



Dealing with non-stationarity in sub-daily stochastic rainfall models

Lionel Benoit¹, Mathieu Vrac², and Gregoire Mariethoz¹

¹Institute of Earth Surface Dynamics (IDYST), University of Lausanne, Lausanne, Switzerland

²Laboratory for Sciences of Climate and Environment (LSCE-IPSL), CNRS/CEA/UVSQ, Orme des Merisiers, France

Correspondence: Lionel Benoit (lionel.benoit@unil.ch)

Abstract.

Understanding the stationarity properties of rainfall is critical when using stochastic weather generators. Rainfall stationarity means that the statistics being accounted for remain constant over a given period, which is required for both inferring model parameters and simulating synthetic rainfall. Despite its critical importance, the stationarity of precipitation statistics is often regarded as a subjective choice whose examination is left to the judgement of the modeler. It is therefore desirable to establish quantitative and objective criteria for defining stationary rain periods. To this end, we propose a methodology that automatically identifies rain types with homogeneous statistics. It is based on an unsupervised classification of the space - time - intensity structure of weather radar images. The transitions between rain types are interpreted as non-stationarities. Our method is particularly suited to deal with non-stationarity in the context of sub-daily stochastic rainfall models. Results of a synthetic case study show that the proposed approach is able to reliably identify synthetically generated rain types. The application of rain typing to real data indicates that non-stationarity can be significant within meteorological seasons, and even within a single storm. This highlights the need for a careful examination of the temporal stationarity of precipitation statistics when modelling rainfall at high resolution.

1 Introduction

Stochastic rainfall models are statistical models that aim at simulating realistic random rains. For this purpose, they generate rainfall simulations which reproduce, in a distributional sense, a set of key rainfall statistics derived from an observation dataset. The practical interest of stochastic rainfall models is notably to complement numerical weather models for the simulation of rainfall heterogeneity at fine scales, and to quantify the uncertainty associated with rainfall reconstructions. Indeed, numerical weather models face challenges for reproducing rainfall heterogeneity in space and time, in particular at fine scales (Bauer et al., 2015; Bony et al., 2015). Some of the main applications of stochastic rainfall models are therefore the fast generation of synthetic rainfall inputs for local impact studies related for instance to hydrology (Paschalis et al., 2014; Caseri et al., 2016) or agronomy (Mavromatis and Hansen, 2001; Qian et al., 2011), and the downscaling of aggregated precipitation products such as rain observations (Allcroft and Glasbey, 2003; Bárdossy and Pegram, 2016) or numerical model outputs (Wilks, 2010; Vaittinada Ayar et al., 2016). In all cases, the target is the transposition of observed rain statistics into synthetic rain simulations. Recently, a considerable attention has been paid to increasing the resolution of stochastic rainfall models so that they can mimic rainfall at sub-daily time scales. Currently, several high resolution stochastic rainfall models are able to deal with precipitation



data at typical resolutions of 1 min to 1 h in time and of 100 x 100 m² to 1 x 1 km² in space (see e.g. (Leblois and Creutin, 2013; Paschalis et al., 2013; Benoit et al., In Press). At such scales, not only the marginal distribution of observed rain intensity matters, but the space-time dependencies within rain fields are also important features of the rain process (Emmanuel et al., 2012; Marra and Morin, 2018). In particular, the impact of the advection and diffusion of spatial rainfall patterns (e.g. rain
5 cells or rain bands) have to be modelled (Lepioufle et al., 2012; Creutin et al., 2015). In consequence, most sub-daily stochastic rainfall models consider rainfall as a space-time random process.

An underlying hypothesis in stochastic rainfall modelling is that of stationarity: the statistics of rainfall are supposed to be constant over a given (space-time) modelling domain. This enables (1) the inference of rainfall statistics from an observation dataset, and (2) the reproduction of these statistics in simulations. The definition of stationary domains can be regarded as a
10 modelling choice, often subjective and left to the judgment of modelers (Journel, 1993). It consists of defining pools of data that are considered similar enough (in a statistical sense) to perform model inference. In the case of stochastic rainfall modelling, the identification of stationary datasets or sub-datasets relies on some phenomenological guesses about rainfall, which serve as fuzzy guidelines to delineate stationary domains. Depending on the application and modeling choices, the statistical structure used for sub-daily stochastic rainfall is considered as changing at scales ranging from seasons (Paschalis et al., 2013; Bárdossy
15 and Pegram, 2016; Peleg et al., 2017) to single rain storms (Casari et al., 2016; Benoit et al., In Press).

One possible approach to delineate pools of homogeneous rain observations in a more quantitative way is to classify them prior to modelling. A set of predefined criteria is used to build a metric of similarity between the observations, and a classification algorithm is applied to the resulting similarity measures in order to define clusters of closely related rain observations. The result of such a classification procedure, often referred to as rain typing, is the identification of a limited number of rain types
20 which gather rain observation sharing similar properties. Until recently, rain typing mainly focused on classifications based on rain intensity, with the aim to assess the physical processes responsible of rain generation (e.g. distinguish convective and stratiform rains) (Rosenfeld and Amitai, 1995; Biggerstaff and Listemaa, 2000; Llasat, 2001). In the last years, the emergence of metrics characterizing rainfall spatial or space-time behavior (Vrac et al., 2007; Ramirez-Cobo et al., 2010; Aghakouchak et al., 2011; Zick and Matyas, 2016) paved the way to new rain typing methods based on the arrangement of rain fields in space
25 and in time (Leblois, 2012; Lagrange et al., 2018).

In this context, we propose a rain typing framework based on the space - time - intensity statistical signature of rainfall to automatically identify and delineate domains for which rainfall statistics are as stationary as possible. In the case of sub-daily stochastic rainfall models, the property considered is the stationarity of space-time statistics, for a given area of interest and modelling period. However, the question of stationarity in space is not addressed here because we focus on regional to local
30 areas. Hence, the spatial extent of the modelling domain is small enough to ensure that all locations experience statistically similar precipitations, leading to spatial stationarity. This is a strong assumption that may not hold in case of heterogeneous topography (e.g. coastline, changing orography) or for large areas. However, in this study, we restrict our investigation to temporal non-stationarities. The main aim of this paper is therefore to design a quantitative method to identify changes in the space-time statistical structure of rainfall at high temporal resolution (10 min in the proposed test application). The proposed
35 framework relies on the classification of radar images based on their space-time features. The resulting classes are then used to



define rain types that group rain fields with similar statistical signatures. Finally, the transition between rain types is interpreted as a break in the temporal stationarity of rainfall statistics.

The remainder of this paper is structured as follows: section 2 gives a general overview of rainfall space-time patterns visible in radars images. Section 3 describes a rain typing method based on the previously identified patterns, and explains how the resulting rain types can be used to identify stationary periods. Then, section 4 assesses the performance of this method for both synthetic and real case studies. Finally, Section 4.3 discusses the implications of the observed patterns of non-stationarity for sub-daily stochastic rainfall modelling and gives general conclusions.

2 Overview of rainfall space-time patterns observed in radar images

Prior to the design of a quantitative method to identify non-stationarities in rainfall statistics, the current section seeks to illustrate with some typical examples the diversity of space-time patterns that can be observed in rain fields, and to give an overview of their temporal evolutions.

For reasons of data availability, we focus in this paper on summer rains observed from 2017 July, 1 to 2017 August, 31. During this period, only periods corresponding to rain events are considered. A rain event is defined as a rainy period isolated by at least 30min of dry conditions before and after, and resulting in at least 2mm of cumulated rain height (in average over the area of interest). The dataset is comprised of 17 rain events causing around 250mm of cumulative rain height. An area of interest located in the Vaud Alps has been selected to encompass a network of high-resolution rain gauges whose data are used later for validation (Fig. 1a). The size of the area (60 x 60 km²) has been selected to be large enough to capture prominent features of local rain fields, but at the same time small enough to ensure spatial stationarity (as mentioned above, this work focuses only on temporal non-stationarity).

Figure 1 displays different rain fields observed by the Swiss weather radar network operated by the MeteoSwiss weather agency (Germann et al., 2006). Weather radars are remote sensing devices providing comprehensive images of rain fields at the regional scale (coverage up to about 200km from the radar device), with a high resolution (in the present case, 10 min in time and 1km in space). The resulting rain rate estimates are known to be biased (Berndt et al., 2014), but in counterpart radar images are currently the most reliable and exhaustive source of information about the spatial structure of rain fields and their temporal evolution (Emmanuel et al., 2012; Thorndahl et al., 2017; Marra and Morin, 2018): this is why radar images are used in the current section to illustrate rainfall structure, and in the following to extract rainfall space-time statistics.

A visual inspection of the rain fields displayed in Fig. 1 allows understanding the interest of characterizing jointly the spatial and the temporal behavior of high resolution rain fields. At the scale of interest, all rain fields shown in Fig. 1b-c have in common strong space-time interactions structured by:

- A distribution of intensities that is often skewed (Vrac and Naveau, 2007), with a variable amount of zero values due to within-storm rain intermittency (Schleiss et al., 2011; Mascaro et al., 2013).
- Well defined spatial patterns (Guillot, 1999; Emmanuel et al., 2012; Marra and Morin, 2018), which can be linked to the processes responsible for rainfall production such as rain cells, rain bands or rain storms.

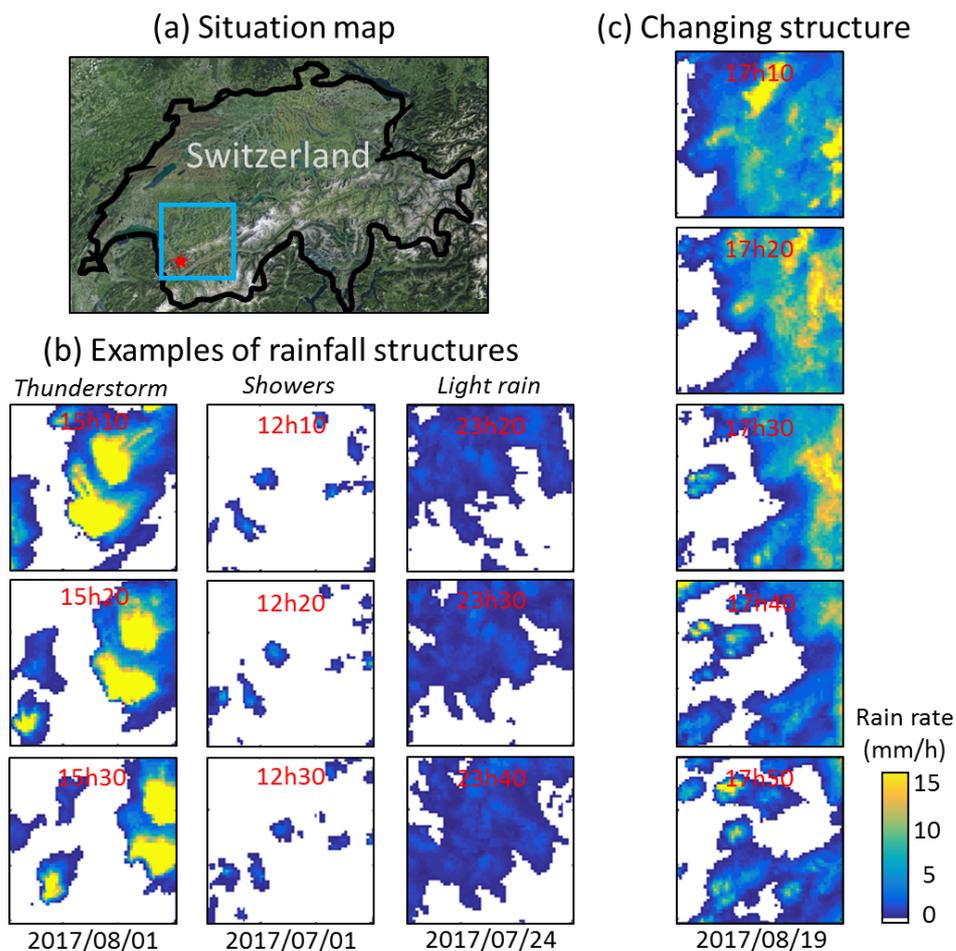


Figure 1. Examples of rain fields over Vaud Alps, Switzerland. (a) Situation map. The area of interest is delineated by the blue square. The red star denotes the location of the rain gauge network used for validation (Sect. 4.3). (b) Examples of rain structures for three rain storms with different space-time behaviors. (c) One example of rain field with a temporally changing structure.

- A temporal behavior shaped by the advection and diffusion of spatial patterns along time (Lepioufle et al., 2012; Creutin et al., 2015).

Despite common space-time behaviors, the three rain events of Fig. 1b look very different. For example, the ‘thunderstorm’ event is characterized by a strong spatial intermittency (i.e. dry and wet locations coexist in the same radar image) combined with well-defined cells generating intense rainfall, while the ‘light rain’ event shows a lower fraction of dry locations, is less spatially structured and generates low rain rates. Hence, it should be possible to find statistical metrics able to distinguish these three rain types based on their spatio-temporal characteristics, as well as transitions between them.

It is worth noting that not only the space-time statistics of rainfall change between rain events, as shown in Fig. 1b, but also



these statistics can change within a single rain event, as illustrated in Fig. 1c. In this case, a widespread and spatially continuous rain field (2 first images) is replaced by disconnected rain cells (2 last images). This change in the space-time features is very rapid and takes place in less than 30 min. Such abrupt changes in the space-time behavior within a rain event are relatively common in our dataset as discussed later.

- 5 Starting from this example, this paper investigates how to detect non-stationarities in rainfall space-time statistics using radar images as primary information.

3 Assessing rain statistics stationarity from radar images

3.1 Extracting space-time information from radar images

To assess the stationarity of rainfall space-time statistics, we propose to start by extracting information on the rainfall space-time behavior from radar images. To this end, 10 statistical metrics are derived for every radar image (Fig. 2), which are split in three categories that reflect the three main characteristics of rain fields identified in Sect. 2:

- Intensity Indices (II), which relate to the probability density function (histogram) of the rain intensities measured in a given radar image. The following indices are used:

- II.1: Fraction of the image covered by rainy pixels (informs the intra-storm rain intermittency).
- 15 – II.2: Mean rain intensity computed over all rainy pixels.
- II.3: Quantile 80% of rain intensities.

- Spatial Indices (SI) that characterize the spatial arrangement of patterns within rain fields. They are selected among the indices proposed by (Aghakouchak et al., 2011) and applied here in the context of rain fields occurring at mid-latitudes under a temperate climate. They are computed based on binary images representing rain masks. Such binary images are obtained by thresholding radar images at a rain intensity of 0.1mm/h, and by assigning a 0 value to the pixels under the threshold and a 1 value otherwise. Then, connected components, hereafter referred to as rain clusters, are identified in every binary image. Their morphological properties are used to derive the following indices:

- SI.1: Fraction of rainy area covered by the largest rain cluster in the image. This is a first indication of how the rain field is split into clusters. Let N_p be the total number of rainy pixels in the binary image and N_m be the number of pixels of the largest cluster; then $SI.1 = \frac{N_m}{N_p}$.
- 25 – SI.2: Connectivity index. It is equal to one if the rain field is fully connected (one single rain cluster) and tends to zero if the rain field is split into many disconnected clusters. Let N_c be the number of rain clusters in the binary image, then: $SI.2 = 1 - \frac{N_c - 1}{\sqrt{N_p + N_c}}$.



- SI.3: Perimeter index, characterizing the sinuosity of the contours of rainy areas. It is equal to 1 if all rain clusters are squares, and tends to 0 if the rain clusters are very sinuous. Let p be the total perimeter of rain clusters, i.e. the sum of the perimeters of all rain clusters, then:

$$SI.3 = \frac{4 \times \sqrt{N_p}}{p} \text{ if } \lfloor \sqrt{N_p} \rfloor = \sqrt{N_p}$$

$$SI.3 = \frac{2 \times (\lfloor 2 \times \sqrt{N_p} \rfloor + 1)}{p} \text{ if } \lfloor \sqrt{N_p} \rfloor \neq \sqrt{N_p}$$

- SI.4: Area index, characterizing the spread of the rain clusters. It is equal to 1 if the radar image contains one single cluster, and tends to zero if the rainy pixels are only in the corners of the image. Let A_{convex} be the area of the convex hull encompassing all the rain clusters, then: $SI.4 = \frac{N_p}{A_{convex}}$.

- Temporal Indices (TI), which characterize the temporal evolution of the rain fields. They assess the advection of rain storms over the ground as well as the temporal deformation of spatial rain patterns. They are obtained by computing the spatial shift that maximizes the correlation between subsequent images, as well as the resulting correlation. Let I_t and I_{t+1} be two subsequent images. In addition, let $r_{i,j}$ be the cross-correlation between I_{t+1} and I_t translated by a vector $\vec{D} = i \cdot \vec{E} + j \cdot \vec{N}$ of coordinates i and j along the Eastward and Northward directions respectively. Finally, let r^{max} be the maximum correlation and $\vec{D}^{max} = i^{max} \cdot \vec{E} + j^{max} \cdot \vec{N}$ the corresponding displacement vector. Then the Temporal Indices (TI) are defined by:

- TI.1: Eastward component of the displacement vector, i.e. i^{max} . This index corresponds to the advection of the rain storm along the West-East direction between times t and $t + 1$.
- TI.2: Northward component of the displacement vector, i.e. j^{max} . This index corresponds to the advection of the rain storm along the South-North direction between times t and $t + 1$.
- TI.3: Correlation coefficient between the two radar images I_t and I_{t+1} after removing advection effects, i.e. r^{max} . This index equals one if the spatial rain patterns remain identical between two subsequent radar images (up to a translation), and tends to zero if the images are completely different.

3.2 Classification of radar images based on rainfall space-time statistics

Next, the 10 indices defined above are used to classify the radar images in order to obtain a limited number of rain types. To ensure the reliability of these indices, only images with a significant proportion of rainy pixels are used for classification. Indeed, if the number of rainy pixels is low, the space indices (SI) are not meaningful and the time indices (TI) cannot be computed because the image correlation procedure fails. We therefore only classify the images with more than 10% rainy pixels.

To define rain types, we adopt an approach based on a Gaussian Mixture Model classifier (GMM) (Fraley and Raftery, 2002). The idea of the GMM is to approximate the joint distribution of the 10 statistical indices. This approximation being a combination of Gaussian functions, those can be considered as representing several discrete categories. In the GMM, the joint

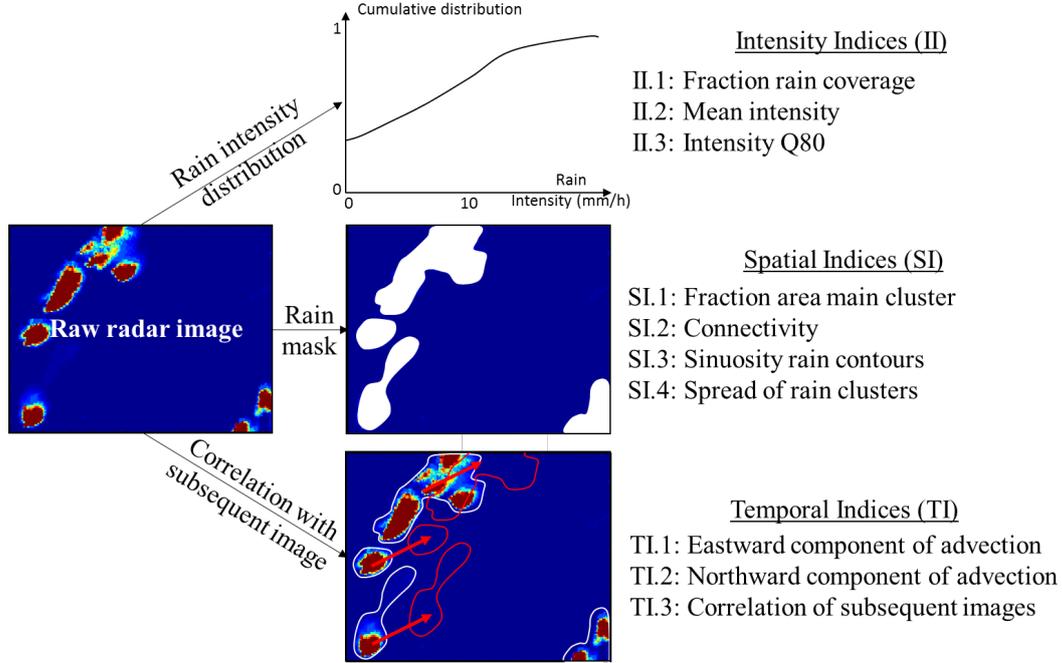


Figure 2. Computation of indices characterizing rainfall space-time statistics for a single radar image. This procedure is repeated for each image with >10% rainy pixels. Note that the temporally subsequent image is required to compute the Time Indices (TI).

distribution $\hat{p}(\mathbf{x})$ of the space-time indices forming a vector $\mathbf{x} \in \mathbb{R}^{10}$ is approximated by a weighted sum of K multivariate Gaussian distributions $\mathcal{N}(\mathbf{x}|\mu_k, \Sigma_k), k = 1, \dots, K$, with respective mean vector μ_k and covariance matrix Σ_k :

$$\hat{p}(\mathbf{x}) = \sum_{k=1}^K \pi_k \times \mathcal{N}(\mathbf{x}|\mu_k, \Sigma_k) \quad (1)$$

The inference of the model parameters (i.e. $\pi_k, \mu_k, \Sigma_k, k = 1, \dots, K$) is performed with an Expectation-Maximization (EM) algorithm (Fraley and Raftery, 2002). A full covariance model is used for the covariance matrices Σ_k . The number K of Gaussian mixtures used in the GMM model is selected by minimization of the BIC criterion derived from EM fits computed for different numbers K (Schwartz, 1978), with the goal of selecting the GMM model resulting in the best fit while maintaining a parsimonious parametrization.

Once fitted, the GMM can be used to derive a probabilistic classification of any vector \mathbf{x} of indices. The probability that a vector belongs to the population G_j whose distribution is the j^{th} mixture component $\mathcal{N}(\mathbf{x}|\mu_j, \Sigma_j)$ is given by:

$$\hat{p}(\mathbf{x} \in G_j) = \frac{\pi_j \times \mathcal{N}(\mathbf{x}|\mu_j, \Sigma_j)}{\sum_{k=1}^K \pi_k \times \mathcal{N}(\mathbf{x}|\mu_k, \Sigma_k)} \quad (2)$$



A classification of the entire image dataset can thus be obtained by assigning to each image I the class that corresponds to the most probable mixture component:

$$G(I) = \max_j (\hat{p}(\mathbf{x} \in G_j)) \quad (3)$$

As mentioned above, the classification procedure can only be applied to radar images with a significant proportion of rainy pixels (10%). Therefore, all images with too few rainy pixels (<10%) remain unclassified. In addition, it can be desired to avoid noisy successions of rain types. To do this, we impose a temporal persistence threshold for the rain types. To this end, all images that lead to temporal clusters of classes that do not reach a certain duration are set to unclassified. Here we use a 60-minute duration threshold. After cleaning the classification, all the time steps that are unclassified (either because less than 10% of the pixels are rainy or because the image belongs to a <60 min cluster) receive the type of the classified image that is temporally the closest (i.e. nearest neighbor interpolation along the time axis). The complete rain typing framework is summarized in Fig. 3.

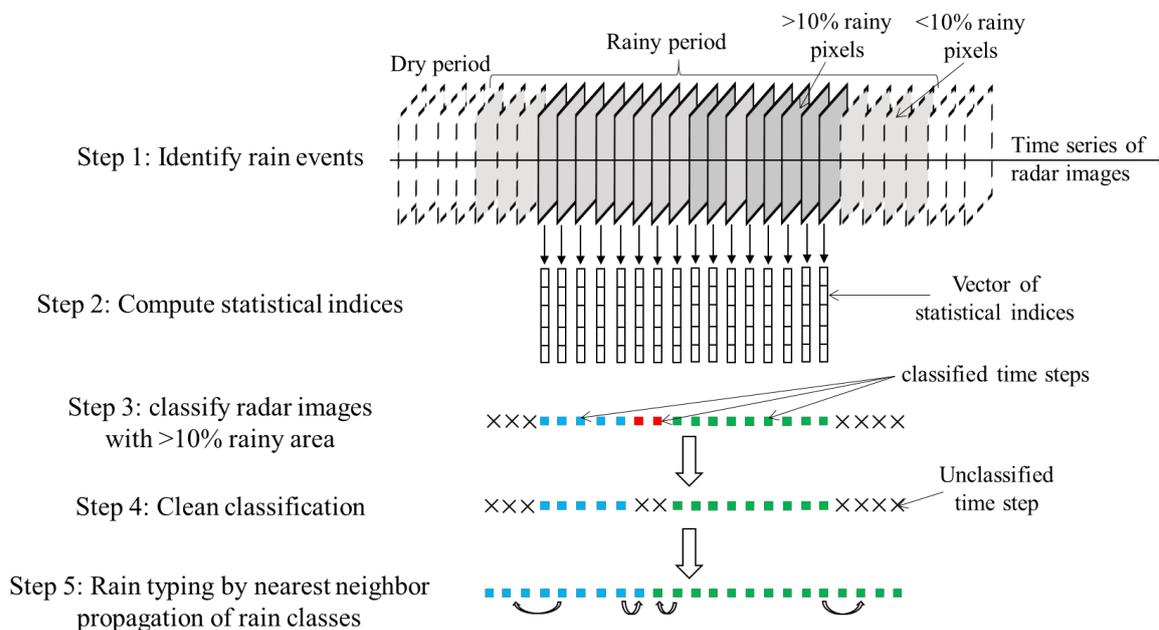


Figure 3. Rain typing framework.

Since the rain typing method presented above aims at defining groups of rain fields sharing similar statistical signatures, the transitions between rain types can be interpreted as non-stationarities. Similarly, periods with a constant rain type are interpreted as stationary periods.



4 Validation and application

In this section, we validate our rain typing approach in the context of stochastic rainfall modelling. The validation study comprises two steps: First, the proposed approach is tested in a synthetic case in order to determine whether we are able to identify known non-stationarities. Then, real data are used to compare our rain typing strategy with two alternative hypotheses of rainfall stationarity: (1) rainfall statistics are stationary at a seasonal scale and (2) rainfall statistics are stationary at a rain storm scale. Prior to the validation itself, the next subsection describes the stochastic rainfall model used for validation.

4.1 Stochastic rainfall model

The validation of the rain typing approach uses a stochastic rainfall model designed for local-scale (a few square kilometer extent) and high-resolution (up to 1 min) data. This model involves 11 parameters and aims at modelling both the marginal distribution of observed rain intensities and the space-time dependencies that exist within rain fields. It is briefly introduced hereafter; for more details the reader is referred to Benoit et al. (In Press). In this model, the marginal distribution of rain rates is accounted for by considering that rain measurements (R) originate from the censoring and power transform (involving parameters a_0, a_1, a_2) of a standardized multivariate Gaussian random field (Z) tainted by an additive measurement noise ($\epsilon \sim \mathcal{N}(0, \sigma_\epsilon)$) (Eq. 4):

$$R = \begin{cases} \left(\frac{Z + \epsilon - a_0}{a_1} \right)^{\frac{1}{a_2}} & \text{if } Z + \epsilon > a_0 \\ R = 0 & \text{if } Z + \epsilon \leq a_0 \end{cases} \quad (4)$$

The multivariate Gaussian latent random field (Z) is characterized by an asymmetric Gneiting space-time covariance (Gneiting, 2002) function ρ which accounts for both the advection and the diffusion of spatial rain patterns (Eq. 5). For two rain observations separated by a spatial lag \mathbf{h} and a temporal lag u , the covariance is given by:

$$\rho(\mathbf{h}, u) = \frac{1}{(u/d)^{2\delta} + 1} \exp \left(\frac{-\left(\|\mathbf{h} + \mathbf{V} \cdot u\| / c \right)^{2\gamma}}{\left((u/d)^{2\delta} + 1 \right)^{\beta\gamma}} \right) \quad (5)$$

In this model, the advection of rain storms is assumed to be constant and linear along a vector \mathbf{V} defined by its amplitude V_S and direction V_θ . The regularity parameters γ (for space) and δ (for time) control the slopes at the origin of the covariance function and thereby regulate the small-scale variability of the rain fields, and ultimately their smoothness. The scale parameters c (for space) and d (for time), in units of distance and time respectively, control the decorrelation distances of rain patterns. Finally, the separability parameter β controls the space-time interactions. When $\beta = 0$, the covariance function is space-time separable.



4.2 Detection of rainfall non-stationarity in a controlled setting

The ability of the rain typing method to detect possible non-stationarities is tested by applying it to synthetic time series of radar images. These images are generated using the stochastic rainfall model presented above, with model parameters changing abruptly. This produces (temporal) non-stationary synthetic rain fields. The rain typing method is then applied to the simulated radar-like images in order to assess if it is able to retrieve the prescribed patterns of temporal non-stationarity. For generating the synthetic images, we use the stochastic rainfall model described in Sect. 4.1 with model parameters corresponding to three typical rain behaviors identified by visual inspection (Table 1).

	a_0	a_1	a_2	σ_ϵ	γ	c	δ	d	β	V_S	V_θ
Type 1	-0.73	0.78	0.41	0.0	0.76	4642	0.89	590	0.97	5.6	18
Type 2	0.0	0.64	0.41	0.0	0.49	6840	0.86	892	0.91	1.2	-11
Type 2	-0.83	1.16	0.45	0.0	0.38	13995	0.81	1702	0.95	0.9	-39

Table 1. Parameters of the stochastic rainfall model used for the generation of synthetic images.

These types correspond to organized convective rains (Type 1), thunderstorm-related heavy showers (Type 2), and stratiform rains (Type 3). Fig. 4a shows an example of simulated rain field for each rain type.

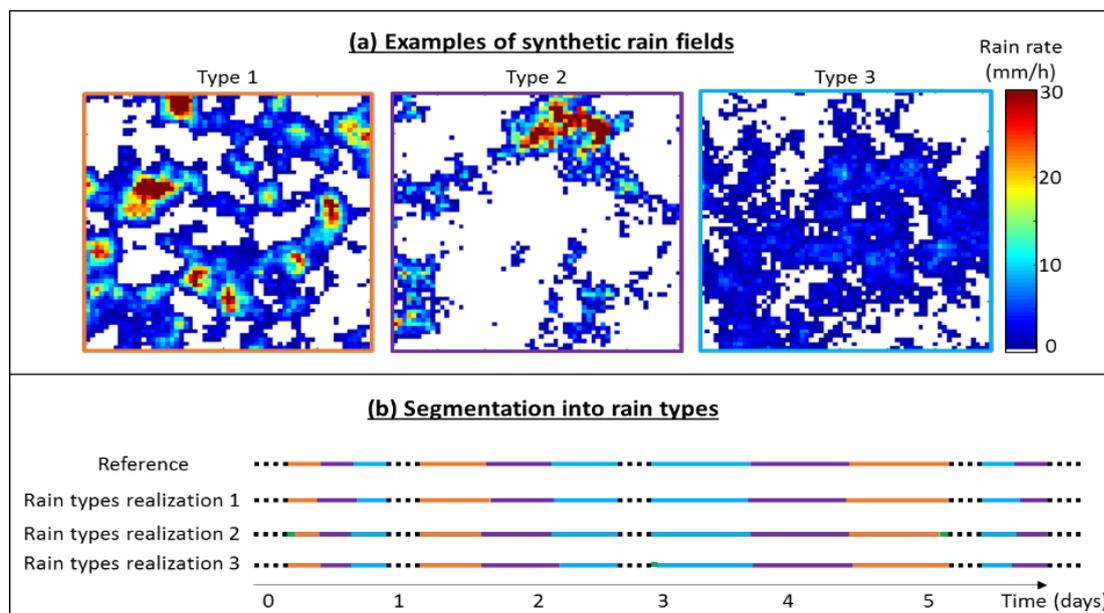


Figure 4. Identification of non-stationarities in rain statistics for a synthetic case study. (a) Examples of synthetic rain fields simulated for each rain type. (b) Segmentation of the time axis into periods with stationary rain statistics. First row: reference. Rows 2-4: Segmentation obtained by rain typing applied to synthetic radar-like images. Dotted black lines represent dry periods.



In Figure 4b, the series titled "reference" shows the rain types prescribed to the stochastic rainfall model for the generation of the synthetic radar images. Based on stochastic simulations, three sets (realizations) of synthetic images are generated. Each realization is classified to determine whether it is possible to identify the reference rain types. Results show that the proposed method can consistently detect the prescribed rain types and their temporal evolution, for all realizations. It also properly estimates the number of rain types prescribed in the reference. The only noticeable difference between the reference and the simulations is the emergence of a very infrequent fourth rain type (accounting for 0%, 1% and 1.6% of the estimated rain types for realizations 1 to 3) at the beginning or at the end of some rain events (in green in Fig. 4b). This is because at these periods, the rain does not cover the whole area of interest, and in certain situations it can produce rain fields with different space-time statistics, which induces this artificial fourth rain type. Except for this infrequent artificial fourth rain type, results show that in this synthetic experiment, the proposed method performs well to detect non-stationarities of rainfall space-time statistics and, in turn, periods during which these statistics remain stationary.

4.3 Impact of rainfall non-stationarity on stochastic modelling of an actual dataset

To further validate our rain typing method, we apply it to a real dataset acquired in the Vaud Alps, Switzerland, during the summer of 2017. In such a real case study, the true succession of rain types is obviously unknown. To assess the performance of the proposed rain typing method, we compare it with two other hypotheses of stationarity that can be found in the literature. We therefore considered three cases, illustrated in Fig. 5:

- Hypothesis H1: the time axis is broken down into rain types interpreted as stationary time periods (the approach proposed in this paper). Applying the rain typing method presented in Sect. 3 to the period of interest leads to 6 rain types.
- Hypothesis H2: the statistical structure of rainfall is constant over meteorological seasons (Paschalis et al., 2013; Bárdossy and Pegram, 2016; Peleg et al., 2017). This lead to one stationary pool of rain events for the period of interest.
- Hypothesis H3: the structure of rainfall is constant within a single rain storm but changes between storms (Casari et al., 2016; Benoit et al., In Press). Here 17 rain events are identified following the definition adopted in Sect. 2.

To compare these three hypotheses, we apply the same stochastic model as above to rain data collected by a dense network of 8 high-resolution rain gauges set up in a small ($3 \times 6 \text{ km}^2$, Fig. 5a) alpine catchment called 'Vallon de Nant' situated within the area of interest presented in Sect. 2. Hence, in the following, radar images will be used only to carry out the rain typing presented in Sect. 3 in order to define the hypothesis H1. The remaining of the validation, i.e. stochastic model calibration and simulation under the three tested hypotheses, will be carried out on rain rate time series acquired by rain gauges, and not on radar images. The goal is indeed to keep the following validation as independent as possible from the radar images used for rain typing. By doing so, we seek to prove that the proposed method captures the stationarity of the rainfall process itself, and not only the stationarity of radar images.

Once the periods of stationarity have been built for each of the three hypotheses, the stochastic model is calibrated for each stationary period. This means that for each hypothesis, a set of model parameters is inferred from observations for each

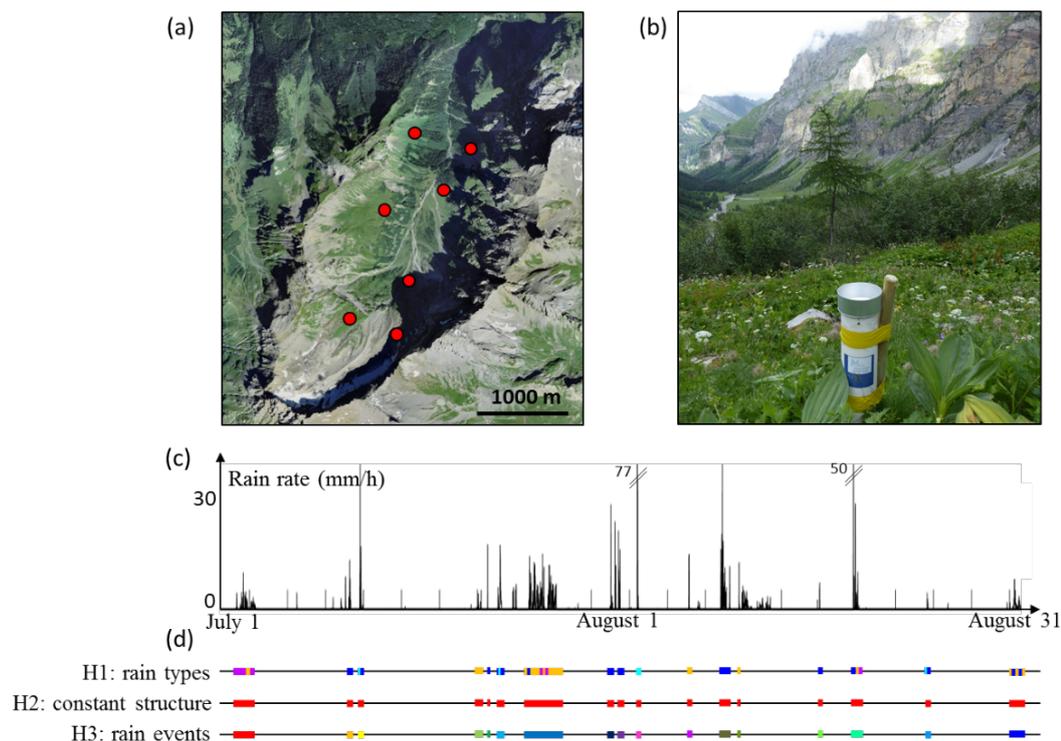


Figure 5. Observation dataset used for validation. (a) Measurement network. Red dots denote rain gauge locations. (b) Picture of a rain gauge with the Vallon de Nant catchment in the background. (c) Observed rain rate averaged over the network. (d) Segmentation of the time axis into stationary periods for the three tested hypotheses. For each hypothesis, segments with the same color denote periods for which rainfall is expected to have similar space-time statistics.

postulated stationary dataset. Then, synthetic rain fields are generated by unconditional simulation under the three hypotheses of stationarity, and in each case 10 realizations (i.e. 10 simulated synthetic rain histories) are compared to actual measurements. To assess the realism of the different scenarios, Fig. 6 shows the simulated cumulative rain heights. Next, Fig. 7 shows quantile-quantile (q-q) plots for four statistics selected to assess the marginal distribution of rain rates and its space-time arrangement:

5 number of rain gauges measuring zero rain at each time step, rain intensity, standard deviation (in time) of rain intensities separated by a time lag of 5min, and standard deviation (in space) of rain intensities at each time step.

Results show that H1 tends to slightly underestimate the cumulative rain due to an underestimation of very high intensities. This underestimation is common to all the three cases, and probably originates from the stochastic model itself, which is not designed to handle extreme rainfalls due to the simple transform function selected in Eq. 4. This could be improved by adopting

10 a transform function accounting for extreme rainfalls (see e.g. (Vrac and Naveau, 2007)) but at the price of a more complex parametrization which is not regarded as essential here because the observed rain rates are mostly low to moderate, and only the 99th centile is affected by rain rate underestimation. Apart for this underestimation of high rain rates, hypothesis H1 allows

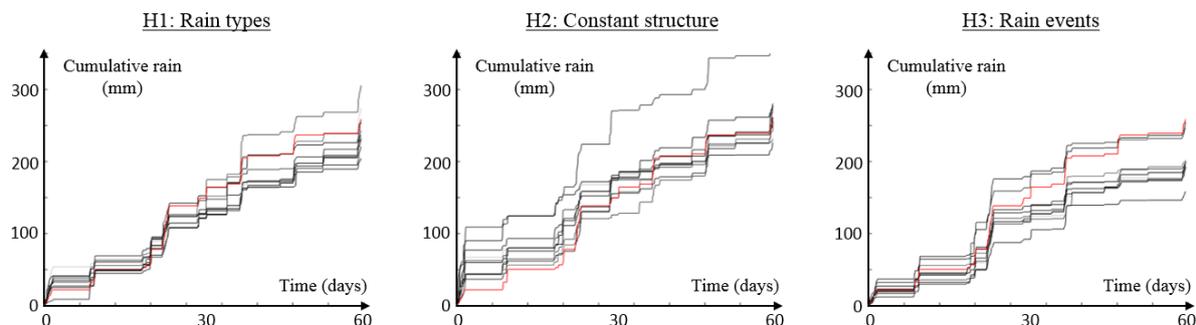


Figure 6. Reproduction of the cumulative rain height (averaged over the whole network) for the three tested hypotheses. Left: rain is stationary by rain type (H1), Center: rain is stationary during the entire summer period (H2), Right: rain is stationary by rain event (H3). The red line corresponds to the observations and the grey lines correspond to simulations.

correctly reproducing the other metrics.

On the contrary to H1, hypothesis H2 leads to a slight overestimation of the simulated rain height, in particular for the first 30 days (Fig. 6). This is due to the overestimation of moderate rain rates that compensates for the underestimation of extremely high values. This bias in the simulated marginal distribution is due to the lack of flexibility of H2 that imposes a single underlying stochastic model for the entire summer period. This does not allow enough flexibility to capture the diversity of structures emerging from high-resolution data. This is also visible for the simulated variability in space and time, which is overestimated for the low centiles but underestimated for the high centiles.

Under hypothesis H3, simulation results are close to the ones of H1, with a slightly higher dispersion and a stronger underestimation of the cumulative rain. This can be explained by the large number of model parameter sets considered in this case. This allows simulating very diverse rainfalls, but in counterpart leads to a high dispersion of the results. In addition, note that in case of short rain events, the hypothesis H3 may result in a poor inference of model parameters due to the low amount of observations available for each stationary period.

To sum up, the proposed method consisting in typing rain fields according to their space-time statistical signature derived from radar images (H1) leads to more realistic rainfall simulations than the other approaches H2 and H3. This result, added to the proper identification of stationary rain periods in a synthetic case shown in Sect. 4.2, indicates that the rain typing method proposed in this paper offers an acceptable identification of non-stationarities in rainfall space-time statistics.

Discussion and conclusion

This paper proposes a quantitative method to identify stationary rainfall periods, that is, periods during which a set of 10 statistics representative of rainfall space-time behavior remains broadly constant. It is based on a classification of radar images into groups of rain fields sharing a similar statistical structure when observed at high resolution. For reasons of data availability,

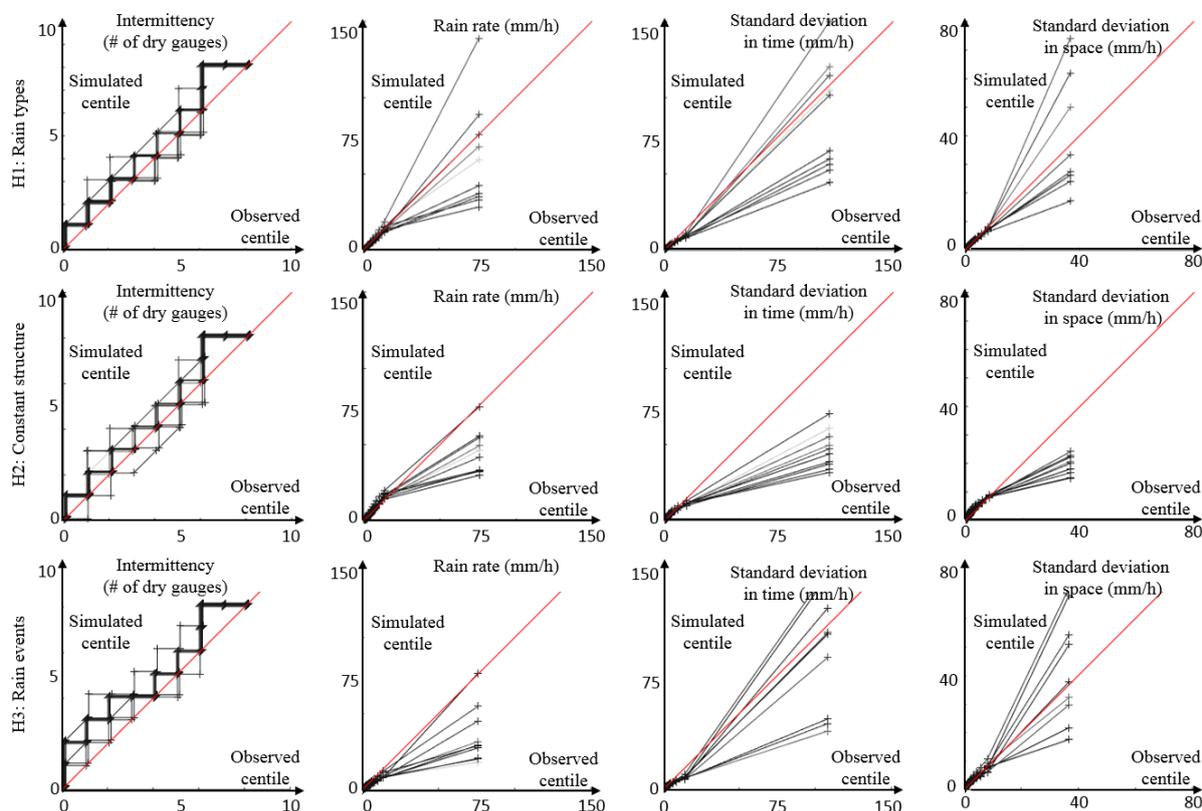


Figure 7. Reproduction of rainfall statistics for the three tested hypotheses. Row 1: rain is stationary by rain type (H1), Row 2: rain is stationary during the entire summer period (H2), Row 3: rain is stationary by rain event (H3). Column 1: quantile-quantile (q-q) plot of simulated vs observed rain intermittency, Column 2: q-q plot of simulated vs observed rain rate, Column 3 q-q plot of simulated vs observed temporal variability of rainfall (at lag 5 min), Column 4: q-q plot of simulated vs observed spatial variability of rainfall. The quantiles used in the q-q plots are centiles. Each centile is denoted by a black cross.

we focused our investigation on summer rains over the Vaud Alps, Switzerland. However most of the results obtained in this context are expected to be transferable to other seasons, as well as to other mid-latitude areas in cases where extratropical rain storms significantly influence the precipitation regime.

- 5 The succession of rain types identified for the dataset used in the present study has two important implications on how stationarity should be regarded in sub-daily stochastic rainfall models. Firstly, a striking observation is that rainfall statistics can change drastically within a single rain event. As a result, the hypothesis of rain stationarity along entire rain events can be invalidated in some instances. At least in our dataset, such non-stationary events seem relatively frequent (7 non-stationary rain events out of 17 have been identified in our data). This observation is not new, since it may lead to temporal asymmetry (or temporal
- 10 irreversibility) in rain rate time series, which is discussed by Müller et al. (2017). Our framework offers a way to deal with this



phenomenon in the context of stochastic rainfall modelling. The proposed solution consists in the identification of stationary periods prior to the stochastic modelling of rain. Then, stochastic modelling is carried out separately for each rain type, and the results are merged afterwards. This allows generating synthetic rain fields presenting a temporal asymmetry, even if the stochastic model itself is only capable of generating symmetric rain fields for a given set of model parameters. The temporal asymmetry is then carried by the temporal arrangement of rain types within a single rain storm. A second consequence of the observed variability in rainfall statistics is that postulating rain stationarity over long periods, such as months or seasons, is likely to impair the characterization of rainfall statistics. This occurs because the model parameters are inferred from a dataset that mixes together different rainfall structures. This results in improper space-time dependencies in the simulated rain fields.

5

10 A possible future work would be to apply the proposed rain typing method to long term archives of radar images in order to investigate the temporal behavior of rain statistics throughout the course of the year as well as their inter-annual variability. Resulting information could be used as the starting point for the design of a statistical model of rain type occurrence, and in turn a stochastic rain type generator. Coupled with already existing high-resolution stochastic rainfall models, this would allow designing high-resolution stochastic rainfall generators able to reproduce local rainfall statistics over long simulation periods.

15

It could also be interesting to use the proposed method to type simultaneously precipitation fields observed by radars and simulated by numerical models. This may provide a new metric to assess the precipitation component of high-resolution numerical weather or climate models. Indeed, the proper reproduction of rain types and rain type successions in model outputs would indicate a correct simulation of the overall space-time behavior of rainfall by the model.

20

Finally, it would be of interest to compare rain type histories derived from radar observations with several meteorological variables derived from climate model reanalyzes. This would allow to link the occurrence of rain types with the state of the atmosphere characterized by a set of meteorological variables. If strong interconnections are found between rain types and climatic conditions, then the future evolution of climate variables simulated by General Circulation Models (GCMs) could

25 provide precious insights on the possible evolution of rain type occurrence in a changing climate.

Competing interests. The authors declare that they have no conflict of interests.

Acknowledgements.



References

- Aghakouchak, A., Nasrollahi, N., Li, J., Imam, J., and Sorooshian, S.: Geometrical Characterization of Precipitation Patterns. *Journal of hydrometeorology*, 12, 274-285, doi:10.1175/2010JHM1298.1, 2011.
- Allcroft, D. J., Glasbey, C. A.: A latent Gaussian Markov random-field model for spatiotemporal rainfall disaggregation, *Applied Statistics*, 52, 487-498, doi:10.1111/1467-9876.00419, 2003.
- Bárdossy, A., Pegram, G. G. S.: Space-time conditional disaggregation of precipitation at high resolution via simulation, *Water Resources Research*, 52, 920-937, doi:10.1002/2015WR018037, 2016.
- Bauer, P., Thorpe, A., Brunet, G.: The quiet revolution of numerical weather prediction, *Nature*, 525, 47-55, doi:10.1038/nature14956, 2015
- Benoit, L., Allard, D., Mariethoz, G.: Stochastic Rainfall Modelling at Sub-Kilometer Scale, *Water Resources Research*, doi:10.1029/2018WR022817, In Press.
- Berndt, C., Rabiei, E., Haberlandt, U.: Geostatistical merging of rain gauge and radar data for high temporal resolutions and various station density scenarios, *Journal of Hydrology*, 508, 88-101, doi:10.1016/j.jhydrol.2013.10.028, 2014.
- Biggerstaff, M. I., Listmaa, S. A.: An Improved Scheme for Convective/Stratiform Echo Classification Using Radar Reflectivity, *Journal of applied meteorology*, 39, 2129-2150, doi:10.1175/1520-0450(2001)040<2129:AISFCS>2.0.CO;2, 2000.
- Bony, S., Stevens, B., Frierson, D., Jakob, C., Kageyama, M., Pincus, R., Shepherd, T. G., Sherwood, S., Siebesma, A., Sobel, A. H., Watanabe, M., Webb, M.: Clouds, circulation and climate sensitivity, *Nature Geoscience*, 8, 261-268, doi:10.1038/ngeo2398, 2015.
- Caseri, A., Javelle, P., Ramos, M. H., Leblois, E.: Generating precipitation ensembles for flood alert and risk management, *Journal of Flood Risk Management*, 9, 402-415, doi:10.1111/jfr3.12203, 2016.
- Creutin, J. D., Leblois, E., Lepioufle, J. M.: Unfreezing Taylor's hypothesis for precipitation, *Journal of Hydrometeorology*, 16, 2443-2462, doi:10.1175/JHM-D-14-0120.1, 2015.
- Emmanuel, I., Andrieu, H., Leblois, E., Flahaut, B.: Temporal and spatial variability of rainfall at the urban hydrological scale, *Journal of Hydrology*, 430-431, 162-172, doi:10.1016/j.jhydrol.2012.02.013, 2012.
- Fraley, C., Raftery, A. E.: Model-Based Clustering, Discriminant Analysis, and Density Estimation, *Journal of the American Statistical Association*, 97, 611-631, doi:10.1198/016214502760047131, 2002.
- Germann, U., Galli, G., Boscacci, M., Bolliger, M.: Radar precipitation measurement in a mountainous region, *Quarterly journal of the royal meteorology society*, 132, 1669-1692, 2006.
- Gneiting, T.: Nonseparable, Stationary Covariance Functions for Space-Time Data, *Journal of American Statistical Association*, 97, 590-600, doi:10.1198/016214502760047113, 2002.
- Guillot, G.: 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, doi:doi.org/10.1007/s004770050034, 1999.
- Journel, A. G.: *Geostatistics: Roadblocks and Challenges*, *Geostatistics Tróia '92, Quantitative Geology and Geostatistics*, Tróia, Springer, 1993.
- Lagrange, M., Andrieu, H., Emmanuel, I., Busquets, G., Loubrié, S.: Classification of rainfall radar images using the scattering transform, *Journal of Hydrology*, 556, 972-979, doi:10.1016/j.jhydrol.2016.06.063, 2018.
- Leblois, E.: *Le bassin versant, système spatialement structuré et soumis au climat*, HDR, University of Grenoble, 2012.
- Leblois, E., Creutin, J. D.: Space-time simulation of intermittent rainfall with prescribed advection field: Adaptation of the turning band method, *Water Resources Research*, 49, 3375-3387, doi: 10.1002/wrcr.20190, 2013.



- Lepioufle, J. M., Leblois, E., Creutin, J. D.: Variography of rainfall accumulation in presence of advection, *Journal of Hydrology*, 464-465, 494-504, doi: 10.1016/j.jhydrol.2012.07.041, 2012.
- Llasat, M. C.: An objective classification of rainfall events on the basis of their convective features: application to rainfall intensity in the northeast of Spain, *International Journal of Climatology*, 21, 1385-1400, doi:10.1002/joc.692, 2001.
- 5 Marra, F., Morin, E. (2018): Autocorrelation structure of convective rainfall in semiarid-arid climate derived from high-resolution X-Band radar estimates, *Atmospheric Research*, 200, 126-138, doi:10.1016/j.atmosres.2017.09.020, 2018.
- Mascaro, G., Deidda, R., Hellies, M. (2013): On the nature of rainfall intermittency as revealed by different metrics and sampling approaches, *Hydrology and Earth System Sciences*, 17, 355-369, doi:10.5194/hess-17-355-2013, 2013.
- Mavromatis, T., Hansen, J. W.: Interannual variability characteristics and simulated crop response of four stochastic weather generators, *Agricultural and Forest Meteorology*, 109, 283-296, doi:10.1016/S0168-1923(01)00272-6, 2001.
- 10 Müller, T., Schütze, M., Bárdossy, A.: Temporal asymmetry in precipitation time series and its influence on flow simulations in combined sewer systems, *Advances in Water Resources*, 107, 56-64, doi:10.1016/j.advwatres.2017.06.010, 2017.
- Paschalis, A., Molnar, P., Fatichi, S., Burlando, P.: A stochastic model for high-resolution space-time precipitation simulation., *Water Resources Research*, 49, doi:10.1002/2013WR014437, 2013.
- 15 Paschalis, A., Fatichi, S., Molnar, P., Burlando, P.: On the effects of small scale space-time variability of rainfall on basin flood response, *Journal of Hydrology*, 514, 313-327, doi:10.1016/j.jhydrol.2014.04.014, 2014.
- Peleg, N., Fatichi, S., Paschalis, A., Molnar, P., Burlando, P.: An advanced stochastic weather generator for simulating 2-D high-resolution climate variables, *Journal of Advances in Modeling Earth Systems*, 9, 1595-1627, doi:10.1002/2016MS000854, 2017.
- Qian, B., De Jong, R., Yang, J., Wang, H., Gameda, S.: Comparing simulated crop yields with observed and synthetic weather data, *Agricultural and Forest Meteorology*, 151, 1781-1791, doi:10.1016/j.agrformet.2011.07.016, 2011.
- 20 Ramirez-Cobo, P., Lee, K. S., Molini, A., Porporato, A., Katul, G., Vidakovic, B.: wavelet-based spectral method for extracting self-similarity measures in time-varying two-dimensional rainfall maps, *Journal of Time Series Analysis*, 32, 351-363, doi:10.1111/j.1467-9892.2011.00731.x, 2010.
- Rosenfeld, D., Amitai, E.: Classification of Rain Regimes by the Three-Dimensional Properties of Reflectivity Fields, *Journal of Applied Meteorology*, 34, 198-211, doi:10.1175/1520-0450(1995)034<0198:CORRBT>2.0.CO;2, 1995.
- 25 Schleiss, M., Jaffrain, J., Berne, A.: Statistical analysis of rainfall intermittency at small spatial and temporal scales, *Geophysical Research Letters*, 38, doi:10.1029/2011GL049000, 2011.
- Schwartz, G.: Estimating the dimension of a model, *The Annals of Statistics*, 6, 461-464, 1978.
- Thorndahl, S., Einfalt, T., Willems, P., Nielsen, J.E., ten Veldhuis, M-C., Arnbjerg-Nielsen, K., Rasmussen, M.R., Molnar, P.: Weather radar rainfall data in urban hydrology, *Hydrology and Earth System Sciences*, 21, 1359-1380, doi:10.5194/hess-21-1359-2017, 2017.
- 30 Vaittinada Ayar, P., Vrac, M., Bastin, S., Carreau, J., Déqué, M., Gallardo, C.: Intercomparison of statistical and dynamical downscaling models under the EURO- and MED-CORDEX initiative framework: present climate evaluations, *Climate Dynamics*, 46, 1301-1329, doi:10.1007/s00382-015-2647-5, 2016.
- Vrac, M., Naveau, P.: Stochastic downscaling of precipitation: From dry events to heavy rainfalls, *Water Resources Research*, 43, doi:10.1029/2006WR005308, 2007.
- 35 Vrac, M., Stein, M., Hayhoe, K.: Statistical downscaling of precipitation through nonhomogeneous stochastic weather typing, *Climate Research*, 34, 169-184, doi:10.3354/cr00696, 2007.



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

Zick, S. E., Matyas, C. J.: A Shape Metric Methodology for Studying the Evolving Geometries of Synoptic-Scale Precipitation Patterns in Tropical Cyclones, *Annals of the American Association of Geographers*, 106, 1217-1235, doi:10.1080/24694452.2016.1206460, 2016.