



Virtual Rain: A Unified Toolkit for High-Resolution Rainfall Simulation and Disaggregation

Cappelli Francesco¹, Salvatore Grimaldi¹, Andrea Petroselli², Emanuele Santinami¹

¹Department for Innovation in Biological, Agro-Food and Forest Systems, Tuscia University, Viterbo, 01100, Lazio, Italy

5 ²Department of Agriculture and Forest Sciences, Tuscia University, Viterbo, 01100, Lazio, Italy

Correspondence to: Cappelli Francesco (francesco.cappelli@unitus.it)

Abstract. Stochastic simulation models are essential for investigating hydrological processes and supporting water resource management. In this study, we introduce Virtual Rain, a two-step toolkit that first generates synthetic daily rainfall time series and then disaggregates them to a user-defined temporal resolution. The toolkit builds on the approaches recently proposed by
10 some of the authors, ensuring a realistic representation of rainfall dynamics while preserving key statistical properties. The framework is implemented through a set of Python and R routines designed to facilitate practical application. In addition to providing observed rainfall time series and the scaling exponent n (Intensity–Duration–Frequency slope), users can configure key modelling components, including the marginal distribution, autocorrelation structure, number of lags, and target temporal resolution. The routines generate both graphical and quantitative outputs, enabling direct comparison between observed and
15 simulated series. Virtual Rain performance is evaluated through a real-world case study, demonstrating satisfactory accuracy and highlighting the robustness, flexibility, and transferability of the proposed toolbox. The toolkit is also available through an interactive web-based platform, facilitating its use by a broad range of users.

1 Introduction

Reliable hydrological modelling, flood risk assessment, and climate variability analyses critically depend on the availability
20 of long-term, high-resolution rainfall records (Grimaldi et al., 2022; Northrop, 2024; Volpi et al., 2024; Fischer et al., 2025). However, sub-daily precipitation observations are often limited in both spatial coverage and temporal continuity, limiting their use for statistically robust analyses and long-term risk assessments. This limitation also hinders the development, calibration, and validation of stochastic precipitation models, which require extensive datasets to reliably reproduce the multiscale structure of precipitation processes (Onof and Arnbjerg-Nielsen, 2009; Volpi et al., 2024).
25 In contrast, hydrologists typically have access to much longer daily precipitation time series, often spanning several decades, as well as non-continuous observations of extreme precipitation events with durations of less than a day (e.g., 1, 3, 6, 12, and 24 hours). These data form the basis of widely used tools such as intensity-duration-frequency (IDF) curves, which play a central role in flood analysis and infrastructure design (Koutsoyiannis, 2025). Exploiting the combined information contained in daily time series and IDF curves, therefore, represents a promising avenue for reconstructing precipitation dynamics at sub-



30 daily scales. However, despite the wide range of stochastic precipitation generators and disaggregation techniques proposed in the literature, choosing an appropriate model for a given application remains a challenge (Koutsoyiannis et al., 2003).

In this paper, we introduce Virtual Rain, a unified toolkit for simulating daily precipitation time series and subsequently disaggregating them to arbitrary temporal resolutions while preserving key statistical features. The toolkit is based on a recently introduced framework (Cappelli et al. 2026) that integrates (Module 1) a daily precipitation simulation model based on the univariate approaches of Papalexiou (2022) and Cappelli et al. (2024), and (Module 2) a sub-daily disaggregation module based on the multifractal cascade model proposed by Cappelli et al. (2025b). Specifically, daily precipitation is simulated using CoSMoS-2s (Papalexiou, 2022), a two-state intermittent model within the broader CoSMoS framework (Papalexiou, 2018), which separates marginal and dependence structures and accurately reproduces intermittency, marginal distribution, and autocorrelation structure (ACS) at the daily scale. The disaggregation module is based on a multiplicative random cascade (MRC) model based on multifractal theory, which exploits the scale-invariant properties of precipitation and provides a parsimonious parameterization using IDF curves (Langousis et al., 2009).

A key advantage of the proposed framework is that it eliminates the need for continuous sub-daily observations during calibration. Daily precipitation data are used to calibrate the simulation module, while IDF curves provide the information needed for disaggregation (Cappelli et al., 2026). This significantly broadens the framework's applicability, particularly in data-sparse regions, and addresses a major limitation of existing approaches that rely on high-resolution observations.

Another important feature of Virtual Rain is its flexibility: the two modules can be used jointly or independently, depending on data availability and modelling objectives. When long-duration daily time series are not available, Module 1 can be used to generate synthetic series of arbitrary length, which can then be disaggregated using Module 2. Conversely, when reliable daily observations exist, users can directly apply the disaggregation module without prior simulation.

50 The proposed toolkit is implemented through modular routines compatible with Python and R environments. In this study, the daily simulation module is implemented in Jupyter Notebook (Pérez and Granger, 2007), while the disaggregation module is developed in R. These open-source environments support interactive workflows and are accessible without local installation such as Anaconda or Binder. Virtual Rain is also available through an interactive web-based platform, facilitating its use by a broad range of users.

55 The proposed framework is validated using a real-world case study involving a single rain gauge station. The results demonstrate its ability to accurately reproduce both daily and 15-minute rainfall time series. The remainder of the paper is structured as follows: Section 2 presents the methodology; Section 3 describes the Virtual Rain framework; Section 4 presents the case studies; Section 5 concludes while in Appendix an overview of the web-based platform is provided.

2 Methodology in brief

60 Reliable hydrological modelling, flood risk assessment, and climate variability analyses critically depend on the availability of long-term, high-resolution rainfall records (Grimaldi et al., 2022; Northrop, 2024; Volpi et al., 2024; Fischer et al., 2025).



The Virtual Rain framework consists of two modules (Figure 1): (i) a daily rainfall simulation module based on CoSMoS-2s and (ii) a multifractal temporal disaggregation module based on an MRC model. Depending on data availability and modelling objectives, the two modules can be used either jointly or independently.

65 The methodological details of the daily simulation and disaggregation procedures are described in Sections 2.1 and 2.2, respectively.

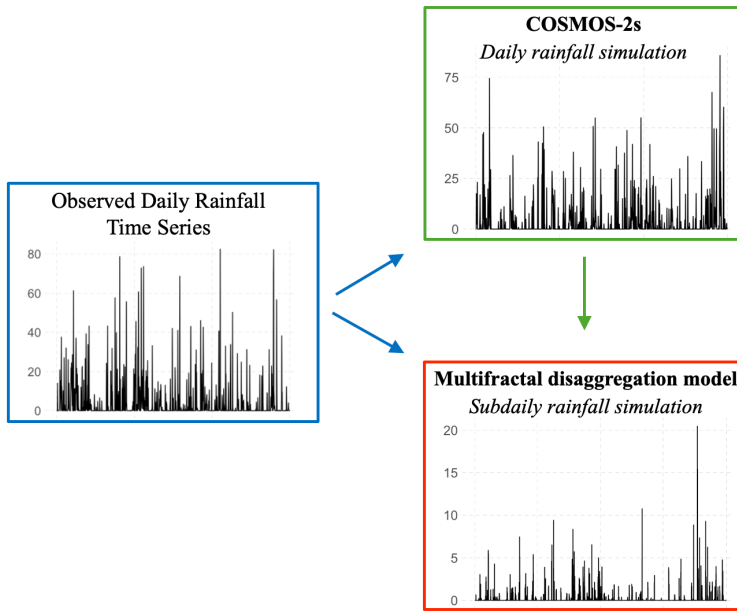


Figure 1: Virtual Rain workflow. Green box: Module 1; red box: Module 2.

2.1 Module 1 – Daily Rainfall Simulation Model

70 Daily rainfall time series are simulated using CoSMoS-2s, a two-state intermittent stochastic model (Papalexiou, 2022; Cappelli et al., 2024). In the single-site configuration adopted here, the model is structured around two coupled components: (i) State 1, a binary wet/dry occurrence process and (ii) State 2, a continuous intensity process defined for wet days (Figure 2). Since rainfall is usually affected by seasonality, it is necessary to group data into subsets (i.e., $m = 1, 2, \dots, 12$ seasons), with the aim of modeling rainfall occurrence and intensity within homogeneous samples. The two processes are then modeled

75 separately. In State 1, the observed daily rainfall is transformed into a binary process $B(t)$. For each season, occurrence probabilities p_0 and empirical ACSs ($\hat{\rho}_B(\tau)$) are estimated and subsequently mapped into Gaussian correlations through the transformation: $\rho_Z(\tau) = 1 - (1 - \rho_B(\tau)^{c_{p_0}})^2$. Temporal dependence is then reproduced using a Gaussian autoregressive (AR) model of order p , and simulated values are converted back into wet/dry states using seasonally varying thresholds. Specifically,

$$80 \quad b_m(t) = \begin{cases} 0, & z_m(t) \leq z_{p_0,m}(t) \\ 1, & z_m(t) > z_{p_0,m}(t) \end{cases}, \quad (1)$$



where $z_{p_{0,m}} = \Phi^{-1}(p_{0,m})$ and Φ^{-1} is the inverse of the cumulative distribution function of the standard normal distribution $\mathcal{N}(0, 1)$.

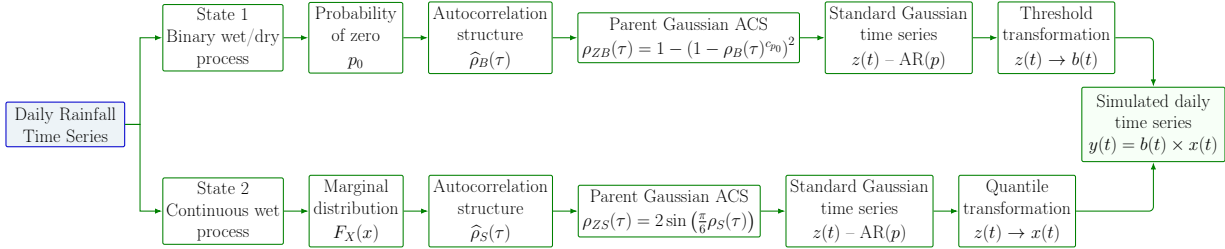


Figure 2: Workflow of Daily Rainfall Simulation Model. Equations details are described in the text.

85 State 2 is devoted to the continuous process $X(t)$ describing the wet-day rainfall amounts. For each season, suitable parametric distributions (e.g., 2- and 3-parameter Gamma, Generalized Gamma, Burr III, Burr XII, and 2- and 3-parameter Weibull distributions) are fitted to non-zero rainfall values, and parametric ACSs are used to model the empirical Spearman ACS ρ_S computed from consecutive wet-day sequences.

These correlations are transformed into Gaussian equivalents through the parent gaussian ACS
90 $\rho_{ZS}(\tau) = 2 \sin\left(\frac{\pi}{6} \rho_S(\tau)\right)$ and reproduced through an autoregressive model of order p . The simulated Gaussian values are then mapped back to rainfall intensities via inverse transformation by applying $x(t) = Q_X(\Phi_Z(z(t)))$ where Q_X is the quantile function of the fitted rainfall distribution and Φ_Z is the standard Gaussian cumulative distribution function.

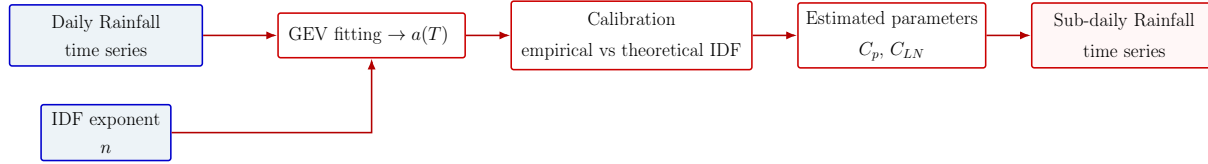
The final synthetic series is obtained by combining the occurrence and intensity components, allowing the model to reproduce key features such as intermittency, marginal distributions, and temporal persistence (Cappelli et al., 2026).

95 2.2 Module 2 – Sub-Daily Rainfall Disaggregation Model

This module is designed to disaggregate a simulated (by Module 1) or observed daily rainfall time series into high-resolution sub-daily time series using a multifractal random cascade (MRC) model (Cappelli et al., 2025a,b) (Figure 3). The approach assumes that rainfall intensity evolves as a multiplicative cascade process, in which the mean intensity at each temporal scale is redistributed through random weights with prescribed statistical properties. These weights follow a beta-lognormal distribution. The MRC model is characterized by three parameters: C_β , controlling intermittency and the proportion of dry
100 intervals; C_{LN} , governing lognormal variability and intensity fluctuations; and D , representing the outer-scale parameter that defines the range of multifractal behavior. Rather than calibrating these parameters using high-resolution rainfall observations, the implemented workflow derives them from IDF relationships, following Cappelli et al. (2025a). IDF curves are typically expressed as $i(d, T) = a(T)d^{n-1}$ where $i(d, T)$ is the rainfall intensity associated with duration d with return period T , and n is
105 a scaling exponent constant with $0 < n < 1$, and $a(T)$ is a function independent of d . By matching theoretical multifractal IDF relationships to empirical IDF curves, it is possible to estimate C_β , C_{LN} , and D without relying on sub-daily calibration



data. The cascade procedure is applied independently to each simulated day, ensuring exact preservation of daily rainfall totals while generating physically consistent sub-daily variability.



110 **Figure 3: Workflow of Sub-Daily Rainfall Disaggregation Model. Blue boxes represent the procedure input.**

3 Virtual Rain: a unified toolkit

In line with the methodological framework described in Section 2, Virtual Rain is organized into two modules: Module 1, dedicated to daily rainfall simulation using CoSMoS-2s, and Module 2, dedicated to temporal disaggregation using the MRC model. Each module consists of a set of routines that guide the user through data analysis, model calibration, simulation, validation, and generation of final outputs.

Table 1 summarizes the main routines implemented in the toolkit and their respective functions.

Module	Routine Name	Goal
M1: Daily rainfall simulation	M1.1: <code>plot_rainfall_timeseries()</code>	Exploratory analysis of observed rainfall data
	M1.2: <code>MargDistr_Fitting()</code>	Estimation of seasonal rainfall occurrence and marginal distributions
	M1.3: <code>ACS_Fitting()</code>	Estimation of occurrence and intensity autocorrelation structures
	M1.4: <code>SimulatedMultiple_DailyTS()</code>	Generation of multiple synthetic time series for diagnostic analyses
	M1.5: <code>Diagnostic_plots()</code>	Comparison between observed and simulated rainfall statistics
	M1.6: <code>SimulatedTS()</code>	Generation of the final synthetic daily rainfall series
M2: Sub-hourly disaggregation	M2.1: <code>GEV_Fitting()</code>	Estimation of extreme rainfall statistics and scaling parameters
	M2.2: <code>MRC_calibration()</code>	Calibration of the multifractal cascade model using IDF information
	M2.3: <code>Ts_Disaggregation()</code>	Generation of high-resolution rainfall series through temporal disaggregation

Table 1: Main routines implemented in Virtual Rain.

The following sections describe each routine, including its goal, required inputs, main operations, and outputs.

120 3.1 Implementation of the Daily Rainfall Simulation Module

This module includes six routines designed to guide the user through the complete workflow of stochastic daily rainfall simulation, from exploratory analysis of the observed series to the generation and validation of synthetic rainfall data. The



125 routines allow the user to reproduce key statistical rainfall properties, including marginal distributions, intermittency, seasonality, and temporal dependence, while providing graphical and quantitative tools to support model selection and validation. The routines are described below.

*** M1.1: `plot_rainfall_timeseries(data)`**

Goal: The routine provides a preliminary exploratory analysis of the observed rainfall series and helps the user visually inspect the main rainfall characteristics before model fitting.

Input:

- 130
- **data:** observed daily rainfall time series.

Operations: The function preprocesses the input series and produces three diagnostic plots:

1. Daily rainfall during the first observation year, useful for visualizing intermittency and wet/dry alternation;
 2. Annual maximum daily rainfall, allowing the user to inspect interannual variability and extreme events;
 3. Monthly seasonal components, highlighting the seasonal rainfall regime and supporting the choice of the seasonal grouping.
- 135

Outputs:

- plots of rainfall time series, annual maxima and monthly seasonal components.

*** M1.2: `MargDistr_Fitting(data, seasonality_option = (1,2,3,4,6,12), start_month=1)`**

Goal: The routine identifies the probability distribution that best describes rainfall intensities in each season.

140 *Inputs:*

- **data:** observed daily rainfall time series.
- **seasonality_option:** number of seasonal groups used to partition the data (e.g., annual = 1, semiannual = 6, quarterly = 3, monthly = 12);
- **start_month:** starting month of the seasonal partition.

145 *Operations:* For each seasonal group, the routine:

1. estimates the probability of dry days (p_0);
 2. fits seven candidate distributions (2- and 3-parameter Gamma, Generalized Gamma, Burr III, Burr XII, and 2- and 3-parameter Weibull distributions) to non-zero rainfall values;
 3. estimates model parameters through the Nelder–Mead optimization algorithm by minimizing discrepancies between empirical and theoretical exceedance probabilities.
- 150

Outputs:

- Estimated parameters for all candidate distributions;
- Diagnostic plots comparing empirical and fitted distributions;



- Summary table containing Relative Mean Absolute Error values (RMAE) estimated on empirical and fitted distributions.

The resulting object (resM1.2) is used as input by the `SimulatedMultiple_DailyTS()` and `SimulatedTS()` routines.

*** M1.3:** `ACS_Fitting(data, lags, seasonality_option = (1,2,3,4,6,12), start_month = 1)`

Goal: Estimate the autocorrelation structure of rainfall occurrence and intensity processes.

Inputs:

- **data:** observed rainfall time series;
- **lags:** maximum lag considered in the autocorrelation analysis;
- **seasonality_option:** seasonal grouping window (same of M1.2);
- **start_month:** starting month of the seasonal partition.

Operations: For each seasonal group and both the binary occurrence process $B(t)$ and the continuous intensity processes $X(t)$, the routine:

1. computes empirical Spearman autocorrelations;
2. fits candidate parametric ACSs (Pareto II, Weibull, Burr XII);
3. estimates model parameters by minimizing the Mean Absolute Error (MAE) between empirical and theoretical ACS values.

Outputs:

- Estimated ACS parameters;
- Diagnostic plots comparing empirical and fitted ACS for each season and process.

The resulting object (resM1.3) is used as input by the `SimulatedMultiple_DailyTS()` and `SimulatedTS()` routines.

*** M1.4:** `SimulatedMultiple_DailyTS(data, n_simulations, resM1.2, resM1.3, p, seasonality_option, start_month, marg_distr = ('gamma', 'gamma3', 'ggamma', 'burrIII', 'burrXII', 'weibull', 'weibull3'), acf_distr_BT = ('pare-toII', 'weibull', 'burrXII', 'empirical'), acf_distr_CT = ('paretoII', 'weibull', 'burrXII', 'empirical'))`

Goal: Generate an ensemble of synthetic rainfall series for model evaluation and uncertainty assessment.

Inputs:

- **data:** observed rainfall time series;
- **n_simulations:** number of synthetic time series to generate;
- **resM1.2:** output from `MargDistr_Fitting()` routine;
- **resM1.3:** output from `ACS_Fitting()` routine;
- **p:** autoregressive order;



- 185
- `seasonality_option`: number of seasonal groups used in the calibration.
 - `start_month`: starting month of the seasonal partition.
 - `Marg_distr`: selected marginal distribution for rainfall intensities.
 - `Acf_distr_BT`, `acf_distr_CT`: selected ACS for the binary $B(t)$ and continuous $X(t)$ processes.

Operations: The routine generates multiple synthetic daily rainfall series having the same length as the observed dataset

190 *Outputs*:

- Ensemble of synthetic rainfall time series.

The resulting object (`resM1.4`) is used as input for the `Diagnostic_plots()` routine.

★ **M1.5**: `Diagnostic_plots(data, resM1.4)`

195 *Goal*: The routine evaluates the ability of the stochastic model to reproduce the main statistical characteristics of observed rainfall.

Inputs:

- `data`: observed rainfall time series;
- `resM1.4`: ensemble of synthetic rainfall series generated by `SimulatedMultiple_TS()` routine.

Operations: The routine compares observed statistics against simulated distributions using four diagnostic groups:

- 200
1. autocorrelation structure;
 2. seasonal component;
 3. ranked annual maxima;
 4. summary statistics, including dry-day probability (p_0), mean wet-day rainfall (mean), standard deviation (sd), extreme percentiles (P95, P99, P99.9).

205 *Outputs*:

- Diagnostic plots comparing observed statistics with the distribution of the simulated ensemble.

★ **M1.6**: `SimulatedTS(data, resM1.2, resM1.3, p, n_years, seasonality_option, start_month, marg_distr, acf_distr_BT = ('paretoII', 'weibull', 'burrXII', 'empirical'), acf_distr_CT = ('paretoII', 'weibull', 'burrXII', 'empirical'))`

210 *Goal*: Generates a synthetic daily rainfall time series using the selected modelling configuration.

Inputs:

- `data`: observed rainfall time series;
 - `resM1.2` and `resM1.3`: output from `MargDistr_Fitting()` and `ACS_Fitting()` routines;
 - `p`: autoregressive order;
 - `n_years`: desired simulation length
- 215



- `seasonality_option`: number of seasonal groups used in the calibration.
- `start_month`: starting month of the seasonal partition.
- `marg_distr`: selected marginal distribution for rainfall intensities.
- `acf_distr_BT`, `acf_distr_CT`: selected ACS for the binary $B(t)$ and continuous $X(t)$ processes.

220 *Operations*: The routine simulates the wet/dry occurrence process and the wet-day intensity process separately using the calibrated seasonal probabilities, marginal distributions, and autocorrelation structures. The two components are subsequently combined to generate a continuous synthetic daily rainfall series of the specified length.

Outputs:

- Synthetic daily rainfall time series with user-defined duration.

225 3.2 Implementation of the Sub-Hourly Rainfall Disaggregation Module

This module includes three routines designed to guide the user through the complete workflow of rainfall temporal disaggregation, from the estimation of extreme rainfall properties to the generation and validation of high-resolution synthetic rainfall series. The routines allow the user to disaggregate daily rainfall while preserving scale-invariant properties and reproducing the imposed IDF scaling exponent through a stochastic MRC approach.

230 The routines are described below.

★ **M2.1:** `GEV_Fitting(data, starting_day = '2000-01-01', num_of_days_for_truncation = 183, n_par = 0.3)`

Goal: Estimate the extreme rainfall statistics and scaling parameters required for MRC calibration.

Inputs:

- 235
- `data`: observed rainfall time series;
 - `starting_day`: starting date of the time series used to construct the temporal index;
 - `num_of_days_for_truncation`: threshold used to exclude incomplete initial or final years;
 - `n_par`: scaling exponent (n), representing the slope of IDF curves.

Operations: The routine:

- 240
1. extracts annual maximum rainfall time series;
 2. transforms annual maxima into scaling-based a -IDF coefficients using the imposed scaling exponent n ;
 3. fits a GEV distribution to the a -coefficients using three estimation methods: Maximum Likelihood (*mle*), Probability Weighted Moments (*pwm*), L-moments (*L-mom*);
 4. estimates mean annual rainfall intensity;
 - 245 5. computes a -coefficients associated with return periods ranging from 2 to 999 years using both: parametric approaches (GEV-based) and empirical estimates.



Outputs:

- GEV parameter estimates for the three fitting methods;
- a -coefficients associated with different return periods;
- Annual maxima time series;
- Diagnostic plots including: 1) daily rainfall during the first year, 2) annual maximum daily rainfall over the observation period, 3) Gumbel plots comparing empirical and fitted distributions.

The resulting object (`resM2.1`) is used as input for the `MRC_calibration()` routine.

```
* M2.2: MRC_calibration(resM2.1, lower_bounds = c(Cb_Lb, Cln_Lb, D_Lb), upper_bounds = c(Cb_Ub,
Cln_Ub, D_Ub), Return_periods = c(2, 4, 6, 8, 10), n_par, method = c('mle', 'pwm', 'L-mom',
'empirical'), Num_Iterations, a_Rp = NULL)
```

Goal: Calibrate the MRC model parameters required to reproduce the imposed rainfall scaling properties and IDF relationships.

Inputs:

- `resM2.1`: output from `GEV_Fitting()` routine.
- `lower_bounds`, `upper_bounds`: admissible parameter ranges for calibration.
- `Return_periods`: return periods used in the calibration.
- `n_par`: scaling exponent (n) of IDF curves.
- `method`: method used to estimate the a -coefficients (*mle*, *pwm*, *L-mom*, or *empirical*).
- `Num_Iterations`: maximum number of optimization iterations;
- `a_Rp` (optional): user-defined a -coefficients.

Operations: The routine estimates the MRC parameters (C_β , C_{ln} , D) through numerical optimization. The optimal parameter set is defined by minimizing discrepancies between theoretical and empirical IDF curves associated with the imposed scaling exponent n and the selected a -coefficients for the chosen return periods.

Outputs:

- Calibrated MRC parameters (C_β , C_{ln} , D);
- Objective function value at convergence;
- Diagnostic plots comparing empirical and theoretical IDF curves.

The resulting object (`resM2.2`) is used as input for the `Ts_Disaggregation()` routine.

```
* M2.3: Ts_Disaggregation(resM2.2, ending_time_resolution = 15, N_simulations = NULL,
preserve_rainfall_volume = c('Yes', 'No'), seed = c('Yes', 'No'))
```



Goal: Generate high-resolution rainfall series through temporal disaggregation of daily rainfall using the calibrated MRC model.

Inputs:

- 280 • `resM2.2`: output from `MRC_calibration()` routine.
- `ending_time_resolution`: target temporal resolution (1–60 min).
- `N_simulations`: number of disaggregated time series to generate.
- `preserve_rainfall_volume`: option to preserve daily rainfall totals.
- `seed`: option to initialize the random number generator for reproducibility.

285 *Operations:* The routine applies the calibrated MRC model to disaggregate daily rainfall into sub-daily intervals. The cascade is applied only to wet days, recursively redistributing daily rainfall volumes while preserving the prescribed scaling properties and, optionally, the original daily totals. One or more stochastic realizations can be generated at the selected temporal resolution.

Outputs:

- 290 • Disaggregated rainfall time series;
- Plots of the observed daily time series and the corresponding disaggregated one for the first simulated year;
- Summary tables reporting key rainfall statistics, ACS coefficients, annual maxima over multiple durations, and IDF-related indicators for each simulation.

The resulting outputs allow users to visually inspect the generated series and evaluate whether the simulated scaling and IDF
295 characteristics are consistent with the calibration targets, namely the prescribed scaling exponent n and the prescribed α -coefficients for the considered return periods.

This module includes six routines designed to guide the user through the complete workflow of stochastic daily rainfall simulation, from exploratory analysis of the observed series to the generation and validation of synthetic rainfall data. The routines allow the user to reproduce key statistical rainfall properties, including marginal distributions, intermittency,
300 seasonality, and temporal dependence, while providing graphical and quantitative tools to support model selection and validation. The routines are described below.

3.3 Interactive digital platform

The routines described in the previous sections are mainly intended for advanced users with programming experience in Python and R. To make the proposed toolkit accessible to a wider scientific and technical audience, an interactive web-based platform,
305 *Virtual Rain* (<https://www.hydrolab.unitus.it/virtualrain>), has been designed and developed. Its main interfaces are reported in the Appendix.

The platform follows the same logical structure as the routines presented above, guiding users' step by step through the modelling workflow, from input data processing to model calibration, simulation, and diagnostic evaluation. The main model



parameters introduced in the methodological framework can be modified directly by the user, providing a flexible environment for both operational applications and methodological exploration.

By combining computational routines with an intuitive graphical interface, Virtual Rain reduces the technical barrier associated with coding while preserving the flexibility of the proposed modelling framework. The platform therefore supports not only practical model implementation, but also a clearer understanding of the underlying stochastic processes, parameter choices, and diagnostic outputs. Overall, Virtual Rain represents a comprehensive tool that bridges methodological flexibility and practical usability, making the proposed framework suitable for research applications, advanced professional analyses, and educational purposes.

4 Case study: Arno Rain

To illustrate the effectiveness and flexibility of the proposed framework, we present a case study based on a daily rainfall time series (Station code TOS01004779) selected from a dataset of more than 70 rain gauges distributed across the Arno River Basin. The Arno is one of the longest rivers in Italy and it is the main river system in Tuscany, Central Italy. The dataset spans the period from January 1, 2001, to December 31, 2020.

The proposed modelling framework and its underlying procedures have already been extensively tested in previous studies on several rainfall datasets and climatic conditions, yielding satisfactory results in terms of reproducing rainfall intermittency, temporal dependence, scaling properties, and extreme-event statistics (Cappelli et al., 2024; 2025a; 2025b; 2026). Building upon these previous validations, the present case study is intended to illustrate the practical implementation and operational use of the two modules within the Virtual Rain platform.

4.1 Results of the Daily Rainfall Simulation Module

As described in Section 3.1, The `plot_rainfall_timeseries()` routine of the Daily Rainfall Simulation Module provides an overview of the observed daily time series (Figure 4).

It produces three plots: daily rainfall for the first year, annual maximum rainfall over the full observation period, and monthly mean seasonal component. In particular, the plot of the seasonal component (third panel in Figure 4) provides a first assessment of the temporal variability of rainfall throughout the year, highlighting the presence of three main seasonal groups and supporting the selection of April as the starting month.

Non-zero rainfall values are modeled using seven candidate distributions (Gamma, three-parameter Gamma, Generalized Gamma, Weibull, three-parameter Weibull, Burr III, and Burr XII). The optimal parameters distribution is selected by minimizing the distance between empirical and theoretical exceedance probabilities using the Nelder–Mead optimization algorithm. Model performance is evaluated using the RMAE computed over the upper tail (values above the 50th percentile), as lower values are generally well reproduced by all candidate distributions. The calibration results obtained using



MargDistr_Fitting() routine (Figure 5 and Table 2) indicate that the Gamma distribution provides the best overall performance.

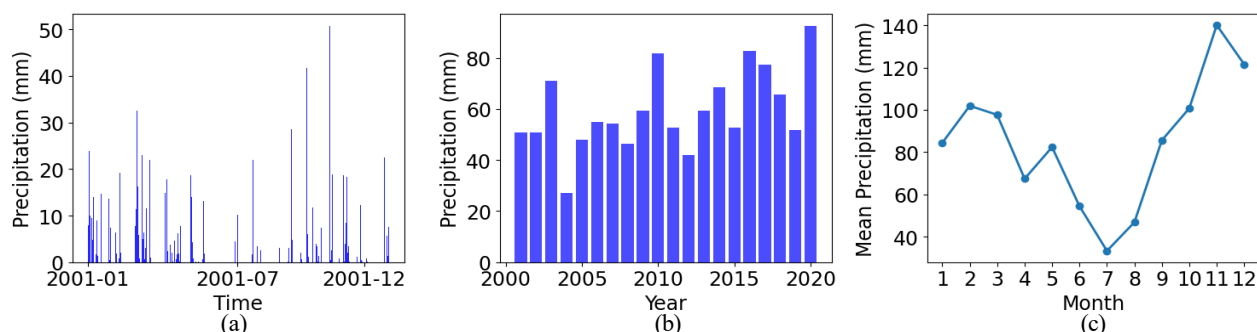


Figure 4: Overview of the input observed daily time series: (a) daily rainfall for the first year, (b) annual maximum daily rainfall, and (c) monthly mean seasonal component.

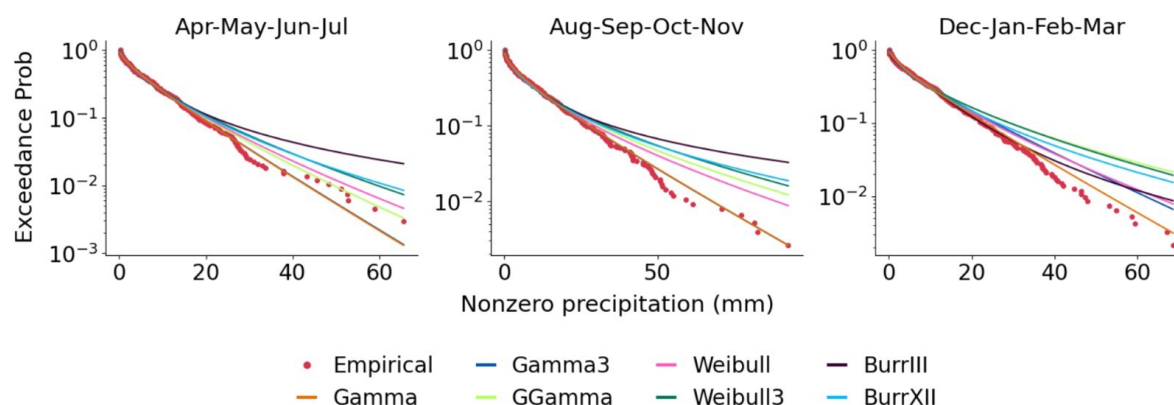


Figure 5: Calibration results of the seven candidate marginal distributions fitted on non-zero rainfall values.

Season	Gamma	Gamma3	GGamma	Weibull	Weibull3	BurrIII	BurrXII
Apr-May-Jun-Jul	0.052	0.053	0.081	0.105	0.160	0.251	0.150
Aug-Sep-Oct-Nov	0.046	0.048	0.191	0.143	0.256	0.356	0.252
Dec-Jan-Feb-Mar	0.059	0.167	0.353	0.148	0.347	0.087	0.242

Table 2: Relative Mean Absolute Error (RMAE) associated with the candidate marginal distributions for each seasonal group. Lower RMAE values indicate better agreement between empirical and theoretical exceedance probabilities, particularly for the upper tail of the rainfall distribution (values above the 50th percentile).

The autocorrelation structure is then modeled using the ACS_Fitting() routine, which estimates the Spearman ACS accounting for seasonality and distinguishing between $B(t)$ and $X(t)$ processes.

After selecting the maximum lag, parametric ACS models (Pareto II, Weibull, and Burr XII) are fitted separately for the two processes $B(t)$ and $X(t)$ by minimizing the MAE between empirical and parametric ACSs. Figure 6 compares the empirical and fitted ACSs, providing a visual assessment of the adequacy of the candidate models. For the binary process, both the Weibull and Pareto II ACS provide a satisfactory representation of the empirical Spearman ACS. For the continuous process,



all three candidate ACSs accurately reproduce the empirical Spearman ACS. In this case study application $p = 3$ and the empirical ACS are adopted.

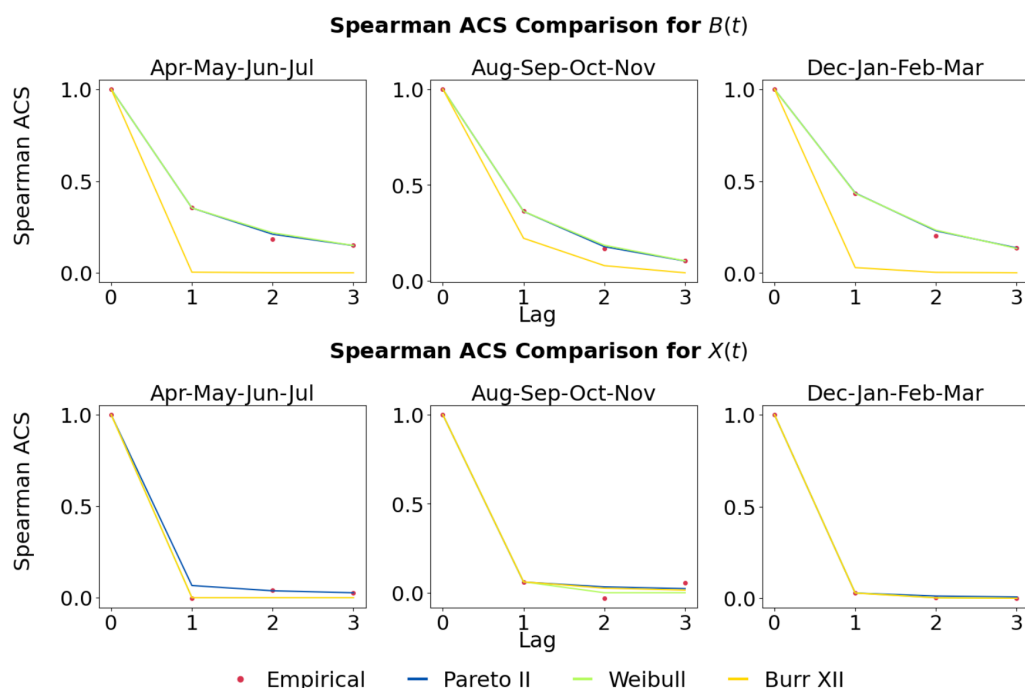


Figure 6: Comparison between empirical and fitted parametric ACSs for the $B(t)$ and $X(t)$ processes across the seasonal groups. The fitted curves are obtained by minimizing the Mean Absolute Error between empirical and theoretical Spearman ACSs.

Based on the results of the previous calibration steps, the selected marginal distributions and ACS models are subsequently used within the `SimulatedMultiple_DailyTS()` routine to generate an ensemble of 50 synthetic daily rainfall series having the same length as the observed record. The resulting ensemble is then provided as input to the `Diagnostic_plots()` routine, which evaluates model performance by comparing observed and simulated rainfall characteristics, including autocorrelation structure, seasonality, dry-day probability (p_0), mean and standard deviation of wet-day rainfall, and extreme percentiles. The results (Figure 7) show that the model successfully reproduces the main statistical properties of the observed data, including dry frequency and central distribution characteristics, with a slight overestimation in the upper tail. Figure 8 compares seasonal component and the Spearman ACS between observed and simulated time series. The model reproduces the main seasonal variability and the temporal dependence structure reasonably well across lags, although some discrepancies emerge in specific months. This behavior is expected, since the adopted seasonal partitioning groups three months together in order to increase the sample size available for distribution and autocorrelation fitting. In this framework, a slight reduction in



the accuracy of the seasonal-cycle reproduction is intentionally accepted in exchange for a more robust estimation of the statistical model parameters, which is considered more relevant for the simulation of daily rainfall series.

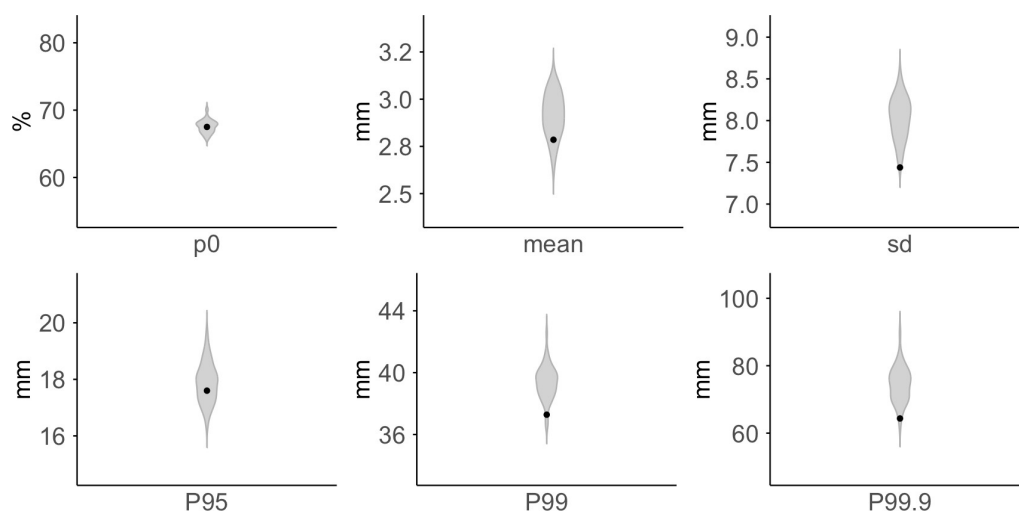


Figure 7: Comparison between observed (black dots) and 50 simulated (gray violin plots) summary statistics of non-zero rainfall: (a) probability of zero p_0 , (b) mean, (c) standard deviation, (d) 95th percentile, (e) 99th percentile, and (f) 99.9th percentile.

375 Figure 9 presents the comparison of ranked annual maxima at different durations. The observed values generally fall within the variability range of the simulated extremes, confirming the model's capability to reproduce the overall multi-day extreme rainfall behavior.

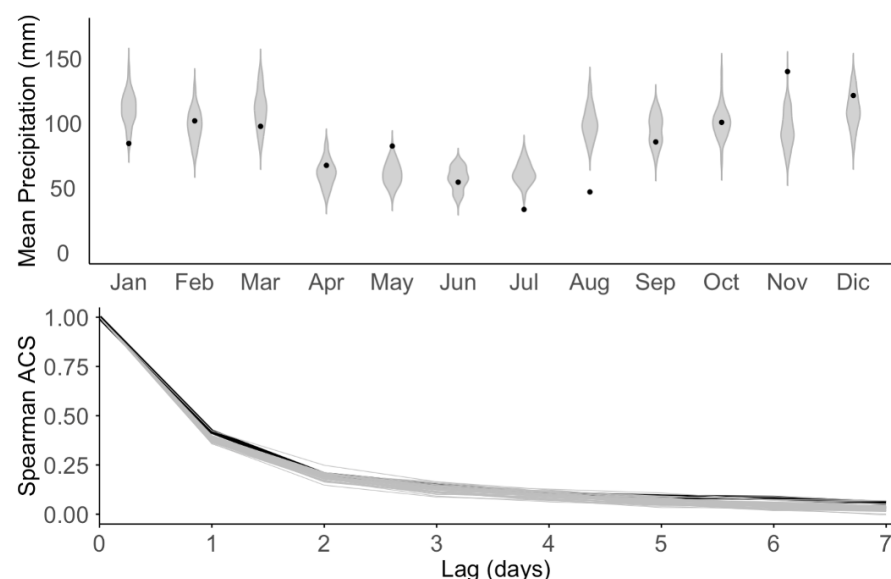


Figure 8: Comparison of: a) seasonal components between observed (black dots) and 50 simulated (gray violin plots) series (upper panel); b) Spearman autocorrelation functions between observed (black line) and 50 simulated series (gray lines) (bottom panel).

380

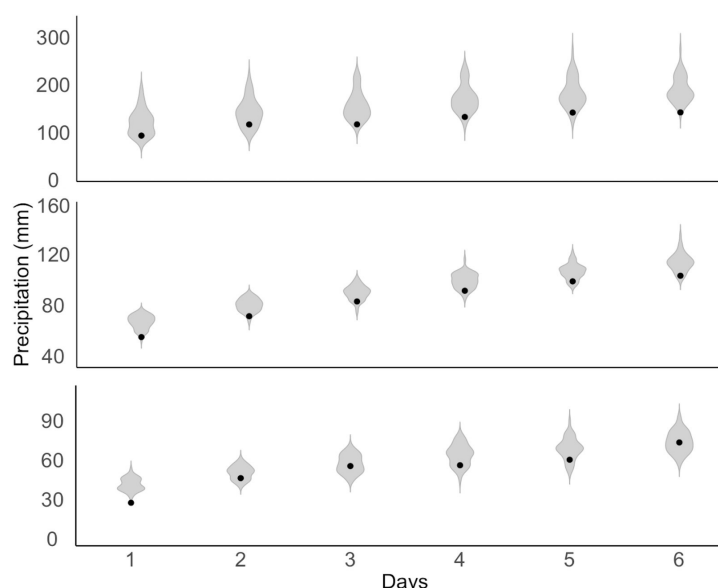


Figure 9: Comparison of (a) highest, (b) median, and (c) lowest ranked annual maxima between observed (black dots) and 50 simulated series (gray violin plots).

Overall, the simulated time series adequately reproduce the statistical properties of observed rainfall across multiple temporal scales.

If the model performance is deemed satisfactory, the user can proceed with the final simulation using the `SimulatedTS()` routine, specifying the desired length. The generated series can then be downloaded or directly used as input for Module 2 for temporal disaggregation.

Figures 10 and 11 present the outputs of `GEV_Fitting()` routine. Figure 10 provides a general overview of the rainfall dataset, including first year of the simulated daily rainfall series (panel a) and the annual maxima extracted from the entire simulated time series (panel b). Figure 11 shows the Gumbel plot that compares the performance of the alternative GEV fitting methods. The plot supports the selection of the most appropriate method for estimating the a -coefficients used in the MRC calibration. Both *mle* and *pwm* provide a good fit to the empirical data.

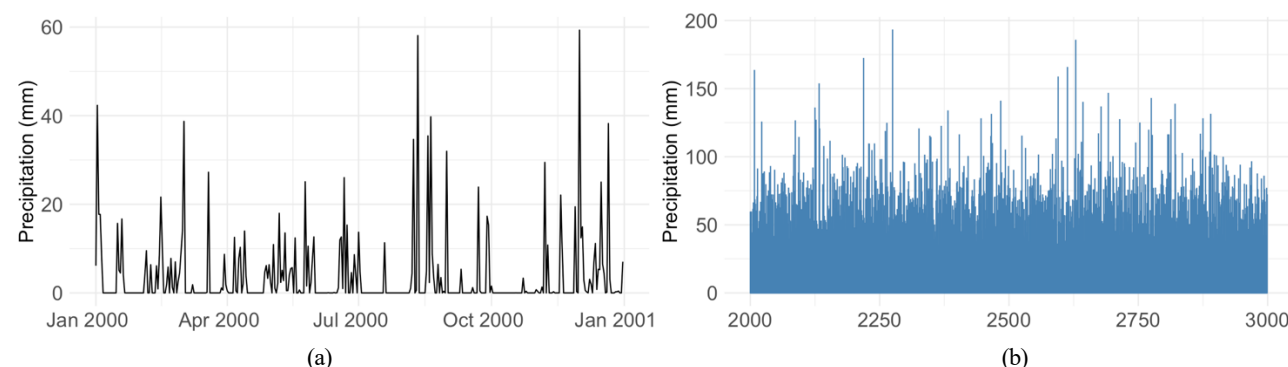


Figure 10: (a) First year of the simulated daily rainfall time series; (b) annual maximum of the simulated daily time series.

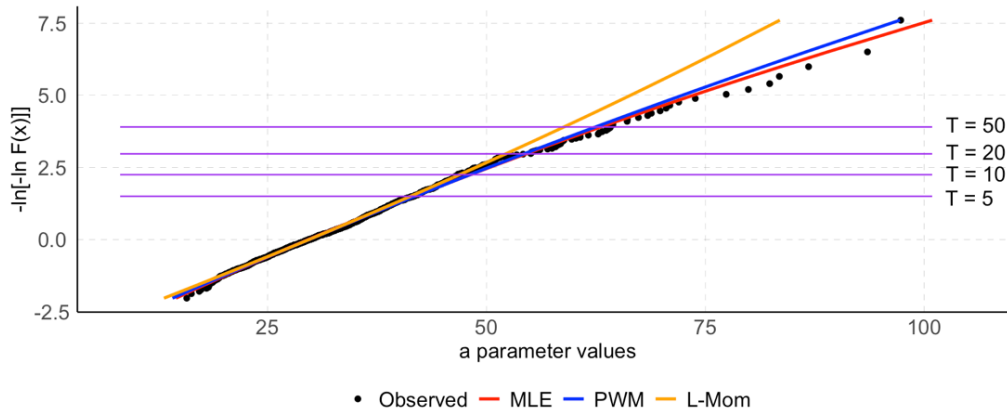


Figure 11: Gumbel plot comparing empirical and theoretical α -coefficients obtained from the fitted GEV distributions using Maximum Likelihood (MLE), Probability Weighted Moments (PWM), and L-moments (L-mom).

400 The `MRC_calibration()` routine is then used to calibrate the MRC model. In this study, return periods of 10, 20, 50, 100, and 200 years are considered, and the *mle* method is selected. Figure 12 presents the calibration results. Panel (a) compares the empirical and theoretical IDF curves obtained after calibration, while panel (b) reports the corresponding parameter estimates, including the GEV parameters, the calibrated MRC parameters (C_β , C_{Ln} , and D), the mean annual rainfall intensity, and the value of the objective function. The low objective function value, together with the close agreement between empirical and theoretical IDF curves shown in panel (a), indicates that the calibrated MRC model successfully reproduces the target IDF relationships over the selected range of return periods.

405

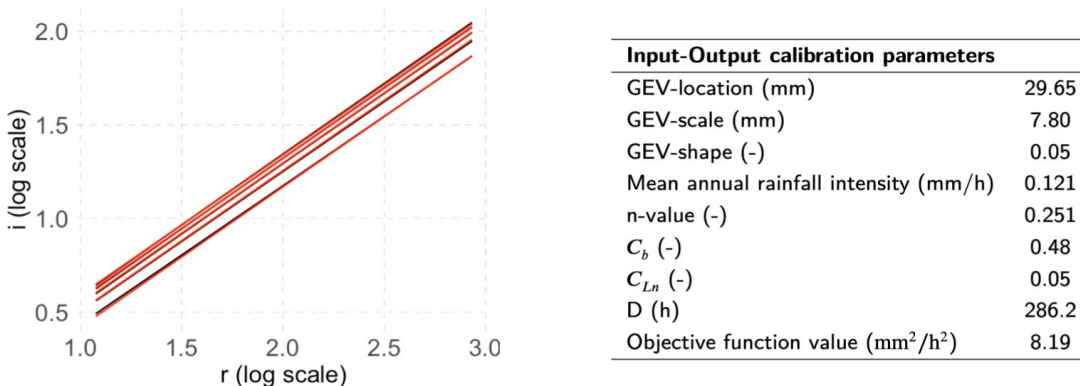


Figure 12: MRC model calibration: (a) comparison between empirical (red line) and theoretical IDF curves (black line); (b) estimated input and output parameters.

Once the calibration parameters have been estimated, they are used as input to the `Ts_Disaggregation()` routine, which implements the MRC disaggregation procedure. For illustrative purposes, the simulated daily rainfall time series is disaggregated to a 15-minute temporal resolution, generating a single stochastic realization. Figure 13 shows the resulting 15-minute rainfall series for the first simulated year.

410

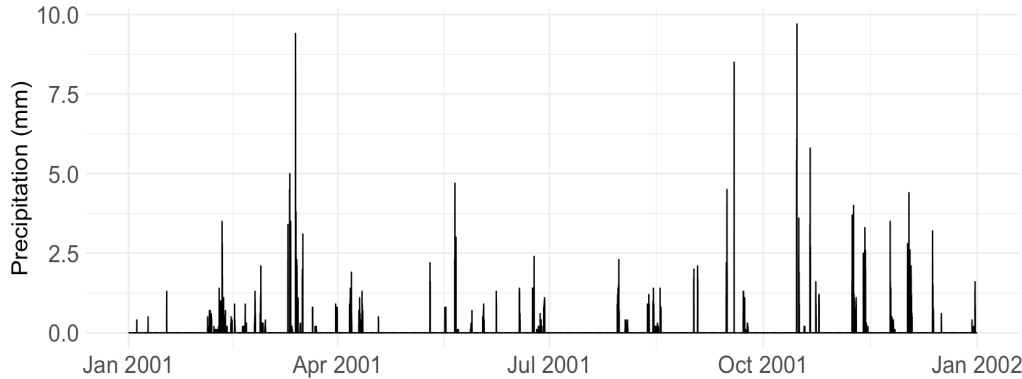


Figure 13: Disaggregated rainfall time series at 15-minute resolution (first year).

415 Finally, a comprehensive summary of the disaggregation results is provided in Table 3. The comparison between the observed
15-minute rainfall series (used here as benchmark and not for calibrating the model) and the disaggregated series shows that
the MRC model successfully reproduces the main statistical characteristics imposed during calibration. In particular, the
scaling exponent n , dry frequency (p_0), mean rainfall, and the central and upper portions of the rainfall distribution (e.g., P99
and P99.99) are reproduced with good accuracy. The autocorrelation structure is also satisfactorily preserved, although a slight
420 underestimation of the first autocorrelation coefficients can be observed.

Statistic	Observed Ts	Disaggr. Ts	Statistic	Observed Ts	Disaggr. Ts
<i>General statistics</i>			<i>Accumulated maxima</i>		
n	0.251	0.251	$\text{Max}_{30\text{min}}$	44.4	120.6
p_0	0.955	0.956	$\text{Max}_{1\text{h}}$	56.2	125.4
Mean	0.03	0.03	$\text{Max}_{3\text{h}}$	78.6	148.5
Stand. dev.	0.26	0.283	$\text{Max}_{6\text{h}}$	81.6	172.0
<i>Extreme values</i>			$\text{Max}_{12\text{h}}$	92.6	189.9
P99	0.8	0.8	$\text{Max}_{24\text{h}}$	92.8	195.8
P99.9	3	3.7	<i>IDF parameters</i>		
P99.99	9.6	9.5	a_{Rp10}	41.65	53.5
Max	29	120.6	a_{Rp20}	46.45	61.99
<i>Autocorrelation</i>			a_{Rp50}	-	77.97
$\rho(1)$	0.527	0.505	a_{Rp80}	-	84.99
$\rho(2)$	0.323	0.294	a_{Rp100}	-	88.29
$\rho(3)$	0.258	0.210			

Table 3: Comparison of summary statistics, ACS coefficients, accumulated maxima, and IDF parameters between the observed 20-year rainfall series at 15-minute resolution and the 1000-year synthetic rainfall series generated by the MRC disaggregation model.

As expected, larger differences emerge when considering extreme rainfall properties, due to the length differences between
the benchmark and the simulated time series. The disaggregated time series exhibits higher maximum rainfall accumulations
425 across all aggregation durations, and the empirically estimated a -coefficients are generally larger than those derived from the



observed record. Nevertheless, the key calibration targets are reproduced remarkably well. Both the dry-frequency (p_0) and the scaling exponent (n)—the latter being the only parameter directly used to calibrate the MRC model—are identical to those observed, confirming the ability of the two modules to preserve rainfall intermittency and scaling properties while remaining consistent with the target IDF relationships.

Overall, the results confirm that the calibrated MRC model provides a statistically consistent representation of sub-daily rainfall variability while preserving the prescribed scaling behavior and extreme-value characteristics.

5 Conclusion

In this paper, we introduced Virtual Rain, a two-module toolkit designed to integrate complementary methods for simulating rainfall time series at arbitrary resolutions. The toolkit is designed to generate univariate time series that reproduce key hydroclimatic features of observed data, including marginal distributions, autocorrelation structure, and intermittency.

A central aspect of this work is the development of a flexible and operational workflow capable of functioning under limited data availability. By combining daily rainfall observations with IDF curve information, the toolkit enables the reconstruction of sub-daily dynamics without requiring continuous high-resolution records.

The proposed approach has been assessed through a real-world case study, which confirms the reliability of the underlying methodological components—already extensively validated in recent studies—and demonstrates the effectiveness of the implemented routines. In particular, the modular structure of the toolkit provides advanced users with a flexible set of routines that allow full control over key modeling choices, enabling adaptation to different datasets, objectives, and hydroclimatic conditions.

Beyond the methodological contribution, this study also emphasizes usability and practical implementation. The proposed routines have been integrated into an interactive web-based platform that reproduces the logical structure of the modeling workflow, allowing users to explore, configure, and apply the methodology without requiring advanced programming skills. By enabling full control over the modeling parameters, the platform supports both operational use and a deeper understanding of the underlying processes and assumptions.

In this perspective, Virtual Rain can be seen not only as a computational tool, but also as an operational and educational environment, facilitating its adoption across research, professional, and teaching contexts. The availability of the routines and datasets further supports transparency, reproducibility, and future developments of the proposed approach.

Appendix A Virtual Rain platform: implementation of the case study workflow

This Appendix reports the main interface screens of the Virtual Rain digital platform used to replicate the case study presented in Section 4. The screenshots follow the same logical sequence adopted in the methodological workflow, guiding the user from the upload and preprocessing of the observed rainfall series to stochastic simulation, diagnostic evaluation and temporal disaggregation.



The Appendix is organized according to the two modules introduced in Section 3.1:

- Module 1: Daily rainfall stochastic simulation;
- Module 2: Sub-hourly rainfall disaggregation.

Figure A.1 shows the initial access interface of the platform, from which the user can select one of the two rainfall models and start the modelling workflow.



Figure A.1: Initial user interface of the Virtual Rain digital platform.

A.1 Module 1 – Daily rainfall simulation

Figure A.2 reports the upload interface for the observed rainfall data together with the first exploratory diagnostics, including daily rainfall variability and annual maxima. These preliminary outputs allow the user to visually inspect the rainfall series before starting the stochastic modelling workflow.

Figure A.3 illustrates the interface used to define the seasonal partitioning of the rainfall series based on the monthly seasonal component plot automatically generated by the platform. This step represents a key modelling choice, since the selected seasonal aggregation directly affects both the fitting of marginal distributions and the calibration of the autocorrelation structures.

Figure A.4 shows the calibration results of the candidate marginal distributions fitted to wet-day rainfall values for each seasonal group. Both graphical and tabular outputs are provided to support the user in selecting the most appropriate marginal distribution according to the fitting performance and the reproduction of the upper-tail behavior.

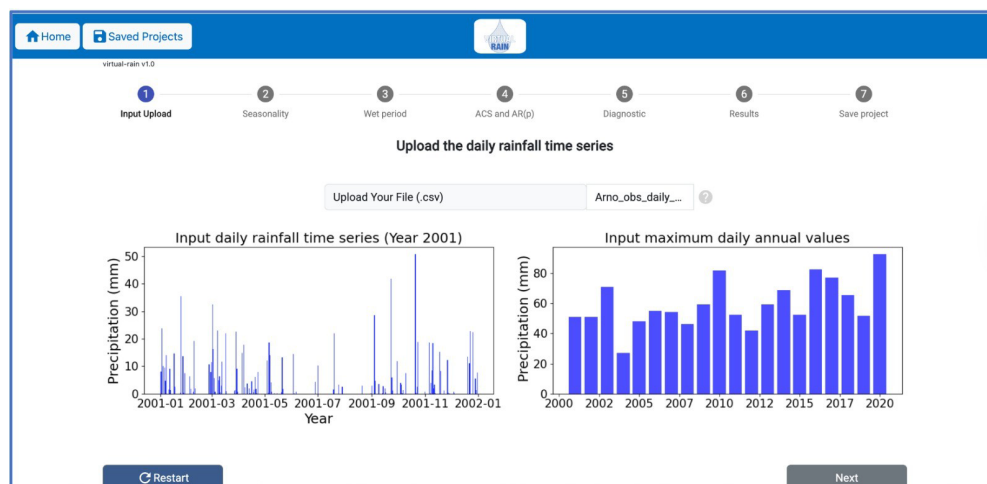


Figure A.2: Input interface for uploading the observed daily rainfall time series, including visualization of the first year and annual daily maxima.

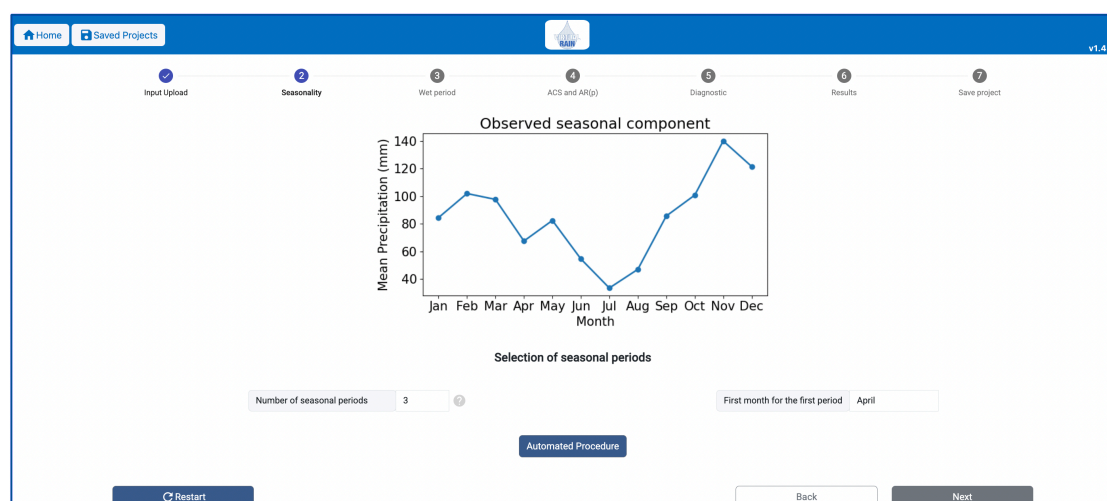


Figure A.3: Seasonality interface for selecting the number of seasonal periods and the starting month.

Figure A.5 reports the ACS calibration for both the binary and continuous rainfall processes based on the number of lags specified by the user. The interface allows the comparison between empirical and fitted Spearman ACSs obtained after calibration. At this stage, the user can select either the empirical ACS or the preferred parametric representation. The interface also allows the specification of the autoregressive order (p). It is important to note that, when the empirical ACS is selected, the number of lags necessarily coincides with the order (p) of the AR model.

Figure A.6 illustrates the diagnostic interface used to evaluate the consistency between observed and simulated rainfall statistics through ensemble-based comparisons. In this step, the user first specifies the number of synthetic daily time series to be generated. The quality of the simulations is then assessed by comparing observed and simulated properties, including: i) autocorrelation structure, describing temporal dependence; ii) seasonal cycle based on monthly mean rainfall; iii) ranked

annual maxima (highest, median and lowest); iv) summary statistics, including dry-day probability (p_0), mean wet-day rainfall, standard deviation, and extreme percentiles (P95, P99 and P99.9). These diagnostics support model validation and help the user refine the modelling configuration when necessary. Figure A.7 presents the final output interface of Module 1, including the simulated rainfall series and a summary of the selected stochastic-model configuration. From this interface, the user can directly download the generated rainfall series or automatically transfer the simulated dataset to Module 2 for temporal disaggregation into sub-hourly resolutions.



Figure A.4: Wet period module: calibration results of marginal distributions for non-zero rainfall values.



Figure A.5: ACS and AR(p) interface: calibration results of the autocorrelation structure.

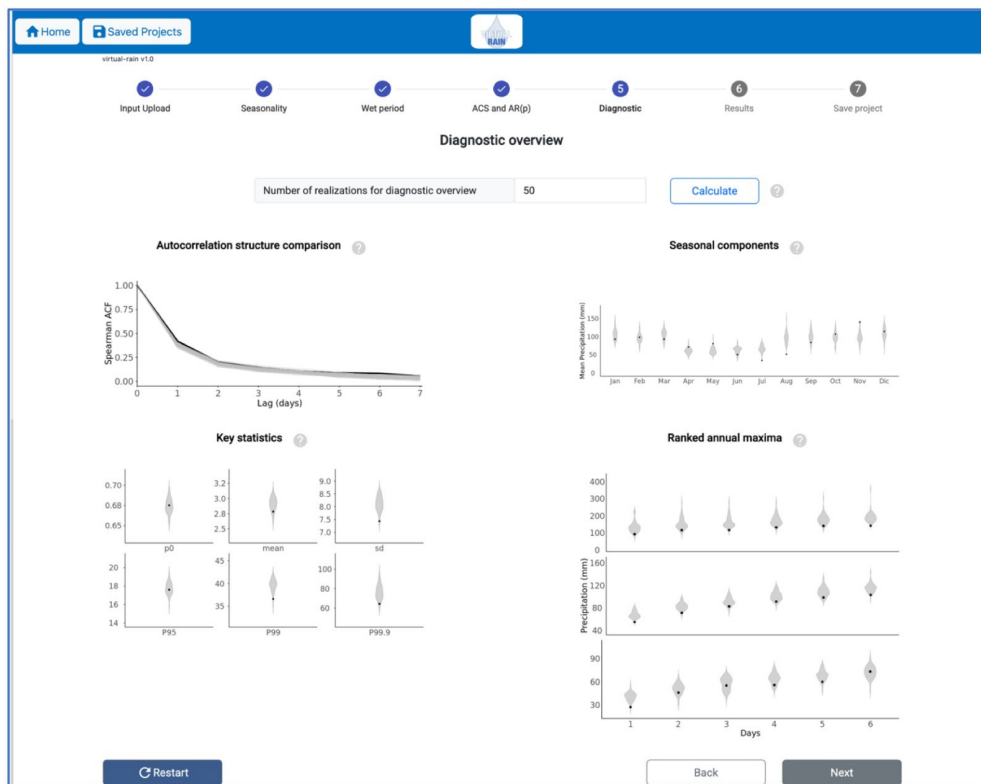


Figure A.6: Diagnostic interface: comparison between observed and simulated statistical properties for model evaluation.

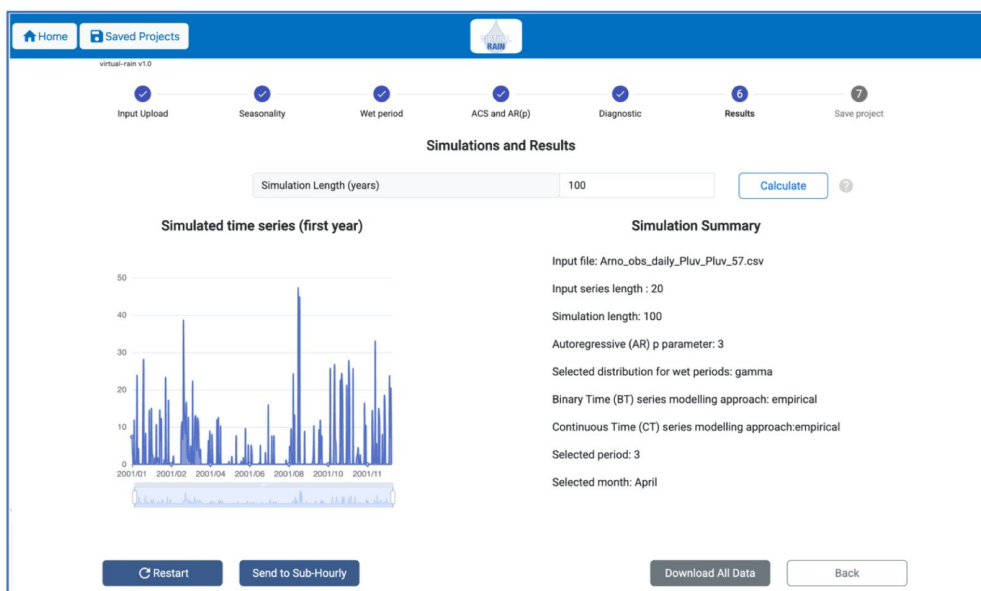
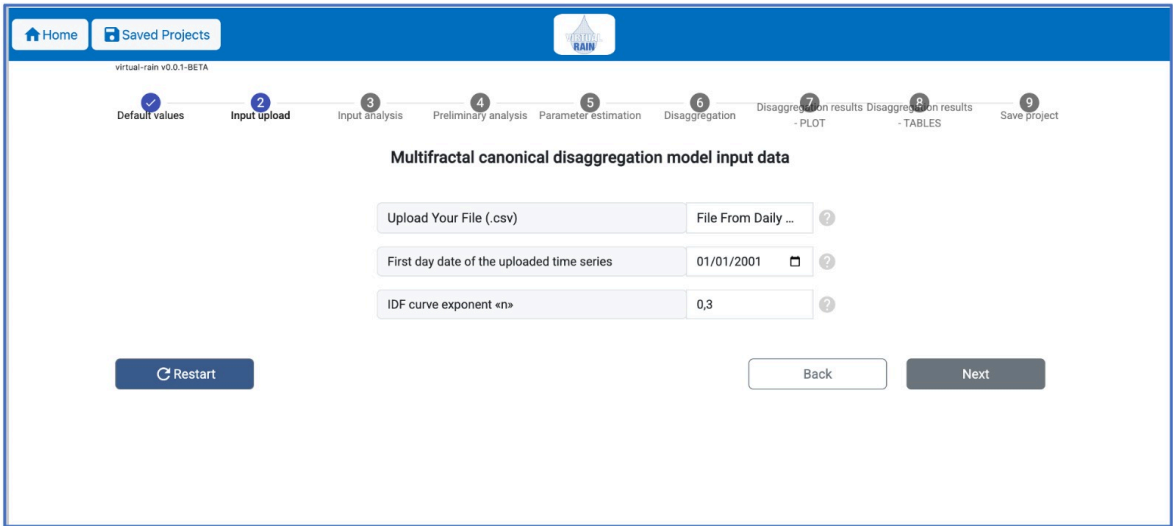


Figure A.7: Results interface: visualization of the simulated series (first year) and summary of simulation settings.



510 **A.2 Module 2 – Sub-hourly rainfall disaggregation**

Figure A.8 shows the input interface of the disaggregation module, where the user specifies the starting date of the uploaded rainfall time series together with the scaling exponent n , representing the slope of the IDF curves in logarithmic scale. This parameter is required for the calibration of the MRC model and controls the scaling relationship across temporal resolutions.



515 **Figure A.8: Input interface of the Sub-hourly Disaggregation Module showing the required inputs, including the starting date of the rainfall series and the IDF scaling exponent n .**

Figure A.9 reports the preliminary extreme-rainfall analysis, including the visualization of the first year of the rainfall series and the annual maximum daily rainfall values. The interface also provides the Gumbel plot comparing the empirical distribution of the α -coefficients with the fitted GEV distributions obtained using three estimation methods: Maximum Likelihood (MLE), Probability Weighted Moments (PWM), and L-moments (L-mom).

Figure A.10 presents the estimated GEV parameters according to the method selected in the previous step, together with the mean annual rainfall intensity and the corresponding α -coefficients used as inputs for the MRC calibration procedure. The platform also allows the user to directly provide custom α -coefficients instead of estimating them from the observed rainfall series.

525 Figure A.11 illustrates the calibration results of the MRC model, including the comparison between theoretical and empirical IDF curves for the selected return periods. The interface also reports the estimated MRC parameters resulting from the numerical optimization procedure and the corresponding objective-function value.

Figure A.12 shows the interface used to configure the temporal disaggregation process, including the target temporal resolution (up to 1 minute) and the number of stochastic realizations to be generated.

530 Figure A.13 presents the graphical comparison between the observed daily rainfall series and the corresponding disaggregated sub-hourly realization generated by the MRC model.

Figure A.14 reports a comprehensive summary of the statistical properties of the disaggregated rainfall series. The interface includes key indicators such as sample size, dry-period frequency, mean and standard deviation, extreme percentiles, maximum rainfall values at multiple aggregation scales, autocorrelation coefficients, and IDF-related parameters. These outputs provide a complete characterization of the generated time series and support the evaluation of its statistical consistency. This final step concludes Module 2 and allows the user to directly download the disaggregated rainfall series.



Figure A.9: Preliminary analysis interface showing daily rainfall for the first year, annual maximum daily rainfall and the Gumbel plot comparing GEV fits obtained through MLE, PWM and L-mom methods.

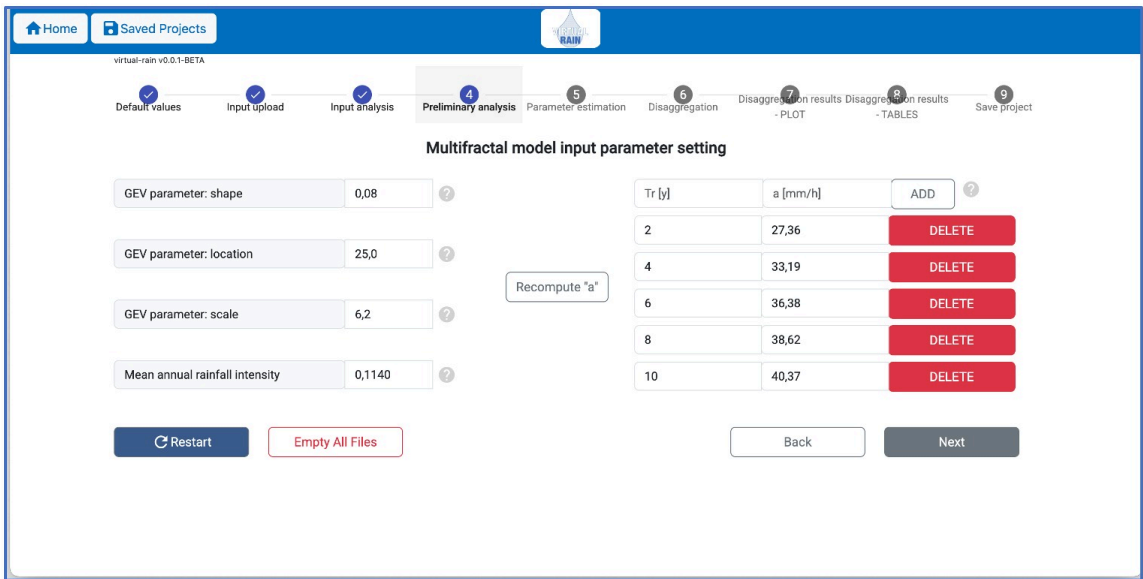


Figure A.10: Interface reporting estimated GEV parameters, mean annual rainfall intensity and the α -coefficients associated with the selected return periods.



545 Figure A.11: MRC calibration interface showing the comparison between imposed and calibrated IDF curves for the selected return periods, together with the estimated MRC parameters and objective-function value.

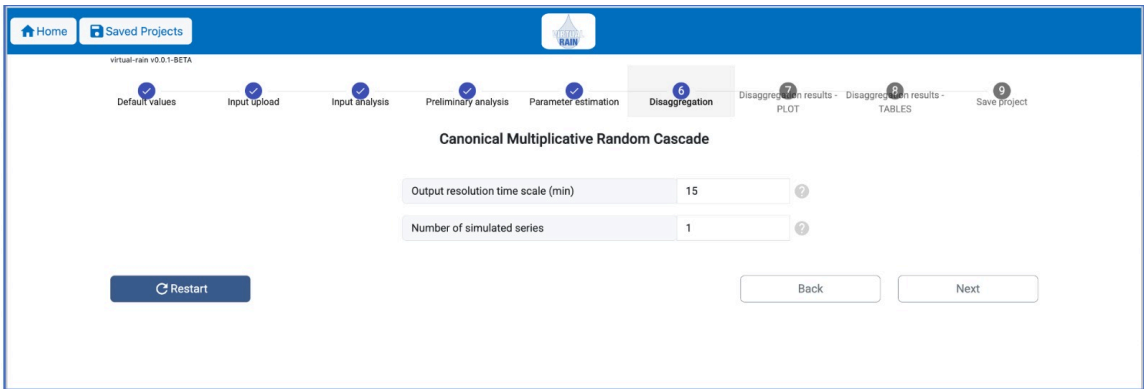


Figure A.12: Disaggregation setup interface used to define the target temporal resolution and the number of stochastic simulations.

550

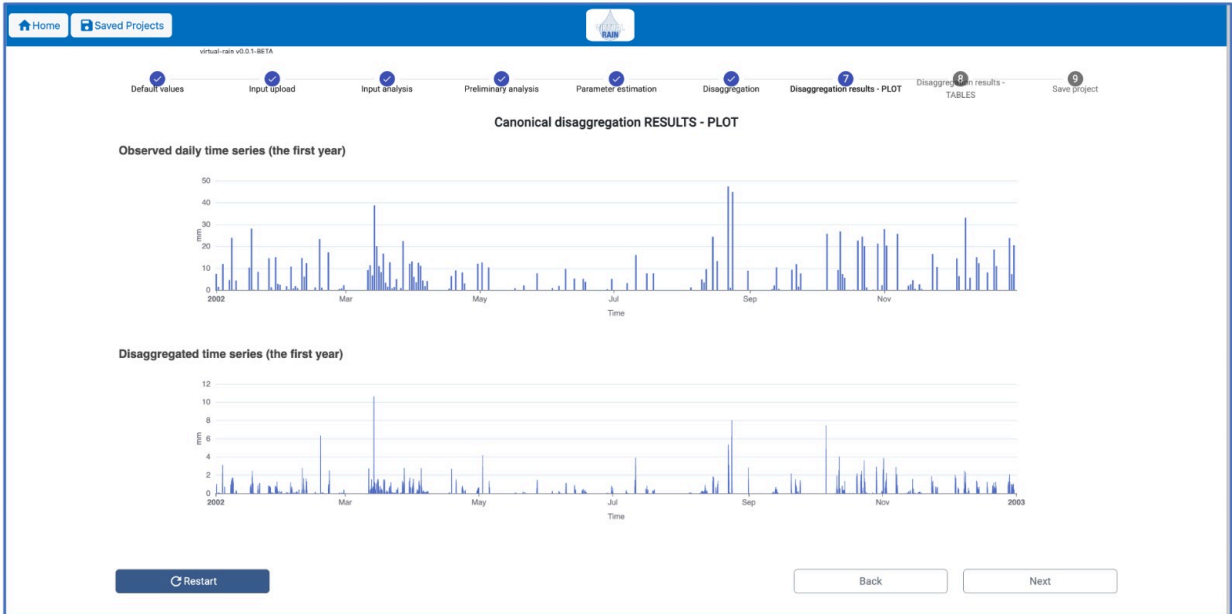


Figure A.13: Disaggregation results interface showing the comparison between the observed daily rainfall series and the generated sub-hourly disaggregated series for the first simulation year.

555

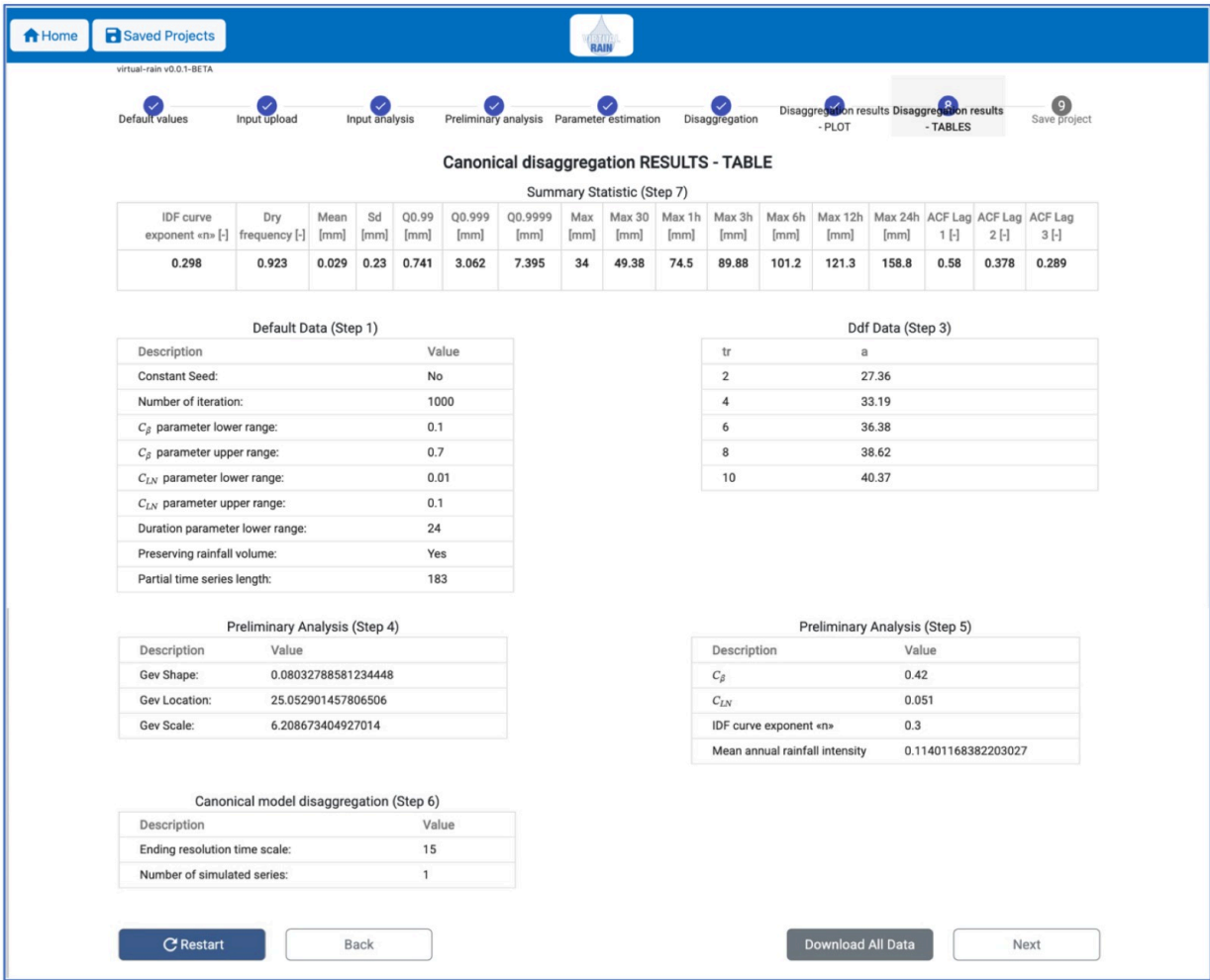


Figure A.14: Summary table interface reporting the statistical properties of the disaggregated rainfall series together with the calibrated MRC parameters and IDF-related statistics.

560 Code and data availability

The routines and data used in the paper can be downloaded from Zenodo: [10.5281/zenodo.20741916](https://zenodo.org/record/20741916)

Team list

Research Team:

- Francesco Cappelli¹ (Corresponding Author; francesco.cappelli@unitus.it)
 - Salvatore Grimaldi¹
 - Andrea Petroselli²
- 565



- Emanuele Santinami¹

¹ Department for Innovation in Biological, Agro-Food and Forest Systems, Tuscia University, Viterbo, 01100, Lazio, Italy

² Department of Agriculture and Forest Sciences, Tuscia University, Viterbo, 01100, Lazio, Italy

570 Correspondence to: Francesco Cappelli (francesco.cappelli@unitus.it)

Author contributions

Francesco Cappelli: Conceptualization, Methodology, Software, Validation, Data Curation, Formal Analysis, Writing – Original Draft.

Salvatore Grimaldi: Conceptualization, Supervision, Methodology, Funding Acquisition, Writing – Review & Editing.

575 Andrea Petroselli: Data Curation, Investigation, Validation, Resources, Writing – Review & Editing.

Emanuele Santinami: Software, Formal Analysis, Validation, Data Curation, Writing – Review & Editing.

All authors approved the final manuscript.

Competing interests

The authors declare no competing interests.

580 Disclaimer

“Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.”

585 Acknowledgements

The authors also wish to thank Innosystem S.r.l., particularly Fabrizio Piergentili and Stefano Bianchini, for their valuable support and collaboration during the development of this work.

Financial support

This work was funded by the WE-FLOOD project (FIS-2024-01376), supported by the Italian Ministry of University and
590 Research (MUR) under the Italian Science Fund (*Fondo Italiano per la Scienza*, FIS), Call FIS 3 (CUP J53C25002290001).
Further financial support was provided through the collaboration agreement between the Autorità di Bacino Distrettuale



dell'Appennino Meridionale and the Department for Innovation in Biological, Agri-food and Forest Systems (DIBAF) of Università degli Studi della Tuscia (CUP F54J16000030001; CUP F52G16000010001).

Review statement

- 595 The review statement will be added by Copernicus Publications listing the handling editor as well as all contributing referees according to their status anonymous or identified.

References

- Cappelli, F., Papalexiou, S. M., Markonis, Y., and Grimaldi, S.: PyCoSMoS: An advanced toolbox for simulating real-world hydroclimatic data, *Environ. Model. Softw.*, 178, 106076, <https://doi.org/10.1016/j.envsoft.2024.106076>, 2024.
- 600 Cappelli, F., Volpi, E., Langousis, A., Deidda, R., Perdios, A., Furcolo, P., and Grimaldi, S.: Sub-daily rainfall simulation using multifractal canonical disaggregation: A parsimonious calibration strategy based on intensity–duration–frequency curves, *Stoch. Environ. Res. Risk Assess.*, 39, 1–19, <https://doi.org/10.1007/s00477-024-02865-6>, 2025a.
- Cappelli, F., Volpi, E., Langousis, A., Deidda, R., Perdios, A., and Grimaldi, S.: 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>, 2025b.
- 605 Cappelli, F., Volpi, E., Langousis, A., Papalexiou, S. M., Deidda, R., Perdios, A., and Grimaldi, S.: Simulating sub-daily rainfall time series in the absence of sub-daily observations, *J. Hydrol.*, <https://doi.org/10.1016/j.jhydrol.2026.135660>, 2026.
- Fischer, S., Dallan, E., Fiori, A., Grimaldi, S., Kochanek, K., Prieto, C., Reis Jr., D. S., and Volpi, E.: Hydrological design in the helping decade – inspiring the community to innovate the hydrological design concept, *Hydrol. Sci. J.*, 70, 375–389, <https://doi.org/10.1080/02626667.2024.2436634>, 2025.
- 610 Grimaldi, S., Volpi, E., Langousis, A., Papalexiou, S. M., De Luca, D. L., Piscopia, R., Nerantzaki, S. D., Papacharalampous, G., and Petroselli, A.: 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>, 2022.
- Koutsoyiannis, D.: Stochastics of Hydroclimatic Extremes: A Cool Look at Risk, Kallipos, <https://doi.org/10.57713/kallipos-1>, 2025.
- 615 Koutsoyiannis, D., Onof, C., and Wheeler, H. S.: Multivariate rainfall disaggregation at a fine timescale, *Water Resour. Res.*, 39, 1173, <https://doi.org/10.1029/2002WR001600>, 2003.
- Langousis, A., Veneziano, D., Furcolo, P., and Lepore, C.: Multifractal rainfall extremes: Theoretical analysis and practical estimation, *Chaos Soliton. Fract.*, 39, 1182–1194, <https://doi.org/10.1016/j.chaos.2007.06.004>, 2009.
- 620 Northrop, P. J.: Stochastic models of rainfall, *Annu. Rev. Stat. Appl.*, 11, 363–389, <https://doi.org/10.1146/annurev-statistics-040622-023838>, 2024.



- Onof, C. and Arnbjerg-Nielsen, K.: Quantification of anticipated future changes in high resolution design rainfall for urban areas, *Atmos. Res.*, 92, 350–363, <https://doi.org/10.1016/j.atmosres.2009.01.014>, 2009.
- Papalexiou, S. M.: 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>, 2018.
- Papalexiou, S. M.: 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>, 2022.
- Pérez, F. and Granger, B. E.: IPython: A system for interactive scientific computing, *Comput. Sci. Eng.*, 9, 21–29, <https://doi.org/10.1109/MCSE.2007.53>, 2007.
- Volpi, E., Grimaldi, S., Aghakouchak, A., Castellarin, A., Chebana, F., Papalexiou, S. M., Aksoy, H., Bárdossy, A., Cancelliere, A., Chen, Y., and co-authors: The legacy of STAHY: Milestones, achievements, challenges, and open problems in statistical hydrology, *Hydrological Sciences Journal*, 69, 1913–1949, <https://doi.org/10.1080/02626667.2024.2405847>, 2024.