Modeling Intensity-Duration-Frequency curves for the whole range of precipitation: A comparison of models

Abubakar Haruna [1], Juliette Blanchet [2], Anne-Catherine Favre [1]

[1]Univ. Grenoble Alpes, Grenoble INP, CNRS, IRD, IGE, 38000 Grenoble, France
[2]Univ. Grenoble Alpes, CNRS, IRD, Grenoble INP, IGE, 38000 Grenoble, France

Key Points:

- We build IDF curves using all the non-zero precipitation data and model the intensities with the Extended Generalized Pareto Distribution (EGPD).
- We consider three approaches to building IDF curves: scale invariance, the general IDF formulation of Koutsoyiannis et al. (1998) and a data-driven method.
- We compare and select the best model in a cross-validation framework.

Corresponding author: Abubakar Haruna, abubakar.haruna@univ-grenoble-alpes.fr

Abstract

Intensity-Duration-Frequency curves are useful in water resources engineering for the planning and design of hydrological structures. As opposed to the common use of only extreme data to build IDF curves, here, we use all the non-zero rainfall intensities, thereby making efficient use of the available information. As a parametric model, we use the Extended Generalized Pareto Distribution (EGPD) for the non-zero intensities. We consider three commonly used approaches to build the IDF curves. The first approach is based on the scale-invariance property of rainfall, the second relies on the general IDF formulation of Koutsoyiannis et al. (1998) while the last approach is purely data-driven (Overeem et al., 2008). Using these three approaches, and some extensions around them, we build a total of 10 models for the IDF curves and then we compare them in a split-sampling cross-validation framework. We consider a total of 81 stations at 10 min resolution in Switzerland. The results reveal the model based on the data-driven approach as the best model. It is able to correctly model the observed intensities across duration while being reliable and robust. It is also able to reproduce the space and time variability of extreme rainfall across Switzerland.

1 Introduction

Intensity-Duration-Frequency (IDF) curves provide the link between precipitation intensity, duration, and non-exceedance frequency (or rather the return period). It is a very common and useful tool in the area of water resources engineering. IDF curves are practically used to infer high return levels of rainfall intensities for the hydrological designs of structures such as sewer lines, culverts, drains, dams, dykes, etc. They are also used to calibrate/validate stochastic weather generators (Willems, 2000; Ritschel et al., 2017).

IDF curves are traditionally modeled by fitting a statistical model, *e.g* a Gumbel distribution, to extreme data of each duration separately, Secondly, selected return levels, e.g. 2, 5, 10 years are obtained for each duration using the fitted distribution. And lastly, the inferred return levels are linked to duration by some empirical formulation (e.g. Sherman, 1931; Bernard, 1932; Chow, 1962; Carreteras, 1987; Meylan et al., 2012). Although common and easy to implement, there are several drawbacks to this approach. The method lacks parsimony because several parameters have to be fitted (a set for each return period). It is not robust in the sense that IDF curves are only available for specific return levels, each time a new return level is needed, the process has to be repeated. Another major drawback is that uncertainty in the return levels obtained at the initial steps is not taken into account in the last step. Lastly, there can be intersections between curves of different return levels that cannot be theoretically justified.

To overcome the outlined limitations of the traditional parametric methods, novel approaches were considered to link the different durations together in IDF curves. In general, in spite of the approach, two choices have to be made: the specific form of the IDF curves and the parametric model for the rainfall intensities. Regarding the specific form of the IDF curves, many formulations that are based on different approaches have been proposed in the literature. Here we identify and focus on three major approaches.

The first approach involves the general formulation of Koutsoyiannis et al. (1998), a generalization of the various empirical formulations for modeling IDF. This formulation has the key advantage of being a separable function of return levels and duration. It is also consistent with both probabilistic theories and the physical constraints of scaling across duration. Several application of this formulation to build IDF curves can be found in the literature (e.g. Koutsoyiannis et al., 1998; Van de Vyver & Demarée, 2010; Blanchet et al., 2016; Sane et al., 2018; Ulrich et al., 2020; Fauer et al., 2021; Roksvåg et al., 2021).

The second approach is based on scale invariance. It has been shown that rainfall exhibits this property within some scales (see Schertzer & Lovejoy, 1987; Gupta & Waymire, 1990, 1993; Over, 1995; Harris et al., 1997; Lima, 1998; Molnar & Burlando, 2008; Veneziano & Lepore, 2012; Paschalis, 2013). This property provides the physical justification for modeling IDF, and thus the possibility of inferring return levels of interest across scales. This approach is arguably the most commonly used approach, possibly because of its rich theoretical background, physical basis, and ease of application in regions with scarce availability of sub-daily rainfall series. IDF curves based on this approach can be found in several applications (Burlando & Rosso, 1996; Menabde et al., 1999; Willems, 2000; Van de Vyver & Demarée, 2010; Blanchet et al., 2016; Innocenti et al., 2017; Sane et al., 2018).

The last approach is termed by Overeem et al. (2008) as data-driven. The method involves fitting a parametric model, for example, GEV, to data of each duration. A particular regression model is then fitted for each parameter as a function of duration. As a consequence, the return level of any duration can be inferred from the inverse of the distribution, with parameters obtained from the regression model. This approach imposes neither the assumption/existence of scaling nor the separability condition in the case of the general formulation of Koutsoyiannis et al. (1998). Interestingly, both approaches can be seen as data-driven approaches with particular functional relationships imposed on the parameters.

There exists also nonparametric approaches, which rather than imposing a parametric model on the intensities, use stochastic rainfall models to estimate the IDF curves (for a brief review, see Langousis & Veneziano, 2007; Veneziano et al., 2007; Tyralis & Langousis, 2019). Here, we focus on the class that uses parametric models for the intensities.

Coming back to the choice of parametric model for the intensities, extreme value distributions are usually used. For example, Generalized Extreme Value (GEV) (e.g. Blanchet et al., 2016; Innocenti et al., 2017; Van de Vyver, 2018; Sane et al., 2018; Mélèse et al., 2018; Ulrich et al., 2020; Jurado et al., 2020; Fauer et al., 2021), Gumbel (Yu et al., 2004; Agbazo et al., 2016; Chang et al., 2016; Ghanmi et al., 2016) for annual maximum series or Generalized Pareto Distribution (GPD) for the peaks over thresholds (e.g. Madsen et al., 1995; Ben-Zvi, 2009; Van de Vyver & Demarée, 2010).

Although the use of these distributions is justified by the practical use of IDF, which is to infer high return levels for hydrological designs, there is however, a major drawback resulting from the poor utilization of already scarce data, and the delicate issue of threshold choice in the GPD case. To address these issues, Naveau et al. (2016) recently proposed the Extended Generalized Pareto Distribution (EGPD) to model all the non-zero precipitation. It has the advantage of using all the information present in the sample of non-zero rainfall data and not only one value per block (like GEV distribution) or only values above a given threshold (as in GPD distribution). It doesn't require the choice of GPD threshold and has the advantage of being compliant with extreme value theory in both the lower and upper tails. It has been used recently in many applications (e.g. Evin et al., 2018; Blanchet et al., 2019; Tencaliec et al., 2020; Rivoire et al., 2021; Le Gall et al., 2022; Haruna et al., 2022). However, to the best of our knowledge, it has never been used for modeling IDF curves.

The goal of this article is to use the EGPD to build IDF curves for the non-zero precipitation intensities, based on the three outlined approaches, *i.e.*, scale-invariance, the general formulation of Koutsoyiannis et al. (1998) and the data-driven approach. We will compare and select the best model based on a split-sampling cross-validation framework.

The rest of the paper is organized as follows: Section 2 introduces the data and the area understudy, and Section 3 presents the EGPD, the models for the IDF curves, and

the evaluation framework. Section 4 presents the results and discussion and finally Section 5 draws the conclusions and gives relevant perspectives.

2 Data and area under study

The study area is Switzerland, a small country by size with an area of 41,285 km^2. It however presents a complex topography with elevations ranging from 191 m to 4127 m above mean sea level. Around 30% of the area is located above the elevation of 1500 m above sea level. This results in marked spatial variability both in intensity and occurrence of precipitation.

Point precipitation data from a total of 81 stations, with a minimum record length of 20 years, are available for this study. They are spread across Switzerland and their location is shown in Fig. 1. Out of this total, 71 stations belong to the SwissMetNet of the Swiss Federal Office for Meteorology and Climatology (MeteoSwiss) while 10 belong to the canton of Lucerne, a partner network of MeteoSwiss. The precipitation data is measured with a tipping-bucket gauge of 0.1 mm depth resolution at a sampling resolution of 10 minutes. The tipping gauge is heated in order to account for snow. The sample data has a variable length ranging from a minimum of 20 years to a maximum of 40 years from 1981 to 2020. The stations are located at elevations ranging from a minimum of 203 m, an average of 952.4 m, and a maximum of 3294 m.

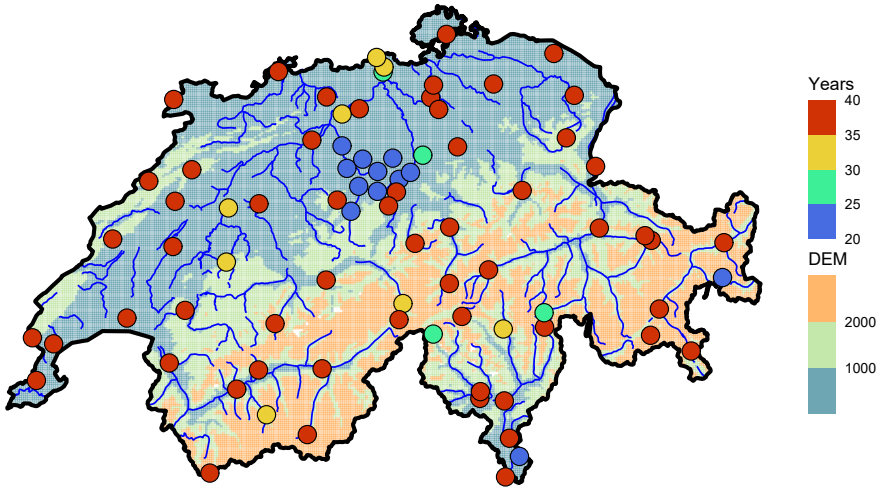

Figure 1. Map of Switzerland showing the location of the 81 stations. The color of the points indicates the length of the precipitation data in years. The background color shows the elevation above sea level in meters.

Due to the marked seasonality of precipitation in Switzerland, we divided the data into four seasons of three months each. Winter includes Dec-Jan-Feb, Spring Mar-Apr-May, Summer Jun-Jul-Aug while Autumn includes Sep-Oct-Nov. A similar seasonal approach was used by Molnar and Burlando (2008); Fukutome et al. (2015); Evin et al. (2018); Haruna et al. (2022), in the study area.

3 Methodology

In this section, we start by presenting the parametric model for the non-zero precipitation intensities, then we present the various IDF models, and finally the inference strategy to estimate the parameters.

3.1 Marginal distribution of non-zero precipitation intensities

As our target is to model the IDF curves using all the non-zero precipitation intensities, we choose the EGPD of Naveau et al. (2016) as the parametric model. The model is compliant with extreme value theory in both its upper and lower tails while providing a smooth transition in-between. It gives an alternative to the light-tailed distributions such as Gamma, which underestimates extremes (Katz et al., 2002). Four parametric families of this model have been proposed by Naveau et al. (2016) to model the transition, however, the simplest of the families is parsimonious and can adequately model precipitation intensities without the need for GPD threshold selection (see Evin et al., 2018; Le Gall et al., 2022; Haruna et al., 2022, for application in the study area). We, therefore, use it in our study.

If I is a random variable representing non-zero daily precipitation intensity that is distributed according to the EGPD, then the cumulative distribution function (CDF) is given by:

$$F(i) = \mathbb{P}(I \leq i) = G\left[H_\xi\left(\frac{i}{\sigma}\right)\right], \tag{1}$$

where G is any CDF that ensures a smooth transition between the EVT compliant upper and lower tails and satisfies the conditions given in Naveau et al. (2016), and:

$$H_\xi\left(\frac{i}{\sigma}\right) = \begin{cases} 1-(1+\xi\frac{i}{\sigma})_+^{-1/\xi} & \text{if } \xi \neq 0 \\ 1-\exp\left(-i/\sigma\right) & \text{if } \xi = 0 \end{cases}, \tag{2}$$

with $a_+ = \max(a, 0)$.

For the parsimonious model we use, the function G is simply defined as $G(v) = v^k$. Therefore the model is given as:

$$F(x) = \left[H_\xi\left(\frac{i}{\sigma}\right)\right]^k \tag{3}$$

The model thus has three parameters. $k > 0$ controls the lower tail, $\xi \geq 0$ controls the upper tail, and $\sigma > 0$ is the scale parameter.

3.2 IDF models

We define the random variable I_d as the average non-zero precipitation intensity over the duration d. It is described by the CDF, $F_d(i)$, such that $F_d(i) = \mathbb{P}(I_d < i)$. The exceedance frequency is defined as $p_d(i) = 1 - F_d(i)$. The return period of any non-zero intensity i, as a function of p_d is given by $T(I_d \geq i) = \frac{1}{p_d \times \delta_d}$, with δ_d the average number of non-zero precipitation intensities per year. Conversely, the T-year return level over duration d, $i(T, d)$, is defined as the $(1 - \frac{1}{T \times \delta_d})$ quantile of F_d.

Accordingly, IDF is a mathematical function $(T, d) \mapsto i(T, d)$ that relates non-zero rainfall intensity i with its duration d, and the frequency of exceedance p_d (or rather the return level T). In this article, the CDF of I_d, $F_d(i)$ is defined by the EGPD presented in Section 3.1. All the different formulations considered here simply differ by how they define this mathematical relationship between i, T and d, while taking $F_d(i)$ as an EGPD model.

In the following subsections, we present the different IDF-EGPD models based on the three outlined approaches, *i.e.*, scale-invariance, the general formulation of Koutsoyiannis

et al. (1998), and data-driven approaches. For sake of simplicity, we drop the "EGPD" term and simply refer to the IDF models as $\text{IDF}_{modelname}$, where the subscript "*modelname*" refers to the approach used to build the model.

For all the models, the IDF curves, corresponding to the $(1 - \frac{1}{T \times \delta_d})$ quantile of the EGPD is defined in Eq. 4. The choice of the model determines whether each of the three parameters; κ, σ, and ξ depends on d or not, and the form of the relationship.

$$i(T,d) = \frac{\sigma_d}{\xi_d} \left\{ \left(1 - \left[1 - \frac{1}{T \times \delta_d} \right]^{\frac{1}{\kappa_d}} \right)^{-\xi_d} - 1 \right\} \tag{4}$$

We consider thirteen durations, *i.e.*, d = 30 min, 40 min, 1, 2, 3, 6, 10, 12, 16, 18, 24, 48 and 72 hours. We use a fixed window to aggregate the data from the gauge resolution of 10 mins, to the various durations. For instance, the 24hr intensities correspond to amounts accumulated from 00h00 to 24h00 of every day, divided by 24.

3.2.1 Scaling IDF

Scale invariance in the strict sense of Gupta and Waymire (1990) refers to the property where the probability distribution of I_d can be inferred from the distribution of I_{d_0} at the reference duration d_0 through:

$$I_d \overset{dist}{=} C_\lambda I_{d_0}, \tag{5}$$

where the parameter C_λ determines the type of scaling; simple-scaling or multi-scaling.

A weaker assumption, the so-called "wide sense scaling" (Gupta & Waymire, 1990), is when the scaling is in the moments according to:

$$\mathbb{E}\left[I_d^q\right] = \left(\frac{d}{d_0}\right)^{-k(q)} \mathbb{E}\left[I_{d_0}^q\right], \tag{6}$$

where q is the order of the moment, $k(q)$ is called the moment scaling function, d_0 is the reference duration. Moment scaling analysis as described by Gupta and Waymire (1990) is used to determine the type of scaling.

Strict sense simple-scaling occurs when the scale ratio C_λ in Eq. 5 is a scalar that depends only on the ratio of the scales as expressed in Eq. 7.

$$I_d \overset{dist}{=} \left(\frac{d}{d_0}\right)^{-H} I_{d_0} \tag{7}$$

Wide sense simple-scaling is when the moment scaling function in Eq. 6 is linear in q, *i.e.* $k(q) = Hq$, as expressed in Eq. 8.

$$\mathbb{E}\left[I_d^q\right] = \left(\frac{d}{d_0}\right)^{-Hq} \mathbb{E}\left[I_{d_0}^q\right] \tag{8}$$

It can be shown that, under the strict sense simple-scaling, only one parameter of the EGPD is scaling, which is σ, whereas κ and ξ are independent of duration. For the rest of the paper, we drop the term "strict-sense", and simply use "simple-scaling" for convenience.

The simple-scaling EGPD model, IDF_{ss}, is defined such that: $\kappa_d = \kappa_{d_0}$, $\xi_d = \xi_{d_0}$, and σ_d is a power law given as:

$$\sigma_d = \left(\frac{d}{d_0}\right)^{-H} \sigma_{d_0} \tag{9}$$

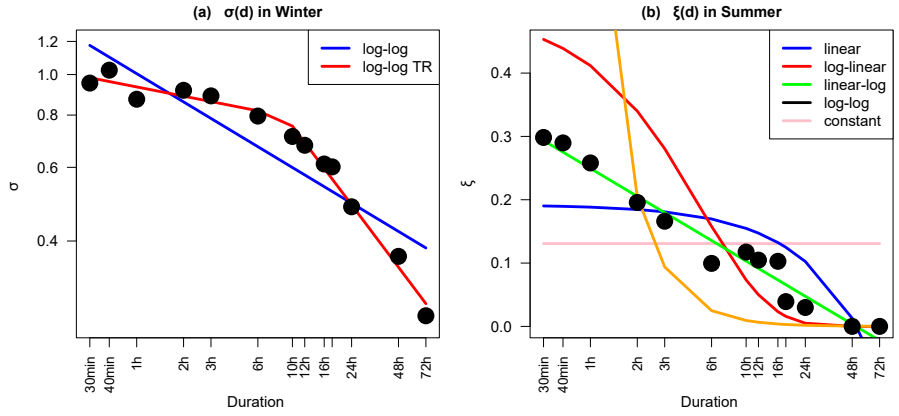

Figure 2. Illustration of **a)** Break in scaling of the σ parameter in winter at a station, Robbia in Graubünden. The points colored in black are the estimated σ for each duration separately. **b)** Dependence of ξ on duration in summer at a station in Zurich. The black colored points are the estimated ξ for each duration separately. The lines are the fitted linear models. The pink line is the mean of the estimated ξ.

An important issue is the existence of multiple scaling regimes in precipitation. This means that different scaling behaviors (scaling exponents) exist for different ranges of duration. IDF then have to be modeled considering the existence of this change in scaling (e.g. Yu et al., 2004; Bougadis & Adamowski, 2006; Courty et al., 2019). An illustration of this behavior is given in Fig. 2a for the σ parameter in winter at a station, Robbia in Graubünden. Here, a single power law (log-log given in Eq. 9) in blue is not enough to explain the scaling. The two regime model (log-log TR in Eq. 10) in red is necessary.

To account for this break in the scaling relationship, we define the two-regime simple-scaling EGPD IDF model, IDF_{ss_TR} as:

$$\sigma_d = \begin{cases} \left(\frac{d}{d_0}\right)^{-H_1} \sigma_{d_0} & \text{if } d \leq K, \\ \left(\frac{d}{d_0}\right)^{-H_2} \sigma_{d_0} \times K^{H_2 - H_1} & \text{if } d > K, \end{cases} \tag{10}$$

where K is the duration of the scaling break, and σ is continuous in $d = K$. The other parameters, H_1 and H_2 are the scaling exponents of the first and second regimes. The other two parameters, κ, and ξ remain independent of durations. Hence for this model, a total of six parameters have to be estimated, *i.e.*, κ_{d_0}, σ_{d_0}, ξ_{d_0}, H_1, H_2 and K.

Lastly, although the simple-scaling EGPD model imposes a constraint on the dependence of ξ with d, i.e, $\xi_d = \xi_{d_0}$, we however notice some of the stations to show apparent dependence of ξ with respect to d. To account for this, we impose a functional relationship of ξ with respect to duration through a linear-log form as expressed in Eq. 11. Fig. 2b gives an example of the dependence of ξ on duration at a station in Zurich, and how the linear-log model fits correctly the points.

$$\xi_d = a_\xi + b_\xi \log(d) \tag{11}$$

Where a_ξ and b_ξ are the intercepts and slopes respectively. This leads to two additional IDF models, with $\xi = f(d)$, namely:

- $\text{IDF}_{ss\_\xi(d)}$: an extension of the basic simple-scaling model IDF_{ss}, to allow ξ to depend on d according to Eq. 11.
- $\text{IDF}_{ss\_TR\_\xi(d)}$: an extension of the two-regime simple-scaling model IDF_{ss_TR}, to allow ξ to depend on d according to Eq. 11.

3.2.2 General IDF formulation

Koutsoyiannis et al. (1998) proposed a general formulation for the different traditional formulations of the IDF curves in the literature. He showed that all of them can be simplified into the form:

$$i(T,d) = \frac{a(T)}{b(d)} \tag{12}$$

where $b(d) = (d+\theta)^H$. The parameter θ is the duration offset, and H is the duration exponent. Both θ and H are non-negative constants. *a(T)* is the $(1-\frac{1}{T})$ quantile of the re-scaled intensities $I_d b(d)$. *a(T)* is independent of d and completely determined by the statistical model considered for I_d, in our case, the EGPD.

This formulation has the key advantage of being a separable function of return levels *a(T)*, and duration *b(d)* that is consistent with both probabilistic theories and the physical constraints of scaling across duration. Menabde et al. (1999) showed that this formulation is the same as the scale-invariant model θ is set to zero.

When applied to the EGPD, IDF_{koust} is defined such that: $\kappa_d = \kappa_{d_0} \quad \sigma_d = \left(\frac{d+\theta}{d_0+\theta}\right)^{-H} \sigma_{d_0}$ $\xi_d = \xi_{d_0}$. Five parameters, κ, σ, ξ, θ and H have to be inferred.

Following the same arguments discussed in Section 3.2.1 regarding the existence of a break in the scaling relationship, and the dependence of ξ with d, we propose three extensions to this model:

- IDF_{koust_TR}: Allowing for a break in the scaling regime. This model is defined as:

$$\sigma_d = \begin{cases} \left(\frac{d+\theta}{d_0+\theta}\right)^{-H_1} \sigma_{d_0} & \text{if } d \leq K \\ \left(\frac{d}{d_0}\right)^{-H_2} \sigma_{d_0} \times K^{H_2-H_1} & \text{if } d > K \end{cases} \tag{13}$$

  where κ and ξ are independent of duration.
- $\text{IDF}_{koust\_\xi(d)}$: an extension of the basic model IDF_{koust}, to allow ξ to depend on d according to Eq. 11.
- $\text{IDF}_{koust\_TR\_\xi(d)}$: an extension of the two-regime model IDF_{koust_TR}, to allow ξ to depend on d according to Eq. 11.

3.2.3 Data-Driven IDF

The scaling theory and the specific form of Eq. 12 impose particular functions for the relation between the scale parameter, σ of the EGPD with respect to duration, d. However, in the case of the data-driven models, the expression of the relationship for each of the three EGPD parameters is entirely determined by the data itself. To guide our choice of the appropriate functional relationship, we inspected how each locally estimated parameter varies with duration. Fig. 2 gives an example for the σ and ξ parameters at two stations. We finally settled on the following functions to model the three parameters with respect to duration:

$$\kappa_d = \begin{cases} \exp\left[a_\kappa + b_{1,\kappa} \log\left(d\right)\right] & \text{if } d \leq K_\kappa \\ \exp\left[a_\kappa + b_{2,\kappa} \log\left(d\right) + \left(b_{1,\kappa} - b_{2,\kappa}\right) \log\left(K_\kappa\right)\right] & \text{if } d > K_\kappa \end{cases} \tag{14}$$

$$\sigma_d = \begin{cases} \exp\left[a_\sigma + b_{1,\sigma} \log\left(d\right)\right] & \text{if } d \leq K_\sigma \\ \exp\left[a_\sigma + b_{2,\sigma} \log\left(d\right) + \left(b_{1,\sigma} - b_{2,\sigma}\right) \log\left(K_\sigma\right)\right] & \text{if } d > K_\sigma \end{cases} \tag{15}$$

For the first two parameters, κ and σ, the function is a continuous two linear piecewise model in log space. K_* is the duration of the breakpoint (σ continuous for $d = K$). a_*, $b_{1,*}$,$b_{2,*}$ are the intercepts and slopes of the first and second lines respectively. In the case of xi, the function os given in Eq. 11.

Note that, by keeping κ and σ independent of duration, and using either $\sigma_d = \left(\frac{d}{d_0}\right)^{-H} \sigma_{d_0}$ or $\sigma_d = \left(\frac{d+\theta}{d_0+\theta}\right)^{-H} \sigma_{d_0}$, the simple-scaling or the general formulations of Koutsoyiannis et al. (1998) presented in Section 3.2.1 and 3.2.2 respectively can be obtained from this data-driven approach.

We consider two IDF models in this class, both impose the same type of functional relationships (Eq. 14, 15, 11), but simply differ in the way the regression parameters are estimated. For the first model, we follow a two-step approach (as implemented by Overeem et al., 2008). First, we fit, for a given station, the EGPD on the data of each duration separately. Second, we fit for each fitted parameter, the chosen regression model as a function of duration. We call this model $\text{IDF}_{DD_{local}}$.

The second model involves a one-step global fitting procedure. We pool all the data from the different durations to estimate the best EGPD with parameters from Eq. 14 to 11), the duration being the covariate. We call this model $\text{IDF}_{DD_{global}}$.

Note again that both models have the same number of free parameters, *i.e.* 10, but differ in the inference strategy.

The different models compared in this study are summarized in Table 1.

Table 1. Summary of the IDF models that are compared in this study.

	Model	No. of Parameters	Name of Approach	Ref. Section
1	IDF_{ss}	4	Simple-scaling	3.2.1
2	IDF_{ss_TR}	6	Simple-scaling	3.2.1
3	$\text{IDF}_{ss\_\xi(d)}$	5	Extension of Simple-scaling	3.2.1
4	$\text{IDF}_{ss\_TR\_\xi(d)}$	7	Extension of Simple-scaling	3.2.1
5	IDF_{kouts}	5	Koutsoyiannis et al. (1998)	3.2.2
6	IDF_{kouts_TR}	7	Extension of Koutsoyiannis et al. (1998)	3.2.2
7	$\text{IDF}_{koust\_\xi(d)}$	6	Extension of Koutsoyiannis et al. (1998)	3.2.2
8	$\text{IDF}_{koust\_TR\_\xi(d)}$	8	Extension of Koutsoyiannis et al. (1998)	3.2.2
9	$\text{IDF}_{DD_{local}}$	10	Data-driven	3.2.3
10	$\text{IDF}_{DD_{global}}$	10	Data-driven	3.2.3

3.3 Inference

For the models of Koutsoyiannis et al. (1998), the authors have proposed two different estimation strategies; the so-called 'robust estimation', and the 'one-step least square method'. The robust estimation is a two-step procedure that involves the estimation of the parameters of $b(d)$, and then those of $a(T)$ (see Eq. 12), through the minimization of the Kruskal-Wallis statistic. The one-step least square method involves the joint estimation of all the parameters of Eq. 12 that minimizes the squared error of the observed

and modeled quantiles from the IDF model. In the case of the simple-scaling models, a two-step procedure has been used (see Nhat et al., 2008; Panthou et al., 2014; Innocenti et al., 2017) where the scaling exponent in Eq. 7 is first obtained through moment scaling analysis, then all re-scaled intensities from all the durations are used to fit the IDF model.

In our case, however, we follow a global maximum likelihood estimation for all the models, as done by Blanchet et al. (2016). This involves pooling, for each station, all the data from the thirteen durations to estimate the model parameters. The duration d is used as a covariate. We note here that by pooling all the data, the dependence between the time steps and durations is neglected.

The log-likelihood (ll) that is maximized here (given in Eq. 16) takes left censoring into account. The importance of using left censoring in fitting rainfall data by maximum likelihood has been pointed out by Naveau et al. (2016), and he showed that better performance is obtained by taking it into account.

$$ll_{EGPD}(\kappa_d, \sigma_d, \xi_d) = ll_{censored}(\kappa_d, \sigma_d, \xi_d) + ll_{uncensored}(\kappa_d, \sigma_d, \xi_d), \tag{16}$$

where $ll_{censored}$ and $ll_{uncensored}$ are the contributions of the censored and uncensored data, given in Eq. 17 and 18 respectively, as

$$ll_{censored}(\kappa_d, \sigma_d, \xi_d) = \sum_d \sum_{j:i_d<c_d} \kappa \log \left[1 - \left(1 + \frac{\xi_d c_d}{\sigma_d}\right)^{-\frac{1}{\xi_d}}\right], \tag{17}$$

$$ll_{uncensored}(\kappa_d, \sigma_d, \xi_d) = \sum_d \sum_{j:i_d\geq c_d} \log \kappa_d - \sum_d \sum_{j:i_d\geq c_d} \log \sigma_d - \sum_d \sum_{j:i_d\geq c_d} \left[1 + \frac{\xi_d i_{d,j}}{\sigma_d}\right]^{\left[1+\frac{1}{\xi_d}\right]} + \sum_d \sum_{j:i_d\geq c_d} \left[1 - \left(\left(1 + \frac{\xi_d i_{d,j}}{\sigma_d}\right)^{-\frac{1}{\xi_d}}\right)\right]^{[\kappa_d - 1]}, \tag{18}$$

where c_d is the left censoring threshold applied to the data of duration d and the summation is done over the 13 durations. Many authors have taken this into account but they usually take a uniform threshold value for all the stations (e.g. Tencaliec et al. (2020) used 2 mm for daily rainfall). Here we didn't find the use of a common threshold over the 81 stations sufficient. We had to select, for each station and duration, the lower threshold c that minimizes the Normalized Root Mean Square Error (NRMSE) of Eq. 19 in Section 3.4.

In both Eq. 17 and 18, the choice of the IDF model specifies the function linking the EGPD parameters to duration. For instance, in the case of the simple-scaling , $\text{IDF}_{DD_{ss}}$, $\kappa_d = \kappa_{d_0} \quad \sigma_d = \left(\frac{d}{d_0}\right)^{-H} \sigma_{d_0} \quad \xi_d = \xi_{d_0}$.

For the specific case where two linear models are fitted (see Eq. 10, 13, 14, 15, 11), we start by estimating the break-point, K, using the functions in the `segmented` R package (Muggeo et al., 2008). The package implements the algorithm of Muggeo (2003) which, given the initial value of the break-point, estimates simultaneously the optimal break-point and the other parameters of the segmented linear model. The estimation is done through computational iteration of a reparameterized form of the standard segmented regression (see Muggeo, 2003, for details). We then plug the estimated value in the likelihood functions to estimate the parameters.

Finally, in the case of the $\text{IDF}_{DD_{local}}$, the parameters in Eq. 14 and 15 are estimated by segmented regression (described in the preceding paragraph), while those of Eq. 11 by least-squares.

3.4 Evaluation framework

We evaluate the performance of the models in two aspects. First, in calibration, that is how well a given model predicts the data that was used in training it. Secondly, we evaluate their predictive performance in a cross-validation framework.

3.4.1 Calibration

To evaluate the performance of the models in calibration, we compute the NRMS. The normalization, which here is done by the mean, allows the comparison of intensities of different duration across different stations.

For each station s, and duration d, we compute the NRMSE over the non-zero precipitation intensities as:

$$\text{NRMSE}_s(d) = \frac{\left\{\frac{1}{n_s(d)} \sum_{j=1}^{n_s(d)} \left(r_{s,T_j}(d) - \hat{r}_{s,T_j}(d)\right)^2\right\}^{1/2}}{\overline{r_s(d)}} \tag{19}$$

where $\text{NRMSE}_\text{s}(d)$ denotes the score computed at station s, and duration d, $n_s(d)$ is the number of non-zero precipitation intensities for duration d , $r_{s,T_j}(d)$ is the empirical quantile with return period T_j, $\hat{r}_{s,T_j}(d)$ is the corresponding T_j-year return level estimated from the fitted model. The denominator is the average rainfall at site s and duration d, given as $\frac{1}{n_s(d)} \sum_{j=1}^{n_s(d)} r_{s,T_j}(d)$. The best model according to this criteria is the one with the lowest $\text{NRMSE}_\text{s}(d)$.

3.4.2 Cross-validation

We follow a split sampling procedure in a cross-validation framework. For each station s, we divide the 10 min precipitation intensities into two equal sub-samples of the same length but on different years that are randomly chosen. We then aggregate the data into intensities of various duration, d = 30 min, 40 min, 1,2, 3, 6, 10, 12, 16, 18, 24, 48, 72 hours. Then we fit each of the 10 IDF models.

We then evaluate the performance of the models fitted on sub-sample 1 on the observations in sub-sample 2 and *vice versa*, by computing the relevant cross-validation criteria. We repeat this procedure 10 times.

In the following, we present three different criteria we use to evaluate the models. These criteria have been used in several studies to evaluate and compare competing models, (see Garavaglia et al., 2011; Renard et al., 2013; Blanchet et al., 2015; Evin et al., 2016; Haruna et al., 2022).

- The Robustness criteria, SPAN, measures the stability of the estimate of a high return level when the training data is changed. It is computed as:

$$\text{SPAN}_{s,T}(d) = \frac{2\left|\hat{r}^{(1)}_{s,T}(d) - \hat{r}^{(2)}_{s,T}(d)\right|}{\left(\hat{r}^{(1)}_{s,T}(d) + \hat{r}^{(2)}_{s,T}(d)\right)} \tag{20}$$

  where $\hat{r}^{(1)}_{s,T}(d)$ and $\hat{r}^{(2)}_{s,T}(d)$ are the T-year return levels estimated from sub-sample 1 and 2 respectively at station s and duration d.

For each duration d, a regional score over all the N stations ($N = 81$) is computed as $\text{SPAN}_{\text{reg},T}(d) = 1 - \frac{1}{N}\sum_{s=1}^{N}\text{SPAN}_{s,T}(d)$ and a perfect model in terms of robustness according to this criteria should have $\text{SPAN}_{\text{reg},T}(d) = 1$.

- The reliability of the model fitted on sub-sample 1 in predicting the maxima in sub-sample 2 and *vice versa* is measured by the FF criteria:

$$\text{FF}_s^{(12)}(d) = \left[\hat{\text{F}}_s^{(1)}(d)\left(\max_s^{(2)}(d)\right)\right]^{n_s^{(2)}(d)} \tag{21}$$

where $\text{FF}_s^{(12)}(d)$ is the cross-validation criteria computed at station s, and duration d, by predicting the probability of the maximum value in sub-sample 2, of sample size $n_s^{(2)}(d)$ using the model $\hat{\text{F}}_s^{(1)}(d)$ fitted on the sub-sample 1. $\text{FF}_s^{(21)}(d)$ is computed symmetrically.
For a given duration, Renard et al. (2013) and Blanchet et al. (2015) showed that each $FF_s^{(12)}(d)$ at a station should be a realization of a uniform distribution. So the difference in the area, $diff$ between a theoretical uniform distribution and that of the N set of $\text{FF}_s^{(12)}(d)$ (computed over the N stations), should be close to zero. $\text{FF}_{\text{reg}}(d)$ at the regional scale, given as $1 - diff$, should therefore take a value of 1 for a reliable model; the lower the value the less reliable the model is.

- The reliability/accuracy of the model in predicting the entire observations in cross-validation is measured by the NRMSE_CV.

$$\text{NRMSE\_CV}_s^{(12)}(d) = \frac{\left\{\frac{1}{n_s^{(2)}(d)}\sum_{j=1}^{n_s^{(2)}(d)}\left(r_{s,T_j}^{(2)}(d) - \hat{r}_{s,T_j}^{(1)}(d)\right)^2\right\}^{1/2}}{\overline{r_s^{(2)}(d)}} \tag{22}$$

where $\text{NRMSE\_CV}_s^{12}(d)$ is the score computed at station s, and duration d, $n_s^{(2)}(d)$ is the sample size, $r_{s,T_j}^{(2)}(d)$ is the empirical quantile with return period T_j in sub-sample 2 , $\hat{r}_{s,T_j}^{(1)}(d)$ is the corresponding T_j return level estimated from $\hat{\text{F}}_s^{(1)}(d)$. The denominator is the average daily rainfall in sub-sample 2 at site s given as $\frac{1}{n_s^{(2)}(d)}\sum_{j=1}^{n_s^{(2)}(d)} r_{s,T_j}^{(2)}(d)$.
Finally, for each duration d, the regional score computed over the N stations is given as: $\text{NRMSE\_CV}_{\text{reg}}^{(12)}(d) = 1 - \frac{1}{N}\sum_{s=1}^{N}\text{NRMSE\_CV}_s^{(12)}(d)$.
$\text{NRMSE\_CV}_{\text{reg}}^{(21)}(d)$ is computed in similar way. $\text{NRMSE\_CV}_{\text{reg}} = 1$ means a perfect model, and the closer the value is to 1, the more accurate the model is.

4 Results and discussion

We present the results in the following order: first we investigate the appropriateness of the EGPD to fit the data of each duration. Then we present the results of the comparison of the IDF models in calibration, and then in cross-validation. Finally, we show some IDF curves modeled with the best IDF.

4.1 Assessment of EGPD goodness of fit

The first issue is to investigate whether EGPD is an appropriate model for the precipitation data at hand. To check this, we fitted the model at each station and for each duration, independently. We call this EGPD model fitted on each data separately as the "base" model. We then assess the quality of the resulting fits by computing the NRMSE given in Eq. 19. The seasonal boxplots of the score for each duration are shown in Fig. 3. The higher the score, the better the model.

In spring and summer, the quality of the fit is less good for durations lower than 2 hours. In winter, on the other hand, the fit is less good for $d = 48$ and 72 hours. Over-

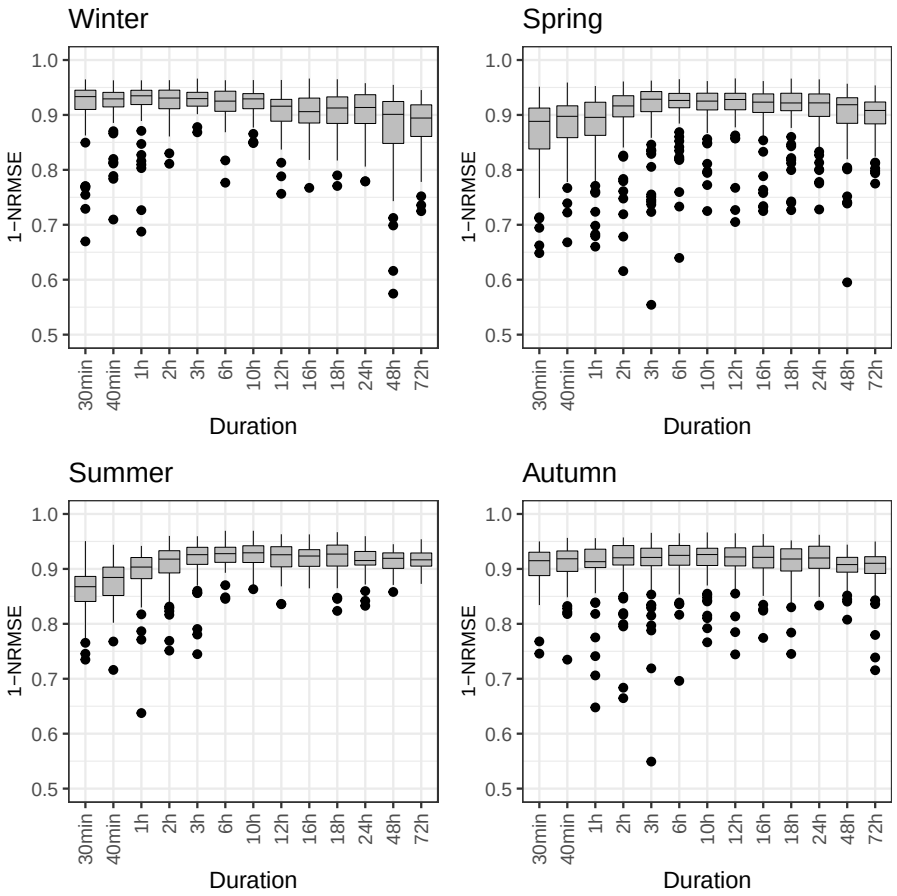

Figure 3. Boxplots of $1 - \text{NRMSE}_s(d)$ versus duration for the base EGPD model, *i.e.*, fitted on data of each duration separately. Each boxplot contains 81 points, with each point corresponding to one station.

all, more than 74% of the scores fall above 0.9 and 96% above 0.8. We, therefore, consider the EGPD to be a reasonable model for the data.

The fitted shape parameter ξ with respect to duration is shown on Fig. 4. Each boxplot contains 81 values, one for each station. We can strong dependence this parameter on duration, especially in summer. For this season, while 75% of the stations have a $\xi > 0.17$ for $d = 1$hr, only 25% have $\xi > 0.06$ at $d = 24$hr. In winter, however, the dependence is not very strong, as judged by the large variability of the boxplots.

4.2 Comparison of models

Results of the model comparison are presented under two frameworks, first in calibration, and then secondly in cross-validation based on split sampling.

4.2.1 Evaluation in calibration

Figure 5 presents the seasonal boxplots of the 1-NRMSE for the 10 IDF models and the base model. Each of the boxplot contains 1053 points, summarizing the score over 81 stations and 13 durations. In the case of the base model (in yellow), the scores are the same as those in Fig. 3, but here we merge the scores for all the durations together.

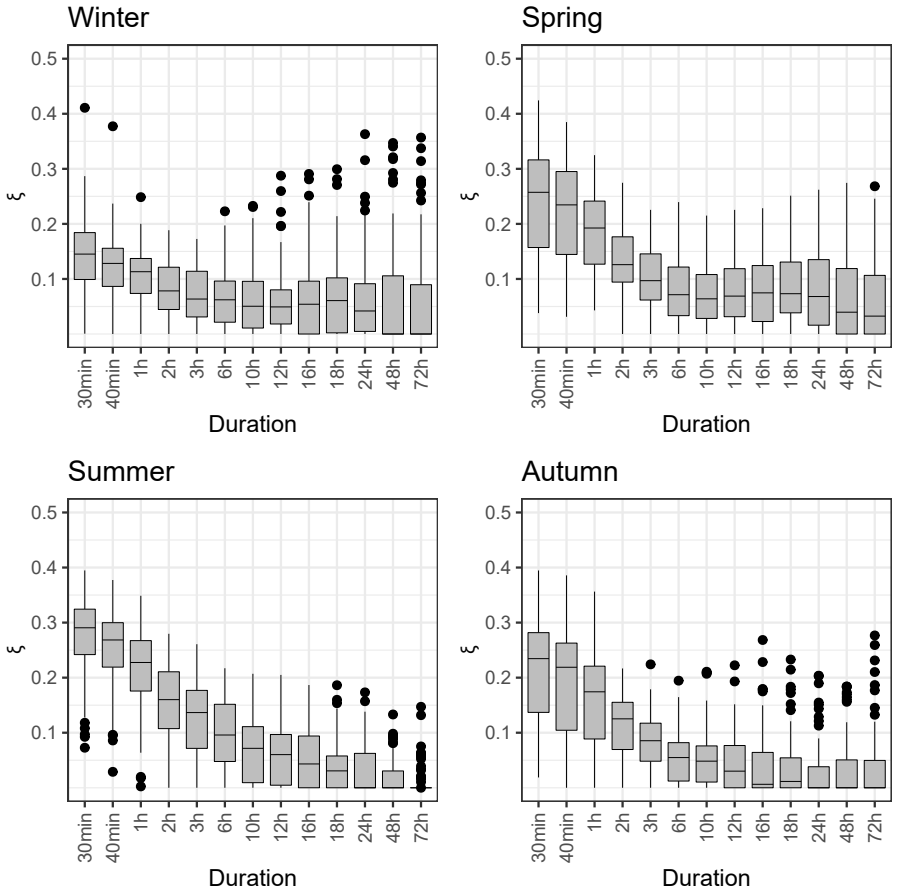

Figure 4. Boxplots of the fitted ξ versus duration obtained with the base model. Each boxplot contains 81 points, with each point corresponding to one station. The closer to 1, the better the model.

For all seasons, the two data-driven IDF models, $\text{IDF}_{DD_{local}}$ and $\text{IDF}_{DD_{global}}$ always show the best performance compared to the others. When looking at the two, the $\text{IDF}_{DD_{global}}$ generally outperforms the $\text{IDF}_{DD_{local}}$. This means that the global fitting of the model improves the estimation performance compared to the simple interpolation of the locally estimated parameters.

Comparing the IDF_{ss} and the IDF_{kouts} (white *vs* red boxplots), the results show that for all seasons, the IDF_{kouts} has a better performance compared to the IDF_{ss}. Recall that the two models differ by the additional parameter θ in the former to account for curvature for short durations.

Allowing for $\xi = f(d)$ increased the performance of the models mainly in summer, where all the models without this addition showed very poor performance. For the other seasons, the gain in performance is not as pronounced.

Lastly, the models allowing for scaling break (those with subscript $\_TR$), show improved performance compared to those with the single regime for all the seasons, except summer (*e.g.* IDF_{ss} vs IDF_{ss_TR}, *i.e.* the white and violet boxplots).

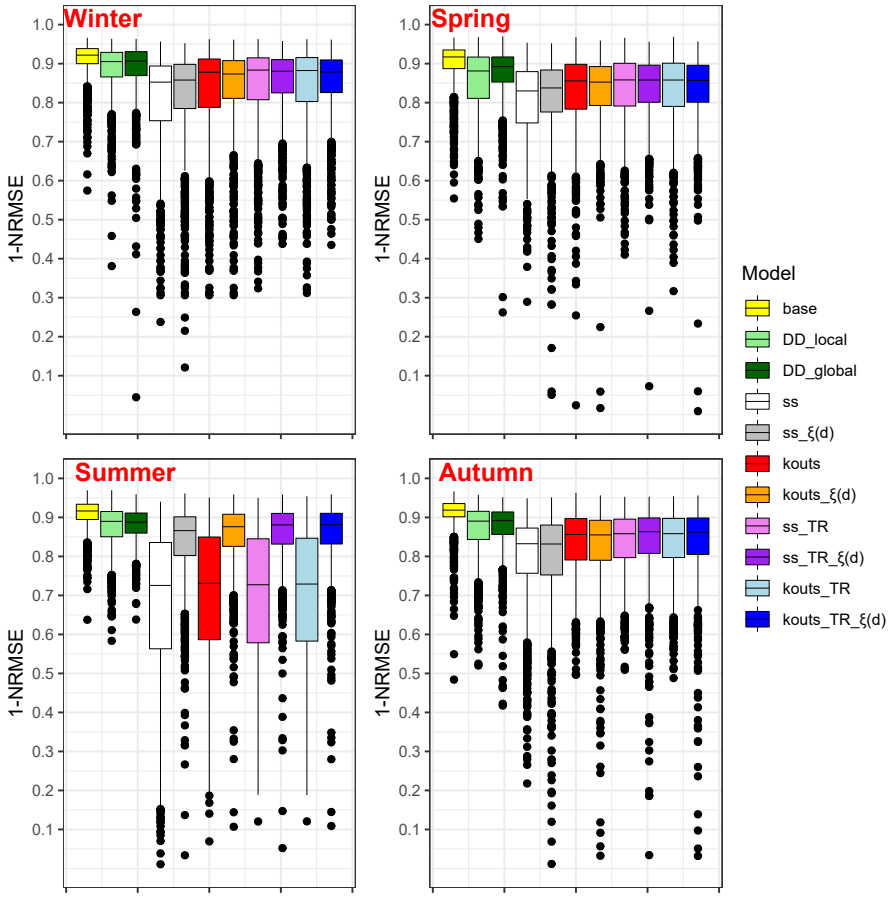

Figure 5. Boxplots of the (1-NRMSE) in calibration. Each boxplot contains 1053 points, each point corresponding to one station and duration.

4.2.2 Evaluation in cross-validation

The split-sampling procedure allows for the comparison of the models in a cross-validation framework. We use three regional criteria: NRMSE_CV, FF, and SPAN (see Section 3.4.2), to enable the comparison of the models based on their predictive capabilities. We want to select a model, which in addition to being able to fit the data used to train it, is able to perform reliably and robustly in the presence of new data.

In the following, we present the results in three paragraphs, first according to the reliability/accuracy of the model in predicting all the observations as measured by NRMSE_CV, then the reliability in predicting the maxima as measured by the FF criterion, and lastly, the robustness of the model in predicting the 100-year return level as measured by SPAN100. Figure 6 presents the results for the four seasons. For all the criteria, the model with a regional score equal to 1 is the best model.

For all seasons, the NRMSE_CV shows the data-driven models, specifically the $\text{IDF}_{DD_{global}}$, to be the most accurate/reliable in predicting the entire observations compared to the other models. In winter, however, the difference in the performance of the models is not very clear. Looking at the summer results, the models without accounting for $\xi = f(d)$ always have the worst performance.

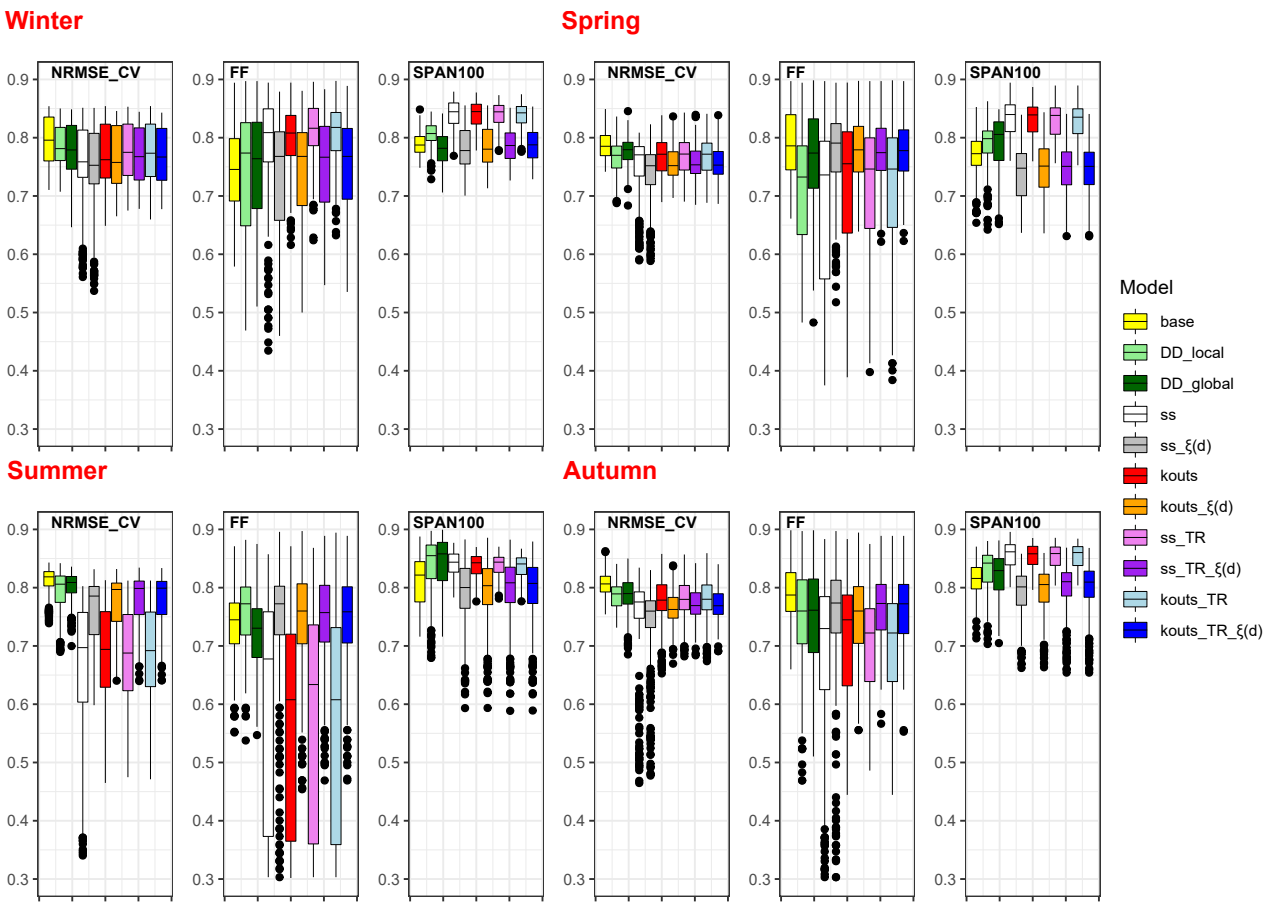

Figure 6. Boxplots of the regional cross-validation criteria, NRMSE_CV, FF, and SPAN. For the first two criteria, each boxplot contains 2 × 130 points, corresponding to one regional score for each of the 13 durations and 10 repetitions of the split-sampling. For the SPAN, each boxplot contains 130 points. The optimal value for each criterion is equal to 1.

In terms of the FF criterion, the best performance in predicting the maxima in winter is shown by the IDF_{ss_TR} model. In fact, all the models with no allowance for $\xi = f(d)$ happens to be the most reliable models in this season. The converse is however true in the case of the remaining seasons. In summer, while $\text{IDF}_{DD_{local}}$, is the best model, $\text{IDF}_{ss\_\xi(d)}$ is the best in spring and autumn.

The robustness criteria, SPAN100 shows the models with no allowance for $\xi = f(d)$ to be the most robust models. An exception to this is in summer, where the $\text{IDF}_{DD_{global}}$ model is the most robust model. Also, higher robustness is found for the models not accounting for $\xi = f(d)$ compared to their counterparts, for example IDF_{ss} *vs* $\text{IDF}_{ss\_\xi(d)}$. This is despite the fact that the former performs poorly in calibration, and is the least performing according to the other cross-validation criteria of reliability. This confirms the previous comments of Garavaglia et al. (2011) that a robust model can completely fail to model/predict the data. Hence the robustness criteria should only be used alongside other reliability criteria, such that the most robust model is only selected among models of similar reliability.

To summarize the results, the best IDF model should perform well in calibration, and should not be very sensitive to the data used to train it. In calibration, the data-

driven model $\text{IDF}_{DD_{global}}$ showed the best performance compared to all the other nine models, it also remain accurate and reliable at predicting the entire observations in the split-sampling cross-validation (as measured by the NRMSE_CV), especially in summer. This is an important feature since we are interested in the complete range of intensities. Finally, it generally showed more robustness compared to the other models of similar reliability.

4.3 IDF curves

Figure 7a shows the IDF curves from two models, IDF_{ss} and $\text{IDF}_{DD_{global}}$, in summer, at a station in Zurich which is located in the Northeast of Switzerland. In this region, summer is the main season of heavy rainfall. As a reminder, the IDF_{ss} allows scaling only in the scale parameter, σ of the EGPD, the other two parameters (κ and ξ), are independent of duration. The $\text{IDF}_{DD_{global}}$ on the other hand allows each of the three parameters to vary with duration. The curves are for return periods $T = 2$, 5, 10, 40, and 100 years, while points are the empirical levels for $T = 2$, 5, and 10 years.

The IDF_{ss} performed poorly at predicting the empirical quantiles. The curves modeled by the $\text{IDF}_{DD_{global}}$ on the other hand are in agreement with the empirical levels.

Similar IDF curves for autumn are shown in Fig. 7b for a station in Locarno which is located in the Ticino area in the south of Switzerland. The Ticino area is subject to the heaviest precipitation compared to the other regions in Switzerland. Again, the $\text{IDF}_{DD_{global}}$ was able to model empirical levels correctly both for the small and long durations.

In Fig. 7, the curves of the simple-scaling model (IDF_{ss}) are not parallel. This behavior resulted from the definition of IDF models for non-zero precipitation in Eq. 4. From this equation, we see that the T-year return level is defined as $(1 - \frac{1}{T \times \delta_d})$. The term δ_d, representing the average number of non-zero precipitation varies across the durations leading to a non-constant slope for the different curves.

We finally show, in Fig 8 and 9 respectively, the seasonal 100-year return level maps for $d =$ 1hr and 24h. The levels were obtained with the best performing model, *i.e.* $\text{IDF}_{DD_{global}}$. Looking at the return levels for d = 1hr (Fig. 8), we see that the levels in winter are the lowest, with no specific spatial pattern or variability. In spring, the levels in the north and Ticino starts to increase. Summer has the highest levels, and similar levels are obtained all along the northern plateau. In autumn, while the levels in the north are comparable to those in winter, those in Ticino are comparable to those in summer. A different spatial pattern is however observed for the 100-year return level at d =24hr. Specifically for summer, the levels in the plateau are lower than those along the north regime. The exhibited spatial pattern of the levels produced by this model is similar to those observed in earlier studies (see Fukutome et al. (2015) for the hourly, and Haruna et al. (2022) for the 24hr precipitation).

4.4 Discussion

In the following paragraphs, we briefly discuss some of our choices in terms of the functional forms of the data-driven models, taking into account the varying shape parameter with duration, and the issue of correlation in the data.

First, for the the data-driven models, we limited our choice of the functional relationships to simple parametric models, specifically to piece-wise linear models. Other choices would be possible such as smooth regression splines (e.g. Youngman, 2019, 2020). This choice has its advantage and drawback. The advantage is that the splines are able to automatically adjust to fit any form of relationship. The main drawback is that it is inherently non-parametric, and so the mapping of the IDF models, to allow predictions at ungauged locations, is not directly possible. One can only map the three EGPD pa-

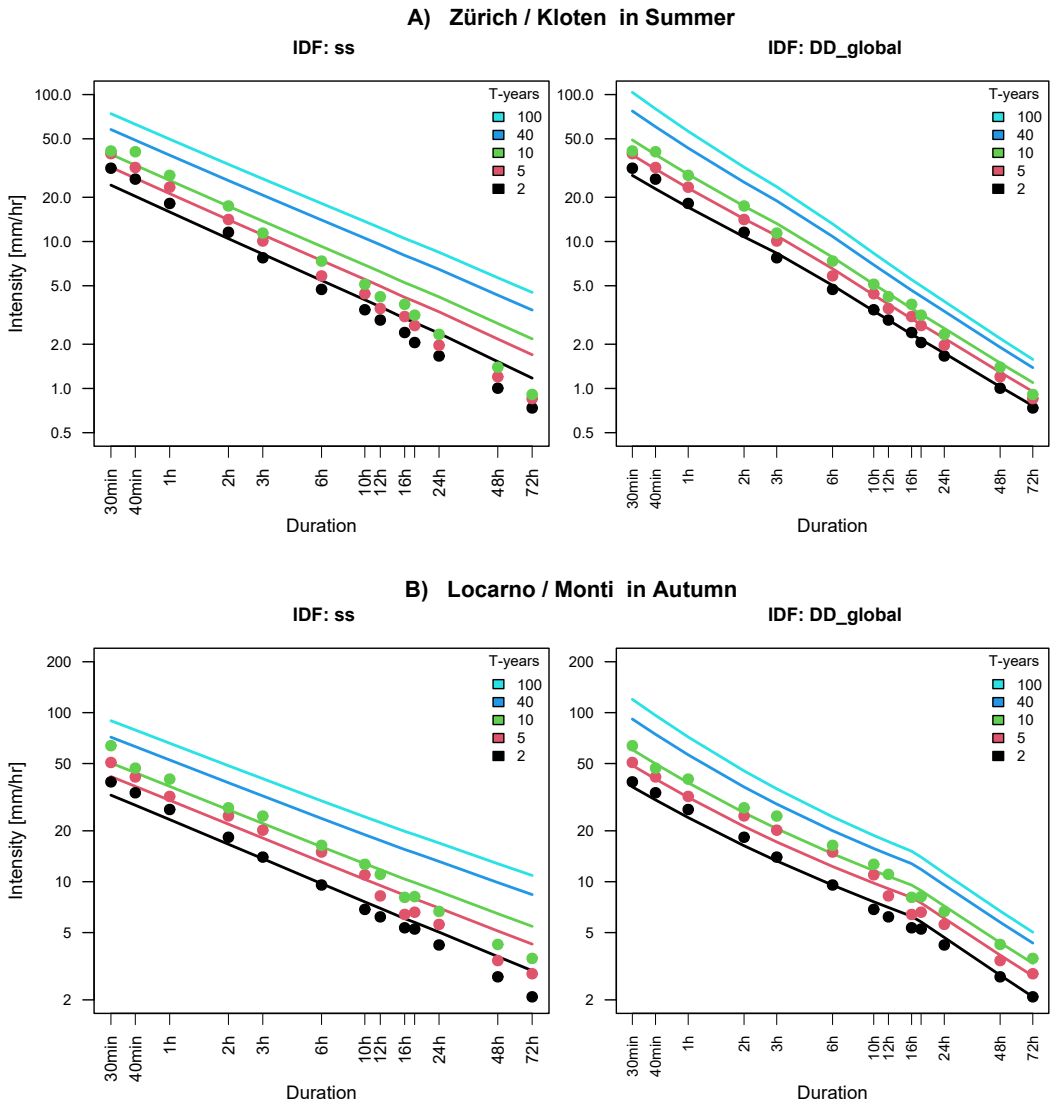

Figure 7. Simple-scaling (IDF_{ss}) and data-driven ($\text{IDF}_{DD_{global}}$) curves **a)** in summer at a station in Zurich (North-east). **b)** in autumn at a station in Locarno (Ticino area in the south). The curves are for the return periods $T = 2$, 5, 10, 40 and 100 years. The points are the empirical quantiles corresponding to T = 2, 5 and 10 years

rameters for a particular duration. For instance, for 13 durations, this means $3 \times 13 = 39$ maps. For our choice of linear functions, 10 parameters are able to describe the IDF curves at each station, and hence 10 maps for the whole area under study.

Regarding the variation of the shape parameter ξ with respect to duration d, some earlier studies did observed or discussed it *e.g.* Veneziano et al. (2007) and Fauer et al. (2021). They however did not model it, either due to the weak form of the relationship, or because the IDF model did not allow for it. Here, especially, in summer, we found very strong dependence, and the results have shown that taking it into account is invaluable.

Finally it should be mentioned that throughout this work, we estimated the IDF models through the independence likelihood, thus omitting the correlation between dif-

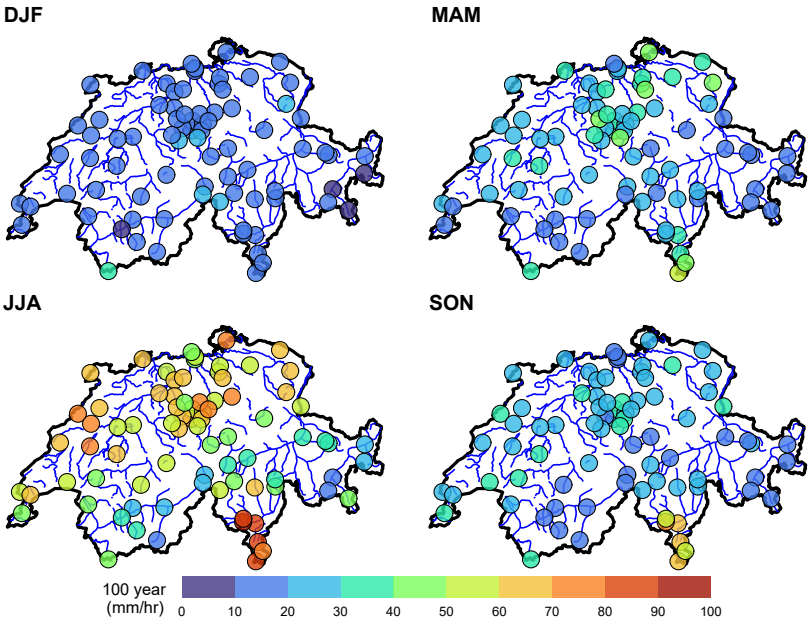

Figure 8. Map of Switzerland showing the seasonal 100-year return level in mm/hr for $d = 1$ hr. Levels predicted with $\text{IDF}_{DD_{global}}$.

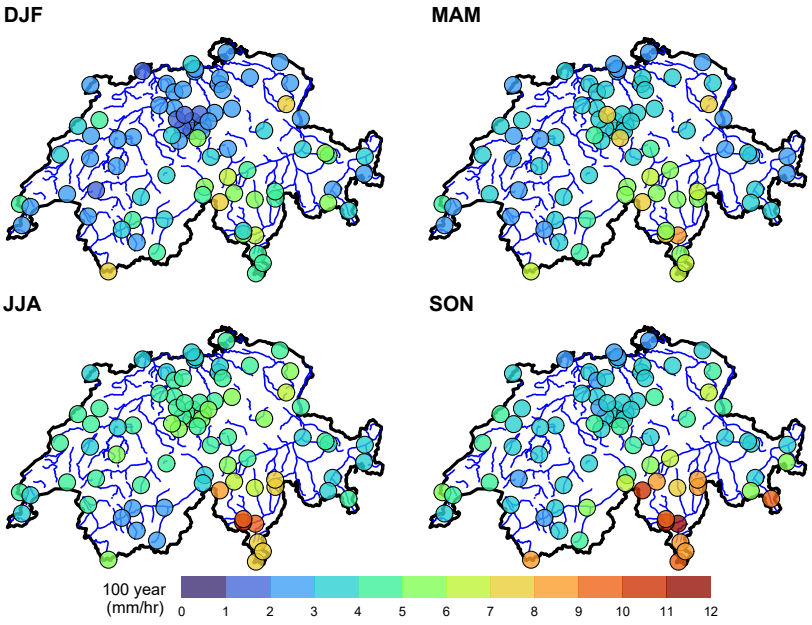

Figure 9. Map of Switzerland showing the seasonal 100-year return level in mm/hr for $d = 24$ hr. Levels predicted with $\text{IDF}_{DD_{global}}$.

ferent times and durations. Nadarajah et al. (1998) has modeled this using multivariate extreme value distributions (MEVD), and Tyralis and Langousis (2019) followed suit by using max-stable processes. Later, Jurado et al. (2020), investigated the impact of accounting for the dependence in extremes and showed that there is little gain in performance, in addition to the added complexity of using max-stable processes.

5 Conclusions

Our aim in this paper was to build IDF curves using all the non-zero precipitation data in Switzerland. To achieve this, we used the EGPD model as the parametric model for the precipitation intensities. The literature presents various approaches to link the different durations together in IDF curves. We considered three of these approaches to build the IDF curves while using the EGPD as the parametric model. The first is the data-driven approach, where each parameter can vary with duration. The form of the relationship is fully determined by the data at hand. The second approach is based on the scale invariance theory, here IDF curves are built based on the scaling behavior of precipitation. The last approach is based on the general IDF formulation of Koutsoyiannis et al. (1998), which generalizes the various traditional IDF formulations.

We started from these three approaches and added some extensions to account for scaling break and varying shape parameter. We ended up with a total of ten IDF models. We then compared them, first in calibration, and then in a split-sample cross-validation approach.

The results showed that, given the EGPD as the parametric model, the data-driven IDF-EGPD, particularly the $\text{IDF}_{DD_{global}}$, is the best model for the data at hand. This is despite being less parsimonious in terms of its number of free parameters. The IDF curves based on simple-scaling and the general formulation of Koutsoyiannis et al. (1998), did not perform as efficiently even with the added extensions in terms of scaling break and in the way the shape parameter varies with duration. The fact that the simple-scaling IDF models performed poorly in summer confirms the previous findings of Molnar and Burlando (2008) and (Paschalis, 2013) that in Switzerland, precipitation in summer shows multiscaling behavior.

In terms of perspectives, it would be interesting to produce maps of the parameters to allow for predictions at ungauged sites. This could be achieved by simple interpolation of the local IDF parameters as done by Blanchet et al. (2016), or through quantile regression methods (Ouali & Cannon, 2018), or by global estimation using spatial covariates (e.g. Ulrich et al., 2020). Another possibility is to use a regionalization technique, such as the method of Hosking and Wallis (2005) and then interpolate the index flood to allow predictions at the ungauged sites (e.g. Mascaro, 2020).

Lastly, consideration of the effect of climate change in building IDF curves is invaluable. For instance, Cheng and AghaKouchak (2014) showed that by neglecting nonstationarity in modeling IDF curves, there could be up to 60% underestimation of extreme precipitation, especially for short durations. It would therefore be interesting to model the curves while accounting for a warmer climate (e.g. Mirhosseini et al., 2013; Cheng & AghaKouchak, 2014; Ragno et al., 2018; Ouarda et al., 2019; Kristvik et al., 2019).

6 Open Research

The data used in this study are maintained by Swiss Federal Office of Meteorology and Climatology, MeteoSwiss (MeteoSwiss, 2021). It is available upon request at https://gate.meteoswiss.ch/idaweb/ (last accessed 10 December 2021).

Acknowledgments

This research has been supported by the Bundesamt für Energie (grant no. SI/502150-01) and the Bundesamt für Umwelt (grant no. SI/502150-01), through the *Extreme floods in Switzerland* project.

References

Agbazo, M. N., Koto N'Gobi, G., Kounouhewa, B., Alamou, E., Afouda, A., & Akpo, A. (2016, April). Estimation of IDF Curves of Extreme Rainfall by Simple Scaling in Northern Oueme Valley, Benin Republic (West Africa). *Earth Sciences Research Journal*, *20*(1), 1–7. Retrieved 2021-09-09, from http://revistas.unal.edu.co/index.php/esrj/article/view/49405 doi: 10.15446/esrj.v20n1.49405

Ben-Zvi, A. (2009). Rainfall intensity–duration–frequency relationships derived from large partial duration series. *Journal of Hydrology*, *367*(1-2), 104–114. doi: https://doi.org/10.1016/j.jhydrol.2009.01.007

Bernard, M. M. (1932). Formulas for rainfall intensities of long duration. *Transactions of the American Society of Civil Engineers*, *96*(1), 592–606. (Publisher: American Society of Civil Engineers) doi: https://doi.org/10.1061/taceat.0004323

Blanchet, J., Ceresetti, D., Molinié, G., & Creutin, J.-D. (2016, September). A regional GEV scale-invariant framework for Intensity–Duration–Frequency analysis. *Journal of Hydrology*, *540*, 82–95. doi: 10.1016/j.jhydrol.2016.06.007

Blanchet, J., Paquet, E., Vaittinada Ayar, P., & Penot, D. (2019, February). Mapping rainfall hazard based on rain gauge data: an objective cross-validation framework for model selection. *Hydrology and Earth System Sciences*, *23*(2), 829–849. doi: 10.5194/hess-23-829-2019

Blanchet, J., Touati, J., Lawrence, D., Garavaglia, F., & Paquet, E. (2015). Evaluation of a compound distribution based on weather pattern subsampling for extreme rainfall in Norway. *Nat. Hazards Earth Syst. Sci.*, 15. doi: https://doi.org/10.5194/nhess-15-2653-2015

Bougadis, J., & Adamowski, K. (2006, November). Scaling model of a rainfall intensity-duration-frequency relationship. *Hydrological Processes*, *20*(17), 3747–3757. Retrieved 2021-09-06, from https://onlinelibrary.wiley.com/doi/10.1002/hyp.6386 doi: 10.1002/hyp.6386

Burlando, P., & Rosso, R. (1996, December). Scaling and muitiscaling models of depth-duration-frequency curves for storm precipitation. *Journal of Hydrology*, *187*(1), 45–64. Retrieved 2021-09-08, from https://www.sciencedirect.com/science/article/pii/S0022169496030867 doi: 10.1016/S0022-1694(96)03086-7

Carreteras, M. T. (1987). Cálculo hidrometeorológico de caudales máximos en pequeñas cuencas naturales. *Textos de la Direccion General de Carreteras*(12), 1987.

Chang, K. B., Lai, S. H., & Othman, F. (2016). Comparison of annual maximum and partial duration series for derivation of rainfall intensity-duration-frequency relationships in peninsular malaysia. *Journal of Hydrologic Engineering*, *21*(1), 05015013. doi: 10.1061/(ASCE)HE.1943-5584.0001262

Cheng, L., & AghaKouchak, A. (2014). Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. *Scientific reports*, *4*(1), 1–6. doi: https://doi.org/10.1038/srep07093

Chow, V. (1962). Hydrologic determination of waterway areas for drainage structures in small drainage basins, Engrg. *Experimental Station, Univ. of Illinois, Urbana I*, *11*.

Courty, L. G., Wilby, R. L., Hillier, J. K., & Slater, L. J. (2019, August). Intensity-duration-frequency curves at the global scale. *Environmental Research Letters*,

14(8), 084045. Retrieved 2022-01-18, from https://doi.org/10.1088/1748-9326/ab370a (Publisher: IOP Publishing) doi: 10.1088/1748-9326/ab370a

Evin, G., Blanchet, J., Paquet, E., Garavaglia, F., & Penot, D. (2016, October). A regional model for extreme rainfall based on weather patterns subsampling. *Journal of Hydrology*, *541*, 1185–1198. Retrieved 2021-08-25, from https://linkinghub.elsevier.com/retrieve/pii/S0022169416305145 doi: 10.1016/j.jhydrol.2016.08.024

Evin, G., Favre, A.-C., & Hingray, B. (2018). Stochastic generation of multi-site daily precipitation focusing on extreme events. *Hydrol. Earth Syst. Sci.*, 18. doi: https://doi.org/10.5194/hess-22-655-2018

Fauer, F. S., Ulrich, J., Jurado, O. E., & Rust, H. W. (2021, July). *Flexible and Consistent Quantile Estimation for Intensity-Duration-Frequency Curves* (preprint). Hydrometeorology/Modelling approaches. Retrieved 2021-11-26, from https://hess.copernicus.org/preprints/hess-2021-334/ doi: 10.5194/hess-2021-334

Fukutome, S., Liniger, M. A., & Süveges, M. (2015, May). Automatic threshold and run parameter selection: a climatology for extreme hourly precipitation in Switzerland. *Theoretical and Applied Climatology*, *120*(3-4), 403–416. Retrieved 2021-04-19, from http://link.springer.com/10.1007/s00704-014-1180-5 doi: 10.1007/s00704-014-1180-5

Garavaglia, F., Lang, M., Paquet, E., Gailhard, J., Garçon, R., & Renard, B. (2011). Reliability and robustness of rainfall compound distribution model based on weather pattern sub-sampling. *Hydrol. Earth Syst. Sci.*, 15. doi: https://doi.org/10.5194/hess-15-519-2011

Ghanmi, H., Bargaoui, Z., & Mallet, C. (2016). Estimation of intensity-duration-frequency relationships according to the property of scale invariance and regionalization analysis in a mediterranean coastal area. *Journal of Hydrology*, *541*, 38–49. doi: 10.1016/j.jhydrol.2016.07.002

Gupta, V. K., & Waymire, E. (1990). Multiscaling properties of spatial rainfall and river flow distributions. *Journal of Geophysical Research: Atmospheres*, *95*(D3), 1999–2009. doi: 10.1029/JD095iD03p01999

Gupta, V. K., & Waymire, E. C. (1993, February). A Statistical Analysis of Mesoscale Rainfall as a Random Cascade. *Journal of Applied Meteorology and Climatology*, *32*(2), 251–267. doi: 10.1175/1520-0450(1993)032⟨0251:ASAOMR⟩2.0.CO;2

Harris, D., Seed, A., Menabde, M., & Austin, G. (1997, September). Factors affecting multiscaling analysis of rainfall time series. *Nonlinear Processes in Geophysics*, *4*(3), 137–156. Retrieved 2021-11-15, from https://npg.copernicus.org/articles/4/137/1997/ doi: 10.5194/npg-4-137-1997

Haruna, A., Blanchet, J., & Favre, A.-C. (2022). Performance-based comparison of regionalization methods to improve the at-site estimates of daily precipitation. *Hydrology and Earth System Sciences*, *26*(10), 2797–2811. doi: https://doi.org/10.5194/hess-26-2797-2022

Hosking, J. R. M., & Wallis, J. R. (2005). *Regional frequency analysis: an approach based on L-moments*. Cambridge ; New York: Cambridge University Press.

Innocenti, S., Mailhot, A., & Frigon, A. (2017, November). Simple scaling of extreme precipitation in North America. *Hydrology and Earth System Sciences*, *21*(11), 5823–5846. Retrieved 2021-09-08, from https://hess.copernicus.org/articles/21/5823/2017/ (Publisher: Copernicus GmbH) doi: 10.5194/hess-21-5823-2017

Jurado, O. E., Ulrich, J., Scheibel, M., & Rust, H. W. (2020, December). Evaluating the Performance of a Max-Stable Process for Estimating Intensity-Duration-Frequency Curves. *Water*, *12*(12), 3314. Retrieved 2022-01-27, from https://www.mdpi.com/2073-4441/12/12/3314 (Number: 12 Publisher: Multidisciplinary Digital Publishing Institute) doi: 10.3390/w12123314

Katz, R. W., Parlange, M. B., & Naveau, P. (2002, August). Statistics of extremes in hydrology. *Advances in Water Resources*, *25*(8-12), 1287–1304. Retrieved 2021-07-26, from https://linkinghub.elsevier.com/retrieve/pii/S0309170802000568 doi: 10.1016/S0309-1708(02)00056-8

Koutsoyiannis, D., Kozonis, D., & Manetas, A. (1998, April). A mathematical framework for studying rainfall intensity-duration-frequency relationships. *Journal of Hydrology*, *206*(1-2), 118–135. Retrieved 2021-09-06, from https://linkinghub.elsevier.com/retrieve/pii/S0022169498000973 doi: 10.1016/S0022-1694(98)00097-3

Kristvik, E., Johannessen, B. G., & Muthanna, T. M. (2019). Temporal downscaling of idf curves applied to future performance of local stormwater measures. *Sustainability*, *11*(5), 1231. doi: https://doi.org/10.3390/su11051231

Langousis, A., & Veneziano, D. (2007, February). Intensity-duration-frequency curves from scaling representations of rainfall. *Water Resources Research*, *43*(2). Retrieved 2021-09-17, from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2006WR005245 (Publisher: John Wiley & Sons, Ltd) doi: 10.1029/2006WR005245

Le Gall, P., Favre, A.-C., Naveau, P., & Prieur, C. (2022). Improved regional frequency analysis of rainfall data. *Weather and Climate Extremes*, *36*, 100456. doi: https://doi.org/10.1016/j.wace.2022.100456

Lima, M. I. P. d. (1998). *Multifractals and the temporal structure of rainfall* (PhD Thesis). Wageningen Agricultural University, Wageningen, Netherlands.

Madsen, H., Rosbjerg, D., & Harremoöes, P. (1995, March). Application of the Bayesian approach in regional analysis of extreme rainfalls. *Stochastic Hydrology and Hydraulics*, *9*(1), 77–88. Retrieved 2021-07-09, from http://link.springer.com/10.1007/BF01581759 doi: 10.1007/BF01581759

Mascaro, G. (2020). Comparison of local, regional, and scaling models for rainfall intensity–duration–frequency analysis. *Journal of applied meteorology and climatology*, *59*(9), 1519–1536. doi: https://doi.org/10.1175/jamc-d-20-0094.1

Menabde, M., Seed, A., & Pegram, G. (1999, January). A simple scaling model for extreme rainfall. *Water Resources Research*, *35*(1), 335–339. Retrieved 2021-09-17, from http://doi.wiley.com/10.1029/1998WR900012 doi: 10.1029/1998WR900012

MeteoSwiss. (2021). *Meteoswiss, federal office of meteorology and climatology.* Retrieved 2021-12-10, from https://gate.meteoswiss.ch/idaweb/login.do

Meylan, P., Favre, A.-C., & Musy, A. (2012). *Predictive hydrology: a frequency analysis approach.* CRC Press.

Mirhosseini, G., Srivastava, P., & Stefanova, L. (2013). The impact of climate change on rainfall intensity–duration–frequency (idf) curves in alabama. *Regional Environmental Change*, *13*(1), 25–33. doi: https://doi.org/10.1007/s10113-012-0375-5

Molnar, P., & Burlando, P. (2008). Variability in the scale properties of high-resolution precipitation data in the Alpine climate of Switzerland. *Water Resources Research*, *44*(10). doi: http://doi.wiley.com/10.1029/2007WR006142

Muggeo, V. M. (2003). Estimating regression models with unknown break-points. *Statistics in medicine*, *22*(19), 3055–3071. doi: https://doi.org/10.1002/sim.1545

Muggeo, V. M., et al. (2008). Segmented: an r package to fit regression models with broken-line relationships. *R news*, *8*(1), 20–25.

Mélèse, V., Blanchet, J., & Molinié, G. (2018, March). Uncertainty estimation of Intensity–Duration–Frequency relationships: A regional analysis. *Journal of Hydrology*, *558*, 579–591. Retrieved 2022-02-15, from https://linkinghub.elsevier.com/retrieve/pii/S002216941730519X doi: 10.1016/j.jhydrol.2017.07.054

Nadarajah, S., Anderson, C., & Tawn, J. (1998). Ordered multivariate extremes.

Journal of the Royal Statistical Society: Series B (Statistical Methodology), *60*(2), 473–496. doi: https://doi.org/10.1111/1467-9868.00136

Naveau, P., Huser, R., Ribereau, P., & Hannart, A. (2016, April). Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection. *Water Resources Research*, *52*(4), 2753–2769. Retrieved 2021-08-25, from http://doi.wiley.com/10.1002/2015WR018552 doi: 10.1002/2015WR018552

Nhat, L. M., Tachikawa, Y., Sayama, T., & Takara, K. (2008). Estimation of sub-hourly and hourly idf curves using scaling properties of rainfall at gauged site in asian pacific region. *Annual of Disaster Prevention Research Institute. B*, *51*(B), 63–73. doi: https://doi.org/10.1016/j.atmosres.2004.10.025

Ouali, D., & Cannon, A. (2018). Estimation of rainfall intensity–duration–frequency curves at ungauged locations using quantile regression methods. *Stochastic Environmental Research and Risk Assessment*, *32*(10), 2821–2836. doi: https://doi.org/10.1007/s00477-018-1564-7

Ouarda, T. B., Yousef, L. A., & Charron, C. (2019). Non-stationary intensity-duration-frequency curves integrating information concerning teleconnections and climate change. *International Journal of Climatology*, *39*(4), 2306–2323. doi: https://doi.org/10.1002/joc.5953

Over, T. M. (1995). *Modeling space-time rainfall at the mesoscale using random cascades* (PhD Thesis). University of Colorado, Boulder.

Overeem, A., Buishand, A., & Holleman, I. (2008, January). Rainfall depth-duration-frequency curves and their uncertainties. *Journal of Hydrology*, *348*(1-2), 124–134. Retrieved 2021-11-30, from https://linkinghub.elsevier.com/retrieve/pii/S0022169407005513 doi: 10.1016/j.jhydrol.2007.09.044

Panthou, G., Vischel, T., Lebel, T., Quantin, G., & Molinié, G. (2014, December). Characterising the space–time structure of rainfall in the Sahel with a view to estimating IDAF curves. *Hydrology and Earth System Sciences*, *18*(12), 5093–5107. Retrieved 2021-11-04, from https://hess.copernicus.org/articles/18/5093/2014/ doi: 10.5194/hess-18-5093-2014

Paschalis, A. (2013). *Modelling the Space-Time Structure of Precipitation and its Impact on Basin Response* (Doctoral dissertation, ETH Zurich). doi: 10.3929/ETHZ-A-009917135

Ragno, E., AghaKouchak, A., Love, C. A., Cheng, L., Vahedifard, F., & Lima, C. H. (2018). Quantifying changes in future intensity-duration-frequency curves using multimodel ensemble simulations. *Water Resources Research*, *54*(3), 1751–1764. doi: https://doi.org/10.1002/2017wr021975

Renard, B., Kochanek, K., Lang, M., Garavaglia, F., Paquet, E., Neppel, L., ... Auffray, A. (2013, February). Data-based comparison of frequency analysis methods: A general framework. *Water Resources Research*, *49*(2), 825–843. Retrieved 2021-06-29, from https://agupubs.onlinelibrary.wiley.com/doi/10.1002/wrcr.20087 doi: 10.1002/wrcr.20087

Ritschel, C., Ulbrich, U., Névir, P., & Rust, H. W. (2017). Precipitation extremes on multiple timescales–bartlett–lewis rectangular pulse model and intensity–duration–frequency curves. *Hydrology and Earth System Sciences*, *21*(12), 6501–6517. doi: https://doi.org/10.5194/hess-21-6501-2017

Rivoire, P., Martius, O., & Naveau, P. (2021). A Comparison of Moderate and Extreme ERA-5 Daily Precipitation With Two Observational Data Sets. *Earth and Space Science*, *8*(4), e2020EA001633. Retrieved 2022-03-22, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2020EA001633 doi: 10.1029/2020EA001633

Roksvåg, T., Lutz, J., Grinde, L., Dyrrdal, A. V., & Thorarinsdottir, T. L. (2021). Consistent intensity-duration-frequency curves by post-processing of estimated bayesian posterior quantiles. *Journal of Hydrology*, *603*, 127000. doi:

https://doi.org/10.1016/j.jhydrol.2021.127000

Sane, Y., Panthou, G., Bodian, A., Vischel, T., Lebel, T., Dacosta, H., ... Diop Kane, M. (2018, July). Intensity–duration–frequency (IDF) rainfall curves in Senegal. *Natural Hazards and Earth System Sciences*, *18*(7), 1849–1866. (Publisher: Copernicus GmbH) doi: 10.5194/nhess-18-1849-2018

Schertzer, D., & Lovejoy, S. (1987). Physical Modeling and Analysis of Rain and Clouds by Anisotropic Scaling Multiplicative Processes. *Journal of Geophysical Research*, *92*, 22. doi: https://doi.org/10.1029/jd092id08p09693

Sherman, C. W. (1931). Frequency and intensity of excessive rainfalls at Boston, Massachusetts. *Transactions of the American Society of Civil Engineers*, *95*(1), 951–960. (Publisher: American Society of Civil Engineers) doi: https://doi.org/10.1061/taceat.0004286

Tencaliec, P., Favre, A., Naveau, P., Prieur, C., & Nicolet, G. (2020). Flexible semi-parametric generalized Pareto modeling of the entire range of rainfall amount. *Environmetrics*, *31*(2). doi: 10.1002/env.2582

Tyralis, H., & Langousis, A. (2019, January). Estimation of intensity–duration–frequency curves using max-stable processes. *Stochastic Environmental Research and Risk Assessment*, *33*(1), 239–252. Retrieved 2022-01-27, from `https://doi.org/10.1007/s00477-018-1577-2` doi: 10.1007/s00477-018-1577-2

Ulrich, J., Jurado, O. E., Peter, M., Scheibel, M., & Rust, H. W. (2020, November). Estimating IDF Curves Consistently over Durations with Spatial Covariates. *Water*, *12*(11), 3119. Retrieved 2021-12-10, from `https://www.mdpi.com/2073-4441/12/11/3119` doi: 10.3390/w12113119

Van de Vyver, H. (2018). A multiscaling-based intensity–duration–frequency model for extreme precipitation. *Hydrological Processes*, *32*(11), 1635–1647. (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hyp.11516) doi: 10.1002/hyp.11516

Van de Vyver, H., & Demarée, G. R. (2010). Construction of Intensity–Duration–Frequency (IDF) curves for precipitation at Lubumbashi, Congo, under the hypothesis of inadequate data. *Hydrological Sciences Journal*, *55*(4), 555–564. doi: 10.1080/02626661003747390

Veneziano, D., & Lepore, C. (2012). The scaling of temporal rainfall. *Water Resources Research*, *48*(8). doi: 10.1029/2012WR012105

Veneziano, D., Lepore, C., Langousis, A., & Furcolo, P. (2007, October). Marginal methods of intensity-duration-frequency estimation in scaling and nonscaling rainfall. *Water Resources Research*, *43*(10). doi: 10.1029/2007WR006040

Willems, P. (2000). Compound intensity/duration/frequency-relationships of extreme precipitation for two seasons and two storm types. *Journal of Hydrology*, *233*(1), 189–205. doi: 10.1016/S0022-1694(00)00233-X

Youngman, B. D. (2019, October). Generalized Additive Models for Exceedances of High Thresholds With an Application to Return Level Estimation for U.S. Wind Gusts. *Journal of the American Statistical Association*, *114*(528), 1865–1879. Retrieved 2021-07-26, from `https://www.tandfonline.com/doi/full/10.1080/01621459.2018.1529596` doi: 10.1080/01621459.2018.1529596

Youngman, B. D. (2020, November). evgam: An R package for Generalized Additive Extreme Value Models. *arXiv:2003.04067 [stat]*. Retrieved 2021-04-19, from `http://arxiv.org/abs/2003.04067` (arXiv: 2003.04067)

Yu, P.-S., Yang, T.-C., & Lin, C.-S. (2004, August). Regional rainfall intensity formulas based on scaling property of rainfall. *Journal of Hydrology*, *295*(1), 108–123. Retrieved 2021-11-17, from `https://www.sciencedirect.com/science/article/pii/S0022169404001362` doi: 10.1016/j.jhydrol.2004.03.003