RESEARCH ARTICLE

Key Points:
- A stochastic rainfall model appropriate for very high-resolution rainfall features is proposed
- An efficient Bayesian estimation is used to infer model parameters
- The proposed framework is applied to a very dense rain gauge network

Supporting Information:
- Supporting Information S1
- Figure S1
- Figure S2
- Figure S3
- Figure S4

Correspondence to:
L. Benoit,
lionel.benoit@unil.ch

Stochastic Rainfall Modeling at Sub-kilometer Scale

Lionel Benoit[1] [ID], **Denis Allard[2], and Gregoire Mariethoz[1]** [ID]

[1]Institute of Earth Surface Dynamics (IDYST), University of Lausanne, Lausanne, Switzerland, [2]Biostatistics and Spatial Processes, INRA, Avignon, France

Abstract New measurement devices allow observing rainfall with unprecedented resolution. Such observations often reveal new features of rainfall occurring at the local scale (areas of about 1–25 km^2). In particular, the joint effects of the advection of rain storms over the ground, and the deformation of spatial rain patterns along time, generate a complex rain field dependence structure characterized by strong space-time interactions. When a high-resolution is desired, stochastic rainfall models must therefore be upgraded to account for these new features of rain fields. In this paper, we propose to improve the meta-Gaussian framework, which is typically used to model space-time rain fields, to the specific case of sub-kilometer rainfall. Particular attention is paid to the reproduction of the main features of local scale rainfall, namely: (1) a skewed distribution of rain intensities with the presence of intraevent intermittency and (2) a space-time dependency structure with strong and complex space-time interactions. The resulting model, able to generate high-resolution, continuous and space-time rain fields at the local scale, is validated and applied to a real data set collected by a network of drop-counting rain gauges recording rainfall at a 1 min frequency. The combination of these data with the proposed model results in a complete framework that allows resolving the features of high-resolution rainfall (1 min temporal resolution, 100 m spatial resolution) over a small alpine catchment in Switzerland.

1. Introduction

Stochastic weather generators are statistical models that aim at simulating realistic random sequences of atmospheric variables such as precipitation, temperature, and wind (Ailliot et al., 2015; Wilks & Wilby, 1999). These models seek to reproduce, in a distributional sense, some key features of atmospheric variables such as their spatiotemporal dynamics, temporal persistence, and natural variability. Among stochastic weather generators, stochastic rainfall models are widely used because they allow reproducing the stochastic behavior of rain and its natural variability, which is challenging when using numerical weather or climate models, especially at fine scales (Bauer et al., 2015; Bony et al., 2015).

The ability of stochastic rainfall models to reproduce a large variety of rain statistics makes them a preferred tool for applications that require realistic and statistically sound rainfall reconstruction. Therefore, stochastic rainfall models are widely used for interpolation of sparse rain gauge measurements (Ailliot et al., 2009; Pardo-Igúzquiza et al., 2006; Vischel et al., 2009), disaggregation of rain rate time series or remote sensing rainfall images (Allard & Bourotte, 2015; Allcroft & Glasbey, 2003; Bárdossy & Pegram, 2016), or for downscaling precipitation fields derived from numerical model outputs (Burton et al., 2010; Vaittinada Ayar et al., 2016; Wilks, 2010). They can also be used as a component of larger modeling workflows, for example, in the generation of ensembles of synthetic rain fields with a given statistical structure, which are used as inputs for impact studies aiming at characterizing the impact of rainfall on different parts of the environment such as hydrology (Caseri et al., 2016; Mandapaka et al., 2009; Paschalis et al., 2014), rain-related hazards (von Ruette et al., 2014), or plant growth (Mavromatis & Hansen, 2001; Qian et al., 2011).

Since perceptible features of rainfall vary with the available measurement data sets, and because the features to transpose in model outputs are application dependent, a large variety of stochastic rainfall models have been developed to capture and reproduce different aspects of rain statistics. For instance, if the focus is on rainfall generation mechanisms, object based rainfall models (Onov et al., 2000; Willems, 2001) allow reproducing related characteristics (e.g., rain cells or rain bands). When the focus is rather on the space-time dependencies within rain fields, two main approaches are considered depending on whether the main

interest is the temporal or the spatial behavior of rainfall. If the temporal evolution is of main interest, Richardson-type models (Richardson, 1981; Wilks & Wilby, 1999) allow modeling rainfall time series at a finite set of locations. They first simulate rain occurrence and persistence using a Markov Chain, and then sample rain rates conditionally to wet spells. If the focus is on spatial rainfall patterns, geostatistical models based on meta-Gaussian random fields allow generating spatialized rain fields at discrete time steps (Benoit & Mariethoz, 2017). Recently, meta-Gaussian models have been extended to allow for space-time rainfall modeling (Baxevani & Lennartsson, 2015; Leblois & Creutin, 2013; Sigrist et al., 2012). Finally, when the focus is on the scaling properties of rainfall, (multi)fractal models (Deidda, 2000; Molnar & Burlando, 2005) use scaling relationships in observations to generate rain fields at different scales.

To date, the highest resolution stochastic rainfall models have focused on scales that are compatible with that of operational weather radars, corresponding to a typical resolution of 1×1 km^2 in space and 10 min to 1 h in time (e.g., Allcroft & Glasbey, 2003; Caseri et al., 2016; Paschalis et al., 2013). However, the emergence of very high-resolution rain measurements, using for instance X-band radars (Anagnostou et al., 2010; Lengfeld et al., 2014; Thorndahl et al., 2017) or high-resolution in situ instruments (Jaffrain et al., 2011), imposes new requirements when designing stochastic rainfall models. In particular, the local and short-time features of rainfall become visible and must be accounted for in the models. In addition, spatial and temporal statistics become more and more interlinked when the modeling resolution increases (Creutin et al., 2015; Lepioufle et al., 2012), making space-time models required at high resolution.

In this context, the present paper proposes a new stochastic rainfall model specially designed for local scale studies requiring high-resolution rainfall reconstructions. This model is able to handle high-resolution rainfall measurements as input and to generate dense space-time rain fields as output. One possible application, shown in this paper as an illustration of the proposed model, is the interpolation of spatially sparse but high frequency rain measurements carried out by a network of drop-counting rain gauges over a small alpine catchment in order to derive high-resolution (100 m) and high frequency (1 min) rain maps.

The remainder of this article is structured as follows: first, section 2 describes the stochastic rainfall model designed for the study of local rain fields. Then section 3 assesses the ability of the model to capture and reproduce key statistics of various rain events observed over a small Alpine catchment during summer 2017. Finally, section 4 concludes about the capabilities of the proposed model and depicts possible applications.

2. High-Resolution Stochastic Rainfall Model

2.1. Overall Modeling Approach

The targeted stochastic rainfall model must be able to handle local and high-resolution rain measurements. These measurements can be either high spatial resolution images collected at discrete times (e.g., by an X-band radar; Anagnostou et al., 2010; Lengfeld et al., 2014), or high temporal resolution rain rate time series collected in situ at discrete locations (e.g., by a network of disdrometers (Jaffrain et al., 2011) or by high-resolution rain gauges (see Appendix A)). In the following, the model is described and tested in the second case (spatially sparse time series), but it is expected to be easily extended to the case of successive images or to mixed images and time series.

When dealing with high-resolution measurements, rainfall exhibits the following key features, which must be reproduced by the stochastic model:

1. Rain events embed dry periods reflecting intraevent rain intermittency, and the marginal distribution of positive rain rates is positively skewed (Allcroft & Glasbey, 2003; Guillot, 1999).
2. The transitions between dry and wet locations are smooth, which reflects the correlation between rain intensity and the presence of rain. This phenomenon has been called "dry drift" in Schleiss et al. (2014a, 2014b) and we will adopt this terminology in this paper.
3. Space-time dependencies within rain fields are shaped by the combined effects of rain advection and diffusion (Creutin et al., 2015; Lepioufle et al., 2012).
4. The space-time statistics of rainfall differ widely depending of the local climate and the nature of the rain event of interest (Emmanuel et al., 2012; Krajewski et al., 2003).

To account for the above features, rainfall is regarded as a temporally varying space-time random process. An underlying hypothesis of this framework is that of stationarity: rainfall statistics embedded into the stochastic model are considered constant over the whole modeling domain. In the present case, it means that the modeling domain must be defined as a period during which the space-time statistics of rainfall remain as constant as possible. To ensure stationarity, an event-based modeling is adopted, i.e., a given parametrization of the model is assumed to hold throughout a single rain event, but to vary between events. The definition of stationary rain events is for now left to the expertise of users, which have to qualitatively check that the spatial patterns of rainfall, their temporal evolution as well as their advection remain broadly constant along time and space. This can be done based on the data set used for modeling, or using additional data coming for instance from C-band or S-band radar images provided by the national weather services of many countries. The development of a more quantitative and automatic framework to delineate stationary rain events is acknowledged to be of great interest for users but is let out of the scope of the present paper that focuses on the stochastic model itself in the case where the stationarity of the space-time structure of rainfall is postulated.

The space-time random process representing rainfall is modeled using a meta-Gaussian framework. This framework states that the space-time dependencies within rain fields can be modeled by a latent spatio-temporal Gaussian random field, and that the actual rain rates can be linked to the latent field through a parametric transform function. The latent field can therefore be seen as a potential of rain (Sigrist et al., 2012) which is realized through the transform function.

Our high-resolution stochastic rainfall model draws from previous studies using the meta-Gaussian framework at coarser resolutions, more specifically Allard and Bourotte (2015); Allcroft and Glasbey (2003); Baxevani and Lennartsson (2015); Guillot (1999); Leblois and Creutin (2013); Sigrist et al. (2012); Vischel et al. (2011). However, our model focuses on local and high-resolution rain field reconstructions and therefore differs from these previous works at coarser resolution in three main aspects: first, the current model adopts an event-based approach, which enables the change of space-time statistics between rain events, and therefore accounts for the variability of the structure of rain events. Second, the measurement noise is explicitly modeled in order to permit the use of high temporal resolution rainfall measurements, which tend to be noisy due to short integration times. Third, we use a versatile space-time covariance function that accounts for short lags variability and rain advection, prominent characteristics of the local scale rainfall variability.

2.2. Stochastic Model

To mimic local rainfall statistics, our model is made of four components that will be detailed in this section:

1. A transform function Ψ which models the marginal distribution of rain rates and involves three parameters a_0, a_1, and a_2.
2. A measurement noise ϵ that represents instrumental errors and involves a single parameter σ_ϵ.
3. An advection vector $\mathbf{V}$ that models the shift of rain storms above the ground and involves two parameters S_V and θ_V.
4. A covariance function ρ_L that models the space-time dependencies within the rain field and involves five parameters γ, δ, c, d, and β.

2.2.1. Marginal Model

We suppose that rainfall measurements $R(\mathbf{s}, t)$ arise from a latent, standardized, stationary Gaussian random field $Z(\mathbf{s}, t)$, with $\mathbf{s}$ being the spatial coordinates and t the time stamp. The Gaussian random field Z is latent (also sometimes referred to as hidden) because it is never observed. It is related to the observations through a transformation such that if the latent process is less than a given threshold a_0, no rainfall is measured. When it is above that threshold, the measured rainfall is modeled as a transformation of it, which will be described later.

The measurement device is not perfect (see Appendix A). Not only rainfall measurements are not perfectly measured in case of non-zero rainfall (i.e., when the latent field value is over a given threshold), but it may also occur that the device reports a non-zero value in the absence of rainfall. Conversely, the device may fail to report rainfall, and indicate a null value when there is in fact some rainfall. Notice in addition that the device never reports negative rainfall. Finally, we expect the relative error to change with the rainfall amount (see Appendix A). In order to fulfill these requirements, we model the observed values $R(\mathbf{s}_i, t_i)$, with

$i = 1, \ldots, n$ as the transformation of the sum of a latent correlated process $Y(\mathbf{s}_i, t_i)$ and a measurement error ϵ. The measurement error, ϵ, is spatially and temporally independent, i.e., $\epsilon_i \perp \epsilon_j$ when $i \neq j$. The relationship between the observed rain and the latent field is therefore:

$$R(\mathbf{s}_i, t_i) = \Psi(Z(\mathbf{s}_i, t_i)) = \Psi(Y(\mathbf{s}_i, t_i) + \epsilon_i), \quad i = 1, \ldots, n. \tag{1}$$

Several possible transform functions can be found in the literature. It can be a power transform (Ailliot et al., 2009; Katz & Parlange, 1995), a quadratic form of the power transform (Allcroft & Glasbey, 2003) or a powered exponential of a truncated Gaussian distribution (Allard & Bourotte, 2015). Here a power transformation of a truncated Gaussian distribution is used because of its good trade-off between parameter parsimony and good fit to commonly observed empirical marginal distributions of rainfall. Finally, the resulting transformation function for non-zero rain values is

$$R(\mathbf{s}_i, t_i) = \Psi(Y(\mathbf{s}_i, t_i) + \epsilon_i) = \left(\frac{Y(\mathbf{s}_i, t_i) + \epsilon_i - a_0}{a_1} \right)^{1/a_2}, \quad i = 1, \ldots, n, \tag{2}$$

where a_0 is the truncation threshold given by the normal inverse cumulative distribution of the proportion of dry measurements, a_1 is an intensity parameter, a_2 is a shape parameter, and $\epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2)$, with the additional assumption that $E[\epsilon_i \epsilon_j] = 0$ when $i \neq j$.

2.2.2. Spatiotemporal Covariance Model

Rain advection generates a time-shift between rain rate time series at distinct locations which makes the correlation function asymmetric. This can be captured by using, for the computation of the covariance, a Lagrangian reference frame moving with the rain storm (Leblois & Creutin, 2013; Lepioufle et al., 2012). It consists in projecting all data in a reference frame corrected from advection effects in which the covariance function is symmetric. Here the advection is assumed to be constant during each rain event and over the whole area of interest. Advection is therefore modeled by a single vector $\mathbf{V}$ which is described by its polar coordinates $\mathbf{V} = (S_V, \theta_V)$. Note that here we consider the advection as uniform because it makes it possible to infer it from sparse measurements. However, more complex advection fields can be considered in the same framework if it is informed by data (for example, weather radars).

The covariance of the latent field in the Lagrangian reference frame, denoted ρ_L, is linked to the advection vector $\mathbf{V}$ and the covariance ρ in the Earth fixed reference frame by

$$\rho(||\mathbf{h} - \mathbf{V}.u||, u) = \rho_L(||\mathbf{h}||, u), \tag{3}$$

where $\mathbf{h}$ is the separation vector between two observation locations $\mathbf{s}_i$ and $\mathbf{s}_j$, and u is the time period separating two observation times t_i and t_j. Gneiting (2002) proposed a type of univariate fully symmetric space-time covariances, which has now become the standard class of models for univariate space-time Gaussian random fields in geostatistical applications; see also Gneiting et al. (2007) and references therein. Here this covariance model is adopted for its flexibility to model most of the empirical covariance structures appearing in the experimental rain rate time series (see section 3), and in particular the short lags (Bourotte et al., 2016). Among this class of Gneiting's covariance functions, we choose a specific model leading to a spatial covariance of power exponential type and a temporal covariance of Cauchy type. Finally, we get

$$\rho_L(\mathbf{h}, u) = \frac{1}{(u/d)^{2\delta} + 1} \exp \left(\frac{-(||\mathbf{h}||/c)^{2\gamma}}{\left((u/d)^{2\delta} + 1 \right)^{\beta\gamma}} \right) \quad (\mathbf{h}, u) \in \mathbb{R}^2 \times \mathbb{R}, \tag{4}$$

where $c, d > 0$ are scale parameters (respectively in space and time), $0 < \gamma, \delta < 1$ are regularity parameters (respectively in space and time) and $0 \leq \beta \leq 1$ is a separability parameter (if $\beta = 0$, the covariance is space-time separable). Scale parameters, in units of distance and time respectively, control the correlation distances and in turn the average size and duration of rain patterns in the final rain fields. Regularity parameters control the differentiability at the origin of the covariance function and thereby regulate the small scale variability of the derived rain fields, and ultimately their smoothness. Values of γ and δ larger than 0.5 correspond to random fields smoother than mean-squared continuity, whereas values lower than 0.5 correspond to random fields that are not mean-squared continuous.

2.3. Model Calibration

2.3.1. Bayesian Setting

A Bayesian approach is used for parameter inference because it allows assessing the uncertainty associated to the parameters of the stochastic rainfall model (Diggle & Ribeiro, 2002). Here the vector of parameters to be inferred is $\theta = (a_1, a_2, c, d, \beta, \gamma, \delta, S_V, \theta_V, \sigma_\epsilon)$. Note that a_0 is excluded from this vector because this parameter is inverted directly from the proportion of dry measurements. Due to the complexity of the proposed stochastic model and the interactions between its parameters, the Bayesian inference is carried out by Markov Chain Monte Carlo (MCMC) sampling (Gilks et al., 1996) using a Metropolis-Hastings algorithm. More details can be found in the above reference, and the adaptation of the Metropolis-Hastings algorithm to the present context is detailed in Appendix B. The following briefly recalls the main principle.

The MCMC approach consists of iteratively proposing new values of the parameters from a prior distribution $\pi(\theta)$ according to some proposal probabilities, and accepting the modifications in such a way that the stationary distribution of the Markov Chain is precisely the distribution $\pi(\theta|\mathbf{R})$ of the model parameters knowing the calibration data set $\mathbf{R} = (R_i)_{i=1,\ldots,n}$. This stationary distribution corresponds to the distribution of model parameters after calibration. In the present case, the Metropolis-Hastings algorithm converges to $\pi(\theta|\mathbf{R})$ if a new state θ^{k+1} is accepted with the following acceptance probability p_A (Gilks et al., 1996):

$$p_A = \min\left\{1, \frac{\ell(\theta^{k+1}|\mathbf{R})\pi(\theta^{k+1})}{\ell(\theta^k|\mathbf{R})\pi(\theta^k)}\right\},\tag{5}$$

where $\ell(\theta|\mathbf{R})$ denotes the likelihood of the observations $\mathbf{R}$ given the model parameters θ. Therefore, standard implementations of the Metropolis-Hastings algorithm can be used for the calibration of the present model, as long as the likelihood $\ell(\theta|\mathbf{R})$ can be derived from the observation data set as described hereafter.

2.3.2. Likelihood

Let $R(\mathbf{s}_1, t_1), \ldots, R(\mathbf{s}_n, t_n)$ be the rainfall measurements corresponding to a rainfall event recorded by high-resolution rain measurement devices. For given parameters (a_0, a_1, a_2) of the transform function Ψ, we shall define

$$Z(\mathbf{s}_i, t_i) = \Psi^{-1}(R(\mathbf{s}_i, t_i)) = a_0 + a_1 R(\mathbf{s}_i, t_i)^{a_2}; \quad i = 1, \ldots, n\tag{6}$$

when $R(\mathbf{s}_i, t_i) > 0$. When $R(\mathbf{s}_i, t_i) = 0$, the corresponding value $Z(\mathbf{s}_i, t_i)$ is undefined.

Let us first consider that all rainfall measurements are non-zero, i.e., that all values Z_i are well defined. Since $Z_i = Y_i + \epsilon_i$, it is clear that the vector $\mathbf{Z} = (Z(\mathbf{s}_1, t_1), \ldots, Z(\mathbf{s}_n, t_n))^t$ is a multivariate Gaussian vector with mean 0 and covariance matrix Σ whose elements are

$$\Sigma[i,j] = (1 - \sigma_\epsilon^2).\rho_L(\mathbf{s}_j - \mathbf{s}_i + \mathbf{V}(t_j - t_i), t_j - t_i) \quad 1 \leq i \neq j \leq n\tag{7}$$

and $\Sigma[i,i] = 1$ for all $i = 1, \ldots, n$. On this set, the log likelihood of $\mathbf{R}$ is

$$L(\theta|\mathbf{R}) = \log(\ell(\theta|\mathbf{R})) = -0.5\left(\log|\Sigma| + \mathbf{Z}^t\Sigma^{-1}\mathbf{Z} + n\,\log(2\pi)\right) - |J_{\Psi^{-1}}(\mathbf{R})|,\tag{8}$$

where $J_{\Psi^{-1}}$ is the Jacobian matrix of the inverse transform function Ψ^{-1} and appears in the likelihood formula because of the mapping between the observed rainfall $\mathbf{R}$ and the latent field $\mathbf{Z}$.

Now consider the general case where some rainfall measurements are equal to 0 and let us denote I^0 the corresponding set of indices with N_0 elements. Let us further denote I the set of indices for which $R(\mathbf{s}_i, t_i) > 0$ and we write $\mathbf{R} = (\mathbf{R}_{I^0}, \mathbf{R}_I)$. The likelihood of the full data, $\ell(\theta|\mathbf{R})$, is the product of the likelihood of the data restricted to I with the probability that $\mathbf{R}_{I^0} = \mathbf{0}$ given that $\mathbf{R}_I$ has been observed:

$$\ell(\theta|\mathbf{R}) = \ell(\theta|\mathbf{R}_I).P(\mathbf{R}_{I^0} \leq \mathbf{0}|\mathbf{R}_I; \theta).\tag{9}$$

Hence, using equation (8), the complete log likelihood is

$$L(\theta|\mathbf{R}) = -0.5(\log|\Sigma_{I,I}| - \mathbf{Z}_I^t\Sigma_{I,I}^{-1}\mathbf{Z}_I - N_I\,\log(2\pi)) - |J_{\Psi^{-1}}(\mathbf{R}_I)|$$
$$+ \log\Phi(a_0; \Sigma_{I^0,I}\Sigma_{I,I}^{-1}\mathbf{Z}_I,\ \Sigma_{I^0,I^0} - \Sigma_{I_0,I}\Sigma_{I,I}^{-1}\Sigma_{I,I^0}),\tag{10}$$

where $\Phi(a_0; \boldsymbol{\mu}, \Sigma)$ is the N_0 dimensional cumulative probability function of a Gaussian random vector with mean $\boldsymbol{\mu}$ and covariance matrix Σ computed at a_0 for all coordinates. The third term is therefore the logarithm of the simultaneous absence of measured rainfall at all locations in I^0 given the observed values in I.

When $N_0 = 0$ it is equal to 0. When $N_I = 0$ the first two terms of equation (10) are set to 0 and the last term reduces to $\log \Phi(a_0; \mathbf{0}, \Sigma_{I^0, I^0})$.

2.4. Simulation of Dense Rain Fields

Once the parameters of the stochastic rainfall model have been inferred, the model can be used for generation of synthetic rain fields through geostatistical simulations. The uncertainty on model parameters is accounted for in the simulation process by using different samples θ^k of the parameter vector to generate different stochastic realizations. The intrinsic variability arising from the stochastic model itself is accounted for by the inherent randomness of geostatistical simulations. In practice, rain fields are generated from the stochastic rainfall model in a two-step simulation procedure:

1. The latent rain field Z is first simulated by space-time geostatistical simulation. It consists in sampling the distribution of the multivariate random variable Z defined in equation (1). The basic method to sample this distribution relies on Choleski decomposition of the covariance matrix Σ of the latent field Z at points $J = \{(\mathbf{s}_j, t_j), j = 1..N_S\}$ to simulate (Le Ravalec et al., 2000). Realizations of the latent field are obtained by multiplying the resulting triangle matrix $\mathbf{L}$ by an independent and identically distributed vector $\mathbf{V_{iid}}$ of length N_S: $\mathbf{Z}_J = \mathbf{\mu} + \mathbf{L} . \mathbf{V_{iid}}$.

   Unconditional realizations are generated by applying this sampling equation directly to the distribution of Z obtained at the parameter inference step, i.e.: $\mu = (0 \ldots 0)^t$ and $\Sigma = \Sigma_{J,J}$ (covariance matrix at simulation points).

   In case of conditioning to a set of observations $\mathbf{R}_c = R(\mathbf{s}_c, t_c)$ at points $C = \{(\mathbf{s}_c, t_c), c = 1..N_C\}$, μ and Σ must be modified to make the realizations of $\mathbf{Z}_J$ honor the measurements $\mathbf{R}_c$, which are beforehand transformed to ensure a standardized normal marginal distribution $Z(\mathbf{s}_c, t_c) = \Psi^{-1}(R(\mathbf{s}_c, t_c))$. It leads to computing the conditional distribution of Z knowing the transformed observations $\mathbf{Z}_c$. The mean vector μ and covariance matrix Σ of this conditional distribution $Z|\mathbf{Z}_c$ become

$$\mu = \Sigma_{C,J}^t \left(\Sigma_{C,C} \right)^{-1} \mathbf{Z}_C, \tag{11}$$

$$\Sigma = \Sigma_{J,J} - \Sigma_{C,J}^t \left(\Sigma_{C,C} \right)^{-1} \Sigma_{C,J}, \tag{12}$$

where $\Sigma_{J,J}$ is the covariance matrix of the latent rain field Z at points J to simulate, $\Sigma_{C,C}$ is the covariance matrix of the transformed conditioning observations $\mathbf{Z}_C$, and $\Sigma_{C,J}$ is the cross covariance matrix between the transformed observations at points C and the latent rain field at points J.

2. Once the latent field Z has been simulated, rain fields R_s are obtained at all points $(\mathbf{s}_j, t_j) \in J$ by censoring and transforming the realizations of the latent field:

$$R(\mathbf{s}_j, t_j) = \Psi(Z(\mathbf{s}_j, t_j)) = \left(\frac{Z(\mathbf{s}_j, t_j) - a_0}{a_1} \right)^{\frac{1}{a_2}} \text{if } Z(\mathbf{s}_j, t_j) > a_0$$
$$R(\mathbf{s}_j, t_j) = \Psi(Z_j(\mathbf{s}_j, t_j)) = 0 \qquad \text{if } Z(\mathbf{s}_j, t_j) \leq a_0 \tag{13}$$

2.5. Processing of Large Data Sets

The proposed model is quite complex due to the ambition to mimic as closely as possible the statistical features of local rain fields. This leads to a high computational cost for both parameter inference and rainfall simulation. To overcome this limitation, we propose hereafter approximations for parameter inference and simulation in order to make high-resolution stochastic rainfall modeling feasible on large data sets, as is often the case in real case studies.

2.5.1. Limitations for Large Data Sets

For parameter inference, computing the likelihood requires the inversion of the covariance matrix of the whole calibration data set and the computation of its determinant, which are demanding operations both in terms of CPU time and memory storage when the data set is larger than a few thousands. It also requires the computation of a multivariate Gaussian cumulative probability of dimension N_0, where N_0 is the number of data with 0 rainfall. This computation can be performed with the help of preexisting packages (e.g., mvtnorm in R or mvncdf in Matlab), but it is limited to moderate dimensions, with at most $N_0 \leq 1,000$.

Here again, CPU time increases dramatically with dimension. Similarly, the simulation of high-resolution space-time rain fields is also computationally demanding. Indeed, the use of high-resolution simulation grids leads to a large number of points to simulate. This requires the Choleski factorization of a large covariance matrix, which becomes computationally cumbersome or even infeasible when the number N_S of grid points to simulate exceeds some thousands. Since calibration data sets as well as simulation grids are by far larger than few thousands data points in the applications foreseen, approximations must be considered to make model calibration and rainfall simulation feasible. Here three approximations are proposed to boost the processing of large data sets:

1. First, for model calibration, all values that are lower than a_0 are included as latent variables to be simulated along with θ. To this end, the MCMC scheme is extended to the vector $\mathbf{Z}_{I_0}$ corresponding to space-time coordinates with null rainfall, and the vector $\mathbf{Z}_{I_0}$ is sampled by a Gibbs sampler within the Metropolis-Hastings algorithm.
2. Second, in the Metropolis-Hastings algorithm, a composite likelihood (Lindsay, 1988) is used to approximate the full likelihood. Composite likelihood methods have been proposed to perform efficient estimation with less computational cost, while still guaranteeing good asymptotic properties.
3. Third, for simulation of dense rain fields a temporal multigrid simulation approach is used. It mixes a block simulation in space with the idea of restricted neighborhood and sequential simulation in time.

These three approximations are detailed in the following sections.

2.5.2. Metropolis-Hastings Algorithm Extended to Latent Variables

In this first approximation, the difficult evaluation of the probability $\Phi(a_0; \Sigma_{I^0,I}\Sigma_{I,I}^{-1}\mathbf{Z}_I, \Sigma_{I^0,I^0} - \Sigma_{I_0,I}\Sigma_{I,I}^{-1}\Sigma_{I,I^0})$ appearing in equation (10) is skipped and replaced by the simulation of a truncated Gaussian vector using a Gibbs sampler within the iterations of the Metropolis-Hastings algorithm. The resulting Metropolis-Hastings algorithm extended to latent variables is detailed in Appendix B. It boosts the parameter inference by skipping the computation of this probability but still requires the inversion of the covariance matrix of the whole calibration data set for each iteration k, for derivation of the conditional distributions that are required by the Gibbs sampler. However, this constraint will be relaxed hereafter by the use of limited neighborhoods in the Gibbs sampler when using a blockwise likelihood.

2.5.3. Metropolis-Hastings With Blockwise Likelihood

In this second approximation, the full likelihood is replaced by a composite likelihood. See Varin et al. (2011) for an overview, and Bevilacqua et al. (2012), Bevilacqua and Gaetan (2015), and Bourotte et al. (2016) for the use of weighted pairwise likelihoods in the context of spatiotemporal and multivariate spatiotemporal data sets, in which it is shown that the loss in accuracy is mild for the estimation of the parameters. Here we divide the data set into a series of subdata sets gathering all observations carried out during given time spans, such that each subdata set has a manageable size. We call this approach blockwise likelihood; it is a special case of marginal composite likelihood. It is worth mentioning that the use of composite likelihoods in a Bayesian framework has been recognized to introduce problems in posterior estimation. In particular, several studies have shown that a pairwise likelihood approach leads to imprecise posteriors, thereby misestimating the uncertainty on model parameters (Pauli et al., 2011; Ribatet et al., 2012). However, in the present case, the block likelihood gathers a significant number of observations, which mitigates the above mentioned effect. This has been tested in a synthetic case (Appendix C) and no adverse effect has been detected.

To build the blockwise likelihood, the original data set is divided into a series of N_τ subdata sets of τ consecutive measurements at all sites, such that $\tau N_\tau = N_T$, where N_T is the total number of time steps in the data set and such that each subdata set has the largest manageable size. Then, the approximate blockwise log likelihood, $L_{BL}(\mathbf{R}; \theta)$ is

$$L_{BL}(\mathbf{R}; \theta) = \sum_{p=1}^{N_\tau} \left[-0.5 \left(|\Sigma_{I_p,I_p}| + \mathbf{Z}_{I_p}^t \Sigma_{I_p,I_p}^{-1} \mathbf{Z}_{I_p} + N_{I_p} \log(2\pi) \right) - |J_{\Psi^{-1}}(\mathbf{R}_{I_p})| \right.$$
$$\left. + \log \Phi(a_0; \Sigma_{I_p^0,I_p} \Sigma_{I_p,I_p}^{-1} \mathbf{Z}_{I_p}, \Sigma_{I_p^0,I_p^0} - \Sigma_{I_p^0,I_p} \Sigma_{I_p,I_p}^{-1} \Sigma_{I_p,I_p^0}) \right] \tag{14}$$

where I_p is the index set of all positive rainfall measurements in block p and I_p^0 is the index set of null measurements in the same block, and where all matrices are built on these set of indices. It must be noted that the set of null measurements, I_p^0, is different in each block p.

In addition to splitting the likelihood in small blocks, the evaluation of the joint cumulative probability $\Phi(a_0$; $\Sigma_{I_p^0,I_p} \Sigma_{I_p,I_p}^{-1} \mathbf{Z}_{I_p}$, $\Sigma_{I_p^0,I_p^0} - \Sigma_{I_p^0,I_p} \Sigma_{I_p,I_p}^{-1} \Sigma_{I_p,I_p^0})$ is replaced by a Gibbs sampler inserted within the Metropolis-Hastings algorithm as described in section 2.5.2. When a blockwise likelihood approach is used, the Gibbs sampler can be carried out over local neighborhoods of the same size as the blocks used for the blockwise likelihood computation, which considerably reduces the size of matrices to invert preliminary to the Gibbs sampler, thereby reducing the computational cost.

2.5.4. Simulation by Temporal Multigrids

The third approximation aims at accelerating the simulation of space-time rain fields on large simulation grids. To the best of our knowledge, none of the methods proposed previously for fast geostatistical simulation on large grids (i.e., the turning bands method (Emery et al., 2016; Emery & Lantuéjoul, 2006; Mantoglou & Wilson, 1982), FFT simulations (Le Ravalec et al., 2000), or Markov Gaussian Random Fields (Lindgren et al., 2011)) is directly transferable to the present model due to the complex space-time covariance function involved. Therefore, we propose a dedicated simulation approach to generate space-time rain fields over local areas. The basic idea is to generate rainfall sequentially over the entire spatial area of interest for one single time step by Choleski factorization. This simulation is conducted conditionally to neighboring simulated time steps, like in Sequential Gaussian Simulation (SGS) algorithm (Gómez-Hernández & Journel, 1993). A multigrid approach (Comunian et al., 2012; Tran, 1994) is used to define the simulation temporal path in order to reproduce as well as possible the temporal dependencies while minimizing the effect of using a local neighborhood along the time axis. Nussbaumer et al. (2017) has shown that in the context of SGS, the multigrid path is an optimal choice to minimize approximations due to limited neighborhoods. Finally, the following simulation algorithm is used:

1. Define a multigrid simulation path in time. This path starts to visit the extremal epochs (first and last time steps), and then visits sequentially the time steps that are the farthest from the previously visited ones, until all time steps have been visited.
2. Simulate rainfall at the time step of interest for all spatial locations (i.e., block simulation over the whole area of interest) following the previously defined simulation path. The simulation is carried out by Choleski factorization, conditional to the M closest time steps already simulated where M is a parameter of the simulation algorithm to be specified by the user.

2.6. Model Summary

The whole framework developed above for local scale stochastic rainfall modeling is summarized in Figure 1. Its core is a stochastic rainfall model specially designed to handle the features of rainfall that appear at sub-kilometer scale in space and at minute scale in time. Model calibration based on high-resolution rainfall observations allows for inference of actual rainfall statistics. After calibration, the stochastic model can be used to simulate synthetic but realistic rain fields.

3. Validation and Application

The local and high-resolution stochastic rainfall model presented above is now tested on a rain data set collected by a network of eight drop-counting rain gauges set up over a small alpine catchment (3 km $\times$ 6 km, Figure 2). Then the use of the whole framework is illustrated in the case of the estimation of high-resolution rain fields.

3.1. Observation Data Set

In this validation study, we use an observation data set collected during the summer of 2017 by a network of eight rain gauges set up at seven locations in an alpine catchment called Vallon de Nant, Vaud Alps, Switzerland (Figure 2). Two rain gauges are colocated at the southernmost location, and measurement locations are distant from 700 to 3,100 m. The measurement devices are drop-counting rain gauges which are featured by a resolution of 0.01 mm of rain depth measurement (Appendix A). In this study, the integration time has been set to 1 min, leading to a rain rate resolution of 0.6 mm/h. On site, the rain gauges have been set in locations minimizing the risks of funnel obstruction and wind effects. However, due to the relatively harsh environment, measurement errors produced by the local conditions cannot be excluded. To ensure a functional network, all sites have been visited every 2 weeks to check the devices and download the data. Two months of continuous measurements are available (July and August 2017), with six to eight gauges recording reliable data. The main causes of malfunction are funnel obstruction, and the vibration of the gauges due to an insufficient implantation which produces virtual rain records. Unreliable measurements are visually tracked, and the corresponding data are discarded in case of doubt.
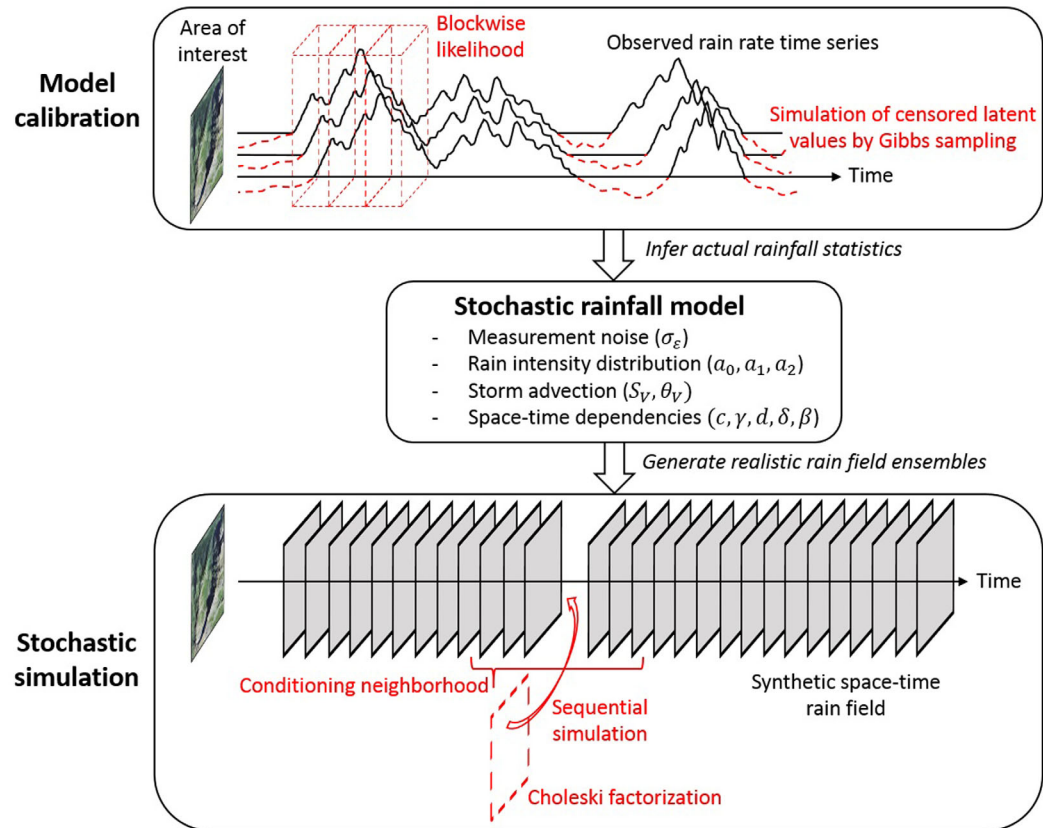
Figure 1. Schematic overview of the proposed stochastic rainfall model.

Hereafter we focus on four rain events that we believe representative of summer rains over the Swiss Alps, in view of the complete time series shown in Figure 2c. These events have been delineated from radar images (not shown here) such that the space-time properties of the corresponding rain fields are as stationary as possible along the whole events. In particular, we visually checked that the advection remains constant, that the spatial patterns of rain intensities overflying the area of interest remain similar, and that the temporal morphing of these patterns does not change much during the rain event. In addition to these qualitative checks, we also verified that the mountainous topography of the study area does not break rainfall stationarity through orographic rain enhancement. To this end, we compared the cumulative rain heights measured by the different gauges over the four events of interest. Results show a 12.7% difference between the driest and the wettest locations, which confirms the limited impact of topography on rain accumulation over this area. By visual inspection of the recorded time series (Figure 2) and additional knowledge from weather radar images, the following typology of the selected rain events can be drawn:

1. Event 1 (24 July, 7h00 to 24 July, 19h30): Moderate rain showers coming from the West. Rain amount: 24 mm. Rain fraction: 61%. Peak intensity: 21 mm/h.
2. Event 2 (8 August, 15h30 to 8 August, 16h00): Convective front with intense rainfall coming from the South-West. Rain amount: 35 mm. Rain fraction: 76%. Peak intensity: 51 mm/h.
3. Event 3 (18 August, 15h30 to 19 August, 3h30): Thunderstorm with intense rain showers coming from the West - South-West. Rain amount: 25 mm. Rain fraction: 51%. Peak intensity: 67 mm/h.
4. Event 4 (31 August, 5h00 to 31 August, 21h00): Stratiform front with low to moderate rain intensities coming from the South-West. Rain amount: 20 mm. Rain fraction: 81%. Peak intensity: 20 mm/h.

3.2. Validation Settings
The model validation consists of first calibrating the stochastic model for the four rain events of interest, and then to use the calibrated models for stochastic generation of space-time rain fields for both
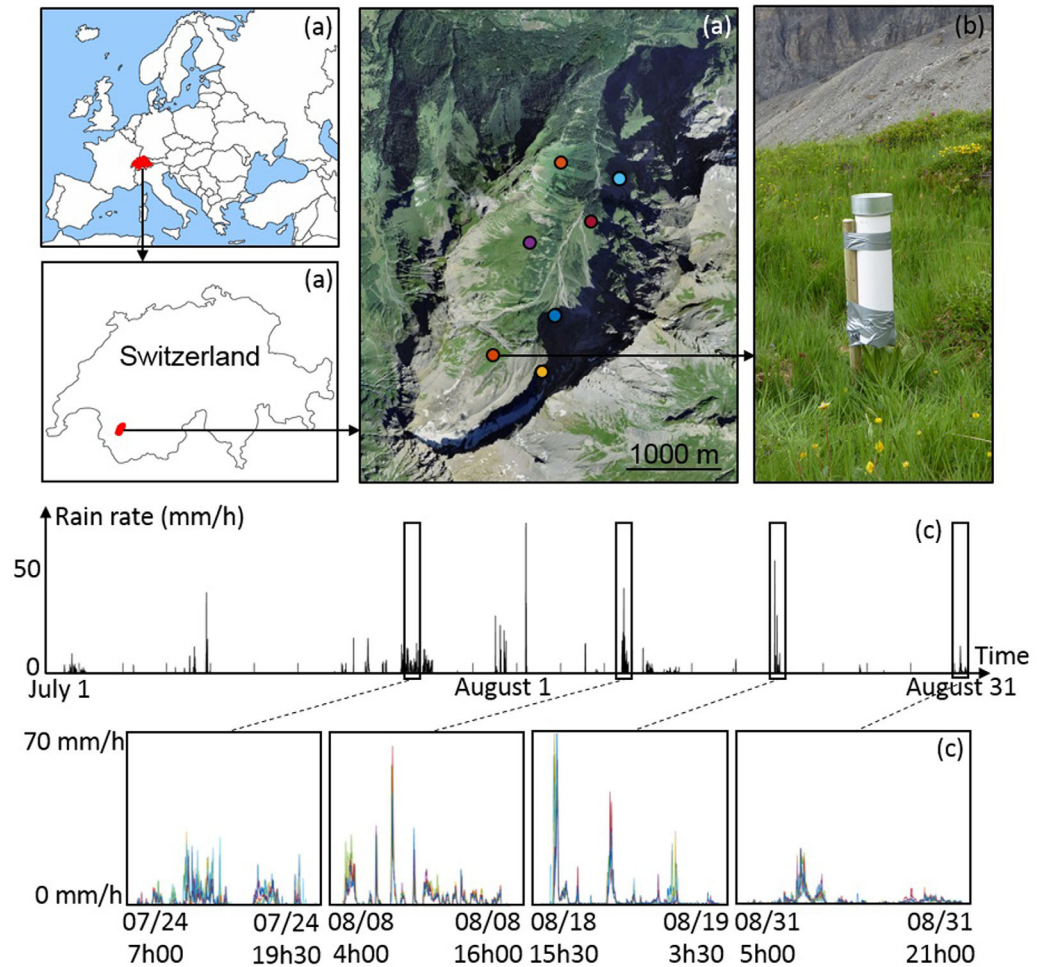
Figure 2. Observation data set used for validation. (a) Observation network, (b) measurement device, (c, top) observed rainfall averaged over the whole network, July and August 2017, and (c, bottom) rain events of interest selected for validation. Each colored line corresponds to a different rain gauge.

unconditional and conditional simulation. The relevance of the parametrization designed to capture the space-time dependencies within local rain fields is tested by comparing our model with two simpler parametrizations that roughly correspond to methodologies applied in previous studies (although applied here at higher resolutions). All three models share the same meta-Gaussian approach with the marginal model specified in section 2.2.1, but they differ in terms of the covariance model involved:

1. Version A corresponds to daily stochastic rainfall models which neglect advection effects (e.g., Baxevani & Lennartsson, 2015; Kleiber et al., 2012). Therefore, this version A considers a symmetric space-time covariance. In addition, a separable covariance function of exponential type in space and of Cauchy type in time has been selected, leading to the following covariance model:

$$\rho(\mathbf{h}, u) = \frac{1}{(u/d) + 1} \exp\left(\frac{-(\|\mathbf{h}\|)}{c}\right) \quad (\mathbf{h}, u) \in \mathbb{R}^2 \times \mathbb{R}. \tag{15}$$

2. Version B corresponds to subdaily stochastic rainfall models which account for advection effects but consider simple separable convariance functions for ρ_L in order to preserve model parsimony (e.g., Bárdossy & Pegram, 2016; Leblois & Creutin, 2013; Lepioufle et al., 2012). Here a separable covariance function of exponential type in space and of Cauchy type in time has been selected, leading to the following covariance model:

$$\rho(\mathbf{h}, u) = \frac{1}{(u/d) + 1} \exp\left(\frac{-(\|\mathbf{h} + \mathbf{V}.u\|)}{c}\right) \quad (\mathbf{h}, u) \in \mathbb{R}^2 \times \mathbb{R}. \tag{16}$$

3. Version C corresponds to the stochastic rainfall model proposed in this paper to model local rain fields at high space and time resolutions. In this case, we consider that advection has a significant impact on rain field statistics, and that a flexible covariance structure is required to account for the diversity of space-time dependencies appearing in high resolution rain measurements. Note that not only the covariance function used for ρ_L is made nonseparable to account for interactions between space and time statistics, but also regularity parameters γ and δ are introduced to account for small scale rainfall variability in space and time respectively. The corresponding covariance model of version C is therefore:

$$\rho(\mathbf{h}, u) = \frac{1}{(u/d)^{2\delta} + 1} \exp\left(\frac{-(\|\mathbf{h} + \mathbf{V}.u\|)/c)^{2\gamma}}{\left((u/d)^{2\delta} + 1\right)^{\beta\gamma}}\right) \quad (\mathbf{h}, u) \in \mathbb{R}^2 \times \mathbb{R}. \tag{17}$$

As a preliminary step, the model parameters are inferred using the procedure detailed in sections 2.3 and 2.5.3. In practice, a burn-in period of 5,000 iterations is selected for initial convergence of the Metropolis-Hastings algorithm, and then 100 vectors of model parameters separated by 100 iterations each are obtained by running the algorithm for 10,000 more iterations. Vague uniform priors are used for all parameters and these priors are scanned by random walk proposals (see Appendix B). For the computation of the blockwise likelihood, we use 40 time steps (i.e., 40 min) long blocks as a compromise between the speed of the inference procedure and the accuracy of the resulting posterior distribution of model parameters (see Appendix C).

3.3. Evaluation of Unconditional Simulations

Once the model has been calibrated for the four events of interest, unconditional simulations are carried out to assess how well the different tested versions of the model capture and reproduce the statistical structure of local rain fields without having to resort to conditioning on actual measurements. Hereafter 100 realizations are generated for each rain event and for each model version. The temporal multigrid approach presented in section 2.5.4 is used for simulation with a conditioning neighborhood of size 4.

A set of four key statistical indicators describing the local structure of rainfall has been selected to assess the quality of the unconditional simulations:

1. The empirical histogram which assesses the marginal distribution of rain rates. It reflects both the rain intermittency and the distribution of positive rain rates.
2. The average decay of rain rates at the vicinity of dry measurements. It assesses the dry drift at the edges of rainy periods.
3. The normalized temporal variogram of the rain rate time series located at the center of the area of interest (dark blue dot in Figure 2a). It quantifies the variability of rain rates measured at a single location as a function of time separation between pairs of rain measurements.
4. The spatial correlation of rain rate time series across the area of interest. Due to the effect of advection, this statistic is expected to be anisotropic. The spatial correlation of rain rate time series is therefore evaluated along two perpendicular directions (East-West and North-South), thereby leading to two substatistics. These statistics allow to quantify the spatial covariance of rainfall in an Earth fixed reference frame, as well as its anisotropy caused by advection.

The results of the unconditional simulation validation experiments are displayed in Figure 3. Focusing first on the marginal distribution of rainfall, the histograms show that the proposed marginal model performs well for the four test rain events, which involve rain rates ranging from 0 mm/h to around 70 mm/h. Note that all versions of the model perform equally well to reproduce rain rate histograms because the same meta-Gaussian framework with the same transform function is used in all cases. In contrast, one can notice that version C reproduces more accurately the dry drift statistics, probably due to a better reproduction of the space-time dependencies within rain fields as discussed hereafter.

Focusing now on the space-time dependencies within rain fields, rows 3–5 of Figure 3 show that the more complex the model, the better the reproduction of the dependencies. Comparing first version A (black curves) against versions B and C (blue and green curves), it appears that considering advection improves the reproduction of the spatial correlation, and in particular allows for anisotropic spatial correlations into simulations. Next, comparing versions B and C, it appears that adding flexibility into the covariance function further improves the joint reproduction of space and time statistics. In particular, analyzing together the temporal variograms and
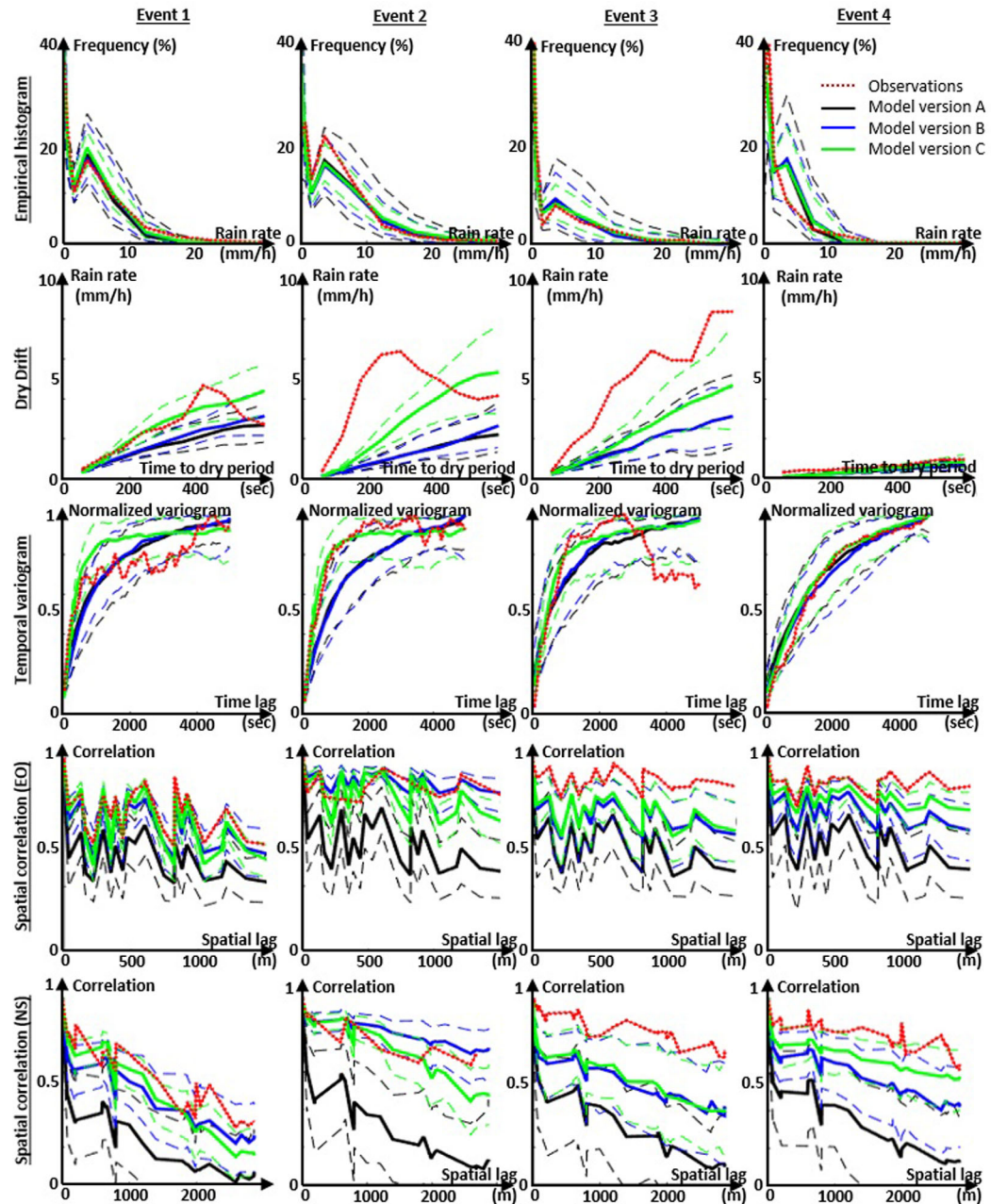
Figure 3. Reproduction of rainfall statistics by unconditional simulation. Rows correspond to the test statistics: row 1, histogram; row 2, dry drift; row 3, temporal variogram; rows 4 and 5, spatial correlogram (along East-West and North-South direction respectively). Columns correspond to the four rain events of interest. Colors correspond to different data sets: red dots, observation data set; black curves, simulations using model version A; blue curves, simulations using model version B; green curves, simulations using model version C. For simulations, the solid lines are the medians of realizations while the dashed lines are Q10 and Q90 of realizations.

the spatial correlograms shows that that the simpler covariance model of version B is not sufficient to reproduce both the spatial and temporal dependencies. Therefore, version B reproduces correctly either the spatial or the temporal dependencies (depending on the rain event), but rarely both together. Conversely, the flexibility added by our model (version C) improves the joint modeling of the space-time dependencies.

To sum up, our model (version C) outperforms simpler model parameterizations (versions A and B) in terms of reproduction of rainfall statistics into unconditional simulations. This proves that the complexity of the covariance function embedded into our model is needed to reproduce the complex space-time

dependencies that are present in local rain fields when observed at high resolution. One can also notice that this model performs well to reproduce most of the selected statistics, except for rain event 3. In this case, the spatial correlation is clearly underestimated. This can be explained by the strong rainfall intermittency during this rain event, which produces many zero measurements and in turn decreases the amount of information available to infer space-time dependencies. This is reflected by the increased uncertainty of the simulated statistics, although this uncertainty is still underestimated and thus fails to encompass the observed statistics.

3.4. Evaluation of Conditional Simulations

In addition to unconditional simulations, the performance of the proposed model for rainfall prediction at ungauged locations is tested through conditional simulations. To this end, a cross-validation experiment is carried out. It consists in predicting rain rate time series at all locations sequentially, while excluding from the conditioning data set the time series corresponding to the location at which rain is predicted. This cross-validation is carried out based on 100 conditional realizations for each rain event and for each model version.

Estimated rain rates are first derived from the conditional simulations in a deterministic way by taking the mean of the 100 realizations. Estimated rain rates are then compared to actual measurements to derive prediction errors. Here the prediction error is defined as the discrepancy between the estimated and the observed rain. Histograms of prediction errors are displayed in Figure 4. Results show that rain predictions are almost unbiased (maximum biases are 0.16, 0.05, and 0.09 mm/h for versions A, B, and C, respectively). In addition, the prediction errors are low, with a proportion of errors under the rain measurement resolution (i.e., 0.6 mm/h) reaching 59%, 71%, and 72% for versions A, B, and C, respectively.

To investigate the performance of the different rain predictions in a probabilistic point of view, the mean Continuous Ranked Probability Score (CRPS; Gneiting & Raftery, 2007) is computed for all cases (Figure 4). This score, expressed in the same unit as the tested variable, assesses the accuracy of an ensemble forecast by summarizing in a single index: (1) the agreement of the observations with the forecast distribution and (2) the sharpness of the forecast distribution. Therefore, an observation falling in the middle of a narrow forecast pdf leads to a low CRPS value, while this score increases as the forecast pdf inflates or as the observed value moves away from the ensemble forecast. In the end, the lower the CRPS, the better the forecast. Computing the mean CRPS for the different cases being tested here shows that the prediction skills improve with the complexity of the model used for the conditional simulation. Indeed the CRPS is reduced by 29% for version B compared to version A, and by 4% for version C compared to version B.

Overall, the comparison of prediction skills for the three versions of the model show that versions B and C accounting for advection clearly outperform version A, which confirms the importance of this phenomenon at the scales of interest. In contrast, versions B and C lead to quite similar prediction skills. This indicates that although the complexity of version C is required to improve the realism of unconditional simulations as demonstrated previously, it does not bring much in term of rainfall prediction by conditional simulation if the conditioning data set is relatively dense (as is the case here).
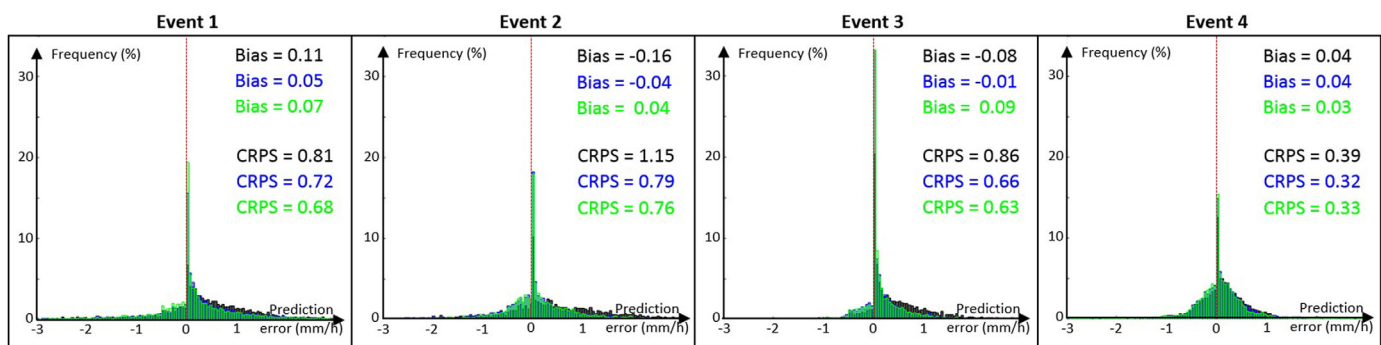

Figure 4. Distribution of prediction errors. Histograms and score statistics related to version A (respectively B and C) are displayed in black (respectively blue and green).
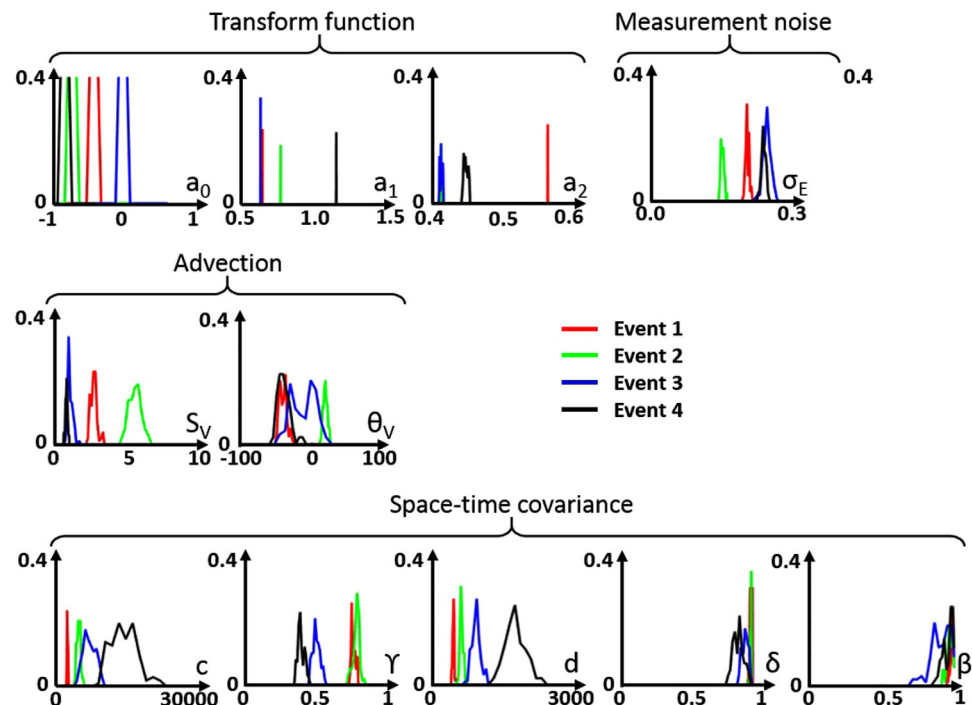
Figure 5. Inferred marginal posterior distributions of model parameters. All posteriors are estimated by MCMC sampling except for parameter a_0 which is directly estimated from the proportion of dry measurements.

3.5. Illustration: Stochastic Rainfall Modeling Over a Small Alpine Catchment

Finally, the full model (version C), demonstrated to be the most appropriate to model high-resolution rain fields, is applied for the stochastic rainfall modeling of the four rain events observed over Vallon de Nant catchment during the summer of 2017 described in section 3.1.

Model calibration provides joint posterior distributions of model parameters for each observation data set (Figure 5). These posterior distributions constitute a quantitative measure of local rainfall statistics for the different rain events. Subsequently, the calibrated stochastic rainfall models can be used to generate conditional realizations at high space and time resolutions which honor rain measurements at gauged locations while reproducing the space-time statistical structure of rainfall inferred from the observation data set. Figure 6 displays samples of such realizations for the four test events. Movies of the full simulations are available in supplemental material.

The statistical structure of the different rain events can consequently be analyzed, by (1) comparing the posterior distributions of model parameters and (2) comparing how the different sets of model parameters translate into simulated rain fields. Results shown in Figures 5 and 6 give the following insights about the observed rain events:

1. Regarding the distribution of rain intensities, shower events (events 1 and 3) are characterized by a higher intermittency (a_0 higher) than frontal events. In addition, convective events (events 2 and 3) are characterized by heavier rainfalls than the stratiform one (event 4). One can also notice that the simulation method is able to generate a dry drift at the transition between dry and wet areas. Note that the inferred marginal distribution of rain intensities is not reproduced accurately in Figure 6 because this figure displays only a small sample of the rain events of interest. In contrast, rain intensity distributions are honored at the scale of the whole rain events, as demonstrated in section 3.3.
2. Regarding the measurement noise amplitude, events 3 and 4 display the lowest signal-to-noise ratios (i.e., the highest amplitudes of measurement noise σ_ϵ). In the case of event 3, this is due to the presence of isolated rain measurements within long rain pauses between thunderstorm-related rainfalls. Such records force the latent field to emerge over the truncation threshold in order to honor the nonzero measurements, which increases the inferred noise amplitude to accommodate it. These isolated rain
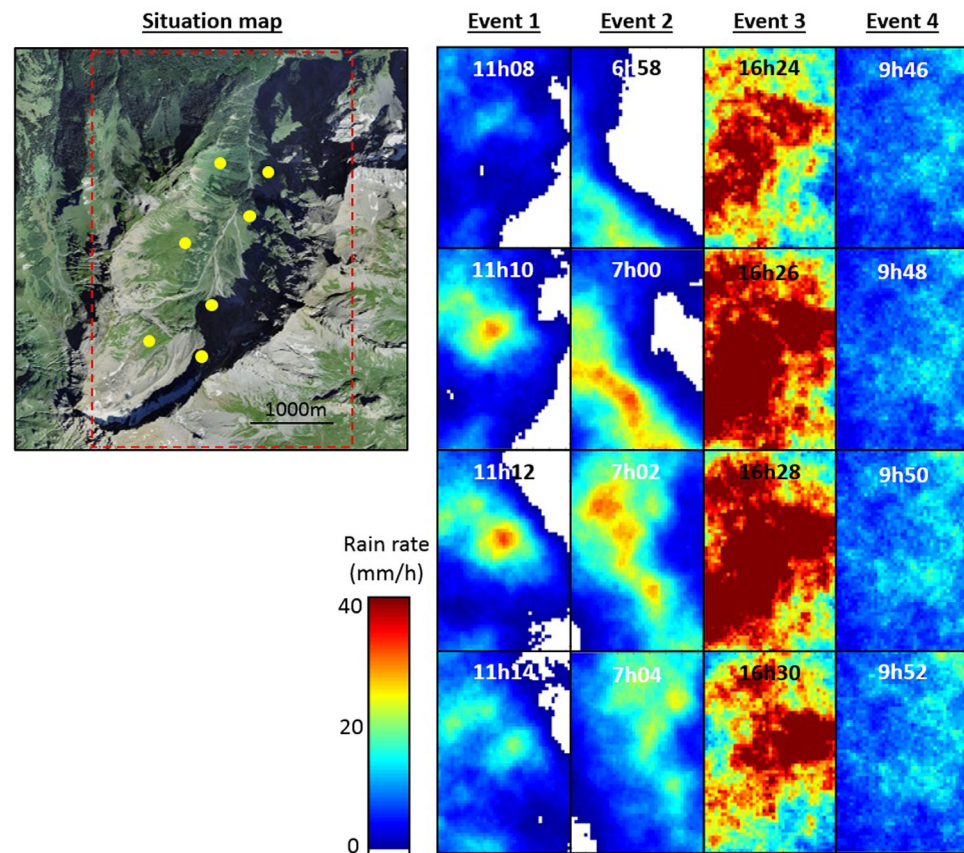
Figure 6. Short samples of space-time rain fields generated by conditional simulation for the four rain events of interest. Spatial resolution, 100 m; temporal resolution, 2 min. (left) Location map. Yellow dots, rain gauge locations. Dashed red lines, footprint of the simulations. (right) Simulated rain fields, from left to right: events 1–4. Movies of the full simulations are available in supporting information.

records are probably due to the presence of a very light rain between the intense rain pulses. The corresponding rain intensities are far below the resolution of the rain gauges (0.6 mm/h), and therefore produce the apparently uncorrelated rain records. In the case of event 4, the low signal-to-noise ratio is due to the overall low rain intensities. This makes the rain signal close to the resolution of the rain gauges, and generates a quantization effect leading to a degraded signal-to-noise ratio. Note that the inferred measurement noise is not translated into the simulations in Figure 6 because we chose to set the simulated measurement noise to zero in these simulations in order to visualize only the rain signal.

3. Regarding storm advection, the inferred advection velocity parameter (S_v) shows that event 2 displays a higher advection velocity than the other events, which is also visible in radar images (not shown here). The estimated advection directions are overall in agreement with the ones derived from radar images (see description of the rain event of interest, section 3.1) except for event 4 where the inferred local advection comes from North-West while the regional advection derived from radar images comes from South-West. This can be due to the low signal-to-noise ratio for this rain event which impairs the estimation of the advection vector. Focusing now on the simulated rain fields (Figure 6), one can notice that the simulations reproduce well the inferred origins of the rain storms. This is clearly visible by inspecting the temporal shift of spatial rain patterns in the simulations.

4. Finally, regarding the space-time dependencies in the Lagrangian reference frame, one can see that very different behaviors are inferred (Figure 5) and in turn translated into the simulated rain fields shown in Figure 6. Focusing first on the scale parameter in space, it appears that the largest value is inferred for the stratiform case (event 4). This translates into large spatial patterns in the corresponding simulations. Focusing next on the regularity parameter in space, it appears that the largest values are inferred for events 1 and 2 traducing a lower short range variability. This leads to smoother simulated rain fields for

these rain events. The temporal evolution of spatial patterns, driven by parameters d and δ, is less diverse than the spatial structure because the regularity parameter in time is similar for all events. Nevertheless, event 4 and in a lesser extent event 3 display a larger scale parameter in time than other events, which translates into more persistent patterns in the corresponding simulations. Finally, one can notice that the inferred values for the separability parameter β range from 0.6 to 1, indicating significant interactions between the spatial and the temporal components.

Overall, the above results show that our stochastic rainfall model, when calibrated using high-resolution rain measurements, is able to capture important features of local rain fields (Figure 5). The use of the calibrated model in a conditional simulation mode leads to realistic synthetic rain fields that are dense in space and time (Figure 6). The final result is an ensemble of equally probable realizations that are consistent with the sparse measurements coming from the network of rain gauges.

4. Conclusion

This paper describes a complete framework for the quantification of high-resolution rainfall. It includes high-resolution data acquisition, parameter inference and the generation of conditional or unconditional rain fields. An important aspect of the proposed model is that it considers rain cells advection and morphing, which are phenomena that become clearly observable at the scales considered. This additional information allows extracting detailed characteristics of the spatiotemporal structure of rainfall, which is not possible with more traditional settings.

The proposed local stochastic rainfall model has been tested on four rain events of different nature that are representative of summer rains over the Swiss Alps. Results show that the model is able to capture in great details the differences of rainfall space-time statistics, and to accurately reproduce these statistics in simulated rain fields. The flexibility of the proposed parametrization allows accounting for the large variety of space-time dependencies arising from high-resolution measurements of local rainfall.

Several types of applications can be considered. First, as illustrated in this paper, our model can be used to measure local rainfall statistics through model parametrization, and to estimate rainfall over the area of interest through conditional simulation. In practice, it can be used to derive high-resolution space-time rain fields over small areas. Due to the moderate cost of the drop-counting rain gauges used in this study (around $500 per device), measuring space-time rain fields using a network of such devices may be feasible even in areas where radar data are not available due to cost issues or wave emission policies. Such rain field estimations are particularly useful to measure rainfall over small urban or mountain catchments because their hydrological response is expected to strongly depend on rainfall variability (Bruni et al., 2015; Peleg et al., 2017; Zoccatelli et al., 2010). Due to the nonlinear behavior of many geomorphological processes, local landscape modeling is also expected to benefit from high-resolution distributed rain fields inputs (Coulthard & Skinner, 2016). A second potential application is the disaggregation of radar images by extending to the context of space-time simulations the idea of conditioning to averaged measurements developed by, e.g., Allcroft and Glasbey (2003) and Onibon et al. (2004). Finally, a third application would be to use the proposed stochastic rainfall model in an unconditional simulation mode as a first step to design a high-resolution stochastic rainfall generator. The generated rain fields can then be used in impact studies focusing on local environments. However, to obtain a complete stochastic rainfall generator, the proposed model still need be nested into another stochastic simulator able to simulate storm arrival and storm duration.

Appendix A: High-Resolution Drop-Counting Rain Gauges: Pluvimates

In this study, the devices used to measure rainfall at high temporal resolution are drop-counting rain gauges from Dryptich called Pluvimates (http://www.driptych.com). They derive from acoustic drip loggers originally used to measure water drop volume at speleothem drip sites (Collister & Mattey, 2008).

A Pluvimate is a drop-counting rain gauge with a resolution of 0.01 mm of rain depth, which is one order of magnitude more resolute than standard tipping bucket rain gauges (resolution 0.1–0.256 mm; Lanza & Vuerich, 2009) and similar to other high resolution but more complex punctual rain measurement devices (Colli et al., 2013; Stow et al., 1998). Pluvimates are based on the following measurement principle: rain water is collected by a funnel (diameter 127 mm) and flows by gravity to its base. At this point, water forms
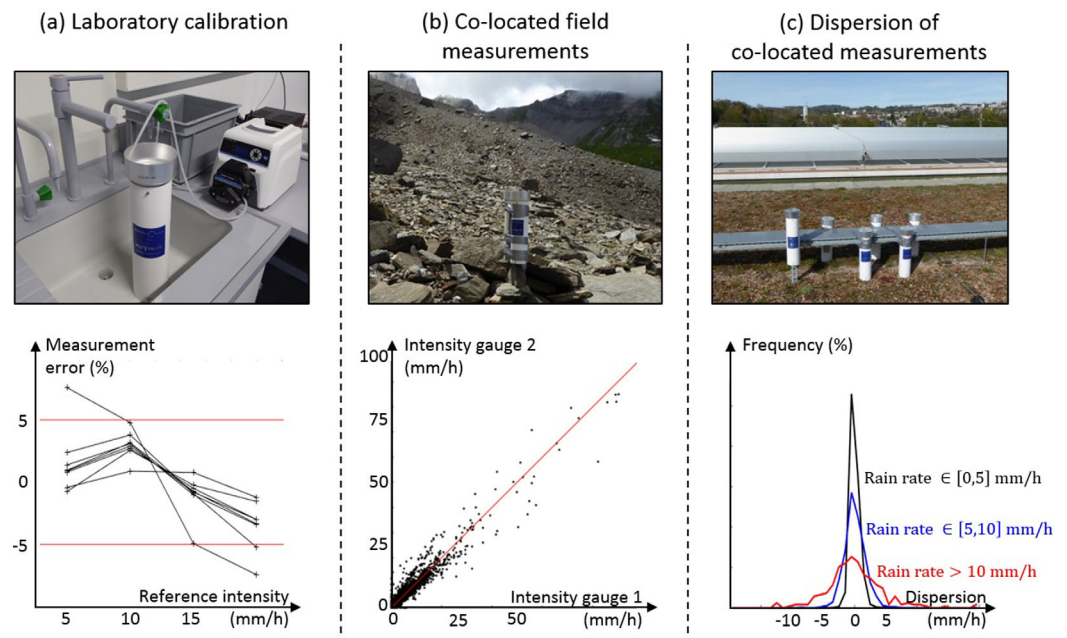
Figure A1. Calibration of Pluvimates. (a) Laboratory calibration. Red lines denote WMO requirements. (b) Colocated field measurements. (c) Dispersion of colocated measurements.

a drop which is suspended to the funnel by capillary forces. When the drop reaches a threshold weight, gravity forces exceed the capillarity tension and the drop falls under the funnel. It then hits an acoustic sensor located at the vertical of the funnel and which records the fall. Because the volume of the drops is constant, the counting of drops is the equivalent of the bucket tips in a standard tipping bucket rain gauge. The volume of the drops generated at the base of the funnel is linked to the geometry of the funnel, and equals 0.125 mL. This volume collected by a 127 mm diameter funnel leads to a 0.01 mm rain depth resolution. If a 1 min integration time is chosen (as is the case in this study) it leads to a rain intensity resolution of 0.6 mm/h.

The Pluvimates have been calibrated prior to their installation in the field. First, the bias of the instruments has been assessed in laboratory by passing a known amount of water through the funnel and by comparing the recorded water depth to the known input. For the 8 calibrated devices, measured biases range from 2.8% to 23% with a mean bias of 13.9%. These biases are corrected for the following of the calibration procedure as well as for field measurements. After bias correction, measurement uncertainty has been assessed through a laboratory experiment. To this end, a controlled and constant flow is injected in the funnel of the rain gauges for four different intensities, following the recommendations of the World Meteorological Organization (WMO; World Meteorological Organization, 2008). Here the reference flow is generated by a high precision rotary pump (Masterflex easy-load2) and is unfortunately limited to intensities ranging from 5 to 20 mm/h due to the limited flow rate of the calibrating pump. Results shown in Figure A1a show that the measurement uncertainty fulfills the WMO recommendations for almost all devices. However, only low rain rates have been tested here due to the limited flow rate of the available calibrating pump, and a nonlinear response of the Pluvimates for higher rain intensities cannot be excluded. According to the Pluvimate manufacturer, the devices are supposed to work properly for rain rates below 10 mm/h, and to possibly saturate

Table A1
Capability of the Pluvimate System for Rain Intensity Measurement

	Range (mm/h)	Resolution (mm/h)	Measurement uncertainty	Averaging time (min)
WMO requirement	0.02–2,000	0.1	5% for >2 mm/h	1
Pluvimate	0.6–80	0.6	5% (tested only for <20 mm/h)	1

for heavier rains. To verify the proper work of the devices during the field experiment presented in this paper (see section 3.1), two rain gauges have been colocated at one measurement location. Plotting measurements of one gauge against the other (Figure A1b) shows good agreement between both devices, and no footpring of saturation for rain intensities under 80 mm/h. No higher intensities have been recorded during this two months experiment, thus no conclusion can be drawn about the operation of Pluvimates under heavier rain conditions. Table A1 summarizes the ability of Pluvimates for rain intensity measurements and compares it to WMO requirements. It shows that Pluvimates perform well for low rain intensities, but their performance is not assessed for intensities > 20 mm/h and is supposed to be poor for intensities > 80 mm/h due to possible saturation.

In addition to the assessment of the measurement uncertainty, the dispersion of Pluvimate measurements was assessed by a relative calibration consisting of setting six colocated Pluvimates recording natural rain events cumulating 100 mm of rain depth with rain intensities ranging from 0 to 40 mm/h. Figure A1c presents the histogram of measurement dispersion d_i (i.e., the deviation of single measurements r_i from the mean of rain measurements $\bar{r} = \frac{1}{6}\sum_{i=1..6} r_i$ over all gauges: $d_i = r_i - \bar{r}$), across the six colocated Pluvimates and for three ranges of rain intensity: 0–5, 5–10, and >10 mm/h. For this calibration data set, the mean measurement dispersion is 0.6 mm/h and the 90% quantile is 1.2 mm/h. One can also notice that the distribution of measurement dispersion has a roughly Gaussian shape, with an amplitude growing with rain intensity. This explains the noise model selected in section 2.2.1.

Appendix B: Metropolis-Hastings Algorithm for MCMC Sampling of Model Parameters

To infer the parameters of the stochastic rainfall model following a Bayesian approach, a Markov Chain Monte Carlo (MCMC) sampling strategy has been selected due to the complexity of the stochastic model and the interlinking of its parameters. If the data set used for parameter inference is of limited size ($<$few thousand observations) we propose to use a standard Metropolis-Hastings algorithm (Algorithm B1 hereafter) to carry out this MCMC sampling.

Algorithm B1: Usual Metropolis-Hastings algorithm

Input: Data **R**; prior distribution $\pi(\theta)$; proposal distribution $q(\cdot|\cdot)$; initial parameter vector θ^0

Output: Samples of parameters sets θ representative of the posterior distribution

1. Compute $\ell(\theta^0|\mathbf{R})$ and $\pi(\theta^0)$

2. **for $k = 0$ to $N - 1$**

   (a) $\theta^p \sim q(\cdot|\theta^k)$

   (b) Compute $\ell(\theta^p|\mathbf{R})$ and $\pi(\theta^p)$.

   (c) Compute $p_A = \min\left\{1, \frac{\ell(\theta^p|\mathbf{R})\pi(\theta^p)}{\ell(\theta^k|\mathbf{R})\pi(\theta^k)}\right\}$,

   (d) Draw $U \sim U_{[0,1]}$

   (e) **if** $p_A \leq U$ **then**

   $\quad\quad \theta^{k+1} = \theta^p$

     else

   $\quad\quad \theta^{k+1} = \theta^k$

     end

  end

return $\{\theta^k_{k=0,...,N}\}$

Algorithm B2: Metropolis-Hastings algorithm with latent variables

Input: Data $\mathbf{R}$; prior distributions $\pi(\boldsymbol{\theta})$; proposal distribution $q(\cdot|\cdot)$; initial parameter vector $\boldsymbol{\theta}^0$; initial latent values $\mathbf{Z}_{I^0}$

Output: Samples of parameters sets $\boldsymbol{\theta}$ and $\mathbf{Z}_{I^0}$

1. Set $\mathbf{Z} = (\mathbf{Z}_I, \mathbf{Z}_{I^0})$ and $\mathbf{R} = (\mathbf{R}_I, \Psi(\mathbf{Z}_{I^0}))$. Compute $\ell(\boldsymbol{\theta}^0|\mathbf{R})$ and $\pi(\boldsymbol{\theta}^0)$

2. **for** $k = 0$ **to** $N - 1$

   (a) $\boldsymbol{\theta}^p \sim q(\cdot|\boldsymbol{\theta}^k)$

   (b) Compute $\ell(\boldsymbol{\theta}^p|\mathbf{R})$ and $\pi(\boldsymbol{\theta}^p)$.

   (c) Compute $p_A = \min\left\{1, \frac{\ell(\mathbf{R};\boldsymbol{\theta}^p)\pi(\boldsymbol{\theta}^p)}{\ell(\mathbf{R};\boldsymbol{\theta}^k)\pi(\boldsymbol{\theta}^k)}\right\}$,

   (d) Draw $U \sim U_{[0,1]}$

   (e) **if** $p_A \leq U$ **then**

   $\qquad \boldsymbol{\theta}^{k+1} = \boldsymbol{\theta}^p$

   else

   $\qquad \boldsymbol{\theta}^{k+1} = \boldsymbol{\theta}^k$

   end

   (f) Gibbs sampler

   For $i = 1$ **to** N_{I^0}

   $\qquad$ i. Compute $Z_{I_0,i}^p \sim \mathcal{N}(\Sigma_{-i,i}^t \Sigma_{-i,-i}^{-1} \mathbf{Z}_{-i}, 1 - \Sigma_{-i,i}^t \Sigma_{-i,-i}^{-1} \Sigma_{-i,i})$

   $\qquad$ ii. **if** $Z_{I_0,i}^p > a_0$

   $\qquad\qquad$ go to i.

   $\qquad$ **end**

   end

   Set $\mathbf{Z} = (\mathbf{Z}_I, \mathbf{Z}_{I^0})$ and $\mathbf{R} = (\mathbf{R}_I, \Psi(\mathbf{Z}_{I^0}))$

   end

return $\{\boldsymbol{\theta}^k, \mathbf{Z}_{I_0}^k, k = 0, \ldots, N\}$

If the available calibration data set exceeds few thousand observations, we propose to use a modified version of the Metropolis-Hasting to carry out the inference of model parameters. It consists in simulating the censored values of the latent variable by Gibbs sampling within the Metropolis-Hastings algorithm in order

Table B1
Priors Used in Metropolis-Hasting Sampler

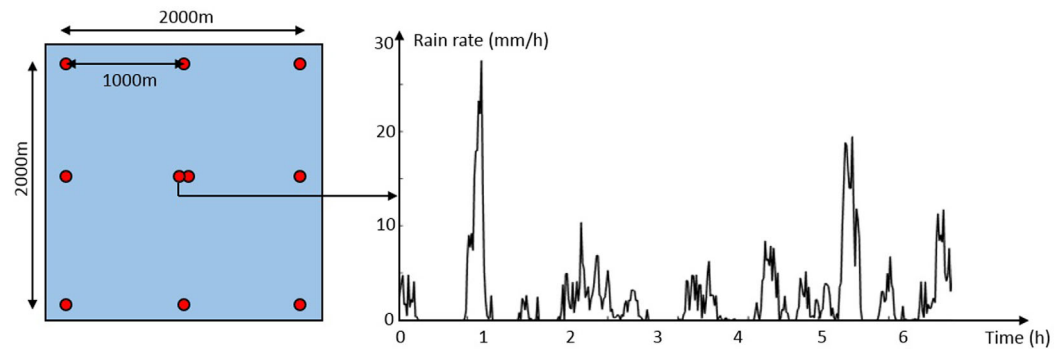|	Anamorphosis		Noise	Advection	
Parameter	a_1	a_2	σ_ϵ	S_V	θ_V
Prior	$\mathcal{U}[0.01, 2]$	$\mathcal{U}[0.1, 1.5]$	$\mathcal{U}[0.01, 1]$	$\mathcal{U}[4, 20]$	$\mathcal{U}[-180, 180]$
Amplitude of proposal step $Q(x, .) \sim \mathcal{U}[x-a, x+a]$	$a = 0.001$	$a = 0.001$	$a = 0.01$	$a = 0.5$	$a = 5.0$
			Covariance		
Parameter	c	γ	d	δ	β
Prior	$\mathcal{U}[5e^2, 1e^5]$	$\mathcal{U}[0.1, 0.9]$	$\mathcal{U}[1e^2, 1.5e^4]$	$\mathcal{U}[0.1, 0.9]$	$\mathcal{U}[0, 1]$
Amplitude of proposal step $Q(x, .) \sim \mathcal{U}[x-a, x+a]$	$a = 1{,}000$	$a = 0.01$	$a = 100$	$a = 0.01$	$a = 0.01$

Figure C1. Synthetic data used for model assessment: (left) schematic map of the network and (right) synthetic time series at the central location.

to avoid the computation of the joint cdf of zero measurements given the positive rain observations. This is done using Algorithm B2 presented hereafter:

In practice, prior and proposal distributions must be specified by the user for each model parameter to be inferred. Here vague uniform priors have been used for all parameters and these priors are scanned by random walk proposals. Table B1 summarizes the priors and proposal distributions used in this study.

Appendix C: Assessment of the Blockwise Likelihood Approach for Model Calibration

To be applicable in practice, the stochastic rainfall model proposed in section 2 requires approximations for parameter inference. The impact of these approximations on model calibration is assessed here through a synthetic case study that mimics as closely as possible real case applications while being of reasonable size to be computationally-manageable when nonapproximate methods are applied for benchmarking. The test
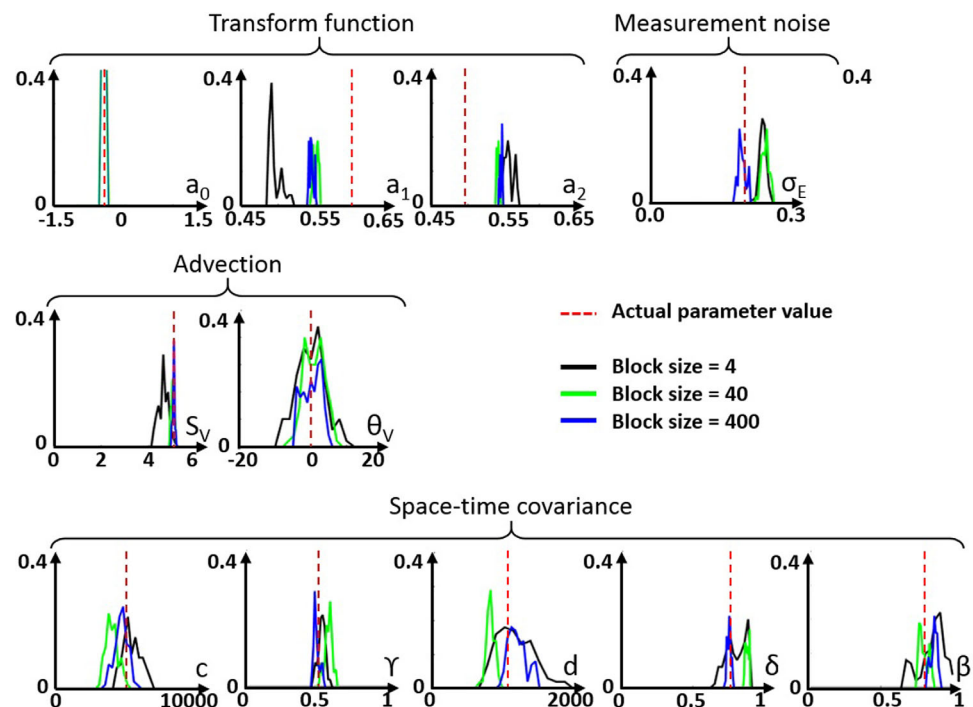

Figure C2. Inferred marginal posterior distributions of model parameters for the three tested block sizes: block size = 4 (black curves); block size = 40 (green curves); block size = 400 = full likelihood (blue curves).

Table C1
Stochastic Rainfall Model Parameters Used for Synthetic Data Generation

c	γ	d	δ	β	σ_ϵ	a_0	a_1	a_2	S_V	S_θ
5,000	0.5	1,000	0.8	0.8	0.2	-0.5	0.6	0.5	5	0.0

consists of calibrating the proposed stochastic rainfall model using a measurement data set composed of synthetic rain rate time series simulated from a model whose parameters are known and summarized in Table C1. Synthetic observations gather ten rain rate time series (duration: 6h40min; temporal resolution: 1 min) spread over a 4 km^2 area. Virtual observation devices are located on a square grid with 1,000 m long edges, with two collocated gauges at the central location of the network (Figure 1).

The parameters of the stochastic rainfall model are inferred using the Bayesian framework proposed in section 2.3. Three different block sizes are tested when computing the likelihood of observations in order to assess the impact of this parameter on posterior estimation: 4 time steps blocks, 40 time steps blocks, and 400 time steps blocks. This last block size (400 time steps) corresponds to the full likelihood. The same parameters of the Metropolis-Hastings algorithm are used for the three cases: burn-in period of 5,000 iterations to ensure initial convergence, and then sampling chain of 10,000 iterations to derive the posterior distribution of model parameters.

Figure 2 shows the inferred marginal distributions of parameter posteriors for the three tested block sizes and compares them to the reference values. It appears that model parameter posteriors present really few deviations from the true values. The small discrepancies that exist between posteriors and theoretical values are probably partly due to the ergodic fluctuations in the simulation process used to generate the simulation data set, which can have produced a realization with statistics departing slightly from their expected values. The posteriors obtained with the three different block sizes are very similar, which confirms the reliability of the blockwise likelihood approximation. Nevertheless one can notice that posteriors derived from the full likelihood are in slightly better agreement with the theoretical parameter values.

Regarding the computing time, using a computer equipped with two Intel Xeon CPU-E5–2699 processors, case 1 (4 time step blocks) required 14 min to complete parameter inference, while case 2 (40 time step blocks) required 1 h, and the full likelihood case required 1 day and 2 h. It is worth noting that the difference of performance between the different methods increases with the size of the data set. Indeed the full likelihood requires inversion and determinant computation of covariance matrices whose size grows as the square of the size of the data set, while the blockwise likelihood involves only small matrices of fixed size. Consequently, considering the mild loss of performance but the considerable speed up obtained using a blockwise likelihood approach, we recommend to use this approximation for large data sets.

References

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

Ailliot, P., Thompson, C., & Thomson, P. (2009). Space-time modelling of precipitation by using a hidden Markov model and censored Gaussian distributions. *Applied Statistics*, *58*, 405–426. https://doi.org/10.1111/j.1467-9876.2008.00654.x

Allard, D., & Bourotte, M. (2015). Disaggregating daily precipitations into hourly values with a transformed censored latent Gaussian process. *Stochastic Environmental Research and Risk Assessment*, *29*, 453–462. https://doi.org/10.1007/s00477-014-0913-4

Allcroft, D. J., & Glasbey, C. A. (2003). A latent Gaussian Markov random-field model for spatiotemporal rainfall disaggregation. *Applied Statistics*, *52*, 487–498. https://doi.org/10.1111/1467-9876.00419

Anagnostou, M. N., Kalogiros, J., Anagnostou, E. N., Tarolli, M., Papadopoulos, A., & Borga, M. (2010). Performance evaluation of high-resolution rainfall estimation by X-band dual-polarization radar for flash flood applications in mountainous basins. *Journal of Hydrology*, *394*, 4–16. https://doi.org/10.1016/j.jhydrol.2010.06.026

Bárdossy, A., & Pegram, G. G. S. (2016). Space-time conditional disaggregation of precipitation at high resolution via simulation. *Water Resources Research*, *52*, 920–937. https://doi.org/10.1002/2015WR018037

Bauer, P., Thorpe, A., & Brunet, A. (2015). The quiet revolution of numerical weather prediction. *Nature*, *525*, 47–55. https://doi.org/10.1038/nature14956

Baxevani, A., & Lennartsson, J. (2015). A spatiotemporal precipitation generator based on a censored latent Gaussian field. *Water Resources Research*, *51*, 4338–4358. https://doi.org/10.1002/2014WR016455

Benoit, L., & Mariethoz, G. (2017). Generating synthetic rainfall with geostatistical simulations. *WIRES Water*, *4*, e1199. https://doi.org/10.1002/wat2.1199

Bevilacqua, M., & Gaetan, C. (2015). Comparing composite likelihood methods based on pairs for spatial Gaussian random fields. *Statistics and Computing*, *25*(5), 877–892. https://doi.org/10.1007/s11222-014-9460-6

Bevilacqua, M., Gaetan, C., Mateu, J., & Porcu, E. (2012). Estimating space and space-time covariance functions for large data sets: A weighted composite likelihood approach. *Journal of American Statistical Association*, *107*(497), 268–280. https://doi.org/10.1080/01621459.2011.646928

Bony, S., Stevens, B., Frierson, D., Jakob, C., Kageyama, M., Pincus, R., et al. (2015). Clouds, circulation and climate sensitivity. *Nature Geoscience*, *8*, 261–268. https://doi.org/10.1038/ngeo2398

Bourotte, M., Allard, D., & Porcu, E. (2016). A flexible class of non-separable cross-covariance functions for multivariate space-time data. *Spatial Statistics*, *18*, 125–146. https://doi.org/10.1016/j.spasta.2016.02.004

Bruni, G., Reinoso, R., van de Giesen, N. C., Clemens, F. H. L. R., & ten Veldhuis, J. A. E. (2015). On the sensitivity of urban hydrodynamic modelling to rainfall spatial and temporal resolution. *Hydrology and Earth System Sciences*, *19*, 691–709. https://doi.org/10.5194/hess-19-691-2015

Burton, A., Fowler, H. J., Blenkinsop, S., & Kilsby, C. G. (2010). Downscaling transient climate change using a Neyman-Scott Rectangular Pulses stochastic rainfall model. *Journal of Hydrology*, *381*, 18–32. https://doi.org/10.1016/j.jhydrol.2009.10.031

Caseri, A., Javelle, P., Ramos, M. H., & Leblois, E. (2016). Generating precipitation ensembles for flood alert and risk management. *Journal of Flood Risk Management*, *9*, 402–415. https://doi.org/10.1111/jfr3.12203

Colli, M., Lanza, L. G., & Chan, P. W. (2013). Co-located tipping-bucket and optical drop counter RI measurements and a simulated correction algorithm. *Atmospheric Research*, *119*, 3–12. https://doi.org/10.1016/j.atmosres.2011.07.018

Collister, C., & Mattey, D. (2008). Controls on water drop volume at speleothem drip sites: An experimental study. *Journal of Hydrology*, *358*, 259–267. https://doi.org/10.1016/j.jhydrol.2008.06.008

Comunian, A., Renard, P., & Straubhaar, J. (2012). 3D multiple-point statistics simulation using 2D training images. *Computers and Geosciences*, *40*, 49–65. https://doi.org/10.1016/j.cageo.2011.07.009

Coulthard, T. J., & Skinner, J. (2016). The sensitivity of landscape evolution models to spatial and temporal rainfall resolution. *Earth Surface Dynamics*, *4*, 757–771. https://doi.org/10.5194/esurf-4-757-2016

Creutin, J. D., Leblois, E., & Lepioufle, J. M. (2015). Unfreezing Taylor's hypothesis for precipitation. *Journal of Hydrometeorology*, *16*, 2443–2462. https://doi.org/10.1175/JHM-D-14-0120.1

Deidda, R. (2000). Rainfall downscaling in a space-time multifractal framework. *Water Resources Research*, *36*, 1779–1794. https://doi.org/10.1029/2000WR900038

Diggle, P. J., & Ribeiro, P. J. (2002). Bayesian inference in Gaussian model-based geostatistics. *Geographical and Environmental Modelling*, *6*, 129–146. https://doi.org/10.1080/1361593022000029467

Emery, X., Arroyo, D., & Porcu, E. (2016). An improved spectral turning-bands algorithm for simulating stationary vector Gaussian random fields. *Stochastic Environmental Research and Risk Assessment*, *30*(7), 1863–1873.

Emery, X., & Lantuéjoul, C. (2006). TBSIM: A computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method. *Computers and Geosciences*, *32*, 1615–1628. https://doi.org/10.1016/j.cageo.2006.03.001

Emmanuel, I., Andrieu, H., Leblois, E., & Flahaut, B. (2012). Temporal and spatial variability of rainfall at the urban hydrological scale. *Journal of Hydrology*, *430–431*, 162–172. https://doi.org/10.1016/j.jhydrol.2012.02.013

Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in practice*. Dordrecht, the Netherlands: Springer Science + Business Media.

Gneiting, T. (2002). Nonseparable, stationary covariance functions for space-time data. *Journal of American Statistical Association*, *97*(458), 590–600. https://doi.org/10.1198/016214502760047113

Gneiting, T., Genton, M., & Guttorp, P. (2007). Geostatistical space-time models, stationarity, separability and full symmetry. In B. Finkenstädt, L. Held, & V. Isham (Eds.), *Statistical methods for spatio-temporal systems* (pp. 151–175). London, UK: Chapman and Hall.

Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, *102*(477), 359–378. https://doi.org/10.1198/016214506000001437

Gómez-Hernández, J. J., & Journel, A. G. (1993). Joint sequential simulation of multi-Gaussian fields. In *Geostatistics Troia'92* (pp. 85–94). Berlin, Germany: Springer. https://doi.org/10.1007/978-94-011-1739-5-8

Guillot, G. (1999). Approximation of Sahelian rainfall fields with meta-Gaussian random functions; Part 1: Model definition and methodology. *Stochastic Environmental Research and Risk Assessment*, *13*, 1000–1112. https://doi.org/10.1007/s004770050034

Jaffrain, J., Studzinski, A., & Berne, A. (2011). A network of disdrometers to quantify the small-scale variability of the raindrop size distribution. *Water Resources Research*, *47*, W00H06. https://doi.org/10.1029/2010WR009872

Katz, R. W., & Parlange, M. B. (1995). Generalizations of chain-dependent processes: Application to hourly precipitations. *Water Resources Research*, *5*, 1331–1341. https://doi.org/10.1029/94WR03152

Kleiber, W., Katz, R. W., & Rajagopalan, B. (2012). Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes. *Water Resources Research*, *48*, W01523. https://doi.org/10.1029/2011WR011105

Krajewski, W. F., Ciach, G., & Habib, E. (2003). An analysis of small-scale rainfall variability in different climatic regimes. *Hydrological Sciences Journal*, *48*, 151–162. https://doi.org/10.1623/hysj.48.2.151.44694

Lanza, L. G., & Vuerich, E. (2009). The WMO field intercomparison of rain intensity gauges. *Atmospheric Research*, *94*, 534–543. https://doi.org/10.1016/j.atmosres.2009.06.012

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

Lengfeld, K., Clemens, M., Münster, H., & Ament, F. (2014). Performance of high-resolution X-band weather radar networks—The PATTERN example. *Atmospheric Measurement Techniques*, *7*, 4151–4166. https://doi.org/10.5194/amt-7-4151-2014

Lepioufle, J. M., Leblois, E., & Creutin, J. D. (2012). Variography of rainfall accumulation in presence of advection. *Journal of Hydrology*, *464–465*, 494–504. https://doi.org/10.1016/j.jhydrol.2012.07.041

Le Ravalec, M., Noetinger, B., & Hu, L. Y. (2000). The FFT Moving Average (FFT-MA) generator: An efficient numerical method for generating and conditioning Gaussian simulations. *Mathematical Geology*, *32*(6), 701–722. https://doi.org/10.1023/A:1007542406333

Lindgren, F., Håvard, R., & Lindstrom, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *Journal of the Royal Statistical Society*, *73*, 423–498. https://doi.org/10.1111/j.1467-9868.2011.00777.x

Lindsay, B. G. (1988). Composite likelihood methods. *Contemporary Mathematics*, *80*(1), 221–239. https://doi.org/10.1090/conm/080/999014

Mandapaka, P. V., Krajewski, W. F., Mantilla, R., & Gupta, V. K. (2009). Dissecting the effect of rainfall variability on the statistical structure of peak flows. *Advances in Water Resources*, *32*, 1508–1525. https://doi.org/10.1016/j.advwatres.2009.07.005

Mantoglou, A., & Wilson, J. L. (1982). The Turning Bands Method for simulation of random fields using line generation by a spectral method. *Water Resources Research*, *18*, 1379–1394. https://doi.org/10.1029/WR018i005p01379

Mavromatis, T., & Hansen, J. W. (2001). Interannual variability characteristics and simulated crop response of four stochastic weather generators. *Agricultural and Forest Meteorology*, *109*, 283–296. https://doi.org/10.1016/S0168-1923(01)00272-6

Molnar, P., & Burlando, P. (2005). Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model. *Atmospheric Research*, *77*, 137–151. https://doi.org/10.1016/j.atmosres.2004.10.024

Nussbaumer, R., Mariethoz, G., Gloagen, E., & Holliger, K. (2017). Which path to choose in sequential Gaussian simulation? *Mathematical Geosciences*, https://doi.org/10.1007/s11004-017-9699-5

Onibon, H., Lebel, T., Afoufa, A., & Guillot, G. (2004). Gibbs sampling for conditional spatial disaggregation of rain fields. *Water Resources Research*, *40*, W08401. https://doi.org/10.1029/2003WR002009

Onov, C., Chandler, R. E., Kakou, A., Northrop, A., Wheater, H. S., & Isham, V. (2000). Rainfall modelling using Poisson-cluster processes: A review of developments. *Stochastic Environmental Research and Risk Assessment*, *14*, 384–411. https://doi.org/10.1007/s004770000043

Pardo-Igúzquiza, E., Grimes, D. I. F., & Teo, C. K. (2006). Assessing the uncertainty associated with intermittent rainfall fields. *Water Resources Research*, *42*, W01412. https://doi.org/10.1029/2004WR003740

Paschalis, A., Fatichi, S., Molnar, P., & Burlando, P. (2014). On the effects of small scale space-time variability of rainfall on basin flood response. *Journal of Hydrology*, *514*, 313–327. https://doi.org/10.1016/j.jhydrol.2014.04.014

Paschalis, A., Molnar, P., Fatichi, S., & Burlando, P. (2013). A stochastic model for high-resolution space-time precipitation simulation. *Water Resources Research*, *49*, 8400–8417. https://doi.org/10.1002/2013WR014437

Pauli, F., Racugno, W., & Ventura, L. (2011). Bayesian composite marginal likelihoods. *Statistica Sinica*, *21*, 149–164.

Peleg, N., Blumensaat, F., Molnar, P., Fatichi, S., & Burlando, P. (2017). Partitioning the impacts of spatial and climatological rainfall variability in urban drainage modeling. *Hydrology and Earth System Sciences*, *21*, 1559–1572. https://doi.org/10.5194/hess-21-1559-2017

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

Ribatet, M., Cooley, D., & Davison, A. C. (2012). Bayesian inference from composite likelihoods, with an application to spatial extremes. *Statistica Sinica*, *22*, 813–845. https://doi.org/10.5705/ss.2009.248

Richardson, C. W. (1981). Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resources Research*, *17*, 182–190. https://doi.org/10.1029/WR017i001p00182

Schleiss, M., Chamon, S., & Berne, A. (2014a). Nonstationarity in intermittent rainfall: The 'dry drift.' *Journal of Hydrometeorology*, *15*, 1189–1204. https://doi.org/10.1175/JHM-D-13-095.1

Schleiss, M., Chamon, S., & Berne, A. (2014b). Stochastic simulation of intermittent rainfall using the concept of 'dry drift.' *Water Resources Research*, *50*, 2329–2349. https://doi.org/10.1002/2013WR014641

Sigrist, F., Künsch, H. R., & Stahel, W. A. (2012). A dynamic nonstationary spatio-temporal model for short term prediction of precipitation. *The Annals of Applied Statistics*, *6*, 1452–1477. https://doi.org/10.1214/12-AOAS564

Stow, C. D., Bradley, S. G., Farrington, K. E., Dirks, K. N., & Gray, W. R. (1998). A rain gauge for the measurement of finescale temporal variations. *Journal of Atmospheric and Oceanic Technology*, *15*, 127–135. https://doi.org/10.1175/1520-0426(1998)015<0127:ARGFTM>2.0.CO;2

Thorndahl, S., Einfalt, T., Willems, P., Nielsen, J. E., ten Veldhuis, M.-C., Arnbjerg-Nielsen, K., et al. (2017). Weather radar rainfall data in urban hydrology. *Hydrology and Earth System Sciences*, *21*, 1359–1380. https://doi.org/10.5194/hess-21-1359-2017

Tran, T. T. (1994). Improving variogram reproduction on dense simulation grids. *Computers and Geosciences*, *20*, 1161–1168. https://doi.org/10.1016/0098-3004(94)90069-8

Vaittinada Ayar, P., Vrac, M., Bastin, S., Carreau, J., Déqué, M., & Gallardo, C. (2016). Intercomparison of statistical and dynamical downscaling models under the EURO- and MED-CORDEX initiative framework: Present climate evaluations. *Climate Dynamics*, *46*, 1301–1329. https://doi.org/10.1007/s00382-015-2647-5

Varin, C., Reid, N., & Firth, D. (2011). An overview of composite likelihood methods. *Statistica Sinica*, *21*(1), 5–42.

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

Vischel, T., Quantin, G., Lebel, T., Viarre, J., Gosset, M., Cazenave, F., et al. (2011). Generation of high-resolution rain fields in West Africa: Evaluation of dynamic interpolation methods. *Journal of Hydrometeorology*, *12*, 1465–1482. https://doi.org/10.1175/JHM-D-10-05015.1

von Ruette, J., Lehmann, P., & Or, D. (2014). Effects of rainfall spatial variability and intermittency on shallow landslide triggering patterns at a catchment scale. *Water Resources Research*, *50*, 7780–7799. https://doi.org/10.1002/2013WR015122

Wilks, D. S. (2010). Use of stochastic weather generators for precipitation downscaling. *WIREs Climate Change*, *1*, 898–907. https://doi.org/10.1002/wcc.85

Wilks, D. S., & Wilby, R. L. (1999). The weather generation game: A review of stochastic weather models. *Progress in Physical Geography*, *23*, 329–357. https://doi.org/10.1177/030913339902300302

Willems, P. (2001). A spatial rainfall generator for small spatial scales. *Journal of Hydrology*, *252*, 126–144. https://doi.org/10.1016/S0022-1694(01)00446-2

World Meteorological Organization. (2008). *Guide to meteorological instruments and methods of observation*.

Zoccatelli, D., Borga, M., Zanon, F., Antonescu, B., & Stancalie, G. (2010). Which rainfall spatial information for flash flood response modelling? A numerical investigation based on data from the Carpathian range, Romania. *Journal of Hydrology*, *394*, 148–161. https://doi.org/10.1016/j.jhydrol.2010.07.019