda:** Writing – review & editing, Methodology, Funding acquisition, Conceptualization. **S.M. Papalexiou:** Writing – review & editing, Methodology, Conceptualization. **A. Perdios:** Writing – review & editing, Methodology, Conceptualization. **S. Grimaldi:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by: Italian Ministry of University and Research (PRIN2022 – DeHySi Project – DEsign HYdrologic Simulation – No. 2022NBXJSL, 2023–2025); the RETURN Extended Partnership and received funding from the European Union Next-GenerationEU (National Recovery and Resilience Plan – NRRP, Mission 4, Component 2, Investment 1.3 – D.D. 1243 2/8/2022, PE0000005); the "Accordo di collaborazione tra l'Autorità di Bacino Distrettuale dell'Appennino Meridionale e il Dipartimento per la innovazione nei sistemi biologici, agroalimentari e forestali (DIBAF) dell'Università degli Studi della Tuscia" – CUP F54J16000030001 – CUP F52G16000010001.

Data availability

Data will be made available on request.

References

Beven, K., 2021. Issues in generating stochastic observables for hydrological models. Hydrol. Process. 35 (6), e14203. https://doi.org/10.1002/hyp.14203.

Cappelli, F., Papalexiou, S.M., Markonis, Y., Grimaldi, S., 2024. PyCoSMoS: An advanced toolbox for simulating real-world hydroclimatic data. Environ. Model. Softw. 178. https://doi.org/10.1016/j.envsoft.2024.106076.

Cappelli, F., Papalexiou, S.M., Markonis, Y., Volpi, E., Grimaldi, S., n.d.. (in review) Bridging Flexibility and Usability: Configuring CoSMoS-2s for Operational Rainfall Simulation in the Arno River Basin. Hydrol. Sci. J..

Cappelli, F., Volpi, E., Langousis, A., Deidda, R., Perdios, A., Furcolo, P., Grimaldi, S., 2025a. Sub-daily rainfall simulation using multifractal canonical disaggregation: a parsimonious calibration strategy based on intensity-duration-frequency curves. Stoch. Env. Res. Risk A. 39 (1), 1–19. https://doi.org/10.1007/s00477-024-02827-8.

Cappelli, F., Volpi, E., Langousis, A., Deidda, R., Perdios, A., Grimaldi, S., 2025b. Rainfall simulation based on parsimonious calibration of a multifractal canonical disaggregation scheme in the Arno river basin, Italy. J. Hydrol.: Reg. Stud. 59, 102447. https://doi.org/10.1016/j.ejrh.2025.102447.

Emmanouil, S., Langousis, A., Nikolopoulos, E.I., Anagnostou, E.N., 2020. Quantitative assessment of annual maxima, peaks-over-threshold and multifractal parametric approaches in estimating intensity-duration-frequency curves from short rainfall records. J. Hydrol. 589, 125151. https://doi.org/10.1016/j.jhydrol.2020.125151.

Fischer, S., Dallan, E., Fiori, A., Grimaldi, S., Kochanek, K., Prieto, C., Reis Jr., D.S., Volpi, E., 2025. Hydrological design in the HELPING decade–inspiring the community to innovate the hydrological design concept. Hydrol. Sci. J. 70 (3), 375–389. https://doi.org/10.1080/02626667.2024.2436634.

Gires, A., Tchiguirinskaia, I., Schertzer, D., 2020. Blunt extension of discrete universal multifractal cascades: development and application to downscaling. Hydrol. Sci. J. 65 (7), 1204–1220. https://doi.org/10.1080/02626667.2020.1736297.

Grimaldi, S., Volpi, E., Langousis, A., Michael, Papalexiou S., De Luca, D.L., Piscopia, R., Nerantzaki, S.D., Papacharalampous, G., Petroselli, A., 2022. Continuous hydrologic modelling for small and ungauged basins: A comparison of eight rainfall models for

sub-daily runoff simulations. J. Hydrol. 610, 127866. https://doi.org/10.1016/j.jhydrol.2022.127866.
Koutsoyiannis D. Stochastics of Hydroclimatic Extremes—A Cool Look at Risk (2022) Kallipos Open Academic Editions: Athens, Greece; 346p. ISBN 978-618-85370-0-2.
Langousis, A., Veneziano, D., 2007. Intensity-duration-frequency curves from scaling representations of rainfall. Water Resour. Res. 43 (2). https://doi.org/10.1029/2006WR005245.
Langousis, A., Veneziano, D., Furcolo, P., Lepore, C., 2009. Multifractal rainfall extremes: Theoretical analysis and practical estimation. Chaos Solitons Fractals 39 (3), 1182–1194. https://doi.org/10.1016/j.chaos.2007.06.004.
Northrop, P.J., 2024. Stochastic Models of Rainfall. Annu. Rev. Stat. Appl. 11 (1), 51–74. https://doi.org/10.1146/annurev-statistics-040622-023838.
Papalexiou, S.M., 2018. Unified theory for stochastic modelling of hydroclimatic processes: Preserving marginal distributions, correlation structures, and intermittency. Adv. Water Resour. 115, 234–252. https://doi.org/10.1016/j.advwatres.2018.02.013.
Papalexiou, S.M., Serinaldi, F., Porcu, E., 2021a. Advancing Space-Time simulation of Random Fields: from Storms to Cyclones and beyond. Water Resour. Res. 57, e2020WR029466. https://doi.org/10.1029/2020WR029466.
Papalexiou, S.M., Serinaldi, F., Strnad, F., Markonis, Y., Shook, K., 2021b. CoSMoS: complete Stochastic Modelling solution. Retrieved from R Package Version 2 (1). https://CRAN.R-project.org/package=CoSMoS.
Papalexiou, S.M., 2022. Rainfall Generation Revisited: introducing CoSMoS-2s and advancing Copula-based Intermittent Time Series Modeling. Water Resour. Res. 58, e2021WR031641. https://doi.org/10.1029/2021WR031641.
Papalexiou, S.M., Serinaldi, F., Clark, M.P., 2023. Large-Domain Multisite Precipitation Generation: Operational Blueprint and Demonstration for 1,000 Sites. Water Resour. Res. 59, e2022WR034094. https://doi.org/10.1029/2022WR034094.
Park, J., Cross, D., Onof, C., Chen, Y., Kim, D. (2021) A simple scheme to adjust Poisson cluster rectangular pulse rainfall models for improved performance at sub-hourly timescales, Journal of Hydrology, Vol. 598, 126296, ISSN 0022-1694, https://doi.org/10.1016/j.jhydrol.2021.126296.
Pui, A., Sharma, A., Mehrotra, R., Sivakumar, B., Jeremiah, E., 2012. A comparison of alternatives for daily to sub-daily rainfall disaggregation. J. Hydrol. 470–471, 138–157. https://doi.org/10.1016/j.jhydrol.2012.08.041.
Read, L.K., Vogel, R.M., 2015. Reliability, return periods, and risk under nonstationarity. Water Resour. Res. 51 (8), 6381–6398. https://doi.org/10.1002/2015WR017089.
Tyralis, H., Langousis, A., 2019. Estimation of intensity–duration–frequency curves using max-stable processes. Stoch. Env. Res. Risk A. 33 (1), 239–252. https://doi.org/10.1007/s00477-018-1577-2.
Veneziano, D., Lepore, C., Langousis, A., Furcolo, P., 2007. Marginal methods of intensity-duration-frequency estimation in scaling and nonscaling rainfall. Water Resour. Res. 43 (10), W10418. https://doi.org/10.1029/2007WR006040.
Volpi, E., Grimaldi, S., Aghakouchak, A., Castellarin, A., Chebana, F., Papalexiou, S.M., Aksoy, H., Bárdossy, A., Cancelliere, A., Chen, Y., Deidda, R., Haberlandt, U., Eris, E., Fischer, S., Francés, F., Kavetski, D., Rodding, K.T., Kochanek, K., Langousis, A., Mediero, O.L., Montanari, A., Nerantzaki, S.D., Ouarda, T.B.M.J., Prosdocimi, I., Ragno, E., Rajulapati, C.R., Requena, A.I., Ridolfi, E., Sadegh, M., Schumann, A., Sharma, A., 2024. The legacy of STAHY: milestones, achievements, challenges, and open problems in statistical hydrology. Hydrol. Sci. J. 69 (14), 1913–1949. https://doi.org/10.1080/02626667.2024.2385686.