



HAL
open science

A stochastic generator of continuous hourly areal rainfall series applied to a wide range of French catchments

Patrick Arnaud, Philippe Cantet

► **To cite this version:**

Patrick Arnaud, Philippe Cantet. A stochastic generator of continuous hourly areal rainfall series applied to a wide range of French catchments. SWGEN 2025 - Conference on Stochastic Weather Generators, Dec 2025, Grenoble, France. 2025. <hal-05547377>

HAL Id: hal-05547377

<https://hal.science/hal-05547377v1>

Submitted on 11 Mar 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



ETALAB - Open licence

A stochastic generator of continuous hourly areal rainfall series applied to a wide range of French catchments

Patrick Arnaud¹, Philippe Cantet²

¹ UMR RECOVER, INRAE, Aix-Marseille University, Aix-en-Provence, 13182, France

² DCSC-EMA, Météo-France, Toulouse, 31000, France



ABSTRACT

This study introduces a stochastic continuous hourly areal rainfall generator designed to simulate rainfall time series for hydrological risk management. This generator, named SCHYPRE for Simulation of Continuous HYetographs for Predictive Risk Estimation, was calibrated using 28.5 years of high-resolution radar rainfall data across 2,108 watersheds in France, covering diverse climatic regimes from continental to Mediterranean and mountainous.

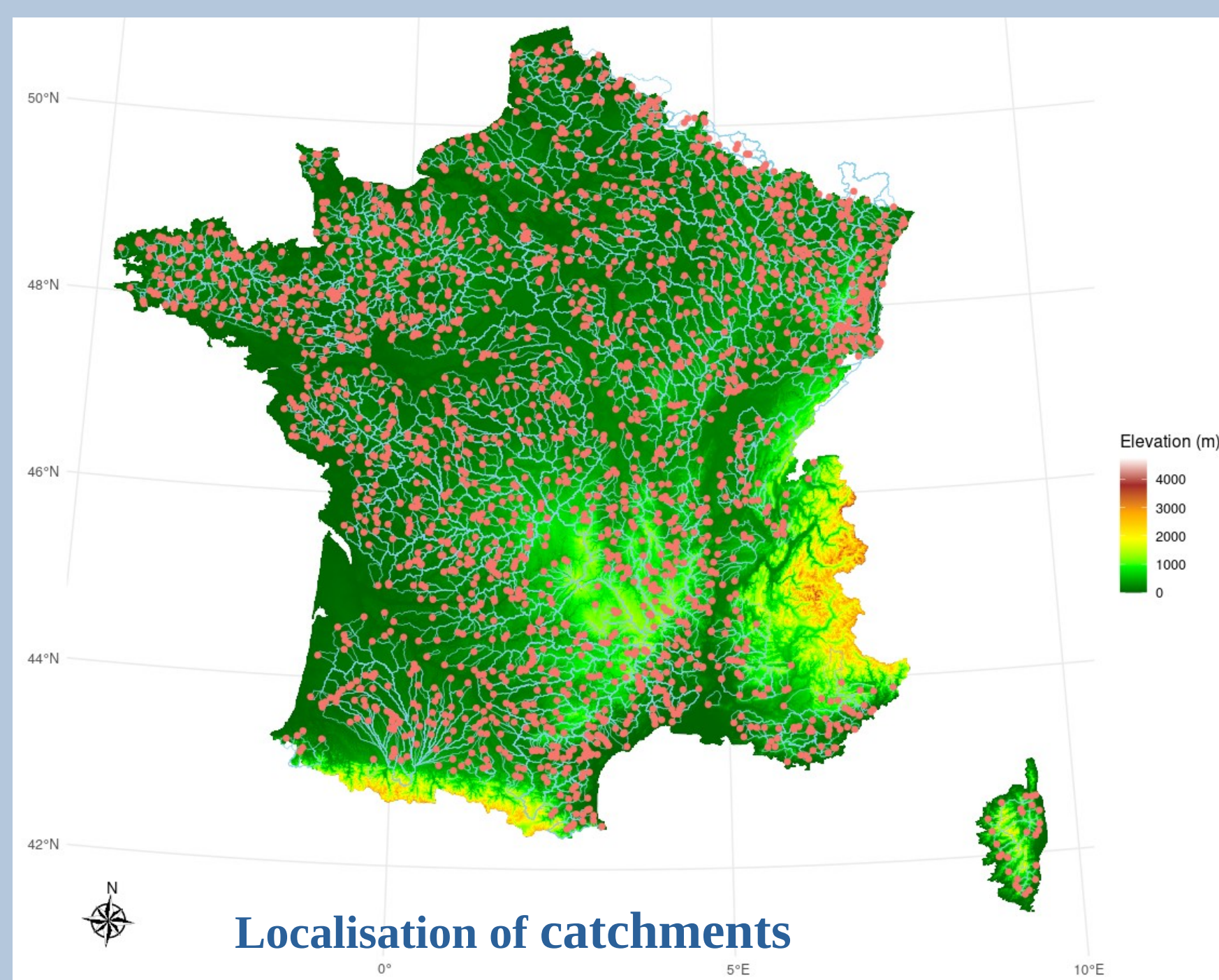
Validation results demonstrate SCHYPRE's ability to faithfully reproduce observed rainfall statistics, including mean and extreme values, seasonality, autocorrelation, and intermittency. Comparative analyses reveal relative errors below 10% for monthly mean rainfall, standard deviation, and dry-day proportions, and seasonal confirms accurate representation of climatic variability. Frequency analyses for durations ranging from 1 hour to 1 year show strong agreement between simulated and observed quantiles, with minimal bias for extreme values.

A key strength of SCHYPRE is its stability in extreme value estimation. Unlike traditional probabilistic methods, which are sensitive to sampling variability, SCHYPRE's Monte Carlo-based approach, calibrated on extensive observational samples, ensures stable quantile estimates even for high return periods. Additionally, the generator inherently avoids quantile crossing inconsistencies, a common issue in independent duration-based probabilistic modeling.

SCHYPRE's operational performance makes it a valuable tool for hydrological applications, including flood simulation, risk assessment, and infrastructure design. Its ability to generate long, coherent, and realistic rainfall time series—without restrictive assumptions—positions it as a reliable framework for contemporary water management challenges. Future work will explore coupling with spatial disaggregation model (GRAFFAS) and distributed hydrological models.

DATA

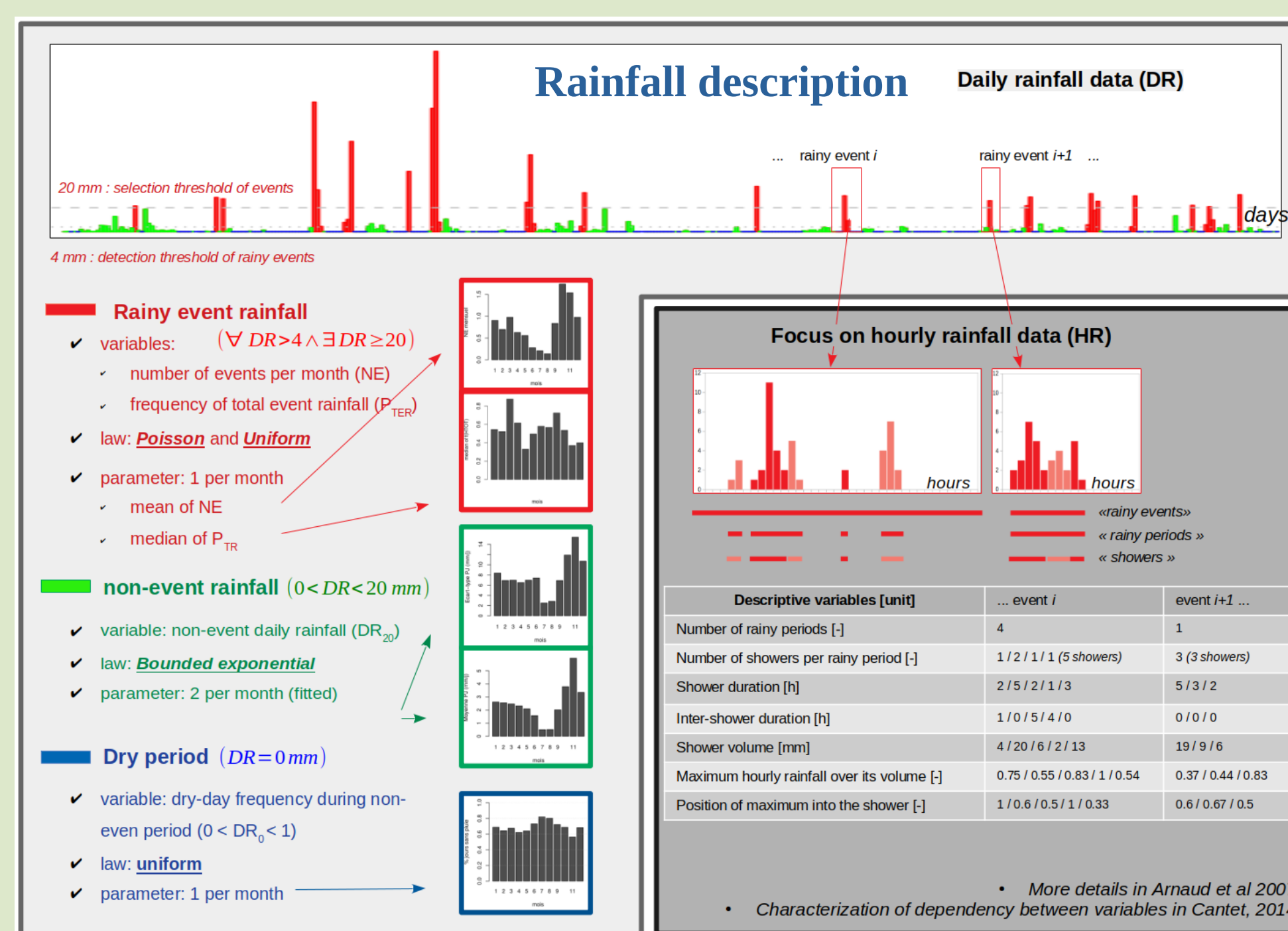
- ⇒ Processing of **2,108 catchments** ranging from 5 to 10,000 km² (with a median area of 220 km²).
- ⇒ Gridded hourly rainfall : **28.5 years of radar-based precipitation data** from the COMEPHORE data (1997–2025), provided on a 1 km × 1 km grid.
- ⇒ Estimation of **hourly areal rainfall** for each catchments



MODEL

Rainfall signal are decomposed into random variables at daily scale to define rainy events characteristics and at hourly scale to define showers characteristics.

- At the daily scale, three period types are identified: rainy events (consecutive days with at least one day exceeding 20 mm rainfall), non-event periods (no rainfall exceed 20 mm), and dry periods (no rainfall). Four monthly variables, NE, P_{TER}, DR_{<20}, HR0 and DR0, describe these periods.
- At the hourly scale, rainy events are analyzed as sequences of rainy periods composed by showers. Each shower is defined by its duration, volume and form. Variables are modeled using probability distributions, with parameters estimated via the method of moments. Seasonality is incorporated, with parameters calibrated for two seasons and monthly adjustments.



- Variables are modeled using probability distributions, for which parameters are estimated via the method of moments. Moreover, copulas are used to model the interdependence between some variables such as duration and volume of showers.

Variables, laws and parameters of the model

Variable	Description of the variable	Unit	Probability law	Parameter estimation ⁽¹⁾	Estimated for ...
Variables required to simulate rainy events					
NE	Number of rainy event per month	-	Poisson	$\theta = \mu$	12 months
NP	Number of rainy period per rainy event	-	Geometric	$\theta = 1/\mu$	2 seasons
NS	Number of shower per rainy period	-	Geometric	$\theta = 1/\mu$	2 seasons
NpS	Number of "principal" shower ⁽²⁾	-	Binomial	$\theta = (1/\mu)^{15}$	2 seasons
IPD	Inter rainy period duration	hours	Truncated geometric	$\theta_1 = 1/\mu$	2 seasons
SD ₁ et SD ₂	Shower duration, for the two types of shower ⁽²⁾	hours	(p) Truncated Poisson (e) Truncated Poisson	$\theta_1 = \mu$ $\theta_2 = \text{Prob}(PD > 9h)$	2 seasons and 2 types
SV ₁ et SV ₂	Shower volume, for the two types of shower ⁽²⁾	mm	(p) Truncated Log-normal (e) Truncated Gamma	$\theta_{1,2,3} = (\mu, \sigma, \text{min})$ $\theta_{1,2,3} = (\mu, \sigma, \text{min})$	2 seasons and 2 types
RX ₁ et RX ₂	Ratio of shower maximum hourly rainfall over its volume, for the two kind of shower ⁽²⁾	-	Normal	$\theta_{1,2,3} = (\mu, \sigma, \text{min})$ $\theta_1 = \mu$ $\theta_2 = \sigma$	2 seasons and 2 types
PX	For a shower, relative position of the maximum hourly rainfall in the shower	-	Normal	$\theta_1 = \mu$	2 seasons
P _{TER}	Probability of the total event rainfall	-	Uniform	$\theta_1 = \text{median}$	12 months
τ_1 and τ_2	Kendall's Tau for copula between duration and volume probabilities	-	(p) Frank's copula (e) Gumbel copula	Kendall's Tau	2 seasons and 2 types
Variables required to simulate "non-event" rainfall periods					
DR ₌₀	Non-event daily rainfall	mm	Exponential	θ_1, θ_2 optimised	12 months
HR0	Proportion of hourly rainfall equal to zero within a non-event daily rainfall	-	Uniform	$\theta = \mu$	12 months

⁽¹⁾ $\mu = \text{mean of the variable values}$, $\sigma = \text{standard deviation of the variable values}$ and $\text{min} = \text{minimum of the variable values}$
⁽²⁾ Only the variables SD, SV, and RX are considered for the two shower types.
⁽³⁾ The binomial distribution parameter describing the variable NpS is a regionally calibrated function of the variable's mean. Using our dataset, we obtain: $\theta = 0.15 \cdot \ln(\mu_{p0.5})$ for Season 1 (1) and $\theta = 0.122 \cdot \ln(\mu_{p0.5})$ for Season 2.
⁽⁴⁾ The mean and standard deviation of the normal distribution used to randomly sample the variable RX are estimated as functions of shower duration SD. For each randomly drawn duration, the variable RX is generated from a normal distribution, $\mathcal{N}(\theta_1, \theta_2)$, where $\theta_1 = a + b \ln(\text{SD})$ and $\theta_2 = c + d \cdot \ln(\text{SD})$. The parameters a, b, c, d are regionally optimized to match the observed mean and standard deviation of RX. These four parameters are determined separately for both seasons and both shower types.

RESULTS

We compare statistical characteristics between observed time series and simulated series (simulation of 10,000 years). Each boxplot shows the distribution across the 2,108 catchments, and each point in the scatterplots represents an individual catchment.

- Relative errors in monthly characteristics are minimal and month-independent (Fig 1), ensuring rainfall seasonality is accurately maintained.
- Time-aggregated rainfall features (hourly to annual) are well-preserved, and the multi-duration rainfall structure is accurately captured (Fig 2).
- The autocorrelation of aggregated rainfall across multiple time steps is faithfully reproduced at the correct order of magnitude, and rainfall intermittency is successfully preserved (Fig 3).
- These results are independent of basin size.

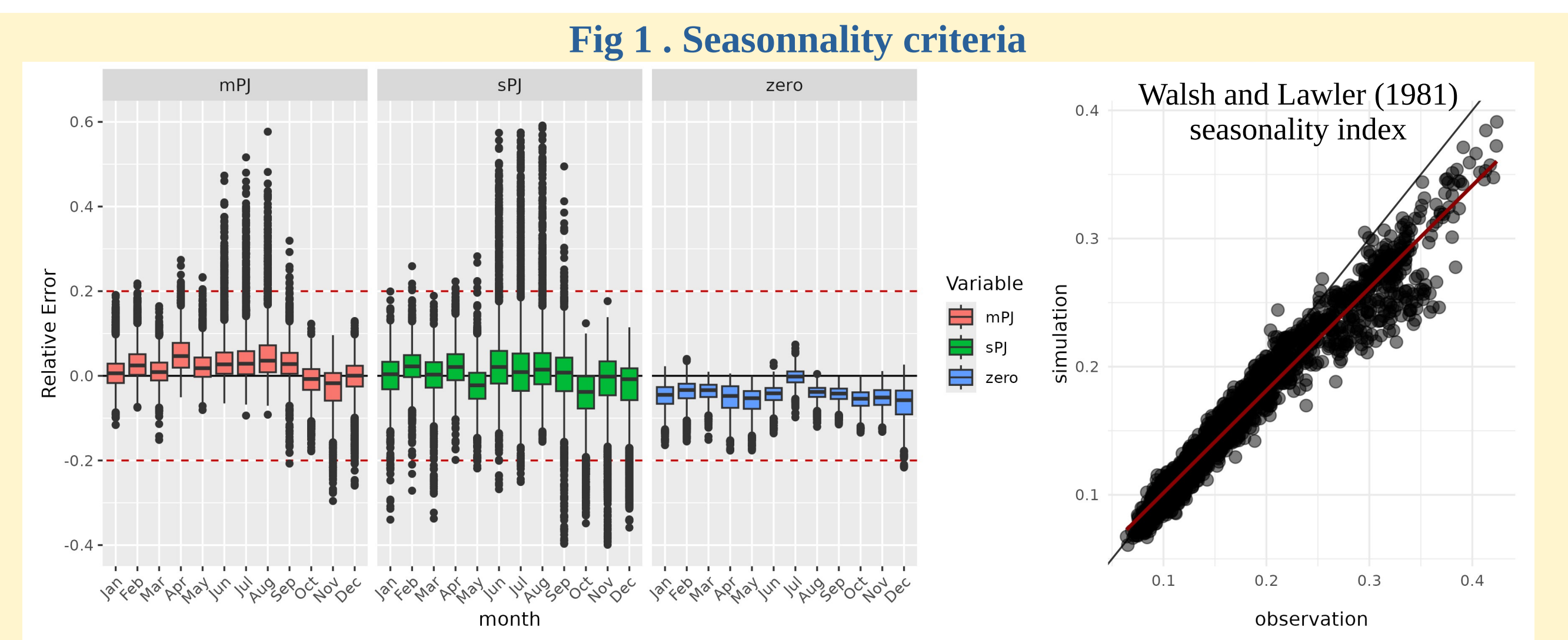


Fig 2. Multi-duration characteristics

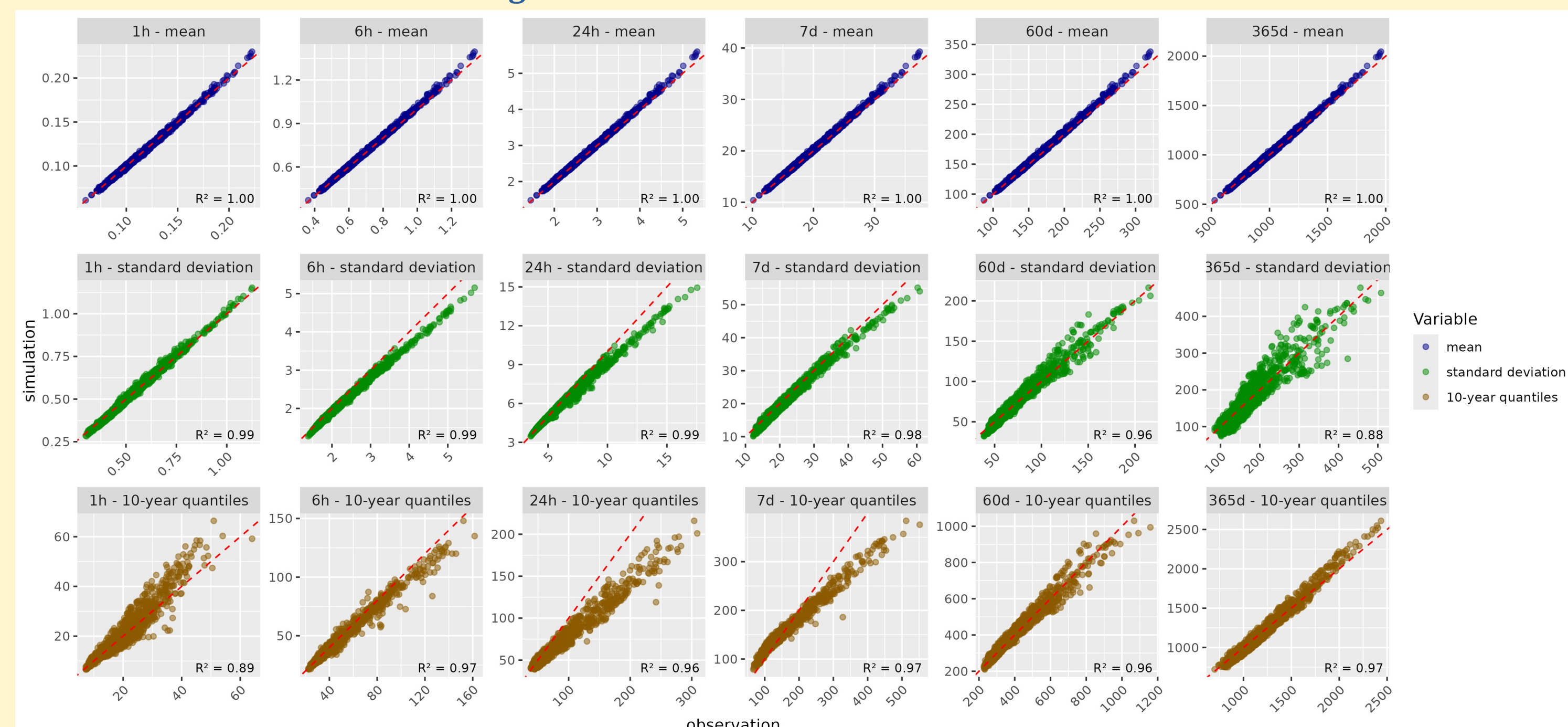
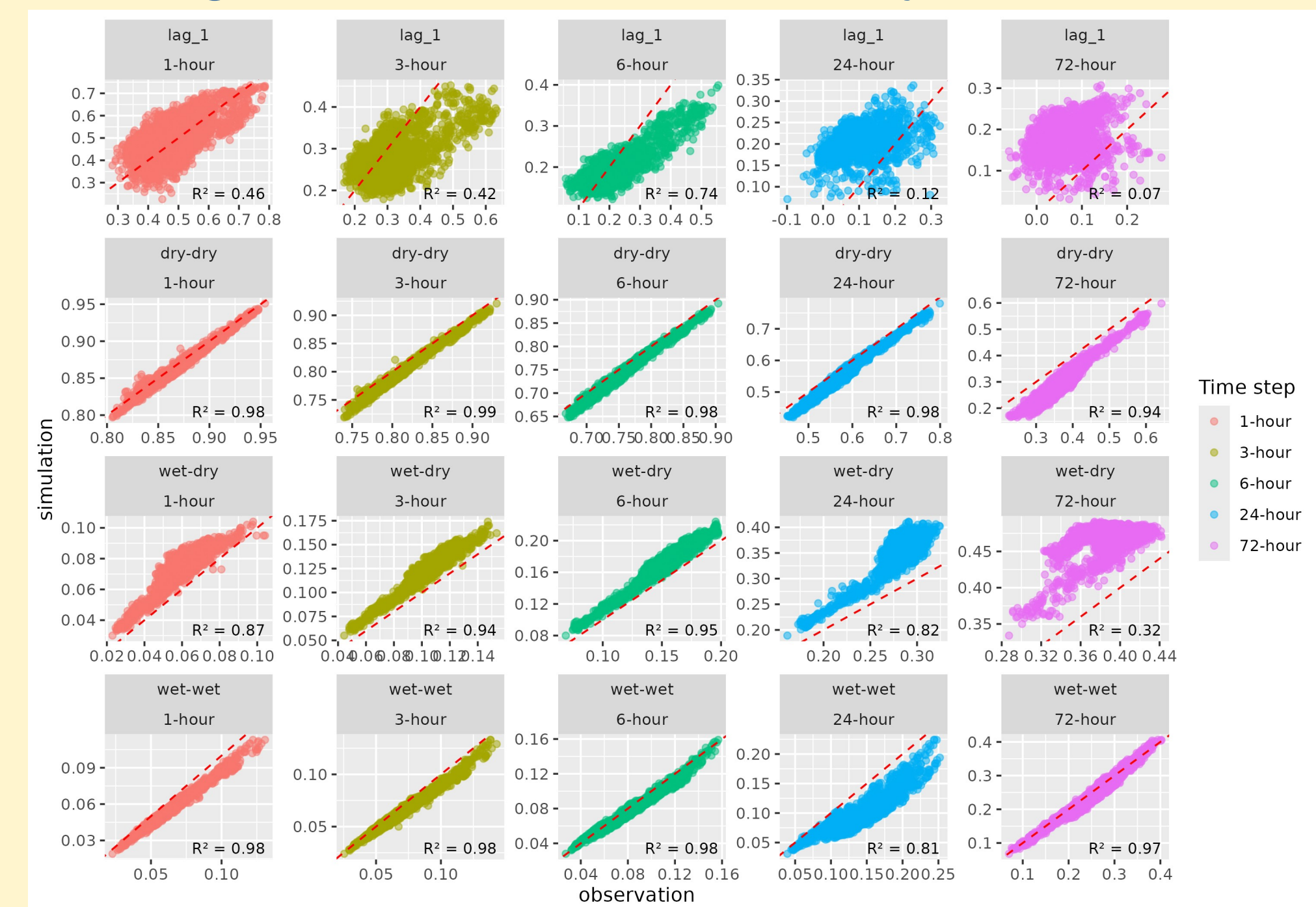


Fig 3. Autocorrelation and intermittency characteristics



CONCLUSION The SCHYPRE generator reproduces various rainfall characteristics across a wide range of basin sizes, over aggregation time-scales ranging from hourly to annual, and for a diverse set of rainfall regimes. We now have a readily calibratable tool capable of simulating very long synthetic hourly areal rainfall time series that are statistically similar to observed data. Our objective is to use these time series either to drive lumped hydrological models or to feed a spatial disaggregation model for areal rainfall (GRAFFAS), thereby providing hourly rainfall field time series to drive distributed hydrological models.