Modeling Areal Precipitation Hazard: A Data-driven Approach to Model Intensity-Duration-Area-Frequency Relationships using the Full Range of Non-Zero Precipitation in Switzerland

Abubakar Haruna[1], Juliette Blanchet[2], and Anne-Catherine Favre[3]

[1]Université Grenoble Alpes
[2]Universite Grenoble Alpes
[3]CNRS, IRD, IGE, Grenoble INP

January 24, 2024

Estimation of Intensity-Duration-Area-Frequency Relationships based on the Full Range of Non-Zero Precipitation from Radar-Reanalysis Data

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 develop seasonal IDAF models at every pixel location in Switzerland.
- We use all the non-zero precipitation data and model the intensities using the extended generalized Pareto distribution.
- We highlight the complexity of modeling areal precipitation in mountainous regions

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

Abstract

Intensity-Duration-Area-Frequency (IDAF) models provide the mathematical link between precipitation intensities (I), durations (D), areas (A), and frequency of occurrence (F). They play a critical role in hydrological design, areal rainfall hazard quantification, storm characterization, and early warning system development. IDAF models extend the conventional Intensity-Duration-Frequency (IDF) models by accounting for the spatial extent of precipitation(*i.e.*, the area). In this study, we develop IDAF models using the entire non-zero precipitation intensities, not only the extremes. We use the extended generalized Pareto distribution (EGPD) to model the precipitation intensities. To build the IDAF models, we adopt a data-driven approach that allows the linkage of EGPD parameters with duration and area, based on empirically determined parametric relationships. The inference of model parameters is done using a global maximum likelihood estimation, and uncertainties are assessed by the bootstrap method. The study area is Switzerland, a topographically complex region of 42,000 km^2 with regional precipitation variability and clear seasonality. The study utilizes 17 years of data from `CombiPrecip`, a radar-reanalysis product developed by geostatistically merging radar and rain gauge data in an operational setting. We build the IDAF models for the spatiotemporal range of 1 to 72 hours and 1 to 1,089 km^2 at each pixel in the study area. To the best of our knowledge, our study is the first attempt to use the EGPD in IDAF curve modeling. It discusses the use and limitations of `CombiPrecip` in extreme value analysis and highlights the challenges of modeling areal precipitation in a complex topographical environment.

1 Introduction

In the face of escalating threats posed by climate change and increasingly volatile weather patterns, understanding and predicting extreme precipitation is necessary, now more than ever, in safeguarding communities and infrastructure. One of the key factors driving flood generation is the spatial aggregate of precipitation over a given area, rather than just the precipitation at a specific point location. This is because watersheds and river basins integrate the precipitation falling over their respective areas, leading to the accumulation of runoff and subsequent flood generation. Additionally, extreme precipitation events, manifesting at varying scales, contribute differently to flood dynamics. Short and small-scale intense precipitation may induce rapid, localized flash flooding, whereas longer and larger-scale precipitation events can lead to sustained fluvial flooding (Sikorska et al., 2015). However, the interactions and synergies between these scales are crucial in shaping the overall flood risk landscape. As a consequence, it is vital to consider multiple spatiotemporal scales in the modeling of extreme precipitation. This will enhance our ability to better predict and manage the impacts on communities and infrastructure, ensuring their resilience in an ever-changing climate.

Intensity Duration Area Frequency (IDAF) curves summarize the main statistical characteristics of extreme precipitation (return level, return period, duration, and area.) They provide the mathematical link between precipitation intensities (I), durations (D), areas (A), and frequency of occurrence (F). They are useful tools for engineers and hydrologists in hydrological design (see Bertini et al., 2020, for example), quantification of areal rainfall hazard (Overeem et al., 2010; Panthou et al., 2014; Mélèse et al., 2019; Zhao et al., 2023), storm characterization (Ramos et al., 2005; Ceresetti et al., 2012; Blanchet & Mélèse, 2020), and development of early warning systems (Panziera et al., 2016). IDAF models extend the well-known Intensity Duration Frequency curves (IDF) by incorporating the spatial extent of precipitation (*i.e.*, the area).

IDAF curves are commonly built by coupling IDF models and a coefficient, the areal reduction factor (ARF) that transforms point rainfall of a given duration and return period to areal return levels of the corresponding duration and return period. Applications of the ARF-based IDAF models can be found in the literature, for example, De Michele et al. (2001) derived an ARF formulation based on the concept of dynamic scaling of rainfall and used it to model IDAF curves in Milan. Later, Ceresetti et al. (2012) used the ARF of De Michele et al. (2001) to model IDAF curves for storm severity assessment in southern France. Panthou et al. (2014) also used the same ARF formulation to characterize areal rainfall in West Africa. Ramos et al. (2005) used an empirical ARF formulation to model IDAF curves for storm severity assessment in Marseille. Bertini et al. (2020) used another empirical ARF formulation to build IDAF curves and used it to design a dam in Italy. Mélèse et al. (2019) and Blanchet and Mélèse (2020) used an extension of the ARF formulation of De Michele et al. (2001) to build IDAF curves respectively for areal hazards and storm severity assessment in southern France. The extension was to cope with the significant spatiotemporal variability in the mountainous area

Beyond the ARF-based IDAF curves modeling approach, Overeem et al. (2010) proposed a purely data-driven approach to model IDAF curves. This involves modeling the parameters of the statistical distribution of the precipitation intensities as a function of duration and area. The type of relationship is empirically determined from the data, with no underlying physical hypothesis such as spatial correlation as done in Rodriguez-Iturbe and Mejía (1974) or scaling (as done in De Michele et al., 2001). As highlighted by Mélèse et al. (2019), this approach has the advantage of being flexible and applicable in cases where the assumptions of the analytical ARF formulations cannot be verified.

In spite of the chosen method of building the IDAF curves, whether ARF-based or purely data-driven, the previously cited works have one thing in common; the precipitation data they used and by extension, the underlying parametric distribution. To elaborate more, all the authors used only extreme data in the form of block maxima and, as the distribution, the generalized extreme value (GEV) distribution (Overeem et al., 2010; Ceresetti et al., 2012; Panthou et al., 2014) or its special case, the Gumbel distribution (Nhat et al., 2007; Mélèse et al., 2019; Blanchet & Mélèse, 2020; Bertini et al., 2020), or log-normal distribution (De Michele et al., 2011). A rare application of generalized Pareto distribution (GPD) for threshold excesses is found in Zhao et al. (2023) for IDAF curve modeling. A major drawback of such approaches is the inefficient use of the data since only one value is retained in a block (usually the maximum in a year) or the excesses of a threshold (a tiny fraction of the data), and all the other data in the block is discarded. This can result in significant uncertainty in estimation, especially in cases where the length of the data series is not sufficiently long. The problem of short record length is more apparent with radar and radar reanalysis products, which are usually used in IDAF curve modeling (Overeem et al., 2010; Mélèse et al., 2019; Blanchet & Mélèse, 2020; Zhao et al., 2023), due to the required spatial information they provide.

To address this issue, our approach here is to make efficient use of information by including all the non-zero precipitation intensities, instead of only the block maxima, in modeling the IDAF curves. We then use the extended generalized Pareto distribution (EGPD) of Naveau et al. (2016) as the parametric model for the intensities. This distribution is compliant with extreme value theory in both tails (an advantage over the gamma distribution), it models the entire distribution of non-zero precipitation and does not require the choice of the threshold as in the generalized Pareto distribution (GPD). It has been shown in many applications to be able to adequately model precipitation (Naveau et al., 2016; Evin et al., 2018; Le Gall et al., 2022; Haruna et al., 2022, 2023). In particular, Haruna et al. (2023) showed that it is possible to model IDF curves (without Area) with the EGPD, and we intend to extend their work to model IDAF curves with the EGPD. To our knowledge, this is the first time the EGPD has been used in modeling IDAF curves.

Modeling IDAF curves using all the non-zero data has two potential advantages. First, by using all the non-zero data, estimation uncertainty is expected to reduce, resulting in more accurate predictions. Secondly, in addition to having IDAF curves, we will have robust marginal distributions for the entire non-zero precipitation that can be used in stochastic weather generators for simulations, or verification of weather and climate models.

We apply the model in Switzerland, a topographically complex location with seasonality, regional variability, and multiple precipitation regimes. Following the work of Mélèse et al. (2019) which underscores the complex spatiotemporal variability of precipitation in mountainous areas, we use the more flexible data-driven method of Overeem et al. (2010) to model the IDAF curves.

The data and study area are presented in Section 2. The EGPD, the methodology for building the IDAF curves, and the method for uncertainty assessment are explained in Section 3. Results on the goodness of fit of the model and areal rainfall hazard assessment in the study area are presented and discussed in Section 4. Finally, conclusions and perspectives are given in Section 5

2 Study area and Data

2.1 Study Area

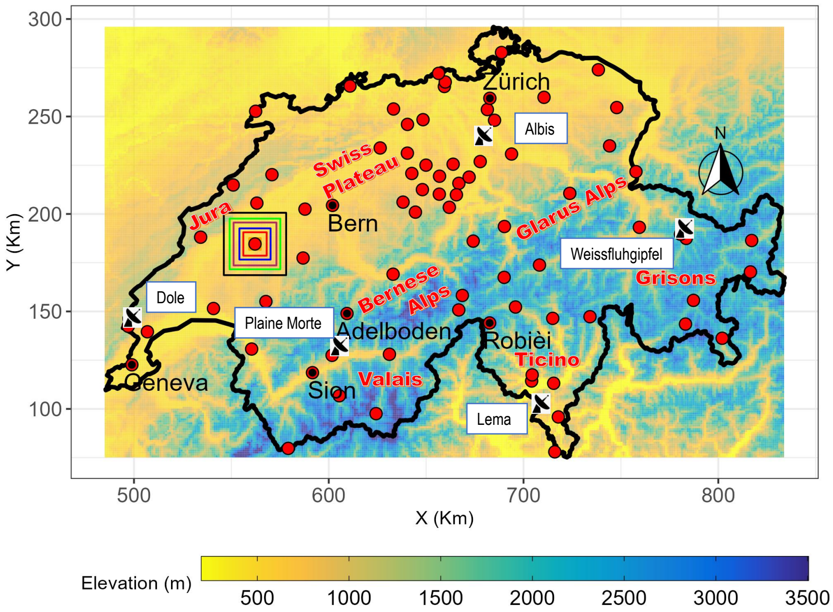

Figure 1. Map of Switzerland, the study area. The background color denotes the elevation (meters) above mean sea level. The Radar symbols show the location of the five radars in Switzerland, with their names in the white boxes. The name of some cities is shown in black and the name of some mountains and regions are shown in red. The colored embedded squares show exemplarily the extent of 7 out of the 10 square windows used for data aggregation. They cover areas of 25, 49, 169, 289, 529, 729, and 1089 km^2, from the innermost to the outermost. The red colored circles show the location of 79 rain gauges used for evaluation of the gridded product (CPC).

Our study focuses on Switzerland, a country covering 41,285 km^2. Despite its small size, Switzerland exhibits a complex topography, ranging in elevation from 191 to 4,127 m above mean sea level. Figure 1 shows the map of the study area. Approximately 30% of the land is situated above 1,500 m elevation, resulting in pronounced spatial variability in both the intensity and occurrence of precipitation. The climate of Switzerland is influenced by multiple factors, such as the Alps, the Atlantic Ocean, and the Mediterranean Sea, and these contribute to the seasonal and spatial variability of precipitation, as documented in previous studies (Sodemann & Zubler, 2009; Giannakaki & Martius, 2015; Scherrer et al., 2016). Precipitation patterns show distinct regional differences, with the highest annual sums exceeding 2,000 mm in the Alps, the Jura region (northwest), and the Ticino region (south of the Alps). Conversely, the inner valleys such as the Rhône and Inn receive the lowest annual precipitation, less than 700 mm. Summer is the primary season for precipitation throughout Switzerland, except in Ticino, where autumn dominates. Conversely, winter experiences the least amount of precipitation across all regions. In terms of heavy precipitation, defined as the average seasonal maxima, the spatial distribution varies according to accumulation duration (Panziera et al., 2018). For short-duration accumulations (*e.g.*, 1 hr), the heaviest precipitation occurs in summer across the entire country, with maximum intensities reaching up to 30 mm/hr in Ticino, Jura, and the northern rim. For longer accumulations (1 day and more), Ticino receives the most intense precipitation, with autumn experiencing a maximum 24 hr total exceeding 130 mm. In other regions, heavy precipitation predominantly occurs during summer.

2.2 `CombiPrecip`

`CombiPrecip` (CPC) is a radar-reanalysis product resulting from the geostatistical merging of radar and rain gauge in an operational setting (Sideris et al., 2014a). It combines the high accuracy of rain gauge with the high spatial coverage of radar. The geostatistical merging is through co-kriging with external drift, where the rain gauge data is treated as the primary source, and the radar data as the external drift. Information from the rain gauge comes from more than 250 automatic stations at 10 minutes resolution, and that from the radar comes from five polarimetric C-band Doppler radars that are suitably located to provide the reliable coverage required in the topographically complex area (see Figure 1). Since CPC is produced operationally, only rain gauge data within Switzerland are used in the algorithm, As a result, an algorithm for the treatment of extrapolation is used in which some radar pixels outside the Switzerland border are used as virtual rain gauges in the merging. Additionally, a convection control scheme is implemented to overcome the limited representativeness of rain gauges during convection events, especially in summer see Sideris et al. (2014b, for details)

The data from both the rain gauge and radar are subjected to substantial quality control before being employed in the CPC algorithm. The gauge data is checked to ensure that recorded values are within climatologically physical limits, they are consistent with those from nearby gauges, they satisfy inter-parameter consistency, and variability tests (MeteoSwiss, 2017). Treatment of the radar data (Germann et al., 2006) involves clutter elimination through a robust algorithm designed for this purpose, visibility correction resulting from orthographic shielding, correction for vertical profile of reflectivity, and bias correction. This is in addition to an automatic hardware calibration of the radars to check the stability/accuracy of the components and a tailored operational scan strategy (20 elevation sweeps every five minutes) crucial in mountainous regions such as Switzerland (Germann et al., 2015).

CPC is available at hourly temporal resolution and a spatial grid of 1 km by 1 km and extends 100-150 km beyond the borders of Switzerland. It has been available since 2005, and 17 years of data is available for this study, from 1$^{\text{st}}$ January 2005 to 31$^{\text{st}}$ December 2021. It has been used in several applications in Switzerland for extreme value analysis (Panziera et al., 2016), climatological studies (Panziera et al., 2018), meteorological forcing of hydrological model (Andres et al., 2016), and has been evaluated in several aspects (Gabella et al., 2017; Panziera et al., 2018; Gugerli et al., 2020). Known limitations of CPC involve the limited length of the data, non-homogeneity of the series due to radar upgrades and evolution of the number of radars, and conditional bias (MeteoSwiss, 2017). Despite these limitations, it is the only sub-daily gridded data available in the study area, and producing a gridded product is beyond the scope of the present study. We note that these limitations are not unique to CPC alone, but common to other radar and radar-reanalysis products, and notwithstanding, they have been used in IDAF modeling e.g. Overeem et al. (2010); Mélèse et al. (2019); Blanchet and Mélèse (2020), or extreme value analysis (Durrans et al., 2002; Allen & DeGaetano, 2005; Wright et al., 2014; Goudenhoofdt et al., 2017; Panziera et al., 2018). This is due to the detailed spatial representativeness they provide, especially in mountainous areas, which is practically unobtainable with rain gauge networks alone.

3 Methodology

3.1 Marginal Distribution for Non-Zero Precipitation Intensities

We use the three-parameter EGPD of Naveau et al. (2016) as the marginal distribution for the non-zero rainfall intensities in the IDAF model. The model is an extension of the classical GPD (which applies only to the excesses of a chosen threshold) to model the entire distribution of precipitation intensities (the low, medium, and extremes). The first advantage of EGPD is that since it is an extension of GPD, it is compliant with extreme value theory, so it behaves like the GPD in the upper tail of the distribution, *i.e.* the same shape parameter see Tencaliec et al. (2020) for demonstration. Secondly, since it makes use of all the non-zero precipitation data, one does not need to worry about the delicate issue of threshold selection that is known with the GPD. Finally, it models the whole range of non-zero precipitation, which has several practical applications in cases where the interest is not only in the largest values but in the medium and low values as well (*e.g.* in simulation frameworks or climatological studies).

We define the random variable I to represent non-zero rainfall intensities. We assume that it follows the EGPD whose cumulative distribution function (CDF) is defined as:

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

with

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

and the probability density function (PDF) is given as

$$f(i) = \frac{\partial}{\partial i}F(i) = \begin{cases} \frac{\kappa}{\sigma}\left[1 - (1 + \xi\frac{i}{\sigma})_+^{-1/\xi}\right]^{\kappa-1}(1 + \xi\frac{i}{\sigma})_+^{-1/\xi-1} & \xi \neq 0, \\ \frac{\kappa}{\sigma}\left(1 - e^{-\frac{i}{\sigma}}\right)^{\kappa-1} e^{-\frac{i}{\sigma}} & \xi = 0. \end{cases}, \tag{3}$$

where $a_+ = \max(a, 0)$, $\sigma > 0$ is the scale parameter, and $\xi \geq 0$ is the shape parameter that controls the upper tail of the distribution. The flexibility parameter, $\kappa > 0$ controls the lower tail. With the addition of only one parameter, κ, compared to the GPD, the distribution is able to model the full range of non-zero precipitation see applications in Evin et al. (2018); Le Gall et al. (2022); Haruna et al. (2022, 2023).

3.2 Space-time aggregation of the data

The total area of Switzerland is 41,285 km^2, and so we have hourly time series of precipitation at 41,285 CPC pixels, each of size 1 km^2. We take each time series and stratify it into four seasons, with winter (Dec-Jan-Feb), spring (Mar-Apr-May), summer (Jun-Jul-Aug), and autumn (Sep-Oct-Nov). This seasonal approach is done to account for the pronounced seasonality in the study area, as done in several studies in the same area (Molnar & Burlando, 2008; Fukutome et al., 2015; Panziera et al., 2018; Evin et al., 2018; Haruna et al., 2022, 2023).

To produce the areal precipitation for use in modeling the IDAF relationships, we aggregate the data into 9 additional spatial scales (area) that includes 9, 25, 49, 81, 169, 289, 529, 729 and 1,089 km^2, . The area corresponds to squares of sides 3, 5, 7, 9, 13, 17, 23, 27, and 33 km, which are illustrated in Figure 1. This leads to a total of 10 series of areal precipitation with area ranging from 1 to 1,089 km^2 centered around each pixel in the study domain. Since CPC is available beyond the borders, it allows us to have spatially aggregated rainfall everywhere in Switzerland (including the pixels close to the border). We comment here that the choice of the squared area is for simplicity and convenience since the CPC data is originally in this geometry. Other choices are possible such as circular or elliptical shapes as discussed in Mélèse et al. (2019).

Next, to build the time series for higher accumulation durations, we use a moving window to aggregate the hourly areal precipitation series into 9 additional durations, that include 2, 3, 6, 10, 12, 16, 24, 48 and 72 hr. We consider durations up to 72 hr (3 days) because according to Froidevaux et al. (2015), while studying catchments larger than 10 km^2, these time scales are the most relevant for flood-triggering precipitation accumulations in Switzerland. The intermediate durations are meant to ensure a good spread on a logarithmic scale. We also apply temporal declustering to reduce serial dependence in the time series as done in several studies (e.g. Naveau et al., 2016; Le Gall et al., 2022; Haruna et al., 2022, 2023). To achieve this, we retain every 3$^{\text{rd}}$ observation in the 1 hr time series, and every 4$^{\text{th}}$, 5$^{\text{th}}$, 8$^{\text{th}}$, 10$^{\text{th}}$, 12$^{\text{th}}$, 16$^{\text{th}}$, 24$^{\text{th}}$, 48$^{\text{th}}$, 72$^{\text{nd}}$, respectively in each time series of 2, 3, 6, 10, 12, 16, 24, 48, and 72 hr. A ratio plot (not shown) of the maximum intensity with and without declustering over all pixels revealed a median value between 0.8 and 0.95, without any systematic evolution with duration and season. This indicates that in certain cases, the highest intensities in each duration were left out as a result of the declustering process. However, given that we are using all the non-zero precipitation intensities, retaining the aggregated time series from the moving window would result in significant serial dependence in the time series. A plot of the lag-1 autocorrelation after declustering (Figure not shown) showed a large reduction in the autocorrelation especially in summer and the transition seasons, whereas it remains relatively high (median of 0.44) in winter, especially for the 1 hr series. Nonetheless, we retain the declustering steps to decrease the potential of omitting the highest intensities.

At the end of the aggregation, we have a total of 100 time series of areal precipitation at each pixel, each for a pair (D, A). Unlike in the case where only block maxima will be used for modeling the IDAF relationships, here, we retain and use all the nonzero precipitation intensities in modeling the IDAF relationships. Although we have the areal precipitation at all the pixels, for computational reasons (an average of 260,000 nonzero observations in summer, at each pixel location), we fit the IDAF model only at a subset of the pixels, by considering every second and third pixel along the latitude and longitude respectively. This results in a total of 7,056 pixels.

3.3 EGPD-IDAF Model

Our assumption is that the random variable of non-zero precipitation intensities for any duration D and area A, $I(D, A)$ follows the EGPD, *i.e.*:

$$I(D, A) \sim \text{EGPD}[\kappa(D, A), \sigma(D, A), \xi(D, A)], \tag{4}$$

where $\kappa(D, A) > 0$, $\sigma(D, A) > 0$ and $\xi(D, A) \geq 0$ are the three EGPD parameters for the duration D and area A.

Let $F_{D,A}(i)$ be the CDF of $I(D, A)$, such that $F_{D,A}(i) = \mathbb{P}(I_{D,A} < i)$, then IDAF curve, which is the T-year return level over duration D and area A is defined by the quantile function of $F_{D,A}$, *i.e.*:

$$i(T, D, A) = \frac{\sigma(D, A)}{\xi(D, A)} \left\{ \left(1 - \left[1 - \frac{1}{T \times \delta(D, A)}\right]^{\frac{1}{\kappa(D,A)}}\right)^{-\xi(D,A)} - 1 \right\}, \tag{5}$$

where $\kappa(D, A) > 0$, $\sigma(D, A) > 0$ and $\xi(D, A) \geq 0$ are the three EGPD parameters for the duration D and area A. T is the return period in years, $\delta(D, A)$ is the average number of non-zero precipitation intensities for duration D and area A per year. We estimate $\delta_{D,A}$ based on the long-term average of the non-zero precipitation intensities per year.

As already highlighted in Section 1, we use the data-driven approach of Overeem et al. (2010) to model the IDAF relationships. The approach involves empirically finding the appropriate regression model to explain the relationship between each of the three EGPD parameters as a function of duration and area. We will now explain our methodology to determine the appropriate regression model.

We begin by considering each pixel and fitting EGPD separately to the 100 aggregated time series of scales (D, A) at that pixel location. We then examine how the three EGPD parameters change with A and D. To model the relationships, we test various regression models using A, D, their transformations; $\log(A)$, $\log(D)$, $\sqrt{A}$, $\sqrt{D}$, as well as some interactions terms. To avoid having a different regression model at each pixel, we compare competing models regionally by assessing their predictive performance in cross-validation. In the end, we retain the following regression models for the EGPD parameters:

$$\begin{aligned}
\log[\kappa(D, A)] &= \beta_{0,\kappa} + \beta_{1,\kappa} A + \beta_{2,\kappa} D + \beta_{3,\kappa} \log(A) + \beta_{4,\kappa} \sqrt{D} + \beta_{5,\kappa} \sqrt{D} \log(A) + \beta_{6,\kappa} D \log(A) \\
\log[\sigma(D, A)] &= \beta_{0,\sigma} + \beta_{1,\sigma} A + \beta_{2,\sigma} D + \beta_{3,\sigma} \log(A) + \beta_{4,\sigma} \sqrt{D} + \beta_{5,\sigma} \sqrt{D} \log(A) + \beta_{6,\sigma} D \log(A) \\
\xi(D, A) &= \beta_{0,\xi} + \beta_{1,\xi} D + \beta_{2,\xi} \log(A) + \beta_{3,\xi} \sqrt{D} + \beta_{4,\xi} \sqrt{D} \log(A) + \beta_{5,\xi} D \log(A),
\end{aligned} \tag{6}$$

where D is in hours and A is in km^2. $\beta_{i,*}$ for $i = 0, 1, ..6$ are the regression coefficients. The scale (σ) and flexibility parameter (κ) both have a log link transformation because of their positive support. They both have seven regression parameters (β_i for $i = 0, 1, ..6$)). The shape parameter ξ has six parameters, making a total of 20 parameters for the complete EGPD-IDAF model for each season and pixel location. We note here that while the number of parameters might appear large, the model is still parsimonious compared to fitting EGPD separately for each time series of (D, A), which amounts to a total of 300 parameters (three (3) EGPD parameters for the 100-time series in our case). In the result Section we will show additional performance comparisons between the 20-parameter EGPD-IDAF model, and the 300-parameter base model. In addition to this, the relative complexity of the model (in terms of parameterization), highlights the inherent difficulty of modeling areal precipitation in mountainous regions, where areal rainfall is less homogeneous in space compared to relatively flat regions. A similar attempt to model IDAF curves in southern France (Massif Central) by Mélèse et al. (2019) highlights similar complexity.

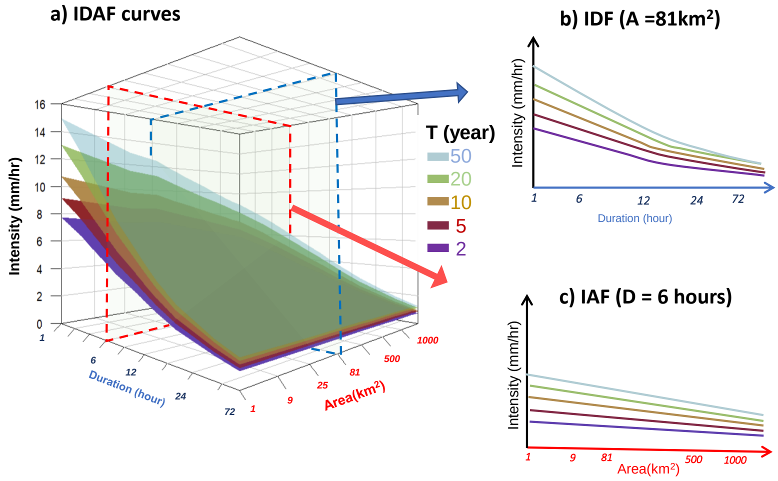

Figure 2. a) Conceptual illustration of IDAF curves in 3-dimension. IDF curves for $A = 81$ km^2 (shown in panel b) are obtained by cutting a plane on the IDAF curves in panel **a** at $A = 81$ km^2 (red-colored broken lines). The IAF curves on panel c) are obtained by cutting a plane at $D = 6$ hr on panel **a** (blue-colored broken lines).

To conclude this section, we illustrate a conceptual plot of IDAF curves in Figure 2. A plot of IDAF curves is 3-dimensional (Figure 2a), with Intensity (I) along the vertical axis, duration (D) along the horizontal axis, and area (A) along the third axis which is perpendicular to the other two axes. For each specific return period (*e.g.* 2-year, 10-year, or 50-year), a curve is shown to visualize how the intensity changes across A and D. However, a much simpler approach is to decouple the 3-dimensional plot into two subplots, each in 2-dimension. The first one shows how the intensities of specific return periods change across durations for a fixed area, i.e. IDF curves (Figure 2b), and the second one, a plot of Intensity-Area-Frequency (IAF) curves, shows how the intensities change across areas for a fixed duration (Figure 2c).

3.4 Model Estimation

Let us call $\boldsymbol{\theta}$ the vector of 20 regression parameters of the EGPD-IDAF model to be estimated at a given pixel location. We estimate $\boldsymbol{\theta}$ by maximizing the censored log-likelihood of the EGPD-IDAF model $l(\boldsymbol{\theta})$, which is given by:

$$l(\boldsymbol{\theta}) = \sum_A \sum_D \sum_{j:i(D,A,j)<C(D,A)} \log\{F_{D,A}\left[C(D,A)\right]\} + \sum_A \sum_D \sum_{j:i(D,A,j)\geq C(D,A)} \log\{f_{D,A}\left[i(D,A,j)\right]\}, \tag{7}$$

where $\boldsymbol{\theta}$ is the vector of the 20 regression parameters to be estimated. $F_{D,A}$ and $f_{D,A}$ are the CDF and PDF of the EGPD associated with (D, A), $i_{(D,A,j)}$ is the precipitation intensity for (D, A) and time step j. $C_{(D,A)} \geq 0$ is the left censoring threshold applied to the data of (D, A). The log-likelihood is finally expressed in Equation 8 as:

$$
\begin{aligned}
l(\boldsymbol{\theta}) = \sum_A \sum_D \sum_{j:i(D,A,j)<C(D,A)} \kappa(D,A) \log\left[1 - \left(1 + \frac{\xi(D,A)C(D,A)}{\sigma(D,A)}\right)^{-\frac{1}{\xi(D,A)}}\right] + \\
\sum_A \sum_D \sum_{j:i(D,A,j)\geq C(D,A)} \log \kappa(D,A) - \sum_A \sum_D \sum_{j:i(D,A,j)\geq C(D,A)} \log \sigma(D,A) - \\
\sum_A \sum_D \sum_{j:i(D,A,j)\geq C(D,A)} \left[1 + \frac{\xi(D,A)i(D,A,j)}{\sigma(D,A)}\right]^{\left[1+\frac{1}{\xi(D,A)}\right]} + \\
\sum_A \sum_D \sum_{j:i(D,A,j)\geq C(D,A)} \left[1 - \left(\left(1 + \frac{\xi(D,A)i(D,A,j)}{\sigma(D,A)}\right)^{-\frac{1}{\xi(D,A)}}\right)\right]^{[\kappa(D,A)-1]},
\end{aligned}
\tag{8}
$$

where $\kappa(D, A) > 0$, $\sigma(D, A) > 0$ and $\xi(D, A) \geq 0$, are the EGPD parameters for (D, A) and the other variables retain their earlier definitions.

The use of the censored likelihood (Equation 8) is mainly to improve the parameter estimation by reducing the influence of the small intensities (Naveau et al., 2016). Without censoring, the smaller intensities influence the parameter estimation, thereby resulting in a gross overestimation of the upper tail shape parameter (ξ). In the equation, both the data above and below the censoring threshold C contribute to the likelihood, albeit in two different ways. The data above C is believed to be observed and so the density function f (Equation 3) is applied to them. For the data below C, it is assumed that their precise magnitude is not known, although they have been observed. All that is known is that they are less than C, and so the distribution function F (Equation 1) is applied. The need for the censored likelihood is likely due to the insufficient flexibility of the three-parameter EGPD model to adequately model the left tail of the distribution or the associated uncertainty in the instrumental recording of very small intensities. A usual censoring approach is to apply a uniform censoring threshold (*e.g.* 2 mm for all the daily data, or 0.5 mm for all the hourly intensities), but as highlighted by Haruna et al. (2023), this is not usually sufficient, and so, an appropriate censoring threshold has to be obtained for each time series. We follow their footstep and estimate a threshold, for each time series of (D, A) that minimizes the squared error between the modeled and empirical quantiles (see Equation 10). This approach usually results in an adequate fit of the model. . We comment that the censoring only applies during inference, afterwards, the model is applied to all the non-zero precipitation intensities, even to the data below the threshold. Additionally, all the goodness of fit criteria is computed on the whole non-zero precipitation, and not only the data above the censoring threshold.

Furthermore, Equation 7 is based on the independence likelihood, which assumes independence in the data. This assumption is unlikely to hold given that we have three levels of dependence in the data; serial dependence within time series of the same (D, A), dependence between time series of different durations (*e.g.* time series of 1 hr and 1 km^2, *versus* time series of 2 hr and 1 km^2), and lastly the dependence between time series of different spatial scales (*e.g.* time series of 1 hr and 1 km^2, *versus* time series of 1 hr and 3 km^2). Despite these, since our target is on the marginal (univariate) return levels, the violation of the independence assumption is unlikely to induce bias in our estimates (Sebille et al., 2017). Additionally, within the framework of modeling IDF curves using generalized extreme value (GEV) distribution, Jurado et al. (2020) showed that little gain in performance is achieved by explicitly modeling the dependence between the data of different durations, in addition to the added complexity. Since their application is with GEV rather than EGPD, an interesting perspective is to investigate this effect with the EGPD. Here we retain the independence assumption to avoid additional complexity to our model which already has 20 free parameters. . Finally, to avoid underestimating uncertainties in our model, which is one of the main consequences of the independence assumption, we resort to block-bootstrapping for uncertainty assessment (see Section 3.5).

3.5 Uncertainty Assessment

In order to assess uncertainty in the EGPD-IDAF model, we use the block bootstrap approach (Kunsch, 1989). The principle of the block bootstrap involves dividing the time series into blocks of consecutive observations. Resamples are then generated by randomly selecting blocks with replacements and concatenating them to create a bootstrap sample. By preserving the block structure, the block bootstrap can capture the dependence structure of the original data. This approach is suitable for uncertainty estimation in our case, where we made the independence assumption in the likelihood estimation of the parameters. The block bootstrap method was used for uncertainty estimation by Overeem et al. (2010) in IDAF curves modeling, and by Overeem et al. (2009); Haruna et al. (2023) in IDF curves modeling.

To apply the block bootstrap approach, we take the seasonal time series at each pixel and estimate the uncertainty by following the outlined steps below:

1. Aggregate the time series into the 10 durations and 10 areas, resulting in a total of 100 time series, each for a pair of duration and area (D, A). Decluster each of the series according to the declustering procedure explained in Section 3.4. We call this sample M_{orig}.
2. Randomly select blocks of size 2 weeks with replacement, G times, to form the resampled time series (M_{boot}). Both M_{orig} and M_{boot} have the same dimensions. The block bootstrapping ensures that we keep the data of the different durations D and areas A together, and hence the dependence structure. We use a block size of 2 weeks, beyond which the autocorrelation in the data does not decrease, as done in Haruna et al. (2023) for the same study area in the case of IDF curve modeling.
3. Fit the EGPD-IDAF model to the data in M_{boot} and estimate the intended return levels.
4. Repeat steps 2 to 3 a total of 300 times to obtain the bootstrap distribution of the return levels. Finally, compute the 95% Confidence Interval (CI) of the return levels by the percentile method. This is done by taking the empirical 0.025 and 0.975 quantiles of the bootstrap distribution of the return levels obtained in step 4.

As a measure of model precision, we compute the normalized width of the 95% CI of a T-year return level estimate (Shehu & Haberlandt, 2023). For a given pixel location s, it is computed from:

$$\mathrm{n95CI}_{\mathrm{width}\ ,s} = \frac{r_{T,97.5\%} - r_{T,2.5\%}}{\bar{r}_{T,}} \tag{9}$$

where $r_{T,p\%}$ is the p% quantile of the 300 bootstrap estimates of the T-year return level (r_T) and $\bar{r}_T$ denotes the average of the 300 estimates. The normalization is to enable the comparison of uncertainty width across intensities of different scales and return periods.

3.6 Goodness of fit of the EGPD IDAF model

To assess the goodness of fit of the EGPD-IDAF model, we compute the normalized root-mean-square error (NRMSE) and the normalized bias (NBias) at each pixel s

and spatiotemporal scale (D, A). To focus on the high intensities, the criteria are computed only on the exceedances above a 1-year return level, computed using the Weibull plotting position, defined as $\frac{j}{n+1}$ with j being the rank (from largest to smallest) and n is the sample size. The normalization allows comparison of the score across intensities of different spatiotemporal scales (D, A). For a given pixel s, the two criteria are given as:

$$\text{NRMSE}_s = \frac{\left\{\frac{1}{n_s}\sum_{j=1}^{n_s}\left(\hat{r}_{s,T_j} - r_{s,T_j}\right)^2\right\}^{1/2}}{\overline{r_s}} \tag{10}$$

$$\text{NBias}_s = \frac{\frac{1}{n_s}\sum_{j=1}^{n_s}\left(\hat{r}_{s,T_j} - r_{s,T_j}\right)}{\overline{r_s}} \tag{11}$$

where n_s is the sample size, r_{s,T_j} is the j^{th} largest empirical quantile with return period $T_j = \frac{n_s+1}{j\times\delta}$, δ is the average number of non-zero precipitations for (D, A) per year, $\hat{r}_{s,T_j}$ is the corresponding T_j return level estimated from the EGPD-IDAF model. The denominator is the average of the exceedances

NRMSE measures the accuracy of a given model in predicting the empirical quantiles. A good model should have NRMSE $= 0$, and the smaller the score, the better the model. NBias measures the ability of the model to avoid systematic underestimation (NBias < 0) or overestimation (NBias > 0) of the empirical quantiles. NBias $= 0$ means an unbiased model.

3.7 Cross validation

A natural question to ask is whether the EGPD-IDAF model which links the EGPD parameters with duration and area is a better model, in terms of some performance indicators, compared to fitting the EGPD model separately to each time series of spatiotemporal scale (D, A). The two models will henceforth be referred to as the global model and the base model, respectively. To answer this, we compare the two models in a split-sample cross-validation framework. We will start by describing the cross-validation framework, and then introduce the criteria for measuring the performance.

In the split sampling cross-validation, we consider each pixel and divide the time series into two subsamples of the same length but on different randomly chosen years. We consider the first sub-sample, aggregate the data into the 10 durations and 10 areas, and fit the two competing models, *i.e.*, the base model and the global model. We then assess how the two models perform on the second sub-sample (validation sample). A good predictive model should perform well in the data not used in training it. We do the same on the second sub-sample (use it as the training sample, and the first sub-sample as the validation sample). Since the split sampling is done randomly, we repeat the procedure 40 times to address sampling bias. We apply the same procedure to all the pixels in the study area. We then select the method that has the best regional performance (average of the scores over all the pixels.)

We use some well-chosen predictive performance criteria to measure the performance of the models. The criteria have seen wide applications in the literature (see Garavaglia et al., 2011; Renard et al., 2013; Blanchet et al., 2015; Evin et al., 2016; Haruna et al., 2022, 2023). We give a brief overview of the criteria, while details can be found in the cited references.

- Robustness: The Robustness criteria, SPAN, measures the ability of a model to give similar estimates of a high return level when data from two different calibration periods are used to train the model (Garavaglia et al., 2011). At a given pixel (s) and for a spatiotemporal scale (D, A), SPAN is computed as:

$$\mathrm{SPAN}_{s,T} = \frac{2\left|\hat{r}^{(1)}_{s,T} - \hat{r}^{(2)}_{s,T}\right|}{\left(\hat{r}^{(1)}_{s,T} + \hat{r}^{(2)}_{s,T}\right)} \tag{12}$$

where $\hat{r}^{(1)}_{s,T}$ and $\hat{r}^{(2)}_{s,T}$ are the T-year return levels estimated from sub-sample 1 and 2 respectively at pixel s. A SPAN of 0.5 means that the absolute difference between the two return levels is half of their average.
A regional value of SPAN, over Switzerland, is computed as $\mathrm{SPAN}_{\mathrm{reg},T} = 1 - \frac{1}{N}\sum_{s=1}^{N}\mathrm{SPAN}_{s,T}$, where $N = 7,056$ is the total number of pixels. A perfectly robust model should have $\mathrm{SPAN}_{\mathrm{reg},T} = 1$.

- Reliability in predicting the maximum value: At a given pixel (s) and for a given spatiotemporal scale (D, A), 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 as follows:

$$\mathrm{FF}^{(12)}_{s} = \left[\hat{\mathrm{F}}^{(1)}_{s}\left(\max{}^{(2)}_{s}\right)\right]^{n^{(2)}_{s}} \tag{13}$$

where $\mathrm{FF}^{(12)}_{s}$ is the cross-validation criteria computed at pixel s, by predicting the probability of the maximum value in sub-sample 2, of sample size $n^{(2)}_{s}$ using the model fitted on the sub-sample 1. $\mathrm{FF}^{(21)}_{s}$ is computed symmetrically.
As noted by Renard et al. (2013) and Blanchet et al. (2015), if the fitted model is a good estimate of the true distribution of the data, $\mathrm{FF}^{(12)}_{s}$ should be a realization of a uniform distribution. Hence, the difference in the area, noted $diff$, between a theoretical uniform distribution and that of the $N = 7,056$ values of $\mathrm{FF}^{(12)}_{s}$ (computed over the N pixels), should be close to zero. $\mathrm{FF}_{\mathrm{reg}}$ at the regional scale, given as $1 - diff$, should therefore take a value of 1 for a reliable model and 0 for a completely unreliable model; the lower the value the less reliable the model is.

- The reliability/accuracy over the entire observations: While the previous reliability score (FF), and SPAN focus on extremes only, it is important that the model is also reliable in the bulk of the distribution, especially given that we use the EGPD. To measure the reliability of a model in predicting all the observations in cross-validation, we use the normalized root mean square error (NRMSE_CV), which is expressed as:

$$\mathrm{NRMSE\_CV}^{(12)}_{s} = \frac{\left\{\frac{1}{n^{(2)}_{s}}\sum_{j=1}^{n^{(2)}_{s}}\left(r^{(2)}_{s,T_j} - \hat{r}^{(1)}_{s,T_j}\right)^2\right\}^{1/2}}{\overline{r^{(2)}_{s}}} \tag{14}$$

where $\mathrm{NRMSE\_CV}^{12}_{\mathrm{s}}$ is the score computed at pixel s, $n^{(2)}_{s}$ is the sample size of the second sub-sample, $r^{(2)}_{s,T_j}$ is the empirical quantile with return period $T_j = \frac{n_s+1}{j\times\delta}$, δ is the average number of non-zero precipitations for (D, A) per year in sub-sample 2, $\hat{r}^{(1)}_{s,T_j}$ is the corresponding T_j return level estimated from the model fitted on sub-sample 1. The denominator is the mean of non-zero precipitation in sub-sample 2 at pixel s computed as $\frac{1}{n^{(2)}_{s}}\sum_{j=1}^{n^{(2)}_{s}} r^{(2)}_{s,T_j}$.
Similar to the other criteria, the regional score for each spatiotemporal scale (D,A), computed over the N pixels, is given as $\mathrm{NRMSE\_CV}^{(12)}_{\mathrm{reg}} = 1 - \frac{1}{N}\sum_{s=1}^{N}\mathrm{NRMSE\_CV}^{(12)}_{s}$. The other score, $\mathrm{NRMSE\_CV}^{(21)}_{\mathrm{reg}}$ is computed symmetrically. $\mathrm{NRMSE\_CV}_{\mathrm{reg}} = 1$ indicates a perfectly accurate model (the model accurately predicts the empirical return levels).

4 Results and Discussion

4.1 Evaluation of CPC data

We begin by checking how the statistics from the CPC data compares with those from the rain gauge through a point-to-pixel comparison. This involves comparing the time series from a gauge to the time series from a CPC pixel at the location of the gauge. We consider 79 stations, with no missing data from 2005 to 2020 (the period of overlap of both datasets), and in each case, the data from the gauge is considered the "truth". The location of the stations is shown in Figure A3. We comment here that the comparison is not entirely independent since most of the stations (69 out of the 79) are already utilized for correcting the radar data to produce CPC. However, since some differences remain between the two, it would be interesting to see how the statistics from CPC compare to those from the rain gauge before using them in modeling the IDAF relationships. The difference arises mainly due to the nugget effect in the variogram model, the inherent scale differences between radar and rain gauge measurements, and the convection control scheme in summer (Sideris et al., 2014a).

The comparison is in two steps, in the first step, we compare the two time series using some chosen criteria and in the second step, we fit EGPD to both time series and compare the 20-year return level estimate. The result of the comparison is presented in the following subsections.

4.1.1 Comparison on the empirical values

i) Criteria on all observations

Following the work of Zambrano-Bigiarini et al. (2017), we use the three sub-components of the Kling-Gupta-Efficiency (KGE) criterion (Kling et al., 2012) to compare the two datasets (see Equation A1 in Appendix A). The first is the bias (the tendency of CPC to under or overestimate the gauge data). The second is the variability ratio, which measures the under or over-dispersion of CPC data compared to the gauge. The third component measures the linear correlation between the two time series. For a perfect match between the gauge and CPC, all the criteria should be equal to 1. The criteria are computed based on all the data, including zeros.

The boxplot of the correlation coefficient is shown in Figure 3a for the four seasons and eight aggregation durations (1, 2, 3, 6, 12, 24, 48, and 72 hr). Generally, there is a good temporal correlation between the two data sets for all seasons and durations (median > 0.9). For all seasons, the correlation increases with the aggregation duration. Summer generally exhibits the lowest correlation irrespective of the duration, due to the localized and isolated nature of convective events that are likely to be missed by the rain gauge. The bias and variability scores are given in Appendix A. There is generally a tendency toward overestimation of the data (Figure A1a by the CPC for all seasons (median > 0), again the bias is more pronounced in summer compared to the other seasons. Lastly, the dispersion bias (Figure A1b is generally negative with a median of 5% for all seasons

ii) Criteria on extremes

Next, we evaluate the ability of CPC to correctly detect extreme precipitation as measured by the gauge. Extremes here are defined as the exceedances of a 2-year return level within the 16 year record. We compute three criteria similar to Panziera et al. (2018). The first criterion measures the bias in extreme precipitation totals. The second criterion computes the probability of detecting extremes (POD), *i.e.*, the ability of CPC to classify events as extremes, given that they are also extremes according to the gauge. Lastly, we compute the false alarm ratio (FAR), which measures the rate at which CPC classifies events as extremes when they are not extremes according to the gauge. For a per-

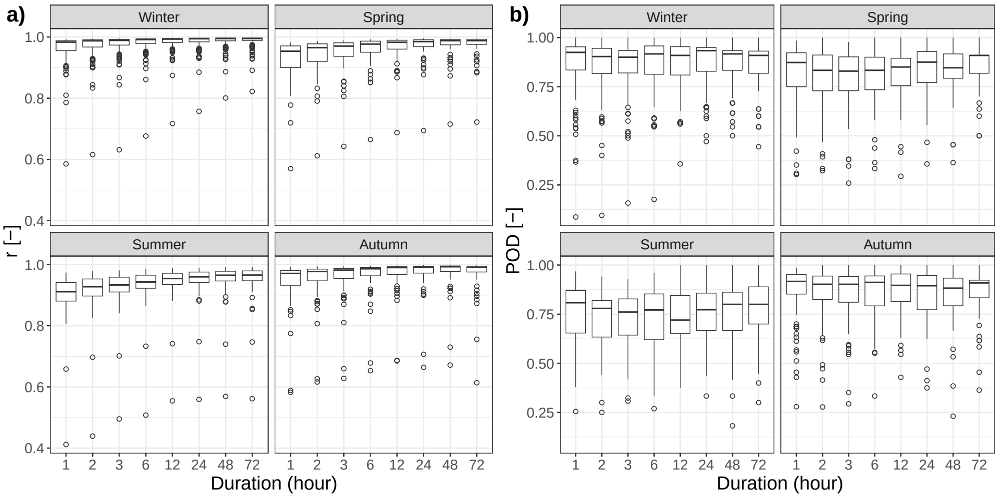

Figure 3. Boxplots of linear correlation (a) and probability of detection (POD) (b) for the four seasons. Each boxplot contains 79 points, 1 point for each pair of rain gauge and the underlying CPC pixel.

fect agreement, bias should be equal to 0, POD should be equal to 1, and FAR should be equal to 0.

Figure 3b shows the seasonal POD scores. The median of the score ranges from 0.7 to 0.99, which means that 70% to 99% of the gauge extreme events are correctly classified as extremes by the CPC. Again, summer shows the lowest values compared to the other seasons.The seasonal boxplots of FAR are shown in Appendix A (Figure A2b) and the median FAR decreases with duration, which shows the agreement improves as the intensities are aggregated to higher durations. The median of the bias in the extremes precipitation totals (Figure A2a is less than 5% for all cases. Summer in this case has the lowest bias but shows the most spread in the case of short durations.

4.1.2 Comparison of return level estimates

In the final phase, we compare the 20-year return level estimates from the two datasets. We fit EGPD to each dataset and estimate the 20-year return level. Figure 4 shows the relative bias in the 20-year return level estimates. A positive bias indicates that the CPC estimates are higher than the gauge estimate. In general, the bias for durations greater than 6 hr is close to zero. For the 1 hr duration, however, there is a tendency to have lower estimates with CPC for all seasons, except summer which shows the opposite.

The map of the relative bias, as well as correlation coefficient (r), POD, and bias in total precipitation (β), is shown in Figure A3. It can be observed that in general, there is no distinct spatial pattern, except for β, which shows overestimation in the north and underestimation in the south in all seasons except summer. In addition, three more stations show a quite large disagreement with the CPC data. They are located at La Dole (elevation 1669 m), Col Du Grand Saint-Bernard (2472 m), and Säntis (2501 m) in the west, southwest, and northeast respectively. MeteoSwiss indicates that rain gauge measurements at these stations are subjected to large uncertainties since the stations are not shielded (*e.g.* influence of wind and snow drift) (MeteoSwiss, 2023).

In conclusion, the result shown in these sections is aimed at checking how the statistics of CPC data compares with those from the rain gauge and to understand which di-

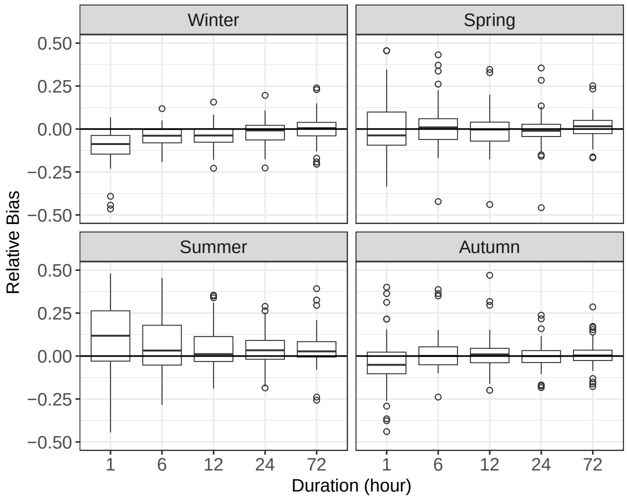

Figure 4. Boxplots of relative bias in a 20-yr return level estimate for the four seasons. Each boxplot contains 79 points, 1 point for each pair of rain gauge and the underlying CPC pixel.

rection they take before using the in modeling the IDAF relationships. Despite the noticeable disagreements, there is generally, a good agreement between the two datasets, given the inherent uncertainties in both databases (gauge *versus* radar reanalysis). As mentioned before, although CPC is corrected using the rain gauge data, some differences remain, mainly due to the nugget effect in the variogram model, the convection control scheme in summer (Sideris et al., 2014a), and the measurement scale difference between radar and raingauge. As emphasized in Section 2.2, it is beyond the scope of this study to develop a new gridded dataset for this topographically complex study area. CPC presents the only dataset at the sub-daily temporal resolution in the study area, and it brings the required spatial information needed for modeling IDAF, which cannot be obtained from rain gauges due to their limited spatial representativity. In the remainder of the article, only CPC is used to build the IDAF models.

4.2 EGPD parameters as a function of Duration and Area

The purpose of this section is to show the complex relationship that exists between the EGPD parameters and duration D, and area A. Moreover, it aims to showcase that the EGPD-IDAF model is flexible enough to adequately capture this complexity.

As an illustration, we focus on a single pixel located at an elevation of 1,351 m in Adelboden, west of the Bernese Alps (see Figure 1). The estimated EGPD parameters as a function of D and A for the four seasons are shown in Figure 5. In each panel. the lines represent the modeled relationship using the EGPD-IDAF model, and the points show the parameter estimates using the base model. It can be observed from the figure that there is a clear season-dependent relationship of the parameters with D and A. We will focus on winter and summer since the other two seasons present behavior in between the two.

Starting with the top row, the flexibility parameter κ that controls the bulk and lower tail of the distribution shows a clear relationship with both D and A. For large A, it shows a positive monotonic relationship with D, while for small A, it shows a non-monotonic relationship, decreasing and then increasing with D. This non-monotonic relationship with D was also observed by Haruna et al. (2023) while modeling IDF curves in the study area using rain gauge data. Next, looking at the middle row, the scale pa-

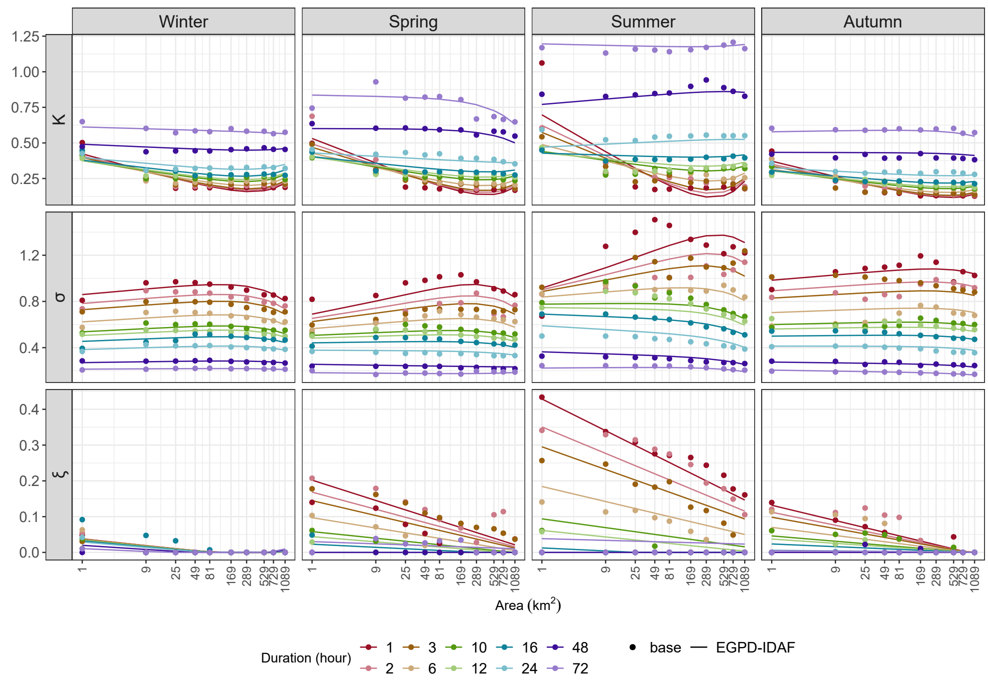

Figure 5. EGPD parameters as a function of duration D and area A at a pixel located in Adelboden (elevation of 1354 m, see Figure 1), for the four seasons (columns). The first row is for $\kappa(D, A)$, the second row is for $\sigma(D, A)$, and the last row is for $\xi(D, A)$. In each panel. the lines represent the modeled relationship using the EGPD-IDAF model, and the points show the parameter estimates using the base model. The lines and points are colored by duration.

rameter σ decreases with an increase in D for all A in both seasons. It however shows a non-monotonic relationship with A, which also varies with D. Finally, in the bottom row, the upper tail shape parameter ξ shows a season-dependent relationship with D and A. The strongest relationship is observed in summer, where it decreases with both D and A. While it shows exponential tail ($\xi \approx 0$) for $D > 24$ hr irrespective of A, it shows a heavy tail ($\xi > 0.1$) for $D = 1$ hr even at $A = 1089$ km^2. In winter, however, it shows an exponential tail for all D and A.

We highlight here that the pattern of relationship observed at this pixel location is not general all over Switzerland, and our aim is just to illustrate the complexity of the relationship by focusing on this pixel. For instance, for some locations, σ can show a positive-monotonic relationship with A for all D. The shape parameter ξ can remain positive for all D and A in winter, increases with D, or increases with A. This intricate relationship of the parameters with D and A underscores the difficulty and complexity of modeling relationships of areal precipitation in topographically complex locations, due to the regional heterogeneity of the rainfall process. Despite this, as seen in Figure 5, the proposed regression models in Equation 6 are flexible enough to capture the observed trends in the points corresponding to the base model estimates. In the next section, we will present the results of the goodness of fit of the EGPD-IDAF model at all the pixel locations in Switzerland.

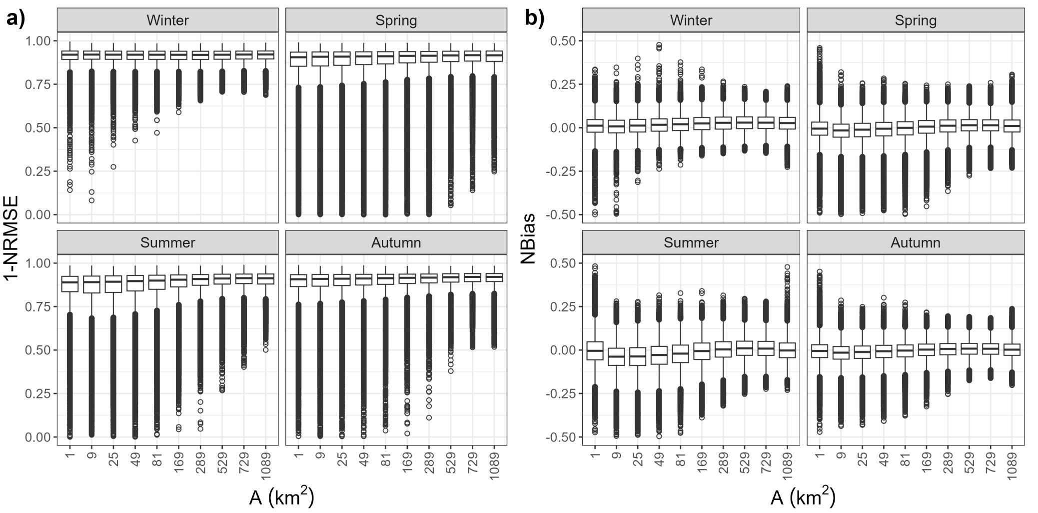

Figure 6. Goodness of fit of the EGPD-IDAF model computed on extremes, defined as the exceedances of a 1-year return level for each (D, A). a) Boxplots of (1-NRMSE). Note that the vertical axis is cut at zero although negative values exist. The negative values account for less than 0.06 % of the scores b) NBias for the four seasons. Each boxplot contains 7,056 by 10 points, 1 point each for a pixel and a spatiotemporal scale (D, A).

4.3 Goodness of fit of the EGPD-IDAF model for extremes

We fitted the 20-parameter EGPD-IDAF at each pixel and for each of the four seasons and assessed the goodness of fit of the model using NRMSE and NBias (see Section 3.6). To assess the model's performance on the extreme values, we computed the two criteria on extremes only, defined as the exceedances of a 1-year return level. The normalization allows comparison of the score across intensities of different spatiotemporal scales (D, A). Figure 6 shows the results for the two criteria. In both figures, each of the four panels shows the score for a given season. The results are shown as a function of area (A), and so each boxplot contains the results of 7,056 pixels for the 10 aggregation durations of a given A. Figure 6a shows the result for 1 - NRMSE and so the ideal score is 1. For all seasons, the median of the score is greater than 0.8 and the score gets better as A increases, possibly because as we aggregate the process over larger spatial domains, the variability decreases and the fit of the model gets better. While the score is relatively the same across seasons, summer shows slightly lower scores for smaller A (as seen from the width of the boxplot). These smaller scales in summer largely correspond to those experiencing more intense and skewed rainfall due the convective events. As such the shape parameter is heavy, and so the fit becomes more difficult. In Figure 6b, the median of the NBias remains close to zero which means that the model does not consistently overestimate or under-estimate the empirical quantiles. As with the other score, the variability around zero decreases as A increases.

These two scores show that the model is able to adequately reproduce the areal precipitation across durations in the study area. It shows good predictive performance as judged by the NRMSE, and doesn't show a systematic tendency to overestimate or underestimate the empirical values (NBias).

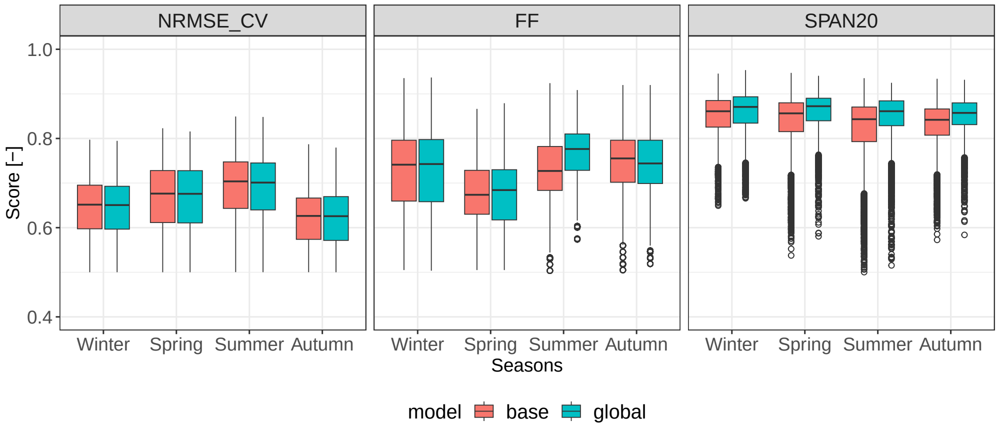

Figure 7. Boxplots of the cross-validation criteria for the four seasons. Each boxplot contains $2 \times 100 \times 40$ points for NRMSE_CV and FF (2 regional scores (*i.e.* $\mathrm{FF}_{\mathrm{reg}}^{(12)}$ and $\mathrm{FF}_{\mathrm{reg}}^{(21)}$) for each pair of ($D$, A), and 40 resamplings). In the case of SPAN20, each boxplot contains 100×40 points (1 regional score for each pair of (D, A) and 40 resamplings)

4.4 Comparison of the EGPD-IDAF model with the base model

4.4.1 Cross validation results

The result of the split sampling cross-validation for the comparison of the EGPD-IDAF model (global model) and the base model is shown in Figure 7. This Figure shows the three cross-validation scores (NRMSE_CV, FF, and SPAN20), one panel for each criterion. As a reminder, the 20-parameter global model allows the linkage of the EGPD parameters with duration and area, the base model fits a separate EGPD model to each of the 100-time series of spatiotemporal scales (D, A). The best model in each case has a score of 1. Starting with the first panel from the left, NRMSE_CV is nearly the same for both models, which means that both models have the same accuracy in predicting the whole non-zero precipitation. Next, the FF criterion also shows similar performance by the two models. A noticeable exception is in summer, where the global model shows better performance. Hence, according to this criterion, while the models have similar reliability in predicting the maximum value, the global model is slightly better in summer. Finally, SPAN20 shows that a better performance is obtained with the global model for all seasons compared to the base model. This means that the global model gives a more stable estimate of a 20-year return level when the calibration sample is changed. Finally, the heat maps in Figure B1 show the median scores for each (D, A) pair. Poor scores are typically obtained for short spatiotemporal scales.

In summary, both models have similar reliability in their predictive ability (NRMSE_CV and FF), however, the global model is more robust in 20-year return level estimations (SPAN20). The robustness of the global model can be explained by the fact that the model is trained with much more data (all the 100-time series are pooled in the parameter estimation), compared to the base model.

4.4.2 Uncertainty

Since the two models have similar predictive performance, we also go a step further to compare the models in terms of their uncertainty estimates. While a good model should give correct predictions, the uncertainty of the prediction should not be too large. Figure 8 shows the $\mathrm{n95CI}_{width}(\%)$) (Equation 9) of a 50-year return level estimate with

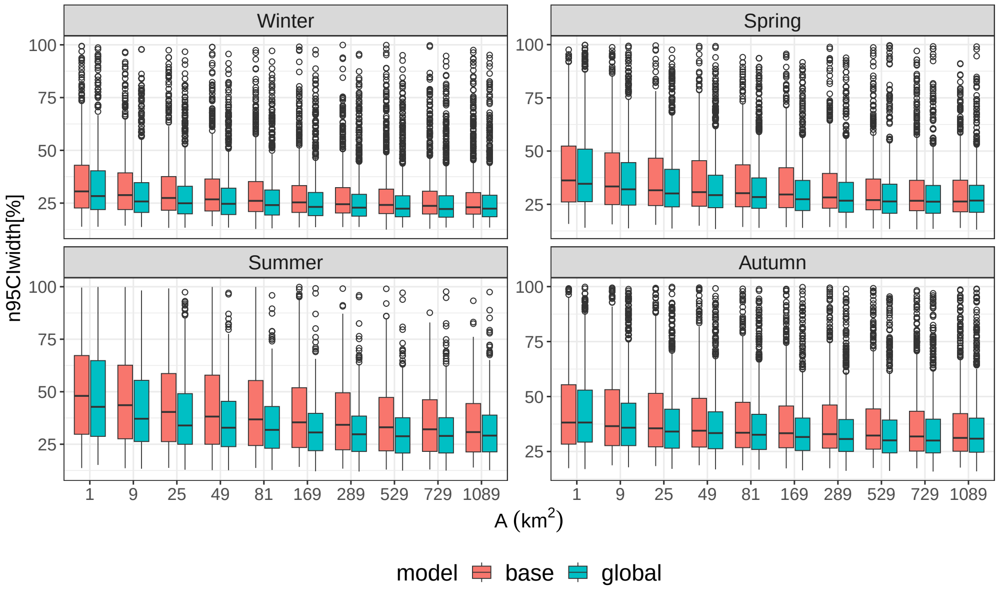

Figure 8. Boxplots of n95CI$_{width}$(%) for a 50-year return level estimate using the base model and the global model. Each boxplot contains 7,056×10 points (7,056 pixels, 10 durations)

both the global and base model. The smaller the score, the better the preciseness of the model (less uncertainty). Each panel in this Figure shows the result for a given season. The results are shown as a function of area (A), and so each boxplot contains the results of 7,056 pixels for the 10 aggregation durations of a given A. For all seasons, the global model has the smallest values of the n95CI$_{width}$ as seen from the median and width of the boxplots. The lower values of the global model mean less uncertainty compared to the base model, which can be explained by the fact that the global model is trained with more data, and this translates to less uncertainty (narrower confidence intervals). Two more comments can be made from Figure 8. First, for all seasons, the uncertainty decreases with A, which can be a result of the smoothing effect due to spatial averaging. Secondly, some inter-seasonal differences are noticeable, with summer (winter) having the highest (lowest) uncertainty. A possible explanation is that since more extremes are observed in summer (especially at sub-daily time scales), the uncertainty is expected to be larger. For a given return period, the magnitude of the return levels in summer at the small scale is larger compared to the other seasons, and so will the uncertainty.

To conclude, the results shown so far demonstrate that the modeled EGPD-IDAF can be used in the study area. It has adequate goodness of fit, is reliable and robust in prediction, and has relatively low uncertainty in estimation. With this validation, we will now proceed to showcase examples of IDAF curves constructed from the EGPD-IDAF model at some pixel locations in the next section.

4.5 IDAF curves

Figure 9 shows an application of the EGPD-IDAF model to build summer IDF and IAF curves at the pixel located in Adelboden. This pixel has been introduced in Section 4.2 and is shown in Figure 1. Starting with the top row (Figure 9a), IDF curves are shown in the case of four aggregation areas $i.e$ $A = 1$, 25, 529, and 1,089 km^2 (1 column each). In each column, the colored lines represent the T-year return level estimate using the EGPD-IDAF model across duration for $T = 2$, 10, and 20 years. The corresponding empirical

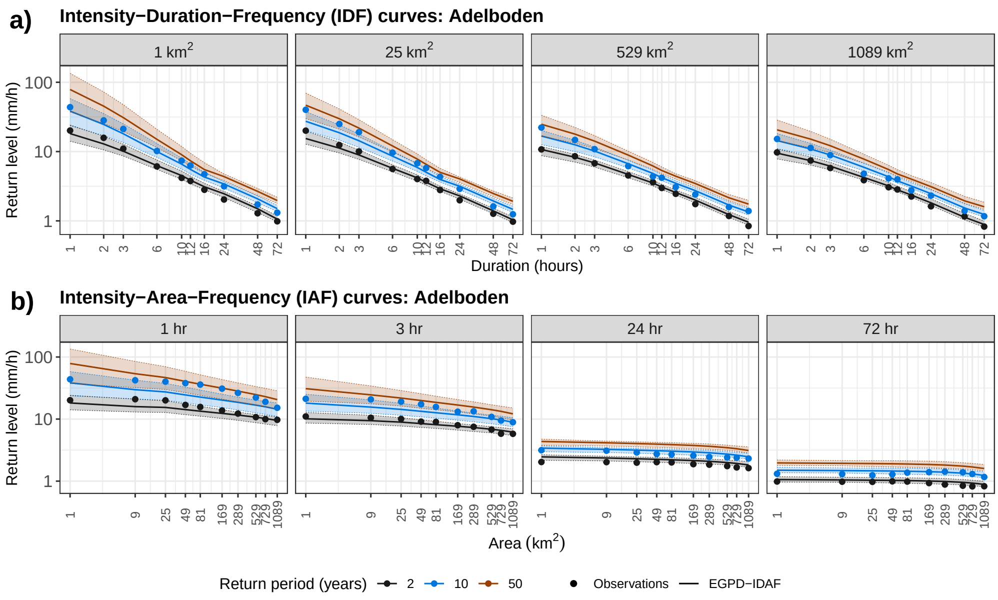

Figure 9. Application of the fitted EGPD-IDAF model at a pixel location in Adelboden for the summer season. The top row (a) shows some IDF curves for four spatial scales (one per column). The bottom row (b) shows the IAF curves for four temporal scales (one per column). The lines and the points show the modeled and empirical levels respectively, colored by their return periods. The colored envelopes are the 95% confidence intervals of the model estimates obtained by block bootstrap. The 50-year empirical values are not shown due to the short record length of the data

levels, computed using the Weibull plotting position, are shown by the colored points. It can be seen that the EGPD-IDAF model correctly predicts the observation as they are within the 95% CI (shown by the colored envelopes). We also see that the uncertainty (indicated by the width of the bounds) is higher for shorter durations. Finally, irrespective of the spatial scale (A), the return levels decrease as the duration increases.

The second row (Figure 9b) shows the IAF curves for four temporal scales, $D =$ 1, 3, 24, and 72 hr. While the model shows an adequate performance for longer durations ($D \geq 24$ hr), the fit is less good in the case of shorter durations, especially for higher return periods. Looking at the IAF curves for short durations, we see that the return levels tend to decrease with an increase in the spatial scale. For longer durations, however, the return levels have nearly the same magnitude (flat IAF curves) irrespective of the spatial scale. A possible explanation is that at short durations, the rainfall events are more localized (typical of convective events) and so the magnitude decreases due to spatial averaging. For longer durations, however, the rainfall is more homogeneous in space (typical of frontal events), with no significant variations in rainfall intensity, leading to similar marginal distributions for the considered areal rainfall.

To explore the regional and seasonal variability of the IDAF relationships, Figure 10 shows the autumnal IDF curves (top row) and IAF curves (bottom row) at a location in Sion, in the inner valleys, southwestern Switzerland. This location is at a relatively low elevation of 482 m and experiences low-intensity rainfall due to the shielding effect of the Alps on both sides. Remarkably, the IDF and IAF curves at this pixel location exhibit a distinctive behavior, diverging from the conventional trend of decreasing return levels with increasing spatial scales. The IAF curves (bottom row) highlight this feature. It can be seen that the IAF curves for 1 hr are nearly flat, and the IAF curves for $D \geq 24$ hr have positive slopes. A plausible explanation of this behavior is that rainfall, of short and long duration, is less intense at the pixel location compared to its neighborhood locations, which are at a higher altitude (see Figure 6 to 8 of Panziera et al. (2018)). As a consequence, more intense rainfall is observed as the rainfall is spatially aggregated around the pixel location. Figure 1 shows that a spatial window of 1,089 km^2, centered around the pixel (elevation of 480 m), extends well beyond the valley into the Bernese alpine slopes (elevation up to 2,400 m). This seasonal and regional variability highlights the complexity of modeling areal precipitation in the study area due to the complex topography.

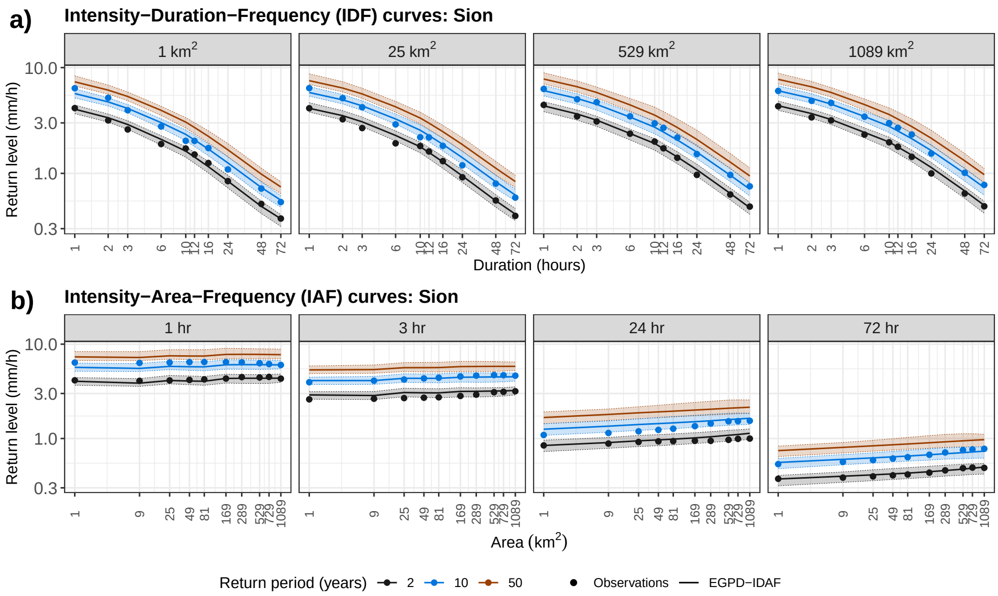

Figure 10. Same as Figure 10 but for autumn at a location in Sion in the Canton of Valais (see Figure 1).

4.6 Areal rainfall hazard in Switzerland

In this last section, we use the EGPD-IDAF model to assess areal rainfall hazards in the study area. We investigate the 20-year return level for two spatiotemporal scales, specifically the scales ($D = 1$ hr, $A = 1$ km^2) and ($D = 24$ hr, $A = 1{,}089$ km^2). The corresponding maps of the seasonal 20-year return level are shown in Figure 11a and Figure 11b respectively. For the scale ($D = 1$ hr, $A = 1$ km^2), we observe that the highest return levels occur during the summer months, while the lowest values are observed in winter. This can be attributed to the prevalence of convective rainfall during summer. We also see significant regional variability across all seasons, particularly during summer. The Ticino region in the south of the Alps, the Bernese Alps in the north, and the Jura Mountains consistently exhibit the highest return levels. Conversely, the inner valleys in Valais and Grisson, due to their location between mountains, depict the lowest values as they are shielded from both directions.

Moving to the scale ($D = 24$ hr, $A = 1{,}089$ km^2), we see a shift in the seasonal and regional variability of the 20-year return level. The black colored square shows the spatial coverage of $A = 1{,}089$ km^2, centered around a pixel in Adelboden. The map in Figure 11b shows that the largest values are observed in Ticino, regardless of the season.

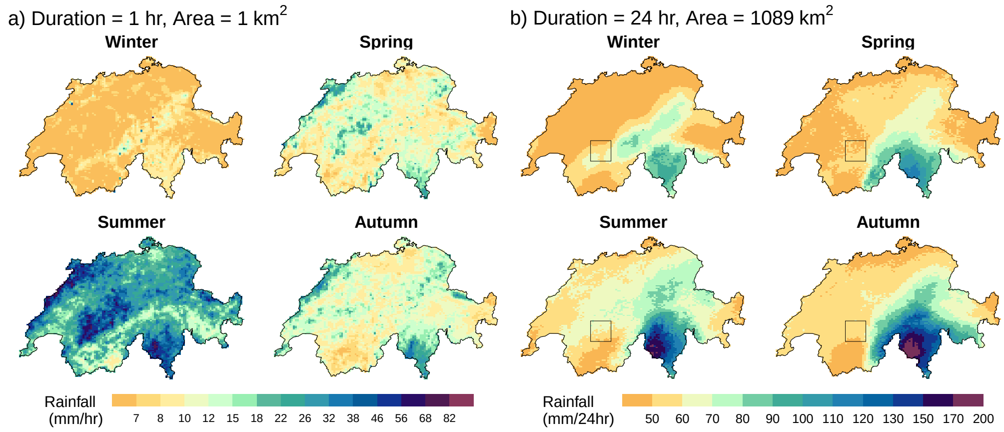

Figure 11. Map of seasonal 20-year return level obtained with the EGPD-IDAF model for the spatiotemporal scales a) ($D = 1$ hr, $A = 1$ km^2) and b) ($D = 24$ hr, $A = 1{,}089$ km^2). The black-colored square in b) shows exemplarily the maximum extent of the square window used for data aggregation, *i.e.* 1089 km^2.

The Ticino region consistently exhibits the highest levels of extreme precipitation in Switzerland. In the north of the Alps, the plateau displays lower levels compared to the pre-Alps (along the Glarus Alps). These results emphasize the influence of spatiotemporal scale on the seasonality and regional patterns of rainfall hazard in Switzerland. Smaller scales show a higher hazard during summer, while larger scales demonstrate a higher hazard during autumn, particularly in the Ticino region. It is important to note that the Ticino region consistently remains at a higher hazard of extreme precipitation, irrespective of the scale. Conversely, the inner valleys in Valais and Grisson exhibit lower susceptibility to extreme precipitation events.

In conclusion, this result provides insights into the seasonal and regional patterns of rainfall hazards in Switzerland, highlighting the importance of considering spatiotemporal scales when assessing extreme precipitation hazards. It is important to note that while this assessment focuses on the hazard of extreme precipitation, it is essential to consider other factors such as exposure and vulnerabilities at specific locations to fully evaluate the overall risk.

5 Conclusions

This paper focused on modeling the relationship of extreme precipitation across duration and area through Intensity-Duration-Frequency (IDAF) curves in Switzerland. We proposed a novel approach to model IDAF curves, by using all the non-zero (low, medium, and extremes) precipitation data, instead of only the extremes. To build the IDAF curves, we used the EGPD as the parametric distribution for the precipitation intensities. The EGPD has the key advantage of adequately modeling the entire distribution of non-zero precipitation while being compliant with extreme value theory in both tails. We followed the footsteps of Overeem et al. (2010) to model the IDAF curves through a data-driven approach. This approach involves modeling the EGPD parameters as a function of area and duration, with the form of the relationship being empirically determined from the data. We used 17 years of data from the radar-reanalysis product, `CombiPrecip` (CPC) (Sideris et al., 2014a) to build the EGPD-IDAF model at each pixel location in the study area.

We used the model to assess areal rainfall hazards for some spatiotemporal scales in Switzerland. More than any region, the results showed that Ticino, located south of the Alps, is the most exposed to extreme precipitation for all the scales considered. Overall, the result provided insights into the seasonal and regional patterns of rainfall hazards in Switzerland, highlighting the importance of considering multiple spatiotemporal scales when assessing extreme precipitation hazards. We comment here that although we used the EGPD-IDAF model for areal rainfall hazard assessment, it can be used in several applications, such as the design of hydraulic structures (Bertini et al., 2020), or the determination of thresholds for use in early-warning systems (Panziera et al., 2016). Another potential application is that since the EGPD models the whole distribution of non-zero precipitation, not only the upper tail, it can be used as a marginal distribution in stochastic weather generators for areal rainfall generation. The model will provide for a robust marginal distribution, given the quantity of data used to train it.

Additional results through a point-to-pixel comparison showed that both CPC and rain gauge data provided similar return level estimates, especially for longer durations. While this can be seen as a sort of validation of the CPC in extreme value analysis, the inferred return levels have to be interpreted with caution, mainly due to the limited length of the data. Notwithstanding, our work still provided a framework for further analysis in the presence of longer time series, *e.g.* from simulated series using weather generators. Another limitation concerning the use of EGPD is that the data has to be temporally declustered to reduce the serial dependence in the time series. For example, for durations higher than 10 hr (see Section 3.2) the temporal declustering means the data are taken in blocks. This can undoubtedly lead to the omission of high-intensity events that might result in the underestimation of the return levels. Potential methods to correct this have to be explored, since existing methods, to our knowledge, are for annual maxima series (e.g. Hershfield, 1962; Blanchet et al., 2016).

Some perspectives for the present work involve using splines to model the relationships in the EGPD-IDAF model rather than regression forms. A possibility is to use Generalized Additive Models (GAMs) as implemented in Youngman (2020), or its extension that uses censored likelihood as used in Haruna et al. (2022). While splines can be promising due to their flexibility, a likely drawback is the enormous computational time required for inference of the model when using the EGPD, which uses all non-zero data. Our experience in Haruna et al. (2022) shows that the model requires significant time before convergence. The problem will be more complicated in this case where 100 time series is used and for more than 7,000 pixels. Another avenue for further research involves developing an Areal-Reduction-Factor (ARF)-based IDAF model and comparing it with the data-driven approach used in this model. While empirical (e.g. Mineo et al., 2018) and analytical (e.g. De Michele et al., 2001) ARF formulations exist in the literature (see Svensson & Jones, 2010, for a review), our suggestion is to empirically develop an ARF model that works in the study area. This is because previous research by Mélèse et al. (2019) showed that in mountainous regions, ARF formulations can exhibit unusual behavior (e.g. increasing value of ARF with an increase in Area, or ARF > 1). Furthermore, from an inference point of view, it will be interesting to explicitly account for dependence in the likelihood of the EGPD model (Equation 8). Beyond addressing the potential of likelihood misspecification, it will allow the possibility to estimate the conditional probability of observing an extreme event of a particular spatiotemporal scale, given that an extreme of another scale has been observed. This kind of information is invaluable in practice for risk management and planning.

Finally, an avenue for further research is to make an objective comparison of the performance of the EGPD and other distributions, such as GEV or GPD, or the recently proposed meta-statistical extreme value (MEV) distribution (Marani & Ignaccolo, 2015) in modeling IDAF relationships. The MEV distribution, in particular, has become increasingly popular in hydrological applications (e.g. Schellander et al., 2019; Gründemann et al., 2023) because it does not require the asymptotic assumption, and it uses more data compared to GEV and GPD. The evaluation framework and criteria used in this thesis (see Section 3.7) can be applied to compare these distributions. Additionally, the criterion proposed by Gründemann et al. (2023) to measure the heaviness of the tail of the distribution can also be computed. This will allow a thorough evaluation of the advantages and potential drawbacks of using the EGPD when the interest is only in the extremes. In any case, the EGPD has an edge over the GPD/GEV/MEV distributions since it models the entire non-zero precipitation (low, medium, and extremes), while the latter distributions only model the extremes.

6 Open Research

The `CombiPrecip` (CPC) data used in this study are freely available from the Swiss Federal Office of Meteorology and Climatology, MeteoSwiss, and can be freely obtained from IDAWEB (2023). The software used in this research is freely available as an R package Haruna (2023) with GPL-3 license.

Acknowledgments

This research is part of the Ph.D. thesis of Abubakar Haruna. It has been supported by the Swiss Confederation: 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.

Appendix A Comparison of CPC to raingauge data

The Kling-Gupta Efficiency (KGE) (Kling et al., 2012) is computed from:

$$KGE = 1 - \sqrt{(r-1)^2 + (\beta - 1)^2 + (\gamma - 1)^2} \tag{A1}$$

Where $r = \frac{\mathbb{C}\mathrm{ov}(i_{CPC}, i_{Gauge})}{\sigma^2_{CPC}\sigma^2_{Gauge}}$ is the Pearson correlation coefficient between the CPC data (i_{CPC}) and the station data (i_{Gauge}), $\mathbb{C}$ov is the co-variance between the two time series and σ denotes the standard deviation. $\beta = \frac{\mu_{CPC}}{\mu_{Gauge}}$, evaluates the bias between the two time series, with μ being the mean. $\gamma = \frac{\mathrm{CV}_{CPC}}{\mathrm{CV}_{Gauge}}$ is the variability ratio, that is the ratio between the coefficient of variations of the two time series.

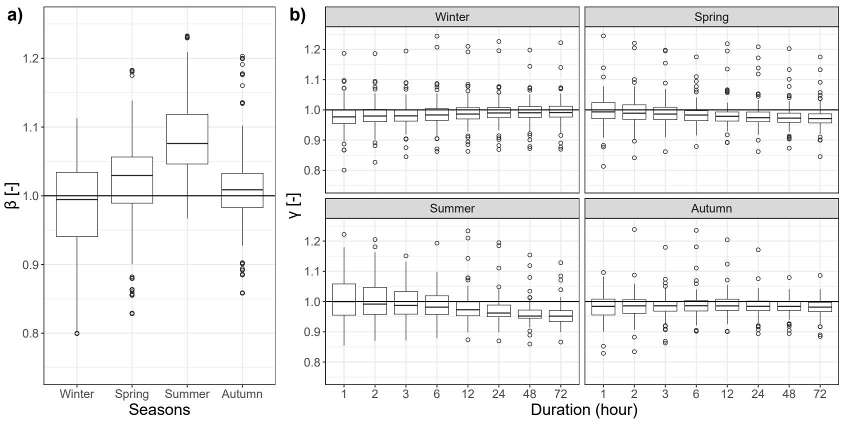

Figure A1. a): Boxplots of bias (β) for the four seasons. b): Boxplots of variability ratio (γ) for the four seasons. Each boxplot contains 79 points, 1 point for each pair of gauge and the underlying CPC pixel.

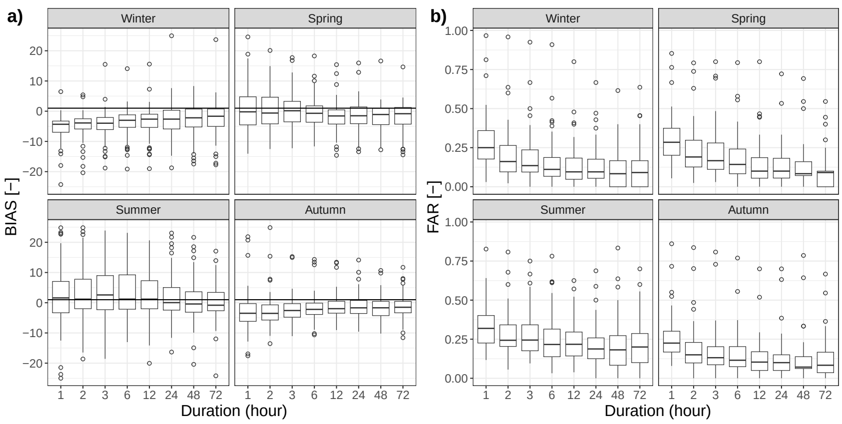

Figure A2. a): Boxplots of the bias in extreme precipitation totals for the four seasons. b): Boxplots of the false alarm ratio (FAR) for the four seasons. Each boxplot contains 79 points, 1 point for each pair of gauges and the underlying CPC pixel.

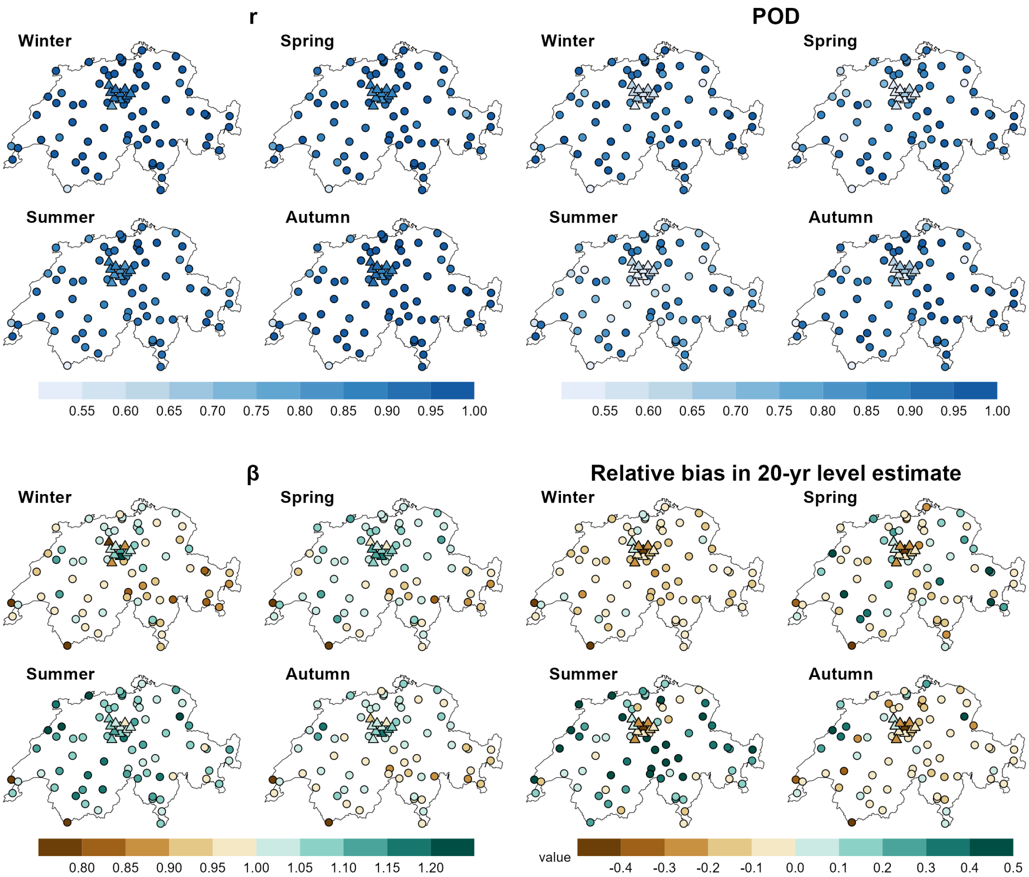

Figure A3. Map of Switzerland showing the location of the 79 stations used for comparing CPC against rain gauge time series. The locations shown by the circles are those managed by MeteoSwiss, while those shown by the triangles are managed by the canton of Lucerne. The shapes are colored according to the value of the criterion. From top left in clockwise direction: linear correlation (r), probability of detection (POD), bias (β) relative bias in a 20-yr return level estimate

Appendix B Cross validation criteria for various spatiotemporal scales

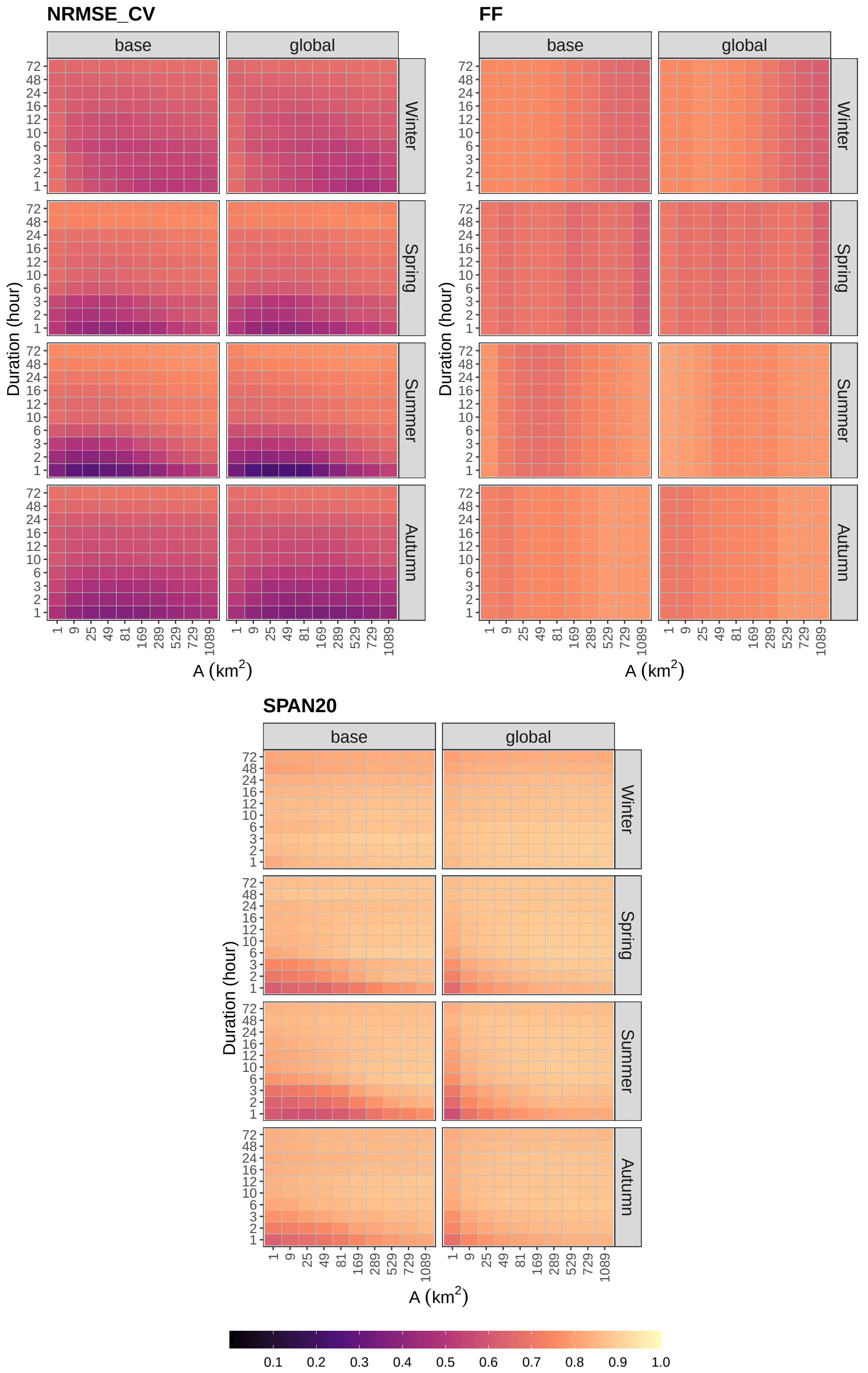

Figure B1. a): Seasonal heat maps of the median score over 80 values (2 × 40 resamplings) for various spatiotemporal scales (D, A). Top left: NRMSE_CV, top right: the FF criterion. The bottom panel shows the same maps for the case of SPAN20 over 40 resamplings.

References

Allen, R. J., & DeGaetano, A. T. (2005). Considerations for the use of radar-derived precipitation estimates in determining return intervals for extreme areal precipitation amounts. *Journal of Hydrology*, *315*(1-4), 203–219. doi: https://doi-org.insu.bib.cnrs.fr/10.1016/j.jhydrol.2005.03.028

Andres, N., Lieberherr, G., Sideris, I. V., Jordan, F., & Zappa, M. (2016). From calibration to real-time operations: an assessment of three precipitation benchmarks for a Swiss river system. *Meteorological Applications*, *23*(3), 448–461. Retrieved 2022-09-08, from `http://onlinelibrary.wiley.com/doi/abs/10.1002/met.1569` (_eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/met.1569) doi: 10.1002/met.1569

Bertini, C., Buonora, L., Ridolfi, E., Russo, F., & Napolitano, F. (2020, November). On the Use of Satellite Rainfall Data to Design a Dam in an Ungauged Site. *Water*, *12*(11), 3028. Retrieved 2021-11-22, from `https://www.mdpi.com/2073-4441/12/11/3028` (Number: 11 Publisher: Multidisciplinary Digital Publishing Institute) doi: 10.3390/w12113028

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., & Mélèse, V. (2020, January). A Bayesian Framework for the Multiscale Assessment of Storm Severity and Related Uncertainties. *Journal of Hydrometeorology*, *21*(1), 109–122. Retrieved 2021-04-15, from `https://journals.ametsoc.org/view/journals/hydr/21/1/jhm-d-18-0254.1.xml` (Publisher: American Meteorological Society Section: Journal of Hydrometeorology) doi: 10.1175/JHM-D-18-0254.1

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

Ceresetti, D., Anquetin, S., Molinié, G., Leblois, E., & Creutin, J.-D. (2012, February). Multiscale Evaluation of Extreme Rainfall Event Predictions Using Severity Diagrams. *Weather and Forecasting*, *27*(1), 174–188. Retrieved 2022-09-14, from `https://journals.ametsoc.org/view/journals/wefo/27/1/waf-d-11-00003_1.xml` (Publisher: American Meteorological Society Section: Weather and Forecasting) doi: 10.1175/WAF-D-11-00003.1

De Michele, C., Kottegoda, N. T., & Rosso, R. (2001, December). The derivation of areal reduction factor of storm rainfall from its scaling properties. *Water Resources Research*, *37*(12), 3247–3252. Retrieved 2021-11-30, from `http://doi.wiley.com/10.1029/2001WR000346` doi: 10.1029/2001WR000346

De Michele, C., Zenoni, E., Pecora, S., & Rosso, R. (2011, March). Analytical derivation of rain intensity–duration–area–frequency relationships from event maxima. *Journal of Hydrology*, *399*(3), 385–393. Retrieved 2022-09-29, from `https://www.sciencedirect.com/science/article/pii/S0022169411000436` doi: 10.1016/j.jhydrol.2011.01.018

Durrans, S. R., Julian, L. T., & Yekta, M. (2002, September). Estimation of Depth-Area Relationships using Radar-Rainfall Data. *Journal of Hydrologic Engineering*, *7*(5), 356–367. Retrieved 2022-10-04, from `https://ascelibrary.org/doi/10.1061/%28ASCE%291084-0699%282002%297%3A5%28356%29` (Publisher: American Society of Civil Engineers) doi: 10.1061/(ASCE)1084-0699(2002)7:5(356)

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

Froidevaux, P., Schwanbeck, J., Weingartner, R., Chevalier, C., & Martius, O. (2015). Flood triggering in switzerland: the role of daily to monthly preceding precipitation. *Hydrology and Earth System Sciences*, *19*(9), 3903–3924. doi: https://doi.org/10.5194/hess-19-3903-2015

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

Gabella, M., Speirs, P., Hamann, U., Germann, U., & Berne, A. (2017, November). Measurement of Precipitation in the Alps Using Dual-Polarization C-Band Ground-Based Radars, the GPM Spaceborne Ku-Band Radar, and Rain Gauges. *Remote Sensing*, *9*(11), 1147. Retrieved 2022-09-09, from https://www.mdpi.com/2072-4292/9/11/1147 (Number: 11 Publisher: Multidisciplinary Digital Publishing Institute) doi: 10.3390/rs9111147

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

Germann, U., Boscacci, M., Gabella, M., & Sartori, M. (2015). Peak performance: Radar design for precipitation in the Swiss Alps. In *Meteorological Technology International,* (pp. 42–45). United kingdom: UKIP Medai and Events. Retrieved 2022-09-12, from http://www.ukipme.com/pub-meteorological.php

Germann, U., Galli, G., Boscacci, M., & Bolliger, M. (2006). Radar precipitation measurement in a mountainous region. *Quarterly Journal of the Royal Meteorological Society*, *132*(618), 1669–1692. Retrieved 2022-09-12, from http://onlinelibrary.wiley.com/doi/abs/10.1256/qj.05.190 doi: 10.1256/qj.05.190

Giannakaki, P., & Martius, O. (2015). Synoptic-scale flow structures associated with extreme precipitation events in northern Switzerland. *Int. J. Climatol.*(40), 19. doi: https://doi.org/10.1002/joc.4508

Goudenhoofdt, E., Delobbe, L., & Willems, P. (2017, October). Regional frequency analysis of extreme rainfall in Belgium based on radar estimates. *Hydrology and Earth System Sciences*, *21*(10), 5385–5399. Retrieved 2022-11-10, from https://hess.copernicus.org/articles/21/5385/2017/hess-21-5385-2017-discussion.html (Publisher: Copernicus GmbH) doi: 10.5194/hess-21-5385-2017

Gründemann, G. J., Zorzetto, E., Beck, H. E., Schleiss, M., van de Giesen, N., Marani, M., & van der Ent, R. J. (2023, June). Extreme precipitation return levels for multiple durations on a global scale. *Journal of Hydrology*, *621*, 129558. Retrieved 2023-05-05, from https://www.sciencedirect.com/science/article/pii/S0022169423005000 doi: 10.1016/j.jhydrol.2023.129558

Gugerli, R., Gabella, M., Huss, M., & Salzmann, N. (2020, December). Can Weather Radars Be Used to Estimate Snow Accumulation on Alpine Glaciers? An Evaluation Based on Glaciological Surveys. *Journal of Hydrometeorology*, *21*(12), 2943–2962. Retrieved 2022-09-09, from https://journals.ametsoc.org/view/journals/hydr/21/12/jhm-d-20-0112.1.xml (Publisher: American Meteorological Society Section: Journal of Hydrometeorology) doi: 10.1175/JHM-D-20-0112.1

Haruna, A. (2023, apr). *egpdidf: Modeling intensity-duration-frequency (idf) curves* *using the extended generalized pareto distribution (egpd) [software].* Zenodo. Retrieved from https://doi.org/10.5281/zenodo.7828750 doi: 10.5281/zenodo.7828750

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

Haruna, A., Blanchet, J., & Favre, A.-C. (2023). Modeling Intensity-Duration-Frequency Curves for the Whole Range of Non-Zero Precipitation: A Comparison of Models. *Water Resources Research*, *59*(6), e2022WR033362. Retrieved 2023-05-26, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2022WR033362 doi: 10.1029/2022WR033362

Hershfield, D. M. (1962). Extreme rainfall relationships. *Journal of the Hydraulics Division*, *88*(6), 73–92.

IDAWEB. (2023). *The data portal of meteoswiss for research and teaching [dataset].* Retrieved 2023-07-21, from https://gate.meteoswiss.ch/idaweb/login.do

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

Kling, H., Fuchs, M., & Paulin, M. (2012). Runoff conditions in the upper danube basin under an ensemble of climate change scenarios. *Journal of hydrology*, *424*, 264–277. doi: https://doi.org/10.1016/j.jhydrol.2012.01.011

Kunsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The annals of Statistics*, 1217–1241. doi: https://www.jstor.org/stable/2241719

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

Marani, M., & Ignaccolo, M. (2015, May). A metastatistical approach to rainfall extremes. *Advances in Water Resources*, *79*, 121–126. Retrieved 2023-08-30, from https://www.sciencedirect.com/science/article/pii/S0309170815000494 doi: 10.1016/j.advwatres.2015.03.001

Mélèse, V., Blanchet, J., & Creutin, J.-D. (2019). A regional scale-invariant extreme value model of rainfall intensity-duration-area-frequency relationships. *Water Resources Research*, *55*(7), 5539–5558. doi: https://doi.org/10.1029/2018WR024368

MeteoSwiss. (2017, November). *Documentation of meteoswiss grid-data products hourly precipitation estimation through rain-gauge and radar: Combiprecip.* Retrieved 2022-08-31, from https://www.meteoswiss.admin.ch/dam/jcr:2691db4e-7253-41c6-a413-2c75c9de11e3/ProdDoc_CPC.pdf

MeteoSwiss. (2023, Dec). *Measurement values and measuring networks.* Retrieved 2023-12-24, from https://www.meteoswiss.admin.ch/services-and-publications/applications/measurement-values-and-measuring-networks.html#lang=en&swisstopoApiKey=cpZJOL3HuO5yENksi97q&param=messwerte-lufttemperatur-10min

Mineo, C., Ridolfi, E., Napolitano, F., & Russo, F. (2018, May). The areal reduction factor: A new analytical expression for the Lazio Region in central Italy. *Journal of Hydrology*, *560*, 471–479. Retrieved 2022-09-30, from https://www.sciencedirect.com/science/article/pii/S0022169418301999 doi: 10.1016/j.jhydrol.2018.03.033

Molnar, P., & Burlando, P. (2008). Variability in the scale properties of high-resolution precipitation data in the Alpine climate of Switzerland. *Water Re-*

sources Research, *44*(10). doi: http://doi.wiley.com/10.1029/2007WR006142

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. (2007). A simple scaling charateristics of rainfall in time and space to derive intensity duration frequency relationships. *Proceedings of Hydraulic Engineering*, *51*, 73–78. Retrieved 2022-05-09, from `http://www.jstage.jst.go.jp/article/prohe1990/51/0/51_0_73/_article` doi: 10.2208/prohe.51.73

Overeem, A., Buishand, T. A., & Holleman, I. (2009). Extreme rainfall analysis and estimation of depth-duration-frequency curves using weather radar. *Water Resources Research*, *45*(10). Retrieved 2021-11-30, from `https://onlinelibrary.wiley.com/doi/abs/10.1029/2009WR007869` (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2009WR007869) doi: 10.1029/2009WR007869

Overeem, A., Buishand, T. A., Holleman, I., & Uijlenhoet, R. (2010, September). Extreme value modeling of areal rainfall from weather radar. *Water Resources Research*, *46*(9). Retrieved 2021-11-30, from `http://doi.wiley.com/10.1029/2009WR008517` doi: 10.1029/2009WR008517

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

Panziera, L., Gabella, M., Germann, U., & Martius, O. (2018). A 12-year radar-based climatology of daily and sub-daily extreme precipitation over the swiss alps. *International Journal of Climatology*, *38*(10), 3749–3769. doi: https://doi.org/10.1002/joc.5528

Panziera, L., Gabella, M., Zanini, S., Hering, A., Germann, U., & Berne, A. (2016, June). A radar-based regional extreme rainfall analysis to derive the thresholds for a novel automatic alert system in Switzerland. *Hydrology and Earth System Sciences*, *20*(6), 2317–2332. Retrieved 2022-09-13, from `https://hess.copernicus.org/articles/20/2317/2016/` (Publisher: Copernicus GmbH) doi: 10.5194/hess-20-2317-2016

Ramos, M. H., Creutin, J.-D., & Leblois, E. (2005, December). Visualization of storm severity. *Journal of Hydrology*, *315*(1), 295–307. Retrieved 2022-09-14, from `https://www.sciencedirect.com/science/article/pii/S0022169405001939` doi: 10.1016/j.jhydrol.2005.04.007

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

Rodriguez-Iturbe, I., & Mejía, J. M. (1974). On the transformation of point rainfall to areal rainfall. *Water Resources Research*, *10*(4), 729–735. doi: https://doi-org.insu.bib.cnrs.fr/10.1029/WR010i004p00729

Schellander, H., Lieb, A., & Hell, T. (2019). Error structure of metastatistical and generalized extreme value distributions for modeling extreme rainfall in Austria. *Earth and Space Science*, *6*(9), 1616–1632.

Scherrer, S. C., Begert, M., Croci-Maspoli, M., & Appenzeller, C. (2016, September). Long series of Swiss seasonal precipitation: regionalization, trends and influence of large-scale flow. *International Journal of Climatology*, *36*(11), 3673–3689. Retrieved 2021-04-19, from `https://onlinelibrary.wiley.com/doi/10.1002/joc.4584` doi: 10.1002/joc.4584

Sebille, Q., Fougères, A.-L., & Mercadier, C. (2017, August). Modeling extreme rainfall A comparative study of spatial extreme value models. *Spatial Statistics*, *21*, 187–208. Retrieved 2021-04-19, from https://linkinghub.elsevier.com/retrieve/pii/S2211675316300975 doi: 10.1016/j.spasta.2017.06.009

Shehu, B., & Haberlandt, U. (2023). Uncertainty estimation of regionalised depth–duration–frequency curves in germany. *Hydrology and Earth System Sciences*, *27*(10), 2075–2097. doi: https://doi.org/10.5194/hess-27-2075-2023

Sideris, I. V., Gabella, M., Erdin, R., & Germann, U. (2014a). The combiprecip experience: development and operation of a real-time radar-raingauge combination scheme in switzerland. In *2014 international weather radar and hydrology symposium* (pp. 1–10).

Sideris, I. V., Gabella, M., Erdin, R., & Germann, U. (2014b). Real-time radar–rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland. *Quarterly Journal of the Royal Meteorological Society*, *140*(680), 1097–1111. Retrieved 2022-09-08, from http://onlinelibrary.wiley.com/doi/abs/10.1002/qj.2188 (_eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2188) doi: 10.1002/qj.2188

Sikorska, A. E., Viviroli, D., & Seibert, J. (2015). Flood-type classification in mountainous catchments using crisp and fuzzy decision trees. *Water Resources Research*, *51*(10), 7959–7976.

Sodemann, H., & Zubler, E. (2009). Seasonal and inter-annual variability of the moisture sources for Alpine precipitation during 1995-2002. *International Journal of Climatology*, 947–961. Retrieved 2021-04-19, from http://doi.wiley.com/10.1002/joc.1932 doi: 10.1002/joc.1932

Svensson, C., & Jones, D. (2010). Review of methods for deriving areal reduction factors. *Journal of Flood Risk Management*, *3*(3), 232–245. Retrieved 2022-09-30, from http://onlinelibrary.wiley.com/doi/abs/10.1111/j.1753-318X.2010.01075.x (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1753-318X.2010.01075.x) doi: 10.1111/j.1753-318X.2010.01075.x

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

Wright, D. B., Smith, J. A., & Baeck, M. L. (2014). Flood frequency analysis using radar rainfall fields and stochastic storm transposition. *Water Resources Research*, *50*(2), 1592–1615. Retrieved 2023-01-23, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2013WR014224 (_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2013WR014224) doi: 10.1002/2013WR014224

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)

Zambrano-Bigiarini, M., Nauditt, A., Birkel, C., Verbist, K., & Ribbe, L. (2017, March). Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile. *Hydrology and Earth System Sciences*, *21*(2), 1295–1320. Retrieved 2022-11-22, from https://hess.copernicus.org/articles/21/1295/2017/ (Publisher: Copernicus GmbH) doi: 10.5194/hess-21-1295-2017

Zhao, W., Abhishek, Takhellambam, B. S., Zhang, J., Zhao, Y., & Kinouchi, T. (2023). Spatiotemporal variability of current and future sub-daily rainfall in japan using state-of-the-art high-quality datasets. *Water Resources Research*, e2022WR034305. doi: https://doi-org.insu.bib.cnrs.fr/10.1029/2022WR034305