https://doi.org/10.5194/hess-2021-453
Preprint. Discussion started: 8 September 2021

Hydrology and Earth System Sciences
Discussions

Stochastic daily rainfall generation on tropical islands with complex topography

Lionel Benoit[1,2], Lydie Sichoix[2], Alison D. Nugent[3], Matthew P. Lucas[4], Thomas W. Giambelluca[1]

[1]Water Resources Research Center, University of Hawai‘i at Mānoa, 96822 Honolulu, Hawai‘i, USA
[2]GePaSud Laboratory, University of French Polynesia, 98702 Tahiti, French Polynesia
[3]Department of Atmospheric Sciences, School of Ocean and Earth Science and Technology, University of Hawai‘i at Mānoa, 96822 Honolulu, Hawai‘i, USA
[4]Department of Geography, University of Hawai‘i at Mānoa, 96822 Honolulu, Hawai‘i, USA

Correspondence to: Lionel Benoit (benoitlionel2@gmail.com)

Abstract. Stochastic rainfall generators are probabilistic models of rainfall space-time behavior. During parameterization and calibration, they allow the identification and quantification of the main modes of rainfall variability. Hence, stochastic rainfall models can be regarded as probabilistic conceptual models of rainfall dynamics.

As with most conceptual models in Earth Sciences, the performance of stochastic rainfall models strongly relies on their adequacy in representing the rain process at hand. On tropical islands with high elevation topography, orographic rain enhancement challenges most existing stochastic models because it creates localized rains with strong spatial gradients, which break down the stationarity of rain statistics. To allow for stochastic rainfall modeling on tropical islands, despite non-stationarity, we propose a new stochastic daily rainfall generator specifically for areas with significant orographic effects.

Our model relies on a preliminary classification of daily rain patterns into rain types based on rainfall space and intensity statistics, and sheds new light on rainfall variability at the island scale. Within each rain type, the spatial distribution of rainfall through the island is modeled following a meta-Gaussian approach combining empirical spatial copulas and a Gamma transform function, which allows us to generate realistic daily rain fields.

When applied to the stochastic simulation of rainfall on the islands of O‘ahu (Hawai‘i, United States of America) and Tahiti (French Polynesia) in the tropical Pacific, the proposed model demonstrates good skills in jointly simulating site specific and island scale rain statistics. Hence, it provides a new tool for stochastic impact studies in tropical islands, in particular for watershed water resources management and downscaling of future precipitation projections.

1 Introduction

Stochastic rainfall generators are probabilistic tools aiming at simulating synthetic rains that mimic as closely as possible the statistical signature of rain observations [*Richardson*, 1981] [*Wilks and Wilby*, 1999] [*Ailliot et al.*, 2015]. More specifically, stochastic rainfall modeling consists of statistical learning (i.e., inference) of the joint space-time probability density function (pdf) of rainfall at all sites and times of interest, and sampling this pdf to generate synthetic rains. This empirical approach bypasses the detailed physical modeling of rain generation processes [*Bauer et al.*, 2015], which enables fast and computationally efficient simulations.

The ability of stochastic rainfall generators to emulate long and realistic rainfall sequences makes them an appropriate tool for the simulation of design storms [*Niemi et al.*, 2016]. Simulated rains can then be used as inputs for impact models assessing the effects of rainfall on different environmental processes including hydrology [*Paschalis et al.*, 2014], water resources [*Cappelaere et al.*, 2020], geomorphology [*Peleg et al.*, 2020], and agronomy [*Mavromatis and Hansen*, 2001]. The probabilistic approach followed by stochastic rainfall generators enables a comprehensive study of rainfall variability and, in turn, the assessment of uncertainty propagation along the whole modeling chain [*Gabellani et al.*, 2007]. This makes stochastic rainfall generation a key tool for management of rain-induced risk, in particular, for flood [*Caseri et al.*, 2016] and drought risks [*Supit et al.*, 2012]. In addition, the focus of stochastic rainfall models on the statistical signature of rainfall creates new ways to characterize rainfall space-time behavior [*Marra and Morin*, 2018], and assess the impact of rainfall variability on the hydrosphere [*Morin et al.*, 2019]. Finally, when conditioned to climate model outputs, stochastic rainfall generation can be used for the downscaling of future precipitation projections, resulting in local-scale and high-resolution scenarios of the possible evolution of rainfall in the context of climate change [*Jha et al.*, 2014] [*Volosciuk et al.*, 2017].

To capture and reproduce rainfall statistics and space-time variability, stochastic rainfall models embed a significant part of our conceptual knowledge about rainfall behavior in their parameterization. However, rainfall properties [*Krajewski et al.*, 2003] and, in turn, the performance of stochastic rainfall generators [*Breinl et al.*, 2017] [*Vu et al.*, 2018] strongly depend on the climate of the area of interest. Hence, different models have been proposed for different climates with each model focusing on a specific aspect of rainfall, for instance: rainfall seasonality in monsoonal climates [*Greene et al.*, 2011]; rainfall spatial-temporal correlation in temperate climates [*Paschalis et al.*, 2013]; or rainfall occurrence and extreme intensities in arid regions [*Wilcox et al.*, 2021].

On high tropical islands, or islands with high elevations and significant topography, rainfall is strongly location dependent due to complex interactions between atmospheric circulation and island topography, which trigger different mechanisms of orographic rain enhancement [*Houze*, 2012]. This makes tropical island rain statistics non-stationary in space [*Benoit et al.*, 2021] because the fixed topography of the islands induces the orographic lifting of relatively steady trade winds, which generates well defined rain patterns [*Lyons*, 1982]. This leads to wetter windward slopes than leeward sides, and wetter highlands than lowlands [*Giambelluca et al.*, 2013] [*Laurent et al.*, 2019]. To this first order quasi-static picture is added the important variability of daily rainfall patterns associated with processes ranging from synoptic-scale disturbances [*Hopuare et al.*, 2018] [*Longman et al.*, 2021] to large-scale atmospheric circulations [*Hopuare et al.*, 2015] [*Frazier et al.*, 2018] [*Brown et al.*, 2020]. This variability brings stochasticity on top of the relatively deterministic long-term patterns of orographic rain enhancement.

To account for both the long-term quasi-static patterns of rain accumulation and the day-to-day fluctuations of the rainfall spatial distribution, this paper proposes a new stochastic rainfall model dedicated to high tropical islands with significant and complex topography. The goal is to develop a daily resolution stochastic rainfall generator able to simulate: (1)

https://doi.org/10.5194/hess-2021-453
Preprint. Discussion started: 8 September 2021

site specific rain occurrence, persistence, intensity and seasonality; (2) spatial patterns of daily rain accumulation; and (3) areal rain statistics at the island scale.

To achieve these objectives, the remainder of the article is structured as follows. Section 2 briefly reviews the main features of tropical island rainfall and describes our stochastic rainfall model. Section 3 illustrates the performance of the model for the island of O‘ahu (Hawai‘i, USA) in the tropical Pacific, and a similar test study is repeated in supplementary material for the island of Tahiti (French Polynesia) to demonstrate the versatility of the model. Finally, section 4 discusses how the focus on orographic rain enhancement has influenced the design of the model and provides concluding remarks.

2 Data and methods

2.1 Rainfall features of interest

Because stochastic rainfall models are data-driven, their structure depends on the rain features one wants to reproduce in simulations. Hence, the identification of the main features of daily rainfall in high tropical islands is a prerequisite for the design of the present model. For illustration purposes, we focus throughout the main text on the island of O‘ahu, Hawai‘i (lon = 158°W, lat = 21.5°N, area = 1545 km^2, max altitude = 1220 m). The available rain gauge observation dataset consists of daily records from a network of 86 rain gauges spread over the island (Fig. 1a), and covers a 20-year period 1991–2011. It corresponds to a compilation of quality controlled and gap-filled daily observations [*Longman et al.*, 2018]. To contextualize the observed rain patterns, several meteorological covariates (e.g., pressure, temperature, humidity and wind) are investigated at the island scale. We use the ERA5 reanalysis dataset [*Hersbach et al.*, 2018] at 12:00 PM HST to inform these covariates and average the values of the 12 grid cells (pixel size = 0.25° x 0.25°) encompassing the island of O‘ahu.

Figure 1 displays the main features of daily rainfall over the island of O‘ahu. It shows the strong impact of trade wind induced orographic rain enhancement on the spatial distribution of annual rains (Fig. 1a), with windward (northeast) sides significantly wetter than leeward (southwest) ones, and highlands generally wetter than lowlands. In addition to prevailing orographic rains triggered by the interactions of trade winds with island topography (east-northeasterly trade winds blow more than 280 days per year over the Hawaiian archipelago [*Longman et al.*, 2015]), the island of O‘ahu also experiences widespread rain events, mostly triggered by regional atmospheric disturbances such as cold fronts originating from mid-latitudes and Kona storms [*Longman et al.*, 2021]. These atmospheric disturbances mostly occur during (boreal) winter, which corresponds to the local rainy season (spanning from October–March). They represent the main source of precipitation for dry leeward locations and are responsible for the enhanced seasonality of rain accumulation in these areas (Fig. 1a).

The diversity of rain generation mechanisms (e.g., orographic lifting, cold fronts, or Kona lows) coupled with the steep island topography of volcanic origin result in a complex distribution of rainfall in space, which produces highly variable island-scale rain statistics (i.e., statistics summarizing rain behavior throughout the island for a given day). Figure 1 b–d shows that at the scale of the island of O‘ahu, daily rainfall is strongly intermittent in space (only 3% of the days record rain at all gauge locations, and half of the time at least 20% of the gauges measure no rain, Fig. 1b), highly skewed (island-scale rain accumulation average < 2.25mm/day 50% of the time, but island-scale maximum accumulation > 15mm/day 50% of the time and reaches 500 mm/day, Fig 1c), and strongly variable in space (coefficient of variation > 1.3 50% of the time, and > 2.9 10% of the time).

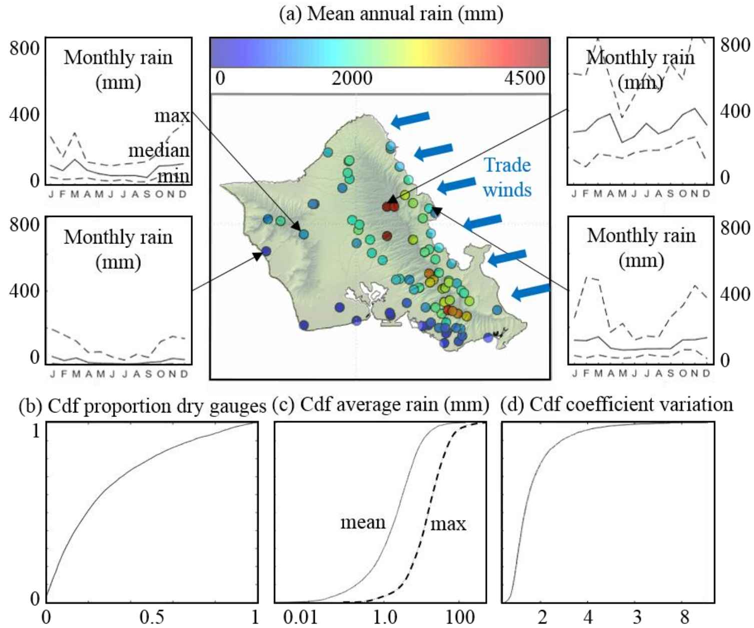

Figure 1: Main features of rainfall observed over the island of O'ahu. (a) Mean annual rainfall (central panel) and seasonality of rain accumulation for four specific rain gauges (outer panels). (b) Cumulative distribution function (cdf) of the proportion of gauges measuring no rain for a given day. (c) Cdf of the mean and maximum daily rain accumulation computed over the whole observation network (abscissa is in log-scale). (d) Cdf of the coefficient of variation (i.e., standard deviation/mean) of daily rain accumulation throughout the rain gauge network.

2.2 Model description

2.2.1 Model overview

To account for the above features of daily rainfall, the proposed model splits rainfall behavior into three components: temporal variability; spatial distribution; and intensity (i.e., marginal distribution). Figure 2 summarizes the structure of the model, which will be discussed in detail later.

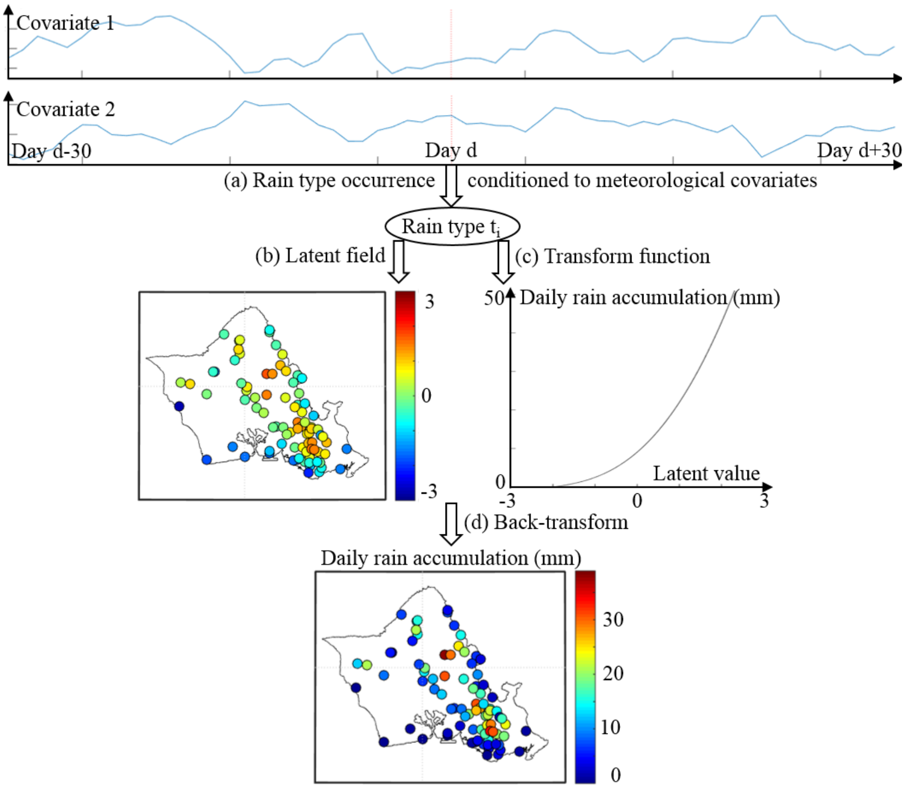

Figure 2: Overview of the structure of the stochastic rainfall model. (a) Meteorological conditions driving the occurrence of rain types, which summarize daily rain statistics. (b) Latent field modeling of the spatial distribution of rainfall across the island. (c) Transform function linking latent values with actual rain accumulations. (d) Back-transform combining (b) and (c) to obtain daily rain simulations.

The temporal variability of rain statistics and its relationships with the state of the atmosphere is modeled following a rain typing approach (Fig. 2a). In this framework [*Ailliot et al.*, 2015] [*Benoit et al.*, 2018a], days with similar rain statistics are pooled together in a finite number of rain types. Rain types represent summaries of island-scale daily rain statistics. To preserve climatological consistency and convey rainfall seasonality and interannual variability, rain type occurrence is conditioned to meteorological covariates [*Benoit et al.*, 2020].

Conditional to each rain type, the distribution of rain across the island and site-specific rain intensity are modeled following a meta-Gaussian approach [*Allard and Bourotte*, 2015] [*Baxevani and Lennartsson*, 2015]. In this framework, rain accumulation at rain gauge locations is modeled as a non-linear transform (Fig. 2b) of a latent field (with standardized normal marginal distribution, Fig. 2c) whose spatial dependencies are used to encode the spatial distribution of rainfall throughout the island. This leads to a realistic representation of the complex distribution of daily rain accumulation across the island (Fig. 2d) and, in particular, rain intermittency at leeward locations and high daily accumulations in windward and mountain areas.

https://doi.org/10.5194/hess-2021-453
Preprint. Discussion started: 8 September 2021

2.2.2 Meta-Gaussian representation of island-scale daily rainfall

As introduced in Fig. 2b–c, rain intensity and spatial distribution are modeled jointly following a meta-Gaussian approach. For a given day, the observed rain accumulations $R_{i=1 \dots N_T}$ across a network of N_T gauges are linked to their latent counterparts Z_i (which follow a standardized Gaussian marginal distribution, i.e., $Z \sim \mathcal{N}(0,1)$) through a non-linear transform function ψ. This transformation is performed by first assuming that non-zero rain accumulations observed throughout the island in a given day follow a Gamma distribution:

$$Z_i = \psi(R_i) = \Phi^{-1}\left(\frac{N_d}{N_T} + \frac{N_w}{N_T} \times Gamma(R_i; k, \theta)\right) \; if \; R_i > 0 \qquad (1)$$

where N_d, N_w are the number of dry and wet gauges, Φ^{-1} is the inverse cumulative distribution function (cdf) of the univariate standardized Gaussian distribution, and $Gamma(R_i; k, \theta)$ is the cdf of the Gamma distribution with shape parameter $k > 0$ and scale parameter $\theta > 0$.

In many instances, gauges measuring no rain (i.e., R_i=0) represent a significant part of the network, which creates a concentration of zero values in rain accumulation distribution, and prevents a correct Gaussian transform using the function of Eq. (1). To circumvent this problem, the latent values corresponding to dry gauges are assigned based on the distance of the dry gauges to the closest wet gauge, such as the marginal distribution of the latent values matches the left portion of a standardized normal distribution:

$$Z_i = \psi(R_i) = \Phi^{-1}\left(\left(1 - \frac{Dw_i}{max_{j=1:N_d}(Dw_j)}\right) \times \frac{N_d}{N_T}\right) \; if \; R_i = 0 \qquad (2)$$

where Dw_i is the distance of the gauge i observing no rain to the closest gauge measuring non-zero rain. This transformation has the advantage of creating spatial patterns of censored latent values (i.e., corresponding to dry gauges) that are coherent with the ones of non-censored latent values (i.e., corresponding to wet gauges), and create smooth transitions between wet and dry domains.

Once latent values (Z_i) are derived from rain observations (R_i), the spatial distribution of rain across the island is defined by the copulas of the latent field [*Bárdossy and Pegram*, 2009], i.e., the joint cdf of Z_i. As mentioned in section 2.1, the spatial distribution of daily rainfall in high tropical islands is complex and strongly non-stationary due to orographic effects, which prevents the use of a simple parametric form (such as the multivariate Gaussian distribution used in most meta-Gaussian models of precipitation [*Benoit et al.*, 2018b] [*Papalexiou and Serinaldi*, 2020]) for the spatial copulas. Hence, in the present case, empirical copulas are used to model the spatial distribution of rainfall [*Rüschendorf*, 2009].

2.2.3 Rain typing

Based on the above meta-Gaussian representation of daily rain fields, days with similar rain statistics are pooled into rain types (Fig. 2a) using a non-supervised clustering applied on the six-dimensional feature-space defined by the following:

- The three parameters of the transform function (ψ) (i.e., $p_0 = \frac{N_d}{N_T}$, k, θ), which inform the marginal distribution of daily rainfall.
- The first three components of the Karhunen-Loève expansion [*Huang et al.*, 2001] of the latent field Z (PC_1, PC_2, PC_3), which inform the spatial distribution of rainfall across the island.

Based on this feature-space $\boldsymbol{Y} = (p_0, k, \theta, PC_1, PC_2, PC_3)^T$, the clustering is performed using a Gaussian Mixture Model (GMM, [*Fraley and Raftery*, 2002]) which approximates the pdf of $\mathbf{Y}$ as a weighted sum of multivariate Normal distributions:

$$p_{\boldsymbol{Y}}(\mathbf{Y} = \mathbf{y}) = \sum_{l=1:N_C} b_l \times \mathcal{N}(\boldsymbol{y}|\, \boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l) \tag{3}$$

where $p_{\boldsymbol{Y}}$ is the joint pdf of the random vector $\boldsymbol{Y}$, N_C is number of components in the GMM, b_l is a weight assigned to the l^th^ component, and $\boldsymbol{\mu}_l$ and $\boldsymbol{\Sigma}_l$ are the mean vector and covariance matrix of the multivariate normal distribution of the l^th^ component. Here, the parameters embedded in the vector $\boldsymbol{Y}$ are assumed to be only slightly correlated and the covariance matrices ($\boldsymbol{\Sigma}_l$) are therefore assumed to be diagonal. The number of components of the GMM (N_C) is selected by minimization of the Bayesian Information Criterion (BIC [*Schwartz*, 1978]) estimated for different numbers of components in order to select a parsimonious classification (i.e., with as few rain types as possible) while properly fitting the pdf of $\mathbf{Y}$ (i.e., $p_{\boldsymbol{Y}}$). Once the pdf $p_{\boldsymbol{Y}}$ is known, the probability that an observed vector $\boldsymbol{y}_{obs}$ belongs to the l^th^ component C_l is given by:

$$p(\boldsymbol{y}_{obs} \in C_l) = \frac{b_l \times \mathcal{N}(\boldsymbol{y}_{obs}|\, \boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l)}{\sum_{k=1}^{N_C} b_k \times \mathcal{N}(\boldsymbol{y}_{obs}|\, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)} . \tag{4}$$

And the classification is obtained by assigning each day (d_i) with a rain type (RT) that corresponds to the most probable mixture component:

$$RT(d_i) = max_{l \in 1..N_C}(p(\boldsymbol{y}_i \in C_l)). \tag{5}$$

2.2.4 Rain type occurrence

Once rain types have been defined based on rainfall statistical properties, their occurrence is conditioned to the vector $\boldsymbol{MC}_d$ of meteorological covariates observed at day d (Fig. 2a) is modeled by a non-homogeneous Markov Chain of order 1 [*Vrac et al.*, 2007]:

$$p(RT_d = j|RT_{d-1} = i, \boldsymbol{MC}_d) = \gamma_{ij} \exp\left(-\frac{1}{2}\left(\boldsymbol{MC}_d - \boldsymbol{\mu}_{ij}\right)\boldsymbol{\Sigma}_{ij}^{-1}\left(\boldsymbol{MC}_d - \boldsymbol{\mu}_{ij}\right)^T\right) \tag{6}$$

Where RT_d is the state of the Markov chain (i.e., the rain type) at day d, $p(RT_d = j|RT_{d-1} = i, \boldsymbol{MC}_d)$ is the probability to transition from rain type i to rain type j, $\boldsymbol{\Sigma}_{ij}$ and $\boldsymbol{\mu}_{ij}$ are the covariance matrix and the mean vector of the meteorological covariates when the transition from type i to type j occurs, and γ_{ij} is the baseline (i.e., long term average) probability of transition from type i to type j. This model allows the transition probability $p(RT_d = j|RT_{d-1} = i, \boldsymbol{MC}_d)$ to vary proportionally to the conditional density of $\boldsymbol{MC}_d$ given the transition and conditions the occurrence of rain types to the state of the atmosphere characterized by the covariates. Conditioning rain type occurrence to meteorological covariates informs the seasonality and the interannual variability of rain type occurrence.

https://doi.org/10.5194/hess-2021-453
Preprint. Discussion started: 8 September 2021

2.3 Model implementation

2.3.1 Selection of meteorological covariates

The set of meteorological covariates used for the conditioning of the non-homogeneous Markov Chain must be chosen so that: (i) the covariates are only weakly correlated to each other, which ensures model parsimony (i.e., minimal redundancy between covariates); and (ii) the temporal variations of the covariates are correlated with variations in rain type occurrence, which informs the seasonality and interannual variability of rainfall patterns. Note that the conditioning to covariates (i.e., the non-homogeneous part of the Markov chain) is used to inform the low frequency fluctuations of rain type occurrence (seasonal to interannual time scales), with higher frequencies (weekly to daily time scales) being informed by the baseline transition probabilities (γ_{ij}). Hence, meteorological covariates are aggregated at the monthly scale prior to use for the conditioning of the non-homogeneous Markov chain. The monthly-aggregated covariates inform monthly anomalies in atmospheric conditions and, in turn, the likelihood of rain types to occur during a given month. In addition to linking monthly atmospheric circulation conditions to daily rain patterns, this aggregation leads to a conditioning scheme that is compatible with the temporal resolution of General Circulation Model (GCM) projections [*Eyring et al.*, 2016] [*Copernicus*, 2021], which paves the way for the use of the present model for stochastic precipitation downscaling of GCM projections.

In the present case, we selected the meteorological covariates according to our initial knowledge about rain generation mechanisms in high tropical islands, and their links with the state of the atmosphere [*Elison Timm et al.*, 2014] [*Réchou et al.*, 2019] [*Sanfilippo*, 2020]; this led to the following five covariates.

1) Geopotential height at 700 hPa ($m^2.s^{-2}$). This covariate is correlated with the presence of synoptic-scale weather systems at the vicinity of the island and identifies regional atmospheric disturbances.
2) Temperature difference between 950 hPa and 700 hPa (K). This covariate is correlated with the lower atmospheric instability and identifies days prone to shallow convection.
3) Specific humidity at 700 hPa ($kg.kg^{-1}$). This covariate informs the presence of humidity above the height of the trade wind inversion and is negatively correlated with the strength of the inversion and positively correlated with the potential for deep convection and cold rain.
4) Meridional and 5) longitudinal humidity fluxes at 950 hPa (i.e., specific humidity multiplied by the *u* (east-west) or *v* (north-south) components of the wind field, $m.s^{-1}.kg.kg^{-1}$). These covariates provide the amount of moisture crossing over the mountain barrier available for precipitation and are a proxy for orographic precipitation.

2.3.2 Model calibration

The model is calibrated from a training dataset made of N days of rain accumulation recorded by a network of N_T rain gauges (Fig. 1a). Data must be available for all stations and all days of the calibration period, and a preliminary gap-filling step is required in case of incomplete data [*Longman et al.*, 2018] [*Oriani et al.*, 2020]. Once a complete training dataset is available, the first step of model calibration consists of inferring the parameters of the transform function (ψ) for each day of the training period. This is performed by calculating the proportion of dry gauges and then estimating the parameters of the gamma distribution of the wet gauges using a maximum likelihood approach. Once the three parameters of ψ are known, this function can be inverted to derive the latent values at each gauge location.

After calibration of the transform function and derivation of the latent values for each day of the calibration dataset, days with similar rain statistics are pooled together by rain typing. The first three principal components of the latent field are preliminarily derived from the Karhunen-Loève transform of all latent values. Next, the parameters of the GMM model are inferred using an expectation-minimization approach [*Fraley and Raftery*, 2002]. Finally, rain typing (i.e., clustering) is performed by assigning to each day the type that corresponds to the most probable component of the GMM model.

After rain typing, the time series of observed rain types is analyzed in relation to observations of the meteorological covariates to calibrate the non-homogeneous Markov chain. The baseline transition matrix (γ_{ij}) is first estimated by counting the transitions between each pair of rain types occurring during the calibration period and normalizing the result by the total number of transitions. Next, the parameters of the mean vector ($\boldsymbol{\mu}_{ij}$) and the covariance matrix ($\boldsymbol{\Sigma}_{ij}$) used to make the Markov chain non-homogeneous are estimated by the method of moments applied to covariates observations.

Conditional to each rain type, the joint distribution of the parameters of ψ is inferred by multivariate kernel density estimation using a trivariate Gaussian kernel. The bandwidth of the kernel is selected following the Scott's rule [*Scott*, 2010], i.e., in the present case:

$$\sqrt{\boldsymbol{H}_{ii}} = N^{-\frac{1}{7}} \times \sigma_i \tag{7}$$

where $\boldsymbol{H}$ is the bandwidth matrix of the kernel, N the number of days in the calibration dataset, and σ_i the standard deviation of the ith parameter (here i=1..3). Finally, because the spatial copulas of the latent field are simulated using an analog approach (cf next sub-section for details), they do not require formal estimation of their pdf.

2.3.3 Stochastic rainfall generation

After model calibration, stochastic rainfall generation is performed following the steps summarized in Fig. 2. Starting from a time series of meteorological covariates, rain types are first simulated using the non-homogeneous Markov chain described in Eq. (6). Next, conditional to this simulated rain type time series, the parameters of the transform function are sampled from their joint distribution defined by Eq. (7). Then, the spatial copulas of the latent field are simulated by randomly picking the empirical copulas of a day belonging to the same rain type as the day to simulate from the calibration dataset. Finally, the simulated rain field is obtained by back-transformation of the simulated latent field (Eq. 1–2) using the simulated parameters of the transform function.

2.4 Model assessment

The ability of the model to identify climatologically relevant rain types is first assessed qualitatively by applying rain typing to the full study dataset of section 2.1 and scrutinizing the emergent spatial-temporal rainfall patterns for each type. The

https://doi.org/10.5194/hess-2021-453
Preprint. Discussion started: 8 September 2021

Hydrology and Earth System Sciences
Discussions

resulting classification is subsequently interpreted in terms of rain generation processes by confronting rain types with co-occurring meteorological covariates. However, in doing so, one should keep in mind that the rain typing procedure is fully statistical and that the rain type description is based on emerging statistical patterns, not on physical modeling (e.g., using a numerical weather model to reproduce the observed patterns).

When discussing rain types and their link to rain generation processes, special attention is paid to:

(1) The emergence of spatial patterns in relation with orographic effects;
(2) The seasonality of rain type occurrence in relation with the regional annual rain cycle;
(3) The relationship of rain types with the state of the atmosphere quantified by the set of climate covariates described in section 2.4 and used here at a daily resolution (i.e., not aggregated at the monthly scale as is the case for the conditioning of the non-homogeneous Markov chain).

After the qualitative assessment of the climatological realism of rain types, the ability of the model to stochastically generate rainfall is assessed quantitatively using a leave-one-year-out cross-validation procedure. Data from one year are iteratively removed from the study dataset of section 2.1 and the stochastic model is calibrated using the remaining data (i.e., 19 years of data are used for model calibration). The model is fully recalibrated, which includes rain typing, inference of the transform function, and creation of a training dataset of spatial copulas. After model calibration, daily rainfall is simulated for each day of the target year, i.e., the year excluded from the calibration dataset. Fifty simulations are generated to assess the uncertainty associated with stochastic rainfall generation. The same procedure is repeated for each year of the study dataset, which leads to a 20-year long validation set made of 50 simulations for each gauge of the O‘ahu rain-monitoring network. Finally, simulation results are compared to observations. The following evaluation statistics are used to assess the ability of the model to simulate daily rainfall.

(1) Site-specific rainfall time series. The following statistics are considered for the four target stations of Fig. 1a: quantiles 10%, 50% and 90% of monthly rain accumulation to assess seasonality; annual rain accumulation to assess interannual variability; quantile-quantile (q-q) plot of the percentiles of daily rain accumulation to assess the probability distribution of daily rainfall; and q-q plot of the percentiles of wet-spell duration to assess rain persistence.
(2) Spatial patterns of rain distribution across the island. The following statistics are mapped to investigate the spatial distribution of rainfall: quantiles 10%, 30%, 50%, 70% and 90% of daily rain to assess how the probability distribution of rainfall varies in space.
(3) Areal rain statistics. Q-q plots of the percentiles of (i) the proportion of dry rain gauges, (ii) mean and (iii) max of daily rain, and (iv) the coefficient of variation of rain accumulation across the island to assess island-scale statistics.

https://doi.org/10.5194/hess-2021-453
Preprint. Discussion started: 8 September 2021

3 Results

3.1 Rain types in O‘ahu

Figure 3 displays the 22 rain types identified for O‘ahu Island during the period 1991–2011. The key attribute of the resulting classification is that although no information is given to the classifier about geographical coordinates, time of occurrence, or meteorological covariates, the identified rain types display well-defined patterns of spatial rain distribution (Fig. 3a), seasonality of occurrence (Fig. 3a), and correlation with the regional state of the atmosphere (Supplementary Material 1).

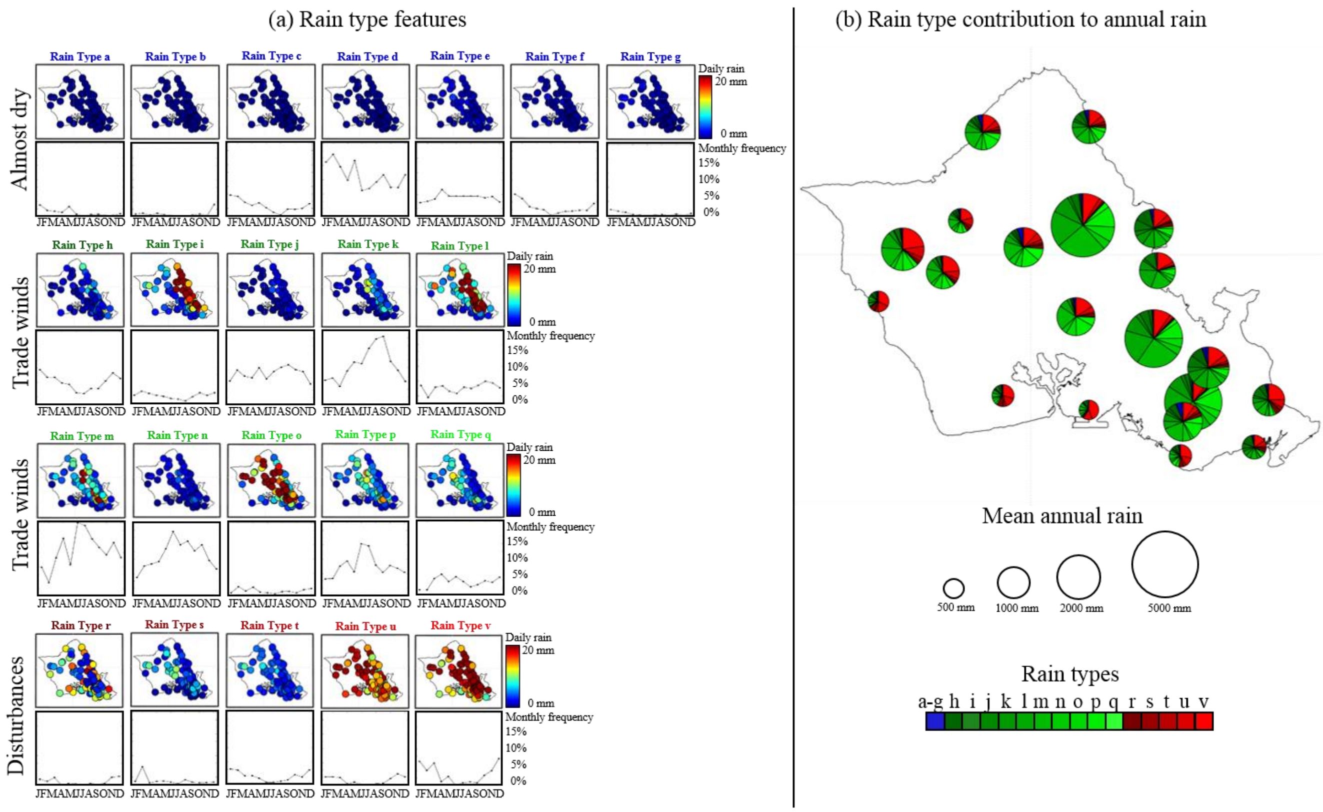

Figure 3: Rain types identified for the island of O‘ahu. (a) Spatial distribution of daily rain and frequency of occurrence of each rain type. (b) Contribution of each rain type to the annual rain accumulation for a selection of 20 gauges spread throughout the island. The color code of the pie charts in (b) is the same as the names of the types in (a).

To better identify the main modes of rainfall variability over O‘ahu, rain types are pooled into three hyperclasses (H1-3) that can be linked to the three main rain generation processes in the area (Fig. 3):

- (H1) Almost dry days (Fig. 3, rain types a–g). During these days, most rain gauges report no rain, and no gauge reports more than 5 mm/day on average. These types of weather conditions are associated with a stable atmosphere and a low moisture flux (Fig. SM 1.1).

https://doi.org/10.5194/hess-2021-453
Preprint. Discussion started: 8 September 2021

Hydrology and Earth System Sciences
Open Access
Discussions

- (H2) Trade wind days (Fig. 3, rain types h–q). This category displays well-defined spatial patterns of rain accumulation caused by orographic lifting, and are associated with a stable atmosphere, a well-defined trade wind inversion, and an important influx of moisture below the inversion layer under the influence of east-northeasterly trade winds (Fig. SM 1.1). When scrutinizing inter-type variability within this category, note that the location of the rain maximum shifts westward with increasing moisture flux, likely due to stronger trade winds causing an overshoot of orographic rain enhancement whereby rain forms over the mountains but falls further downwind on the leeward side [*Daly et al.*, 2017]. In addition, for similar wind conditions and, therefore, spatial patterns (compare for instance types j, k and l), rain intensity is correlated to the instability of the atmosphere (Fig. SM 1.1).
- (H3) Regional atmospheric disturbance days (Fig. 3, rain types r-v). These types display either unstructured (types r–t) or relatively homogeneous (types u–v) spatial patterns of rain accumulation and are associated with low pressure, unstable atmosphere, and absent (or weak) trade wind inversion. This allows high moisture content at high altitude (Fig. SM 1.1). These rain types mostly occur during winter, i.e., the local rainy season. When scrutinizing inter-type variability within this category note that rain intensity increases with atmospheric instability and the presence of humidity at high altitude, and that the spatial patterns tend to become more structured when the low-level moisture influx increases (probably due to stronger and more uniform winds).

Hence, rain typing provides new insights on island-scale rain climatology (Fig. 3b). In particular, this step helps us gain a better understanding of how different atmospheric conditions lead to different rain generation processes that, when interacting with island topography, generate contrasting orographic effects. In the case of the island of O‘ahu, orographic rain enhancement occurring during days influenced by trade winds is the main explanation for the high annual rain accumulations in the Ko‘olau mountains (up to 5000 mm annual rainfall), while widespread rainfall linked to regional atmospheric disturbances is the main source of rain at leeward locations despite their relative temporal scarcity.

3.2 Simulation of site-specific rainfall time series

Figure 4 displays the results of the cross-validation procedure (50 realizations are drawn) for four rain gauges experiencing different rainfall climatologies (Fig. 1).

https://doi.org/10.5194/hess-2021-453
Preprint. Discussion started: 8 September 2021

Hydrology and Earth System Sciences Discussions

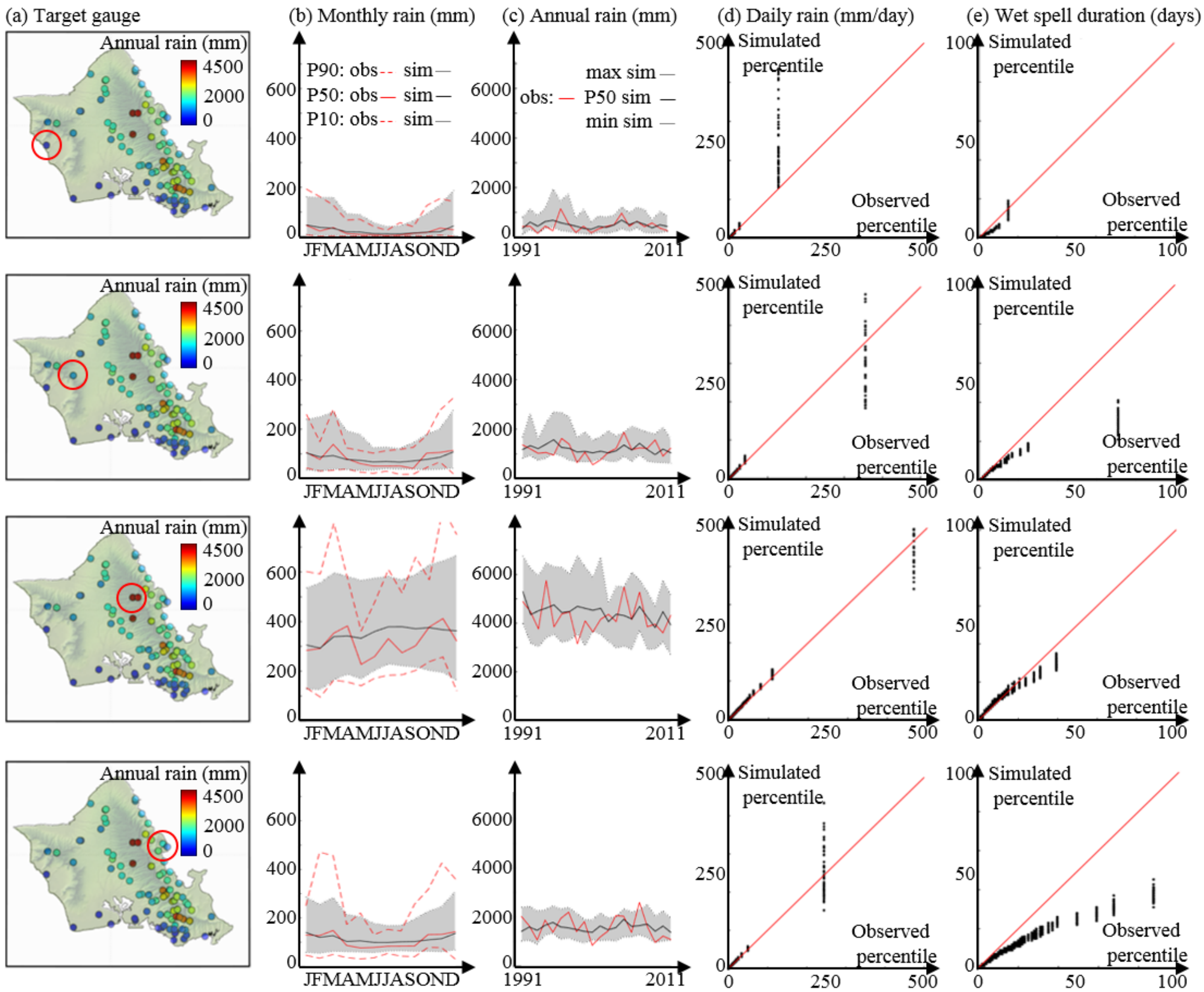

Figure 4: Ability of the model to simulate site-specific rain statistics on O‘ahu. (a) Target locations. (b) Observed (black) and simulated (red) monthly rain accumulation. Dashed lines denote quantiles 10% and 90%, and solid lines denote the median value. For simulated values, each statistic is estimated as the median across the 50 realizations. (c) Observed (black) and simulated (red) annual rain accumulation. For simulations, dashed lines denote the minimum and maximum of the 50 realizations, and the solid line denotes the median of simulations. (d) Q-q plot of daily rain percentiles. (e) Q-q plot of wet-spell duration percentiles. Black dots line up vertically in q-q plots (d–e) because for each percentile, 50 simulations are compared to a single observation.

The results in Fig. 4 show that the proposed model correctly simulates rainfall seasonality (Fig. 4b) and interannual variability (Fig. 4c). Note that simulations capture both the stronger seasonality at leeward locations (compared to windward locations) as well as the near absence of seasonality at the wettest gauge located in Ko‘olau Mountains (Fig. 4, third row). The

https://doi.org/10.5194/hess-2021-453
Preprint. Discussion started: 8 September 2021

interannual variability of rain accumulation is also properly simulated, in particular, at leeward locations where the impact of winter storms is the highest. These results suggest that the non-homogeneous Markov chain of order 1 conditioned to monthly-aggregated meteorological covariates adequately models the long-term variability of rain accumulation, and that the selected covariates properly capture rain type occurrence in a tropical marine climate.

However, rain persistence is slightly underestimated at some locations, especially for the high percentiles, i.e., long-lasting wet spells (Fig. 4e). This result exposes limitations in the use of the non-homogeneous Markov chain of order 1 for modeling weekly- to monthly-scale temporal variability of rainfall. This may be explained by the fact that daily-scale and seasonal-scale rainfall fluctuations are informed, respectively, by the Markov chain of order 1 and conditioning to monthly-aggregated meteorological covariates, but that the weekly- to monthly-scale is not explicitly included in the model. Nevertheless, the resulting errors are of low amplitude and the simplicity of the selected order 1 non-homogeneous Markov chain model justifies this small underestimation of persistence.

The simulations properly reproduce site-specific marginal distributions of daily rain accumulation (Fig. 4d). The satisfactory simulation of rainfall distribution at several sites suggests that a type-dependent gamma distribution is an adequate model for the non-zero daily rain accumulations across the island. It is noteworthy that all percentiles of the marginal distribution of rain accumulation are properly reproduced in simulations (for all four gauges), which suggests that our model is able to simulate the whole spectrum of daily rains, from dry days to intense rains.

3.3 Simulation of island-scale rain fields

Figure 5 displays the results of the cross-validation procedure focusing on island-scale features. Figure 5a compares observed and simulated spatial patterns for five quantiles of daily rain accumulation across the island of Oʻahu. Results show very good model performance in reproducing the spatial patterns of daily rainfall. This result was expected because the use of empirical copulas combined with rain typing is almost equivalent to resampling the observed spatial patterns conditional to meteorological covariates. However, satisfactory simulation results ensure that the rain-type-based resampling of spatial copulas is unbiased and that the choice and calibration of the meta-Gaussian model are relevant for the study island.

Hydrology and Earth System Sciences
Open Access
Discussions

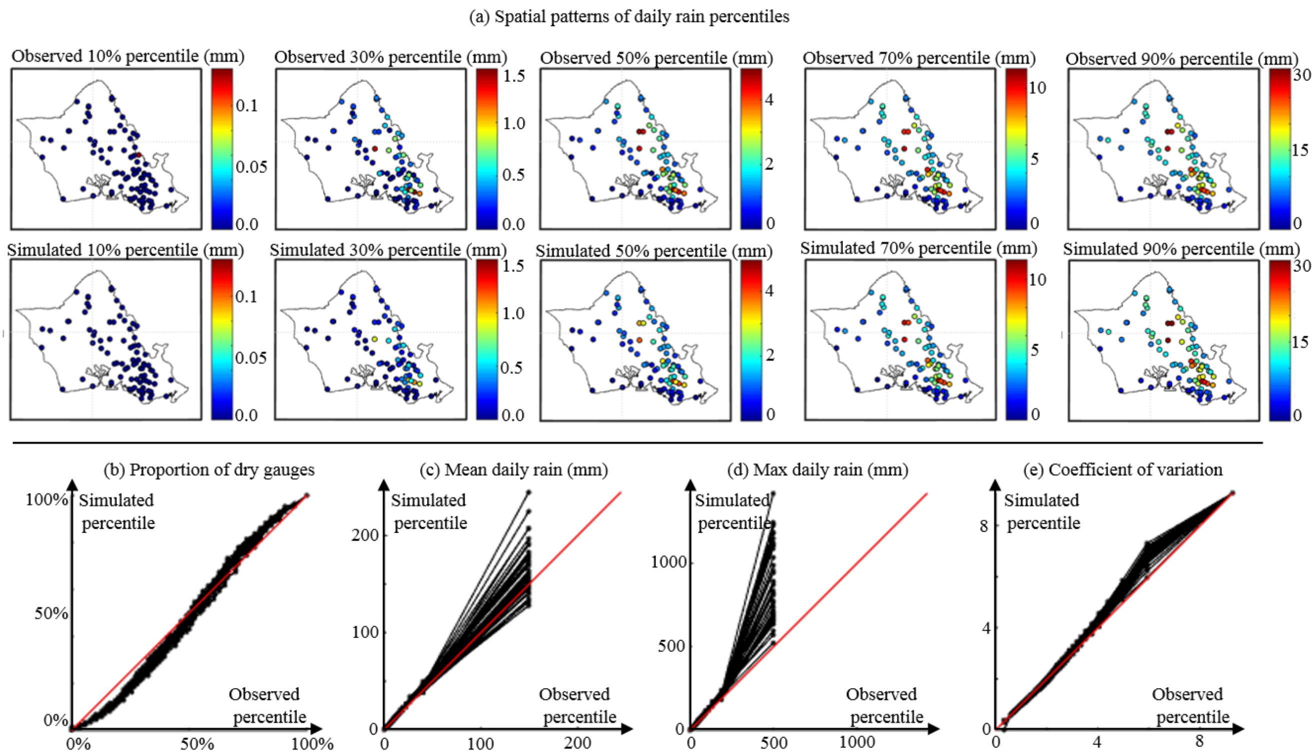

Figure 5: Assessment of island-scale statistics simulation in O'ahu. (a) Spatial patterns of observed (upper row) and simulated (lower row) percentiles of daily rain accumulation. From left to right: 10%, 30%, 50%, 70% and 90% percentiles. (b–d) Q-q plots of key rain statistics aggregated over the whole rain gauge network: (b) proportion of dry gauges; (c–d) mean and max daily rain; (e) coefficient of variation.

Figure 5b–e assesses the ability of the model to simulate four key rain statistics—the proportion of dry gauges, mean and max of daily rain accumulation, and coefficient of variation of daily rain across the island—aggregated over all rain gauges of the rain monitoring network of O'ahu. Results show a slight underestimation of the low percentiles of the proportion of dry gauges, which is compensated by the slight overestimation of the high percentiles (Fig. 5b). This level of accuracy in the simulation of the rain fraction shows that a truncated Gaussian latent field is an appropriate model for rain intermittency. In addition, the correct simulation of the spatial patterns of dry locations in Fig. 5a suggests that the distance-based modeling of the censored latent values (Eq. 2) coupled with empirical copulas is a proper model for the spatial distribution of dry locations. Similarly, the good agreement between observed and simulated coefficients of variation (Fig. 5e) coupled with the correct simulation of spatial patterns of non-zero daily rain accumulation in Fig. 5a suggest that the selected meta-Gaussian framework captures the spatial distribution of non-zero rain accumulations.

Finally, Fig. 5c–d shows that island-scale daily mean and maximum rain accumulation are properly simulated, despite an overestimation of the last percentile of the maximum, i.e., the 20 year maximum observed over the whole island. This result suggests that the meta-Gaussian framework coupled with the kernel estimation of the transform function parameters performs reasonably well to reproduce the marginal distribution of island-scale rain accumulation. However, the attempt to reproduce both island scale statistics and site-specific marginal distributions (from dry days to heavy rains) results in an inaccurate simulation of the island-scale 20-years extreme precipitation. This limitation calls for additional developments before the proposed model can be used for simulating extremes in a spatial context [*Opitz et al.*, 2021].

3.4 Model versatility

To investigate the flexibility of the above model, the case study performed in sections 3.1–3.3 for the island of O‘ahu (Hawai‘i, USA) located in the North Pacific was repeated in supplementary material 2 for the island of Tahiti (French Polynesia) located in the South Pacific. This additional cross-validation shows that our model also performs very well for Tahiti, despite a wetter (annual rain reaches 10 000 mm in Tahiti) and more seasonal climate than the O‘ahu case study. In addition, the model adapts automatically to different dataset sizes (86 rain gauges x 21 years for O‘ahu, 26 gauges x 11 years for Tahiti) due to the selection of different numbers of rain types. The above results suggest that our model may be adapted to most high tropical islands across the globe.

4 Discussion and conclusion

4.1 Discussion: stochastic modeling of orographic rainfall patterns

Validation results in section 3 show that the proposed model is able to accurately reproduce site-specific and island-scale daily rain statistics for two different tropical islands. This has been made possible by a hierarchical model structure with two main components (rain typing and meta-Gaussian representation of island-scale daily rainfall), which replicates the spatial rainfall patterns caused by orographic effects.

The first component consists of rain types, which summarize island-scale rain statistics. Unlike weather type based approaches [*Ailliot et al.*, 2015] [*Réchou et al.*, 2019], we define rain types based on rain features only, i.e., no information about meteorological covariates or large-scale circulation are included during the classification step. This leads to a classification centered on rainfall intensity and spatial distribution, which allows us to explore how island-scale rainfall variability is impacted by orographic effects (section 3.1). The links between rain types and local climate are established in a second step by conditioning the non-homogeneous Markov model of rain type occurrence to meteorological covariates. We conceptualize rain types as the main modes of island-scale daily rainfall variability, which is assumed to be primarily influenced by orographic effects caused by interactions between changing atmospheric conditions and fixed island topography. In this context, one interesting contribution of this study is the refinement of the meteorological predictors proposed by [*Sanfilippo*, 2020] for rain type occurrence in a tropical marine climate, in particular, to distinguish between shallow convection occurring during typical trade wind situations and deeper convection in the vicinity of atmospheric disturbances.

The second component of the model consists of a meta-Gaussian representation of island-scale daily rainfall. By explicitly separating rain intensity and spatial distribution, this representation contributed to the performance of the rain typing procedure detailed above and in the identification of rain types with well-defined spatial patterns. When used for stochastic rainfall generation, the adopted meta-Gaussian representation performed well in simulating site-specific rain statistics as well as island-scale spatial patterns of daily rain accumulation. This good performance can be explained by two factors. First, the determination of the censored latent values based on the distance to the closest wet gauge (Eq. 2) generates realistic spatial patterns of dry areas and dry-wet transition [*Schleiss et al.*, 2014]. This contributes to the proper modeling of the spatial intermittency of daily rain fields in tropical islands, which is caused by the drying effect of sinking air masses after crossing mountains. The second innovation of the model is the joint use of empirical copulas and a parametric transform function to model the spatial patterns of non-zero rains. It has the advantage of faithfully preserving the spatial rainfall patterns while generating unobserved values through the kernel density estimation of the transform function parameters distribution. The choice of mimicking the observed spatial rainfall patterns as closely as possible is justified by the complexity of orographic effects and associated rain gradients in tropical islands [*Giambelluca et al.*, 2013] [*Laurent et al.*, 2019] [*Benoit et al.*, 2021].

4.2 Concluding remarks

In this paper we presented a new stochastic daily rainfall generator dedicated to high tropical islands. The combination of (i) a hierarchical approach based on rain typing, (ii) a non-homogeneous Markov model of rain type occurrence conditioned to meteorological covariates, and (iii) a meta-Gaussian representation of the spatial distribution of daily rainfall allowed us to generate realistic daily rain fields honoring both site-specific and island-scale rain statistics. The performance of the model was carefully tested and illustrated for the islands of O‘ahu (Hawai‘i, USA) and Tahiti (French Polynesia), both located in the tropical Pacific. Cross-validation results prove the ability of the model to capture and simulate the main features of daily rainfall over these two high tropical islands.

The main strength of our model is its ability to simulate diverse spatial patterns of daily rainfall, as well as their linkage with regional atmospheric conditions. It represents a new tool for stochastic investigation and modeling of orographic rain enhancement on tropical islands with complex topography. The main limitation is the imperfect simulation of spatial extremes, which calls for caution when using our model for flood risk assessment.

Because of the above strengths and limitations, the main envisioned applications relate to impact studies that require detailed knowledge of daily precipitation in tropical islands, in particular, when the spatial distribution of rainfall plays an important role. This includes watershed water resources management and eco-hydrological studies. Our model can also be used for the stochastic downscaling of future precipitation projections and can contribute to the current efforts to better understand, manage, and secure tropical island water resources in a changing climate.

https://doi.org/10.5194/hess-2021-453
Preprint. Discussion started: 8 September 2021

Code and data availability.

The implementation of the proposed stochastic rainfall model is open source (MATLAB implementation) and freely available in the following repository (https://github.com/LionelBenoit/StochasticRainfallGenerator_TropicalIslands). The dataset of daily rainfall observations on Oʻahu is open data and freely available on the Hawaiʻi Climate Data Portal (https://www.hawaii.edu/climate-data-portal/data-portal/). An extract of this dataset is available in MATLAB format as a code demo in the same repository as the source code of the model. The dataset of daily rainfall observations on Tahiti is available upon request from Météo France (contact.polynesie-francaise@meteo.fr) and Groupement d'Etudes et de Gestion du Domaine Public de Polynésie Française (secretariat@equipement.gov.pf).

Author contributions.

LB, LS and TWG designed the experiment. MPL and LS compiled the daily rainfall datasets of Oʻahu and Tahiti respectively. LB and ADN selected the meteorological covariates and designed the non-homogeneous Markov Chain. LB and MPL designed the meta-Gaussian model and the rain typing method. LB implemented the model and performed the numerical experiments. LB wrote the paper with input and corrections from all co-authors.

Acknowledgments.

The work of Lionel Benoit is funded by the Swiss National Science Foundation (SNSF), grant number P2LAP2_191395. The work of Lydie Sichoix is supported by the Government of French Polynesia - Ministère de la Recherche through the project E-CRQEST, grant number 05832 MED 08/26/2019. The authors are grateful to the Hawaiʻi Climate Data Portal for providing the daily rainfall dataset of the island of Oʻahu, and to the French Weather Agency (Direction Interrégionale en Polynésie française - Météo France) and the Polynesian public service named Direction de l'Equipement (Groupement d'Etudes et de Gestion du Domaine Public de Polynésie Française - GEGDP) for providing the daily rainfall dataset of the island of Tahiti. The authors are grateful to May Izumy from the Publication services of the School of Ocean and Earth Science and Technology, University of Hawaiʻi for proofreading this manuscript.

References

Ailliot, P., D. Allard, V. Monbet, and P. Naveau (2015), Stochastic weather generators: an overview of weather type models, Journal de la Société Française de Statistiques, 156(1), 101-113.

Allard, D., and M. Bourotte (2015), Disaggregating daily precipitations into hourly values with a transformed censored latent Gaussian process, Stochastic Environmental Research and Risk Assessment, 29, 453-462.

Bárdossy, A., and G. G. S. Pegram (2009), Copula based multisite model for daily precipitation simulation, Hydrology and Earth System Sciences, 13, 2299-2314.

Bauer, P., A. Thorpe, and G. Brunet (2015), The quiet revolution of numerical weather prediction, Nature, 525, 47-55.

Baxevani, A., and J. Lennartsson (2015), A spatiotemporal precipitation generator based on a censored latent Gaussian field, Water Resources Research, 51, 4338-4358.

Benoit, L., M. Vrac, and G. Mariethoz (2018), Dealing with non-stationarity in sub-daily stochastic rainfall models, Hydrology and Earth System Sciences, 22, 5919–5933.

Benoit, L., D. Allard, and G. Mariethoz (2018b), Stochastic Rainfall Modeling at Sub-kilometer Scale, *Water Resources Research*, *54*, 4108-4130.

Benoit, L., M. Vrac, and G. Mariethoz (2020), Nonstationary stochastic rain type generation: accounting for climate drivers, Hydrology and Earth System Sciences, 24, 1-14.

Benoit, L., M. Lucas, H. Tseng, Y.-F. Huang, Y.-P. Tsang, A. D. Nugent, T. W. Giambelluca, and G. Mariethoz (2021), High Space-Time Resolution Observation of Extreme Orographic Rain Gradients in a Pacific Island Catchment, Frontiers in Earth Sciences, 8, 546246.

Breinl, K., G. Di Baldassarre, M. Girons Lopez, M. Hagenlocher, G. Vico, and A. Rutgersson (2017), Can weather generation capture precipitation patterns across different climates, spatial scales and under data scarcity? Scientific reports, 7, 5449.

Brown, J. R., M. Lengaigne, B. R. Lintner, M. J. Widlansky, K. van der Wiel, C. Dutheil, B. K. Linsley, A. J. Matthew, and J. Renwick (2020), South Pacific Convergence Zone dynamics, variability and impacts in a changing climate, Nature Reviews Earth and Environment 1, 530-543.

Cappelaere, B., et al. (2020), Modeling Land Surface Fluxes from Uncertain Rainfall: A Case Study in the Sahel with Field-Driven Stochastic Rainfields, Atmosphere, 11, 465.

Caseri, A., P. Javelle, M. H. Ramos, and E. Leblois (2016), Generating precipitation ensembles for flood alert and risk management, Journal of Flood Risk Management, 9, 402-415.

Copernicus (2021), CMIP6 climate projections, edited by E. Copernicus (ESA, ECMWF), Copernicus.

Daly, C., M. E. Slater, J. A. Roberti, S. H. Laseter, and L. W. Swift (2017), High-resolution precipitation mapping in a mountainous watershed: ground truth for evaluating uncertainty in a national precipitation dataset, International Journal of Climatology, 37, 124-137.

Elison Timm, O., T. W. Giambelluca, and H. F. Diaz (2014), Statistical downscaling of rainfall changes in Hawai‘i based on the CMIP5 global model projections, Journal of Geophysical Research: Atmospheres, 120, 92-112.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor (2016), Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937-1958.

Fraley, C., and A. E. Raftery (2002), Model-Based Clustering, Discriminant Analysis, and Density Estimation, Journal of the American Statistical Association, 97(458), 611-631.

Frazier, A. G., O. E. Timm, T. W. Giambelluca, and H. F. Diaz (2018), The influence of ENSO, PDO and PNA on secular rainfall variations in Hawai‘i, Climate Dynamics, 51, 2127-2140.

Gabellani, S., G. Boni, L. Ferraris, J. Von Hardenberg, and A. Provenzale (2007), Propagation of uncertainty from rainfall to runoff: A case study with a stochastic rainfall generator, Advances in Water Resources, 30, 2061-2071.

Giambelluca, T. W., Q. Chen, A. G. Frazier, J. P. Price, Y.-L. Chen, P.-S. Chu, J. K. Eischeid, and D. M. Delparte (2013), Online Rainfall Atlas of Hawai'i, Bulletin of the American Meteorological Society, 94, 313-316.

Greene, A. M., A. W. Robertson, P. Smyth, and S. Triglia (2011), Downscaling projections of Indian monsoon rainfall using a non-homogeneous hidden Markov model, Quarterly Journal of the Royal Meteorological Society, 137, 347-359.

Hersbach, H., et al. (2018), Operational global reanalysis: progress, future directions and synergies with NWP, ERA Rep. Ser. 27, 1-63 pp, doi: 10.21957/tkic6g3wm.

Hopuare, M., M. Guglielmino, and P. Ortega (2018), Interactions between intraseasonal and diurnal variability of precipitation in the South Central Pacific: The case of a small high island, Tahiti, French Polynesia, International Journal of Climatology, 39, 670-686.

Hopuare, M., M. Pontaud, J.-P. Céron, P. Ortega, and V. Laurent (2015), Climate change, Pacific climate drivers and observed precipitation variability in Tahiti, French Polynesia, Climate Research, 63, 157-170.

Houze, R. A. (2012), Orographic effects on precipitating clouds, Reviews of Geophysics, 50, RG000365.

Huang, S. P., S. T. Quek, and K. K. Phoon (2001), Convergence study of the truncated Karhunen–Loeve expansion for simulation of stochastic processes, International Journal for Numerical Methods in Engineering, 52, 1029-1043.

Jha, S. K., G. Mariethoz, J. Evans, M. F. McCabe, and A. Sharma (2014), A space and time scale-dependent nonlinear geostatistical approach for downscaling daily precipitation and temperature, Water Resources Research, 51, 6244-6261.

Krajewski, W. F., G. Ciach, and E. Habib (2003), An analysis of small-scale rainfall variability in different climatic regimes, Hydrological Sciences Journal, 48, 151-162.

Laurent, V., K. Maamaatuaiahutapu, I. Brodien, S. Lombardo, M. Tardy, and P. Varney (2019), Atlas climatologique de la Polynésie française, 232 pp., Météo France, Délégation Interrégionale de Polynésie Française.

Longman, R. J., H. F. Diaz, and T. W. Giambelluca (2015), Sustained Increases in Lower-Tropospheric Subsidence over the Central Tropical North Pacific Drive a Decline in High-Elevation Rainfall in Hawaii, Journal of climate, 28, 8743-8759.

Longman, R. J., O. Elison Timm, T. W. Giambelluca, and L. Kaiser (2021), A 20-Year Analysis of Disturbance-Driven Rainfall on O‘ahu, Hawai‘i, Monthly Weather Review, 6, 1767-1783.

Longman, R. J., et al. (2018), Compilation of climate data from heterogeneous networks across the Hawaiian Islands, Scientific Data, 5, 180012.

https://doi.org/10.5194/hess-2021-453
Preprint. Discussion started: 8 September 2021

Lyons, S. W. (1982), Empirical Orthogonal Function Analysis of Hawaiian Rainfall, Journal of applied meteorology, 21, 1713-1729.

Marra, F., and E. Morin (2018), Autocorrelation structure of convective rainfall in semiarid-arid climate derived from high-resolution X-Band radar estimates, Atmospheric Research, 200, 126-138.

Mavromatis, T., and J. W. Hansen (2001), Interannual variability characteristics and simulated crop response of four stochastic weather generators, Agricultural and Forest Meteorology, 109, 283-296.

Morin, E., T. Ryb, I. Gavrieli, and Y. Enzel (2019), Mean, variance, and trends of Levant precipitation over the past 4500 years from reconstructed Dead Sea levels and stochastic modeling, Quaternary Research, 91, 751-767.

Niemi, T. J., J. H. A. Guillaume, T. Kokkonen, T. M. T. Hoang, and A. W. Seed (2016), Role of spatial anisotropy in design storm generation: Experiment and interpretation, Water Resources Research, 52, 69-89.

Opitz, T., D. Allard, and G. Mariethoz (2021), Semi-parametric resampling with extremes, Spatial statistics, 42, 100445.

Oriani, F., S. Stisen, M. C. Demirel, and G. Mariethoz (2020), Missing Data Imputation for Multisite Rainfall Networks: A Comparison between Geostatistical Interpolation and Pattern-Based Estimation on Different Terrain Types, Journal of Hydrometeorology, 21, 2325-2341.

Papalexiou, S. M., and F. Serinaldi (2020), Random Fields Simplified: Preserving Marginal Distributions, Correlations, and Intermittency, With Applications From Rainfall to Humidity, Water Resources Research, 56, e2019WR026331.

Paschalis, A., P. Molnar, S. Fatichi, and P. Burlando (2013), A stochastic model for high-resolution space-time precipitation simulation, Water Resources Research, 49, doi:10.1002/2013WR014437.

Paschalis, A., S. Fatichi, P. M. Molnar, P. S., and P. Burlando (2014), On the effects of small scale space–time variability of rainfall on basin flood response, Journal of Hydrology, 514, 313-327.

Peleg, N., C. Skinner, S. Fatichi, and P. Molnar (2020), Temperature effects on the spatial structure of heavy rainfall modify catchment hydro-morphological response, Earth Surface Dynamics, 8, 17-36.

Réchou, A., O. Flores, G. Jumaux, V. Duflot, O. Bousquet, C. Pouppeville, and F. Bonnardot (2019), Spatio-temporal variability of rainfall in a high tropical island: Patterns and large-scale drivers in Réunion Island, Quarterly Journal of the Royal Meteorological Society, 145, 893-909.

Richardson, C. W. (1981), Stochastic simulation of daily precipitation, temperature, and solar radiation, Water Resources Research, 17, 182-190.

Rüschendorf, L. (2009), On the distributional transform, Sklar's theorem, and the empirical copula process, Journal of Statistical Planning and Inference, 139, 3921-3927.

Sanfilippo, K. M. (2020), Predictor selection and model evaluation for future rainfall projection in Hawai'i, 124 pp, University of Hawaiʻi at Mānoa, Honolulu.

Schleiss, M., S. Chamon, and A. Berne (2014), Nonstationarity in Intermittent Rainfall: The ''Dry Drift'', Journal of Hydrometeorology, 15, 1189-1204.

Schwartz, G. (1978), Estimating the dimension of a model, The Annals of Statistics, 6, 461-464.

https://doi.org/10.5194/hess-2021-453
Preprint. Discussion started: 8 September 2021

Scott, D. W. (2010), Scott's rule, WIREs Computational Statistics, 2, 497-502.

Supit, I., C. A. van Diepen, A. J. W. de Wit, J. Wolf, P. Kabat, B. Baruth, and F. Ludwig (2012), Assessing climate change effects on European crop yields using the Crop Growth Monitoring System and a weather generator, Agricultural and Forest Meteorology, 164, 96-111.

Volosciuk, C., D. Maraun, M. Vrac, and M. Widmann (2017), A combined statistical bias correction and stochastic downscaling method for precipitation, Hydrology and Earth System Sciences, 21, 1693-1719.

Vrac, M., M. Stein, and K. Hayhoe (2007), Statistical downscaling of precipitation through nonhomogeneous stochastic weather typing, Climate Research, 34, 169-184.

Vu, T. M., A. K. Mishra, G. Konapala, and D. Liu (2018), Evaluation of multiple stochastic rainfall generators in diverse climatic regions, Stochastic Environmental Research and Risk Assessment, 32, 1337-1353.

Wilcox, C., C. Aly, T. Vischel, G. Panthou, J. Blanchet, G. Quantin, and T. Lebel (2021), Stochastorm: A Stochastic Rainfall Simulator for Convective Storms, Journal of Hydrometeorology, 22, 387-404.

Wilks, D. S., and R. L. Wilby (1999), The weather generation game: a review of stochastic weather models, Progress in Physical Geography, 23, 329-357.