Hydrology and Earth System Sciences

Mapping snow depth return levels: smooth spatial modeling versus station interpolation

J. Blanchet and M. Lehning

WSL Institute for Snow and Avalanche Research SLF, Davos, Switzerland

Abstract. For adequate risk management in mountainous countries, hazard maps for extreme snow events are needed. This requires the computation of spatial estimates of return levels. In this article we use recent developments in extreme value theory and compare two main approaches for mapping snow depth return levels from in situ measurements. The first one is based on the spatial interpolation of pointwise extremal distributions (the so-called Generalized Extreme Value distribution, GEV henceforth) computed at station locations. The second one is new and based on the direct estimation of a spatially smooth GEV distribution with the joint use of all stations. We compare and validate the different approaches for modeling annual maximum snow depth measured at 100 sites in Switzerland during winters 1965–1966 to 2007–2008. The results show a better performance of the smooth GEV distribution fitting, in particular where the station network is sparser. Smooth return level maps can be computed from the fitted model without any further interpolation. Their regional variability can be revealed by removing the altitudinal dependent covariates in the model. We show how return levels and their regional variability are linked to the main climatological patterns of Switzerland.

1 Introduction

Heavy snow events are among the most severe natural hazards in mountainous countries. In the European Alps, one of the most exceptional avalanche winters occurred in winter 1998–1999, mainly due to continued and heavy snowfall events in February 1999. A multitude of large avalanches released in the northern sectors of the alpine regions in Austria, Italy, France, and Switzerland. In Switzerland 36 people died

in avalanches over the whole winter season, 12 of them alone in Evolène on 21 February 1999 (SLF, 2000). Avalanches in February 1999 damaged around 230 houses and many other buildings, vehicles, etc. (SLF, 2000). An even more disastrous avalanche winter occurred in 1950–1951 and resulted in 98 fatalities in Switzerland (Bründl et al., 2004).

Risk management in mountainous regions such as the Swiss Alps requires to establish return level maps depicting dangerous areas and to take them into account for land-use planning (Lateltin and Bonnard, 1999). The well-founded framework for the computation of return levels is extreme value theory (Coles, 2001). It has been widely used among others in hydrology (Katz et al., 2002; Vasiliev et al., 2007; Reiss and Thomas, 2007) and climatology (Naveau et al., 2005; Brown and Katz, 1995; Palutikof et al., 1999). However, its use for snow events is limited with a few exceptions (Bocchiola et al., 2006, 2008; Blanchet et al., 2009). We revealed in a previous article (Blanchet et al., 2009) how extreme snowfall is spatially distributed over Switzerland and argued that this spread is determined by the main climatological patterns. Nevertheless, the methodology developed in Blanchet et al. (2009) is based on univariate extreme value theory and does not allow the calculation of *spatial* return levels. This article can be seen as a next step towards a spatial modeling of extreme snow events, allowing spatial return levels to be computed.

Quite a wide range of literature exists on the issue of spatial mapping (or spatial interpolation) of snow depth. A broad range of spatial interpolation techniques are compared in Erxleben et al. (2002) and Molotch et al. (2005), including geostatistics, binary regression trees and combined methods of both techniques. Other statistical methods include generalized additive models (e.g., López-Moreno and Nogués-Bravo, 2005) allowing a non-linear dependence between snow depth and topographical variables. Several studies use a different approach by taking advantage of remotely sensed data for estimating snow depth (georadar in Marchand et al.,

2003, for example). More recently, interpolation has been constrained by remotely sensed estimates of snow-covered area. For the special case of snow depth mapping in Switzerland, Foppa et al. (2007) described a first practical method and Harshburger et al. (2010) proposed to mix a multiple regression method with satellite-based information. It is important to stress at this point that all of the aforementioned articles deal with interpolation of *daily* or *monthly* snow depth and not of *annual maximum* snow depth as we do in this article. The difference between the two problems is two-fold. First here only the largest events (namely the annual maxima) are studied and the framework for that is Extreme Value Theory. Second, the process of annual maxima does in general not correspond to a one-day event since annual maxima in Switzerland usually do not occur simultaneously. They may occur simultaneously on some neighboring stations but not across larger areas. The problem of how to model the spatial characteristics of snow depth applies to mean as well as extreme values: just as daily snow depth, annual maximum snow depth is likely to vary smoothly over space. Investigating its spatial mapping on the basis of Extreme Value theory is the subject of this article.

Studies on the spatial mapping of extreme events in general can be divided into two main groups. The first one is based on the spatial interpolation of in-situ estimates in order to enable the construction of return level maps. Kohnová et al. (2009) and Beguería and Vicente-Serrano (2006) for example interpolated the in-situ extremal distributions, whereas Loukas et al. (2001) and Weisse and Bois (2001, 2002) interpolated directly the in-situ (100-year) return levels. A comparison of different interpolation methods for mapping extreme precipitation can be found in Szolgay et al. (2009). The second group of studies is based on the direct estimation of the spatial extremal distribution, without requiring any interpolation. This is a well-founded approach that should theoretically be preferred to any interpolation method. Cooley et al. (2007) proposed a Bayesian modeling whereas Padoan et al. (2010) made use of max-stable modeling for spatial extremes. Return level maps are also obtained in Gardes and Girard (2010) based on nearest neighbor estimators. However, all the aforementioned statistical papers aim at proposing new methods for the spatial modeling of extreme events, but none of them make a comparison with more naive interpolation routines that are practical for operational applications. In this article, we propose to compare these two main approaches for the first time – the interpolation-based approach and the spatial statistics-based approach – for mapping extreme snow depth in Switzerland. More precisely, throughout this paper we make use of the Generalized Extreme Value (GEV) distribution and compare the GEV parameter interpolation approach with a smooth GEV modeling approach, i.e. a GEV modeling in which the parameters are modeled as smooth functions in space through the use of spatial covariates. Compared to Padoan et al. (2010) where smooth GEV distributions are also used, we will use in

this article more sophisticated response surfaces for modeling the GEV parameters but within a less complicated statistical framework in which the property of max-stability will not be accounted for. Note that this is also the underlying assumption made in all papers on spatial interpolation of extremes. We argue that this simplification does not affect the computation of return level maps, which is the goal of this paper. Smooth GEV modeling is already a practical improvement over simple interpolation as is commonly done in application oriented work.

The article is organized as follows. We first recall the principle of extreme value theory and show in particular how return levels can be computed from it. We then present the data under study in Sect. 3 and perform an analysis of extreme snow depth at station locations in Sect. 4. This allows us to derive pointwise estimates of return levels in Switzerland. In Sect. 5 we present a first method for obtaining spatial estimates of return levels, based on the interpolation of the individual GEV distributions obtained in Sect. 4. We give some results for the Swiss snow depth data and show the limitations of this methodology. In Sect. 6 we develop a better approach based on the estimation of a smooth GEV distribution using all stations jointly. Finally, a comparison of the different methods based on a validation data set is used to draw conclusions as to which method should be preferred in practical applications.

2　Extreme value theory

We only provide a short overview of the statistical theory of extreme values. We focus here on "block maxima approach" in the univariate case. For more details, we refer to Coles (2001), chapter 3, for example. Extreme value theory focuses then on the asymptotic behavior of the so-called *block maxima*

$$Z_L = \max\{Y_1, \ldots, Y_L\}, \tag{1}$$

where $Y_1, \ldots, Y_L$ is a sequence random variables. Here "asymptotic" means that the theory predicts the behavior as the block length L goes to infinity, but one usually requires L to be "large enough". In practice, $Y_1, \ldots, Y_L$ is usually a time series, of daily observations for example. Z_L is then the maximum of the measured process over a block of L observations (a year for example, $L = 365.25$). Ideally, extreme value theory requires the (e.g. daily) observations $Y_l, l = 1, \ldots, L$, to be independent. However this hypothesis can be relaxed to the case of shortly dependent Y_l. More specifically, extreme value theory still applies if the "$D(u_n)$ condition" of Leadbetter et al. (1983) is satisfied, i.e. if distant maxima of the (e.g. daily) process $Y_1, \ldots, Y_L$ are near-independent. For independent or shortly dependent Y_l and if the length L of the sample is large enough, the probability that the sample maximum Z_L does not exceed a certain level z, i.e. $P(Z_L \leq z)$, is
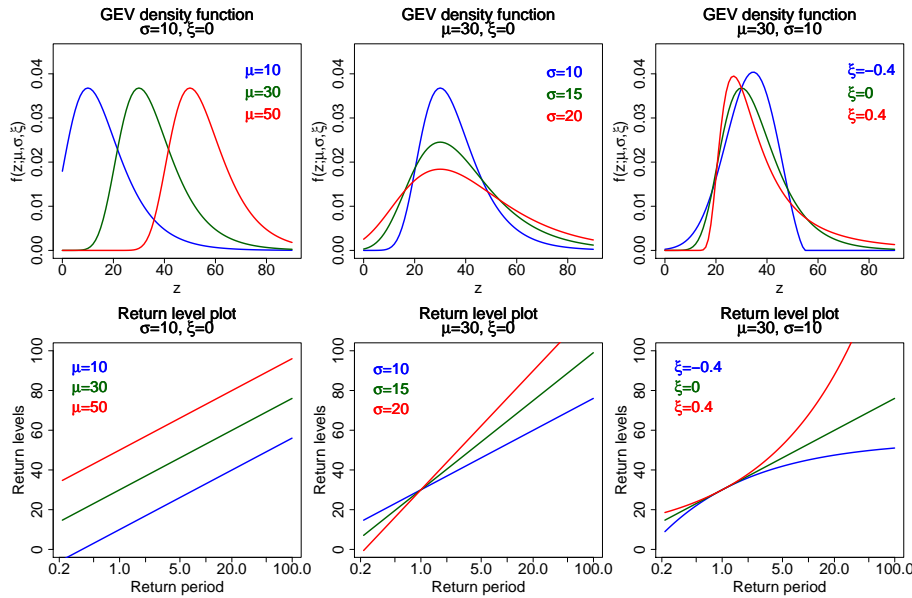
Fig. 1. Illustration of the GEV distribution. Upper panel: example of GEV densities $f(x; \mu, \sigma, \xi)$ when varying the location μ parameter (left plot), the scale σ parameter (middle plot) or the shape ξ parameter (right plot). Lower panel: corresponding return level plots.

approximately given by the GEV distribution, with cumulative distribution function:

$$G(z; \mu, \sigma, \xi) = \exp\left\{-\left[1 + \xi\left(\frac{z - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right\} \quad (2)$$

if $1 + \xi\left(\frac{z - \mu}{\sigma}\right) > 0$, 0 otherwise.

The GEV distribution has three parameters (Eq. 2): a *location* parameter μ, a *scale* parameter $\sigma > 0$ and a *shape* parameter ξ. The location specifies where the distribution is centered and the scale its spread. The shape parameter ξ describes the tail behavior of the distribution, leading to three types of GEV distributions:

– when $\xi > 0$, a heavy-tailed (or Fréchet) distribution,

– when $\xi = 0$, a light-tailed (or Gumbel) distribution,

– when $\xi < 0$, a bounded (or Weibull) distribution.

The Gumbel distribution with $\xi = 0$ is interpreted in Eq. (2) as the limit when $\xi \to 0$, leading to the distribution function

$$G(z; \mu, \sigma, \xi = 0) = \exp\left\{-\exp\left[-\left(\frac{z - \mu}{\sigma}\right)\right]\right\}.$$

In case of a bounded distribution (Weibull, $\xi < 0$), the variable of interest Z_n has a finite upper point, meaning that theoretically no value above this upper bound can be observed. A light tailed distribution (Gumbel, $\xi = 0$) has an infinite upper point and any value could theoretically be observed. Nevertheless, very extreme values (i.e. far from average observations) are very rare. In a heavy-tailed distribution (Fréchet,

$\xi > 0$), such extremes are still rare but more probable. An illustration of the influence of the three GEV parameters is depicted in Fig. 1, upper panel, for arbitrary snow-like GEV parameters.

In practice return levels are commonly used for operational purposes. The *return level* q_p associated with the *return period* $\frac{1}{p}$ ($0 < p \leq 1$) is the $(1 - p)$-th quantile of the GEV distribution; it is expected to be exceeded on average once every $\frac{1}{p}$ years. Estimates of return levels are obtained by setting in Eq. (2) $G(q_p; \mu, \sigma, \xi) = 1 - p$ and by inverting it:

$$q_p = \begin{cases} \mu - \frac{\sigma}{\xi}[1 - \{-\log(1 - p)\}^{-\xi}], & \text{for } \xi \neq 0, \\ \mu - \sigma \log\{-\log(1 - p)\}, & \text{for } \xi = 0. \end{cases} \quad (3)$$

The graph of q_p against $-\log(1 - p)$ on a logarithm scale (i.e. the plot of q_p against $\log\{-\log(1 - p)\}$) is a *return level plot*. It is particularly convenient for interpreting extreme value models. It gives, for any return period r on the x-axis, the associated return level, i.e. roughly speaking the highest value expected to be exceeded once every r years (for yearly maxima data). From Eq. (3), if $\xi < 0$ the plot is convex with asymptotic limit as $p \to 0$ (infinite return period, i.e. $r \to \infty$) at $\mu - \frac{\sigma}{\xi}$; if $\xi > 0$ the plot is concave and has no finite bound; if $\xi = 0$ it is linear. An illustration is given in Fig. 1, lower panel. It is usually long return periods, corresponding to small values of p, that are of greatest interest. Cases when ξ is positive are of particular concern for risk management because very extreme events may occur.
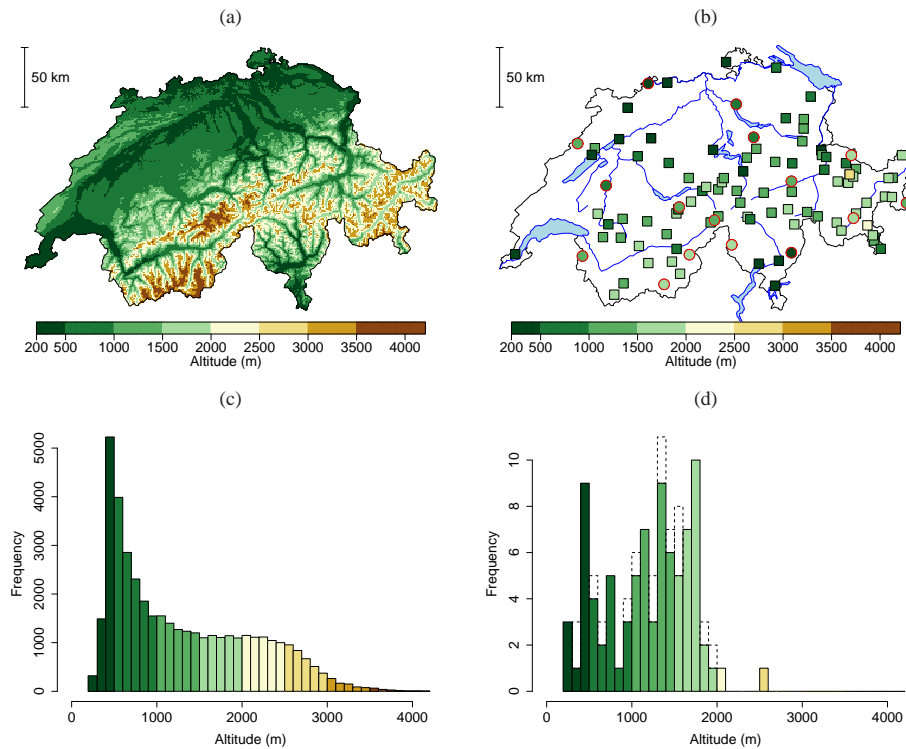
Fig. 2. Upper row: **(a)** elevation map of Switzerland and **(b)** station locations. Lower row: **(c)** histogram of elevations in Switzerland in a 1 km grid spacing and **(d)** of the stations. Color indicates elevation in meters above sea level. Among the 100 stations, 16 are excluded from the analysis for validation (red circles in the right map corresponding to dashed part of the right histogram).

3 Data

We shall consider annual maximum snow depth from the 100 sites in Switzerland shown in Fig. 2. The stations we consider belong to two manual networks run by SLF (WSL Institute for Snow and Avalanche Research) and MeteoSwiss (Swiss Federal Office for Meteorology and Climatology). Annual maxima are extracted from daily snow depth measured manually on a stake at around 07:30 a.m. during the winter season, i.e. between 1 November and 30 April, for the winters 1965–1966 to 2007–2008. The study area covers all of Switzerland with a higher density in the alpine part; see maps of Fig. 2. The area is characterized by a high density of population, tourism infrastructure and traffic during winter. The elevations of the stations range between 250 m and 2500 m a.s.l. (above sea level), with only two stations above 2000 m. 16 of these 100 stations are excluded from the analysis for validation, and thus 84 are used for inference. These 16 stations have been chosen to cover most of Switzerland and are located at various elevations between 300 and 2000 m. Some of these chosen stations are purposely climatologically unique. For example, the easternmost of the validation stations is the only one in a valley system with a special local climatology; or the westernmost of the validation stations is located at least 500 m higher than all the

surrounding stations in a larger region. The choice of these "unique" stations was made in order to assess performance of the spatial model in the most difficult case. This has to be taken into account when interpreting results for the validation stations in Sects. 5 and 6.

Let $Z(s)$ denote the annual maximum snow depth at site s of Switzerland, i.e.

$$Z(s) = \max \{Y_1(s), \ldots, Y_L(s)\}, \tag{4}$$

where $Y_l(s) \geq 0$ denotes the snow depth at site s the l-th day of the winter ($l \in \{1, \ldots L\}$) and $L = 181$ or 182 denote the number of days in the six winter months from November to April (note that for shortness, index L in $Z(s)$ is omitted). The process $Z = \{Z(s), s \in \mathcal{S}\}$ where $\mathcal{S}$ denotes Switzerland can be assumed to be a continuous process, i.e. a smoothly varying process over space. In the context of extreme value theory, this means that $Z(s)$ can be approximated through a GEV distribution where parameters $\mu(s)$, $\sigma(s)$ and $\xi(s)$ describing respectively the location, scale and shape parameters over Switzerland should be modeled as smoothly varying functions. Doing so, it is straightforward to see from Eq. (3) that return levels will also be smoothly varying over space. The rest of this paper will be devoted to the specification of smooth functions for $\mu(s)$, $\sigma(s)$ and $\xi(s)$. Applying Eq. (3)
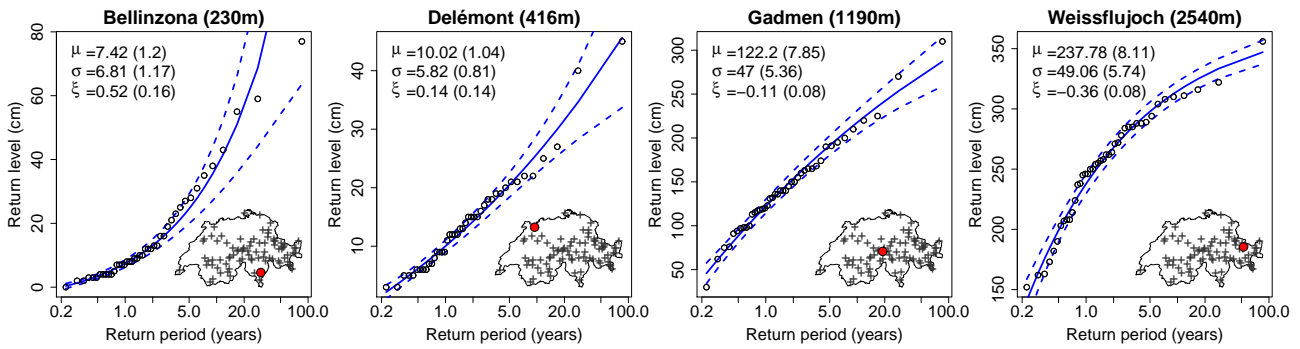
Fig. 3. Snow depth return level plots for four different stations. The blue curve is the GEV-based curve with standard errors (dashed line). Points are empirical estimates. Locations of the stations are indicated by the red circle in the lower-right Swiss map. Maximum likelihood estimates of the GEV parameters are indicated the upper-left corner (with standard errors).

will allow us to compute return levels for every s in Switzerland, and therefore to build smooth return level maps.

4 Pointwise estimation of the GEV distributions

Let s_1, ..., s_N denote the $N = 84$ station locations at hand (see Fig. 2). We start by studying the pointwise distributions of $Z(s_i)$, $i \in \{1, ..., N\}$. Spatial estimates of these distributions will be derived in Sects. 5 and 6. As previously stated, $Z(s_i)$ is a block maxima random variable given by Eq. (4) where location s is replaced by location s_i. The L daily snow depths $Y_1(s_i)$, ..., $Y_L(s_i)$ are dependent random variables due to the strong temporal dependence of snow depth. Nevertheless, a separate analysis (not shown) reveals that, for every winter and every location s_i, $i \in \{1, ..., N\}$, the time-series of snow depths exhibit a short-range dependence, which suggests that the $D(u_n)$ condition mentioned in Sect. 2 is satisfied. Furthermore, the block size L on which maxima $Z(s_i)$ are retrieved is 181 (or 182) corresponding to the six winter months. This size seems to be large enough to assume that the statistical theory of extreme values presented in Sect. 2 applies. Annual maximum snow depth at a given location s_i is then expected to follow a GEV distribution (Eq. 2) with parameters (μ_i, σ_i, ξ_i) to be estimated.

We adopt a maximum likelihood approach. Let $z_i^{(1)}$, ..., $z_i^{(K)}$ denote the $K = 43$ annual snow depth maxima measured at location s_i, i.e. realizations of the random variable $Z(s_i)$. The log-likelihood for the GEV parameters at station i is given by:

$$l(\mu_i, \sigma_i, \xi_i) = -K \log \sigma_i - \left(1 + \frac{1}{\xi_i}\right) \sum_{k=1}^{K} \log \qquad (5)$$

$$\left[1 + \xi_i \left(\frac{z_i^{(k)} - \mu_i}{\sigma_i}\right)\right] - \sum_{k=1}^{K} \left[1 + \xi_i \left(\frac{z_i^{(k)} - \mu_i}{\sigma_i}\right)\right]^{-\frac{1}{\xi_i}}$$

Maximization of Eq. (5) with respect to the parameter vector (μ_i, σ_i, ξ_i) leads to the maximum likelihood estimate,

denoted by $(\hat{\mu}_i, \hat{\sigma}_i, \hat{\xi}_i)$. There is no analytical solution but the maximization is straightforward using standard numerical optimization algorithms. R functions fgev in package evd or fit.gev in package ismev for example perform this maximum likelihood estimation. When $\xi_i > -0.5$, the maximum likelihood estimate has the usual asymptotic properties (Smith, 1985): $\hat{\mu}_i$, $\hat{\sigma}_i$ and $\hat{\xi}_i$ are asymptotically unbiased and standard errors are approximately given by the square root of the diagonal of the inverse observed information matrix (Coles, 2001, chapter 3).

Return levels for station i can then be computed by Eq. (3) where (μ, σ, ξ) are replaced by the maximum likelihood estimate $(\hat{\mu}_i, \hat{\sigma}_i, \hat{\xi}_i)$. Standard errors and confident intervals can also be obtained with the delta method (Coles, 2001, chapter 3). For illustration, return levels plots for four stations at low, middle and high altitude are depicted in Fig. 3, together with empirical estimates of return levels. These plots can also be used for model validation. If the GEV model is suitable for the data, the model-based curve and empirical estimates should be in reasonable agreement. Figure 3 suggests a reasonably good adequacy of the GEV model, even at low altitudes where daily snow depth time-series show a longer time dependence.

An interesting result is that the shape parameters ξ_i are usually positive at low altitudes, close to 0 at middle altitudes (about 1000 m) and negative at higher altitudes, as illustrated in Fig. 3. The positivity of ξ_i means that at low altitudes, the distribution of annual maximum snow depth is heavy tailed, i.e. that very large snow depth compared to "usual" snow depth can occur. This heaviness is even more pronounced in the low altitude region of Ticino in southern Switzerland (see Bellinzona compared to Delémont in Fig. 3) where very intense snowfalls sometimes occur due to the vicinity of the Mediterranean sea and the steep topography. By contrast, at high altitudes the GEV distribution is bounded (see Weissflujoch in Fig. 3), meaning that no real extremes occur with regards to other events. Similar results were found in Blanchet et al. (2009) regarding extreme snowfall in Switzerland.

www.hydrol-earth-syst-sci.net/14/2527/2010/

Hydrol. Earth Syst. Sci., 14, 2527–2544, 2010

Figure 4 depicts the 50-year return levels obtained from the fitted GEV distributions at the $N = 84$ station locations. Such a map is nevertheless difficult to interpret and can only give information for the few locations where data are available. In practice spatial return levels rather than pointwise estimates would be of much higher value. The rest of this paper will be devoted to this issue.

5 Interpolating the GEV parameters

With the aim of producing return level maps, we would like to have estimates of the smooth surfaces $\mu(s)$, $\sigma(s)$, $\xi(s)$ for every location s in Switzerland. The pointwise analysis of Sect. 4 allowed us to have estimates at isolated locations. The most naive way of deriving spatial estimates for μ, σ and ξ is to spatially interpolate the point estimates $\hat{\mu}_i$, $\hat{\sigma}_i$, $\hat{\xi}_i$, $i \in \{1, \ldots, N\}$. This is the subject of this section where different interpolation techniques are compared. This is also the approach adopted in Kohnová et al. (2009) and Beguería and Vicente-Serrano (2006) both regarding precipitation. Here the functions to be interpolated are then the three GEV parameters μ, σ, ξ. They are assumed to be known at all station locations s_i, $i \in \{1, \ldots, N\}$ with values $\hat{\mu}_i$, $\hat{\sigma}_i$, $\hat{\xi}_i$ of Sect. 4. The interpolation problem consists of specifying values at arbitrary locations $s \in \mathcal{S}$. In the following η will denote one of the three functions μ, σ and ξ to be interpolated and $\hat{\eta}_i$ its known value at location s_i. The goal is then to get smooth estimates $\tilde{\eta}(s)$ for all locations s in Switzerland, based on values $\hat{\eta}_i$, $i \in \{1, \ldots, N\}$.

5.1 Interpolation methods

We give in this section a general overview of the different interpolation techniques that will be used for interpolating the GEV parameters.

5.1.1 Inverse distance weighted

In inverse distance weighting (IDW), interpolated values are a function of the distance to surrounding locations. The inverse distance weight is used to attenuate the influence of distant points. The interpolated value η at location s is given by:

$$\tilde{\eta}(s) = \frac{\sum_{i=1}^{N} \frac{\hat{\eta}_i}{||s - s_i||}}{\sum_{i=1}^{N} \frac{1}{||s - s_i||}}, \tag{6}$$

where $||s - s_i||$ is the distance from the interpolating location s_i to the interpolated location s. Often the squared distance is used. IDW is an exact interpolation: at station location s_i, interpolated value $\tilde{\eta}(s_i)$ given by Eq. (6) is equal to the known value $\hat{\eta}_i$ used in the interpolation.
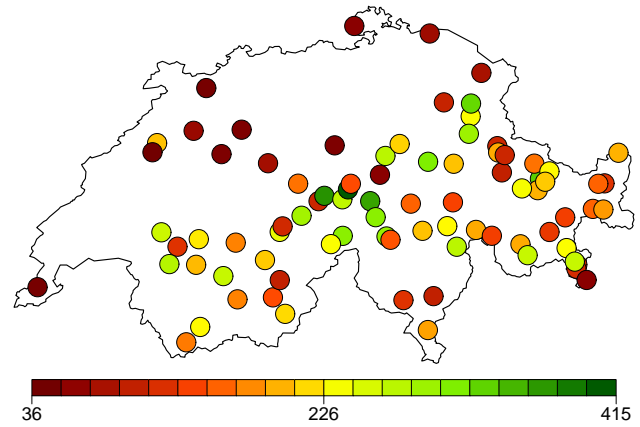
50–year return level

Fig. 4. Pointwise 50-year snow depth return level map (in centimeters).

5.1.2 Linear regression models

Let $x_{1,s}, \ldots, x_{p,s}$ denote p exploratory variables known for the whole of Switzerland (i.e. for each $s \in \mathcal{S}$). For example the $p = 3$ exploratory variables of longitude, latitude and elevation could be used, or polynomials of these variables. To predict η at site s given $x_{1,s}, \ldots, x_{p,s}$ one might consider the following model:

$$\eta(s) = \beta_0 + \beta_1 x_{1,s} + \ldots + \beta_p x_{p,s} + \epsilon_s, \tag{7}$$

where β_0, β_1, $\ldots$, β_p are the regression parameters to be estimated and ϵ_s is an error term. Errors ϵ_s are supposed to follow the ordinary least squares assumptions, in particular to be i.i.d. The model is estimated by minimizing with respect to the βs the least squares error at locations s_i, $i \in \{1, \ldots, N\}$, where values $\hat{\eta}_i$ of $\eta(s_i)$ are known. This gives least square estimates $\tilde{\beta}_0$, $\tilde{\beta}_1$, $\ldots$, $\tilde{\beta}_p$. The predicted value at unknown location s is then given by

$$\tilde{\eta}(s) = \tilde{\beta}_0 + \tilde{\beta}_1 x_{1,s} + \ldots + \tilde{\beta}_p x_{p,s} \tag{8}$$

This is not an exact interpolation: at station location s_i, predicted value $\tilde{\eta}(s_i)$ given by Eq. (8) and known value $\hat{\eta}_i$ are usually not equal (the error is given by ϵ_{s_i} of Eq. 7).

5.1.3 Spline-based regression model

Consider the nonparametric regression model defined as

$$\eta(s) = F\left(x_{1,s}, \ldots, x_{p,s}\right) + \epsilon_s \tag{9}$$

where F is a function and ϵ_s is an error term. Errors ϵ_s are supposed to follow the classical regression assumptions, in particular to be i.i.d. If F is linear with respect to each variable x, then Eq. (9) is the linear regression model of Eq. (7). For a more complex behavior than a linear dependence, one

may model F as a smooth non-linear function of the covariates x's. A particularly convenient model results when F is taken as a penalized spline (P-spline henceforth) with radial basis function of order p, p being odd (Marx and Eilers, 1998). Estimation of Eq. (9) based on known values $\hat{\eta}_i$ is then performed by minimizing the sum of squared errors subject to some constraints to avoid overfitting. Technical details can be found in Appendix A.

A drawback of model Eq. (9) is that estimation may involve a very large number of free parameters, even for a low number p of covariates $x_{1,s}$, ..., $x_{p,s}$. In order to reduce the number of parameters to be estimated while still using the p covariates $x_{1,s}$, ..., $x_{p,s}$, one may combine the approaches Eqs. (7) and (9) by considering a partially linear model of the form

$$\eta(s) = \beta_0 + \beta_1 x_{1,s} + \ldots + \beta_q x_{q,s} + F\left(x_{q+1,s}, \ldots, x_{p,s}\right) + \epsilon_s \tag{10}$$

where F is a P-spline. Equation (10) belongs then to the family of generalized additive models (GAM) of Hastie and Tibshirani (1990). Estimation and prediction of such a model can be performed similarly to the spline regression model Eq. (9) by using a straightforward modification of matrix $\mathbf{X}$ of Appendix A. If $\tilde{\beta}$s denote the estimate of βs and $\tilde{F}$ the estimated P-spline in Eq. (10) then predicted value of η at location s is given by

$$\tilde{\eta}(s) = \tilde{\beta}_0 + \tilde{\beta}_1 x_{1,s} + \ldots + \tilde{\beta}_q x_{q,s} + \tilde{F}\left(x_{q+1,s}, \ldots, x_{p,s}\right). \tag{11}$$

This is usually not an exact interpolation (error at location s_i is given by ϵ_{s_i} of Eq. 10).

5.1.4 Kriging

Kriging is a stochastic interpolation method that computes the best linear unbiased estimator $\tilde{\eta}(s)$ of $\eta(s)$ based on a Gaussian model of the spatial dependence. Different kinds of kriging methods exist depending on the assumptions about the mean structure $\mathbb{E}[\eta(s)]$ of the model. The most general case, universal kriging, assumes that the mean is unknown but depends linearly on p covariates $x_{1,s}$, ..., $x_{p,s}$

$$\mathbb{E}[\eta(s)] = \beta(s) \equiv \beta_0 + \beta_1 x_{1,s} + \ldots + \beta_p x_{p,s}, \tag{12}$$

where coefficients βs have to be estimated. In other words, η at location s is modeled as

$$\eta(s) = \beta_0 + \beta_1 x_{1,s} + \ldots + \beta_p x_{p,s} + F(s) \tag{13}$$

where $\{F(s), s \in \mathcal{S}\}$ is a zero-mean Gaussian process. Model Eq. (13) belongs to the family of generalized linear geostatistical models as described by Diggle and Ribeiro (2007). Equations (10) and (13) are similar but differ in that in the former F is deterministic (a P-spline) whereas in the latter

F is stochastic (a Gaussian process). A formal connection between spline and kriging exists, see for example Cressie (1993), p. 180 and references therein.

Estimation of Eq. (13) involves estimating the βs parameters and the variogram describing the dependence structure in the Gaussian process F, which can be realized by maximum likelihood (see Diggle and Ribeiro, 2007, chapter 5). The minimum mean square error predictor $\tilde{\eta}$ of η at location s is then given by an equation of the form (Diggle and Ribeiro, 2007, chapter 6)

$$\tilde{\eta}(s) = \tilde{\beta}(s) + \sum_{i=1}^{N} w_i(s)\{\hat{\eta}_i - \tilde{\beta}(s)\} \tag{14}$$

where $\tilde{\beta}(s)$ is the estimated mean at location s and $w_i(s)$ are called prediction weights. This shows that the predicted value $\tilde{\eta}(s)$ is basically a weighted mean of the known values $\hat{\eta}_i$. However, unlike IDW of Sect. 5.1.1, weights $w_i(s)$, $i \in \{1, \ldots, N\}$, depend on the target location s. These weights can be positive, zero or negative depending on the correlation between locations s and s_i. When no nugget effect is considered (i.e. when the variogram is supposed to be continuous at the origin), kriging is an exact interpolation method.

5.2 Choice of covariates

The interpolation methods presented in the previous section are applied for interpolating the three GEV parameters, i.e. with η being either the location μ parameter, the scale σ parameter, or the shape ξ parameter. The first question to answer regards the choice of the covariates $x_{\cdot,s}$ to be used. As the final goal of this work is the mapping of return levels, maps of these covariates would be needed, or at least gridded values in Switzerland.

It is natural to consider the three geographical coordinates (longitude, latitude, altitude) as covariates in the interpolation methods. For snow events, altitude plays the most important role. As shown in Table 1, there is a strong linear dependence of the GEV parameters with altitude. The magnitude (modeled by μ) and spread (modeled by σ) of extreme snow depth strongly increase with elevation. On the contrary, as discussed in Sect. 4, ξ basically decreases with elevation, with positive values (heavy tailed distributions) at low altitudes and negative values (bounded distributions) at high altitudes. Similar results were found in Blanchet et al. (2009) regarding extreme snowfall in the Swiss Alps.

Additional covariates may help the prediction of extreme snow depth. Table 1 reveals that for the location and scale parameters, the mean snow depth is even more informative than elevation (higher R^2). The importance of mean snow depth shows that snow distribution in Switzerland cannot be completely described by the three geographical coordinates. Mean snow depth may contain additional information on local maxima and minima in snow distribution such as caused by individual mountain ranges. The positive

Table 1. Summary of linear dependence of the GEV parameters of Sect. 4 with altitude (columns 3 and 4) and with the mean snow depth (columns 5 and 6). In all cases, covariates are very significant (p-values lower than 10^{-6}). Altitude ranges between 230 and 2540 m, mean snow depth between 0.5 and 146 cm. The range of the $N = 84$ estimated GEV parameters is indicated in the first column.

	range of the GEV param.	versus altitude R^2	slope/km	versus mean snow depth R^2	slope/cm
Location μ	[4, 238]	0.68	84	0.98	1.65
Scale σ	[5, 61]	0.51	21.8	0.78	0.44
Shape ξ	[−0.38, 0.88]	0.40	−0.23	0.28	−0.003

correlation between mean snowfall and the location parameter μ of extreme snowfall has already been discussed by Blanchet et al. (2009) and it is therefore reasonable to have a positive correlation between mean and extreme snow depth as well. It is also reasonable to have a generally wider distribution of snow depth values if snow depth is large, i.e. a positive correlation between mean snow depth and the scale parameter, σ. Note, however, that for both snowfall and snow depth the shape parameter, ξ, is negatively correlated to the mean as discussed in Blanchet et al. (2009). This is partly a result of having many zero-events (rain instead of snow and zero snow depths) and a few much larger events for stations with a low mean.

To use it as a possible covariate in the spatial interpolation, a map of the mean snow depth is needed, which is not available. However, the mean snow depth is a very smooth process, even much smoother than a one-day event process such as maximum annual values or daily values. Its spatial interpolation is therefore easier than mapping daily snow depth as in Erxleben et al. (2002), Molotch et al. (2005) or Foppa et al. (2007) for example. A universal kriging interpolation (see section 5.1.4) with a linear trend on altitude gives very accurate interpolated mean values, with a root mean squared error (RMSE) of less than 5 cm (not shown). Using a digital elevation model (DEM henceforth) of Switzerland, gridded maps of mean snow depth can then be obtained and used as covariate for the location and scale GEV parameters.

Other meteorological variables, such as wind speed, wind direction and temperature, are also measured at the station locations. Nevertheless, a separate analysis did not reveal any significant influence of these variables on the GEV parameters. This may be partly due to the poor quality of these data and to the fact that many values are missing. Topographical variables such as slope, aspect, net solar radiation or vegetation used in Erxleben et al. (2002), Molotch et al. (2005), or Grünewald et al. (2010) could also be used as supplementary information, but are not considered in this work. Since local topographical variables are already of limited use in fine-scale snow depth analysis (Grünewald et al., 2010), we do not expect them to help explain our snow depth data, which are collected on flat fields. We will return to this point in the discussion (Sect. 7). Here the considered covariates for interpolating the GEV parameters are then the four covariates longitude, latitude, altitude and mean snow depth.

5.3 Considered models

We detail below the models used for interpolating the GEV parameters. As in Sect. 5.1, we denote η any of the three functions μ, σ or ξ to be interpolated. For sake of conciseness, we only detail in this section models when the four covariates longitude, latitude, altitude and mean snow depth are considered. Embedded models (in particular models when the mean snow depth is not accounted for) are particular cases which are straightforward to derive.

5.3.1 Inverse distance weighting

In order to account for the strong dependence of the GEV parameters with altitude and mean snow depth, one might rather use IDW with gradient correction as proposed in e.g. Nalder and Wein (1998):

$$\tilde{\eta}(s) = \frac{\sum_{i=1}^{N} \frac{\hat{\eta}_i + \beta_a (a_s - a_{s_i}) + \beta_m (m_s - m_{s_i})}{||l_s - l_{s_i}||}}{\sum_{i=1}^{N} \frac{1}{||l_s - l_{s_i}||}}, \quad (15)$$

where a_s is the altitude of location s, m_s its mean snow depth, l_s is the two-dimensional coordinate of this site (longitude, latitude), β_a and β_m are parameters. Compared to Eq. (6), Eq. (15) has the advantage of giving more weight to altitude and mean snow depth. This corrected IDW method is still an exact interpolation. Parameters β_a and β_m can be estimated by cross-validation by choosing values minimizing the score

$$CV(\beta_a, \beta_m) = \sum_{i=1}^{N} (\hat{\eta}_i - \tilde{\eta}_{-i}(s_i))^2,$$

where $\tilde{\eta}_{-i}(s_i)$ is the interpolated value at site s_i when this station is omitted in Eq. (15).

5.3.2 Linear regression models

Models of the form Eq. (7) are used where the covariates are polynomials of longitude, latitude, altitude and mean snow depth with a maximum degree of 3. We consider all possible combinations of these covariates with a maximum number of covariates (p in Eq. 7) equal to 6. We select the "best" linear regression model with the help of AIC, a penalized likelihood criteria (Akaike, 1974). Results of Sect. 5.5 will correspond to this model.

5.3.3 Spline regression models

We use models of the form

$$\eta(s) = \beta_0 + \beta_1 a_s + \beta_2 m_s + F(l_s) + \epsilon_s. \quad (16)$$

where F is a P-spline of order 3. Spline-regression models require to choose fix knots (see Appendix A). Choosing the best number of knots, and the best locations for them, is a difficult task. Here we fixed the number of knots to 15 which seems to be a good compromise between flexibility (large number of knots) and simplicity (low number of knots) of the model. These 15 knots are considered among the $N = 84$ station locations only. The best position for these 15 knots, from 10 000 of the total of $\binom{84}{15}$ possibilities, is made by generalized cross-validation. More precisely, 10 000 estimations with different choices of positions for the 15 knots are performed. Among the 10 000 estimated models, the one with the lowest GCV value (see Appendix A) is selected as the "best" model and results of Sect. 5.5 will correspond to this model.

5.3.4 Kriging

Model Eq. (16) is used with F being a Gaussian process with mean zero. No nugget effect is considered here, i.e. the variogram modeling the dependence structure is supposed to be continuous at the origin. Computation of this variogram involves the choice of a covariance function for F. Nine of the most commonly used covariance functions are used, namely the spherical, circular, cubic, Gneiting, exponential, Matérn, Gaussian, powered-exponential and Cauchy covariance functions (Schabenberger and Gotway, 2005). These covariance functions have one or two degrees of freedom and the four first ones have an upper-bound. Maximum likelihood estimation is performed with library geoR of R. The "best" model (i.e. the best covariance function) is then selected with the help of the AIC criteria (Akaike, 1974). Note that as no nugget effects are used here, kriging is an exact interpolation method.

5.4 Prediction comparison

To assess quality of the predictions, measures of accuracy will be used. The most stringent comparison is obviously to compute such measures for the validation stations. Note that validation stations were mainly selected for their climatological properties and not in order to achieve a high score in the validation (see Sect. 3). Therefore, the validation tests the reliability and stability of the predictions over all of Switzerland, or at least below 2500 m. However, it may also be of interest to assess the quality of the interpolated distributions for the fitted stations, in particular for the linear and spline regression models which are non-exact methods. Large differences between measures for the $N = 84$ fitting stations and the $M = 16$ validation stations indicate models that are questionable.

Here four measures of accuracy are used: the root mean-squared error (RMSE), the mean absolute error (MAE), the maximum prediction error (MPE) and the bias. These measures could be computed for assessing quality of the

interpolated $\tilde{\mu}(s_i)$, $\tilde{\sigma}(s_i)$ and $\tilde{\xi}(s_i)$ compared respectively to the individual values $\hat{\mu}_i$, $\hat{\sigma}_i$ and $\hat{\xi}_i$ of Sect. 4. This would however result in a comparison between two estimators, and not between an estimator and an observation. Furthermore, strictly speaking, it would not answer the question as to how well the data distribution is captured: the three best models for μ, σ and ξ separately might not be the best triplet of (μ, σ, ξ) since the GEV parameters are not orthogonal parameters. A better comparison is to assess goodness-of-fit of the quantiles of the interpolated GEV distribution compared to the observed ones. Let $z_i^{(1)}$, ..., $z_i^{(K)}$ be the $K = 43$ quantiles (i.e. sorted values) observed at a given station i. The probability associated to the k-th value $z_i^{(k)}$ is usually $p_k = \frac{k-1/2}{K}$ (function ppoints in R). $z_i^{(k)}$ can therefore be compared with the $(1 - p_k)$ quantile of the interpolated GEV distribution at station i, denoted $\tilde{q}_{p_k,i}$. It is given by Eq. (3), where μ, σ and ξ are replaced by their interpolated values $\tilde{\mu}(s_i)$, $\tilde{\sigma}(s_i)$ and $\tilde{\xi}(s_i)$ and p is replaced by p_k. The goodness-of-fit scores for quantile comparison are then given by

$$\text{RMSE} = \sqrt{\frac{1}{NK} \sum_{i=1}^{N} \sum_{k=1}^{K} \left(z_i^{(k)} - \tilde{q}_{p_k,i} \right)^2},$$

$$\text{MAE} = \frac{1}{NK} \sum_{i=1}^{N} \sum_{k=1}^{K} | z_i^{(k)} - \tilde{q}_{p_k,i} |,$$

$$\text{MPE} = \max_{i \in \{1,...,N\}} \max_{k \in \{1,...,K\}} | z_i^{(k)} - \tilde{q}_{p_k,i} |,$$

$$\text{Bias} = \frac{1}{NK} \sum_{i=1}^{N} \sum_{k=1}^{K} \left(z_i^{(k)} - \tilde{q}_{p_k,i} \right).$$

All these criteria involve quantities of the form $(z_i^{(k)} - \tilde{q}_{p_k,i})$ which is the error of predicting the $(1 - p_k)$ quantile of station i when using the interpolated GEV distribution. Quantile comparison should be made for both the $N = 84$ fitting stations and the $M = 16$ validation stations (replacing N by M is the previous scores). Note that, in the context of extremes, an alternative quantile validation score is used in Friederichs and Hense (2007) and in Maraun et al. (2010), still based on differences of the form $(z_i^{(k)} - \tilde{q}_{p_k,i})$ but where cases of overestimation (i.e. when $z_i^{(k)} - \tilde{q}_{p_k,i} < 0$) and of underestimation (i.e. when $z_i^{(k)} - \tilde{q}_{p_k,i} > 0$) of the observed quantiles $z_i^{(k)}$ are differently penalized. There is no reason to use this additional functionality in our case.

5.5 Results

The different interpolation methods presented in Sect. 5.3 are used for interpolating the three GEV parameters. Tables 2 and 3 summarize the scores of Sect. 5.4. In Table 2 only a DEM is used as covariates for the three GEV parameters. In Table 3, the mean snow depth is used as additional covariate

Table 2. Scores of quantile comparison when (a) fitting a GEV to each station separately as in Sect. 4; (b)–(e) interpolating the GEV parameters with a DEM as covariate.

	Fitting stations				Validation stations			
	RMSE	MAE	MPE	Bias	RMSE	MAE	MPE	Bias
(a) Pointwise GEV	6.7	3.6	227.9	0.1	5.4	3.3	53.0	0.2
(b) IDW	6.7	3.6	227.9	0.1	17.7	14.8	62.5	−2.2
(c) Linear regression	34.7	26.7	234.6	0.4	33	27.9	107.3	−4.8
(d) Spline regression	19	14.4	131.2	0.2	27.5	21.7	102.7	−6.4
(e) Kriging	6.7	3.6	227.9	0.1	16.2	12.7	71.6	−1.3

Table 3. Scores of quantile comparison when using a DEM and the mean snow depth as covariates. For the validation stations, either the kriged mean snow depths or the observed mean snow depths (scores in brackets) are used. Methods (a) to (d) are interpolation methods of Sect. 5. Method (e) refers to Sect. 6.

	Fitting stations				Validation stations			
	RMSE	MAE	MPE	Bias	RMSE	MAE	MPE	Bias
(a) IDW	6.7	3.6	227.9	0.1	14.0 (12.8)	10.3 (9.3)	69.9 (68.8)	−1.1 (−0.5)
(b) Linear regression	10.7	6.6	181.7	0.3	13.5 (12.1)	10.1 (8.9)	73.5 (72.6)	−1.2 (−0.7)
(c) Spline regression	9.5	6.0	123.1	0.2	12.8 (11.6)	9.3 (8.1)	87.5 (86.6)	−1.0 (−0.5)
(d) Kriging	6.7	3.6	227.9	0.1	12.9 (11.7)	9.4 (8.2)	61.9 (60.9)	−0.7 (−0.4)
(e) Smooth GEV	8.6	5.7	118.9	0.3	9.2 (8.3)	6.5 (5.4)	50.9 (48.6)	1.0 (0.6)

for the location and scale parameters, using either the kriged mean values (see Sect. 5.2) or the observed ones. For comparison, we also indicate in Table 2 the scores corresponding to fitting a GEV distribution to each station separately, including the validation stations, without any spatial model (see Sect. 4). For the validation stations, these scores thus do not correspond to predictions but to fittings, unlike all the other scores (lines b to e). They can thus only be interpreted as lower bounds of the error that would result from a prediction.

Table 2 suggests that, when using only longitude, latitude and elevation as covariates, kriging performs better as almost all scores are lower. IDW is the second best model. For both methods, results for the validation stations are relatively poor compared to those for the fitting stations, in particular for RMSE and MAE. This suggests that the prediction quickly deteriorates away from the fitting stations. Note that kriging and IDW are exact interpolation methods. This implies for example that the interpolated location $\tilde{\mu}(s_i)$ for the fitting station i is equal to the individual value $\hat{\mu}_i$ used in the interpolation. The same applies for the scale $\sigma(s)$ and shape $\xi(s)$. Interpolated and individual GEV distributions of the fitting stations are then identical (see lines a, b and e of Table 2). These scores are all low, with the exception of a very large MPE value (227.9) due to one single observation at the station Lugano in southern Ticino. Parameter ξ_i is likely to be overestimated which produces a strong overestimation of the largest observation.

Table 3 (lines a to d) compared to Table 2 confirms that using the mean snow depth as covariate for the location and scale parameters is helpful. There is a clear improvement in the spline and linear regression models for both the fitted and validation stations. For kriging and IDW, results for the validation stations are only slightly better and results for the fitted stations are exactly the same since they are exact interpolation techniques. All interpolation methods now have a similar performance but kriging still performs slightly better. Scores for the validation stations when using the observed mean snow depth as a covariate are better than when using the kriged mean snow depth but differences are low. This confirms again that the kriged mean snow depth is a very accurate estimation of the observed mean, as already discussed in Sect. 5.2. Note that even when using the observed mean snow depth, error measures from the smooth model are quite high compared to those when a GEV is fitted to each station separately (first line of Table 2). The errors cannot strictly be compared since the individual GEV fitting uses all available information at the validation stations for parameter estimation, while this information is not used in the parameter estimation of the smooth GEV. However, the difference in the errors shows that further improvements in the spatial aspect of the smooth GEV model may be possible.

Table 3 merely gives a global picture of the goodness-of-fit. A closer look at how well each station is fitted may be interesting. A way of summarizing goodness-of-fit of the quantiles for a given station i is to plot all observed quantiles
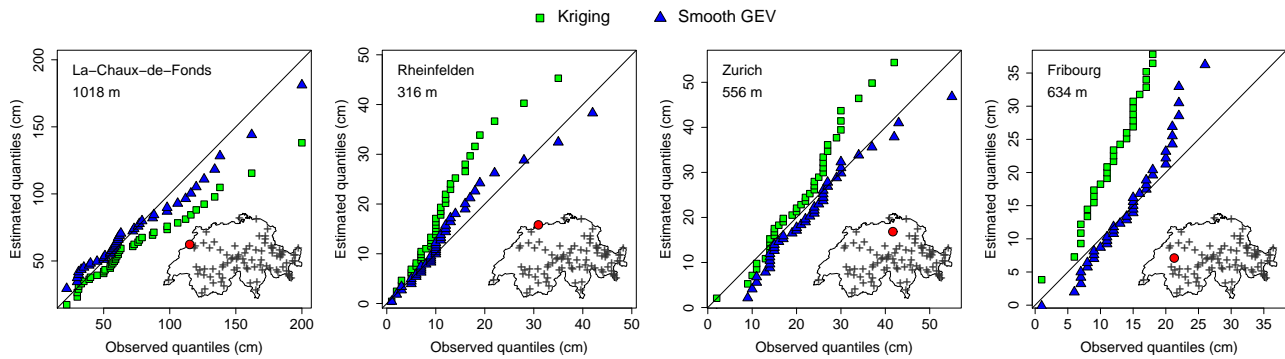
Fig. 5. QQ-plots of all four validation stations located in the Plateau, with kriging interpolation (green squares) and smooth GEV fitting (blue triangles). With both methods, longitude, latitude and elevation are used as covariates for the GEV parameters. The kriged mean snow depth is an additional covariate for the location and scale. Kriging interpolation is related to Sect. 5. Smooth GEV fitting is related to Sect. 6.

$z_i^{(k)}$, $k \in \{1, \ldots, 43\}$, against the modeled quantiles $\tilde{q}_{p_k,i}$: this is a QQ-plot. Figure 5 (green squares) depicts QQ-plots of all four validation stations located on the Swiss Plateau with kriging interpolation of Table 3. With a perfect fit, all points would lie on the diagonal line. In Fig. 5 the kriged mean snow depth is used as a covariate but results with the observed mean are almost similar. The figure reveals that the interpolated GEV distributions in the low elevation Plateau are quite poor. A comparison with QQ-plots of the validation stations located in the Alps (not shown) reveals a better fit in the Alps. One reason is that the station network in the Plateau is much less dense than in the Alpine region (see Fig. 2). The interpolation process will then produce a better fit in the Alps than on the Plateau, which has fewer stations. In addition, the statistically more extreme snow depth values on the Plateau (due to positive shape parameters, see Sect. 4) are by their very nature more difficult to model. However, the main drawback of the methodology is that the interpolation is done independently of the data. Of course, individual estimations $\hat{\mu}_i$, $\hat{\sigma}_i$ and $\hat{\xi}_i$ in Sect. 4 were done based on observed data. However, once the GEV parameters are estimated, they are considered as true values in the interpolation process, as if they were really observed. A bad individual estimate will therefore induce a bad interpolated value and may lead to models that are very unlikely for the data. This is no longer the case when a smooth GEV model is directly fitted to the data, as described in the next section.

6 Fitting a smooth GEV model

6.1 Smooth GEV modeling

There are crucial differences between the interpolation methods of Sect. 5. Unlike kriging and IDW, linear and spline regressions are generic models: once the model parameters have been estimated, prediction does not involve the individual values $\hat{\eta}_i$ anymore, which are only used for inference

(see Eqs. 8 and 11 compared to Eqs. 6 and 14). For these two cases, another approach is to directly estimate the regression parameters from the data, without involving the individual values $\hat{\eta}_i$. This method will be termed "smooth GEV modeling" because the GEV parameters are directly modeled as smooth functions in space. The main difference with the approach of Sect. 5 is that in interpolation methods, the spatial information is derived by interpolating individual GEV estimates whereas in the smooth GEV modeling it can be directly estimated from the data. More precisely, let η denote the surface model for either the location μ, scale σ or scale ξ parameter. We model the surface η at location s with the linear model

$$\eta(s) = \beta_0 + \beta_1 x_{1,s} + \ldots + \beta_p x_{p,s} \qquad (17)$$

as in linear regression prediction (Eq. 8), or with the more general additive model

$$\eta(s) = \beta_0 + \beta_1 x_{1,s} + \ldots + \beta_q x_{q,s} + F\left(x_{q+1,s}, \ldots, x_{p,s}\right) \qquad (18)$$

where F is a P-spline as in spline regression prediction (Eq. 11). Note that compared to the regression models Eqs. (7) and (10), models Eqs. (17) and (18) are deterministic as they do not comprise the stochastic part contained in the (Gaussian) residuals ϵ_s.

Smooth spline-based models similar to Eq. (18) have also been used for example in Hall and Tajvidi (2000), Ramesh and Davison (2002) and Padoan and Wand (2008) but for modeling smooth temporal trends of the GEV parameters at individual locations (i.e. with time as a covariate), rather than smooth spatial surfaces as in this article. In the spatial framework, recently quite simple linear regression models as in Eq. (17) have been used in Padoan et al. (2010) regarding US precipitation, using only latitude and elevation as covariates. The GEV modeling involves there in total only 7 parameters with a constant model $\xi(s) = \xi_0$ for the shape. However,

only 46 gauging stations were used over an area equivalent to 10 times Switzerland, with much flatter topography (maximum elevation around 1500 m). Due to the denser network used in this analysis, the rougher topography and the larger variability of snow depth in Switzerland, the surfaces responses given by Eqs. (17) and (18) to be used on our data are likely to be more complicated, i.e. to involve more covariates $x_{\cdot,s}$.

As in Sect. 5, we will consider as possible covariates the three geographical coordinates (longitude, latitude, elevation) and the mean snow depth, with the GEV parameters μ, σ and ξ being modeled by either Eqs. (17) or (18). Each combination of these three models then leads to a smooth GEV modeling of extreme snow depth in Switzerland. However, considering all possible combinations of models for μ, σ and ξ and all possible covariate choices would clearly be too computationally intensive: if K_μ models are considered for μ, K_σ models for σ and K_ξ for ξ, then in total, by combination of all possible models for each of the three parameters, this means that $K_\mu \times K_\sigma \times K_\xi$ smooth GEV models have to be fitted. This gives many thousands of models in our case. In order to limit the number of considered GEV models, we restrict our analysis to the best combinations of covariates found in the previous section using the linear Eq. (7) and spline Eq. (10) regression models. The linear regression models fitted in Sect. 5 used as possible covariate polynomials of longitude, latitude, altitude and mean snow depth with a maximum degree of 3 (see Sect. 5.3.2). Spline linear regression models used P-spline of order 3 with 15 knots and an additional possible linear dependence in elevation and in the mean (see Sect. 5.3.3). Among all those linear and spline regression models, we only select here as possible models for μ, σ and ξ in the smooth GEV model:

– for the location μ: the best two regression models with a DEM as covariate and the best four regression models with a DEM and the mean snow depth as covariates;

– for the scale σ: the best two regression models with a DEM as covariate and the best four regression models with a DEM and the mean snow depth as covariates. As σ (which models the spread of the GEV distribution) and μ (which models the center) are usually very correlated, we also allow the location to be a covariate for σ by considering the best four regression models with a DEM and the location as covariates;

– for the shape ξ: the best six regression models with a DEM as covariates.

Note that here only the equations of the best models for μ, σ and ξ are used, and not the values of the βs which will in fact be directly estimated from the data (see Sect. 6.2). This gives a total number of $6 \times 10 \times 6 = 360$ smooth GEV models. These models have between 10 and 57 degrees of freedom. The only model with 57 degrees of freedom is

when μ, σ and ξ are all modeled as in Eq. (18) with a linear dependence with the mean (for μ and σ) or with elevation (for ξ) and a smooth dependence in space (modeled in Eq. 18 through the P-spline F of order 3 with 15 knots).

6.2 Model estimation and selection

Unlike in Sect. 5, we wish to estimate the smooth GEV models directly from the data, which are considered jointly, without any individual fitting. We adopt a likelihood approach. This requires to consider the joint distribution of annual maximum snow depth at the N fitting locations. For the sake of simplicity, we will assume here that the N-variate density can be approximated by the product of marginal densities. This is equivalent to considering that the N annual maxima are approximately independent. This approximation is actually very unlikely to be fulfilled in reality due to the spatial dependence of annual maxima. However, it is applicable and gives satisfying results if the marginal distributions only are of interest, which is the case in this study. We will return to this approximation and its limits in the concluding discussion (Sect. 7). The log-likelihood of the N stations is then approximated by

$$l_a(\mu, \sigma, \xi) = \sum_{i=1}^{N} l\{\mu(s_i), \sigma(s_i), \xi(s_i)\} \qquad (19)$$

where μ, σ and ξ are smooth surfaces and $l\{\mu(s_i), \sigma(s_i), \xi(s_i)\}$ is the GEV log-likelihood of Eq. (5) when parameters (μ_i, σ_i, ξ_i) are replaced by $(\mu(s_i), \sigma(s_i), \xi(s_i))$. Approximation Eq. (19) is a special case of composite likelihood (Varin and Vidoni, 2005; Varin, 2008). Maximizing Eq. (19) consists then in finding the best smooth surfaces μ, σ and ξ for the observed data. It involves 10 to 57 unknown parameters. Note that in the individual fitting of Sect. 4, many more parameters were estimated: each individual GEV involves three parameters, leading to a total number of $3 \times N = 252$ parameters.

Let $\check{\boldsymbol{\beta}}$ denote the vector of all estimated parameters when maximizing Eq. (19) (we use notation "$\check{\boldsymbol{\beta}}$" instead of "$\tilde{\boldsymbol{\beta}}$" to differentiate with the estimated parameters of Sect. 5). As Eq. (19) is an approximated likelihood, usual properties of maximum likelihood estimates do not hold for $\check{\boldsymbol{\beta}}$. Nevertheless, theoretical properties are available from the theory of composite likelihood estimation (Varin and Vidoni, 2005; Varin, 2008). Under suitable regularity conditions, $\check{\boldsymbol{\beta}}$ is asymptotically unbiased and normal. Approximate confidence intervals for the GEV parameters can be computed based on the diagonal elements of its covariance matrix, estimable by $H(\check{\boldsymbol{\beta}})^{-1} J(\check{\boldsymbol{\beta}}) H(\check{\boldsymbol{\beta}})$ where $H(\check{\boldsymbol{\beta}})$ is the observed information matrix and $J(\check{\boldsymbol{\beta}})$ the squared score statistics corresponding to l_a in Eq. (19) (see e.g. Cox and Hinkley, 1974 for a definition of the information matrix and of the score statistics). Return levels can be computed at every site s by Eq. (3) with the estimated GEV parameters $\check{\mu}(s)$, $\check{\sigma}(s)$ and

$\check{\xi}(s)$ corresponding to $\check{\boldsymbol{\beta}}$. Approximate confidence intervals can be obtained by the delta method (Coles, 2001).

In our case, 360 smooth GEV models are estimated by maximizing Eq. (19) for different regression models of μ, σ and ξ. Model selection criteria are then needed to decide which of the fitted model should be preferred. We use the Takeuchi Information Criterion, TIC, (Takeuchi, 1976; Varin and Vidoni, 2005) defined as

$$\text{TIC} = -2\,l_a\,(\check{\boldsymbol{\beta}}) + 2\,\text{Tr}\,\{H(\check{\boldsymbol{\beta}})^{-1}\,J\,(\check{\boldsymbol{\beta}})\}, \qquad (20)$$

where l_a is the approximated likelihood Eq. (19). TIC is simply the AIC criterion (Akaike, 1974) extended to a misspecified likelihood function. As with the AIC, the best model will be that having the lowest value of TIC.

6.3 Results

The 360 smooth GEV models are fitted by maximizing the approximated log-likelihood Eq. (19), using R function `optim` initialized with the parameters obtained in Sect. 5. Values of TIC for the 360 estimated models range between 34 590 and 36 255. A clear feature is that models with μ depending only on a DEM always have higher values of TIC than those with μ depending on a DEM and the mean snow depth, regardless of the models for σ and ξ. This is illustrated in Fig. 6 where the 60 TIC values obtained for two different μ models (out of the six considered ones) are depicted as an example: models 1 to 60 use a DEM as covariates for μ; models 61 to 120 use a DEM and the mean snow depth as covariates for μ. As the best model is that having the lowest value of TIC, these results confirm that using the mean snow depth for modeling μ is helpful. Mean values have also been shown to be informative for spatial modeling of extreme precipitation in Cooley et al. (2007) and Blanchet et al. (2009) have pointed out similar regional trends in mean and extreme snowfalls.

Figure 6 also highlights that once a relatively good model has been chosen for μ (i.e. a model with the mean as covariate, models 61 to 120), using the mean or location as covariate for σ is preferable: models 61 to 73 (blue points) have clearly higher values of TIC than models 74 to 120 (green and red points), which use the location as additional information. Using the location as covariate for σ (in red) is usually preferred over using the mean (in green) but values of TIC barely differ: likelihoods are almost similar but less parameters have to be estimated. Last but not least, it seems from Fig. 6 that the model choice for ξ is not determining: each block delimited by two dotted lines corresponds to a given model for μ and σ, when all the six different models for ξ are used (each block then contains six values of TIC). Values of TIC change barely inside each block, meaning that the model for ξ does not really matter. This is partly because all models for ξ use exactly the same covariates (a DEM). A second reason is that actually, in terms of likelihood, it is much more important to fit the center of the distribution well
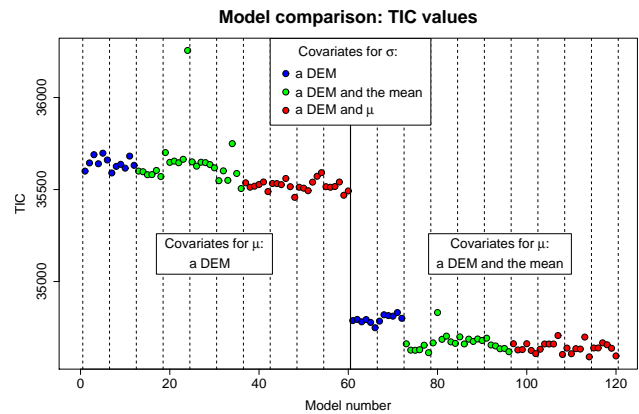

Fig. 6. Model comparison using TIC for two out of the six considered μ models. Models 1 to 60 use a DEM as covariates for μ; models 61 to 120 use a DEM and the mean snow depth as covariate for μ. Each block delimited by two dotted lines corresponds to a given model for μ and σ, when all the six different models for ξ are used.

(i.e. the location μ) than the tail which basically concerns only the largest values. Values of TIC may differ strongly between two models for μ, but differ usually less between two models for ξ. Over the 360 estimated models, the best smooth GEV model (number 114 in Fig. 6 corresponding to the lowest of the 360 values of TIC) has 19 degrees of freedom with linear regression models (Eq. 17) for μ, σ and ξ. More precisely, for this selected model, μ is a polynomial of longitude, latitude, elevation and mean snow depth (9° of freedom), σ is a polynomial of elevation and location μ (5 degrees of freedom) and ξ is a polynomial of longitude, latitude and elevation (5 degrees of freedom).

Table 3, line e, gives the goodness-of-fit measures for this selected model. The smooth GEV shows a better performance than all interpolation methods of Sect. 5 since all scores for the validation stations are lower. When using the observed or kriged mean as covariate, the results are very similar. A closer look at how well each station is fitted reveals that the smooth GEV distribution slightly outperforms the kriged GEV distribution in the Swiss Alps but that its better performance is more pronounced in the Swiss Plateau. This is visible in Fig. 5 depicting QQ-plots of all four validation stations located in the Plateau. The figure confirms a relatively good fit of the stations with the smooth GEV (blue triangles), with the exception of the largest quantiles of station Fribourg (right plot) which are overestimated due to an overestimation of the scale parameter. All the stations (even Fribourg) are clearly better fitted with the smooth GEV than with the kriged GEV. Annual maximum snow depth in the Plateau therefore seems to be better predicted. As previously mentioned (Sect. 5.5), the station network in this region is much less dense than in the Alps. The interpolation method seems to be more sensitive to this issue than our
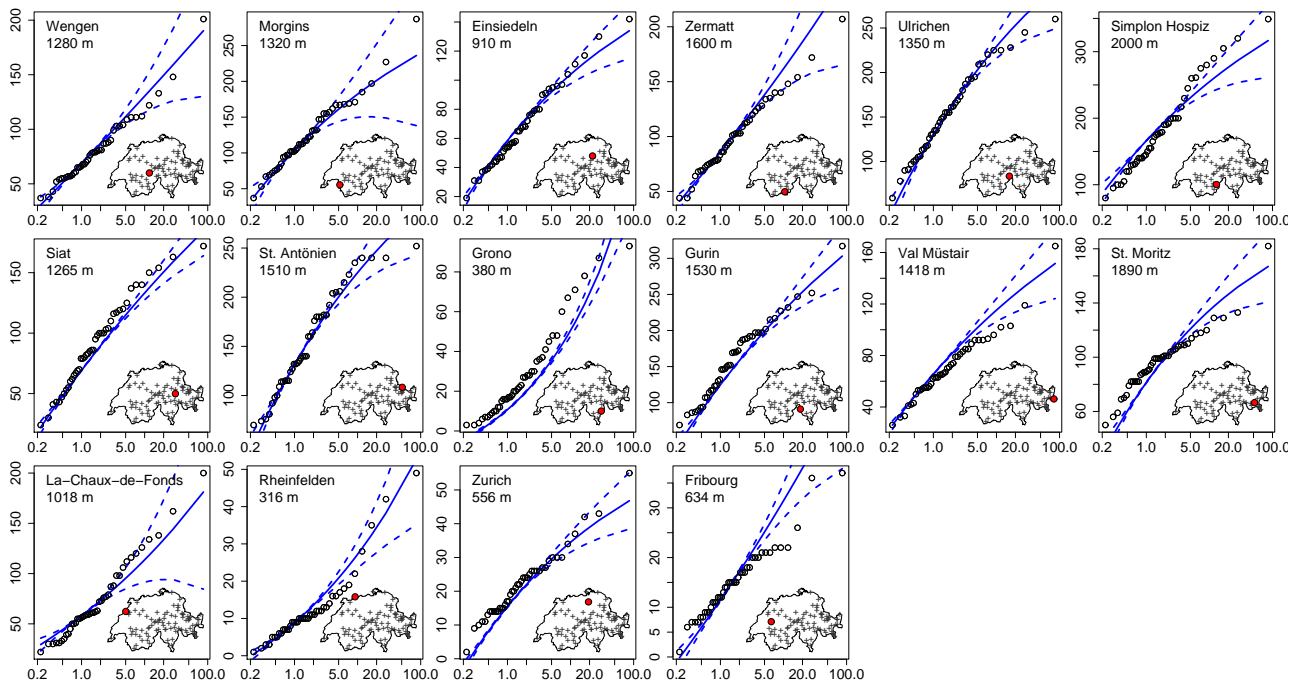
Fig. 7. Snow depth return level plots with the best smooth GEV model for the $M = 16$ validation stations, with 95% confident intervals (dotted lines). The kriged mean snow depth is used as an additional covariate. Red points in the Swiss map indicate the location of the station; its altitude is mentioned in the upper left corner. Crosses in the Swiss map locate the $N = 84$ stations used for fitting. Return levels (y-axis) are in centimeters (note the different scales among the plots); return periods (x-axis) are in years.

new likelihood-based approach in which all stations have the same weight.

Another advantage of smooth GEV modeling is that additional information regarding model uncertainty is available (see Sect. 6.2) which is not the case for all interpolation methods. Figure 7 depicts return level plots for the $M = 16$ validation stations together with 95% confidence intervals obtained by the delta method (see Sect. 6.2). This is also a way of assessing quality of the predicted distributions. Note that some of the validation stations show quite specific features compared to the surrounding stations (see Sect. 3), and therefore Fig. 7 corresponds to a quite difficult case of spatial prediction. The figure confirms a fairly good fit of the predicted distributions even in these difficult cases.

Smooth return level maps for any return period, which was the goal of this paper, can be computed from the smooth GEV model. For illustration, the 50-year return level map is depicted in the left panel of Fig. 8. The lowest 50-year return levels (around 10 cm) are obtained at the lowest elevations in the Plateau region but only for a few locations. The return level is approximately 40–50 cm in the main part of the Plateau (at 300–400 m elevation). The low region of Ticino lies at even lower altitude but its 50-year return level is slightly higher: it is around 70–80 cm in southern Ticino at 200–300 m elevation. The highest return levels (almost 8 m) are in the Gotthard region. This is a mountainous area but not the highest one in Switzerland (maximum elevation

around 3000 m). Even in the altitude range of the analyzed stations (basically below 2500 m), i.e. without relying on extrapolation, this is still the region having the largest return levels.

More information about regional variability of the 50-year return level can be obtained by removing the altitudinal effect from the parametric models of μ, σ and ξ. This is simply done by setting the β coefficients involving altitude in Eq. (17) to 0. This leads to "normalized" GEV parameters (i.e. without altitudinal dependency), from which one can compute "normalized" return levels. The normalized 50-year return level map is depicted in the right panel of Fig. 8. It shows that Grisons and part of southern Valais (reddish regions) have highly negative normalized return levels. This means that the 50-year return level in this region is lower than expected. These regions are indeed inner alpine dry valleys protected by high mountains all around, which usually shade the region from heavy snowfall events, and therefore annual maximum snow depths are usually quite low. In Fig. 8 the southern Valais is divided into three subareas: two areas with low residuals (red) separated by an area with higher residuals (green). However, boundaries between these three regions might be too sharp on the map due to the fact that there is only one station in the green area in between (see Fig. 2). Contrary to southern Valais, the Gotthard region (green area to the north of Ticino in Fig. 8) is open towards the direction of the main precipitation systems originating from the
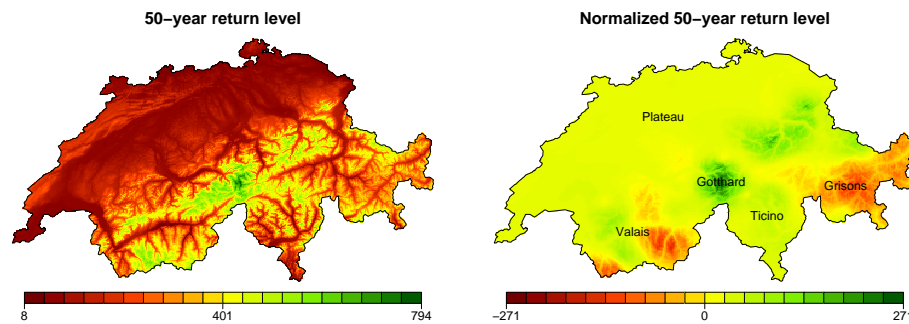
Fig. 8. 50-year return level map (left panel) and regional variability after removing the altitudinal effect (right panel) with the smooth GEV model. Units are centimeters.

north and south of the Alps. Annual maximum snow depths are therefore usually higher than expected at this altitude. Note that similar results have also been found in Blanchet et al. (2009) regarding extreme snowfall, which unsurprisingly appears to have similar regional variability to extreme snow depth.

A comparison with the 50-year return level map obtained by kriging the GEV parameters (Sect. 5) reveals the same main regional patterns. The kriging method seems to underestimate the 50-year return level in the Gotthard region and to overestimate it in the southern Valais. These two regions correspond respectively to the highest and lowest normalized return levels. This then basically means that kriging the GEV parameters gives maps which are too smooth. This is not surprising since the three GEV parameters are smoothed independently of the data. Each kriging is an interpolation of few points (namely the $N = 84$ individual estimates) which usually produces too smooth interpolated fields. Return levels are then computed as combinations of likely too smooth GEV parameters by Eq. (3), which then produces too smooth maps.

7 Discussion and outlook

This paper compares different techniques for mapping extreme snow depth in Switzerland. It suggests a better performance of a smooth GEV fitting than the most commonly used interpolation techniques, in particular where the station network is sparse. Suggestions for further developments are discussed below.

Several studies on snow depth mapping showed that in addition to elevation, using variables such as net solar radiation, slope, aspect or vegetation type can deliver useful information (Erxleben et al., 2002; Molotch et al., 2005). Improvements have also been obtained by the inclusion of variables representing wind redistribution of snow (Molotch et al., 2005). None of these variable has been used in this study. They could however easily be incorporated as covariates in the smooth GEV fitting, in the same way as we used

the mean snow depth. Nevertheless their influence on the GEV parameters might be difficult to assess here because the data used in this article are basically gathered in "ideal" conditions from flat, open and not overly exposed (to the wind) fields. The analysis here aims therefore to assess extreme snow depth and return levels independently from any modification through the small scale local terrain. A way to increase the data basis for our analysis further would be to incorporate the SLF automatic IMIS stations (Interkantonales Mess- und Informationssystem), which are located at higher elevations (typically above 2200 m) and therefore often in more rugged environments but still in locally flat terrain. We have not done this here because of the relatively short-time series available for SLF automatic stations. Only approximately 10 years of data are available, which is short, particularly in the framework of block maxima. Using a model for exceedances over high thresholds (Davison and Smith, 1990) could work in this case. However, care must be taken that extremely snowy winters (such as winter 1999 in the Swiss Alps) or extremely snow-scarce winters do not bias the GEV fitting of short-term stations compared to long-term ones. Another issue is the potential impact of climate change on extreme snow depth, particularly when combining short- and long-term data. A potential trend effect has not been addressed in this article. It could however easily be accounted for by using time as a covariate in the smooth GEV modeling. Many studies show that mean snow levels and snow days have been affected by climate change (Marty, 2008; Beniston et al., 2003; Scherrer et al., 2004; Bavay et al., 2009; Bocchiola and Diolaiuti, 2010). However, the impact of climate change on *extreme* is unclear at present, as shown in Bocchiola et al. (2008). Nevertheless, there is, to our knowledge, no study on long-term trend in extreme snow events based exclusively on extreme value theory and a large potential area of research still remains open on this subject (Katz, 2010).

The smooth GEV modeling intends to model that snow depth return levels at neighboring locations are likely to be similar. This means for example that the maximum snow depth value expected once every 50 years at neighboring locations is likely to be similar. However, the model does not

assume that this maximum is expected to occur the same year, or more generally that in a given year annual values are likely to be similar. On the contrary, the underlying assumption made in the likelihood fitting of Sect. 6 is that annual values are approximately independent, which permitted to write the likelihood as being a sum of GEV likelihoods (Eq. 19). This is a simplifying approximation which is unlikely to be met in reality: if a location has received a large amount of snow in a given year, a neighboring location is also likely to have received a large amount of snow that same year. This type of spatial dependence is not accounted for in the present methodology. Approximation Eq. (19) gives satisfying results for computing local return levels but is not likely to give the best possible results for computing regional return levels, i.e. probabilities of exceeding some specific level anywhere in a region a given year. This is, however, the type of question that often needs to be answered in risk management and land-use planning. For such issues, spatial dependence between annual maximum values would have to be accounted for. This is possible by using the exact framework of *spatial extremes*. The most natural way for the specification of spatial extremes is provided by the theory of max-stable processes which is a current active topic of research in the statistical community. Modeling of spatial dependence of extreme snow depth in Switzerland in the framework of max-stable processes has been provided in Blanchet and Davison (revised), based on normalized time-series to get rid of the GEV margins. Combining both, the smooth GEV modeling of this article and the spatial dependence of Blanchet and Davison (revised) would provide a complete modeling of extreme snow depth in Switzerland and will be attempted in future. Both, the smooth intensities (through the smooth GEV model) and the spatial dependence (through the max-stable process) would be explicitly modeled. This has been achieved in Padoan et al. (2010) for US precipitation using a simpler model and smoother data (due to the flatter topography and the relative sparcity of the stations): simpler GEV models than those of this article and less sophisticated max-stable model than Blanchet and Davison (revised) allowed this to be combined. Future work will show whether this can also be achieved for our more complicated problem. In the meantime, the presented smooth solution to GEV and return period calculations presented here is a practical improvement over simple interpolation as is commonly done in application oriented work.

Appendix A

P-splines with radial basis function

Consider the spline-based regression model

$$\eta(s) = F(x_{1,s}, ..., x_{p,s}) + \epsilon_s,$$

where $x_{1,s}, ..., x_{p,s}$ are p covariates at location s and F is a P-spline with radial basis function of order p, p being odd. For sake of clarity, we will consider in the following the case of one single covariate x_s. The generalization to p covariates is straightforward. The considered spline-based regression model is then

$$\eta(s) = F(x_s) + \epsilon_s.$$

F can be written as

$$F(x) = \beta_0 + \beta_1 x + ... + \beta_{m-1} x^{m-1} \qquad (A1)$$

$$+ \sum_{r=1}^{R} \beta_{m+r-1} \mid x - \kappa_r \mid^{2m-1},$$

where $m = (p+1)/2$, $\kappa_1, ..., \kappa_R$ is a set of fixed knots and $\beta_0, \beta_1, ..., \beta_{m+R-1}$ are coefficients to be estimated. With the notations of Sect. 5, it is assumed that estimates $\hat{\eta}_1, ..., \hat{\eta}_N$ of η at N locations $s_1, ..., s_N$ are available. The goal is to estimate the best βs in Eq. (A1) based on the known $\hat{\eta}$s. Adopting a matrix notation, the sum of squared errors can be written as $||\hat{\boldsymbol{\eta}} - \mathbf{X}\boldsymbol{\beta}||^2$ where $\hat{\boldsymbol{\eta}} = (\hat{\eta}_1, ..., \hat{\eta}_N)^T$ is a known $N \times 1$ vector, $\boldsymbol{\beta} = (\beta_0, \beta_1, ..., \beta_{m+R-1})^T$ is a $(m+R)$-dimensional vector to be estimated, and $\mathbf{X}$ is the $N \times (m+R)$ matrix

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 & ... & x_1^{m-1} & |x_1 - \kappa_1|^{2m-1} & ... & |x_1 - \kappa_R|^{2m-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_N & ... & x_N^{m-1} & |x_N - \kappa_1|^{2m-1} & ... & |x_N - \kappa_R|^{2m-1} \end{bmatrix}.$$

To avoid overfitting, one aims to minimize the sum of squared errors subject on some constraint on the β parameter i.e.

Minimize $||\hat{\boldsymbol{\eta}} - \mathbf{X}\boldsymbol{\beta}||^2$ subject to $\boldsymbol{\beta}^T \mathbf{M} \boldsymbol{\beta} \leq c$ (A2)

for a judicious choice of c and a given matrix $\mathbf{M}$ of dimension $(m+R) \times (m+R)$. One possible choice of $\mathbf{M}$ is $\mathbf{M} = \mathbf{M}_*^T \mathbf{M}_*$ with

$$\mathbf{M}_* = \begin{bmatrix} 0 & ... & 0 & 0 & ... & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & ... & 0 & 0 & ... & 0 \\ 0 & ... & 0 & |\kappa_1 - \kappa_1|^{m-1/2} & ... & |\kappa_1 - \kappa_R|^{m-1/2} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & ... & 0 & |\kappa_R - \kappa_1|^{m-1/2} & ... & |\kappa_R - \kappa_R|^{m-1/2} \end{bmatrix}$$

where the first m rows and m columns of $\mathbf{M}_*$ are zeros. Using the Lagrange multiplier argument, the constraint optimization problem Eq. (A2) is equivalent to choosing β minimizing

$$||\hat{\boldsymbol{\eta}} - \mathbf{X}\,\boldsymbol{\beta}||^2 + \lambda\,\boldsymbol{\beta}^T\,\mathbf{M}\,\boldsymbol{\beta} \qquad (A3)$$

for some $\lambda \geq 0$, fixed, called the *smoothing parameter* as it controls the amount of smoothing. An automatic choice for λ is to minimize the cross-validation score. However, in terms of invariance, it may be preferable (Wood, 2006) to choose λ minimizing the *generalized cross-validation* (GCV) score

$$GCV(\lambda) = N^2 \sum_{i=1}^{N} \left(\frac{\hat{\eta}_i - \tilde{\eta}(s_i)}{\mathrm{Tr}\,(\mathbf{Id} - \mathbf{Q}_\lambda)} \right)^2,$$

where $\mathbf{Q}_\lambda$ is the *smoother matrix* $\mathbf{Q}_\lambda = \mathbf{X}(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{M})^{-1}\mathbf{X}^T$.

For any fixed $\lambda \geq 0$, it can be shown that problem Eq. (A3) has the solution

$$\tilde{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{M})^{-1}\mathbf{X}^T\hat{\boldsymbol{\eta}}.$$

The predicted values at location s is then given by

$$\tilde{\eta}(s) = \tilde{F}(x_s) = \tilde{\beta}_0 + \tilde{\beta}_1\,x_s + \ldots + \tilde{\beta}_{m-1}\,x_s^{m-1}$$

$$+ \sum_{r=1}^{R} \tilde{\beta}_{m+r-1}|x_s - \kappa_r|^{2m-1}.$$

References

Akaike, H.: A new look at the statistical model identification, IEEE T. Automat. Contr., 19, 716–723, doi:10.1109/TAC.1974.1100705, 1974.

Bavay, M., Lehning, M., Jonas, T., and Löwe, H.: Simulations of future snow cover and discharge in Alpine headwater catchments, Hydrol. Process., 23, 95–108, doi:10.1002/hyp.7195, 2009.

Beguería, S. and Vicente-Serrano, S. M.: Mapping the Hazard of Extreme Rainfall by Peaks over Threshold Extreme Value Analysis and Spatial Regression Techniques, J. Appl. Meteorol. Clim., 45, 108–124, doi:10.1175/JAM2324.1, 2006.

Beniston, M., Keller, F., Koffi, B., and Goyette, S.: Estimates of snow accumulation and volume in the Swiss Alps under changing climatic conditions, Theor. Appl. Climatol., 76, 125–140, 2003.

Blanchet, J. and Davison, A. C.: Spatial modelling of extreme snow depth, revised, 2010.

Blanchet, J., Marty, C., and Lehning, M.: Extreme value statistics of snowfall in the Swiss Alpine region, Water Resour. Res., 45, W05424, doi:10.1029/2009WR007916, 2009.

Bocchiola, D. and Diolaiuti, G.: Evidence of climate change within the Adamello Glacier of Italy, Theor. Appl. Climatol., 100, 351–369, doi:10.1007/s00704-009-0186-x, 2010.

Bocchiola, D., Medagliani, M., and Rosso, R.: Regional snow depth frequency curves for avalanche hazard mapping in central Italian Alps, Cold Reg. Sci. Technol., 46, 204–221, 2006.

Bocchiola, D., Bianchi Janetti, E., Gorni, E., Marty, C., and Sovilla, B.: Regional evaluation of three day snow depth for avalanche hazard mapping in Switzerland, Nat. Hazards Earth Syst. Sci., 8, 685–705, doi:10.5194/nhess-8-685-2008, 2008.

Brown, B. G. and Katz, R. W.: Regional analysis of temperature extremes: spatial analog for climate change?, J. Climate, 8, 108–119, 1995.

Bründl, M., Etter, H.-J., Steiniger, M., Klingler, Ch., Rhyner, J., and Ammann, W. J.: IFKIS – a basis for managing avalanche risk in settlements and on roads in Switzerland, Nat. Hazards Earth Syst. Sci., 4, 257–262, doi:10.5194/nhess-4-257-2004, 2004.

Coles, S.: An Introduction to Statistical Modelling of Extreme Values, Springer, New York, 2001.

Cooley, D., Nychka, D., and Naveau, P.: Bayesian Spatial Modeling of Extreme Precipitation Return Levels, J. Am. Stat. Assoc., 102, 824–840, doi:10.1198/016214506000000780, 2007.

Cox, D. and Hinkley, D.: Theoretical Statistics, Chapman and Hall, 1974.

Cressie, N. A.: Statistics for spatial data, Wiley Series in Probability and Statistics, 1993.

Davison, A. C. and Smith, R. L.: Models for exceedances over high thresholds (with discussion), J. Roy. Stat. Soc. Ser. B, 52, 393–442, 1990.

Diggle, P. J. and Ribeiro, P. J.: Model-based Geostatistics, Springer Series in Statistics, New York, 2007.

Erxleben, J., Elder, K., and Davis, R.: Comparison of spatial interpolation methods for estimating snow distribution in the Colorado Rocky mountains, Hydrol. Process., 16, 3627–3649, doi:10.1002/hyp.1239, 2002.

Foppa, N., Stoffel, A., and Meister, R.: Synergy of in situ and space borne observation for snow depth mapping in the Swiss Alps, International J. Appl. Earth Obs., 9, 294–310, doi:10.1016/j.jag.2006.10.001, 2007.

Friederichs, P. and Hense, A.: Statistical Downscaling of Extreme Precipitation Events Using Censored Quantile Regression, Mon. Weather Rev., 135, 2365–2378, doi:10.1175/MWR3403.1, 2007.

Gardes, L. and Girard, S.: Conditional extremes from heavy-tailed distributions: an application to the estimation of extreme rainfall return levels, Extremes, 13, 177–204, 2010.

Grünewald, T., Schirmer, M., Mott, R., and Lehning, M.: Spatial and temporal variability of snow depth and ablation rates in a small mountain catchment, The Cryosphere, 4, 215–225, doi:10.5194/tc-4-215-2010, 2010.

Hall, P. and Tajvidi, N.: Nonparametric Analysis of Temporal Trend When Fitting Parametric Models to Extreme-Value Data, Stat. Sci., 15, 153–167, 2000.

Harshburger, B. J., Humes, K. S., Walden, V. P., Blandford, T. R., Moore, B. C., and Dezzani, R. J.: Spatial interpolation of snow water equivalency using surface observations and remotely sensed images of snow-covered area, Hydrol. Process., 24, 1285–1295, doi:10.1002/hyp.7590, 2010.

Hastie, T. J. and Tibshirani, R. J.: Generalized Additive Models, Chapman & Hall, New York, 1990.

Katz, R. W.: Statistics of extremes in climate change, Climatic Change, 100, 71–76, doi:10.1007/s10584-010-9834-5, 2010.

Katz, R. W., Parlange, M. B., and Naveau, P.: Statistics of extremes in hydrology, Adv. Water Resour., 25, 1287–1304, 2002.

Kohnová, S., Parajka, J., Szolgay, J., and Hlavcová, K.: Bioclimatology and Natural Hazards, chap. Mapping of Gumbel Extreme Value Distribution Parameters for Estimation of Design Precipitation Totals at Ungauged Sites, Springer Verlag, 129–136, 2009.

Lateltin, O. and Bonnard, C.: Hazard assessment and land-use planning in Switzerland for snow avalanches, floods and landslides, in: The Comprehensive Risk Assessment for Natural Hazards – A contribution to the International decade for natural disasters reduction (IDNDR), Technical Reports, World Meteorological Organization, sols, 1999.

Leadbetter, M. R., Lindgren, G., and Rootzén, H.: Extremes and Related Properties of Random Sequences and Processes, Springer Verlag, New York, 1983.

López-Moreno, J. I. and Nogués-Bravo, D.: A generalized additive model for the spatial distribution of snowpack in the Spanish Pyrenees, Hydrol. Process., 19, 3167–3176, doi:10.1002/hyp.5840, 2005.

Loukas, A., Vasiliades, L., Dalezios, N., and Domenikiotis, C.: Rainfall-frequency mapping for Greece, Phys. Chem. Earth Pt. B, 26, 669–674, doi:10.1016/S1464-1909(01)00067-3, 2001.

Maraun, D., Rust, H., and Osborn, T.: Synoptic airflow and UK daily precipitation extremes, Extremes, 13, 133–153, doi:10.1007/s10687-010-0102-x, 2010.

Marchand, W.-D., Killingtveit, A., Wilén, P., and Wikström, P.: Comparison of Ground-Based and Airborne Snow Depth Measurements with Georadar Systems, Case Study, Nord. Hydrol., 34, 427–448, doi:10.2166/nh.2003.025, 2003.

Marty, C.: Regime shift of snow days in Switzerland, Geophys. Res. Lett., 35, L12501, doi:10.1029/2008GL033998, 2008.

Marx, B. D. and Eilers, P. H. C.: Direct generalized additive modeling with penalized likelihood, Comput. Stat. Data An., 28, 193–209, doi:10.1016/S0167-9473(98)00033-4, 1998.

Molotch, N. P., Colee, M. T., Bales, R. C., and Dozier, J.: Estimating the spatial distribution of snow water equivalent in an alpine basin using binary regression tree models: the impact of digital elevation data and independent variable selection, Hydrol. Process., 19, 1459–1479, doi:10.1002/hyp.5586, 2005.

Nalder, I. A. and Wein, R. W.: Spatial interpolation of climatic normals: test of a new method in the Canadian boreal forest, Agr. Forest Meteorol., 92, 211–225, 1998.

Naveau, P., Nogaj, M., Ammann, C., Yiou, P., Cooley, D., and Jomelli, V.: Statistical methods for the analysis of climate extremes, Compt. Rendus Geosci., 337, 1013–1022, 2005.

Padoan, S. and Wand, M.: Mixed model-based additive models for sample extremes, Stat. Probabil. Lett., 78, 2850–2858, doi:10.1016/j.spl.2008.04.009, 2008.

Padoan, S. A., Ribatet, M., and Sisson, S. A.: Likelihood-based inference for max-stable processes, J. Am. Stat. Assoc., 105, 263–277, doi:10.1198/jasa.2009.tm08577, 2010.

Palutikof, J. P., Brabson, B. B., Lister, D. H., and Adcoc, S. T.: A review of methods to calculate extreme wind speeds, Meteorol. Appl., 6, 119–132, 1999.

Ramesh, N. I. and Davison, A. C.: Local models for exploratory analysis of hydrological extremes, J. Hydrol., 256, 106–119, doi:10.1016/S0022-1694(01)00522-4, 2002.

Reiss, R. D. and Thomas, M.: Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields, Birkhäuser, Basel, Switzerland, 3 edition, 2007.

Schabenberger, O. and Gotway, C. A.: Statistical Methods for Spatial Data Analysis, Chapman & Hall/CRC, 2005.

Scherrer, S. C., Appenzeller, C., and Laternser, M.: Trends in Swiss Alpine snow days: The role of local- and large-scale climate variability, Geophys. Res. Lett., 31, L13215, doi:10.1029/2004GL020255, 2004.

SLF: Der Lawinenwinter 1999 Ereignisanalyse, Eidg. Institut für Schnee- und Lawinenforshung SLF, 2000.

Smith, R. L.: Maximum likelihood estimation in a class of nonregular cases, Biometrika, 72, 67–90, doi:10.1093/biomet/72.1.67, 1985.

Szolgay, J., Parajka, J., Kohnová, S., and Hlavcová, K.: Comparison of mapping approaches of design annual maximum daily precipitation, Atmos. Res., 92, 289–307, doi:10.1016/j.atmosres.2009.01.009, 2009.

Takeuchi, K.: Distribution of informational statistics and a criterion of model fitting, Suri-Kagaku (Mathematic Sciences), 153, 12–18, 1976.

Varin, C.: On composite marginal likelihoods, Adv. Stat. Anal., 92, 1–28, doi:10.1007/s10182-008-0060-7, 2008.

Varin, C. and Vidoni, P.: A note on composite likelihood inference and model selection, Biometrika, 92, 519–528, doi:10.1093/biomet/92.3.519, 2005.

Vasiliev, O., Van Gelder, P., Plate, E., and Bolgov, M.: Extreme Hydrological Events: New Concepts for Security, vol. 78 of NATO Science Series IV: Earth & Environmental Sciences, Springer, 2007.

Weisse, A. K. and Bois, P.: Topographic Effects on Statistical Characteristics of Heavy Rainfall and Mapping in the French Alps, J. Appl. Meteorol., 40, 720–740, doi:10.1175/1520-0450(2001)040⟨0720:TEOSCO⟩2.0.CO;2, 2001.

Weisse, A. K. and Bois, P.: A comparison of methods for mapping statistical characteristics of heavy rainfall in the French Alps: the use of daily information, Hydrolog. Sci. J., 47, 739–752, 2002.

Wood, S.: Generalized Additive Models: An Introduction with R, Chapman & Hall, 2006.