STORM v.2: A simple, stochastic decision-support tool for exploring the impacts of climate, and climate change at, and near the land surface in small watersheds

Manuel F. Rios Gaona[1], **Katerina Michaelides**[4,3,5], **and Michael Bliss Singer**[1,2,3]

[1]School of Earth and Environmental Sciences, Cardiff University, CF10 3AT Cardiff, Wales - UK
[2]Water Research Institute, Cardiff University, CF10 3AX Cardiff, Wales - UK
[3]Earth Research Institute, University of California Santa Barbara, Santa Barbara, CA 93106-3060, US
[4]School of Geographical Sciences, University of Bristol, BS8 1QU Bristol, England - UK
[5]Cabot Institute for the Environment, University of Bristol, BS8 1QU Bristol, England - UK

Correspondence: manuel (RiosGaonaM@cardiff.ac.uk)

Abstract. Climate change is projected to have major impacts to land surface and subsurface processes through its expression in the hydrological cycle, but the impacts to any particular basin or region are highly uncertain. Non-stationarities in the frequency, magnitude, duration, and timing of floods and droughts would have important implications for human societies and ecosystems. The conventional approach for assessing the near-surface impacts of climate change is to downscale global climate model output and use it to drive regional and local models that express the climate within hydrology near the land surface. While this approach may be useful for linking general circulation models to the hydrological cycle, it is limited for examining the details of hydrological response to climate forcing for a specific location over timescales relevant to decision makers. For example, management of flood hazard or drought amelioration requires detailed information that includes uncertainty based on variability in storm characteristics, rather than on differences between models within an ensemble. To fill this gap, we present the second version of our STOchastic Rainfall Model (STORM), an open-source, parsimonious and user-friendly modeling framework for simulating climatic expression as rainfall fields over a catchment. This work showcases the use of STORM in simulating ensembles of realistic sequences, and spatial patterns of rainstorms for current climate conditions, and bespoke climate change scenarios that affect the water balance near the Earth's surface. We outline, and detail STORM's new approaches such as: a copula for linking marginal distributions of storm intensity and duration; an orographic stratification (in which intensity-duration copulas can be applied too); a radial decay-rate which takes into consideration potential, but unrecorded, maximum storm intensities; an optional component to simulate storm starting date-times via circular/directional statistics; and a compressed implementation in modelling future climate scenarios. We also introduce a new ingestion module that facilitates the generation of relevant input in the form of probability density functions (PDFs), from historical data, for stochastic sampling. Independent validation exercises showed that the average performance of STORM falls within a 5.5% from the average of all storms in the Walnut Gulch (Arizona, US) ocurred in the current century.

1 Introduction

In earlier research (Singer and Michaelides, 2017; Singer et al., 2018), we introduced the STOchastic Rainstorm Model (STORM)[1], presented the justification for its creation, and demonstrated its application to simulating spatial rainfall fields at Walnut Gulch, Arizona (see Sec. 2.10). In this paper, we introduce STORM v.2 and highlight the novel aspects of the model that warrant a new version number. We made several changes to the model that make it more user-friendly and enhance its capability for simulating water balance over small watersheds under historical climate or un-

[1]https://github.com/blissville71/STORM

der various user-defined scenarios of climate change. Specifically, STORM v.2: a) treats rainstorm intensity and duration as joint variables in copula framework, rather than as independent variables, a major shortcoming in the previous version of the model; b) offers an altitude stratification to account for orographic characteristics influencing precipitation c) improves on the radial decay-rate model to incorporate potential, but unrecorded, maximum storm intensities; d) tackles modelling of storm's starting date-times from directional-statistics perspective; and e) contains a pre-processing module that automatically generates all the input probability density functions (PDFs) required for storm. These advances, which will be discussed in detail below, were required to create a model that is faithful to the underlying hydrological processes (e.g., capturing relationships between rainstorm intensity, duration, and frequency), while also enabling broad uptake and easy use of the model for a range of purposes, and for any small basin with available storm rainfall data.

An individual rainstorm (discrete in space and time) has an intensity that varies spatially from the center of the storm (the so-called storm core) to its margins, and a duration over which an average intensity is expressed. Rainstorm intensity and duration are related in the sense that the highest intensity storms are generally short-lived, while long rainstorms have low average intensity. The functional form of the relationship between rainfall intensity and duration is typically characterized as a negative exponential, where intensity declines with duration (Nicholson, 2011). However, in rain gauge data, there can be dramatic scatter in this relationship, so a single-valued function cannot represent the phase space between intensity and duration. To overcome this limitation, the previous version of STORM fitted the relationship for the upper envelope of the intensity-duration phase space and then used the functional form of the fitted curve to fit additional curves that pass through the entire phase space (Singer et al., 2018). These intensity-duration curves are then treated as a stochastic variable for random selection within the STORM code. To further enable complete sampling of the entire phase space, STORM 1.0 also includes a fuzzy tolerance such that storm intensity for the selected duration can vary up or down away from the selected curve.

This representation of intensity and duration is the crux of STORM 1.0, as it forms the basis for rainstorm characteristics that affect rainfall totals during a storm, over a season, and over the longer term. However, this approach has several weaknesses: a) it is based on debatable, heuristic rules of probability assignation; b) it does not capture the inherent multi-valued relationship between rainfall intensity and duration; c) the functional form of the relationship is assumed based on the upper envelope of the phase space; and d) there is an arbitrary number of curves used to represent the phase space. Notably, we also use the curve number probabilities to represent orography in STORM 1.0. This means that the STORM 1.0's representation of orography contains these same weaknesses.

The relationship between rainfall intensity and duration is a critical attribute of rainstorms that affects the overall delivery of water to the land surface, the balance between infiltration and evapotranspiration, and the corresponding antecedent moisture condition at any point in time and space. Thus, it is critical to characterize the distribution of storm intensity-duration from historical records, as well as the frequency of their occurrence. STORM v.2 now offers a better characterization of storm intensity-duration relationship throughout a copula approach.

Copulas (or *copulae*), from the Latin word for "tie", represent a way forward for characterizing the complex relationship between intensity and duration from the perspective of joint frequency of occurrence (Vandenberghe et al., 2011). A copula is a function that links/couples a multi-variate distribution function to its univariate marginals, regardless any prior knowledge of such marginals (see Sec. 2.5). The copula approach obviates the need for fitting intensity-duration curves, and for the arbitrary assignment of curve probabilities. Once the intensity-duration copula is fit, it can be sampled randomly to simulate the rainstorm characteristics.

Another shortcoming in STORM 1.0 was its reliance on user-developed PDFs as input to the model. We recognize that this requirement may be a major limitation which prevents some users from deciding to use STORM for rainstorm simulation. To make STORM more user-friendly, we added the pre-processing, and visualization modules that respectively allow the automatization in computing the best fit of PDFs on (input) gauge data, and the visualization of the (output-modelled) storms (see Sec. 2.8).

We provide STORM v.2 (and its ingestion, and visualization modules, along with toy/processed input data, and parameters) as open source code[2]. Unlike STORM 1.0, STORM v.2 is uniquely, and entirely written in Python 3 (Van Rossum and Drake, 2009). From here onwards, we will refer to STORM v.2 simply as STORM.

2 Data and Methods

STORM is a stochastic model built upon continuous PDFs for seven variables, i.e., total seasonal rainfall (TOTALP), maximum storm radius/extent (RADIUS), rainfall decay rate from the storm's center outwards (BETPAR), maximum intensity (MAXINT), average duration (AVGDUR), storm's starting date (DOYEAR), and storm's starting time (DATIME). Here we model the relation between the storm's maximum intensity and its average duration via a copula approach (COPULA). STORM allows the stratification of the copula approach based on the orography of the region, that is, one can specify a maximum_intensity–average_duration copulas model for every altitude band in which the catchment is split. This "altitude stratification", along with the storm's starting time are optional features in STORM. In the case

[2]https://github.com/feliperiosg/STORM2

of the former a digital elevation model (DEM) is required; whereas in the case of the latter the user might encounter some difficulties installing the circular statistics libraries.

STORM also preserves STORM 1.0's functionality to simulate the impact of climate change either on the total seasonal rainfall or the storm's maximum intensity. Such functionality is applied through two types of mutually excluding factors: _SC (i.e., Step-Change) which is constantly applied to every and each of the simulated years; and _SF (i.e., Scaling-Factor) which is increasingly/decreasingly applied to all of the simulated years. Hence, for potential/climate impacts on the total seasonal rainfall these factors are dubbed as PTOT_SC, and PTOT_SF; whereas for potential/climate impacts on the maximum rainfall intensity these factors are dubbed as STORMINESS_SC, and STORMINESS_SF (see Sec. 2.7).

2.1 Total Seasonal Rainfall [TOTALP]

STORM stops a given simulation once the median of the cumulative rainfall over the catchment surpasses the sampled TOTALP value for the season under consideration. The sampled TOTALP value comes from a PDF of historical medians of total seasonal storm rainfall. Each of these historical medians represents the spatial median of the cumulative seasonal rainfall recorded by the gauge network spread within the catchment. To avoid sampling negative values of rainfall, the fitting (and the sampling) of the PDF is done in the (natural) logarithmic space, i.e., $\mathrm{TOTALP} = e^{\mathrm{TOTALP}_{(\mathrm{sampled})}}$. STORM 1.0 used too as stopping criteria the TOTALP median. Nevertheless, that sampled value came from a PDF of historical means (or was it maxima?) of total seasonal rainfall. Now, we consider that reaching the (catchment) median sampled from a distribution of historical medians offers a more accurate picture in stochastic modelling of seasonal totals. Figure 1, panel b, shows the spatial distribution of rainfall at the end of one simulation exercise, i.e., once the median of the cumulative rainfall over the catchment is larger than the sampled value for TOTALP.

2.2 Maximum Storm Extent [RADIUS]

Storm radii are defined in STORM as the maximum distance computed for a group a gauges and their centroid. Here, a "group of gauges" means all those gauges for which the time-stamp of any storm's starting time is identical among them. A PDF of radii was computed from groups with at least two rain gauges. We are aware that this assumption does not consider the extent, evolution, and/or trajectory of any storm in particular throughout the gauge data. Nevertheless, by assuming that identical time-stamps (in storm starting times) might imply that the whole storm is being simultaneously captured by the gauge network, one can easily estimate an extension of the storm from plain gauge records. This premise also relies in the assumption of a circular-shape model for storm cells,

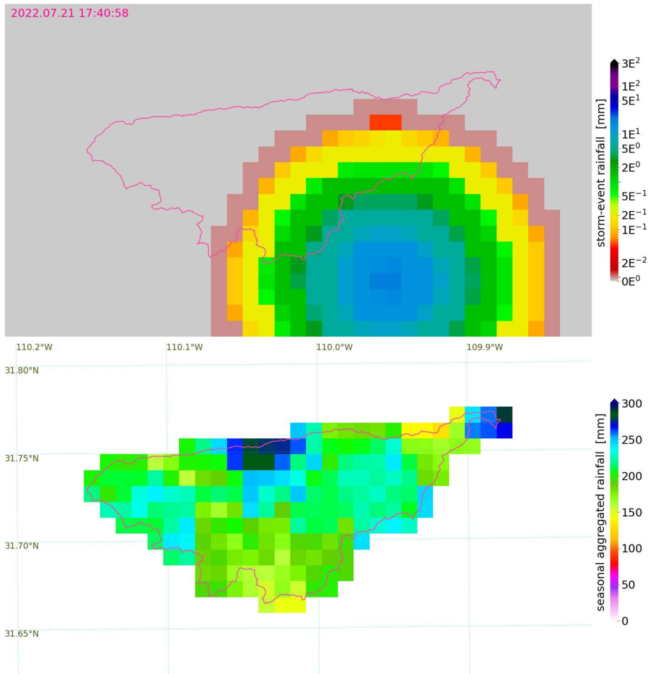

Figure 1. Spatio-temporal distribution of simulated storm rainfall over the Walnut Gulch catchment (see Sec. 2.10). Spatial resolution of $1 \times 1\,\mathrm{km}$. Panel **a** - One large simulated storm starting at $\sim$ 17:41 on July 21st, with a radius of $\sim 11\,\mathrm{km}$, $\sim 2.5\,\mathrm{h}$ of duration, and a maximum intensity of $\sim 19\,\mathrm{mm}$ (i.e., $7.57\,\mathrm{mm \cdot h^{-1}}$). Please note its logarithmic color scale. Panel **b** - Cumulative seasonal distribution of 116 storms for the wet season, i.e., from June through October. Even though the grid is presented in "lat-lon" coordinates (i.e., CRS WGS-84), the actual projection (in both panels) is the 2D-Cartesian coordinate system known as NAD83 / UTM zone 12N (i.e., EPSG:26912; https://epsg.io/26912).

which might not be entirely true, and how reliable might the gauge network be with regard to its spatial density. Figure 1, panel a, shows a simulated storm with a radius of $\sim 11\,\mathrm{km}$. This approach is biased towards spatially-large storms given that small-radii storms, i.e., storms not captured by a single gauge are disregarded in this methodology.

The minimum radius that can be sampled is restricted by the spatial resolution the user might set up the model output to. For instance, for a model output's resolution of $0.5 \times 1.0\,\mathrm{km}$, the minimum possible (sampled) radius would be $1\,\mathrm{km}$. This is achieved by truncating the RADIUS PDF, and then sampling from it. Instead of using a "maximum" criterion for the selection of storm radii, the user can also modify this criterion to be, e.g., the mean (or median or whatever) distance of a group of gauges and their centroid. This change can be implemented by the user, via the `pre_processing.py` script. STORM 1.0 did not use a "radius" approach. Instead, storm area values were sampled from a pre-fixed (GEV?) PDF.

2.3 Rainfall Decay Rate [BETPAR] & Maximum Storm Intensity [MAXINT]

We model individual storms as isotropic circular cells for which maximum intensities (I_{max}) are (always) located at their centres, with a quadratic exponential decay (β^2) as the distance from such centres (r) increases:

$$I(r) = I_{max} \cdot e^{-2 \cdot \beta^2 \cdot r^2}, \tag{1}$$

where $I(r)$ (in $\mathrm{mm} \cdot \mathrm{h}^{-1}$) is the rainfall intensity at a distance r (in km) from the storm centre. β has units of km^{-1}.

We use the quadratic exponential decay model to fit both the decay rate (β), and maximum intensity (I_{max}). This is done via *scipy*'s module `curve_fit`, i.e., a non-linear least squares approach, for which the Trust Region Reflective method is applied, given the constraints we enforce to our minimization problem (Virtanen et al., 2020; Branch et al., 1999). Such constraints/bounds simply refer to the limits for which one intends β and I_{max} (in this case) to be within. For instance, and following Eagleson et al. (1987, Fig. 17), we bound β between 0 and 3; whereas for I_{max} we set 3 times the highest intensity found in the gauge data as the upper limit, and some value slightly above zero as the lower limit ($0.07\,\mathrm{mm} \cdot \mathrm{h}^{-1}$, in our case). Eagleson et al. (1987), and Morin et al. (2005) previously used, for the Walnut Gulch catchment, the same model to only fit the rainfall's decay rate respectively from gauge and radar data. Figure 1, panel a, shows a simulated storm with a steep β of $\sim 0.18\,\mathrm{km}^{-1}$, and $I_{max} = 18.77\,\mathrm{mm}$.

We fit the model for storms simultaneously registered by four or more gauges (i.e., with identical starting time-stamps). Along with the optimal values for which the model is fitted, `curve_fit` also returns the estimated covariance of such optimal values. We only kept optimal values for which their covariance is equal or smaller than 5, and equal or larger than 0. These "clean" optimal values are the ones over which the PDFs (BETPAR and MAXINT) are then constructed upon. Supplementary Figure B6 shows three cases for which the model represented by Eq. (1) offers a got fit/approximation. Now i'm reluctant to use this figurE We obtained similar results (not shown here) to Eagleson et al. (1987), and Morin et al. (2005) for the PDF of β. In our case, $\beta_{mean} \approx 0.1$, whereas for the them is ~ 0.4. This is mainly attributed to our methodology of fitting simultaneously both I_{max} and β. We also hit the $\mu \approx 0.4$ when we only fit for β, using a lot more storm records than they did.

We assume that in the vast majority of the cases, the rainfall recorded by the gauge network does not correspond to the maximum intensity of the storm event; thus, our need to model for a maximum intensity (MAXINT). Eq. (1) is therefore an adequate model that allows us to easily estimate the maximum rainfall intensity from gauge records (given the current computational tools, and the extensive rainfall records). Supplementary Fig. B1 shows the difference between PDFs accounting (and not) for maximum intensity. Accounting for maximum storm rainfall intensity is a feature not present in STORM 1.0.

2.4 Storm Average Duration [AVGDUR]

The AVGDUR PDF is constructed from the correspondent "clean" optimal values for maximum intensity (MAXINT) (see Sec. 2.3). Once a "group of gauges" is established (see Sec. 2.2), we model storm duration as the average of all storm durations registered within such a group. Please recall that here a storm event is that one in which a group of gauges shares the same storm's starting time-stamp. Nevertheless, the storm's total duration registered by each gauge does differ from gauge to gauge, mainly due to the pass/movement of the storm front over the gauge network. Thus, for every fit of Eq.(1) to a group of gauges (for which I_{max} and β are estimated) an average storm duration is also retrieved. And after selecting the best fits, average storm durations included, then we proceed to fit the AVGDUR PDF.

2.5 Copula Approach [MAXINT-AVGDUR COPULA]

The cornerstone of a copula framework is (set on) Sklar's theorem (e.g., Hofert et al. (2018, chap. 1), Joe (2014, chap. 1), Nelsen (2006, chap. 2)), which states that for any d-dimensional (joint) distribution function H with univariate marginals (margins) $F_1, \ldots, F_d$, there exist a d-dimensional copula C such that:

$$H(\boldsymbol{x}) = C\big(F_1(x_1), \ldots, F_d(x_d)\big), \quad \boldsymbol{x} \in \mathbb{R}^d. \tag{2}$$

If the univariate marginals $F_1, \ldots, F_d$ are continuous, then C is uniquely defined on $[0,1]^d$. In simpler terms, a copula is a function that links/couples (thus its etymology) a multivariate (joint) distribution function to its univariate marginals, with no prior knowledge of the actual shape (or type) of such marginals (e.g., Zhang and Singh, 2019; Nelsen, 2006; Hofert et al., 2018; Vandenberghe et al., 2011; Dai et al., 2014).

Elliptical copulas (which show elliptically contoured density level surfaces) refer to copulas from elliptical distributions (e.g., Mai and Scherer (2017, chap. 4), Tjøstheim et al. (2022, chap. 5)). An elliptical distribution represents a linear transformation of spherical distributions (Mai and Scherer, 2017, chap. 4), these latter being extensions of multinormal distributions (Fang et al., 1990, chap. 2). The vast majority of application from elliptical copulas are found in financial sciences (Genest et al., 2009; The Economist, 2009). Nonetheless, there have recent applications of elliptical copulas in hydrometeorology such as modelling radar rainfall uncertainty (Dai et al., 2014), and establishing seasonal correlation between ENSO, PDO and precipitation (Khedun et al., 2014), to name a couple. Zhang and Singh (2019); Chen and Guo

(2019) provide a thorough review of recent advances and applications of copulas (elliptical among others) in several areas of hydrology fields such as extreme analysis, drought(s), rainfall, flood (frequency, forecasting, and risk), streamflow, water quality, and suspended sediment transport. Elliptical copulas are very common and advantageous as they allow the specification of different levels of global correlation between marginals (Tjøstheim et al., 2022, chap. 5). Nevertheless they offer no simple closed-form expressions, that is, they have only implicit analytical expressions/solutions (Mai and Scherer, 2017, chap. 4).

A (d-variate) Gaussian (namely, standard normal) copula belongs to the parametric family of the elliptical copulas (e.g., Mai and Scherer, 2017, Fig. 4.1), and it is described by the functional form (e.g., Mai and Scherer (2017, chap. 4)):

$$C_P^{Ga}\big(\boldsymbol{u}\big) = \boldsymbol{\Phi}_P\big(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_d)\big), \tag{3}$$

where $\boldsymbol{\Phi}_P$ is the joint cumulative distribution function (CDF) of a d-variate Gaussian distribution; Φ^{-1} is the univariate Gaussian inverse CDF (i.e., the quantile function); P is the $d \times d$ correlation matrix of multivariate normal random vector; with C_P^{Ga} denoting the copula is parametrized by the $\frac{1}{2}d(d-1)$ parameters of the correlation matrix (McNeil et al., 2015, chap. 7).

STORM uses a bi-variate Gaussian copula to model the dependence between storm rainfall intensity and duration. In a d-variate Gaussian copula the $d \times d$ correlation matrix could be replaced by a/the covariance matrix (Mai and Scherer, 2017, chap. 4). For the bi-variate case, i.e. $d = 2$, C_P^{Ga} becomes C_ρ^{Ga}, with ρ the (scalar) Pearson correlation coefficient (e.g., Joe (2014, chap. 4), McNeil et al. (2015, chap. 7), Tjøstheim et al. (2022, chap. 5)). In doing so the parameterization is reduced to its minimum (only depending of ρ); thus its (relatively) easy implementation, and therefore its popularity. Still, a bi-variate (or any d-variate, for that matter) Gauss copula does not have a simple closed form, but can be expressed as an integral over the density of a bi-variate normal random vector (e.g., McNeil et al. (2015, chap. 7), Ross (2013, chap. 6)):

$$C_\rho^{Ga}\big(u,v\big) = \int\limits_{-\infty}^{\Phi^{-1}(u)} \int\limits_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \cdot \exp\left\{ -\frac{u^2+v^2-2\rho u v}{2(1-\rho^2)} \right\} dv\, du,$$
$$\text{with } 0 \leqslant u,v \leqslant 1, \text{ and } \rho \in [-1,1]. \tag{4}$$

STORM constructs the bi-variate Gaussian copula via the `GaussianCopula` module from the *statsmodels* package (Seabold and Perktold, 2010; Joe, 2014). First of all, during the pre-processing stage (Sec. 2.8) the Pearson correlation coefficient ρ is obtained through Greiner's equality (Berger, 2016):

$$\tau = \frac{2}{\pi} \cdot \arcsin\big(\rho\big), \tag{5}$$

where τ is Kendall's rank correlation (also known as Kendall's tau) (Kendall, 1945; Virtanen et al., 2020). A rank correlation is a copula-based measure of (strength of) dependence, i.e., only depends on the copula (of a bi-variate distribution), and not on the marginals (McNeil et al., 2015, chap. 7). It is computed from the ranks of the (empirical) data, which means one only needs the ordering of the random variables, and not the actual values, i.e., storm intensity and duration in this case. Eq. (5) generally holds for elliptical copulas (from which the bi-variate Gaussian is a member); offering a simple approach to compute ρ without the estimation of variances and covariances (Langworthy et al., 2021; McNeil et al., 2015, chap. 6). Then, during a simulation (or validation) run, the bi-variate normal distribution is constructed from Eqs. (5) and (4) by using the probability integral transform (Seabold and Perktold, 2010). Once the (bi-variate) Gaussian copula is built, n samples are randomly sampled from it. These samples are drawn from the $0 \leqslant u,v \leqslant 1$ CDF-space; hence, each sample, i.e., (u,v)-point, must be transformed (back) into the intensity-duration space. This transformation is done throughout the marginal PDFs (and their `ppf` objects, from *scipy*'s module `stats`). During the pre-processing stage STORM builds the marginal PDFs for intensity and duration from the input gauge data.

Figure 2 shows a comparison between storm rainfall measured by rain gauges, and simulated from a bi-variate Gaussian copula. From this figure, one can see that for the simulated exercise (Fig. 2, panel b) STORM generates storms with higher (and lower) intensities than those actually observed by the gauge network (Fig. 2, panel a).

2.6 Day of Year [DOYEAR] & Time of Day [DATIME]

Realistic storm's starting dates and times can now be sampled in STORM through a modular implementation of directional (or circular) statistics. Directional statistics takes into consideration the periodicity of random variables that can be distributed in a closed space, e.g., torus, sphere, circle (Breitenberger, 1963). The day of the year (DOY), and the time of the day (TOD), of an occurring storm, belong to such a set of variables.

STORM models storm's starting dates and times throughout a finite mixture of unimodal von Mises (vM) distributions. The vM distribution (also known as the Tikhonov distribution, e.g., Shmaliy (2005)) is a widely used PDF (in the circle space) given its simplistic parameterization, and mathematical tractability (e.g., Mardia and Jupp, 1999; Pewsey et al., 2013). The vM distribution is a close approximation of

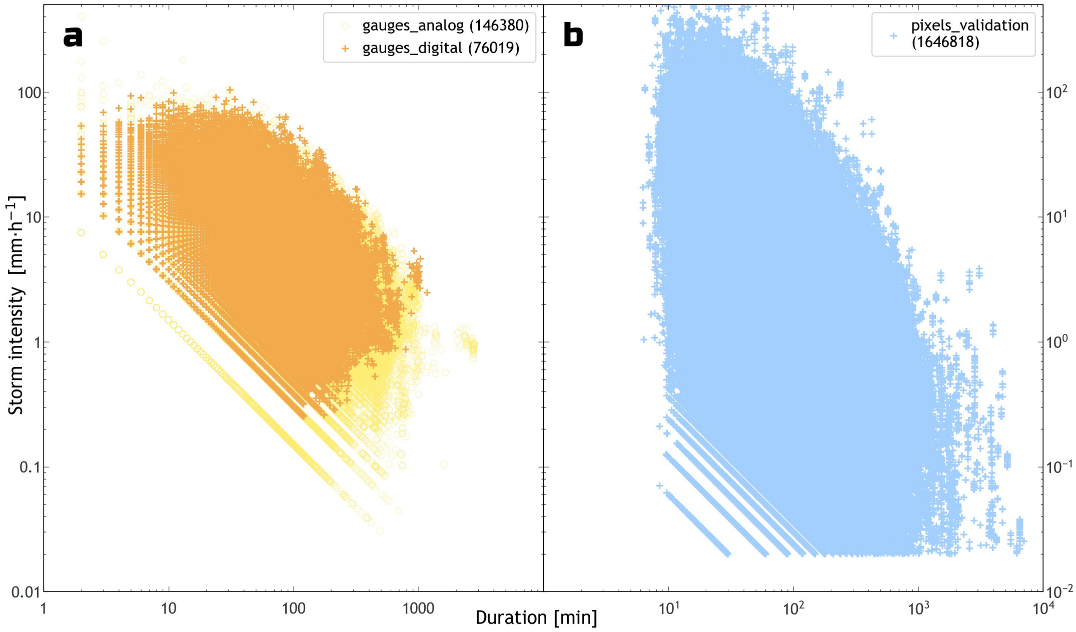

Figure 2. Scatter plots of storm intensity (y-axis, $\mathrm{mm \cdot h^{-1}}$) against storm duration (x-axis, min), in log-log scale, for gauge, and validation datasets. Panel **a** - Recorded storms for the wet season (i.e., June through October) over the Walnut Gulch catchment (see Sec. 2.10). The orange markers/crosses are records from the digital network, i.e., gauges from 2000 onwards (from June 2000 through October 2022, i.e., the validation dataset). The yellow markers/circles are records from the analog network, i.e., gauges prior to 2000 (from August 1953 through October 1999, i.e., the calibration dataset). Panel **b** - 23 years of simulated storm, each year having 30 runs. These storm intensity-duration "pairs" are obtained from the marginal PDFs fitted in the pre-processing module (see Sec. 2.8) for storm maximum intensity (MAXINT), and average duration (AVGDUR).

distributions such as the Cardiod, the wrapped Cauchy, and the wrapped normal. This latter (as its name suggests) is the equivalent of wrapping the normal distribution (from the linear space) into the circular space (Mardia and Jupp, 1999, chap. 3).

The model for a finite mixture of vM (MvM) PDFs (for a random variable θ) is given by (e.g., Jammalamadaka and SenGupta, 2001, chap. 4.3):

$$f\left(\theta \,|\, \{p,\mu,\kappa\}_{i=1}^{\mathrm{M}}\right) = \sum_{i=1}^{\mathrm{M}} p_i \cdot \frac{e^{\kappa_i \cdot \cos(\theta-\mu_i)}}{2\pi \cdot I_0(\kappa_i)},$$
$$\text{with } 0 \leqslant \theta,\, \mu_i < 2\pi,\, 0 \leqslant p_i \leqslant 1, \text{ and } \sum_{i=1} p_i = 1. \tag{6}$$

In Eq. (6), p_i is the mixing proportion of the i-unimodal vM distribution (i.e., everything to the right of p_i); κ $\{$for $\kappa \geqslant 0\}$ is the concentration parameter that quantifies the sparseness/spreadness of the distribution around its mean direction μ; and $I_0\left(\kappa\right)$ is the modified Bessel function of the first kind with order 0, and argument κ. Mardia and Jupp (1999, Eq. 3.5.19), and/or Jammalamadaka and SenGupta (2001, Eq. 2.2.7), for instance, define $I_0\left(\kappa\right)$ as:

$$I_0\left(\kappa\right) = \frac{1}{2\pi}\int_0^{2\pi} e^{\kappa \cdot \cos(\theta)}\, d\theta = \sum_{s=0}^{\infty} \frac{1}{(s!)^2}\left(\frac{\kappa}{2}\right)^{2 \cdot s}. \tag{7}$$

This latter, i.e., the term most to the right in Eq. (7), is the power series expansion (in infinite series form). Parameters μ, and $1/\kappa$ (Eq. (6)) are analogous to the mean μ, and variance σ^2 of the normal distribution.

Eq. (6) has no analytical solution. Hence, STORM uses the `vonMisesMixtures`[3] package, which computes the parameters (μ, κ, p) via Maximum Likelihood Estimators within an Expectation-Maximization framework (e.g., Hornik and Grün, 2013; Dhillon and Sra, 2003). The description of such an algorithm is beyond the scope of this work. At its core, the `vonMisesMixtures` package uses the `iv` object from *scipy*'s module `special` for the Modified Bessel function (Temme, 1975; Virtanen et al., 2020), and the `fsolve` object from *scipy*'s module `optimize` for the root finding (of non-linear functions). `fsolve`, ultimately is a wrapper for a modified Powell's hybrid method (Moré et al., 1980, p. 57-64, 71-78); this latter, an algorithm for nonlinear optimization (Powell, 2009, 1970).

[3]https://framagit.org/fraschelle/mixture-of-von-mises-distributions

Table 1. Mean dates, and times μ (in decimal days of year for DOY, and in decimal hours for TOD, respectively), concentration parameters κ, and mixing proportions p for 1, 3, and 5 mixtures of von Mises (MvM) probability density functions (PDFs). For instance, for the time of day (TOD), and for the 3 MvM PDFs, μ (in radians) are 0.691, 1.707, and 2.557, i.e., (in decimal hours) 14.64, 18.52, and 21.77 (where $0_{\text{rad}} = 12{:}00$, and $-\pi/+\pi = 00{:}00/24{:}00$). The parameters for the 5, and 3 MvM PDFs are respectively the default for the DOY, and TOD models in STORM. These defaults are defined in the `pre_processing.py` script/module (and in the input file *ProbabilityDensityFunctions_ONE–ANALOG.csv*). The fitted PDFs presented in Fig. 3 (and sup. Fig. B2) can be reconstructed by plugging these parameters into Eqs. (7) and (6).

	#-MvM		pdf-1	pdf-2	pdf-3	pdf-4	pdf-5
DOY	1	μ	-	-	223.9547	-	-
		κ	-	-	3.9086	-	-
		p	-	-	1.0000	-	-
	3	μ	-	207.1516	252.9430	-	294.0633
		κ	-	9.3424	9.2696	-	122.8981
		p	-	0.6533	0.3062	-	0.0405
	5	μ	158.0482	201.3853	238.3692	273.6777	293.2942
		κ	287.1728	15.5872	9.4628	129.0467	104.7193
		p	0.0118	0.4657	0.4250	0.0445	0.0531
TOD	1	μ	-	-	-	17.5157	-
		κ	-	-	-	1.0544	-
		p	-	-	-	1.0000	-
	3	μ	-	-	14.6420	18.5217	21.7681
		κ	-	-	6.3909	3.0643	0.4700
		p	-	-	0.2492	0.3245	0.4263
	5	μ	3.3158	8.0591	15.0518	19.0519	22.4794
		κ	4.6405	8.4354	3.7416	5.9302	3.7813
		p	0.0825	0.0424	0.4667	0.2400	0.1683

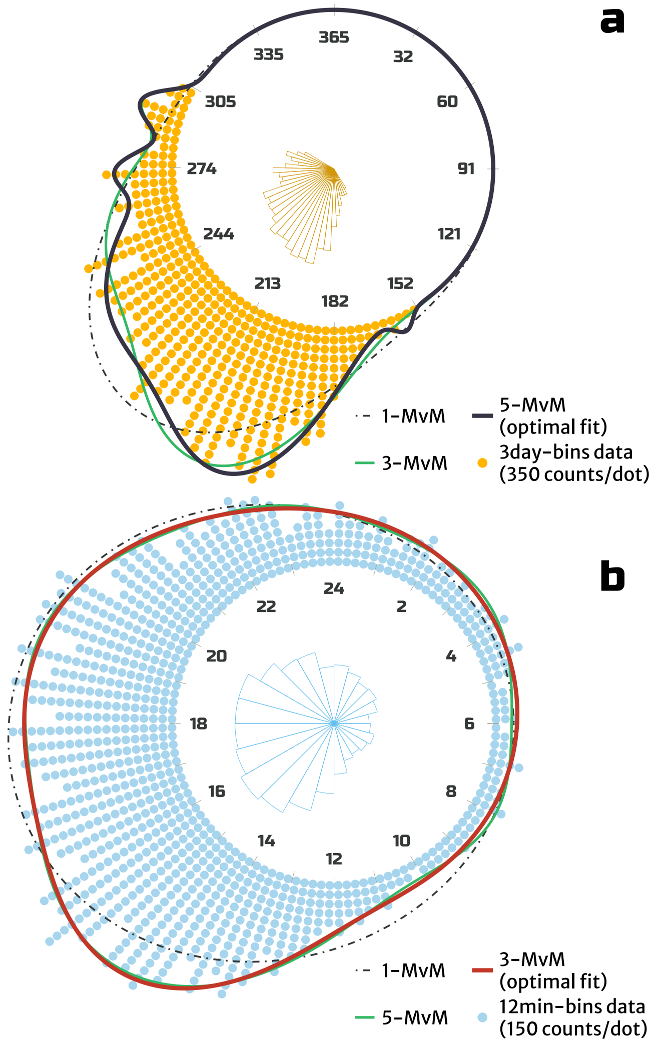

Figure 3. Panel **a** - Circular distribution for 3-day binned-data of (storm's starting) days of year (DOY; orange dots, each dot representing 350 counts). The black continuous curve indicates the optimal mixture of von Mises (MvM) probability density functions (PDFs), a mixture of 5 vM-PDFS, in this case (see Appendix A). The green curve represents a fit for 3 MvM-PDFs. A 5 day-bin circular histogram is also plotted on the inside. Panel **b** - Circular distribution for 12-min binned-data of (storm's starting) times of day (TOD; blue dots, each dot representing 150 counts). The red continuous curve indicates the optimal MvM-PDFs, i.e., 3 MvM-PDFS, in this case. The (almost imperceptible) green curve represents a fit for 5 MvM-PDFs. A 1 h-bin circular histogram is plotted on the inside. In both panels, the dashed black curves represent a fit of just 1 vM-PDF. The size of the sample is ~ 146k values, for both DOY and TOD, encompassing the wet seasons (June through October) from 1953 through 1999, in the Walnut Gulch catchment (see Sec. 2.10). Table 1 (Sec. 2.6) displays the parameters μ, κ, and p which the vM PDFs are constructed from.

Table 1 presents the estimated parameters for mixtures of 1, 3, and 5 vM-PDFs. Given the storm's starting DOY and TOD, STORM transforms those date-time stamps into radians, and feed them to the `vonMisesMixtures` package, along with the number of vM PDFs to compute the mixture. The conversion from decimal-based days (d_{dec}) into radians (d_{rad}), follows: $d_{rad} = \pi(2 \cdot d_{dec}/365 - 1)$; for $0 \leqslant d_{dec} \leqslant 365$, and $-\pi \leqslant d_{rad} \leqslant +\pi$. Similarly, the conversion from decimal-based hours (h_{dec}) into radians (h_{rad}), follows: $h_{rad} = \pi(h_{dec}/12 - 1)$; for $0 \leqslant h_{dec} \leqslant 24$, and $-\pi \leqslant h_{rad} \leqslant +\pi$. Figure 3 shows the fitted mixtures reconstructed from the parameters in Table 1, along with the circular distribution of DOY, and TOD. In this figure (panel b), the optimal (and more parsimonious) fit for TOD is given by 3 MvM-PDFs. A fit for 5 MvM-PDFs is also presented in panel b of Fig. 3, even though it overshadowed by the 3 MvM-PDFs. This shows the preference (and optimality) of the latter model not only to capturing in quite detail the (potential) multimodality of the TOD distribution but also offering a less burdensome/intensive parameter estimation, with regard to the former model (i.e., a 5 MvM-PDFs). Disregarding its circular framework, the TOD histogram presented in Fig. 3 is consistent with that of Eagleson et al. (1987, Fig. 5). Appendix A presents the rationale behind the optimum selection of 5 MvM-PDFs for DOY, and 3 MvM-PDFs for TOD, which are the default settings in STORM. Still, we encourage the user to assess the optimal number of vM PDFs on an case-by-case basis.

The choice to implement an approach like the MvM-PDFs allows the end-users to account for potential multimodality (and asymmetry) in their storm's starting date-times. Nonetheless, in the eventuality that any user encounters some difficulties when installing/running the `vonMisesMixtures` package (as it is not shipped through the `conda` channels (Anaconda Software Distribution, 2023)); or that they simply do not want to follow such an approach, STORM can still run without this feature (once it is turned off). In that case, STORM finds the best fit throughout a set of discrete probability mass functions (PMFs) for the DOY; and samples TOD from a uniform distribution (up-scaled to the $00{:}00 - 24{:}00$ h domain). Supplementary Figure B2 shows the best fit of a PMF for DOY in the Walnut Gulch dataset. In Fig. B2, one can see the advantages of using a more elaborate model. i.e., MvM-PDFs, with regard to a simple PMF model. Having a statistical model for DOY is another improvement over STORM 1.0. Thus, we avoid modelling inter-arrival, and do not contradict the notion of rainfall modelling from a (Poisson) point-process perspective (e.g., Eagleson et al., 1987).

Both TOD, and DOY sampling takes place independently from one another. Then, they are glued together into full date-time stamps (i.e., DOYEAR, and DATIME). Although theoretically possible, the probability of having two storms simulated at the same location with the very same date-time stamp is extremely low.

2.7 Scaling Factors & Stratification

One key feature carried on from its predecessor is STORM's capability to model potential future climate change scenarios throughout two scaling factors (f_1, f_2), applied to TOTALP (total seasonal rainfall), and MAXINT (maximum storm intensity). Equation (8) is a generic equation where U represents the variable to be scaled (i.e., TOTALP or MAXINT), U^* its new value after being modified by factors f_1 or f_2, and k the iterator for the number of years per simulation, namely NUMSIMYRS.

$$U^* = U \cdot \left(1 + f_1 + (f_2 \cdot k)\right), \quad 0 < k \leqslant \text{NUMSIMYRS}. \tag{8}$$

Equation (8) implies that for every simulated year one can apply either a factor f_1, which yields a constant increase (or decrease) for every year throughout the whole span of the simulation, or a factor f_2 which progressively increases (or decreases) with regard to the previous simulated year. For instance, a factor $f_1 = -0.1$ will decrease 10% of every sampled TOTALP in any given n-years simulation; whereas a factor $f_2 = +0.1$ will double the value of sampled TOTALP at the end of a 10-year simulation, for instance. Both factors (f_1, f_2) are expressed as percentages, and are mutually exclusive, i.e., STORM ensures they cannot be applied at the same time, even though Eq. (8) suggests the opposite (this constraint can easily be removed in the source code, though). Otherwise, the effect of each factor in the output becomes somewhat muddy to disentangle.

For TOTALP, $f_1 = $ PTOT_SC, and $f_2 = $ PTOT_SF; whereas for MAXINT, $f_1 = $ STORMINESS_SC, and $f_2 = $ STORMINESS_SF (i.e., variables used in the script `rainfall.py`). A legacy from STORM 1.0, PTOT_SC is a factor that simulates (percentage) step changes in the catchment wetness (seasonal precipitation totals); whereas PTOT_SF is a fractional scaling factor (progressive percentage) that simulates temporal trends in seasonal totals. Similarly, STORMINESS_SC simulates step changes in storminess (increase/decrease in maximum storm intensities); whereas STORMINESS_SF is a fractional scaling trend in maximum intensities. Section 3.2 shows the results for one simulation where PTOT_SC $= +0.5$ (Fig. 8; and Supp. Fig. B5, panel b); and another where STORMINESS_SF $= -0.035$ (Fig. 7; and Supp. Fig. B5, panel a).

STORM now offers the possibility to simulate storm rainfall at different altitude bands, so potential orographic effects are taken into account. The basic (and simplest) setup of STORM only requires the catchment shapefile (SHP) to determine the spatial domain over which the simulation(s) will take place. In this is scenario, it is not possible to determine any altitude bands within/from the SPH, and STORM falls back to sample storm's intensity-duration pairs from the "global" copula, i.e. the copula model retrieved from all gauge data (see Sec. 2.5, and Fig. 2). On the other hand, if the user not only provides a SHP but also its digital elevation model (DEM), STORM can compute as many copulas/copulae? as altitude bands the catchment is split into. To this end, and during the pre-processing stage (see Sec. 2.8.1), the user must define such altitude bands, and STORM will compute one copula per altitude band (as long as as the storm/gauge dataset also provides the altitude of the gauge network, which is almost always the case). During the simulation/validation stage, the storm's extent is defined, then overlapped to the DEM, and STORM calculates the median elevation/altitude, which is ultimately used to infer which copula (band) maximum rainfall intensity must be sampled from. By default, STORM calculates the median altitude of the storm's extent over the DEM. Nevertheless, this metric can be changed to other statistic, for instance, the mean (see Sec. 2.8.1).

2.8 Extras

2.8.1 Pre-Processing Module

This module is divided in two parts: 1) the actual module that processes all gauge data and generates the pdfs that STORM uses as input; and 2) the file *parameters.py*, where all "soft-" and "hard-coded" parameters/variables are placed, and can be read/ingested by STORM.

The standalone script `pre_processing.py` ingests event- and aggregated-based gauge data to best-fit PDFs for several variables (see Tables 2, and 3). These storm variables are: total seasonal rainfall - TOTALP, maximum extent - RADIUS, rainfall decay rate - BETPAR, maximum intensity - MAXINT, average duration - AVGDUR, intensity-duration copula - COPULA, starting date - DOYEAR, and starting time - DATETIME. The PDF parameters are exported to a CSV (Comma-Separated Values) file (stored in the *model_input/data_WG* folder) that is later read during the simulation/validation stage. If the analysis require altitude stratification, STORM generates MAXINT, and AVGDUR PDFs for each altitude band, and appends a "Z#" tag to distinguish them from the all-gauges-based PDFs (see Table 4, rows 6-11 and 13-15). Depending on the number of vM PDFs used in the DOYEAR, and DATETIME variables, STORM appends a "m#" tag (see Table 4, last 8 rows). The number 1 appended to the PDF, RHO, and VMF tags indicates that the preprocessing was done for only one wet season. If analyses are carried out for more than one wet season, STORM replicates the same analyses for every season, appending numerical tags accordingly (e.g., file *ProbabilityDensityFunctions_TWO–ANALOG-py.csv*).

Table 2. First and last four rows of the (sorted) storm event-based gauge data used by the script `pre_processing.py` to compute the best-fit parameters presented in Table 4. In the S column, W indicates a storm occurring within the established wet season, whereas D is for storms out of such a wet season. The complete table/data can be found in the file *gage_data–1953Aug18-1999Dec29_eventh–ANALOG.csv*, located in the folder/path *model_input/data_WG*.

Gage	Year	DOY	Hour	S	Duration (min)	Depth (mm)
RG022	1953	230	13.000	W	20	1.02
RG022	1953	233	13.083	W	29	8.38
RG022	1953	243	8.000	W	24	1.52
RG036	1953	230	0.167	W	24	6.10
⋮	⋮	⋮	⋮	⋮	⋮	⋮
RG100	1999	259	20.400	W	146	3.30
RG100	1999	262	21.133	W	44	0.25
RG100	1999	263	23.250	W	153	2.54
RG100	1999	265	18.550	W	12	1.78

2.8.2 Visualization Tool

GIF (Graphics Interchange Format)[4] animations of selected simulations are created via the script `animation.py` (located in STORM's *xtras* folder/path). STORM's simulations (or validations) are stored in NetCDF (Network Common Data Form)[5] files, i.e., one file per each season containing m-simulations each one of n-years. Once the NetCDF files are produced, and for a given simulation, the user can easily create animations (and/or snapshots) depicting the evolution of storm events during the wet season, along with its seasonal aggregation within the defined catchment. An example of such an animation can be found in the README.md (page) of STORM's repository[6]. The snapshots from which the animation is built upon look like Fig. 1.

Table 3. Twelve rows of the storm aggregated gauge data used by the script `pre_processing.py` to compute the best-fit for total seasonal rainfall (TOTALP in Table 4). In the S column, W indicates a month within the established wet season, whereas D is for months out of such a wet season. The complete table/data can be found in the file *gage_data–1953Aug-1999Dec_aggregateh–ANALOG.csv*, located in the folder/path *model_input/data_WG*.

Gage	Year	month	S	Rain (mm)
⋮	⋮	⋮	⋮	⋮
RG080	1990	1	D	11.94
RG080	1990	2	D	17.78
RG080	1990	3	D	9.65
RG080	1990	4	D	4.57
RG080	1990	5	D	4.32
RG080	1990	6	W	17.53
RG080	1990	7	W	150.88
RG080	1990	8	W	97.54
RG080	1990	9	W	59.69
RG080	1990	10	W	18.29
RG080	1990	11	D	24.13
RG080	1990	12	D	29.97
⋮	⋮	⋮	⋮	⋮

2.9 STORM's skeleton

Starting from the pre-processing module (see Algorithm 1), STORM ingests pre-preprocessed storm data in the format presented in Tables 2, and 3. The output of this pre-processing module is the file *ProbabilityDensityFunctions_ONE–ANALOG.csv*, containing the parameter of several PDFs needed to stochastically model rainfall storms. Table 4 presents the aforementioned file in its entirety.

Algorithm 2 is the cornerstone of STORM. This algorithm shows the main steps required to simulate storm rainfall, relating all the stochastic variables previously described in this section. Algorithm 3 (script `storm.py`) is the wrapper in charge of: 1) gathering the input files/parameters (scripts `parameters.py`, and `parse_input.py`); 2) verify that all the necessary file/parameters, and variables are correctly set, and allocated (script `check_input.py`); and 3) ultimately call Algorithm 2 (i.e., script `rainfall.py`).

[4] software developed by CompuServe (https://www.w3.org/Graphics/GIF/spec-gif87.txt)

[5] software developed by UCAR/Unidata (http://doi.org/10.5065/D6H70CW6)

[6] https://github.com/feliperiosg/STORM2

Algorithm 1 Pre-Processing module

create CSV file {all the processes below write into this file}
read (pre-processed) gauge data and metadata
fit (wet) seasonal PDF
estimate and fit radii PDF
estimate and fit rainfall decay rate and maxima intensity
compute intensity-duration copula {with stratification or not}
compute and fit DOY and TOD PDFs

Table 4. Parameters of PDFs that best fit the Walnut Gulch gauge data for a given random variable. _PDF indicates probability density functions; _RHO refers to the copula ρ-parameter; and _VMF indicates a von Mises PDF. The number next to the aforementioned nomenclature refers to the wet season for which the variable is estimated/fit. In this case theres is only one wet season, thus the number 1. The "Z#" tag refers to the altitude band for which the parameters (of the random variable) are estimated. If the variable does not present such tag (i.e., rows $1-5$, 12, and $16-23$) that means that the parameters were estimated/fit regardless altitude. Except for COPULA, DATIME, and DOYEAR, the end-string indicates the pdf-family to which the parameters belong to; so STORM (via *scipy*) can construct the adequate PDF. For variables built upon PDFs, i.e., rows 1-11, par-1 and par-2 columns are respectively for the mean, and the variance. If the PDF presents more than two parameters (i.e., par-3, and/or par-4) they are for location, and scale. For COPULA, par-1 represents the correlation parameter ρ (see Sec. 2.5). For DOYEAR, and DATIME, "m#" indicates the number of vM-PDFs that make up the mixture, i.e., 5-vM for DOYEAR (see Sec. 2.6), and 3-vM for DATIME; and columns par-1, par-2, par-3 respectively represent their p, μ (in radians), κ parameters (see Table 1). This table is produced by the script `pre_processing.py`, exported as *ProbabilityDensityFunctions_ONE–ANALOG.csv* into the *model_input* folder/path, and later ingested by STORM.

Variable's pdf (or parameter)	par-1	par-2	par-3	par-4
TOTALP_PDF1+gumbel_l	5.512	0.226		
RADIUS_PDF1+johnsonsb	1.519	1.270	-0.279	20.798
BETPAR_PDF1+exponnorm	8.287	0.018	0.010	
MAXINT_PDF1+expon	0.106	6.996		
AVGDUR_PDF1+geninvgauss	-0.090	0.770	2.843	82.079
MAXINT_PDF1+Z1+expon	0.109	5.761		
MAXINT_PDF1+Z2+expon	0.106	7.114		
MAXINT_PDF1+Z3+expon	0.305	7.353		
AVGDUR_PDF1+Z1+geninvgauss	-0.106	0.609	5.046	74.205
AVGDUR_PDF1+Z2+geninvgauss	-0.084	0.812	2.380	83.780
AVGDUR_PDF1+Z3+fisk	1.434	10.178	57.545	
COPULA_RHO1+	-0.316			
COPULA_RHO1+Z1	-0.277			
COPULA_RHO1+Z2	-0.313			
COPULA_RHO1+Z3	-0.440			
DATIME_VMF1+m1	0.249	0.692	6.391	
DATIME_VMF1+m2	0.325	1.707	3.064	
DATIME_VMF1+m3	0.426	2.557	0.470	
DOYEAR_VMF1+m1	0.045	1.570	129.047	
DOYEAR_VMF1+m2	0.012	-0.421	287.173	
DOYEAR_VMF1+m3	0.466	0.325	15.587	
DOYEAR_VMF1+m4	0.425	0.962	9.463	
DOYEAR_VMF1+m5	0.053	1.907	104.719	

Algorithm 2 Computes and exports storm rainfall

for $i \leqslant$ SEASONS **do**
  create NetCDF file
  for $j \leqslant$ NUMSIMS **do**
    for $k \leqslant$ NUMSIMYRS **do**
      TOTALP $\leftarrow$ sample total seasonal rainfall
      TOTALP $\leftarrow$ TOTALP $\cdot \left(1 + f_1 + f_2 \cdot k\right)$
      NUM_S $\leftarrow 40 * 5$ {initial number of storms}
      CUM_S $\leftarrow 0$ {initial cumulative rainfall}
      while CUM_S $<$ TOTALP $\wedge$ NUM_S $\geqslant 2$ **do**
        CENTERS $\leftarrow$ sample center geolocations
        BETPAR $\leftarrow$ sample rainfall decay rates
        RADIUS $\leftarrow$ truncated sampling of radii
        stratification {if requested}
        MAXINT, AVGDUR $\leftarrow$ copula sampling
        MAXINT $\leftarrow$ MAXINT $\cdot \left(1 + f_1 + f_2 \cdot k\right)$
        DOYEAR, DATIME $\leftarrow$ sample of date-times
        rasterisation
        interpolation
        aggregation {CUM_S updated}
        NUM_S $\leftarrow$ NUM_S$/2$
      end while
      write into NetCDF file
    end for
  end for
  close NetCDF file
end for

Algorithm 3 STORM in a nutshell

Require: input parameters {passed to the shell or read from a file}
Ensure: input parameters make sense
  call Algorithm 2 {simulates rainfall}

2.10 Walnut Gulch Catchment

The Walnut Gulch (WG) Experimental Watershed[7] is the selected catchment to calibrate and validate STORM. With an area of $147.75\,\mathrm{km}^2$, and managed by the USDA-ARS[8] Southwest Watershed Research Center (SWRC), it is located near Tombstone, southwestern Arizona, U.S. Stillman et al. (2013) describes the WGEW as having covers of shrub, and grassland as the dominant type of vegetation, with sandy, and gravely loams as (its) predominant soils, and mean annual precipitation of $350\,\mathrm{mm}$ (60% of which falls throughout JAS). Goodrich et al. (2008) documented this value to be $312\,\mathrm{mm}$, and refers to JAS as the "summer monsoon". This thunderstorm precipitation is attributed to convective summer airmasses with moisture originated in the Gulf of Mex-

[7]Historical storm data (among many other hydrological and hydrometeorological data) from the WGEW is freely available at https://www.tucson.ars.ag.gov/dap/

[8]U.S. Department of Agriculture - Agricultural Research Service, https://www.ars.usda.gov/pacific-west-area/tucson-az/southwest-watershed-research-center/

ico, and the Pacific Ocean (Osborn, 1983; Syed et al., 2003). Keefer and coauthors (2007) offer a detailed report on physiography, instrumentation, and different applications on the WGEW.

Dating back from the early/mids 1950s (Meles et al., 2022; Stillman et al., 2013), the WGEW is, according to Moran et al. (2008), “one of the most highly instrumented semiarid experimental watersheds in the world”. Its rain gauge network is one the densest in the world, for watersheds greater than $10\,\mathrm{km}^2$ ($\sim 0.6\,\mathrm{gauges}\cdot\mathrm{km}^{-2}$ (Goodrich et al., 2008); or one gauge per $1.7\,\mathrm{km}^2$ (Meles et al., 2022)). Storm rainfall data dates back from 1953 (Moran et al., 2008), and up to 1999 the entire gauge network was analog. From 2000 to the present, the gauge network was updated to a digital network (Meles et al., 2022; Goodrich et al., 2008). From the dataset used in this work, there were a total of 93 digital stations (as of 2022), averaging 84 stations per year since 2000. Supplemental Fig. B7 shows the gauge network used in this study. We parameterize STORM using 37 years of analog data (i.e., from 1963 to account at least for 80 gauges per year); and we validate the performance of STORM over the 23 years of digital/automatic data (see Sec. 3.1).

3 Results and Discussion

3.1 Evaluation of STORM

We carried out a validation run to evaluate the performance of STORM. In STORM, a “validation” run is equivalent to a “simulation” run (thus we interchangeably use these terms). The difference is that for a “simulation” run the catchment mask is exported along the output file, whereas for the “validation” run the mask of the gauge network (for which the validation exercise is carried onto) is the one stored in the output. We run through STORM 30 simulation runs, each one comprising 23 years. The above is equivalent to having ~ 1.65m storms, compared against the ~ 76k storms (for the wet season) measured by the automatic network from 2000 through 2022, i.e., the validation dataset.

In general terms, STORM does perfectly (and efficiently) well what it was set up to do, that is, to reach the median precipitation over the entire catchment. This can be seen from the box-plots presented in Fig. 4, panel a, where the (pixel/gauge aggregated) median for the validation dataset ($228.3\,\mathrm{mm}$) is just 5% larger than the median for the gauge data ($217.4\,\mathrm{mm}$). This difference is maily due to STORM always stopping after the (sampled) median seasonal total (TOTALP) is reached. Therefore, STORM seasonal aggregates (on average) will always be larger than the sampled value of reference. One advantage of such a stochastic approach is the ability to reach maxima (and minima) seasonal totals (per station/pixel) outside the inter-quartile range of the gauge dataset; thus accounting for un-recorded (but potential) extreme events.

Thanks to the statistical modeling of storm’s starting TOD, STORM is now able to capture some of the intra-seasonal variability of rainfall. This can be seen in the percentile time series of cumulative seasonal rainfall presented in Fig. 4, panel b. This latter shows how (on average) the cumulative rainfall, over the WG catchment, slowly rises to a peak (inflexion point in the solid orange line) halfway through the wet season, from which then follows a slow and steady decline up until November. Such a seasonal intra-variability is replicated by STORM (solid blue line), having a final understimation of 5.5% (i.e., $236.1\,\mathrm{mm}$) with regard to the actual seasonal (cumulative) median of $249.9\,\mathrm{mm}$. In the case that any user does not follow the circular approach (see Sec. 2.6), STORM does also replicate rainfall intra-seasonal variability by using a discrete pmf (dashed black line in Fig. 4). Something missing in this iteration of STORM is its capability to

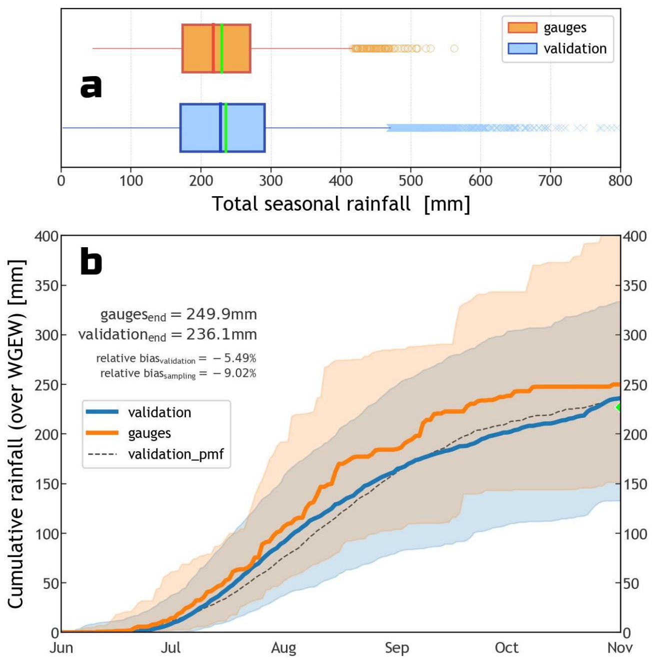

Figure 4. Panel **a** - Distribution of storm rainfall totals (for the wet season) year-by-year, and station/pixel-based, i.e., not spatially averaged over the catchment. Blue is for the validation dataset (~ 50.6k samples), whereas orange is for the gauge dataset (~ 1.9k samples). The bright green line (inside the box-plots) represents the mean of the distribution, i.e., $229.2\,\mathrm{mm}$, and $235.4\,\mathrm{mm}$ respectively for gauge and validation sets. Panel **b** - Percentile time series for the 90^{th}-percentile of all time series from June through October (wet season), for the validation (blue), and gauge (orange) datasets. The solid lines represent the median(s) of each dataset (50^{th}-percentile). The dashed black line represents the median for a validation where DOY was modeled through a discrete pmf (see Supp. Fig. B2). The green marker at the end of the time series indicates the median of the sampled (simulated) values of total seasonal rainfall (TOTALP). Supplementary Fig. B3 shows percentile time series for the 100^{th}-percentile.

model other local hydrometeorological patterns, and global teleconnections (e.g., Philander, 1990; Diaz and Markgraf, 2000; Sarachik and Cane, 2010) that might contribute to intra- and inter-seasonal rainfall variability. The scatter plot presented in Fig. 5 clearly shows STORM's innability to depict extreme stormy seasons, either wetter or drier (i.e., a very low coefficient of determination ($\rho^2 = 0.0028$)). For instance, gauge data tell us that the years 2022, and 2020 had respectively the most and the lest wet seasons of the last two decades. The seasonal averages (for the whole gauge network) were $429.9\,\mathrm{mm}$ for 2022, and $82.5\,\mathrm{mm}$ for 2020. These seasonal (mean) extremes contrast the systematic simulations (30 runs for each year) for which the validation dataset averages $237.0\,\mathrm{mm}$ for 2022, and 222.2 for 2020. Nonetheless, and regardless intra- and inter-annual rainfall variabilities, the seasonal average pixel totals ($235.4\,\mathrm{mm}$) is just 3.3% larger than the seasonal average gage totals ($228.0\,\mathrm{mm}$). The modelling of teleconnection phenomena/patterns into STORM was beyond the scope of this work.

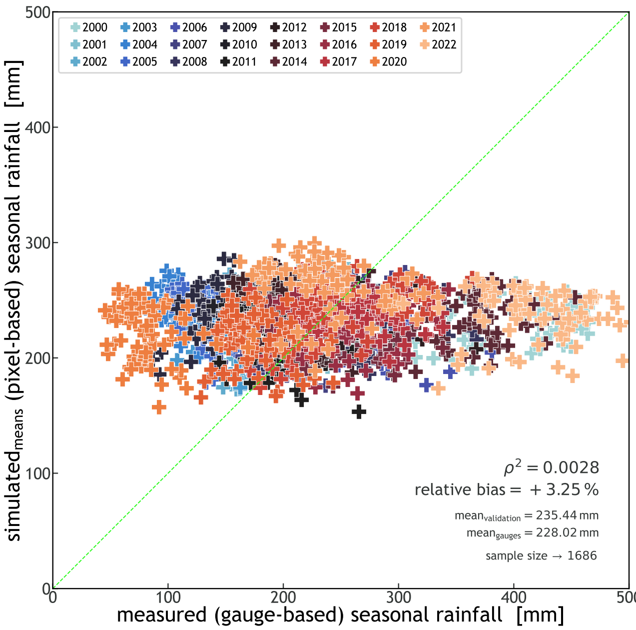

Figure 5. Scatter plot of simulated (means) seasonal rainfall against measured seasonal rainfall. Each marker/cross represents a pixel/station for which the seasonal totals of 30 simulations were averaged (y-axis), and the actual seasonal total recorded (x-axis). The color scale varies for the 23 simulated years (from 2000 through 2022). Within the plot, it is indicated the coefficient of determination (ρ^2, which is the square of the coefficient of correlation); the medians of the datasets; the relative bias between them; and the size of the sample (an average of 73.3 gauges per year). The green line indicates a 1 : 1 line.

The box-plots in Fig. 6, panel a, represent the distribution of number of storms during the wet season for both validation (blue), and gauge (orange) datasets. Once again, one can see how STORM despite being close to the average number of storms in a season (32), fails to account for the inter-annual variability in storm rainfall present in the gauge records. The average number of storms for the gauge data is (39). When disaggregated by year (see Fig. 6, panel a), the maximum average number of storms (66.3) is found for the year 2022 (with a global maxima of 79 storms), whereas the minimum average (20.6) is for 2020 (13 of global minima). As pointed before throughout the scatter plot, 2022, and 2020 match respectively the years for maximum and minimum (average) seasonal totals. Thus implying the direct relationship between the number of storm in a given season, and its total precipitation.

We selected three gauges sparsely located throughout the WG catchment, and compared the temporal distribution of their (mean) storm intensities. The box-plots in Fig. 6, panel b (all three rows), show that the mean yearly storm intensities produced by STORM (blue boxes) are consistently lower than the mean yearly intensities measured by the gauge network (orange boxes). On average, and throughout the whole validation exercise, mean recorded storm intensities ($7.14\,\mathrm{mm\cdot h^{-1}}$) are 16.2% lower than the mean of simulated storm intensities ($8.52\,\mathrm{mm\cdot h^{-1}}$). This is mainly attributed to extremely large simulated storms (see Supp. Fig. B4, panel a). With regard to the medians, storm intensities from gauge data ($3.95\,\mathrm{mm\cdot h^{-1}}$) is 48.8% larger than simulated storm intensities ($2.65\,\mathrm{mm\cdot h^{-1}}$) In spite of its inability to model inter-annual storm variability, the stochasticity imprinted in STORM allows for plausible storm intensities larger and smaller than those (ever) recorded by the gauge network (see Supp. Fig. B4, panel a, where the average of the maximum simulated intensities is $12.6\,\mathrm{mm\cdot h^{-1}}$).

One final validation exercise was to compare the top 10^{th}-percentile of all storm intensities, of both gauge and validation datasets, included simulated maximum intensities. The storms maxima (by design, see Sec. 2.3) are found in the centres of the storms, and can only be retrieved for the simulation dataset. The box-plots presented in Supp. Fig. B4, panel b, show that, despite STORM's ability to simulate (on average) extreme rainfall intensities about twice as large/high as those recorded by the gauge network; the top $10^{\mathrm{th}}\%$ of maxima simulated intensities are 44% larger than the top $10^{\mathrm{th}}\%$ of storm intensities in the gauge set. Supplementary Fig. B4, panel a, shows that on average, mean maxima intensities ($12.6\,\mathrm{mm\cdot h^{-1}}$) are 76.5% larger than the mean of actual/recorded intensities ($7.1\,\mathrm{mm\cdot h^{-1}}$); and 47.9% larger than average simulated intensities. The above suggest the goodness of the methodology here developed to account for maximum intensities when designing the storms.

3.2 Testing Climate Drivers

To evaluate the ability of STORM in accounting for potential future climate change scenarios, we carried out two more/extra validation exercises. One where TOTALP is increased by a fixed scalar throughout the whole period, i.e.,

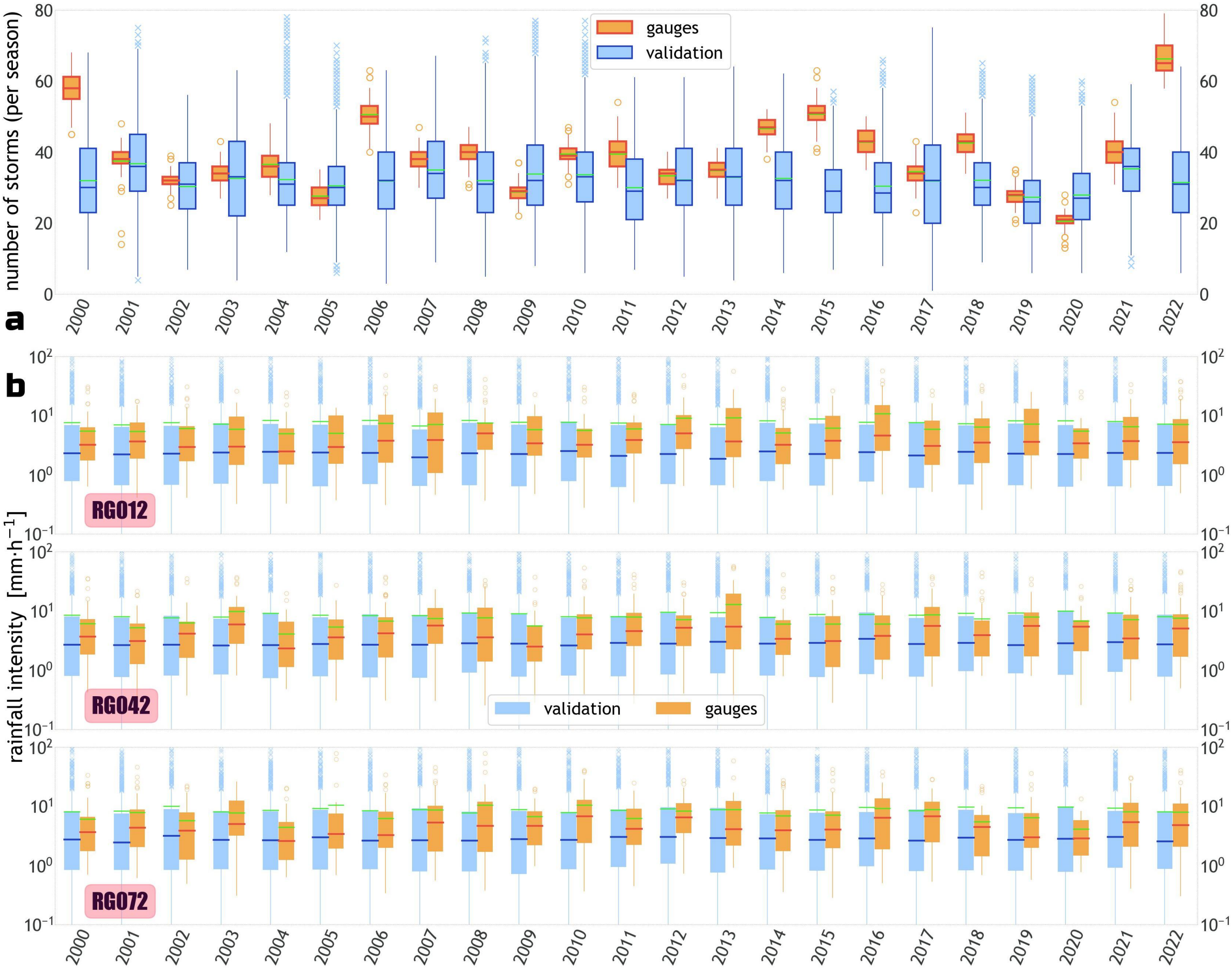

Figure 6. Yearly box-plots for the validation (blue), and gauge (orange) datasets. Panel **a** - Distribution of the number of storms in a wet season. Panel **b** - Distribution of storm intensities of three stations, i.e., RG012, RG042, and RG072 inside the Walnut Gulch catchment. In both panels, the green line within each boxplot represents the mean of the distribution. Please note the logarithmic scale of the y-axes in panel b (i.e., rainfall intensity). Supplementary Fig. B7 shows the (sparse) location of the aforementioned gauges.

PTOT_SC $= +0.5$. The other where MAXINT is reduced by a progressive scalar, i.e., STORMINESS_SF $= -0.035$. We are aware that these two scalars might not be realistic or even at all plausible. Still, we chose those numbers as they enable drastic changes in the final outputs, thus allowing easy comparisons between these “climate-driven” results, and the ones presented for the default validation (i.e., where no climate controls are simulated).

With a progressive factor _SF $= -0.035$, applied to the MAXINT variable, we force the sampled maximum storm intensity of every simulated year to be 3.5% less than the year before. Hence, for a validation run of 23 years, one can expect that in the last simulated year the (mean) decrease in maximum storm intensity would be 77% (i.e., $(23-1)\times 0.035$) less than the first simulated year. The above can be seen in the yearly box-plots presented in Fig. 7. For any of the gauges presented in Fig. 7 (e.g., gauge RG042), one can see how the median rainfall intensity of the validation dataset, i.e., $0.68\,\mathrm{mm}\cdot\mathrm{h}^{-1}$ at the end of the simulated period (2022) is 76% less than the median at the starting of the simulation (2000), i.e., $2.82\,\mathrm{mm}\cdot\mathrm{h}^{-1}$. 86.7% less when compared to the the median at the end of the actual records (i.e., $5.08\,\mathrm{mm}\cdot\mathrm{h}^{-1}$). In STORM 1.0, the progressive factor SF (over the MAXINT variable) is referred as “temporal trend in storminess” (Singer et al., 2018).

With a constant factor _SC $= +0.5$, applied to the TOTALP variable, we force the sampled seasonal total rainfall of every simulated year to be 50% larger than it normally would. Hence, no matter what year of a given validation one is running, the expected (mean) increase in seasonal total will

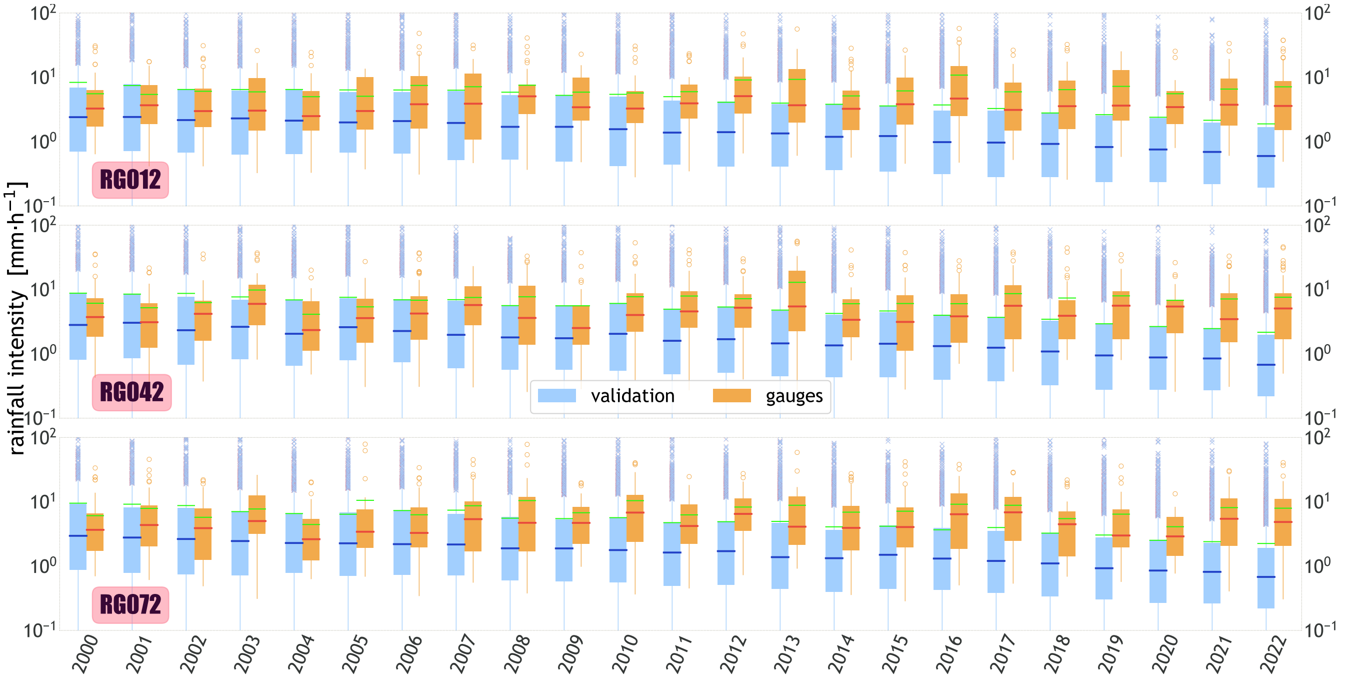

Figure 7. Distribution of storm intensities of three stations, i.e., RG012, RG042, and RG072 inside the Walnut Gulch catchment, for a validation dataset (blue), and the gauge set (orange). This plot is equivalent to Fig. 6, panel b, except that here we force the sampled maximum storm intensity (MAXINT) to be 3.5% lower than the previous year (thorough the whole period of any given simulation).

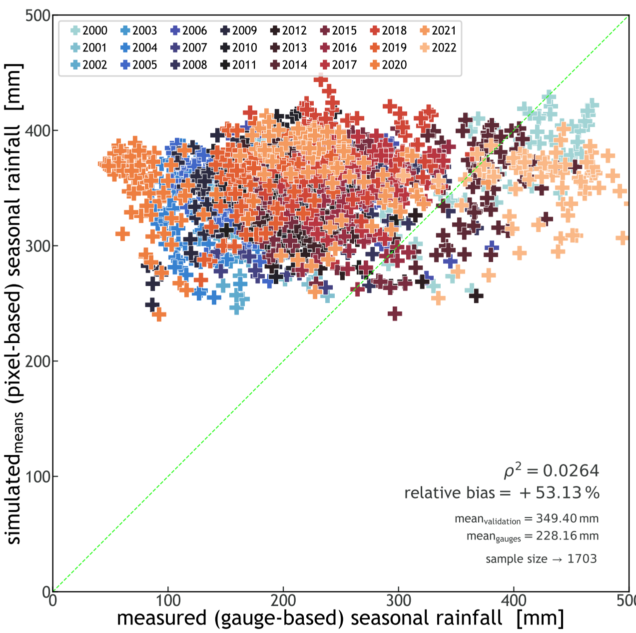

Figure 8. Scatter plot of simulated (means) seasonal rainfall against measured seasonal rainfall. This plot is equivalent to Fig. 5 except that here we force all simulated seasonal totals to be (every time) 50% larger than the sampled total seasonal rainfall (TOTALP).

be roughly constant. The above can be seen in the scatter plot presented in Fig. 8. In this figure, the cloud of points (scatter) has shifted upwards 48.4% of the mean value for simulated seasonal totals presented in Fig. 5; this latter corresponding to a validation were no climate drivers were applied. In STORM 1.0, the constant factor SC (over the TOTALP variable) is referred as "step change in wetness".

Supplementary Fig. B5 shows how the number of storms (in a wet season) are modified due to the (two) above mentioned climated drivers. For the case in which TOTALP is increased by a fixed scalar (i.e., Fig. 8), STORM generates (on average) more storms per season in order to reach the increased total seasonal rainfall. For the case in which MAXINT is progressively reduced by a progressive scalar (i.e., Fig. 7), STORM is forced to continually increase the number of storms in order to reach the median (sampled) seasonal total.

3.3 STORM Applications

These improvements to STORM 1.0 now make STORM suitable as a climate driver of other watershed response models that simulate hydrology between slopes and channels (surface runoff, infiltration, streamflow) (Michaelides and Wainwright, 2002; Michaelides and Wilson, 2007; Michaelides and Wainwright, 2008), groundwater recharge during and after rainfall events (Beven and Freer, 2001), and interactions between streamflow and alluvial aquifers (Evans et al., 2018). It also enables STORM to be useful in water balance models (e.g., Land Surface Models) to assess water availability to plants through dynamic eco-hydrological simulation of plant-climate interactions and water utilization (Caylor et al., 2006; Laio et al., 2006; D'Odorico et al., 2007), as well as energy/carbon fluxes between the land surface and the atmosphere (Bonan, 1996; Best et al., 2011). Finally, STORM can also be used to drive geomorphic models that characterize erosion and deposition processes within drainage basins in response to sequences of rainfall and runoff (Michaelides et al., 2009, 2012; Michaelides and Martin, 2012; Michaelides and Singer, 2014), and even landscape evolution models that simulate landform development over longer timescales (Tucker and Hancock, 2010; Hobley et al., 2017). Coupling STORM to such models would enable a wide range of interdisciplinary scientists to investigate key problems in the environment that have their origin in the climate system. These problems range from which water sources are used by plants (Dawson and Ehleringer, 1991; Singer et al., 2014; Evaristo et al., 2015; Sargeant and Singer, 2016; Evaristo and McDonnell, 2017) to what is the dominant source and timing of groundwater recharge (Scanlon et al., 2006; Wheater et al., 2010; Cuthbert et al., 2016) to the role of climate in shaping landscape morphology (Tucker and Slingerland, 1997; Tucker and Bras, 2000; Singer and Michaelides, 2014; Michaelides et al., 2018).

4 Summary and Conclusions

Built upon STORM 1.0, STORM[9] is an improved Stochastic Rainfall generator focused on small watersheads. This stochastic framework heavily relies on PDFs of total seasonal rainfall (TOTALP), maximum storm radius (RADIUS), decay rate of maximum rainfall from the storm's centre towards its maximum radius (BETPAR), maximum rainfall intensity (MAXINT), average storm duration (AVGDUR), the copula's correlation parameter (COPULA), storm's starting date (DOYEAR), and the (optional) storm's starting time (DATIME). The main modelling features of STORM with regard to its predecessor are: storm intensity and duration via a (bi-variate) Gaussian copula framework; intensity-duration copulas at different altitude bands within the catchment; storm occurrence via a Circular statistics approach (i.e., mixture of von Mises PDF) or via discrete PMFs; storm starting times via a Circular statistics (optional); compressed implementation of future (and very likely) climate scenarios; output compressed into (geo-referenced) NetCDF files, readily available for visualization; and pre-processing module to construct all necessary PDFs from gauge data. Added to STORM, and with a future mindset of its applicability at larger scales, we implemented slick and cool capabilities/tweaks such as: PDFs easily defined by the user (or retrieved from gauge data); storm simulation with regard to altitude (provided a Digital Elevation Model - DEM); customizable spatial resolution (and Coordinate Reference System - CRS); spatial operations under a raster framework, thus adding speed, versatility, and scalability; and optimal output storing in NetCDF format.

To develope the stochastic model, we derived and calibrated all PDFs to 37 years of storm data, collected by an analog network of 118 gauges sparsely deployed over the Walnut Gulch (WG) catchment ($148\,\mathrm{km}^2$). To test the performance of the model, we carried out one validation exercise consisting of 30 runs, each one having 23 simulated years (i.e., 690 simulation-years in total). The output of such a validation run was compared against 23 years of storm data, collected by the digital network of 94 gauges located within the WG. To evaluate the STORM's ability to model storm rainfall under potential future varying-climate scenarios, we carried out two more validation runs, each one comprising 690 simulation-years too. These results were also compared against the digital/automatic gauge network.

Results showed that the seasonal total rainfall reached by STORM is 5.5% lower than the actual records, when accounted as the spatial median of all the stations/pixels within the WG (see Fig. 4, panel b). If accounted on a temporal basis, i.e., without any spatial averaging, this relative difference amounts to $+5\%$ (see Fig. 4, panel a). $+3.3\%$ when accounted on a station/pixel basis (see Fig. 5). This general small but positive difference is mainly attributed STORM seasonal aggregates being always larger than the sampled value of reference as STORM stops only after the median seasonal total is reached. On a seasonal basis, the storm rainfal intensity recorded by the gauge network is (on average) 16.2% smaller than simulated storm intensities (see Supp. Fig. B4, panel a). Nevertheless, the stochasticity embedded in our model allows for un-recorded but very plausible, either larger and/or smaller, storm intensities (see Fig. 6, panel b, and Supp. Fig. B4).

STORM's Achilles' Heel is its innability to account for other local hydrometeorological patterns, and global teleconnections that may contribute to intra- and inter-seasonal rainfall variability (see Figs. 5, and 6, panel a). This is something expected as STORM (by design) does not incorporate any PDF that might describe (even remotely) the behaviour of such inter-annual variability. On the bright side, results obtained for the varying-climate simulations showed that STORM is able to imprint seasonal varibility to storm rainfall (either in intensities or totals), on long-term analyses. This seasonal varibility might be likely attributed to change in climate-drivers such an increse/decrease in (air) temperature, thus reflecting an increase/decrease (so called feedbacks?) in storminess or seasonal totals (see Figs. 7, and 8).

[9]https://github.com/feliperiosg/STORM2

5 Constraints and Recommendations

The choice of a bi-variate Gaussian copula was mainly driven on its simplicity/easy-configuration, and applicability. Nevertheless, a further improvement (at least conceptually) might be the implementation of a more elaborate copula models (and truly applicable to the intensity-duration case) like Extreme-Values, Archimedean, etc. (e.g., Zhang and Singh (2019); Chen and Guo (2019)).

Another key area of future work would be to investigate how temporal resolution of rainfall data affects the signal of observed trends in rainfall (e.g., Barbero et al., 2017) and how these might yield different watershed responses. this is entirely's Michael's

Code and data availability. The STORM code and its pre-, and post-processed data can be found in STORM's repo | https://github.com/feliperiosg/STORM2. Documentation to run the model, and tools for its output visualization are also provided in the aforementioned link. The DOI for the STORM v.2.X is doi:10.???/zenodo.???.

Appendix A: BIC Estimation

The Bayesian information criterion (BIC; also known as Schwarz's Bayesian criterion - SBC) is a metric used for the unbiased assessment of the optimal number of M-unimodal vM distributions (e.g., Rios Gaona and Villarini, 2018; Lark et al., 2014). Such a criterion allows the selection of the least complex of all the models in consideration, that is, the one with the lowest BIC. From a mathematical point of view (Eq. (A1)), BIC (or similar models, i.e., AIC - Akaike's information criterion) combines the maximized log likelihood of the fitted model with a penalization term that is related to the number of estimated parameters (Pewsey et al., 2013, Eq. (6.3)).

$$\mathrm{BIC} = \nu \cdot \ln(n) - 2 \cdot \ell_{max}, \tag{A1}$$

where ℓ_{max} is the maximized (full) log-likelihood of a model with ν degrees of freedom, and n the number of observations.

Unfortunately, the `vonMisesMixtures` package does not offer a way to retrieve the maximized log-likelihood from which to compute the BIC of the mixture of M-unimodal vM PDFs. Unlike Python's `vonMisesMixtures` package, R (R Core Team, 2023), jointly with the `movMF` package (Hornik and Grün, 2014), does offer the possibility to easily retrieve BIC estimates for fitted MvM PDFs (Supp. Fig. A1). The implementation of such a feature in STORM was beyond the scope of this work. Nevertheless, STORM does offer the script `pre_processing_circular.R`, which the entire circular analyses (BIC included) can be computed from. Once this analysis is carried out, the user will have all the neceesary elements to discern the optimal fit for their "circular" data.

Figure A1 shows the DOY, and TOD BICs for mixtures ranging from 1 to 9 vM PDFs. Strictly speaking, and for the DOY case, the lowest BIC found in the figure is for a mixture of 9 vM, i.e., -318370.54. One can argue that a 9-MvM model certainly over-fits the multimodality of DOY (see Fig. 3, panel a), without even mentioning its computationally intensive parameter-estimation. Nevertheless, if one looks at the 5-MvM model (BIC equals to -317840.12), one can see that the improvement of the BIC metric is increasingly very small beyond this point in comparison to the 1-, to 4-MvM models. Therefore, we are confident that a 5-MvM model not only accurately describes the multimodality of DOY (for the Walnut Gulch dataset) but also is faster in its parameter-estimation with regard to any larger (i.e., more vM PDFs) model. Hence, a mixture of 5 vM-PDFs is the default configuration for DOY in STORM. Following that train of thought, we found the 3-MvM model the optimal mixture for TOD, and thus its default settings in STORM.

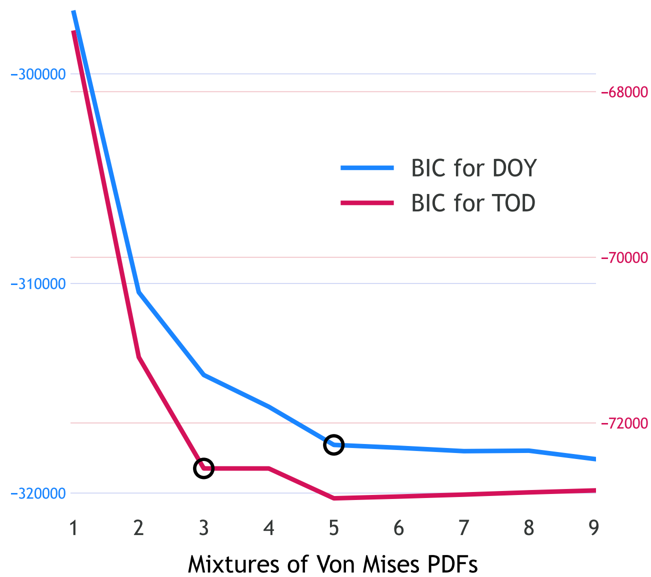

Figure A1. Bayesian information criterion (BIC) for mixtures that go from 1 to 9 von Mises (vM) probability density functions (PDFs). The blue line is for the BIC of day-of-year (DOY); whereas the red line is for the BIC of time-of-day (TOD). The color of the y-axes indicate the values of their respective BICs. The black circles indicate one of the lowest point of the related BIC curve. The lower the BIC the more optimal the number of vM PDFs (in the mixture) that best describes the sample multimodality. Thus, to avoid the selection of a model with too many vM-PDFs, the black circles also indicate where the change, in slope, is more drastic even if they are not global minima.

Appendix B: Supplementary Figures

Supplementary figures start after Acknowledgements!

Author contributions. M.F. Rios Gaona wrote and extensively tested the code, did the analyses and visualizations, and completed the early version of this manuscript. M.B. Singer developed the idea, wrote the early version of this manuscript, revised the finished version of the manuscript, and provided valuable feedback. K. Michaelides revised the finished version of this manuscript.

Competing interests. The authors declare no competing interests.

Disclaimer. The authors take no responsibility for the use or misuse of the provided code.

Acknowledgements. This work was supported by the European Union Horizon 2020 programme (DOWN2EARTH; grant no. 869550). We thank Owen Jones for his suggestions and comments.

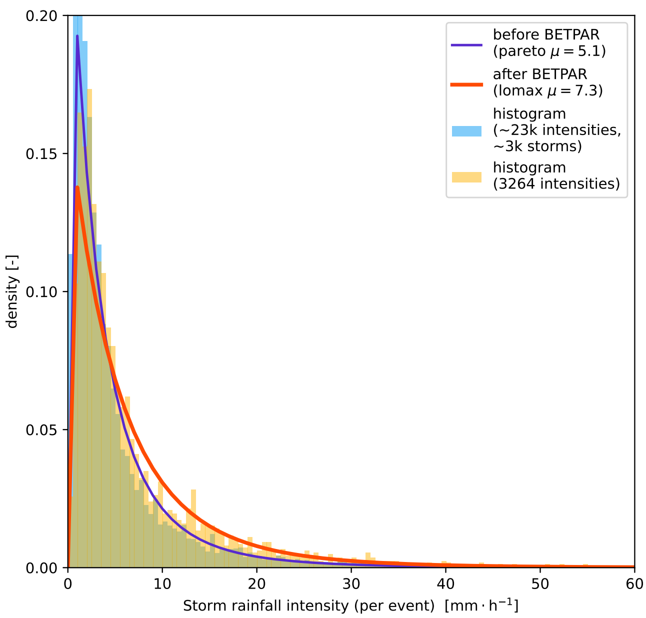

Figure B1. Probability density functions (PDFs) for storm rainfall measured by gauge data (blue curve), and for maximum (estimated) rainfall (orange curve). Maximum intensities are retrieved by fitting an exponential (quadratic) model $I(r) = I_{max} \cdot e^{-2 \cdot \beta^2 \cdot r^2}$ to measured storm rainfall (see Sec. 2.3). The background histograms show the data for which the pdfs are best fitted to. Note how the mean from maximum (estimated) intensities is larger than the mean of rainfall intensities measured by gauges prior any model fitting. I STRONGLY ADVISE AGAINST THIS FIGURE! (FOR VERY PRACTICAL PURPOSES)

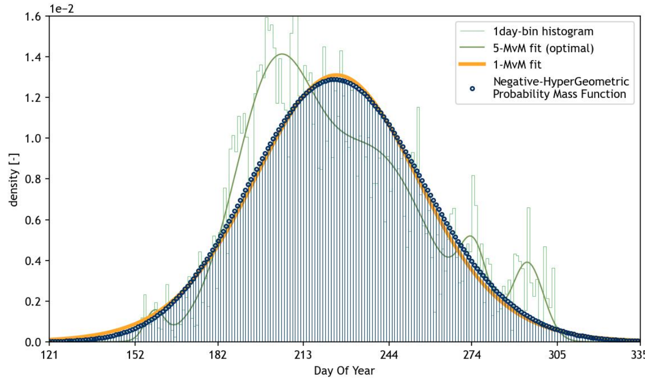

Figure B2. Probability mass function (PMF - vertical lines ending in blue circles) of a negative hyper-geometric family, fitted (best fit) to a distribution of storm's starting day-of-year (DOY - green histogram). The coarser green line represents the optimal fit of a mixture of 5-von Mises (MvM) PDFs for the aforementioned distribution; presented also in Fig. 3 (over a "circular" space; see Sec. 2.6). The orange line is a mixture of just one von Mises PDF. Note its similarity with the PMF, and its poor fit of the underlaying DOY-distribution with regard to the 5 MvM-PDFs fit.

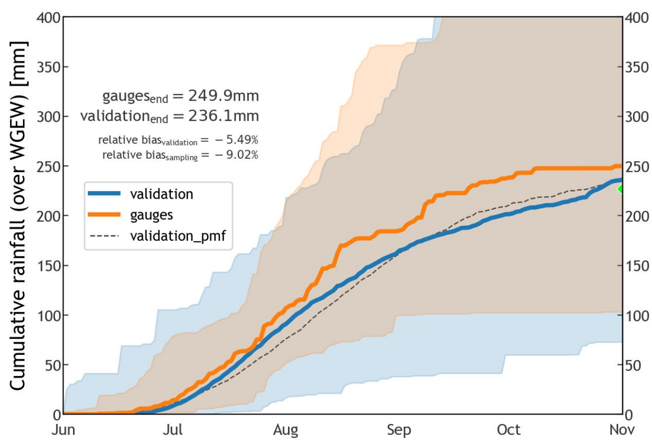

Figure B3. Percentile time series for the 100^{th}-percentile of all time series from June through October (wet season), for the simulation/-validation (blue), and gauge (orange) datasets. The solid lines represent the median(s) of each dataset (50^{th}-percentile). The dashed black line represents the median for a validation where DOY was modeled through a discrete pmf (see Supp. Fig. B2). The green marker at the end of the time series indicates the median of the sampled (simulated) values of total seasonal rainfall (TOTALP). By design, STORM stops once the sampled seasonal total is reached or surpassed (the probability of reaching exactly the sampled value is extremely low). Hence, the actual (median) simulated seasonal total will always be greater than the sampled TOTALP.

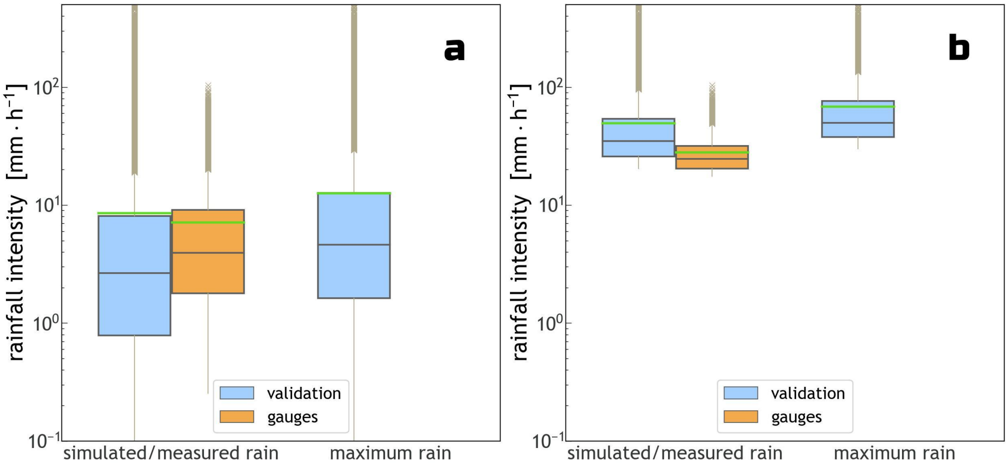

Figure B4. Distribution of storm station/pixel-based intensities. Panel **a** - for all data, i.e., 100^{th}-percentile. Panel **b** - for the top 10^{th}-percentile of all storm intensities. Blues is for the validation dataset, whereas orange is for gauge data. The green lines represent the mean of the distributions. Please note the logarithmic scale of the y-axes in both panels. The column most to the right is for the maxima intesities found in the storm centres (see Sec. 2.3). Such storm centre maxima are only retrieved for the validation dataset (no way to account for them in the gauge set).

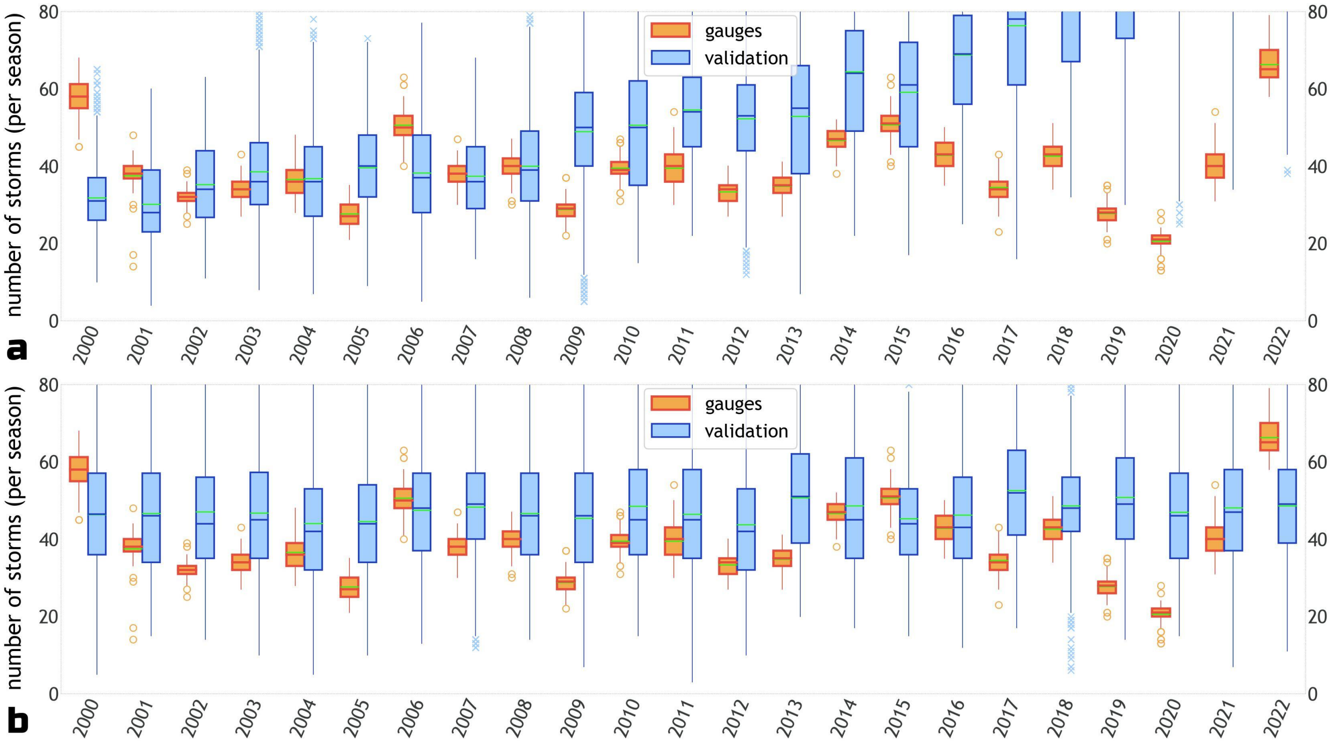

Figure B5. Distribution of the number of storms in a wet season, for the validation (blue), and gauge (orange) datasets. Panel **a** - Validation case for which MAXINT is reduced by a progressive scalar, i.e., STORMINESS_SF $= -0.035$ (see Fig. 7). Panel **b** - Validation case for which TOTALP is increased by a fixed scalar throughout the whole period, i.e., PTOT_SC $= +0.5$ (see Fig. 8). All y-axes are consistent with Fig. 6, panel a, to allow (visually) equivalent comparisons. Note how in panel a STORM generates more storms per season in order to reach the now increased total seasonal rainfall; whereas in panel b the progressive decrease in storm intensity forces STORM to continually increase the number of storms in order to reach the median (sampled) seasonal total.

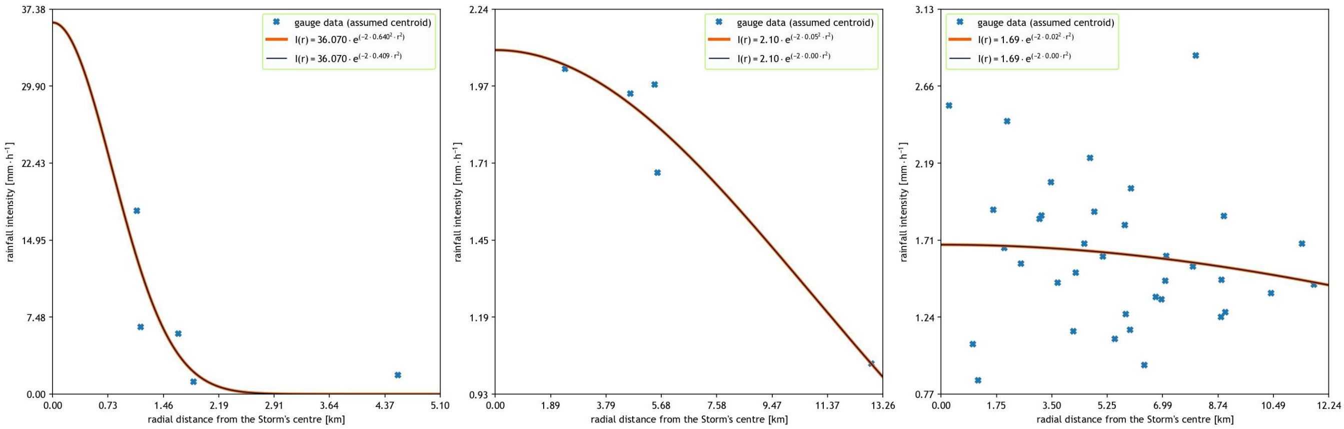

Figure B6. Examples of fitting the exponential (quadratic) model $I(r) = I_{max} \cdot e^{-2 \cdot \beta^2 \cdot r^2}$ to storm rainfall data. The x-axis represents the distance from the storm centre to the gauge registering storm rainfall. The storm centre is assumed to be the centroid of the group of gauges for which the storm starting time is identical among them (see Sec. 2.2). Hence, the blue crosses (in each panel) represent one storm event being registered by multiple gauges. Y-axis is for the rainfall intensity. The orange line(s) represents the best fit of the exponential model (for each case). Note how the model does fit a maximum intensity, not registered by any gauge, at the assumed storm's centre (left and middle panels). Still, there are cases in which the model under-performs in capturing indeed maximum intensities very close to the assumed storm's centre.

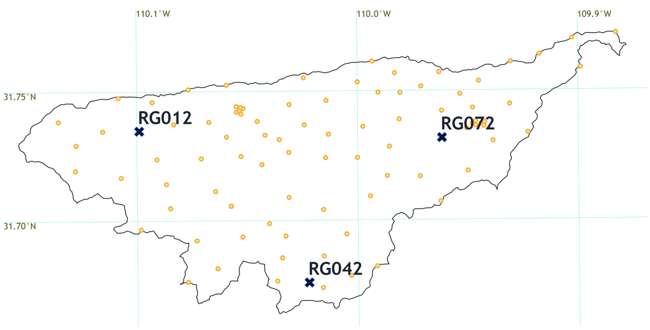

Figure B7. Digital gauge network for the Walnut Gulch catchment (from 2000 through 2022). The 3 bold markers, i.e., gauges/stations RG012, RG042, and RG072, indicate the geo-location of the gauges referred to in Figs. 6, and 7. Even though the grid is presented in "lat-lon" coordinates (i.e., CRS WGS-84), the actual projection (in both panels) is the 2D-Cartesian coordinate system known as NAD83 / UTM zone 12N (i.e., EPSG:26912).

References

Anaconda Software Distribution: Conda, https://anaconda.com, 2023.

Barbero, R., Fowler, H. J., Lenderink, G., and Blenkinsop, S.: Is the intensification of precipitation extremes with global warming better detected at hourly than daily resolutions?, Geophys Res Lett, 44, 974–983, https://doi.org/10.1002/2016GL071917, 2017.

Berger, D.: Kendall's Rank Correlation vs Pearson's Linear Correlation: A Proof Of Greiner's Relation, https://doi.org/10.2139/ssrn.2837712, 2016.

Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., Edwards, J. M., Hendry, M. A., Porson, A., Gedney, N., Mercado, L. M., Sitch, S., Blyth, E., Boucher, O., Cox, P. M., Grimmond, C. S. B., and Harding, R. J.: The Joint UK Land Environment Simulator (JULES), model description - Part 1: Energy and water fluxes, Geosci Model Dev, 4, 677–699, https://doi.org/10.5194/gmd-4-677-2011, 2011.

Beven, K. and Freer, J.: A dynamic TOPMODEL, Hydrol Process, 15, 1993–2011, https://doi.org/10.1002/hyp.252, 2001.

Bonan, G. B.: A Land Surface Model (LSM Version 1.0) for Ecological, Hydrological, and Atmospheric Studies: Technical Description and User's Guide (No. NCAR/TN-417+STR), Tech. Rep. NCAR/TN-417+STR, University Corporation for Atmospheric Research, https://doi.org/10.5065/D6DF6P5X, 1996.

Branch, M. A., Coleman, T. F., and Li, Y.: A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems, SIAM J. Sci. Comput., 21, 1–23, https://doi.org/10.1137/S1064827595289108, 1999.

Breitenberger, E.: Analogues of the normal distribution on the circle and the sphere, Biometrika, 50, 81–88, https://doi.org/10.1093/biomet/50.1-2.81, 1963.

Caylor, K. K., D'Odorico, P., and Rodriguez-Iturbe, I.: On the ecohydrology of structurally heterogeneous semiarid landscapes, Water Resour Res, 42, https://doi.org/10.1029/2005WR004683, 2006.

Chen, L. and Guo, S.: Copulas and Its Application in Hydrology and Water Resources, no. 2364-8198 in Springer Water, Springer Singapore, 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore, 1 edn., https://doi.org/10.1007/978-981-13-0574-0, 2019.

Cuthbert, M. O., Acworth, R. I., Andersen, M. S., Larsen, J. R., McCallum, A. M., Rau, G. C., and Tellam, J. H.: Understanding and quantifying focused, indirect groundwater recharge from ephemeral streams using water table fluctuations, Water Resour Res, 52, 827–840, https://doi.org/10.1002/2015WR017503, 2016.

Dai, Q., Han, D., Rico-Ramirez, M. A., and Islam, T.: Modelling radar-rainfall estimation uncertainties using elliptical and Archimedean copulas with different marginal distributions, Hydrol. Sci. J., 59, 1992–2008, https://doi.org/10.1080/02626667.2013.865841, 2014.

Dawson, T. E. and Ehleringer, J. R.: Streamside trees that do not use stream water, Nature, 350, 335–337, https://doi.org/10.1038/350335a0, 1991.

Dhillon, I. and Sra, S.: Modeling Data using Directional Distributions, Tech. Rep. TR-03-06, University of Texas: Department of Computer Science, Austin, TX, USA, https://www.cs.utexas.edu/users/inderjit/public_papers/tr03-06.pdf, 2003.

Diaz, H. F. and Markgraf, V., eds.: El Niño and the Southern Oscillation: Multiscale Variability and Global and Regional Impacts, Cambridge University Press, https://doi.org/10.1017/CBO9780511573125, 2000.

D'Odorico, P., Caylor, K., Okin, G. S., and Scanlon, T. M.: On soil moisture-vegetation feedbacks and their possible effects on the dynamics of dryland ecosystems, J. Geophys. Res. G: Biogeosci., 112, https://doi.org/10.1029/2006JG000379, 2007.

Eagleson, P. S., Fennessey, N. M., Qinliang, W., and Rodriguez-Iturbe, I.: Application of spatial Poisson models to air mass thunderstorm rainfall, J. Geophys. Res. D: Atmos., 92, 9661–9678, https://doi.org/10.1029/JD092iD08p09661, 1987.

Evans, C. M., Dritschel, D. G., and Singer, M. B.: Modeling Subsurface Hydrology in Floodplains, Water Resour Res, 54, 1428–1459, https://doi.org/10.1002/2017WR020827, 2018.

Evaristo, J. and McDonnell, J. J.: Prevalence and magnitude of groundwater use by vegetation: a global stable isotope meta-analysis, Sci. Rep., 7, 44 110, https://doi.org/10.1038/srep44110, 2017.

Evaristo, J., Jasechko, S., and McDonnell, J. J.: Global separation of plant transpiration from groundwater and streamflow, Nature, 525, 91–94, https://doi.org/10.1038/nature14983, 2015.

Fang, K.-T., Kotz, S., and Ng, K. W.: Symmetric Multivariate and Related Distributions, no. 36 in Monographs on Statistics and Applied Probability, Chapman & Hall/CRC, 6000 Broken Sound Parkway NW, Suite 300 Boca Raton, FL 33487-2742, 1 edn., https://doi.org/10.1201/9781351077040, 1990.

Genest, C., Rémillard, B., and Beaudoin, D.: Goodness-of-fit tests for copulas: A review and a power study, Insurance Math. Econom., 44, 199–213, https://doi.org/10.1016/j.insmatheco.2007.10.005, 2009.

Goodrich, D. C., Keefer, T. O., Unkrich, C. L., Nichols, M. H., Osborn, H. B., Stone, J. J., and Smith, J. R.: Long-term precipitation database, Walnut Gulch Experimental Watershed, Arizona, United States, Water Resour Res, 44, https://doi.org/10.1029/2006WR005782, 2008.

Hobley, D. E. J., Adams, J. M., Nudurupati, S. S., Hutton, E. W. H., Gasparini, N. M., Istanbulluoglu, E., and Tucker, G. E.: Creative computing with Landlab: an open-source toolkit for building, coupling, and exploring two-dimensional numerical models of Earth-surface dynamics, Earth Surf. Dyn., 5, 21–46, https://doi.org/10.5194/esurf-5-21-2017, 2017.

Hofert, M., Kojadinovic, I., Mächler, M., and Yan, J.: Elements of Copula Modeling with R, no. 2197-5744 in Use R!, Springer Cham, Gewerbestrasse 11, 6330 Cham, Switzerland, 1 edn., https://doi.org/10.1007/978-3-319-89635-9, 2018.

Hornik, K. and Grün, B.: On maximum likelihood estimation of the concentration parameter of von Mises–Fisher distributions, Comput. Statist., 29, 945–957, https://doi.org/10.1007/s00180-013-0471-0, 2013.

Hornik, K. and Grün, B.: movMF: An R Package for Fitting Mixtures of von Mises-Fisher Distributions, J Stat Softw, 58, 1–31, https://doi.org/10.18637/jss.v058.i10, 2014.

Jammalamadaka, S. R. and SenGupta, A.: Topics in Circular Statistics, no. 5 in Series on Multivariate Analysis, World Scientific, P O Box 128, Farrer Road, Singapore 912805, https://doi.org/10.1142/4031, 2001.

Joe, H.: Dependence Modeling with Copulas, no. 134 in Monographs on Statistics and Applied Probability, Chapman & Hall/CRC, 6000 Broken Sound Parkway NW, Suite 300 Boca Raton, FL 33487-2742, 1 edn., https://doi.org/10.1201/b17116, 2014.

Keefer, T. O. and coauthors: Southwest Watershed Research Center and Walnut Gulch Experimental Watershed, Tech. Rep. SWRC Publ. Reference No. 1588, Southwest Watershed Research Center, 2000 East Allen Road, Tucson, AZ 85719, http://www.tucson.ars.ag.gov/unit/publications/PDFfiles/1588.pdf, 2007.

Kendall, M. G.: The Treatment of Ties in Ranking Problems, Biometrika, 33, 239–251, https://doi.org/10.1093/biomet/33.3.239, 1945.

Khedun, C. P., Mishra, A. K., Singh, V. P., and Giardino, J. R.: A copula-based precipitation forecasting model: Investigating the interdecadal modulation of ENSO's impacts on monthly precipitation, Water Resour Res, 50, 580–600, https://doi.org/10.1002/2013WR013763, 2014.

Laio, F., D'Odorico, P., and Ridolfi, L.: An analytical model to relate the vertical root distribution to climate and soil properties, Geophys Res Lett, 33, https://doi.org/10.1029/2006GL027331, 2006.

Langworthy, B. W., Stephens, R. L., Gilmore, J. H., and Fine, J. P.: Canonical correlation analysis for elliptical copulas, J. Multivariate Anal., 183, 104 715, https://doi.org/10.1016/j.jmva.2020.104715, 2021.

Lark, R. M., Clifford, D., and Waters, C. N.: Modelling complex geological circular data with the projected normal distribution and mixtures of von Mises distributions, Solid Earth, 5, 631–639, https://doi.org/10.5194/se-5-631-2014, 2014.

Mai, J.-F. and Scherer, M.: Simulating Copulas, no. Vol. 6 in Quantitative Finance, World Scientific, 5 Toh Tuck Link, Singapore 596224, 2 edn., https://doi.org/10.1142/10265, 2017.

Mardia, K. and Jupp, P.: Directional Statistics, Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd., West Sussex, PO19 1UD England, https://doi.org/10.1002/9780470316979, 1999.

McNeil, A. J., Frey, R., and Embrechts, P.: Quantitative Risk Management: Concepts, Techniques and Tools, Princeton Series in Finance, Princeton University Press, 41 William Street, Princeton, New Jersey 08540, revised edn., https://dl.acm.org/doi/10.5555/2811305, 2015.

Meles, M. B., Demaria, E. M. C., Heilman, P., Goodrich, D. C., Kautz, M. A., Armendariz, G., Unkrich, C., Wei, H., and Perumal, A. T.: Curating 62 Years of Walnut Gulch Experimental Watershed Data: Improving the Quality of Long-Term Rainfall and Runoff Datasets, Water, 14, https://doi.org/10.3390/w14142198, 2022.

Michaelides, K. and Martin, G. J.: Sediment transport by runoff on debris-mantled dryland hillslopes, J. Geophys. Res. Earth Surf., 117, https://doi.org/10.1029/2012JF002415, 2012.

Michaelides, K. and Singer, M. B.: Impact of coarse sediment supply from hillslopes to the channel in runoff-dominated, dryland fluvial systems, J. Geophys. Res. Earth Surf., 119, 1205–1221, https://doi.org/10.1002/2013JF002959, 2014.

Michaelides, K. and Wainwright, J.: Modelling the effects of hillslope-channel coupling on catchment hydrological response, Earth Surf Proc Land, 27, 1441–1457, https://doi.org/10.1002/esp.440, 2002.

Michaelides, K. and Wainwright, J.: Internal testing of a numerical model of hillslope-channel coupling using laboratory flume experiments, Hydrol Process, 22, 2274–2291, https://doi.org/10.1002/hyp.6823, 2008.

Michaelides, K. and Wilson, M. D.: Uncertainty in predicted runoff due to patterns of spatially variable infiltration, Water Resour Res, 43, https://doi.org/10.1029/2006WR005039, 2007.

Michaelides, K., Lister, D., Wainwright, J., and Parsons, A. J.: Vegetation controls on small-scale runoff and erosion dynamics in a degrading dryland environment, Hydrol Process, 23, 1617–1630, https://doi.org/10.1002/hyp.7293, 2009.

Michaelides, K., Lister, D., Wainwright, J., and Parsons, A. J.: Linking runoff and erosion dynamics to nutrient fluxes in a degrading dryland landscape, J. Geophys. Res. G: Biogeosci., 117, https://doi.org/10.1029/2012JG002071, 2012.

Michaelides, K., Hollings, R., Singer, M. B., Nichols, M. H., and Nearing, M. A.: Spatial and temporal analysis of hillslope-channel coupling and implications for the longitudinal profile in a dryland basin, Earth Surf Proc Land, 43, 1608–1621, https://doi.org/10.1002/esp.4340, 2018.

Moran, M. S., Holifield Collins, C. D., Goodrich, D. C., Qi, J., Shannon, D. T., and Olsson, A.: Long-term remote sensing database, Walnut Gulch Experimental Watershed, Arizona, United States, Water Resour Res, 44, https://doi.org/10.1029/2006WR005689, 2008.

Moré, J. J., Garbow, B. S., and Hillstrom, K. E.: User Guide for MINPACK-1, Tech. Rep. ANL-80-74, Argonne National Laboratory, Argonne, IL, USA, https://www.math.utah.edu/software/minpack/, 1980.

Morin, E., Goodrich, D. C., Maddox, R. A., Gao, X., Gupta, H. V., and Sorooshian, S.: Rainfall modeling for integrating radar information into hydrological model, Atmos Sci Lett, 6, 23–30, https://doi.org/10.1002/asl.86, 2005.

Nelsen, R. B.: An Introduction to Copulas, no. 2197-568X in Springer Series in Statistics, Springer New York, New York, NY 10013, USA, 2 edn., https://doi.org/10.1007/0-387-28678-0, 2006.

Nicholson, S. E.: Dryland Climatology, Cambridge University Press, The Edinburgh Building, Cambridge CB2 8RU, UK, https://doi.org/10.1017/CBO9780511973840, 2011.

Osborn, H. B.: Timing and duration of high rainfall rates in the southwestern United States, Water Resour Res, 19, 1036–1042, https://doi.org/10.1029/WR019i004p01036, 1983.

Pewsey, A., Neuhäuser, M., and D., R. G.: Circular Statistics in R, Oxford University Press, Great Clarendon Street, Oxford, OX2 6DP, United Kingdom, 1 edn., 2013.

Philander, S. G.: El Niño, La Niña, and the Southern Oscillation., no. 46 in International Geophysics Series, Academic Press, San Diego, California 92101, US, 1990.

Powell, M. J. D.: A Hybrid Method for Nonlinear Equations, in: Numerical Methods for Nonlinear Algebraic Equations, edited by Rabinowitz, P., chap. 6, pp. 87–114, Gordon and Breach Science Publishers, 150 Fifth Avenue, New York, N.Y. 10011, U.S., 1970.

Powell, M. J. D.: On nonlinear optimization since 1959, in: The Birth of Numerical Analysis, edited by Bulthee, A. and Cools, R., pp. 141–160, World Scientific, 5 Toh Tuck Link, Singapore 596224, https://doi.org/10.1142/9789812836267_0009, 2009.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/, 2023.

Rios Gaona, M. F. and Villarini, G.: Characterization of the diurnal cycle of maximum rainfall in tropical cyclones, J Hydrol, 564, 997–1007, https://doi.org/https://doi.org/10.1016/j.jhydrol.2018.07.062, 2018.

Ross, S. M.: Simulation, Academic Press, Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands, 5 edn., https://doi.org/10.1016/C2011-0-04574-X, 2013.

Sarachik, E. S. and Cane, M. A.: The El Niño-Southern Oscillation Phenomenon, Cambridge University Press, The Edinburgh Building, Cambridge CB2 8RU, UK, https://doi.org/10.1017/CBO9780511817496, 2010.

Sargeant, C. I. and Singer, M. B.: Sub-annual variability in historical water source use by Mediterranean riparian trees, Ecohydrology, 9, 1328–1345, https://doi.org/10.1002/eco.1730, 2016.

Scanlon, B. R., Keese, K. E., Flint, A. L., Flint, L. E., Gaye, C. B., Edmunds, W. M., and Simmers, I.: Global synthesis of groundwater recharge in semiarid and arid regions, Hydrol Process, 20, 3335–3370, https://doi.org/10.1002/hyp.6335, 2006.

Seabold, S. and Perktold, J.: Statsmodels: Econometric and Statistical Modeling with Python, in: Proceedings of the 9th Python in Science Conference, edited by Stéfan van der Walt and Jarrod Millman, pp. 92–96, https://doi.org/10.25080/Majora-92bf1922-011, 2010.

Shmaliy, Y. S.: Von Mises/Tikhonov-based distributions for systems with differential phase measurement, Signal Process, 85, 693–703, https://doi.org/https://doi.org/10.1016/j.sigpro.2004.11.008, 2005.

Singer, M. B. and Michaelides, K.: How is topographic simplicity maintained in ephemeral dryland channels?, Geology, 42, 1091–1094, https://doi.org/10.1130/G36267.1, 2014.

Singer, M. B. and Michaelides, K.: Deciphering the expression of climate change within the Lower Colorado River basin by stochastic simulation of convective rainfall, Environ Res Lett, 12, 104 011, https://doi.org/10.1088/1748-9326/aa8e50, 2017.

Singer, M. B., Sargeant, C. I., Piégay, H., Riquier, J., Wilson, R. J. S., and Evans, C. M.: Floodplain ecohydrology: Climatic, anthropogenic, and local physical controls on partitioning of water sources to riparian trees, Water Resour Res, 50, 4490–4513, https://doi.org/10.1002/2014WR015581, 2014.

Singer, M. B., Michaelides, K., and Hobley, D. E. J.: STORM 1.0: a simple, flexible, and parsimonious stochastic rainfall generator for simulating climate and climate change, Geosci

Model Dev, 11, 3713–3726, https://doi.org/10.5194/gmd-11-3713-2018, 2018.

Stillman, S., Zeng, X., Shuttleworth, W. J., Goodrich, D. C., Unkrich, C. L., and Zreda, M.: Spatiotemporal Variability of Summer Precipitation in Southeastern Arizona, J Hydrometeorol, 14, 1944–1951, https://doi.org/10.1175/JHM-D-13-017.1, 2013.

Syed, K. H., Goodrich, D. C., Myers, D. E., and Sorooshian, S.: Spatial characteristics of thunderstorm rainfall fields and their relation to runoff, J Hydrol, 271, 1–21, https://doi.org/10.1016/S0022-1694(02)00311-6, 2003.

Temme, N.: On the numerical evaluation of the modified bessel function of the third kind, J. Comput. Phys., 19, 324–337, https://doi.org/https://doi.org/10.1016/0021-9991(75)90082-0, 1975.

The Economist: In defense of the Gaussian copula, electronic periodical, https://www.economist.com/free-exchange/2009/04/29/in-defense-of-the-gaussian-copula, accessed on: 2022-10-05, 2009.

Tjøstheim, D., Otneim, H., and Støve, B.: Statistical Modeling Using Local Gaussian Approximation, Academic Press, 125 London Wall, London EC2Y 5AS, United Kingdom, https://doi.org/10.1016/C2017-0-02646-0, 2022.

Tucker, G. E. and Bras, R. L.: A stochastic approach to modeling the role of rainfall variability in drainage basin evolution, Water Resour Res, 36, 1953–1964, https://doi.org/10.1029/2000WR900065, 2000.

Tucker, G. E. and Hancock, G. R.: Modelling landscape evolution, Earth Surf Proc Land, 35, 28–50, https://doi.org/10.1002/esp.1952, 2010.

Tucker, G. E. and Slingerland, R.: Drainage basin responses to climate change, Water Resour Res, 33, 2031–2047, https://doi.org/10.1029/97WR00409, 1997.

Van Rossum, G. and Drake, F. L.: Python 3 Reference Manual, CreateSpace, Scotts Valley, CA, https://dl.acm.org/doi/book/10.5555/1593511, 2009.

Vandenberghe, S., Verhoest, N. E. C., Onof, C., and De Baets, B.: A comparative copula-based bivariate frequency analysis of observed and simulated storm events: A case study on Bartlett-Lewis modeled rainfall, Water Resour Res, 47, https://doi.org/10.1029/2009WR008388, 2011.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nat. Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Wheater, H. S., Mathias, S. A., and Li, X.: Groundwater Modelling in Arid and Semi-Arid Areas, Cambridge University Press, https://doi.org/10.1017/CBO9780511760280, 2010.

Zhang, L. and Singh, V. P.: Copulas and their Applications in Water Resources Engineering, Cambridge University Press, University Printing House, Cambridge CB2 8BS, United Kingdom, https://doi.org/10.1017/9781108565103, 2019.