Hydrol. Earth Syst. Sci., 30, 4799–4822, 2026
https://doi.org/10.5194/hess-30-4799-2026

Hydrology and Earth System Sciences
Open Access

A non-stationary trans-Gaussian model for daily rainfall over complex topography

Lionel Benoit[1], **Matthew P. Lucas**[2], **Denis Allard**[1], **Keri M. Kodama**[3], **and Thomas W. Giambelluca**[2]

[1]Biostatistics and Spatial Processes (BioSP), INRAE, 84914 Avignon CEDEX 9, France
[2]Water Resources Research Center, University of Hawai‘i at Mānoa, Honolulu, Hawaii, USA
[3]University of Hawai‘i Sea Grant College Program, University of Hawai‘i at Mānoa, Honolulu, Hawaii, USA

Correspondence: Lionel Benoit (lionel.benoit@inrae.fr)

Received: 9 May 2025 – Discussion started: 19 June 2025
Revised: 6 March 2026 – Accepted: 1 June 2026 – Published: 29 July 2026

Abstract. In mountainous regions orographic effects create strong horizontal gradients of various rainfall statistics such as the frequency of occurrence, the distribution of intensity and the structure of spatial correlation. However, most statistical models of daily rainfall assume spatial stationarity (i.e., the spatial homogeneity of rainfall statistics) and are therefore not well suited for studying the highly non-homogeneous characteristics of orographic rainfall. To overcome this limitation, we design a non-stationary trans-Gaussian geostatistical model for the analysis of daily rainfall fields over complex topography. This framework infers rainfall statistics from sparse rain gauge observations, simulates realistic rainfall fields after calibration and stochastically interpolates rain gauge observations to create rainfall maps. The performance of the model is assessed with data from the Island of Hawai‘i where extreme spatial gradients in rainfall are observed. Results demonstrate that the non-stationary trans-Gaussian model can skillfully reproduce orographic rainfall statistics as well as their variations in space.

1 Introduction

Stochastic rainfall models (SRMs) are probabilistic tools that simulate synthetic rainfall datasets with statistical properties that resemble those from observations (Richardson, 1981; Wilks and Wilby, 1999; Sharma and Mehrotra, 2010; Ailliot et al., 2015). These models first infer the probability distribution of rainfall from observations and then generate synthetic datasets by random sampling of this distribution, a process known as stochastic simulation. Stochastic simulation makes SRMs particularly suitable to assess the uncertainty of rainfall estimates (Vischel et al., 2009; Caseri et al., 2016) and forecasts (Bowler et al., 2006; Nerini et al., 2017), and to assess the natural variability of rainfall (Evin et al., 2018; Peleg et al., 2018). The ability of SRMs to simulate probabilistic ensembles of synthetic but realistic rainfall data makes them an important tool for rainfall-related impact studies, for instance to perform dam or drainage system sizing (Niemi et al., 2016), hydro-meteorological modeling (Mandapaka et al., 2009; Paschalis et al., 2014), or crop yield simulation (Hansen and Ines, 2005; Qian et al., 2011).

A large variety of stochastic models have been applied to rainfall modeling, the choice of which depends on the features of rainfall to reproduce in priority. For instance, point process models have been used to simulate the occurrence of the main meteorological processes responsible for rainfall generation (e.g., rain cells, rain bands, storms) (Onof et al., 2000; Cowpertwait, 2010), models combining Markov chains and parametric distributions of rain intensity (Richardson, 1981; Ailliot et al., 2009) as well as resampling approaches (Rajagopalan and Lall, 1999; Oriani et al., 2014) have been applied to time-series simulation, generalized linear models have been used to capture the relationships between atmospheric conditions and rainfall (Chandler and Wheater, 2002; Chandler, 2020), fractal approaches have shown good performance in simulating rainfall scaling properties (Molnar and Burlando, 2005; Ramanathan et al., 2022; Maloku et al., 2025), and generative AI recently demonstrated promising skills at downscaling and forecasting com-

Published by Copernicus Publications on behalf of the European Geosciences Union.

plex rainfall features (Leinonen et al., 2020; Ravuri et al., 2021; Harris et al., 2022; Srivastava et al., 2024).

When the focus is on spatial patterns and spatial dependencies, models based on Gaussian random fields (GRFs) – also known as geostatistical models – are often chosen because they enable modeling rainfall at any point of the spatial domain while accounting for spatial dependencies (Chilès and Delfiner, 2012). However, the distribution of rainfall intensity is rarely Gaussian at daily to sub-daily resolution, and a parametric transform function is often combined with a GRF to model rainfall at these time scales resulting in so-called trans-Gaussian (or meta-Gaussian) approaches (Leblois and Creutin, 2013; Benoit et al., 2018; Papalexiou and Serinaldi, 2020). Trans-Gaussian models can be used to simulate rainfall in two main settings: unconditional simulations allow for the generation of synthetic rainfall fields (Paschalis et al., 2013; Peleg et al., 2017; Vaittinada Ayar et al., 2020; Wilcox et al., 2021) while conditional simulations are used for mapping rainfall from sparse rain gauge observations (Creutin and Obled, 1982; Lanza, 2000; Kyriakidis et al., 2001) and for radar-rain gauge data fusion (Creutin et al., 1988; Foehn et al., 2018). In both cases the resulting rainfall datasets can be used to assess input errors in distributed hydrological models (Vischel et al., 2009; Renard et al., 2011) or to conduct sensitivity analysis of spatially explicit hydrometeorological modeling chains (Moraga et al., 2022; Liu et al., 2024).

In mountainous regions, the performance of trans-Gaussian models is constrained by the strong variability of rainfall statistics through space. Indeed, air masses flowing across mountains are displaced by the topography and forced into successive uplifts and downdrafts. The vertical displacement of air modulates the condensation within the air column, triggering or enhancing precipitation in the case of uplift and, conversely, inhibiting or attenuating rainfall in the case of downdraft. This impact of topography on precipitation is called the orographic effect (Roe, 2005; Houze, 2012) and results in windward slopes being generally wetter than leeward slopes and the highlands generally wetter than the lowlands (Foresti et al., 2018). Additionally, fluctuations in the direction and intensity of prevailing winds combined with changes of precipitation regimes lead to complex patterns of orographic rainfall with strong spatial gradients between wet and dry areas (Giambelluca et al., 2013; Isotta et al., 2014; Benoit and Sichoix, 2023).

In this context, the present study intends to develop a trans-Gaussian model for daily rainfall over complex topography, which requires to adapt the trans-Gaussian framework to the spatial variation of rainfall statistics caused by orographic effects. This implies making some parameters of the model location-dependent, a condition referred to as spatial non-stationarity (Sampson and Guttorp, 1992). Two approaches have been proposed so far to make trans-Gaussian rainfall models non-stationary: either the marginal distribution (i.e., the parameters modeling rainfall occurrence and intensity) is made non-stationary (Frazier et al., 2016; Benoit et al., 2021; Lucas et al., 2022), or the non-stationarity is modeled in the covariance function (i.e., the parameters modeling the spatial dependencies within rainfall fields) (Paciorek and Schervish, 2006; Fuglstad et al., 2015; Fouedjio et al., 2016). In this paper we aim to combine these two ways of modeling spatial non-stationarity, which leads to what we call a fully non-stationary trans-Gaussian model, i.e., a trans-Gaussian model with all parameters varying in space. This provides a very flexible spatial model for daily rainfall, in which the statistics characterizing (1) rainfall occurrence, (2) the marginal distribution of rainfall intensity and (3) the spatial dependencies within rainfall fields are location-dependent. The performance of the model is illustrated for the Island of Hawai‘i (State of Hawaii, USA) where orographic effects are very strong (Giambelluca et al., 2013), and for which an extensive and high-quality rain gauge dataset enables the calibration of a highly parametrized model. The model is designed and tested with two specific applications in mind: the interpolation of rain gauge data to generate daily rainfall maps for the Hawai‘i Climate Data Portal (HCDP) (Longman et al., 2024), and the stochastic generation of spatially explicit rainfall fields to foster hydrological modeling in Hawai‘i (Huang et al., 2021; Strauch et al., 2024).

The rest of the paper is structured as follows: Sect. 2 illustrates orographic precipitation on the Island of Hawai‘i and motivates the need for a non-stationary rainfall model. Section 3 describes the trans-Gaussian model being used, explains how to make it non-stationary in space and proposes a method for model calibration. Section 4 applies the model to the simulation of orographic rainfall on the Island of Hawai‘i and assesses the performance of the model. Finally, Sect. 5 discusses the advantages and limitations of the model with respect to orographic rainfall modeling and proposes some lines of inquiry for future research.

2 Example dataset: orographic precipitation on the Island of Hawai‘i

Due to the location of the Island of Hawai‘i (State of Hawaii – USA, area = 10 432 km^2, highest point: Mauna Kea – 4207 m) in the middle of the subtropical Pacific (Fig 1a), the climate of the island is shaped by prevailing easterly trade winds which blow 85 %–95 % of the time in summer and 50 %–80 % of the time in winter (Longman et al., 2021). When they reach the Island of Hawai‘i, trade winds produce the orographic lifting of moist oceanic air masses, which triggers shallow convection leading to frequent and abundant rain showers in the east facing slopes and generally dry conditions on the leeward west side of the island (Fig 1b) (Giambelluca et al., 2013). An exception to these generally dry weather conditions is a distinct rainfall maximum in the Kona area (West from Mauna Loa, at around 5 km from the coast) which is the result of mechanically diverted tradewinds being directed upslope, enhanced in the afternoon by strong solar heating of the west-facing slope (Huang and Chen, 2019).

In Hawaiʻi weather conditions are often affected by the presence of the trade wind inversion (TWI, average inversion base height around 2200 m) that caps the moist oceanic air and creates very dry conditions at high altitudes located in the center of the island (Fig 1b) (Longman et al., 2015). This predominant pattern of precipitation is drastically modified when atmospheric disturbances originating from the mid-latitudes (e.g., cold fronts or Kona lows) eliminate the TWI, produce southerly, southwesterly, or westerly winds at the regional scale, and allow deep convection to develop. Under atmospheric disturbance conditions, widespread rainfall occurs throughout the island, and the west-facing slopes as well as the high altitude locations, which are usually dry, receive a significant part of their annual precipitation (Longman et al., 2021).

To account for the diversity of climate patterns in the State of Hawaii, Luo et al. (2024) proposed a partitioning of the archipelago into twelve climate divisions (Fig 2). Climate divisions constitute the basic geographic unit used by the National Oceanic and Atmospheric Administration's National Centers for Environmental Information (NCEI-NOAA) for climate analyses at the scale of the United States of America (Guttman and Quayle, 2005). The Hawaiʻi climate divisions have been defined based on a cluster analysis on monthly rainfall maps and local expert knowledge (Luo et al., 2024), and will be used in the present study to delineate six sub-domains of the island of Hawaiʻi within which the rainfall climatology is assumed to be relatively homogeneous (Fig. 2a).

Figures 1b and 2b–c illustrate the diversity of rainfall climatology over the Island of Hawaiʻi. This diversity translates into a strong heterogeneity of rainfall statistics throughout the island, which calls for a non-stationary trans-Gaussian model for the spatial analysis of daily rainfall. Such a model is introduced in the next section and applied to the Island of Hawaiʻi in Sect. 4.

3 Non-stationary trans-Gaussian rainfall model

In the following, we model daily rainfall as a stochastic process studied over a spatial domain $\mathcal{D}$: $R(\boldsymbol{s})$ with $\boldsymbol{s} \in \mathcal{D} \subset \mathbb{R}^2$. As the focus of the study is on the spatial modeling of rainfall, we model the temporal variability separately (and beforehand) following a rain-typing approach (Ailliot et al., 2015; Blanchet et al., 2019; Vaittinada Ayar et al., 2020). We therefore assume that days can preliminarily be pooled into N_{clust} clusters referred to hereafter as rain types, within which rainfall fields are statistically similar to each other. All the information about the temporal variability of rainfall is therefore encoded into the sequence of rain types, which has been shown to be a reasonable assumption for daily rainfall in tropical islands, and in particular in Hawaiʻi (Benoit et al., 2022). As a result, daily rainfall fields are assumed independent to each other conditionally to rain types. The spatial modeling is therefore performed independently for each rain type, and for the sake of simplicity we assume a single rain type for the description of the model in Sect. 3.1 to 3.5. How to deal with multiple rain types when conditioning daily rainfall maps to monthly totals will be addressed in Sect. 3.6, and the practical delineation of rain types from an actual dataset will be addressed in Sect. 4.1.

3.1 Trans-Gaussian geostatistics applied to rainfall modeling

When observed at a single site and at daily resolution, rainfall is featured by a large number of zeros (i.e. dry days) and a distribution of non-zero rainfall intensities that strongly differs from the normal distribution (Evin et al., 2018; Blanchet et al., 2019; Benoit et al., 2022). In addition, when observed simultaneously at neighboring locations, rainfall is structured spatially with inter-site dependencies decreasing with the separation distance and vanishing at distances of dozens to hundreds of kilometers depending on the regional climate (Creutin et al., 1988; Guillot and Lebel, 1999). In this study we adopt a trans-Gaussian approach to model the above features of rainfall (Allard and Bourotte, 2015; Benoit et al., 2018; Papalexiou and Serinaldi, 2020; Vaittinada Ayar et al., 2020). Trans-Gaussian geostatistics assume that a rainfall observation $R(\boldsymbol{s})$ performed at location $\boldsymbol{s} \in \mathcal{D}$ originates from a latent and standardized (i.e., with zero mean and unit variance) Gaussian random variable $Y(\boldsymbol{s})$ through a transform function Ψ which combines a truncation at threshold a_0 and a monotonic transformation ψ of the non-censored part of the latent variable. The spatial dependencies observed between a set of N rain observations performed at different locations $R(\boldsymbol{s}_i), i = 1, \ldots, N_G$ are modeled by assuming that the associated latent variables $Y(\boldsymbol{s}_i), i = 1, \ldots, N_G$ form a random vector $\boldsymbol{Y} = [Y(\boldsymbol{s}_1), Y(\boldsymbol{s}_2), \ldots, Y(\boldsymbol{s}_{N_G})]$ that follows a standardized multivariate Gaussian distribution with covariance function C_Y.

In the following of Sect. 3.1 and in Sect. 3.2 we develop a Trans-Gaussian rainfall model tailored to the example dataset introduced in Sect. 2, which entails choosing parameterizations of ψ and C_Y that best fit the statistical signature of rainfall in the Island of Hawaiʻi. A different dataset, in particular from a different climate, could require different parameterizations of these functions without calling into question the non-stationary Trans-Gaussian framework deployed in this study.

Here we assume a Gamma distribution for single-site rain intensity (Wilks and Wilby, 1999; Berrocal et al., 2008) because this distribution has easy-to-interpolate parameters (i.e., with high auto-correlation and low cross-correlation) and because when combined with rain types the Gamma distribution has been shown to accurately model daily rainfall intensity (Blanchet et al., 2019). In addition we assume a

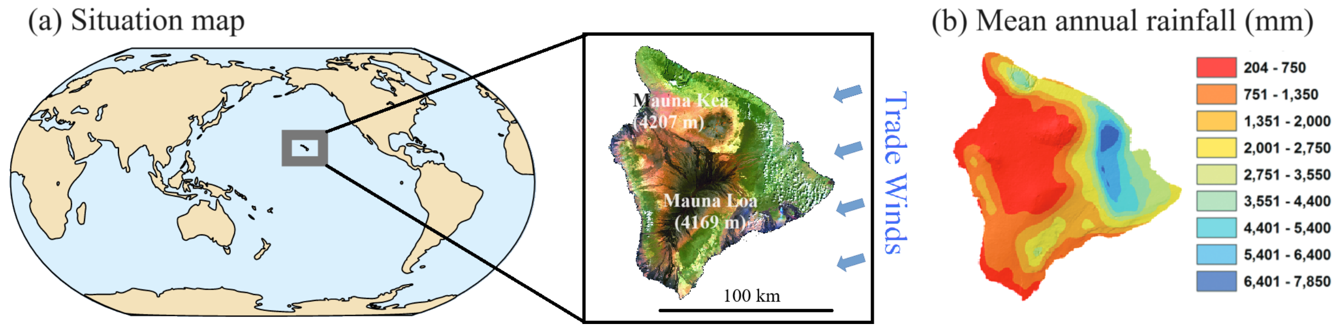

Figure 1. Geographical and climatological context of the Island of Hawai‘i. **(a)** Situation map and main topo-climatic features (adapted from an ETM+ Landsat 7 mosaic produced by NOAA Coastal Services Center as part of the Hawaii Land Cover Analysis project), **(b)** mean annual rainfall (adapted from Giambelluca et al. (2013)).

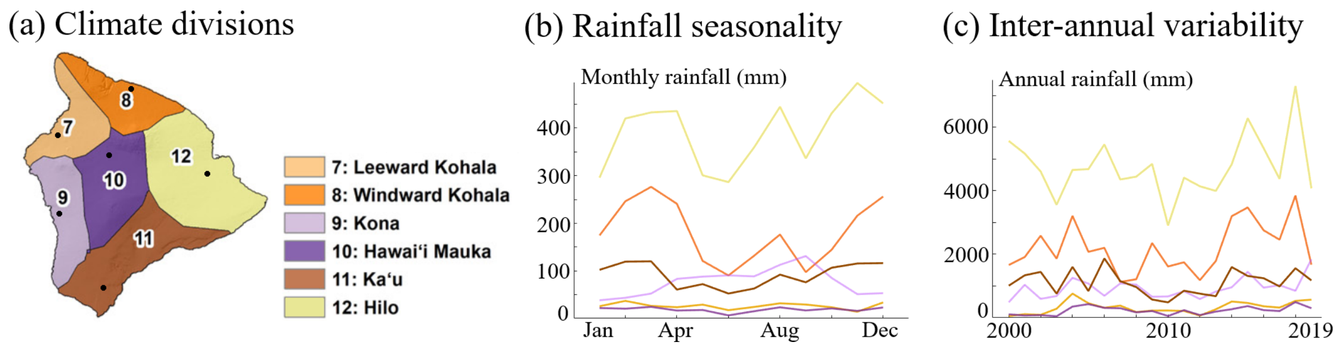

Figure 2. Climate divisions of the Island of Hawai‘i. **(a)** Footprint of the climate divisions (adapted from Luo et al. (2024); the numbering over the Island of Hawai‘i starts at 7 because the original zonation covers the whole State of Hawaii). **(b)** Rainfall seasonality and **(c)** rainfall inter-annual variability for 6 rain gauges spread in the different climate divisions (gauge locations are denoted by black dots in **a**).

Matérn covariance function for the latent Gaussian random field because this model of covariance is very flexible (Porcu et al., 2024) and has shown good skills in modeling the spatial dependencies of most atmospheric variables, including rainfall, at daily resolution (Obakrim et al., 2025). Finally, a geometric anisotropy is added to the model of covariance to allow spatial dependencies to vary with the direction of interest (Allard et al., 2016). Overall, the trans-Gaussian model for daily rainfall is defined as:

$$
\begin{aligned}
R(\boldsymbol{s}) &= \Psi(Y(\boldsymbol{s})) \\
&= \begin{cases} 0 \text{ if } Y(\boldsymbol{s}) \le a_0 \\ \mathrm{Gamma}^{-1}(\Phi(Y(\boldsymbol{s})); k, \theta) \text{ if } Y(\boldsymbol{s}) > a_0 \end{cases}
\end{aligned}
\tag{1}
$$

With $Y \sim \mathcal{N}(0, C_Y)$ and $C_Y(\boldsymbol{h}; \nu, \gamma_{a,1}, \gamma_{a,2}, \rho_b) = \frac{1}{\Gamma(\nu)2^{\nu-1}}\left(\sqrt{Q}\right)^{\nu}\mathcal{K}_\nu\left(\sqrt{Q}\right)$ and with $Q = \boldsymbol{h}^T \times \boldsymbol{\Sigma} \times \boldsymbol{h}$; $\boldsymbol{\Sigma} = \boldsymbol{V} \times \boldsymbol{\Lambda} \times \boldsymbol{V}^T$; $\boldsymbol{V} = \begin{bmatrix} \frac{\gamma_{a,1}}{\sqrt{\gamma_{a,1}^2+\gamma_{a,2}^2}} & -\frac{\gamma_{a,2}}{\sqrt{\gamma_{a,1}^2+\gamma_{a,2}^2}} \\ \frac{\gamma_{a,2}}{\sqrt{\gamma_{a,1}^2+\gamma_{a,2}^2}} & \frac{\gamma_{a,1}}{\sqrt{\gamma_{a,1}^2+\gamma_{a,2}^2}} \end{bmatrix}$; $\boldsymbol{\Lambda} = \begin{bmatrix} \gamma_{a,1}^2 + \gamma_{a,2}^2 & 0 \\ 0 & \rho_b^2 \end{bmatrix}$ where $\mathrm{Gamma}(\cdot; k, \theta)$ is the cumulative distribution function (cdf) of the Gamma distribution with parameters $k > 0$ (shape) and $\theta > 0$ (scale); Φ is the cdf of the standardized normal distribution; $\boldsymbol{h}$ is the separation vector between two locations $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$; Γ is the usual gamma function; $\mathcal{K}_\nu$ is the modified Bessel function of the second kind with parameter $\nu > 0$, and ν is the shape parameter of the Matérn covariance function that controls the regularity of the latent process Y at short separation lags; $\gamma_{a,1} \in \mathbb{R}$ and $\gamma_{a,2} \in \mathbb{R}$ are the two components of the vector defining the major anisotropy axis (whose direction corresponds to the direction of maximum spatial correlation of the latent process Y and whose norm $\rho_a = \sqrt{\gamma_{a,1}^2 + \gamma_{a,1}^2}$ controls the variability of Y in that direction); and $\rho_b > 0$ is the norm of the minor anisotropy axis vector (which is orthogonal to the major anisotropy axis vector).

3.2 Making the trans-Gaussian model non-stationary

The trans-Gaussian model of Eq. (1) is made fully non-stationary in space by allowing the parameters of the marginal distribution and those of the dependence structure to vary through space. To this end, the parameters of the transform function Ψ (i.e., a_0, k and θ) are first defined for each observation location separately, and then interpolated at ungauged locations (Blanchet et al., 2019; Benoit et al., 2021; Maloku et al., 2025). Making the parameters of the covariance function C_Y (i.e., ν, $\gamma_{a,1}$, $\gamma_{a,2}$, and ρ_b) non-stationary is more challenging because the resulting covariance function must be a valid model of covariance over the whole domain of interest (Chilès and Delfiner, 2012). Different approaches have been proposed to obtain valid non-stationary covariance models, for instance the convolution of locally stationary covariance kernels (Paciorek and Schervish, 2006; Fouedjio et al., 2016), the space deformation of the modeling domain (Sampson and Guttorp, 1992; Fouedjio et al., 2015), and the use of a locally varying diffusion operator in stochastic partial differential equations (Fuglstad et al., 2015; Pereira et al., 2022). Here we follow the convolution approach because it provides a direct extension of the Matérn covariance function defined in Eq. (1) to the non-stationary case, and because the use of locally stationary covariance kernels simplifies the calibration of the model based on sparse rain gauge observations (cf. Sect. 3.3). We adopt the parametrization of Paciorek and Schervish (2006) to derive the non-stationary and anisotropic covariance function of Matérn type between two observation locations $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$:

$$C_Y(\boldsymbol{s}_i,\boldsymbol{s}_j;\nu_i,\nu_j,\gamma_{a,1_i},\gamma_{a,1_j},\gamma_{a,2_i},\gamma_{a,2_j},\rho_{b_i},\rho_{b_j}) = \frac{2^{1-\nu_{ij}}}{\sqrt{\Gamma(\nu_i)\Gamma(\nu_j)}}|\boldsymbol{\Sigma}_i|^{\frac{1}{4}}|\boldsymbol{\Sigma}_j|^{\frac{1}{4}}\left|\frac{\boldsymbol{\Sigma}_i+\boldsymbol{\Sigma}_j}{2}\right|^{-\frac{1}{2}}\left(\sqrt{Q_{ij}}\right)^{\nu_{ij}}\mathcal{K}_{\nu_{ij}}\left(\sqrt{Q_{ij}}\right) \quad (2)$$

with $Q_{ij} = \left(\boldsymbol{s}_i-\boldsymbol{s}_j\right)^T\left(\frac{\boldsymbol{\Sigma}_i+\boldsymbol{\Sigma}_j}{2}\right)^{-1}\left(\boldsymbol{s}_i-\boldsymbol{s}_j\right)$, with $\nu_{ij}=\frac{\nu_i+\nu_j}{2}$ and with $\boldsymbol{\Sigma}_i$ and $\boldsymbol{\Sigma}_j$ defined as in Eq. (1).

3.3 Model calibration

Model calibration starts with the estimation of model parameters at the local scale assuming local stationarity, and continues with the interpolation of these parameters to create a non-stationary model. The data used for calibration are daily rainfall measurements performed by a rain gauge network encompassing a set $\mathcal{N}_{\text{etwork}}$ of N_g observation locations $\boldsymbol{s}_{\text{g}}$ positioned within the study domain $\mathcal{D}$: $\boldsymbol{s}_{\text{g}} \in \mathcal{N}_{\text{etwork}} \subset \mathcal{D}$, $g = 1, \ldots, N_g$.

3.3.1 Estimation of model parameters assuming local stationarity

The basic unit for the estimation of locally stationary parameters of the transform function is the rain gauge. Hence, for a given rain gauge at location $\boldsymbol{s}_{\text{g}}$, the truncation threshold $a_0(\boldsymbol{s}_{\text{g}})$ is obtained from the proportion $d(\boldsymbol{s}_{\text{g}})$ of dry measurements observed at this location by simply inverting the cdf of a standardized normal distribution $a_0(\boldsymbol{s}_{\text{g}}) = \Phi^{-1}(d(\boldsymbol{s}_{\text{g}}))$, and then the parameters of the marginal distribution of non-zero daily intensities (i.e. k and θ) are estimated by marginal likelihood maximization. The log-likelihood of the parameters k and θ given a set of N_{obs} non-zero rain observations performed at location $\boldsymbol{s}_{\text{g}}$ (i.e., $R_j(\boldsymbol{s}_{\text{g}}) > 0$, $j = 1, \ldots, N_{\text{obs}}$) is:

$$l(k,\theta) = \sum_{j=1}^{N_{\text{obs}}}\left(\ln\left(\frac{1}{\Gamma(k)\theta^k}R_j(\boldsymbol{s}_{\text{g}})^{k-1}e^{-\frac{R_j(\boldsymbol{s}_{\text{g}})}{\theta}}\right)\right) \quad (3)$$

The estimation of covariance parameters requires observations from several rain gauges simultaneously to capture spatial dependence, and therefore the basic unit used for the estimation of locally stationary parameters of the covariance function is the climate division. Climate divisions form sub-domains of the target area within which the climatology of rainfall is deemed relatively homogeneous (cf. Sect. 2), which allows us to make an assumption of local stationarity for the dependence structure. The size of the climate divisions used as sub-domains where local stationarity is postulated derives from the tradeoff between (i) the number of rain gauges within each division that should be maximized to improve the estimation of the covariance parameters and (ii) the distance between climate divisions that should be minimized to improve the performance of the interpolation used to build the non-stationary covariance model. After the delineation of climate divisions, the parameters of the covariance function C_Y of the latent field are estimated for each division separately using pairwise likelihood maximization, which has been proved to be an efficient and unbiased method for the estimation of covariance parameters (Bevilacqua et al., 2012). For a given pair of rain observations $R(\boldsymbol{s}_{g_1})$ and $R(\boldsymbol{s}_{g_2})$ recorded the same day by two rain gauges at locations $\boldsymbol{s}_{g_1} \in \mathcal{N}_{\text{etwork}}$ and $\boldsymbol{s}_{g_2} \in \mathcal{N}_{\text{etwork}}$ with separation vector $\boldsymbol{h} = \boldsymbol{s}_{g_2} - \boldsymbol{s}_{g_1}$ the log-likelihood of the parameters ν, $\gamma_{a,1}$, $\gamma_{a,2}$ and ρ_b can be written as (Allard and Bourotte, 2015):

$$\begin{aligned}
&l(\nu,\gamma_{a,1},\gamma_{a,2},\rho_b) = \ln(\Phi_2(a_0,a_0;C_Y(\boldsymbol{h};\nu,\gamma_{a,1},\gamma_{a,2},\rho_b)))\\
&\text{if } R(\boldsymbol{s}_{g_1}) = R(\boldsymbol{s}_{g_2}) = 0\\
&l(\nu,\gamma_{a,1},\gamma_{a,2},\rho_b) = \ln(\Phi(a_0) - \Phi_2(a_0,a_0;C_Y(\boldsymbol{h};\nu,\gamma_{a,1},\gamma_{a,2},\rho_b)))\\
&+\ln(\phi(Y(\boldsymbol{s}_{g_2})))\\
&\text{if } R(\boldsymbol{s}_{g_1}) = 0 \text{ and } R(\boldsymbol{s}_{g_2}) > 0\\
&l(\nu,\gamma_{a,1},\gamma_{a,2},\rho_b) = \ln(\phi_2(Y(\boldsymbol{s}_{g_1}),Y(\boldsymbol{s}_{g_2});C_Y(\boldsymbol{h};\nu,\gamma_{a,1},\gamma_{a,2},\rho_b)))\\
&\text{if } R(\boldsymbol{s}_{g_1}) > 0 \text{ and } R(\boldsymbol{s}_{g_2}) > 0
\end{aligned} \quad (4)$$

where Φ_2 and ϕ_2 are respectively the cdf and the probability density function (pdf) of the bivariate standard normal distribution, ϕ is the pdf of the univariate standard normal distribution, and $Y(\boldsymbol{s}_{g_1})$ is the latent counterpart of the rainfall observation $R(\boldsymbol{s}_{g_1})$.

The log-likelihoods of the marginal distribution and covariance function are maximized sequentially using the `fmincon` function in Matlab (using the default interior point algorithm (Byrd et al., 1999)). This sequential estimation scheme has shown good performance in estimating the parameters of trans-Gaussian models of rainfall (Leblois and Creutin, 2013; Caseri et al., 2016; Papalexiou and Serinaldi, 2020).

3.3.2 Spatial interpolation of model parameters

Once the parameters of the model are known locally for each rain gauge and climate division, they can be propagated to the entire domain of interest by spatial interpolation.

To interpolate the parameters of the transform function Ψ we use ordinary Kriging with a nugget term and a Matérn covariance function (Bennett et al., 2018) applied directly to the parameters inferred at rain gauge locations. To interpolate the parameters of the covariance function C_Y we assign the locally stationary values to the barycenter of each climate division, and then interpolate them through space using a squared inverse distance interpolation because in the targeted application the small number of climate divisions does not allow for more complex interpolation techniques such as Kriging (Babak and Deutsch, 2009).

After interpolation, all parameters of the trans-Gaussian model are known at every point of the study domain $\mathcal{D}$. The parameters of the transform function Ψ can be used to transform rainfall values to their latent counterpart (and vice versa), and the parameters of the covariance function C_Y can be used to build a model of non-stationary covariance as defined in Eq. (2).

3.4 Simulating synthetic rainfall fields by stochastic simulation

Synthetic rainfall fields reproducing the statistics of the training dataset are generated by stochastic simulation performed at a set of N_T target locations $\boldsymbol{u}_i$ often located on a regular lattice grid to comply with the requirement of impact models: $\boldsymbol{u}_i \in \mathcal{G}rid \subset \mathcal{D}, i = 1, \ldots, N_\mathrm{T}$. Note that target locations are denoted by $\boldsymbol{u}_i$ to distinguish them from rain gauge locations $\boldsymbol{s}_\mathrm{g}$, and we recall here that model parameters at target locations derive from the spatial interpolation of the parameter values estimated at gauge locations.

To simulate synthetic rainfall fields, a realization of the latent field is first generated by geostatistical simulation. A simple way to draw this realization is by Cholesky decomposition of the covariance matrix $\mathbf{C}_Y$ (Davis, 1987):

$$\boldsymbol{Y}_\mathrm{us} = \mathbf{L} \times \boldsymbol{V}_\mathrm{iid} \tag{5}$$

where $\boldsymbol{Y}_\mathrm{us}$ is the vector containing the realization of the latent field Y_us at target locations $\boldsymbol{u}_i$, $\mathbf{L}$ is the lower triangular matrix resulting from the Choleski decomposition of the covariance matrix $\mathbf{C}_Y$ evaluated at the target locations (i.e., $C_Y = \mathbf{L} \times \mathbf{L}^\mathrm{T}$) and $\boldsymbol{V}_\mathrm{iid}$ is an independent and identically distributed standardized normal vector of size N_T.

Standard algorithms for the Cholesky decomposition of the covariance matrix have a complexity of $\mathcal{O}({N_T}^3)$ point operations and a memory footprint of $\mathcal{O}({N_T}^2)$ (Abdulah et al., 2019), which can be prohibitive when more than a few thousands of simulation points are intended. When more target points are requested the simulation based on the Cholesky decomposition of the covariance matrix can be replaced by an approximate but more scalable simulation method, for instance based on a spectral approach that scales linearly with the number of simulation points (Emery and Arroyo, 2018; Allard et al., 2025).

A synthetic rainfall field is finally obtained by transformation of the latent field:

$$R_\mathrm{us}(\boldsymbol{u}_i) = \Psi_{\boldsymbol{u}_i}(Y_\mathrm{us}(\boldsymbol{u}_i)) \tag{6}$$

where $Y_\mathrm{us}(\boldsymbol{u}_i)$ is the realization of latent field at target location $\boldsymbol{u}_i$ and $\Psi_{\boldsymbol{u}_i}$ is the local transform function at the same location. Note that an ensemble of N_sim realizations (i.e., $R^k_\mathrm{us}(\boldsymbol{u}_i)$, $i = 1, \ldots, N_\mathrm{T}$, $k = 1, \ldots, N_\mathrm{sim}$) can be obtained by iterating N_sim times the Eqs. (5) and (6) with different realizations of the vector $\boldsymbol{V}_\mathrm{iid}$.

3.5 Rainfall mapping by stochastic interpolation of daily rain gauge observations

Rainfall maps and associated uncertainty estimations can be obtained by the stochastic interpolation of the rain gauge observations. To this end, the ensemble of synthetic rainfall fields R^k_us described above is modified by a process called conditioning that forces them to reproduce rain gauge observations. An efficient way to perform conditioning in the framework of Geostatistics is conditioning by Kriging (Chilès and Delfiner, 2012). Since conditioning by Kriging requires a target variable with a multivariate Gaussian distribution, it should be performed on the latent field Y^k_us. Hence, conditioning a synthetic rainfall field requires the computation of the latent counterparts of rainfall observations $Y(\boldsymbol{s}) = \Psi_{\boldsymbol{s}}^{-1}(R(\boldsymbol{s}))$, which are referred to as pseudo-observations. For the observation locations $\boldsymbol{s}_w$ where some rainfall is recorded (i.e., $R(\boldsymbol{s}_w) > 0$), the pseudo-observation is obtained by simply inverting the parametric transform function ψ. In contrast, at observation locations $\boldsymbol{s}_d$ where no rainfall is recorded (i.e., $R(\boldsymbol{s}_d) = 0$), the pseudo-observation $Y(\boldsymbol{s}_d)$ is not explicitly defined because the latent field Y is censored. To circumvent this problem we propose to use a Gibbs sampler to simulate the censored latent values $Y(\boldsymbol{s}_d)$ conditional to the uncensored latent values $Y(\boldsymbol{s}_w)$ and under the constraint that $Y(\boldsymbol{s}_d) < a_0(\boldsymbol{s}_d)$ (Lantuéjoul, 2002; Benoit et al., 2018). The Gibbs sampler algorithm adapted to the simulation of the censored pseudo-observations of rainfall is detailed in Appendix A. Once the pseudo-observations $Y(\boldsymbol{s})$ are known at all observation locations, conditioning by Kriging is performed by adding to the original simulations Y^k_us the result of the simple Kriging of the residuals between the simulations of the latent field $Y^k_\mathrm{us}(\boldsymbol{s})$ and the pseudo-observations $Y(\boldsymbol{s})$ (Chilès and Delfiner, 2012):

$$Y^k_\mathrm{cs}(\boldsymbol{u}) = Y^k_\mathrm{us}(\boldsymbol{u}) + \sum_{i=1}^{N_G} \lambda_i^{SK}(\boldsymbol{u}) \left(Y(\boldsymbol{s}_i) - Y^k_\mathrm{us}(\boldsymbol{s}_i) \right) \tag{7}$$

where $\lambda_i^{SK}(\boldsymbol{u})$ is the solution of the simple Kriging system: $\sum_{j=1}^{N_G} \lambda_j^{SK}(\boldsymbol{u}) C_Y(\boldsymbol{s}_i, \boldsymbol{s}_j) = C_Y(\boldsymbol{u}, \boldsymbol{s}_i), \forall i = 1 \ldots N_G$. Finally the realizations of the interpolated rainfall field are retrieved by applying the local transform function at each target location:

$$R_{\mathrm{cs}}^{k}(\boldsymbol{u}) = \Psi_u(Y_{\mathrm{cs}}^{k}(\boldsymbol{u})) \tag{8}$$

and the uncertainty of the stochastic interpolation is assessed by evaluating the dispersion of the ensemble of N_{sim} realizations.

3.6 Conditioning daily rainfall maps to monthly totals

The uncertainty in daily rainfall maps derived from the stochastic interpolation of rain gauge observations increases with the distance to the rain gauges, and the variance of the interpolation uncertainty tends to the variance of the rainfall signal itself at grid points far from any rain gauge. This leads to a high uncertainty of rainfall interpolation in sparsely gauged areas and at the edges of the interpolation domain. To reduce the uncertainty of daily rainfall maps in data-poor areas, it is possible to generate daily rainfall maps conditioned not only to daily rain gauge observations, but also to monthly rainfall totals derived for instance from monthly rainfall maps incorporating additional observations recorded by rain gauges operating at the monthly resolution, as well as information about long-term rainfall patterns (Giambelluca et al., 2013; Frazier et al., 2016; Lucas et al., 2022).

In the present framework, conditioning daily rainfall maps to monthly totals is made complicated by the fact that daily rainfall can originate from different rain types within a given month. This implies resorting to different statistical models for the spatial modeling of the daily rainfall fields, and makes it difficult to use direct conditioning approaches. We therefore use a Markov Chain Monte Carlo (MCMC) simulation approach to deal with the complex statistical setup imposed by the coexistence of multiple rain types (Bárdossy and Pegram, 2016). Among MCMC approaches we select the Metropolis within Gibbs algorithm to condition the daily rainfall simulations to monthly totals, in which a Gibbs sampler is used for simulating spatial patterns of daily rainfall conditional to rain gauge observations and a Metropolis acceptance rule is leveraged for conditioning the local sum of daily rainfall to a monthly total.

The main elements of this algorithm are introduced hereafter and a full description is provided in Appendix B. Let m be the tag of a target month encompassing N_D days $d_{\mathrm{t}}, t = 1, \ldots, N_D$, and consider a simulation grid encompassing N_{T} target locations $\boldsymbol{u}_i, i = 1, \ldots, N_{\mathrm{T}}$. Assume that the monthly rainfall amount $R(\boldsymbol{u}_i, m)$ is known at each target location, and that daily rainfall observations $R(\boldsymbol{s}_j, d_{\mathrm{t}})$ are available at a set of N_G rain gauge locations $\boldsymbol{s}_j, j = 1, \ldots, N_G$ (distinct from the target locations $\boldsymbol{u}_i$). Then the main steps of the Metropolis within Gibbs algorithm are:

1. *Initialization.*
   a. At each rain gauge location and for each target day, compute pseudo-observations $Y(\boldsymbol{s}_j, d_{\mathrm{t}})$ by inverting Eq. (1) (rainy observations) or by Gibbs sampling following Appendix A (dry observations).
   b. For each target day, simulate separately a latent field $Y_{\mathrm{cs}}(\boldsymbol{u}_i, d_{\mathrm{t}})$ conditional to the pseudo-observations $Y(\boldsymbol{s}_j, d_{\mathrm{t}})$ using Eqs. (5) and (7), and then initialize the MCMC sampler by setting $Y_{\mathrm{iter}=0}^{\mathrm{MCMC}}(\boldsymbol{u}_i, d_{\mathrm{t}}) = Y_{\mathrm{cs}}(\boldsymbol{u}_i, d_{\mathrm{t}})$.
2. *MCMC sampling.* (for a large number of iterations it, and for all days d_{t} and all target locations $\boldsymbol{u}_i$)
   a. Simulate a candidate latent value $Y_{\mathrm{candidate}}^{\mathrm{MCMC}}(\boldsymbol{u}_i, d_{\mathrm{t}})$ by Gibbs sampling (cf. Gibbs sampling step in Appendix B) conditional to the current latent values at all other locations (i.e., $Y_{\mathrm{iter=it}}^{\mathrm{MCMC}}(\boldsymbol{u}_j, d_{\mathrm{t}}), j \neq i$) and conditional to all pseudo-observations $Y(\boldsymbol{s}_j, d_{\mathrm{t}})$ of day d_{t}.
   b. Accept or reject the candidate latent value based on the Metropolis acceptance probability (cf. Metropolis acceptance rule in Appendix B) applied to the difference between observed and simulated monthly rainfall totals, i.e.,
   $$R(\boldsymbol{u}_i, d_{\mathrm{t}}) - \left(\sum_{t'=1..N_D, t' \neq t} (\Psi_{\boldsymbol{u}_i}\ (Y_{\mathrm{iter=it}}^{\mathrm{MCMC}}(\boldsymbol{u}_i, d_{t'}))) + \Psi_{\boldsymbol{u}_i}(Y_{\mathrm{candidate}}^{\mathrm{MCMC}}(\boldsymbol{u}_i, d_{\mathrm{t}})) \right).$$
   In case of acceptance set
   $Y_{\mathrm{iter=it}}^{\mathrm{MCMC}}(\boldsymbol{u}_i, d_t) = Y_{\mathrm{candidate}}^{\mathrm{MCMC}}(\boldsymbol{u}_i, d_{\mathrm{t}})$.
3. *Retrieving daily rainfall maps.*
   a. Select a set of iterations $\mathrm{it}_{\mathrm{keep}}$ that will be used to derive the daily rainfall maps. A warm-up period of a few thousand iterations is set aside to allow for the initial convergence of the MCMC algorithm, and then $\mathrm{it}_{\mathrm{keep}}$ are separated by a thousand iterations from each other to avoid inter-samples correlation (Gilks et al., 1996).
   b. Transform the latent fields $Y_{\mathrm{it}_{\mathrm{keep}}}^{\mathrm{MCMC}}(\boldsymbol{u}_i, d_{\mathrm{t}})$ using Eq. (1) to get an ensemble of daily rainfall fields $R_{\mathrm{it}_{\mathrm{keep}}}^{\mathrm{MCMC}}(\boldsymbol{u}_i, d_{\mathrm{t}}), i = 1, \ldots, N_{\mathrm{T}}, t = 1, \ldots, N_D$ which are finally used to derive daily rainfall maps and the associated uncertainty.

The main limitation of this algorithm is its low computational efficiency, which is typical of MCMC techniques (Gilks et al., 1996). In consequence, the Metropolis within Gibbs algorithm is only applied to a restricted set of locations, which are selected in such a way that the density of conditioning locations becomes relatively homogeneous in space.

4 Model assessment and application to the Island of Hawai‘i

The Non-Stationary Trans-Gaussian rainfall model – hereafter refereed to as NSTG model – is assessed on the Island of Hawai‘i where orographic effects are very strong (cf. Sect 2). The following case study resorts to daily observations from 2000 to 2019 recorded by a network of 79 rain gauges distributed across the island. We use quality-controlled and gap-filled data from Longman et al. (2018), which are now publicly available on the Hawai‘i Climate Data Portal (https://www.hawaii.edu/climate-data-portal/, last access: 25 July 2026). The following processing steps will be illustrated hereafter:

- Sect. 4.1: Identify and delineate rain types and estimate the parameters of the NSTG model for each type.
- Sect. 4.2: Generate an ensemble of synthetic rainfall fields reproducing the statistics of observations.
- Sect. 4.3: Condition the synthetic rainfall fields to rain gauge observations in order to generate daily rainfall maps.
- Sect. 4.4: Condition daily rainfall maps to monthly totals to improve daily rainfall estimation in poorly gauged areas.

4.1 Model settings

To acknowledge the temporal variability of the rainfall patterns on the Island of Hawai‘i the 20-year dataset is split into rain types within which the statistics of rainfall are as homogeneous as possible. The main patterns of daily precipitation are delineated by unsupervised clustering of days based on rain gauge observations. We follow the approach of Benoit et al. (2022) and use the following indicators to characterize daily rainfall fields at the scale of the island: (1) the proportion of rain gauges recording a wet day, (2–3) the shape and skewness parameters of the distribution of daily rainfall intensity recorded across the island, and (4–6) the first three components of the Karhunen–Loève expansion of the (preliminarily standardized) rain gauge observations. Based on these indicators, the clustering is performed using a Gaussian Mixture Model (GMM; see e.g., (Fraley and Raftery, 2002)), which approximates the joint pdf of the indicators as a weighted sum of multivariate normal distributions. In the present case the indicators characterizing rainfall intensity (1–3) and rainfall spatial distribution (4–6) are assumed to be only slightly correlated to each other and the covariance matrices used in the GMM are therefore assumed to be diagonal. The number of clusters, hereafter referred to as rain types, is set to six as a compromise between taking into account the diversity of rainfall patterns and keeping the analysis parsimonious.

The rainfall data from the six rain types are processed separately, and one rainfall model is set up for each cluster. Hence, the observation dataset is split into rain types prior to model calibration, and one set of model parameters is estimated for each rain type separately in order to account for the temporal variability of daily rainfall patterns. Figure 3 displays for each rain type the seasonality of occurrence, the spatial patterns of mean daily rainfall observed by the rain gauge network, as well as the estimated model parameters. Rain typing results show strong contrasts in seasonality and spatial distribution of rainfall between rain types, which translates into contrasting maps of model parameters. It is interesting to notice that the stronger spatial dependencies (denoted by bigger ellipses of anisotropy in Fig. 3, right column) occur in the North-East of the island, in the wet windward areas. In these areas the spatial dependencies tend to be significantly stronger in the direction parallel to the topography than across it, which is in line with strong rainfall gradients being correlated with gradients of topography. One can also notice the prevalence of rain type 5 during summer months, which brings us to link this rain type with trade wind conditions. This view is reinforced by the observed pattern of daily rainfall, which combines all distinctive marks of trade winds induced orographic rainfall, namely moderately wet windward (East facing) slopes, very dry inland areas because of the presence of the TWI, and some thermally driven rain showers on the Kona coast on the west side of the island.

4.2 Stochastic generation of orographic rainfall

4.2.1 Multi-site rainfall simulation

We first assess the ability of the NSTG model to capture and reproduce the statistical signature of rainfall at gauge locations, which corresponds to setting-up and deploying a multi-site stochastic rainfall generator. Synthetic rainfall is generated by stochastic simulation (cf. Sect. 3.4) performed on an irregular grid whose nodes are the locations of the rain gauges. For comparison purposes the simulation is repeated the same number of times as the number of days within each rain type, and the 7305 daily simulations are ordered and concatenated to create a single 20-year time series with the same rain type timeline as the observations. Section S1 in the Supplement displays the seasonality and inter-annual variability of the 20-year time series simulated at the six locations used in Fig. 2 to investigate the diversity of rainfall climatology over the Island of Hawai‘i, and shows that the NSTG model coupled with rain types is able to reproduce rainfall seasonality and inter-annual variability across the island.

To explore the spatial patterns of daily rainfall emerging from the simulations, the statistical distribution of the simulated daily rainfall is subsequently mapped and compared to the distribution of observations. The following statistics are used to assess the distribution of rainfall at each site: the

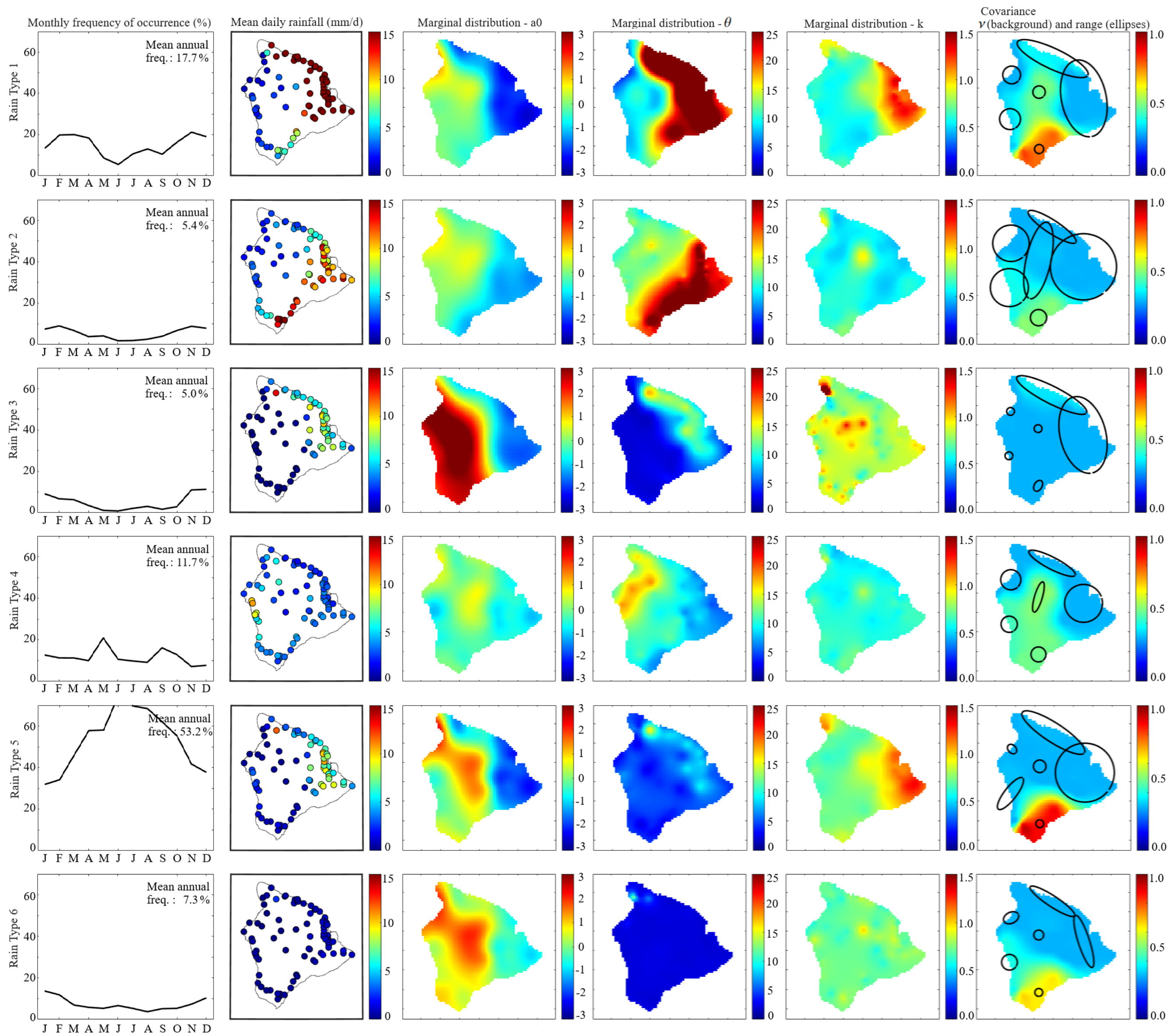

Figure 3. Daily rainfall patterns of the six rain types of the Island of Hawai'i and associated model parameters. Different rows denote different rain types. The first column shows the seasonality of rain type occurrence, column 2 displays the mean daily rainfall, columns 3–5 display maps of the estimated marginal parameters, and column 6 displays maps of the estimated covariance parameters, namely the shape parameter ν of the covariance function (background color) and ellipses of anisotropy (overprint) that denote the value of the range parameters ρ_a and ρ_b for different directions and for each climate division.

mean daily rainfall, the frequency of rainfall occurrence, the median of daily rainfall during wet days (with wet days defined at the rain gauge level, i.e., the day d is considered as wet at location $\boldsymbol{s}$ if $R(\boldsymbol{s},d) > 0$), and the quantile 95 % of daily rainfall during wet days. In addition, the whole distributions of simulations and observations are compared at each location using the 1-Wasserstein distance (Panaretos and Zemel, 2019; Vissio et al., 2020):

$$W_1(\boldsymbol{R}_{\text{obs}}, \boldsymbol{R}_{\text{sim}}) = \sum_{i=1}^{n} |\boldsymbol{R}_{\text{obs}(i)} - \boldsymbol{R}_{\text{sim}(i)}| \qquad (9)$$

where $\boldsymbol{R}_{\text{obs}(i)}$ (resp. $\boldsymbol{R}_{\text{sim}(i)}$) is the element of rank i of the vector $\boldsymbol{R}_{\text{obs}}$ (resp. $\boldsymbol{R}_{\text{sim}}$).

The performance of the NSTG model is benchmarked against the model of Benoit et al. (2022), hereafter refereed to as BSNLG2022. BSNLG2022 is a multi-site stochastic rainfall generator that has been designed specifically for tropical islands with complex topography. It draws on rain types to model the temporal evolution of rainfall statistics and on a combination of empirical copulas and Gamma distribution to model the multi-site distribution of rainfall conditional to rain types (see Appendix C for a comparison between BSNLG2022 and the trans-Gaussian model of this study). In the following the BSNLG2022 model is calibrated using the same training dataset and the same rain type time series as the ones used to calibrate the NSTG model.

Figure 4 compares the pointwise statistical signature of rainfall observations for the period 2000–2019 with the one of the 20-year synthetic rainfall simulations performed by the NSTG model (see Sect. S2 for a comparison stratified by rain types) and by the BSNLG2022 benchmark model. Results show that both models reproduce very well the four diagnostic statistics, and that the NSTG model slightly outperforms BSNLG2022 in terms of 1-Wasserstein distance with observations, in particular in places where the gradients of rainfall statistics are the strongest (i.e., the northern and eastern endpoints of the island). The only area where the proposed model is less effective than BSNLG2022 is on the East side of the island where the density of rain gauges is the highest. In this area one can notice that the NSTG model does not perfectly captures the strong variability of rain statistics between very nearby locations, and tends to slightly over-smooth rainfall statistics in space. This can be explained by the inclusion of a nugget term in the Kriging system used to interpolate the parameters of the transform function Ψ (cf. Sect. 3.3), which is necessary to filter local errors during the interpolation but leads to parameter maps that do not exactly honor point parameter estimates, and therefore results in rainfall simulations that do not perfectly reproduce rainfall statistics at observation locations. Apart for this minor over-smoothing of rain statistics in densely gauged areas, the NSTG model is able to simulate synthetic multi-site daily rainfall events that closely mimic observations.

4.2.2 Rainfall fields simulation

The NSTG model has the ability to simulate rainfall not only at gauge locations, but also across the whole domain of interest $\mathcal{D}$. To assess this ability to simulate spatially continuous rainfall fields, the NSTG model is set-up to simulate rainfall on a 2 km-resolution regular grid. In contrast, the BSNLG2022 benchmark model remains multi-site because it is by construction unable to interpolate rainfall between observation locations (due to the use of empirical copulas – cf. Appendix C – the spatial dependencies can only be modeled between locations where observations are available).

Synthetic rainfall is generated by stochastic simulation (cf. Sect. 3.4). As in Sect. 4.2.1 the simulation is repeated the same number of times as the number of days within each rain type, and the 7305 daily simulations are ordered and concatenated to create a single 20-year time series with the same rain type timeline as the observations. To explore the performance of the simulations to generate realistic spatial patterns of rainfall, we map the similarity of rainfall time series simulated at three specific locations with the time series simulated in the rest of the target area. We also look at scatter-plots of similarity statistics between all pairs of rain gauge locations to assess spatial patterns at the scale of the island. The similarity of time series of rainfall occurrence is quantified by the Jaccard Index of Rainfall Occurrence (JI-RO) (Jaccard, 1901), while the similarity of time series of rainfall intensity is quantified by the Pearson Correlation Coefficient of Rainfall Intensity (PCC-RI, evaluated solely on wet days):

$$\begin{aligned}
&\text{JI-RO}(\boldsymbol{s}_i,\boldsymbol{s}_j)=\frac{M_{11}}{M_{01}+M_{10}+M_{11}}\ \text{with :}\\
&M_{11}=\sum_{t=1}^{N_\mathrm{T}}\mathbb{1}[R(\boldsymbol{s}_i,t)>0]\cdot\mathbb{1}\big[R(\boldsymbol{s}_j,t)>0\big]\\
&M_{01}=\sum_{t=1}^{N_\mathrm{T}}\mathbb{1}[R(\boldsymbol{s}_i,t)=0]\cdot\mathbb{1}\big[R(\boldsymbol{s}_j,t)>0\big]\\
&M_{10}=\sum_{t=1}^{N_\mathrm{T}}\mathbb{1}[R(\boldsymbol{s}_i,t)>0]\cdot\mathbb{1}\big[R(\boldsymbol{s}_j,t)=0\big]
\end{aligned}\tag{10}$$

$$\begin{aligned}
&\mathrm{PCC-RI}(\boldsymbol{s}_i,\boldsymbol{s}_j)\\
&=\frac{\sum_{t_\mathrm{w}=1}^{N_\mathrm{w}}\big(R(\boldsymbol{s}_i,t_\mathrm{w})-\overline{R}(\boldsymbol{s}_i)\big)\big(R(\boldsymbol{s}_j,t_\mathrm{w})-\overline{R}(\boldsymbol{s}_j)\big)}{\sqrt{\sum_{t_\mathrm{w}=1}^{N_\mathrm{w}}\big(R(\boldsymbol{s}_i,t_\mathrm{w})-\overline{R}(\boldsymbol{s}_i)\big)^2}\sqrt{\sum_{t_\mathrm{w}=1}^{N_\mathrm{w}}\big(R(\boldsymbol{s}_j,t_\mathrm{w})-\overline{R}(\boldsymbol{s}_j)\big)^2}}\\
&\text{with } \overline{R}(\boldsymbol{s}_i)=\frac{1}{N_\mathrm{w}}\sum_{t_\mathrm{w}=1}^{N_\mathrm{w}}R(\boldsymbol{s}_i,t_\mathrm{w})\\
&\text{and } \overline{R}(\boldsymbol{s}_j)=\frac{1}{N_\mathrm{w}}\sum_{t_\mathrm{w}=1}^{N_\mathrm{w}}R(\boldsymbol{s}_j,t_\mathrm{w})
\end{aligned}\tag{11}$$

where N_T is the number of days in the dataset (here $N_\mathrm{T} = 7305$), N_w is the number of wet days, and $\mathbb{1}[.]$ is the indicator function that takes the value 1 if the event in brackets is true and 0 otherwise.

Figure 5 compares the spatial structure of rainfall observations with those of the synthetic rainfall datasets simulated by the NSTG model on one hand (see Sect. S2 for a comparison stratified by rain types), and of the BSNLG2022 benchmark model on the other hand. Results show that when focusing on the ability of the models to reproduce the spatial patterns of rainfall occurrence evaluated through JI-RO (Fig. 5a, b, e, f), both models perform equally well and almost perfectly simulate these patterns. Focusing next on the spatial patterns of rainfall intensity during rainy days evaluated through PCC-RI (Fig. 5c, d, g, h), one can notice that the BSNLG2022 benchmark model simulates them very well while the NSTG model presents a few imperfections, with some correlations of moderate intensity being underestimated (Fig. 5g). Underestimated spatial correlations mostly involve pairs of locations that are far from each others, and are apparent at the South-East of the island in Fig. 5c (top row), and also detectable at the North-East of the island in Fig. 5c (top row) and at the East of the island in Fig. 5c (bottom row). This imperfect modeling of some long range correlations is likely due to the use of relatively large climate divisions when estimating the parameters of the covariance function at the local scale, which is necessary in the present setting to ensure that

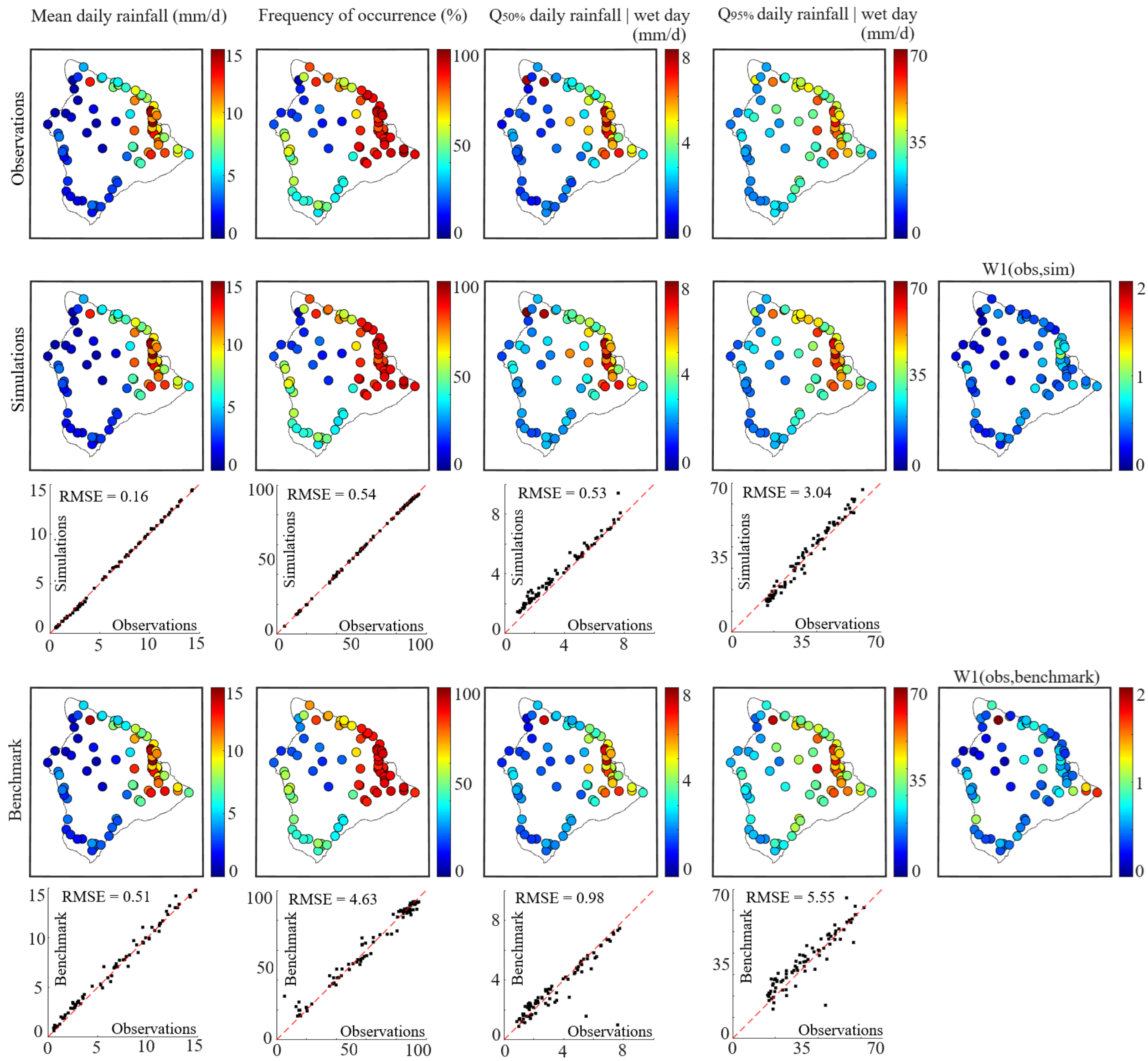

Figure 4. Statistical signature of rainfall derived from observations (first row), stochastic simulations performed by the NSTG model (second row), and stochastic simulations performed by the BSNLG2022 benchmark model (fourth row). The third (resp. fifth) row displays the scatter-plots of the observed versus simulated (resp. benchmark simulation) statistics of interest as well as the associated Root Mean Square Error (RMSE). The first column displays the mean daily rainfall, column 2 displays the frequency of rainfall occurrence, columns 3 and 4 display the quantiles 50 % and 95 % of daily rainfall during wet days, and column 5 displays the 1-Wasserstein distance between observations and simulations.

enough observations are available for a robust estimation of the covariance parameters (cf. Sect. 3.3). The interpolation of these parameters to build an island-scale non-stationary covariance function can misrepresent the actual values in areas where the covariance parameters vary strongly in space as well as at the edge of the climate divisions, leading to an imperfect modeling of the spatial dependencies in these areas. The spatial interpolation of the covariance parameters could be improved by extending the rain gauge network in poorly gauged locations, and in turn defining smaller climate divisions. The densification of the rain gauge network may also provide more detailed insights on the patterns of spatial dependencies, which could call for the use of a more refined covariance model accounting for instance for barrier effects (Bakka et al., 2019) that may be induced by the drying effect of the TWI in the center of the island. On the positive side, the short to medium range correlations are very well simulated by the trans-Gaussian model, including the patterns of non-stationarity and anisotropy. One should in particular notice the ability of the model to simulate sinuous contours of iso-correlation (i.e., not circular nor elliptic) in Fig. 5c, which is made possible by the use of a non-stationary and anisotropic covariance function to model spatial dependencies within the latent field. In addition, Fig. 5g shows that the NSTG model properly simulates medium to strong correlations (Pearson correlation > 0.5) for all pairs of gauges, which is of particular importance for rainfall mapping because they are the ones that influence the most the spatial interpolation.

In summary, Fig. 4 shows that the NSTG model is able to simulate rainfall fields with point statistics reproducing almost perfectly the marginal distribution of the observations. Figure 5 shows that the NSTG model can simulate rainfall fields with realistic patterns of rainfall occurrence and intensity, and is in particular able to reproduce the sinuous contours of iso-correlation emerging from the Hawai‘i observation dataset. The use of a parametric covariance function is the keystone for the design of a stochastic rainfall generator able to simulate spatially continuous rainfall fields, which is a requirement to generate rainfall maps from sparse rain gauge observations as detailed hereafter.

4.3 Mapping daily rainfall over complex topography

Building on the ability of the NSTG model to simulate realistic rainfall fields between observation locations, we next assess its ability to predict rainfall at ungauged locations, which corresponds to rainfall mapping. This is evaluated through a cross-validation procedure, in which we iteratively select a target rain gauge, remove it from the training dataset (as well as the data from all gauges belonging to a 5 km-radius circular area centered on this target gauge in order to mitigate the effect of near co-located gauges), re-estimate model parameters for the new training dataset, predict rainfall at the target location, and finally compare the predicted and observed rainfall values. The same procedure is applied to all gauges sequentially, and ultimately the cross-validation implemented here is equivalent to a leave-one-out approach. The quality of the rainfall prediction is evaluated with respect to the ability of the model to estimate rainfall occurrence as well as rainfall intensity during wet days. The prediction of rainfall occurrence is evaluated at each gauge location through the following metrics: the mean bias of the predicted rainfall occurrence (MB-RO), the mean absolute error of rainfall occurrence (MAE-RO), and the Brier score of rainfall occurrence (BS-RO) that measures the accuracy of the probabilistic prediction (Brier, 1950):

$$\mathrm{MB\text{-}RO}(\boldsymbol{s}) = \frac{1}{N_\mathrm{T}} \sum_{t=1}^{N_\mathrm{T}} \left(\mathbb{1}\left[\hat{R}(\boldsymbol{s},t) > 0\right] - \mathbb{1}[R_\mathrm{obs}(\boldsymbol{s},t) > 0]\right) \tag{12}$$

$$\mathrm{MAE\text{-}RO}(\boldsymbol{s}) = \frac{1}{N_\mathrm{T}} \sum_{t=1}^{N_\mathrm{T}} \left|\mathbb{1}\left[\hat{R}(\boldsymbol{s},t) > 0\right] - \mathbb{1}[R_\mathrm{obs}(\boldsymbol{s},t) > 0]\right| \tag{13}$$

$$\mathrm{BS\text{-}RO}(\boldsymbol{s}) = \frac{1}{N_\mathrm{T}} \sum_{t=1}^{N_\mathrm{T}} \left(\frac{1}{N_\mathrm{sim}} \left(\sum_{\mathrm{sim}=1}^{N_\mathrm{sim}} \mathbb{1}[R_\mathrm{sim}(\boldsymbol{s},t) > 0]\right) - \mathbb{1}[R_\mathrm{obs}(\boldsymbol{s},t) > 0]\right)^2 \tag{14}$$

where $\hat{R}(\boldsymbol{s},t)$ is the prediction of rainfall at location $\boldsymbol{s}$ and day t (here the median of 100 simulations is selected as rainfall prediction for the NSTG model), and N_sim is the number of conditional simulations (here $N_\mathrm{sim} = 100$). To complement the evaluation of rainfall occurrence prediction, we assess the prediction of rainfall intensity conditionally to the occurrence of rainfall (in the following, the occurrence of rainfall is always determined from observations). The prediction of rainfall intensity during wet days is evaluated through the following metrics, evaluated solely on wet days: the mean bias of the predicted intensity (MB-RI), the mean absolute error of rainfall intensity (MAE-RI), and the mean continuous ranked probability score of rainfall intensity (MCRPS-RI, (Gneiting and Raftery, 2007)) that measures the accuracy of the probabilistic prediction and corresponds to the integral of the Brier score associated with all possible rainfall intensity thresholds greater than zero:

$$\mathrm{MB-RI}(\boldsymbol{s}) = \frac{1}{N_\mathrm{w}} \sum_{t_\mathrm{w}=1}^{N_\mathrm{w}} \left(\hat{R}(\boldsymbol{s},t_\mathrm{w}) - R_\mathrm{obs}(\boldsymbol{s},t_\mathrm{w})\right) \tag{15}$$

$$\mathrm{MAE-RI}(\boldsymbol{s}) = \frac{1}{N_\mathrm{w}} \sum_{t_\mathrm{w}=1}^{N_\mathrm{w}} \left|\hat{R}(\boldsymbol{s},t_\mathrm{w}) - R_\mathrm{obs}(\boldsymbol{s},t_\mathrm{w})\right| \tag{16}$$

$$\mathrm{MCRPS-RI}(\boldsymbol{s}) = \frac{1}{N_\mathrm{w}} \sum_{t_\mathrm{w}=1}^{N_\mathrm{w}} \int_{r=0}^{+\infty} \left(\mathrm{F}_{R_\mathrm{sim}(\boldsymbol{s},t_\mathrm{w})}(r) - \mathbb{1}[r \geq R_\mathrm{obs}(\boldsymbol{s},t_\mathrm{w})]\right)^2 \mathrm{d}r \tag{17}$$

where N_w is the number of wet days observed at the target location (i.e., days such that $R_\mathrm{obs}(\boldsymbol{s},t_\mathrm{w}) > 0$), and $\mathrm{F}_{R_\mathrm{sim}(\boldsymbol{s},t_\mathrm{w})}$ is the cdf of the rainfall forecast derived from the 100 simulations $R_\mathrm{sim}(\boldsymbol{s},t_\mathrm{w})$.

The performance of the NSTG model is benchmarked against a Climatically Aided Interpolation (CAI) framework (Willmott and Robeson, 1995; Hunter and Meentemeyer, 2005), which has been applied to map monthly rainfall in Hawai‘i (Lucas et al., 2022) and is customized hereafter to enable the mapping of daily rainfall. CAI is used in place of BSNLG2022 to benchmark rainfall mapping because the latter model only allows for multi-site modeling and therefore lacks the interpolation skills necessary for mapping. CAI involves firstly to derive a mean Climatological Rainfall Pattern (CRP) at each gauge location for the period of interest, secondly to define RainFall Anomalies (RFA) to the CRP, thirdly to interpolate CRP and RFA at ungauged locations, and fourthly to retrieve the interpolated rainfall

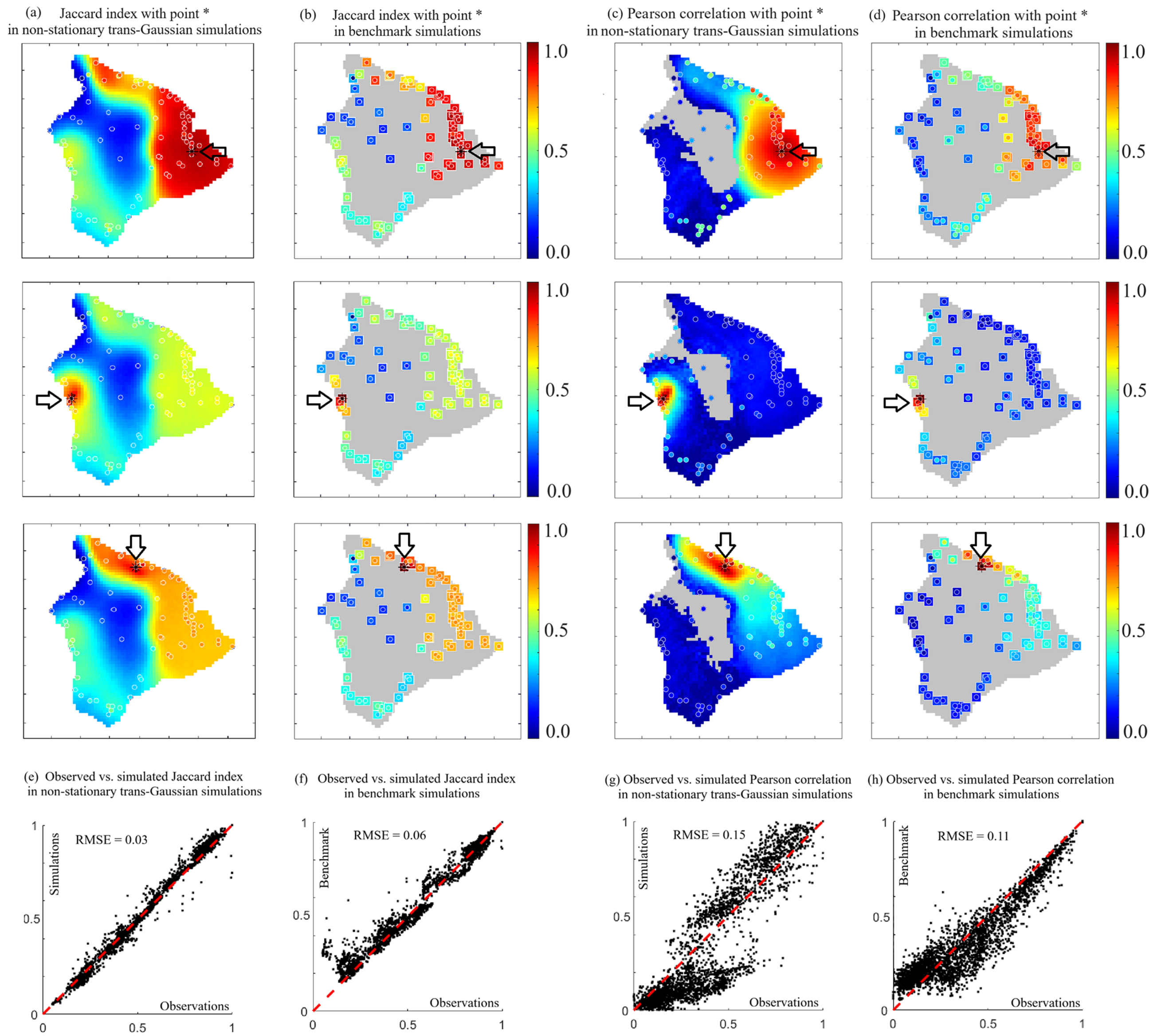

Figure 5. Spatial patterns of rainfall generated by stochastic simulations performed by the NSTG model **(a, c, e, g)** and by the BSNLG2022 benchmark model **(b, d, f, h)**. The panels **(a)** and **(b)** display the Jaccard Index of rainfall occurrence (JI-RO) with respect to three locations denoted by a star (*), and the panels **(c)** and **(d)** display the Pearson Correlation Coefficient of rain intensity evaluated solely on wet days (PCC-RI) with respect to the same locations. The panel **(e)** (resp. **f**) displays scatter-plots of observed versus simulated Jaccard index for all pairs of rain gauge locations for the NSTG model (resp. the BSNLG2022 benchmark model). The panel **(g)** (resp. **h**) displays scatter-plots of observed versus simulated Pearson correlation for all pairs of rain gauge locations and for the NSTG model (resp. the BSNLG2022 benchmark model). In **(c)** the grey color denotes areas where less than 5 % of the days are wet simultaneously with the target location, and where the estimation of PCC-RI is therefore deemed unrealiable. In **(a)**, **(b)**, **(c)**, **(d)** the foreground circles denote observation data, the continuous background color (in **a**, **c**) denotes simulations from NSTG model, and the background squares (in **b**, **d**) denote simulations from the BSNLG2022 benchmark model.

fields by back-transformation of RFA. To allow a fair comparison with the NSTG model, one CAI model is set up for each rain type separately and the exact same dataset is used for model calibration. In this setting and for a given rain gauge location $s_g \in \mathcal{N}_{\text{etwork}} \subset \mathcal{D}$ we define the CRP as the mean daily rainfall during the rain type of interest (i.e., $\text{CRP}(s_g) = \frac{1}{N_{\text{obs}}} \sum_{j=1}^{N_{\text{obs}}} \left(R_j(s_g)\right)$), and the RFA of a given day j as: $\text{RFA}_j(s_g) = \log\left(\frac{R_j(s_g)}{\text{CRP}(s_g)} + 1\right)$. The interpolation of the CRP at ungauged locations is performed by Ordinary Kriging using a stationary Matérn covariance function with Nugget, and the interpolation of RFA is performed by Ordinary Kriging using a stationary Matérn covariance function without

Nugget. With the above definitions the CAI model can be understood as a deterministic spatial interpolator (i.e., it provides a single estimation of daily rainfall at each target location, which is in contrast with stochastic interpolation that provides an ensemble prediction), which accounts for the non-stationarity of rainfall marginal distribution through the CRP, but assumes a stationary covariance for the spatial interpolation.

Figure 6a–d displays the results of the cross-validation focusing on the spatial prediction of rainfall occurrence. It shows that the NSTG model performs very well except for a few stations located in sparsely gauged areas and at the transition between a domain where rainfall is very frequent (North-East coast and mountain slopes) and another domain with very rare precipitation (summits in the center of the island and North-West coast). In that context, the withdrawal of the target rain gauge from the training and conditioning datasets leads to a lack of information about the exact location of the gradient of rainfall occurrence, and in the absence of such information the model generates a smooth transition which misrepresents the actual gradient. This artifact is absent from more densely gauged areas (e.g., the West and North-East coasts), and one can therefore infer that the density of the rain gauge network is crucial to get accurate rain occurrence maps in areas with strong rainfall gradients. When comparing the above results with the CAI, one can notice that the NSTG model outperforms the CAI benchmark model both in terms of bias (Fig. 6b) and magnitude of the prediction errors (Fig. 6c). This can be explained by the tendency of the CAI benchmark model to over-predict very light rain in place of dry conditions because it does not explicitly account for the zero-inflated nature of daily rainfall data. In contrast, the explicit modeling of this feature in the trans-Gaussian framework allows the NSTG model to map rainfall occurrence more effectively.

Figure 6e–h investigates the interpolation of rainfall intensity conditionally to the occurrence of rainfall. One can notice that in contrast with rainfall occurrence, rainfall intensity varies substantially between nearby locations, which is particularly visible in the densely gauged area in the East of the island (Fig 6e). This feature is not perfectly captured by the NSTG model nor by the CAI benchmark model, and both models smooth-out the local variability of rainfall intensity. This is because all data from the gauges located in a 5km radius around the target gauge are discarded from the cross-validation, which prevents the models from learning the local behavior of rainfall and leads to biases of opposite signs for neighbouring gauges (Fig 6f). The second artefact in rainfall intensity mapping is the imperfect depiction of rainfall intensity gradients in sparsely gauged areas, which is caused by the same reasons as discussed for rainfall occurrence. When comparing the skills of the two rainfall mapping approaches, one can notice that the CAI benchmark model marginally outperforms the NSTG model when averaging the MB-RI and MAE-RI metrics over all gauges, but that the patterns and overall magnitude of the interpolation errors are very similar for both models. This suggests that using a non-stationary covariance in the NSTG model does not improve rainfall mapping in sparsely gauged areas, even though the non-stationary covariance effectively improves the realism of the stochastic rainfall field simulations (cf. Sect. 4.2.2) used as a basis for the stochastic interpolation.

On another note, Fig. 6d and h show that a significant advantage of the NSTG model is its ability to provide probabilistic estimates of rainfall by means of an ensemble of conditional simulations, which CAI cannot do. The better scores obtained for the probabilistic forecasts (BS-RO in Fig. 6d, and MCRPS-RI in Fig. 6h) than for their deterministic counterparts (MAE-RO in Fig. 6c, and MAE-RI in Fig. 6g) show the additional information brought by the ensemble approach.

4.4 Conditioning daily rainfall maps to monthly totals

Section 4.3 showed that, like every interpolation method, the NSTG model has decreasing skills in rainfall mapping when the density of the rain gauge network decreases. This is often unavoidable in mountainous regions because the rough topography makes the setup and maintenance of rain gauge networks challenging. To improve the representation of rainfall gradients in poorly gauged areas, we propose conditioning daily rainfall maps not only to daily rain gauge observations as demonstrated in Sect. 4.3, but also to monthly totals derived from monthly rainfall maps that incorporate more information about rainfall patterns than is provided by the daily rain gauge network.

Conditioning daily maps to monthly totals is obtained by adding a Metropolis within Gibbs step in the processing framework (see Sect. 3.6), and hereafter we apply this step at 200 virtual stations spread throughout the Island of Hawai‘i (white crosses in Fig. 7, bottom left) to compensate for the large gaps existing in the rain gauge network. The impact of conditioning daily rainfall maps to monthly totals is illustrated for the months of January, April, July and October 2018, which have been selected to cover a wide range of weather conditions and sample all rain types observed in Hawai‘i.

Figure 7 compares the reference HCDP monthly rainfall maps (Fig. 7, top row) with the monthly accumulations obtained by summing the 30 or 31 daily rainfall maps derived from daily rain gauge observations only (Fig. 7, middle row), and with the sum of daily maps also conditioned to HCDP monthly totals at 200 virtual stations (Fig. 7, bottom row). The daily rainfall fields associated to Fig. 7 middle row and Fig. 7 bottom row are displayed in Sect. S3–S6. Results show that conditioning daily rainfall maps to monthly totals enables to transpose the features of the reference HCDP maps into the daily rainfall maps derived from rain gauge observations, and in their monthly sum displayed in Fig. 7 bottom row. In particular, one can notice in Fig. 7 bottom

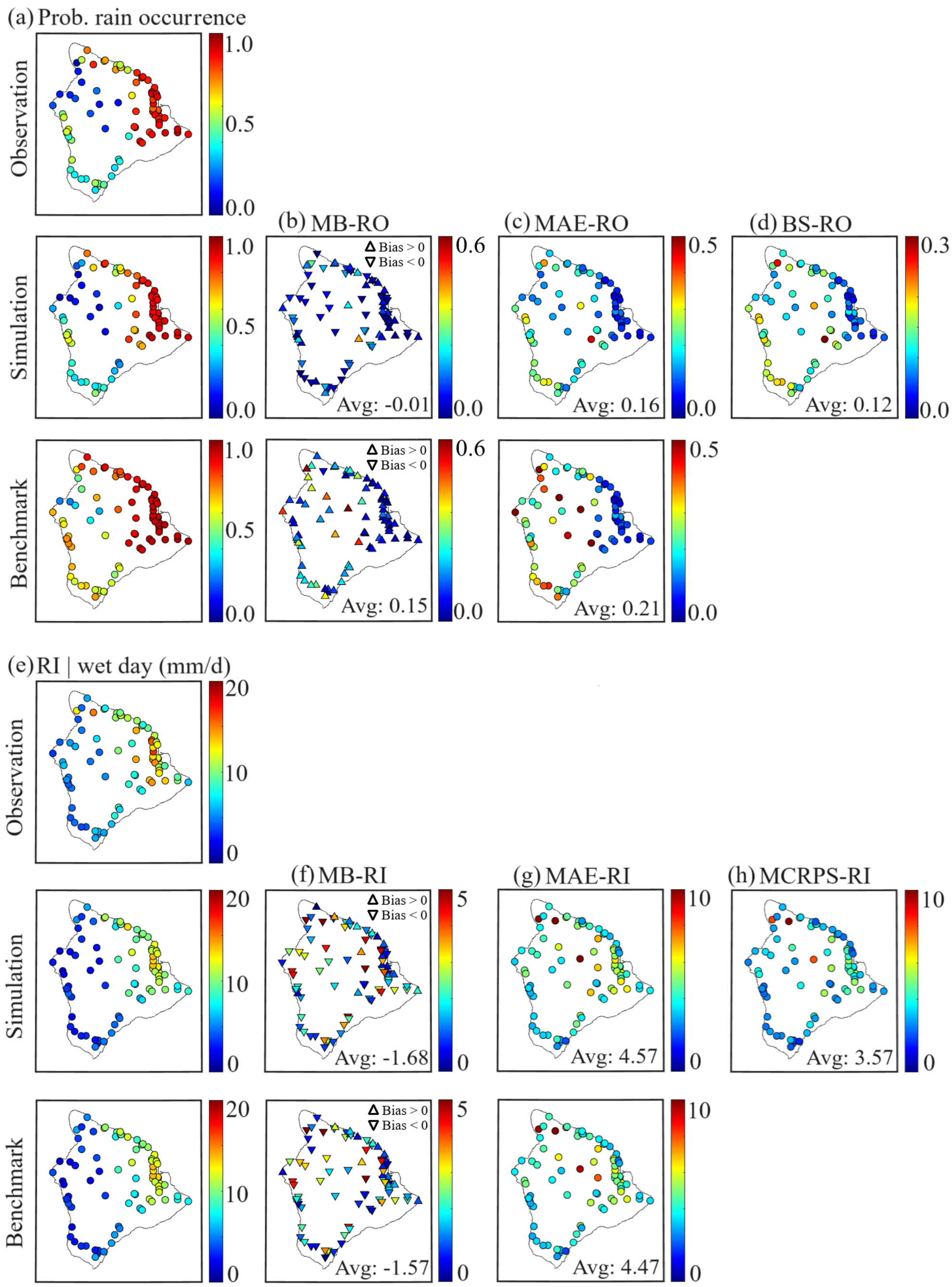

Figure 6. Cross-validation assessment of rainfall mapping performance. **(a–d)** Evaluation of rain occurrence prediction. **(e–h)** Evaluation of the prediction of daily rainfall intensity conditionally to the occurrence of rainfall.

row the emergence of a monthly spatial maximum at an ungauged location that is located on the East side of Mauna Loa for the four months of interest. In addition, the spatial gradients of rainfall occurring in the North and South of this wet area are sharper when the daily maps are conditioned to monthly totals, leading to more realistic rainfall patterns in the South-East and East feet of Mauna Loa. At higher altitudes, the wet-dry gradient occurring at the East of Mauna Kea and Mauna Loa summits is more sinuous when adding the monthly conditioning, which again improves the realism of the monthly rainfall patterns. It is worth mentioning that the improved gradients and patterns of monthly rainfall obtained by conditioning daily maps to monthly totals propagate to the daily rainfall maps displayed in Sect. S3–S6 in supplementary material, and that this improvement does not introduce any spurious smoothing nor a drizzle effect

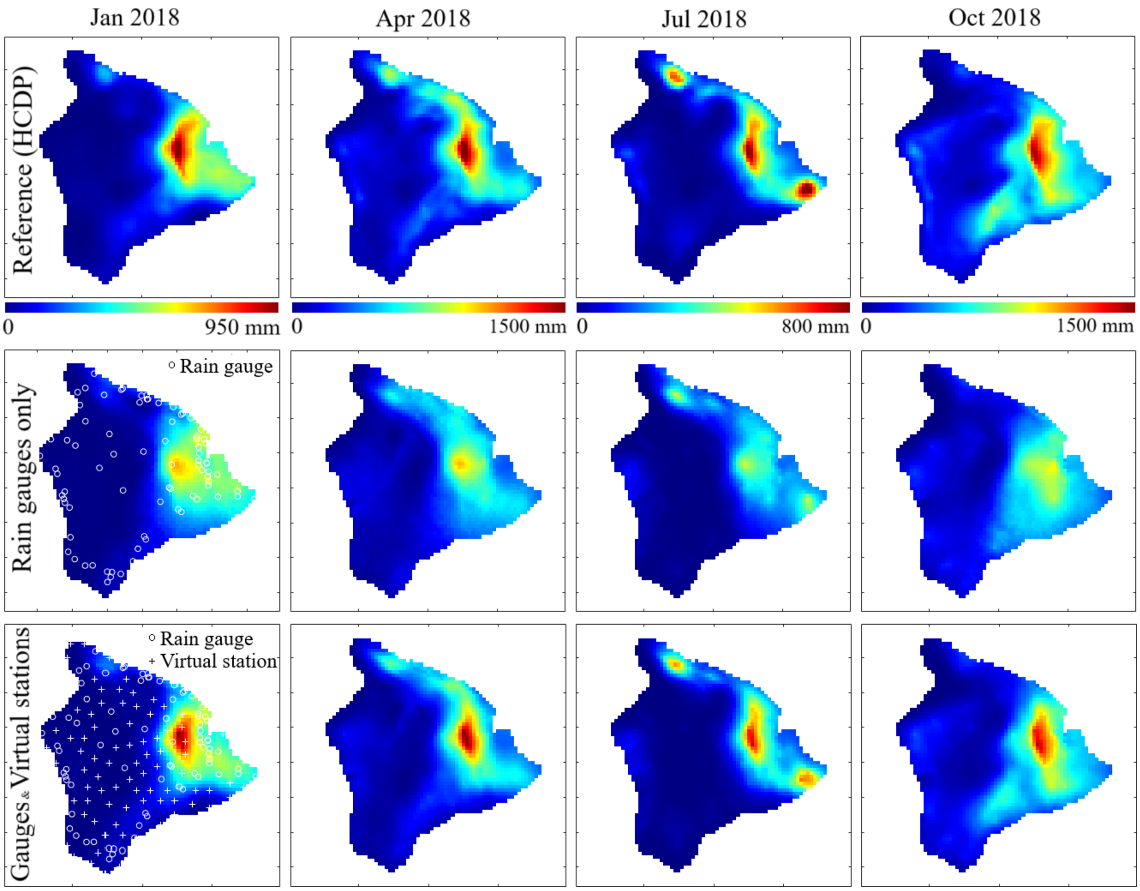

Figure 7. Maps of monthly rainfall for the months of January, April, July and October 2018 derived from the HCDP monthly historical rainfall dataset (top row), the sum of 30 or 31 daily rainfall maps obtained by the stochastic interpolation of rain gauge observations (middle row), and the sum of daily rainfall maps obtained by the stochastic interpolation of rain gauge observations constrained to honor HCDP monthly totals at 200 virtual stations (bottom row). In the bottom plot of January 2018 the dots denote the locations of the 79 rain gauges and the crosses denote the locations of the 200 virtual stations.

(i.e., the under-estimation of dry areas coupled with an over-estimation of very low rainfall intensities) in the daily rainfall maps.

To assess how the conditioning to monthly totals impacts daily rainfall maps, we perform a cross-validation study narrowed to three gauges located on the East side of Mauna Kea and on the months of January, April, July and October 2018. This area has been selected because it experiences a strong spatial gradient of rainfall and because it is only sparsely gauged, which leads to a limited performance in rainfall mapping as discussed in Sect. 4.3. The cross-validation is performed as in Sect. 4.3 and we use the same evaluation metrics (i.e., Mean Bias – MB, Mean Absolute Error - MAE, and Mean Continuous Ranked Probability Score – MCRPS) but for conciseness these metrics are evaluated hereafter on a monthly basis and on the entire time series (i.e., without separation between rainfall occurrence and rainfall intensity). We compare three methods of rainfall mapping: (1) the stochastic interpolation of daily rain gauge observations with the parameters of the NSTG model assumed to be unknown at the target location, which corresponds to the setting tested in Sect. 4.3; (2) the stochastic interpolation of daily rain gauge observations with the parameters of the NSTG model assumed to be known at the target location, which corresponds to the hypothetical case of a rain gauge that has been discontinued but whose past data are used to estimate the parameters of the model at that location; (3) the stochastic interpolation of daily rain gauge observations with the parameters of the NSTG model assumed to be unknown at the target location, and with an additional conditioning to the monthly total derived from the HCDP reference map.

Figure 8 displays the results of the cross-validation for the three settings described above. Results show that the conditioning to monthly totals (black lines in Fig. 8) drastically improves the stochastic interpolation when the parameters of the model are unknown at the target location (dark blue lines in Fig. 8). The improvement is major for the mean bias (90 % overall improvement) because the conditioning to monthly totals forces the mean monthly bias to be near zero, and the improvement is also very substantial for the mean ab-

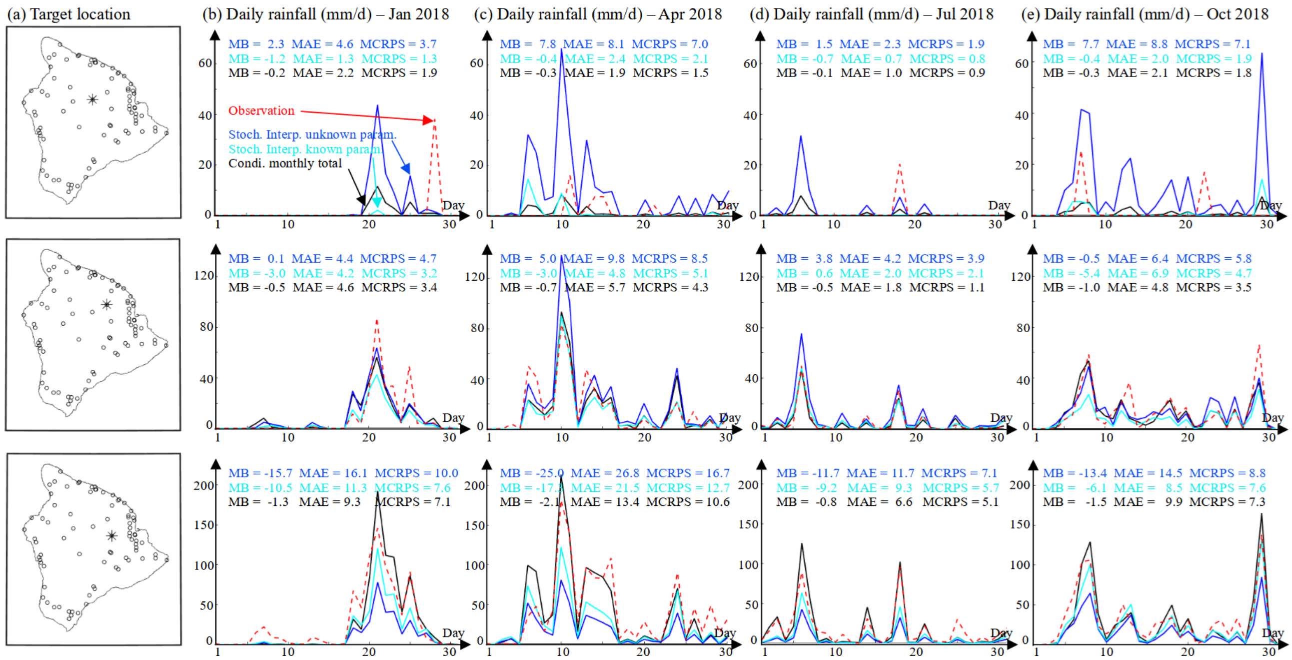

Figure 8. Cross-validation assessment of the conditioning of daily rainfall maps to monthly totals. **(a)** Location of the three gauges of interest. **(b–e)** Time series of observed (dashed red lines) and interpolated (solid lines) daily rainfall for the three target locations and the months of January, April, July and October 2018. In **(b)**–**(e)** the dark blue lines denote results of stochastic interpolation with unknown model parameters at target location, the light blue lines denote results of stochastic interpolation with known model parameters at target location, and the black lines denote results of stochastic interpolation with unknown model parameters at target location and with additional conditioning to the monthly total.

solute error (46 % overall improvement) and the mean CRPS (44 % overall improvement). This proves that conditioning to monthly totals not only improves the spatial patterns of rainfall embedded in monthly and daily maps as shown in Fig. 7 and in Sect. S3–S6, but also substantially improves point prediction – that is, the actual quality of the daily rainfall maps. It is also worth noticing that in the case of a sparsely gauged area the conditioning to monthly totals leads to better results than knowing the true value of model parameters (light blue lines in Fig. 8), with an overall improvement of 83 % for the mean bias, 15 % for the mean absolute error, and 13 % for the mean CRPS. This suggests that knowing the local monthly rainfall accumulation brings more information to the stochastic interpolation of daily observations than knowing the long term rainfall statistics at the target location, and supports the idea that combining different temporal resolutions (here daily and monthly) in rainfall mapping effectively improve the high-resolution maps while ensuring consistency between maps at different resolutions.

5 Conclusions

In this study we designed a fully non-stationary and trans-Gaussian model that is able to accommodate the steep spatial gradients of daily rainfall observed in mountainous regions. The use of a non-stationary transform function enables the simulation of realistic patterns of daily rainfall occurrence and intensity. In addition, the use of an anisotropic and non-stationary covariance function for the latent field paves the way for the simulation of rainfall fields with realistic spatial dependencies. When applied to data from a network of 79 rain gauges on the Island of Hawai‘i, the proposed model is able to infer the spatial distribution of orographic rainfall and simulate gridded rainfall products that skillfully reproduce the observed rainfall statistics. Finally, daily rainfall maps are improved in poorly gauged areas by conditioning rainfall interpolation not only to daily rain gauge observations, but also to monthly totals through a Metropolis within Gibbs approach.

When designing the geostatistical model, we chose to follow a data-driven approach using only observations to inform the non-stationarity of rainfall. Hence, the marginal distribution of rainfall is first estimated at each rain gauge location separately and subsequently interpolated at ungauged locations. Similarly, the covariance function of the latent field is first assumed to be stationary within relatively small climate divisions during the estimation of the covariance parameters, then the covariance parameters are assigned to the barycenter of each climate division, and finally these parameters are interpolated through space and incorporated in a model of non-stationary covariance. The choice of a data-driven non-stationary model has the advantage of removing the need for covariates such as elevation or wind to inform the patterns of rainfall non-stationarity. This is of particular interest over complex topography and for high-resolution rainfall modeling because in this context the link between rainfall and topo-climatic covariates tends to be weak and complex (Giambelluca et al., 2013; Benoit et al., 2024) and therefore challenging to account for in statistical rainfall models. However, data-driven non-stationary models require a lot of observations for calibration, which restricts the use of our model to densely gauged areas and calls for an increased density of rain gauge networks to improve model performance. Such high-density rain gauge network is currently being implemented in the state of Hawaii by the extension of the HCDP weather station network through the construction of Hawaiʻi Mesonet (Longman et al., 2024).

Over the past few decades, automatic rain gauges and weather stations have started to record rainfall data at sub-daily time steps and large observation datasets at a 10 min (and sometimes higher) temporal resolution have become common. Generating high-temporal-resolution rainfall maps at the scale of an island or a mountain range using such rain gauge observations could be addressed using trans-Gaussian stochastic rainfall models. The main difficulty we envision in this endeavor is the modeling of the temporal dependencies that emerge in sub-daily resolution rainfall fields in addition to the spatial dependencies addressed in the present study that focused on daily resolution. Accounting for both spatial and temporal dependencies in the presence of orographic effects will require the development of non-stationary space-time covariance functions to model the latent field (Allard et al., 2025) and robust estimation methods to infer model parameters from rain gauge observations.

In the end, rainfall maps derived from rain gauge observations are expected to complement radar rainfall estimates in mountainous regions where the complex topography challenges this technology with beam blocking and ground clutter, and to substitute the lack of radar products in non-equipped areas. Increasing the accuracy and resolution of rainfall maps is expected to improve hydrological modeling and flood forecasting in small catchments that are widespread in mountainous regions where orographic effects prevail.

Appendix A: Gibbs sampler adapted to the simulation of censored rainfall pseudo-observations

Algorithm A1 Gibbs sampler.

Inputs:
– A vector $\boldsymbol{R}$ of daily rainfall observations sorted so that its first N_d elements correspond to a dry observation (i.e., $\boldsymbol{R}(i) = 0$, $\forall i = 1, \ldots, N_\mathrm{d}$), and the other elements correspond to a non-zero observation (i.e., $\boldsymbol{R}(j) > 0$, $\forall j = N_\mathrm{d} + 1, \ldots, N_\mathrm{s}$).
– The transform function $\Psi_\mathbf{s}$ at each observation location.
– The covariance function C_Y between all pairs of observation locations.
– The number of iterations of the Gibbs sampler N_iter (typically $N_\mathrm{iter} = 2500$).
Instructions:
1. **For** $d = 1$ **to** N_d
  Initialize $\boldsymbol{Y}(d) = a_0(\boldsymbol{s}_d) - 0.1$
 end
2. **For** $w = N_\mathrm{d} + 1$ **to** N_s
  Compute $\boldsymbol{Y}(w) = \psi_{s_\mathrm{w}}^{-1}(\boldsymbol{R}(w))$
 end
3. Pre-compute the covariance matrix $\boldsymbol{\Sigma}$ between all observation locations using the covariance function C_Y
4. **For** iter $= 1$ **to** N_iter
  For $d = 1$ **to** N_d
  (a) Draw $\alpha \sim \mathcal{N}(\boldsymbol{\Sigma}^t_{-d,d}\boldsymbol{\Sigma}^{-1}_{-d,-d}\boldsymbol{Y}_{-d}, 1 - \boldsymbol{\Sigma}^t_{-d,d}\boldsymbol{\Sigma}^{-1}_{-d,-d}\boldsymbol{\Sigma}_{-d,d})$ where $\boldsymbol{\Sigma}_{-d,-d}$ is the matrix $\boldsymbol{\Sigma}$ deprived of its dth row and column, $\boldsymbol{Y}_{-d}$ is the vector $\boldsymbol{Y}$ deprived of its dth element and $\boldsymbol{\Sigma}_{-d,d}$ is the dth column of the matrix $\boldsymbol{\Sigma}$ deprived of its dth element
  (b) **While** $\alpha > a_0(\mathbf{s}_d)$
   Go to 4.a
   end
  (c) Set $\boldsymbol{Y}(d) = \alpha$
  end
end
Output: The vector of pseudo-observations $\boldsymbol{Y}$

Appendix B: Metropolis within Gibbs sampler to simulate an ensemble of meta-Gaussian random fields with prescribed sum

Algorithm B1 Metropolis within Gibbs sampler.

Inputs:
– A set of N_{f} mutually independent random fields Y_{f} characterized by their covariance: $Y_{\mathrm{f}} \sim \mathcal{N}(0, C_{Y_{\mathrm{f}}})$. A transform function Ψ_f is associated to each latent field Y_{f} to link it with the actual rainfall.
– A vector $\boldsymbol{R}$ of daily rainfall observations (length N_{s}).
– A vector $\boldsymbol{R}_{\mathrm{sum}}$ of monthly rainfall totals (length N_{v}) embodying the monthly rainfall data at the virtual stations $\mathbf{u}_j$, associated with an uncertainty vector $\mathbf{U}_{\mathrm{sum}}$ (specified as one standard deviation).
– The number of iterations N_{iter} of the Gibbs sampler (typically $N_{\mathrm{iter}} = 2000$ (warm-up) + 1000× # realizations (sampling)).
Instructions:
for $f = 1$ **to** N_{f} **do**
1. Create a vector $\boldsymbol{Y}_{\mathrm{f,obs}}$ (length N_{s}) so that: $\forall i = 1, \ldots, N_{\mathrm{s}}$: $\boldsymbol{Y}_{\mathrm{f,obs}}(i) = {\Psi_{f,\boldsymbol{s}_i}}^{-1}(R(\boldsymbol{s}_i))$. If $R(\boldsymbol{s}_i) = 0$ the corresponding latent value is undefined and has to be simulated using Algorithm 1.
2. Create a vector $\boldsymbol{Y}_{\mathrm{f,sim}}$ (length N_{v}) of latent values at the virtual station locations $\boldsymbol{u}_j$. These simulations of the latent values are obtained by combining an unconditional simulation step with a conditioning by Kriging (cf. Eq. 7 in Sect. 3.5) to rain gauge pseudo-observations (cf. Sect. 3.4).
end for
for iter $= 1$ **to** N_{iter} **do**
for $f = 1$ **to** N_{f} **do**
for $k = 1$ **to** N_{v} **do**
1. [*Gibbs sampling step*] Draw $\alpha \sim \mathcal{N}(\boldsymbol{\Sigma}_{\mathrm{f}}{}^{\mathrm{T}}_{-k,k} \boldsymbol{\Sigma}_{\mathrm{f}}{}^{-1}_{-k,-k} \boldsymbol{Y}_{\mathrm{f}_{-k}}, 1 - \boldsymbol{\Sigma}_{\mathrm{f}}{}^{T}_{-k,k} \boldsymbol{\Sigma}_{\mathrm{f}}{}^{-1}_{-k,-k} \boldsymbol{\Sigma}_{\mathrm{f}-k,k})$ where $\boldsymbol{\Sigma}_{\mathrm{f}-k,k}$, $\boldsymbol{\Sigma}_{\mathrm{f}-k,-k}$ and $\boldsymbol{Y}_{\mathrm{f}_{-k}}$ are defined in a similar way as in Algorithm 1.
2. Set $\mathbf{Z}_{k,0} = \mathbf{Z}_{k,1} = [\boldsymbol{Y}_1(k), \ldots, \boldsymbol{Y}_{N_{\mathrm{f}}}(k)]$ and assign $\mathbf{Z}_{k,1}(f) = \alpha$
3. Compute the likelihoods:

$$l_0 = \frac{1}{\sqrt{2\pi}\mathbf{U}_{\mathrm{sum}}(k)} \cdot \exp\left(-\frac{1}{2 \cdot \mathbf{U}_{\mathrm{sum}}(k)^2}\left(\left(\sum_{f=1}^{N_{\mathrm{f}}} \Psi_{\mathrm{f}}(\mathbf{Z}_{k,0}(f))\right) - \boldsymbol{R}_{\mathrm{sum}}(k)\right)^2\right)$$

$$l_1 = \frac{1}{\sqrt{2\pi}\mathbf{U}_{\mathrm{sum}}(k)} \cdot \exp\left(-\frac{1}{2.\mathbf{U}_{\mathrm{sum}}(k)^2}\left(\left(\sum_{f=1}^{N_{\mathrm{f}}} \Psi_f(\mathbf{Z}_{k,1}(f))\right) - \boldsymbol{R}_{\mathrm{sum}}(k)\right)^2\right)$$

4. Draw $\beta \sim U_{[0,1]}$
5. [*Metropolis acceptance rule*] **if** $\beta > 1 - \frac{l1}{l0}$
Set $\boldsymbol{Y}_f(k) = \alpha$
end for
end for
end for
Set $\boldsymbol{Y}_{\mathrm{f,iter}}$=$\boldsymbol{Y}_{\mathrm{f}}$
Output: A set of N_{iter} random vectors of latent field pseudo-observations at the virtual station locations.

Appendix C: Comparison between the three rainfall models used in this study

Table C1. Main features of the three models conditionally to a predefined rain type.

	BSNLG2022	CAI	NSTG
Type of interpolation	Stochastic simulation	Kriging (i.e., deterministic)	Stochastic simulation
Simulation locations	Observation sites only	Any location	Any location
Marginal distribution	Atom of zeros + single Gamma across the island	Log-normal at each location (non-stationary)	Atom of zeros + Gamma at each location (non-stationary)
Spatial dependencies	Empirical copulas (non-stationary and non-parametric)	Stationary Matérn covariance	Non-stationary Matérn covariance

Code and data availability. The source code used for this study is freely available on the following repository: https://github.com/LionelBenoit/StochasticRainfallModel_Orography (last access: 25 July 2026), https://doi.org/10.5281/zenodo.21341381 (Benoit , 2026).

The daily and monthly resolution rainfall datasets used to illustrate the study are publicly available on the Hawai‘i Climate Data Portal: https://www.hawaii.edu/climate-data-portal/ (last access: 25 July 2026).

Supplement. The supplement related to this article is available online at https://doi.org/10.5194/hess-30-4799-2026-supplement.

Author contributions. LB, MPL and TWG designed the study. LB and DA developed the stochastic rainfall model. LB implemented the model and performed the numerical experiments. MPL, KMK and TWG compiled the rainfall dataset. LB wrote the paper with contributions and editing from all co-authors.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Financial support. LB and DA acknowledge the financial support of the Chair Geolearning, funded by Andra, BNP Paribas, CCR and the SCOR Foundation for Science. MPL, KMK and TWG acknowledge the financial support of the National Science Foundation ChangeHI EPSCoR Research Infrastructure Improvement Award #OIA-2149133. KMK and TWG acknowledge the financial support of the Cooperative Institute for Research to Operations in Hydrology (CIROH) NA22NWS4320003.

Review statement. This paper was edited by Marie-Claire ten Veldhuis and reviewed by Pradeebane Vaittinada Ayar and two anonymous referees.

References

Abdulah, S., Ltaief, H., Sun, Y., Genton, M. G., and Keyes, D. E.: Geostatistical modeling and prediction using mixed precision tile Cholesky factorization, In 2019 IEEE 26th international conference on high performance computing, data, and analytics (HiPC), IEEE, 152–162, https://doi.org/10.1109/HiPC.2019.00028, 2019.

Ailliot, P., Thompson, C., and Thomson, P.: Space–time modelling of precipitation by using a hidden Markov model and censored Gaussian distributions, J. R. Stat. Soc. C-Appl., 58, 405–426, https://doi.org/10.1111/j.1467-9876.2008.00654.x, 2009.

Ailliot, P., Allard, D., Monbet, V., and Naveau, P.: Stochastic weather generators: an overview of weather type models, Journal de la Societe Francaise de Statistique, 156, 101–113, 2015.

Allard, D. and Bourotte, M.: Disaggregating daily precipitations into hourly values with a transformed censored latent Gaussian process, Stoch. Env. Res. Risk A., 29, 453–462, https://doi.org/10.1007/s00477-014-0913-4, 2015.

Allard, D., Senoussi, R., and Porcu, E.: Anisotropy models for spatial data, Math. Geosci., 48, 305–328, https://doi.org/10.1007/s11004-015-9594-x, 2016.

Allard, D., Benoit, L., and Obakrim, S.: Modeling and simulating spatio-temporal, multivariate and nonstationary Gaussian Random Fields: a Gaussian mixtures perspective, HAL open science, hal-05034982, https://hal.science/hal-05034982v1 (last access: 25 July 2026), 2025.

Babak, O. and Deutsch, C. V.: Statistical approach to inverse distance interpolation, Stoch. Env. Res. Risk A., 23, 543–553, https://doi.org/10.1007/s00477-008-0226-6, 2009.

Bakka, H., Vanhatalo, J., Illian, J. B., Simpson, D., and Rue, H.: Non-stationary Gaussian models with physical barriers, Spat. Stat., 29, 268–288, https://doi.org/10.1016/j.spasta.2019.01.002, 2019.

Bárdossy, A. and Pegram, G. G.: Space-time conditional disaggregation of precipitation at high resolution via simulation, Water Resour. Res., 52, 920–937, https://doi.org/10.1002/2015WR018037, 2016.

Bennett, B., Thyer, M., Leonard, M., Lambert, M., and Bates, B.: A comprehensive and systematic evaluation framework for a parsimonious daily rainfall field model, J. Hydrol., 556, 1123–1138, https://doi.org/10.1016/j.jhydrol.2016.12.043, 2018.

Benoit, L.: StochasticRainfallModel_Orography, Zenodo [code], https://doi.org/10.5281/zenodo.21341381, 2026.

Benoit, L. and Sichoix, L.: Sub-daily rainfall patterns in the mountainous regions of the Island of Tahiti: insights from a one-year rain gauge network expansion, J. Hydrol.: Regional Studies, 50, 101559, https://doi.org/10.1016/j.ejrh.2023.101559, 2023.

Benoit, L., Allard, D., and Mariethoz, G.: Stochastic Rainfall Modeling at Sub-kilometer Scale, Water Resour. Res., 54, 4108–4130, https://doi.org/10.1029/2018WR022817, 2018.

Benoit, L., Lucas, M., Tseng, H., Huang, Y. F., Tsang, Y. P., Nugent, A. D., Giambelluca, T. W., and Mariethoz, G.: High space-time resolution observation of extreme orographic rain gradients in a Pacific Island catchment, Front. Earth Sci., 8, 546246, https://doi.org/10.3389/feart.2020.546246, 2021.

Benoit, L., Sichoix, L., Nugent, A. D., Lucas, M. P., and Giambelluca, T. W.: Stochastic daily rainfall generation on tropical islands with complex topography, Hydrol. Earth Syst. Sci., 26, 2113–2129, https://doi.org/10.5194/hess-26-2113-2022, 2022.

Benoit, L., Koch, E., Peleg, N., and Mariethoz, G.: Precipitation-elevation relationship: non-linearity and space-time variability prevail in the Swiss Alps, J. Hydrol. X, 25, 100186, https://doi.org/10.1016/j.hydroa.2024.100186, 2024.

Bevilacqua, M., Gaetan, C., Mateu, J., and Porcu, E.: Estimating space and space-time covariance functions for large data sets: a weighted composite likelihood approach, J. Am. Stat. Assoc., 107, 268–280, https://doi.org/10.1080/01621459.2011.646928, 2012.

Blanchet, J., Paquet, E., Vaittinada Ayar, P., and Penot, D.: Mapping rainfall hazard based on rain gauge data: an objective cross-validation framework for model selection, Hydrol. Earth Syst. Sci., 23, 829–849, https://doi.org/10.5194/hess-23-829-2019, 2019.

Bowler, N. E., Pierce, C. E., and Seed, A. W.: STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP, Q. J Roy. Meteor. Soc., 132, 2127–2155, https://doi.org/10.1256/qj.04.100, 2006.

Berrocal, V. J., Raftery, A. E., and Gneiting, T.: Probabilistic quantitative precipitation field forecasting using a two-stage spatial model, Ann. Appl. Stat., 2, 1170–1193, https://doi.org/10.1214/08-AOAS203, 2008.

Brier, G. W.: Verification of forecasts expressed in terms of probability, Mon. Weather Rev., 78, 1–3, https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2, 1950.

Byrd, R. H., Hribar, M. E., and Nocedal, J.: An interior point algorithm for large-scale nonlinear programming, SIAM J. Optimiz., 9, 877–900, https://doi.org/10.1137/S1052623497325107, 1999.

Caseri, A., Javelle, P., Ramos, M. H., and Leblois, E.: Generating precipitation ensembles for flood alert and risk management, J. Flood Risk Manag., 9, 402–415, https://doi.org/10.1111/jfr3.12203, 2016.

Chandler, R. E.: Multisite, multivariate weather generation based on generalised linear models, Environ. Modell. Softw., 134, 104867, https://doi.org/10.1016/j.envsoft.2020.104867, 2020.

Chandler, R. E. and Wheater, H. S.: Analysis of rainfall variability using generalized linear models: A case study from the west of Ireland, Water Resour. Res., 38, 10-1, https://doi.org/10.1029/2001WR000906, 2002.

Chilès, J. P. and Delfiner, P.: Geostatistics: modeling spatial uncertainty, Vol. 713, John Wiley and Sons, https://doi.org/10.1002/9781118136188, 2012.

Cowpertwait, P. S.: A spatial-temporal point process model with a continuous distribution of storm types, Water Resour. Res., 46, W12507, https://doi.org/10.1029/2010WR009728, 2010.

Creutin, J. D. and Obled, C.: Objective analyses and mapping techniques for rainfall fields: an objective comparison, Water Resour. Res., 18, 413–431, https://doi.org/10.1029/WR018i002p00413, 1982.

Creutin, J. D., Delrieu, G., and Lebel, T.: Rain measurement by raingage-radar combination: a geostatistical approach, J. Atmos. Ocean. Tech., 5, 102–115, https://doi.org/10.1175/1520-0426(1988)005<0102:RMBRRC>2.0.CO;2, 1988.

Davis, M. W.: Production of conditional simulations via the LU triangular decomposition of the covariance matrix, Math. Geol., 19, 91–98, https://doi.org/10.1007/BF00898189, 1987.

Emery, X. and Arroyo, D.: On a continuous spectral algorithm for simulating non-stationary Gaussian random fields, Stoch. Env. Res. Risk A., 32, 905–919, https://doi.org/10.1007/s00477-017-1402-3, 2018.

Evin, G., Favre, A.-C., and Hingray, B.: Stochastic generation of multi-site daily precipitation focusing on extreme events, Hydrol. Earth Syst. Sci., 22, 655–672, https://doi.org/10.5194/hess-22-655-2018, 2018.

Foehn, A., Hernández, J. G., Schaefli, B., and De Cesare, G.: Spatial interpolation of precipitation from multiple rain gauge networks and weather radar data for operational applications in Alpine catchments, J. Hydrol., 563, 1092–1110, https://doi.org/10.1016/j.jhydrol.2018.05.027, 2018.

Foresti, L., Sideris, I. V., Panziera, L., Nerini, D., and Germann, U.: A 10-year radar-based analysis of orographic precipitation growth and decay patterns over the Swiss Alpine region, Q. J. Roy. Meteor. Soc., 144, 2277–2301, https://doi.org/10.1002/qj.3364, 2018.

Fouedjio, F., Desassis, N., and Romary, T.: Estimation of space deformation model for non-stationary random functions, Spat. Stat., 13, 45–61, https://doi.org/10.1016/j.spasta.2015.05.001, 2015.

Fouedjio, F., Desassis, N., and Rivoirard, J.: A generalized convolution model and estimation for non-stationary random functions, Spat. Stat., 16, 35–52, https://doi.org/10.1016/j.spasta.2016.01.002, 2016.

Fraley, C. and Raftery, A. E.: Model-Based Clustering, Discriminant Analysis, and Density Estimation, J. Am. Stat. Assoc., 97, 611–631, https://doi.org/10.1198/016214502760047131, 2002.

Frazier, A. G., Giambelluca, T. W., Diaz, H. F., and Needham, H. L.: Comparison of geostatistical approaches to spatially interpolate month-year rainfall for the Hawaiian Islands, Int. J. Climatol., 36, 1459–1470, https://doi.org/10.1002/joc.4437, 2016.

Fuglstad, G. A., Simpson, D., Lindgren, F., and Rue, H.: Does non-stationary spatial data always require non-stationary random fields?, Spat. Stat., 14, 505-531, https://doi.org/10.1016/j.spasta.2015.10.001, 2015.

Giambelluca, T. W., Chen, Q., Frazier, A. G., Price, J. P., Chen, Y.-L., Chu, P.-S., Eischeid, J. K., and Delparte, D. M.: Online Rainfall Atlas of Hawai‘i, B. Am. Meteorol. Soc., 94, 313–316, https://doi.org/10.1175/BAMS-D-11-00228.1, 2013.

Gilks, W.R., Richardson, S., and Spiegelhalter, D. J.: Markov Chain Monte Carlo in practice, Dordrecht – the Netherlands, Springer Science – Business Media, https://doi.org/10.1201/b14835, 1996.

Gneiting, T. and Raftery, A. E.: Strictly proper scoring rules, prediction, and estimation, J. Am. Stat. Assoc., 102, 359–378, https://doi.org/10.1198/016214506000001437, 2007.

Guillot, G. and Lebel, T.: Approximation of Sahelian rainfall fields with meta-Gaussian random functions; Part 2: parameter estimation and comparison to data, Stoch. Env. Res. Risk A., 13, 113–130, https://doi.org/10.1007/s004770050035, 1999.

Guttman, N. B. and Quayle, R. G.: A historical perspective of US climate divisions, B. Am. Meteorol. Soc., 77, 293–304, https://doi.org/10.1175/1520-0477(1996)077<0293:AHPOUC>2.0.CO;2, 1996.

Hansen, J. W. and Ines, A. V.: Stochastic disaggregation of monthly rainfall data for crop simulation studies, Agr. Forest Meteorol., 131, 233–246, https://doi.org/10.1016/j.agrformet.2005.06.006, 2005.

Harris, L., McRae, A. T., Chantry, M., Dueben, P. D., and Palmer, T. N.: A generative deep learning approach to stochastic downscaling of precipitation forecasts, J. Adv. Model. Earth Sy., 14, e2022MS003120, https://doi.org/10.1029/2022MS003120, 2022.

Houze Jr, R. A.: Orographic effects on precipitating clouds, Rev. Geophys., 50, https://doi.org/10.1029/2011RG000365, 2012.

Huang, Y. F. and Chen, Y. L.: Numerical simulations of seasonal variations of rainfall over the Island of Hawaii, J. Appl. Meteorol. Clim., 58, 1219–1232, https://doi.org/10.1175/JAMC-D-18-0078.1, 2019.

Huang, Y. F., Tsang, Y., Strauch, A. M., and Clilverd, H. M.: Shifting magnitude and timing of streamflow extremes and the relationship with rainfall across the Hawaiian Islands, J. Hydrol., 600, 126424, https://doi.org/10.1016/j.jhydrol.2021.126424, 2021.

Hunter, R. D. and Meentemeyer, R. K.: Climatologically aided mapping of daily precipitation and temperature, J. Appl. Meteorol., 44, 1501–1510, https://doi.org/10.1175/JAM2295.1, 2005.

Isotta, F. A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassegues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G., and Vertačnik, G.: The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, Int. J. Climatol., 34, 1657–1675, https://doi.org/10.1002/joc.3794, 2014.

Jaccard, P.: Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines, Bulletin de la Societe Vaudoise des Sciences Naturelles, 37, 241–272, 1901.

Kyriakidis, P. C., Kim, J., and Miller, N. L.: Geostatistical mapping of precipitation from rain gauge data using atmospheric and terrain characteristics, J. Appl. Meteorol. Clim., 40, 1855–1877, https://doi.org/10.1175/1520-0450(2001)040<1855:GMOPFR>2.0.CO;2, 2001.

Lantuéjoul, C.: Geostatistical Simulation: Models and Algorithms, Springer Verlag, Berlin Heidelberg, 254 pp., https://doi.org/10.1007/978-3-662-04808-5, 2002.

Lanza, L. G.: A conditional simulation model of intermittent rain fields, Hydrol. Earth Syst. Sci., 4, 173–183, https://doi.org/10.5194/hess-4-173-2000, 2000.

Leblois, E. and Creutin, J. D.: Space-time simulation of intermittent rainfall with prescribed advection field: Adaptation of the turning band method, Water Resour. Res., 49, 3375–3387, https://doi.org/10.1002/wrcr.20190, 2013.

Leinonen, J., Nerini, D., and Berne, A.: Stochastic super-resolution for downscaling time-evolving atmospheric fields with a generative adversarial network, IEEE T. Geosci. Remote, 59, 7211–7223, https://doi.org/10.1109/TGRS.2020.3032790, 2020.

Liu, Y., Wright, D. B., and Lorenz, D. J.: A nonstationary stochastic rainfall generator conditioned on global climate models for design flood analyses in the Mississippi and other large river basins, Water Resour. Res., 60, e2023WR036826, https://doi.org/10.1029/2023WR036826, 2024.

Longman, R. J., Diaz, H. F., and Giambelluca, T. W.: Sustained increases in lower-tropospheric subsidence over the central tropical North Pacific drive a decline in high-elevation rainfall in Hawaii, J. Climate, 28, 8743–8759, https://doi.org/10.1175/JCLI-D-15-0006.1, 2015.

Longman, R. J., Giambelluca, T. W., Nullet, M. A., Frazier, A. G., Kodama, K., Crausbay, S. D., Krushelnycky, P. D., Cordell, S., Clark, M. P., Newman A. J., and Arnold, J. R.: Compilation of climate data from heterogeneous networks across the Hawaiian Islands, Sci. Data, 5, 1–17, https://doi.org/10.1038/sdata.2018.12, 2018.

Longman, R. J., Timm, O. E., Giambelluca, T. W., and Kaiser, L.: A 20-year analysis of disturbance-driven rainfall on O‘ahu, Hawai‘i, Mon. Weather Rev., 149, 1767–1783, https://doi.org/10.1175/MWR-D-20-0287.1, 2021.

Longman, R. J., Lucas, M. P., Mclean, J., Cleveland, S. B., Kodama, K., Frazier, A. G., Kamelamela, K., Schriber, A., Dodge, M., Jacobs, G., and Giambelluca, T. W.: The Hawai‘i Climate Data Portal (HCDP), B. Am. Meteorol. Soc., 105, E1074–E1083, https://doi.org/10.1175/BAMS-D-23-0188.1, 2024.

Lucas, M. P., Longman, R. J., Giambelluca, T. W., Frazier, A. G., Mclean, J., Cleveland, S. B., Huang, Y-F., and Lee, J.: Optimizing automated kriging to improve spatial interpolation of monthly rainfall over complex terrain, J. Hydrometeorol., 23, 561–572, https://doi.org/10.1175/JHM-D-21-0171.1, 2022.

Luo, X., Frazier, A. G., Diaz, H. F., Longman, R., and Giambelluca, T. W.: Routine Climate Monitoring in the State of Hawai‘i: Establishment of State Climate Divisions, B. Am. Meteorol.

Soc., 105, E1045–E1061, https://doi.org/10.1175/BAMS-D-23-0236.1, 2024.

Mandapaka, P. V., Krajewski, W. F., Mantilla, R., and Gupta, V. K.: Dissecting the effect of rainfall variability on the statistical structure of peak flows, Adv. Water Resour., 32, 1508–1525, https://doi.org/10.1016/j.advwatres.2009.07.005, 2009.

Maloku, K., Evin, G., and Hingray, B.: Generation of sub-daily precipitation time series anywhere in Switzerland by mapping the parameters of GWEX-MRC, an at-site weather generator, J. Hydrol.: Regional Studies, 60, 102454, https://doi.org/10.1016/j.ejrh.2025.102454, 2025.

Molnar, P. and Burlando, P.: Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model, Atmos. Res., 77, 137–151, https://doi.org/10.1016/j.atmosres.2004.10.024, 2005.

Moraga, J. S., Peleg, N., Molnar, P., Fatichi, S., and Burlando, P.: Uncertainty in high-resolution hydrological projections: Partitioning the influence of climate models and natural climate variability, Hydrol. Process., 36, e14695, https://doi.org/10.1002/hyp.14695, 2022.

Nerini, D., Besic, N., Sideris, I., Germann, U., and Foresti, L.: A non-stationary stochastic ensemble generator for radar rainfall fields based on the short-space Fourier transform, Hydrol. Earth Syst. Sci., 21, 2777–2797, https://doi.org/10.5194/hess-21-2777-2017, 2017.

Niemi, T. J., Guillaume, J. H., Kokkonen, T., Hoang, T. M., and Seed, A. W.: Role of spatial anisotropy in design storm generation: Experiment and interpretation, Water Resour. Res., 52, 69–89, https://doi.org/10.1002/2015WR017521, 2016.

Obakrim, S., Benoit, L., and Allard, D.: A multivariate and space-time stochastic weather generator using a latent Gaussian framework, Stoch. Env. Res. Risk A., 39, 3677–3701, https://doi.org/10.1007/s00477-024-02897-8, 2025.

Onof, C., Chandler, R. E., Kakou, A., Northrop, P., Wheater, H. S., and Isham, V.: Rainfall modelling using Poisson-cluster processes: a review of developments, Stoch. Env. Res. Risk A., 14, 384–411, https://doi.org/10.1007/s004770000043, 2000.

Oriani, F., Straubhaar, J., Renard, P., and Mariethoz, G.: Simulation of rainfall time series from different climatic regions using the direct sampling technique, Hydrol. Earth Syst. Sci., 18, 3015–3031, https://doi.org/10.5194/hess-18-3015-2014, 2014.

Paciorek, C. J. and Schervish, M. J.: Spatial modelling using a new class of nonstationary covariance functions, Environmetrics, 17, 483–506, https://doi.org/10.1002/env.785, 2006.

Panaretos, V. M. and Zemel, Y.: Statistical aspects of Wasserstein distances, Annu. Rev. Stat. Appl., 6, 405–431, https://doi.org/10.1146/annurev-statistics-030718-104938, 2019.

Papalexiou, S. M. and Serinaldi, F.: Random fields simplified: Preserving marginal distributions, correlations, and intermittency, with applications from rainfall to humidity, Water Resour. Research, 56, e2019WR026331, https://doi.org/10.1029/2019WR026331, 2020.

Paschalis, A., Molnar, P., Fatichi, S., and Burlando, P.: A stochastic model for high-resolution space-time precipitation simulation, Water Resour. Res., 49, 8400–8417, https://doi.org/10.1002/2013WR014437, 2013.

Paschalis, A., Fatichi, S., Molnar, P., Rimkus, S., and Burlando, P.: On the effects of small scale space–time variability of rainfall on basin flood response, J. Hydrol., 514, 313–327, https://doi.org/10.1016/j.jhydrol.2014.04.014, 2014.

Peleg, N., Fatichi, S., Paschalis, A., Molnar, P., and Burlando, P.: An advanced stochastic weather generator for simulating 2-D high-resolution climate variables, J. Adv. Model. Earth Sy., 9, 1595–1627, https://doi.org/10.1002/2016MS000854, 2017.

Peleg, N., Marra, F., Fatichi, S., Paschalis, A., Molnar, P., and Burlando, P.: Spatial variability of extreme rainfall at radar subpixel scale, J. Hydrol., 556, 922–933, https://doi.org/10.1016/j.jhydrol.2016.05.033, 2018.

Pereira, M., Desassis, N., and Allard, D: Geostatistics for Large Datasets on Riemannian Manifolds: A Matrix-Free Approach, J. Data Sci., 20, https://doi.org/10.6339/22-JDS1075, 2022.

Porcu, E., Bevilacqua, M., Schaback, R., and Oates, C. J.: The Matérn model: A journey through statistics, numerical analysis and machine learning, Stat. Sci., 39, 469–492, https://doi.org/10.1214/24-STS923, 2024.

Qian, B., De Jong, R., Yang, J., Wang, H., and Gameda, S.: Comparing simulated crop yields with observed and synthetic weather data, Agr. Forest Meteorol., 151, 1781–1791, https://doi.org/10.1016/j.agrformet.2011.07.016, 2011.

Rajagopalan, B. and Lall, U.: A k-nearest-neighbor simulator for daily precipitation and other weather variables, Water Resour. Res., 35, 3089–3101, https://doi.org/10.1029/1999WR900028, 1999.

Ramanathan, A., Versini, P.-A., Schertzer, D., Perrin, R., Sindt, L., and Tchiguirinskaia, I.: Stochastic simulation of reference rainfall scenarios for hydrological applications using a universal multi-fractal approach, Hydrol. Earth Syst. Sci., 26, 6477–6491, https://doi.org/10.5194/hess-26-6477-2022, 2022.

Ravuri, S., Lenc, K., Willson, M., Kangin, D., Lam, R., Mirowski, P., Fitzsimons, M., Athanassiadou, M., Kashem, S., Madge, S., Prudden, R., Mandhane, A., Clark, A., Brock, A., Simonyan, K., Hadsell, R., Robinson, N., Clancy, E., Arribas, A., and Mohamed, S.: Skilful precipitation nowcasting using deep generative models of radar, Nature, 597, 672–677, https://doi.org/10.1038/s41586-021-03854-z, 2021.

Renard, B., Kavetski, D., Leblois, E., Thyer, M., Kuczera, G., and Franks, S. W.: Toward a reliable decomposition of predictive uncertainty in hydrological modeling: Characterizing rainfall errors using conditional simulation, Water Resour. Res., 47, https://doi.org/10.1029/2011WR010643, 2011.

Richardson, C. W.: Stochastic simulation of daily precipitation, temperature, and solar radiation, Water Resour. Res., 17, 182–190, https://doi.org/10.1029/WR017i001p00182, 1981.

Roe, G. H.: Orographic precipitation, Annu. Rev. Earth Pl. Sc., 33, 645–671, https://doi.org/10.1146/annurev.earth.33.092203.122541, 2005.

Sampson, P. D. and Guttorp, P.: Nonparametric estimation of nonstationary spatial covariance structure, J. Am. Stat. Assoc., 87, 108–119, https://doi.org/10.2307/2290458, 1992.

Sharma, A. and Mehrotra, R.: Rainfall generation, in: Rainfall: state of the science, no. 191 in Geophysical Monograph Series, AGU, Washington, D.C., 215–246, https://doi.org/10.1029/2010GM000973, 2010.

Srivastava, P., Yang, R., Kerrigan, G., Dresdner, G., McGibbon, J., Bretherton, C. S., and Mandt, S.: Precipitation downscaling with spatiotemporal video diffusion, Adv. Neur. In., 37, 56374–56400, https://doi.org/10.52202/079017-1795, 2024.

Strauch, A. M., Huang, Y. F., and Tsang, Y. P.: Characterizing streamflow regimes using a distributed model for sustainable resource management in the humid tropics, J. Hydrol., 636, 131287, https://doi.org/10.1016/j.jhydrol.2024.131287, 2024.

Vaittinada Ayar, P., Blanchet, J., Paquet, E., and Penot, D.: Space-time simulation of precipitation based on weather pattern subsampling and meta-Gaussian model, J. Hydrol., 581, 124451, https://doi.org/10.1016/j.jhydrol.2019.124451, 2020.

Vischel, T., Lebel, T., Massuel, S., and Cappelaere, B.: Conditional simulation schemes of rain fields and their application to rainfall–runoff modeling studies in the Sahel, J. Hydrol., 375, 273–286, https://doi.org/10.1016/j.jhydrol.2009.02.028, 2009.

Vissio, G., Lembo, V., Lucarini, V., and Ghil, M.: Evaluating the performance of climate models based on Wasserstein distance, Geophys. Res. Lett., 47, e2020GL089385, https://doi.org/10.1029/2020GL089385, 2020.

Wilcox, C., Aly, C., Vischel, T., Panthou, G., Blanchet, J., Quantin, G., and Lebel, T.: Stochastorm: A stochastic rainfall simulator for convective storms, J. Hydrometeorol., 22, 387–404, https://doi.org/10.1175/JHM-D-20-0017.1, 2021.

Wilks, D. S. and Wilby, R. L.: The weather generation game: a review of stochastic weather models, Prog. Phys. Geogr., 23, 329–357, https://doi.org/10.1177/030913339902300302, 1999.

Willmott, C. J. and Robeson, S. M.: Climatologically aided interpolation (CAI) of terrestrial air temperature, Int. J. Climatol., 15, 221–229, https://doi.org/10.1002/joc.3370150207, 1995.