A non-stationary extreme value approach for climate projection ensembles: application to snow loads in the French Alps

Erwan Le Roux[1], Guillaume Evin[1], Nicolas Eckert[1], Juliette Blanchet[2], and Samuel Morin[3]

[1]Univ. Grenoble Alpes, INRAE, UR ETNA, Grenoble, France
[2]Univ. Grenoble Alpes, Grenoble INP, CNRS, IRD, IGE, Grenoble, France
[3]Univ. Grenoble Alpes, Univ. Toulouse, Météo France, CNRS, CNRM, CEN, Grenoble, France

Correspondence: Erwan Le Roux (erwan.le-roux@inrae.fr)

Abstract. Anticipating risks related to climate extremes often relies on the quantification of large return levels (values exceeded with small probability) from climate projection ensembles. Current approaches based on multi-model ensembles (MMEs) usually estimate return levels separately for each chain of the MME. By contrast, using MME obtained with different combinations of general circulation model (GCM) and regional climate model (RCM), our approach estimates return levels together from the
5   past observations and all GCM-RCM pairs, considering both historical and future periods. The proposed methodology seeks to provide estimates of projected return levels accounting for the variability of individual GCM-RCM trajectories, with a robust quantification of uncertainties. To this aim, we introduce a flexible non-stationary generalized extreme value (GEV) distribution that includes i) piecewise linear functions to model the changes in the three GEV parameters ii) adjustment coefficients for the location and scale parameters to adjust the GEV distributions of the GCM-RCM pairs with respect to the GEV distribution
10   of the past observations. Our application focuses on snow load at 1500 m elevation for the 23 massifs of the French Alps, which is of major interest for the structural design of roofs. Annual maxima are available for 20 adjusted GCM-RCM pairs from the EURO-CORDEX experiment, under the scenario RCP8.5. Our results show with a model-as-truth experiment that at least two linear pieces should be considered for the piecewise linear functions. We also show, with a split-sample experiment, that eight massifs should consider adjustment coefficients. These two experiments help us select the GEV parameterizations for each
15   massif. Finally, using these selected parameterizations, we find that the 50-year return level of snow load is projected to decrease in all massifs, by -2.9 kN m^{-2} (-50%) on average between 1986-2005 and 2080-2099 at 1500 m elevation and RCP8.5. This paper extends to climate extremes the recent idea to constrain climate projection ensembles using past observations.

1 Introduction

Climate model simulation is one of the main scientific paradigms to anticipate extreme climate events. In particular, multimodel
20   GCM-RCM ensembles are widely used to quantify the changes in climate extremes and their uncertainties (IPCC, 2019). General circulation models (GCMs) represent key processes of the climate system relevant at the global scale, and provide input for regional climate models (RCMs) used to downscale and refine the climate projections at the local to regional scale.

Climate extremes are usually assessed within the statistical framework of extreme value theory (EVT), by focusing either on annual maxima or on values exceeding a high threshold (Coles, 2001). EVT makes it possible to robustly estimate return levels,

25   i.e. extreme quantiles that occur on average once every T years, where T is the corresponding return period. Return levels play
a key role in the design of structures (dams, protections, roofs) to withstand the effects of natural hazards (floods, avalanches,
wildfires, snow loads), see e.g. Rao and Hamed (2000); Eckert et al. (2008); Evin et al. (2018); Le Roux et al. (2020).

Most approaches using EVT to study climate extremes from multi-model ensembles (MMEs) rely on stationary generalized
extreme value (GEV) distributions estimated separately on each chain of the MME, i.e. with each ensemble member (Kharin
30   et al., 2007; Beniston et al., 2007). Specifically, for each ensemble member, annual maxima are assumed stationary for two
time slices of 20/30 years: one in the historical period representing the late 20th century climate, and one in the future period.
For instance, Fowler et al. (2007) opted for two 30-year time slices: 1961-1990 and 2071-2100. Next, stationary 20/30-year
return levels are estimated for each time slice with a GEV distribution. Finally, average changes, i.e. differences of return
levels between time slices averaged on all ensemble members, are usually reported (Kharin et al., 2013; O'Gorman, 2014).
35   However, such approaches based on stationary GEV distributions have several drawbacks. First, the assumption of stationarity
for 20/30 consecutive annual maxima can be debatable, and the possibility of a trend within the 20/30 years time slices is
often not checked (Kharin and Zwiers, 2004). Then, the choice to rely only on 20/30 maxima implies that the estimated GEV
parameters have large uncertainties. In this case, large return levels, e.g. 50-year (or even larger) return periods which are
usually considered to design structures (see e.g., Tab.1 of Cabrera et al., 2012), can be highly uncertain.

40   Temporal non-stationary GEV approaches address these limitations by taking into account all the available annual maxima
for each ensemble member, i.e. all the historical and future annual maxima are fitted with a single statistical model (Kharin
and Zwiers, 2004). Such approaches combine a stationary random component (a fixed extreme value distribution) with non-
stationary deterministic functions that map each temporal covariate (such as the years or the global mean temperatures) to the
changing parameters of the distribution (Montanari and Koutsoyiannis, 2014). Another advantage of temporal non-stationary
45   approaches is that they allow return levels to be estimated conditionally on each temporal covariate (Kharin et al., 2013).

A majority of temporal non-stationary approaches for MMEs rely on the GEV distributions estimated separately with each
chain of the MME (Tab. 1), with some exceptions (Caires et al., 2006; Kyselý et al., 2010; Roth et al., 2014; Winter et al.,
2017), and report return levels (conditionally on a given covariate) averaged on all ensemble members. We believe that such
approaches are sub-optimal because they estimate one non-stationary GEV distribution with each chain of the MME, i.e. with
50   roughly less than 200 maxima, which often implies simple parameterization (linear) for the non-stationary functions (Tab. 1).

Our study follows an alternative approach which relies on temporal non-stationary GEV distributions fitted together on
all ensemble members. This approach enables us to robustly quantify uncertainties using standard tools from non-stationary
extreme value analysis (AghaKouchak et al., 2012). Such an approach has mainly been proposed for initial condition ensembles
(Tab. 1), i.e. ensemble members that consist of replicates from the same GCM-RCM pair (or same GCM for GCM ensembles)
55   simulated with different initial conditions. For initial condition ensembles, this alternative approach estimates a single non-
stationary distribution on all ensemble members by assuming that they are independent and identically distributed (*iid*).

However, this alternative approach is inadequate for GCM-RCM ensembles with several GCMs because the *iid* assumption,
i.e. that all GCM-RCM pairs follow the same non-stationary distribution, is unlikely to hold in all the cases.

Our study fills this gap with a novel non-stationary extreme value approach inspired by the recent trend of statistical methods
60 that constrain climate projections using past observations (Brunner et al., 2020). We propose to fit a non-stationary GEV
distribution together from the past observations and all GCM-RCM pairs, without necessarily assuming that all GCM-RCM
pairs follow the same distribution. To this end, we introduce adjustment coefficients for the location and scale parameters of the
GEV distribution to represent the variability of climate trajectories, i.e. the various GEV distributions of all GCM-RCM pairs.
Specifically, the non-stationary GEV distribution is jointly fitted i) without adjustment coefficients on the past observations
65 to represent the most likely climate trajectory ii) with adjustment coefficients on the GCM-RCM pairs to account for the
variability of climate trajectories. Besides, non-stationary GEV based approaches for climate projections ensembles usually
consider linear functions for the non-stationary functions, with the exception of Um et al. (2017) that uses nonlinear functions
(Tab. 1). In this study, we extend these approaches by considering piecewise linear functions for the non-stationary functions.

Ensemble members are fitted	Reference	Adjustment coefficients for the GEV parameters of ensemble members	Non-stationary functions for the GEV parameters	Extreme variable
Separately	Fowler et al. (2010)	×	Linear	Precipitation
	Hanel and Buishand (2011)	×	Linear	Precipitation
	Kharin et al. (2013)	×	Linear	Temperature & Precipitation
	Brown et al. (2014)	✓	Linear	Temperature & Rainfall
	Um et al. (2017)	×	Nonlinear	Precipitation
	Tramblay and Somot (2018)	×	Linear	Precipitation
Together	Kharin and Zwiers (2004)	×*	Linear	Temperature & Precipitation
	Wang et al. (2004)	×*	Linear	Significant wave height
	Aalbers et al. (2018)	×*	Linear	Precipitation
	Fix et al. (2018)	×*	Linear	Precipitation
	Wehner (2020)	×	Linear	Temperature & Precipitation
	Our approach	✓	Piecewise linear	Snow Load

Table 1. Temporal non-stationary GEV based approaches for GCM ensembles and GCM-RCM ensembles. The symbol "*" means that the
ensemble is an initial condition ensemble, i.e. each ensemble member consists of the same GCM-RCM pair with different initialisation.

Our application focuses on snow loads at 1500 m elevation in the 23 massifs of the French Alps. Snow load corresponds to
70 the pressure exerted by accumulated snow on the ground (proportional to the snow water equivalent) which is of major interest
both for water resource management (Marty et al., 2017) and for the structural design of buildings (Croce et al., 2018).

Section 2 presents our data, i.e. the 20 GCM-RCM pairs for RCP8.5 adjusted from EURO-CORDEX, and the S2M reanalysis
set as the reference observational dataset (Vernay et al., 2019, 2021). In Section 3, we detail our statistical methodology. Finally,
results, discussions and conclusions are introduced in Sects. 4, 5 and 6 respectively.

2 Data

75

Our application relies on the S2M reanalysis (1959-2019) and GCM-RCM pairs (1951-2100) statistically adjusted at 1500 m
for the 23 massifs in the French Alps, i.e. between Lake Geneva to the north and the Mediterranean Sea to the south (Fig. 1).
This region, home to the largest ski areas in the world, is prone to snow-related hazards such as avalanches (Favier et al., 2016;
Dkengne Sielenou et al., 2021) which are heavily impacted by ongoing warming (Eckert et al., 2013; Castebrunet et al., 2014).

80 This study estimates potential changes in snow load hazard for a high emission scenario (RCP8.5) as a case study, although it
could also be applied to other scenarios and variables. Following the block maxima approach to estimate the hazard of snow
load (Sect. 3.1), we compute annual maxima of daily snow load centred on the winter season, e.g. an annual maximum for
the year 1959 is the maximum from the 1st of August 1958 to the 31st of July 1959 (Fig. 1). We remind the reader that the
snow load is equal to the gravitational acceleration (g = 9.81 m s^{-2}) times the snow water equivalent (in kg m^{-2}). Snow water

85 equivalent equals the observed height of accumulated snow (in m) multiplied by the snow density (in kg m^{-3}).
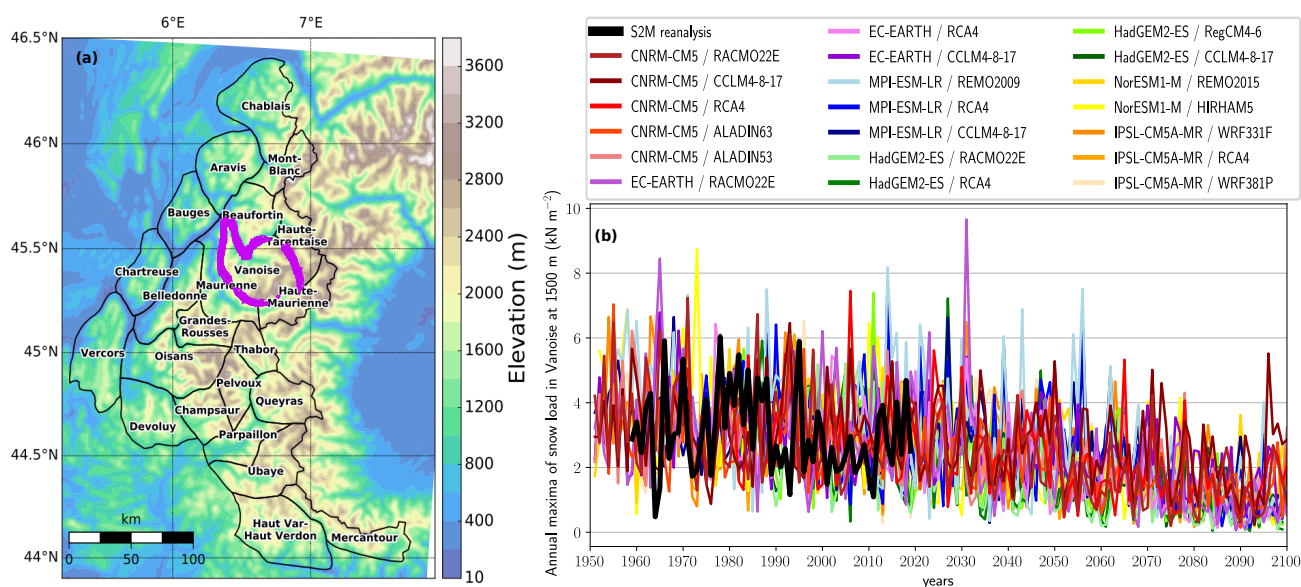
Figure 1. (a) Topography and delineation for the 23 massifs of the French Alps, e.g. the Vanoise massif corresponds to the purple region
(Durand et al., 2009). **(b)** Time series of annual maxima of daily snow load from 1951 to 2100 for the Vanoise massif at 1500 m elevation.
Annual maxima from the S2M reanalysis (1959-2019) are displayed in black, while annual maxima from the 20 adjusted GCM-RCM pairs
(1951-2100) under an historical and a high emission scenario (RCP8.5) are displayed with brighter colors.

The S2M reanalysis (Durand et al., 2009; Vernay et al., 2019, 2021) combines large scale reanalyses and forecasts with
in situ meteorological observations to provide daily values of snow load from 1959 to 2019. The S2M reanalysis has been
both evaluated with in situ temperature and precipitation observations (Durand et al., 2009) and with various snow depth
observations (Vionnet et al., 2016; Quéno et al., 2016; Revuelto et al., 2018; Vionnet et al., 2019; Vernay et al., 2021). The

90 S2M reanalysis focuses on the elevation dependency of meteorological conditions. Indeed, this reanalysis is not produced on a regular grid, but provides data for each massif every 300 m of elevation between 600 m and 3600 m.

Quantile mapping method ADAMONT (Verfaillie et al., 2017) was used to adjust the EURO-CORDEX dataset (Jacob et al., 2014) against the S2M reanalysis to provide daily values of snow load that spans historical (1951–2005) and future (2006–2100) time periods. Specifically, the EURO-CORDEX dataset consists of RCMs forced over Europe by GCMs from the CMIP5

95 ensemble (Taylor et al., 2012) for the historical and several representative concentration pathways (RCP) scenarios (Moss et al., 2010). We focus on the RCP8.5 emission scenario, and consider a total of 20 GCM-RCM pairs, with 6 GCMs and 11 RCMs (see Supplement, Tab. S1). Finally, every 300 m of elevation for each massif, adjusted EURO-CORDEX meteorological data are used as input to Crocus to provide estimates of the time evolution of the snow cover (Verfaillie et al., 2018), enabling us to compute the maximum annual value of snow load. For simplicity, we often refer to S2M reanalysis as our observation

100 reference. We note that we discard the two most southern massifs because many projected annual maxima are equal to zero.

The anomaly of global mean surface temperature (GMST) w.r.t. the pre-industrial period (1850-1900) is chosen as the temporal covariate for our statistical methodology. In practice, we smooth this anomaly with cubic splines to obtain a covariate that does not depend on the internal variability of GMST (Fig. 2). For each GCM-RCM pair we rely on the GMST corresponding GCM as covariate, while we rely on GMST from HadCRUT5 (Morice et al., 2021) as covariate for the observations. For

105 simplicity, we refer to +1 degree of smoothed anomaly of GMST as +1 degree of global warming.
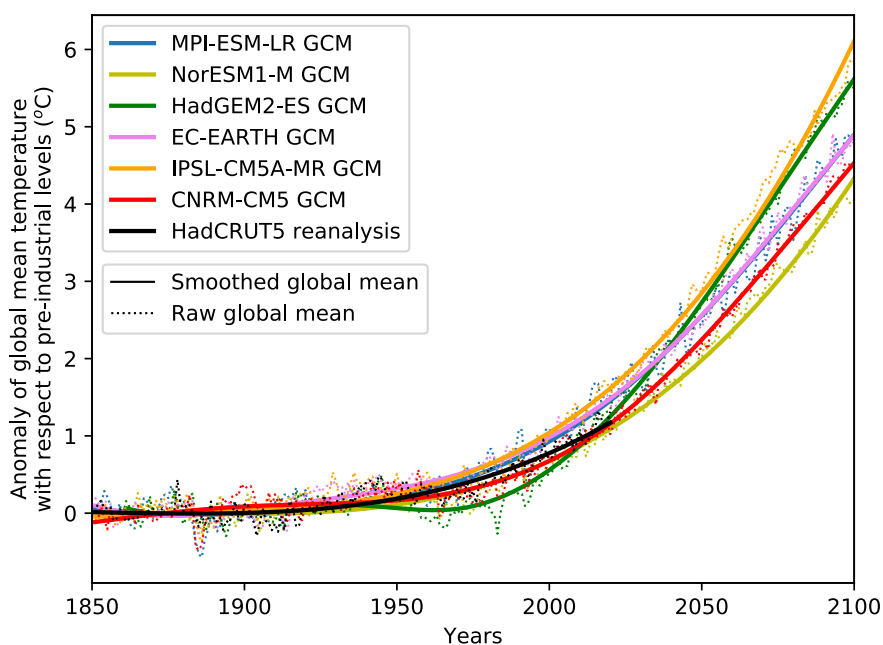
Figure 2. Raw output (dotted lines) and smoothed output (plain lines) for the anomaly of global mean annual temperature with respect to industrial levels (1850-1900). For the 6 GCMs, we show the anomaly of global mean surface temperature using historical emissions until 2005, and projected emissions (RCP8.5). Years correspond to the year centered on winter.

3 Statistical methodology

3.1 Generalized extreme value distribution

Following the block maxima approach of extreme value theory (Coles, 2001), we model annual maxima with the GEV distribution. Indeed theoretically, as the central limit theorem motivates asymptotically sample means modeling with the normal
110 distribution, the Fisher–Tippett–Gnedenko theorem (Fisher and Tippett, 1928; Gnedenko, 1943) encourages asymptotically sample maxima modeling with the GEV distribution. In practice, if Y represents an annual maximum, we can assume that $Y \sim \text{GEV}(\mu, \sigma, \xi)$, which implies that:

$$P(Y \leq y) = \begin{cases} \exp\left[-\left(1 + \xi \frac{y-\mu}{\sigma}\right)_+^{-\frac{1}{\xi}}\right] \text{ if } \xi \neq 0 \text{ and where } u_+ \text{ denotes } \max(u, 0), \\ \exp\left[-\exp\left(-\frac{y-\mu}{\sigma}\right)\right] \text{ if } \xi = 0, \end{cases} \quad (1)$$

where the three parameters are: the location μ, the scale $\sigma > 0$, and the shape ξ. Three subfamilies of distribution (reversed
115 Weibull, Gumbel, Fréchet) can be derived depending on the sign of the shape parameter ($\xi < 0$, $\xi = 0$, $\xi > 0$) respectively.

Due to these theoretical justifications, the GEV distribution enables a robust estimation of return levels (Coles, 2001; Cooley, 2012). The T-year return level, which is defined as a daily value y_p exceeded each year with probability $p = \frac{1}{T}$, corresponds to the $1-p$ quantile of the GEV distribution $P(Y \leq y_p) = 1 - p \leftrightarrow y_p = \mu - \frac{\sigma}{\xi}[1 - (-\log(1-p))^{-\xi}]$. In this study, we set $p = \frac{1}{50}$ as it corresponds to the 50-year return period that is widely used for the design working life of building (Cabrera et al.,
120 2012) notably for the building standard against snow load (Croce et al., 2019).

3.2 Non-stationary distribution

Let Y_t^{obs} denote an observed annual maximum for the year t between 1959 and 2019, and T_t^{obs} represent the smoothed anomaly of global mean surface temperature (GMST) from HadCRUT5 (Sect. 2). We rely on a non-stationary distribution where each GEV parameter is a piecewise linear function of T, the smoothed anomaly of GMST. We note that we rely on a log link
125 function for the scale parameter to ease the numerical optimization.

$$Y_t^{\text{obs}} | \boldsymbol{\theta} \sim \text{GEV}(\mu(T_t^{\text{obs}}), \sigma(T_t^{\text{obs}}), \xi(T_t^{\text{obs}})) \quad \text{with} \quad \begin{aligned} \mu(T) &= \mu_0 + \sum_{i=1}^{L} \mu_i \times (T - \kappa_i)_+, \\ \log \sigma(T) &= \sigma_0 + \sum_{i=1}^{L} \sigma_i \times (T - \kappa_i)_+, \\ \xi(T) &= \xi_0 + \sum_{i=1}^{L} \xi_i \times (T - \kappa_i)_+, \end{aligned} \quad (2)$$

where $\boldsymbol{\theta}$ is a vector of coefficients for the piecewise-linear functions $\mu(.), \sigma(.), \xi(.)$, $1 \leq L \leq 4$ corresponds to the number of linear pieces $\kappa_i = T_{\min} + \frac{(i-1) \times (T_{\max} - T_{\min})}{L}$, and $T_{\min}$ and $T_{\max}$ are the minimum and maximum smoothed anomaly of GMST for the period 1951-2100 (Fig. 2). In other words, $\kappa_2, ..., \kappa_{L-1}$ are fixed and equally spaced between $T_{\min}$ and $T_{\max}$ and correspond
130 to the $L-1$ anomalies of smoothed GMST where the line breaks, i.e. where the slope of the piecewise linear functions changes.

Let $Y_t^{\text{pair k}}$ represent an annual maximum of the GCM-RCM pair k for the year t between 1951 and 2100, and $T_t^{\text{pair k}}$ represent the smoothed anomaly of global mean temperature (Sect. 2) for the GCM of GCM-RCM pair k (with k between 1 and 20).

$$Y_t^{\text{pair k}}|\boldsymbol{\theta} \sim \text{GEV}(\mu(T_t^{\text{pair k}}, \text{GCM}^{\text{pair k}}, \text{RCM}^{\text{pair k}}), \sigma(T_t^{\text{pair k}}, \text{GCM}^{\text{pair k}}, \text{RCM}^{\text{pair k}}), \xi(T_t^{\text{pair k}})), \tag{3}$$

where the five parameterizations for $\mu(T_t^{\text{pair k}}, \text{GCM}^{\text{pair k}}, \text{RCM}^{\text{pair k}})$ and $\sigma(T_t^{\text{pair k}}, \text{GCM}^{\text{pair k}}, \text{RCM}^{\text{pair k}})$ are shown in Table 2,
135 and where $\xi(.)$ is given in Eq. 2. For the 20 GCM-RCM pairs, we consider adjustment coefficients for the location and scale parameters that aim at adjusting the distribution of GCM-RCM pairs w.r.t. the distribution of the past observations. We did not consider adjustment coefficients on the shape parameter because it sometimes leads to prediction failures. We consider five parameterizations: zero adjustment coefficients, one adjustment coefficient for all GCM-RCM pairs, one for each GCM, one for each RCM, and one for each GCM-RCM pair (Tab. 2). Following Brown et al. (2014), we assume that these adjustment
140 coefficients are constant, i.e. the same for historical and future climates.

In Table 2, we show the size of the vector of coefficients $\boldsymbol{\theta}$ for the non-stationary GEV distribution. For each parameterization, $\boldsymbol{\theta}$ contains three coefficients for the intercepts (μ_0, σ_0, ξ_0), the linear pieces coefficients, and the adjustment coefficients.

Parameterization of the adjustment coefficients	GEV parameters for the ensemble members		Size of the vector of coefficients $\boldsymbol{\theta}$
	$\mu(T, \text{GCM}_i, \text{RCM}_j)$	$\log \sigma(T, \text{GCM}_i, \text{RCM}_j)$	
Zero	$\mu(T)$	$\log \sigma(T)$	$3 + 3 \times L$
One for all GCM-RCM pairs	$\mu(T) + \mu_{\text{all}}$	$\log \sigma(T) + \sigma_{\text{all}}$	$3 + 3 \times L + 2$
One for each GCM	$\mu(T) + \mu_{\text{GCM}_i}$	$\log \sigma(T) + \sigma_{\text{GCM}_i}$	$3 + 3 \times L + 2 \times \#\text{GCMs}$
One for each RCM	$\mu(T) + \mu_{\text{RCM}_j}$	$\log \sigma(T) + \sigma_{\text{RCM}_j}$	$3 + 3 \times L + 2 \times \#\text{RCMs}$
One for each GCM-RCM pair	$\mu(T) + \mu_{\text{GCM}_i, \text{RCM}_j}$	$\log \sigma(T) + \sigma_{\text{GCM}_i, \text{RCM}_j}$	$3 + 3 \times L + 2 \times \#\text{GCM-RCM pairs}$

Table 2. The five parameterizations of the adjustment coefficients considered for the location and scale parameters of ensemble members. For each parameterization, we detail the non-stationary functions. L represents the number of linear pieces. "#" means "number of".

3.3 Maximum likelihood estimation

For each massif, a temporal non-stationary GEV distribution, parameterized by a vector of coefficients $\boldsymbol{\theta}$, is estimated using the
145 past observations and all GCM-RCM pairs. Let $\boldsymbol{y} = (y_{1959}^{\text{obs}}, ..., y_{2019}^{\text{obs}}, y_{1951}^{\text{pair 1}}, ..., y_{2100}^{\text{pair 1}}, ..., y_{1951}^{\text{pair 20}}, ..., y_{2100}^{\text{pair 20}})$ represent a vector with all annual maxima of a given massif, i.e. annual maxima from the observations and from the 20 GCM-RCM pairs (Sect. 2). The maximum likelihood method makes it possible to estimate the most likely coefficients $\widehat{\boldsymbol{\theta}}$ with $\boldsymbol{y}$. We obtain the maximum likelihood estimator $\widehat{\boldsymbol{\theta}}$ together from the past observations and all GCM-RCM pairs by maximizing the likelihood $p(\boldsymbol{y}|\boldsymbol{\theta})$:

$$\widehat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\arg\max}\, p(\boldsymbol{y}|\boldsymbol{\theta}) = \underbrace{\prod_{t=1959}^{2019} p(y_t^{\text{obs}}|\boldsymbol{\theta})}_{\text{past observations}} \times \underbrace{\prod_{t=1951}^{2100} \prod_{k=1}^{20} p(y_t^{\text{pair k}}|\boldsymbol{\theta})}_{\text{20 GCM-RCM pairs}}, \quad \text{where} \quad \begin{array}{l} p(y_t^{\text{obs}}|\boldsymbol{\theta}) = \frac{\partial P(Y_t^{\text{obs}} \leq y_t^{\text{obs}}|\boldsymbol{\theta})}{\partial y_t^{\text{obs}}}, \\[2ex] p(y_t^{\text{pair k}}|\boldsymbol{\theta}) = \frac{\partial P(Y_t^{\text{pair k}} \leq y_t^{\text{pair k}}|\boldsymbol{\theta})}{\partial y_t^{\text{pair k}}}. \end{array} \tag{4}$$

3.4 Evaluation experiments

Our first evaluation experiment is a model-as-truth experiment, a.k.a. perfect model experiment, which evaluates long-term predictive performances using future projections (Abramowitz et al., 2019). The observations from the S2M reanalysis (Sect. 2) are discarded for this experiment. Instead, one GCM-RCM pair is chosen as pseudo-observations for the calibration of the non-stationary GEV distribution. The calibration set contains the "historical" data (1959-2019) of the GCM-RCM pair chosen as pseudo-observations, and the 19 remaining GCM-RCM pairs (1951-2100). The predictive performance is evaluated on an evaluation set that contains the future data (2020-2100) of the GCM-RCM pair chosen as pseudo-observations. In detail, each GCM-RCM pair is successively regarded as being pseudo-observations. Thus, a model-as-truth experiment can be roughly regarded as a leave-one-out cross-validation w.r.t. to GCM-RCM pairs. We note that for GCM-RCM pairs with the GCM HadGEM2-ES starts in 1982, while the pairs with the RCM RCA4 starts in 1971. Therefore, we successively regard as pseudo-observations the 12 GCM-RCM pairs (out of 20) that start before 1959, i.e. that have annual maxima for the period 1959-2019.

Our second evaluation experiment is a split-sample experiment, a.k.a. calibration–validation experiment, which enables us to estimate the short-term predictive performance of each parameterization. Specifically, for the calibration of the non-stationary GEV distribution, we rely on the oldest observations from the S2M reanalysis (Sect. 2) and all the GCM-RCM pairs. We validate the predictive performance on the most recent observations. For instance, if we choose to keep 80% of the observations for the calibration (1959-2007), then the remaining 20% of the observations are held-out for the evaluation (2008-2019).

In these two evaluation experiments for GCM-RCM ensembles, we calculate the mean logarithmic score ($\overline{\text{LS}}$) on the evaluation set, the lower the better, to assess the out-of-sample skill of a non-stationary distribution parameterized with $\boldsymbol{\theta}$. For N held-out observations $y_{\text{year}_1}^{\text{obs}}, ..., y_{\text{year}_N}^{\text{obs}}$, we have that $\overline{\text{LS}} = \frac{1}{N} \sum_{n=1}^{N} -\log[p(y_{\text{year}_n}^{\text{obs}} | \boldsymbol{\theta}))]$ where $p(y_t^{\text{obs}} | \boldsymbol{\theta}) = \frac{\partial P(Y_t^{\text{obs}} \leq y_t^{\text{obs}} | \boldsymbol{\theta})}{\partial y_t^{\text{obs}}}$.

3.5 Workflow

First, for a set of past and projected annual maxima, we select one parameterization of the GEV distribution (number of linear pieces, parameterization of the adjustment coefficients) using a two-step selection method: i) we select the number of linear pieces with a model-as-truth experiment using zero adjustment coefficients for the GEV parameters ii) we select the parameterization of the adjustment coefficients with a split-sample experiment using the number of linear pieces selected in the model-as-truth experiment. Then, we study trends in the 50-year return level of snow load. For each massif we rely on the parameterization of the GEV distribution selected using the two-step selection method. We report RL50, the 50-year return level that corresponds to Eq. 2, i.e. to the 50-year return level of the observations and their adjusted projections w.r.t. GCM-RCM pairs. In other words, if the selected parameterization has adjustment coefficients, RL50 is computed without these adjustment coefficients since using these coefficients would provide the 50-year return level of the GCM-RCM pairs. The 90% uncertainty interval is computed using a semi-parametric bootstrap resampling method adapted to non-stationary extreme value distributions (Appendix A). For every anomaly of global mean temperature T, we have that the 50-return level RL50(T) is:

$$\text{RL50}(T) = y_{\frac{1}{50}}(T) = \mu(T) - \frac{\sigma(T)}{\xi(T)}[1 - (-\log{(1 - \frac{1}{50})})^{-\xi(T)}]. \tag{5}$$

4 Results

4.1 Selection of one parameterization of the GEV distribution for each massif

In Figure 3a, for each massif, we illustrate the selected parameterization of the GEV distribution (number of linear pieces,
185 parameterization of the adjustment coefficients). Next, we detail the design of our two-step selection method.
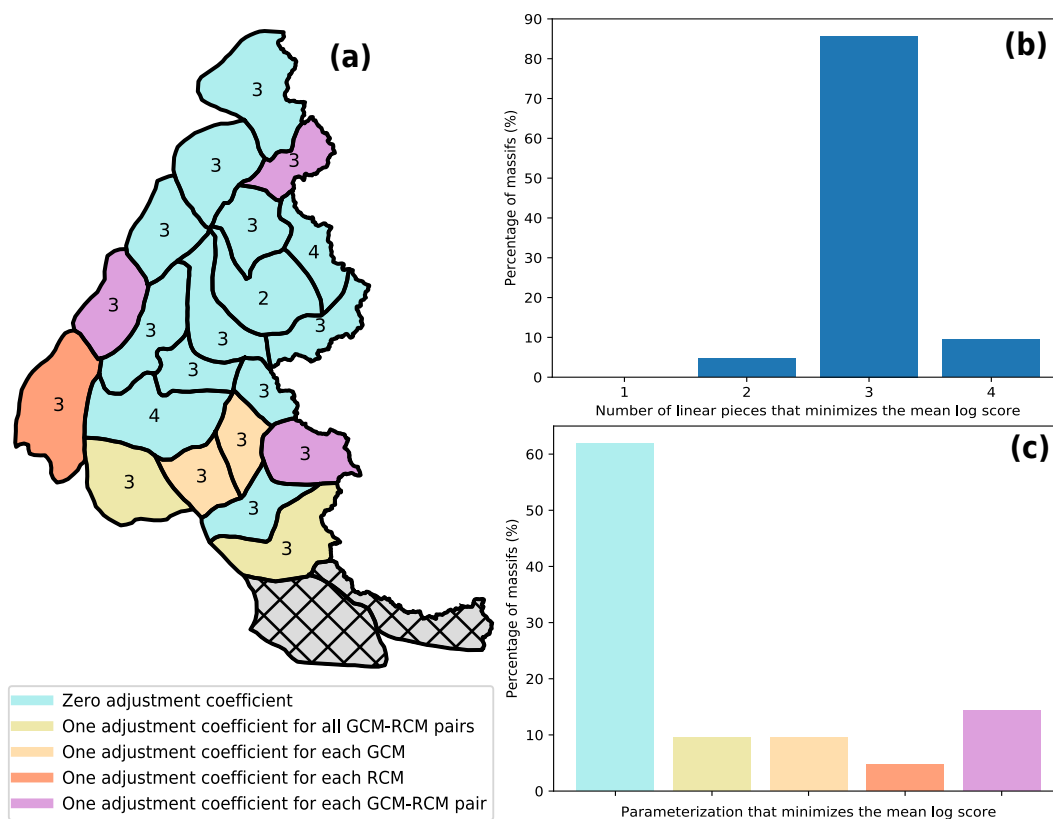
Figure 3. (a) Map of the selected parameterization of the GEV distribution. The selected parameterizations of the adjustment coefficients are illustrated with colors, while the selected numbers of linear pieces are written on the map. **(b)** Distribution of the selected number of linear pieces. **(c)** Distribution of the selected parameterization of the adjustment coefficients.

In the first step, we select the number of linear pieces that minimizes the mean logarithmic score of a model-as-truth experiment using zero adjustment coefficients for the GEV parameters. The mean logarithmic score is averaged on the held-out pseudo-observations (2020-2100) for each of the 12 GCM-RCM pairs (which are set as pseudo-observations, see Sect. 4.1). We find that the parameterization with three linear pieces minimizes the mean logarithmic score for 80% of the massifs, see
190 Fig. 3b. The parameterization with two linear pieces is selected for one massif, and the one with four linear pieces is selected for two massifs. Thus, at least two linear pieces are selected for the piecewise linear functions.

In the second step, we select the parameterization of the adjustment coefficients (Tab. 2) that minimizes the mean logarithmic score for a split-sample experiment using the number of linear pieces selected in the model-as-truth experiment. The mean logarithmic score is averaged on the evaluation observations for three split-sample experiments, where the calibration set corresponds to $60\%, 70\%$, and 80% of the observations. Indeed, we observe that the split-sample experiment is quite sensitive to the size of the calibration set. Thus, we choose to average the mean logarithmic score on three split-sample experiments to obtain more robust results. We find that the parameterization with zero adjustment coefficients minimizes the mean logarithmic score for two thirds of the massifs, see Fig. 3c. Otherwise, the parameterization with one adjustment coefficient for all GCM-RCM pairs is selected for two massifs, the parameterization with one adjustment coefficient for each GCM is selected for two massifs, the parameterization with one adjustment coefficient for each RCM is selected for one massif, and the parameterization with one adjustment coefficient for each GCM-RCM pair is selected for three massifs. Thus, for two thirds of the massifs, adjustment coefficients do not lead to a better predictive performance on the validation periods. This is presumably due to the fact that GCM-RCM pairs have already been statistically adjusted.

For a detailed analysis of the mean logarithmic scores of each parameterization for each massif, see Supplement, Part C.

4.2 Trends in the 50-year return level of snow load

In this section, for each massif we rely on the parameterization of the GEV distribution selected in Sect. 4.1.

In Figure 4, we illustrate changes in the 50-year return level between +1 and +4 degrees of global warming for four massifs where the selected parameterization is composed of three linear pieces with one adjustment coefficient for all GCM-RCM pairs (Fig. 4a), one coefficient for each GCM (Fig. 4b), one coefficient for each RCM (Fig. 4c), or one coefficient for each RCM-RCM pair (Fig. 4d). All 50-year return levels (for the non-stationary GEV distribution fitted on the observations, on each GCM-RCM pair, and on the observations and all GCM-RCM pairs) are decreasing with the anomaly of global mean temperature. We observe that RL50 with adjustment coefficients (shown in a warm color) is closer to the 50-year return level of the observation (in dark grey) than RL50 without adjustment coefficients (in cyan). This figure also shows how adjustment coefficients adjust the distribution toward the distribution of the observations by illustrating the probability density functions (with and without adjustment) at +1 degree of global warming. Nevertheless, we note that adjusted distributions do not perfectly match the distributions of the observation, which entails that the adjusted RL50 do not match the 50-year return levels of the observation. This is probably because we do not consider adjustment coefficients on the shape parameter. For instance, in Figure 4c, we observe that the shape parameter is negative for the distribution of the observation (because the density has an upper bound), while the shape parameter is positive for the adjusted distribution in orange. We choose to not consider adjustment adjustment for the shape parameter because it enables us to constrain predictive distributions on the future period, and to avoid prediction failures (Sect. 5.2). Besides, the 90% uncertainty interval of RL50 is computed using a semi-parametric bootstrap resampling method adapted to non-stationary extreme distributions (Appendix A). We note that uncertainty intervals are widening at the nodes of the piecewise linear functions, i.e. at the anomaly of global temperature where the slope of the GEV parameters changes (κ_i in Eq. 2). This is presumably due to the fact that the variability of the three GEV parameters is more important at the nodes than between them.
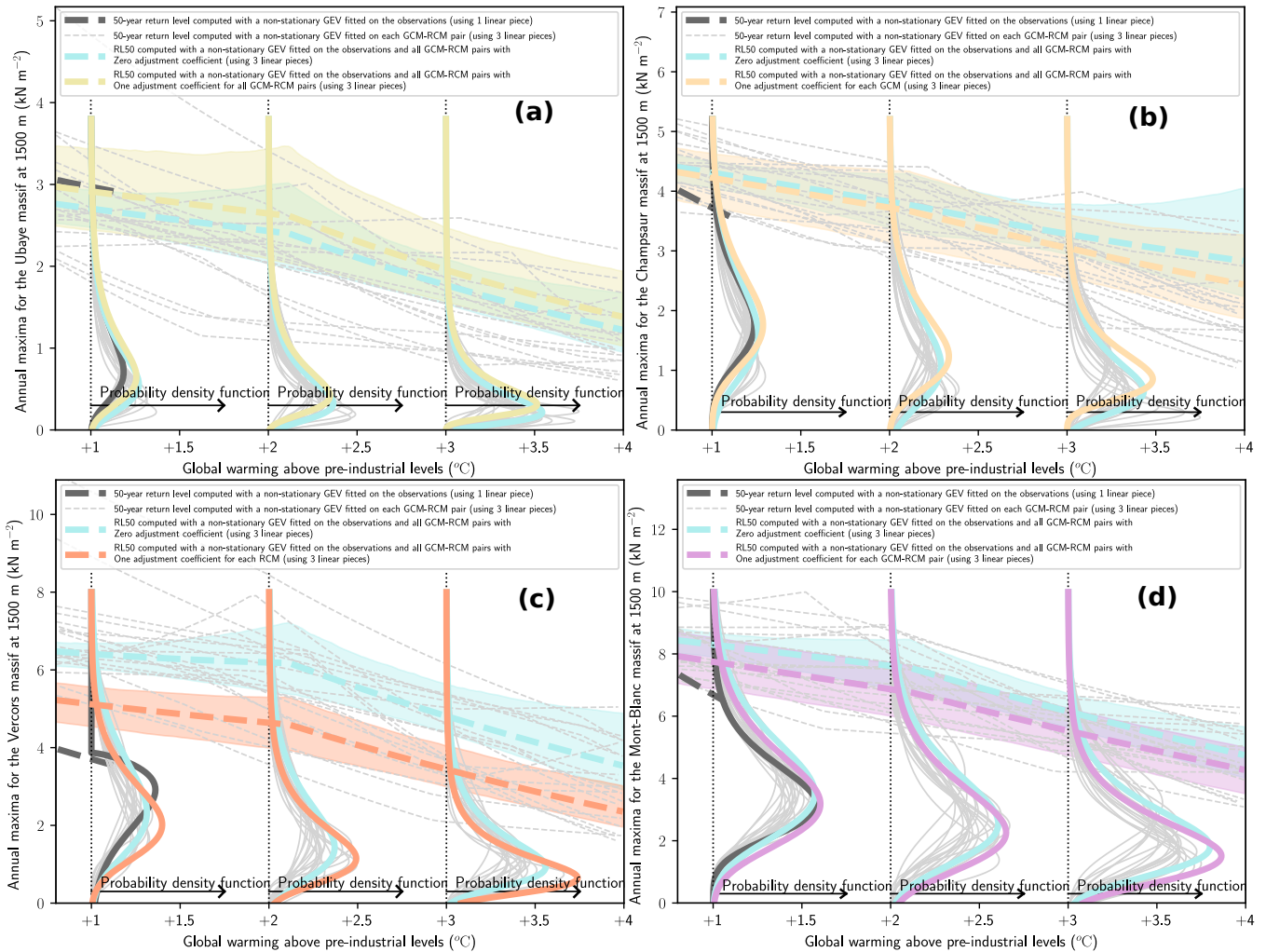
Figure 4. Estimated 50-year return levels between +1 and +4 degrees of global warming at elevation 1500 m under RCP8.5 for four massifs with adjustment coefficients: **(a)** one coefficient for all GCM-RCM pairs **(b)** one coefficient for each GCM, **(c)** one coefficient for each RCM, and **(d)** one coefficient for each RCM-RCM pair. RL50 (Eq. 5) without adjustment coefficients are shown in cyan, and with adjustment coefficients in a warm color. 90% uncertainty intervals are shaded. The 50-year return levels computed for each GCM-RCM pair (using for each GCM-RCM pair a non-stationary GEV distribution with the selected number of linear pieces) and for the observation (using a non-stationary GEV distribution with one linear piece and a constant shape parameter) are displayed with thin gray lines and thick dark lines, respectively. The probability density functions at +1, +2 and +3 degrees exemplify how adjustment coefficients can adjust the distribution.

Figure 5 illustrates RL50 for the 23 massifs of the French Alps at 1500 m elevation for +1, +2, +3, and +4 degrees of global warming, i.e. of smoothed anomaly of global mean surface temperature. The return levels are larger in the northwest of the French Alps, and this pattern persists with global warming. Over the whole French Alps, the average RL50 equals 5.7 kN m^{-2} at +1 degrees of global warming, and 3.3 kN m^{-2} at +4 degrees.
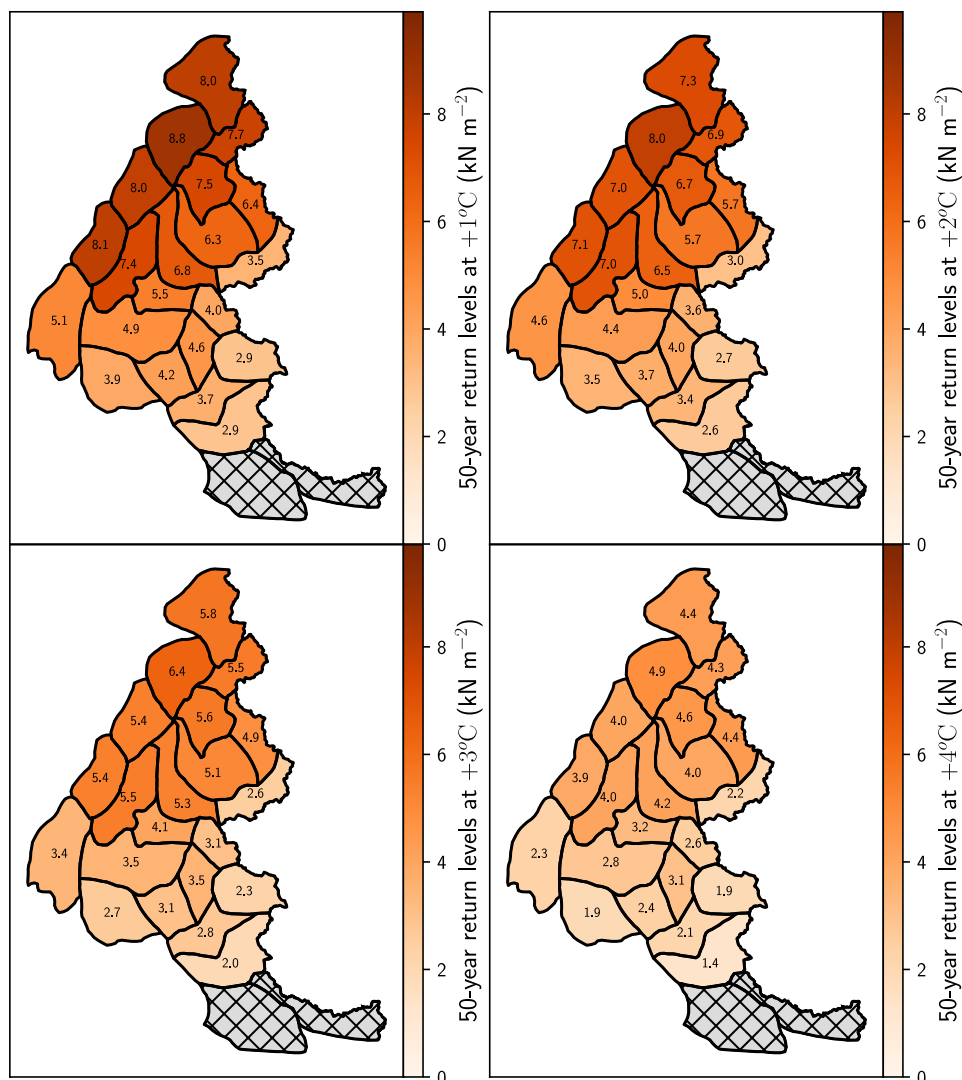
Figure 5. 50-year return levels (RL50) of snow load at 1500 m for +1, +2, +3, and +4 degrees of global warming under RCP8.5.

Figure 6 details the relative change of RL50 for +2, +3, and +4 degrees of global warming at 1500 m elevation w.r.t. +1 degrees, which corresponds roughly to the current level of global warming above industrial levels (see Fig. 2). Over the French Alps, the average change of RL50 is equal to -0.6 kN m^{-2} (-10%), -1.5 kN m^{-2} (-27%), -2.5 kN m^{-2} (-43%) for +2, +3, and +4 degrees of global warming, respectively. These relative changes are different for other elevations, a smaller relative decrease being obtained at 2100 m of elevation, and a larger relative decrease at 900 m of elevation (see Supplement, Part B). This result is consistent with the literature (Fig. 2.3 of IPCC 2019). At 1500 m, the relative decrease is less important in the center east side of the French Alps. For instance, for +4 degrees of global warming, the relative decrease roughly ranges between -33% and -38% in the center east side, while it ranges between -40% and -54% in the rest of the French Alps.
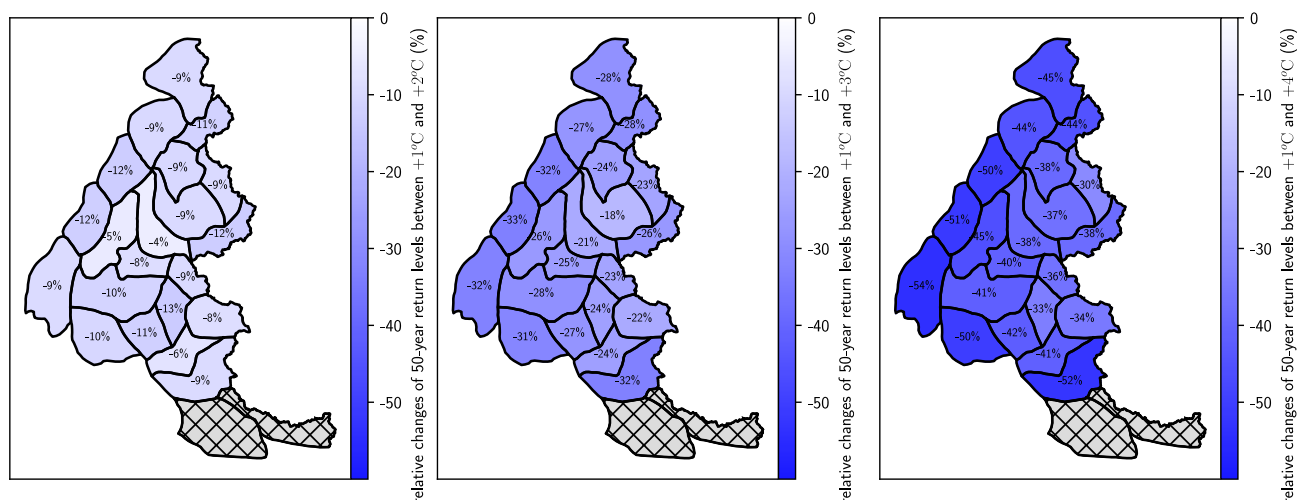
Figure 6. Relative changes in 50-year return levels (RL50) of snow load at 1500 m for +2, +3, and +4 degrees of global warming under the scenario RCP8.5 w.r.t +1 degrees of global warming.

For each massif, it is also possible to compute the average 50-year return level for several time slices: 1986-2005, 2031-2050, and 2080-2099. For instance, for the time slice 1986-2005, the average return level equals the average of the return level found for the years 1986, 1987, ..., 2005. In order to compute the return level of a given year, e.g. 1986, we rely on the relationship between the anomaly of global mean surface temperature (GMST) and the years (Fig. 2). Specifically, we rely on the anomaly of GMST averaged on the six GCMs to compute this relationship. Following this method, we find that on average the 50-year return level is projected to decrease by -0.8 kN m^{-2} (-14%) between 1986-2005 and 2031-2050 and by -2.9 kN m^{-2} (-50%) between 1986-2005 and 2080-2099 under the scenario RCP8.5. Note that this method could also provide the rate of change of other RCPs for various lead times, using their corresponding global temperature values.

5 Discussion

5.1 Comparison of our results with the projected trends at the scale of the European Alps

In Table 3, we compare our results with the Fig. 2.3 of IPCC (2019) that provides the trends in winter mean snow water equivalent (SWE) at the scale of the European Alps between 1000 m and 2000 m under the scenario RCP8.5. As detailed in Sect. 2, the snow load is proportional to the SWE, as it is equal to the SWE times the gravitational acceleration (g = 9.81 m s^{-2}). For the 23 massifs, the average return level for several time slices 1986-2005, 2031-2050, 2080-2099 can be obtained as explained in Sect. 4.2. Likewise, with a similar methodology, the mean annual maxima can be expressed as the expectation of the non-stationary GEV distribution averaged for each year of the time slices. We find a decrease of mean annual maxima of snow load by -30% and -69% for the future periods 2031-2050 and 2080-2099 compared to the reference period 1986-2005.

Source	Variable	Indicator	Location	Reference period	Future period	Trend
IPCC (2019)	SWE	Mean (Dec to May)	European Alps, 1000-1500 m	1986-2005	2031-2050	≈ -35%
					2080-2099	≈ -75%
			European Alps, 1500-2000 m		2031-2050	≈ -25%
					2080-2099	≈ -70%
Our results	Snow Load	Mean annual maxima	French Alps, 1500 m	1986-2005	2031-2050	-30%
					2080-2099	-69%
		50-year return level	French Alps, 1500 m	1986-2005	2031-2050	-14%
					2080-2099	-50%

Table 3. Projected trends in snow water equivalent (SWE), and snow load under the scenario RCP8.5 using the EURO-CORDEX experiment. In the first four rows of the Table, we specify that the result is approximated because the trend was read from the Figure 2.3. of IPCC (2019).

255   Figure 2.3 of IPCC (2019) relies on the raw (without adjustment) EURO-CORDEX data. They also find decreasing trends. For instance, between 1500 m and 2000 m of elevation, the mean winter SWE (proportional to the mean winter snow load) is expected to approximately decrease by -25% and -70% for the periods 2031-2050 and 2080-2099, respectively (Tab. 3). We observe that our mean annual maxima of snow load has a decreasing rate comparable to the decreasing rate of the mean value of snow load. These comparable rates may stem from the fact that i) both approaches rely (directly or indirectly) on
260   the EURO-CORDEX data, ii) the annual maxima of snow load results from an accumulation during the winter (Dec to May), which implies that we can expect that the mean value will roughly decrease with the same rate as the mean annual maxima.

5.2   Methodological choices, assumptions and limitations

For the non-stationarity of the GEV parameters, we choose piecewise linear functions because they can approximate more complex functions with few parameters. This makes our methodology widely applicable. One limitation is that the nodes of
265   the piecewise linear functions are fixed. Yet, we are confident that these functions are well-estimated owing to the high amount of maxima: each of the 20 GCM-RCM pairs provides more than 100 maxima. Otherwise, we rely on the anomaly of global mean temperature as covariate (Sect. 2), like a majority of references cited in Table 1. Indeed, this anomaly is often thought as a good proxy to measure the level of climate change (Fix et al., 2018) which helps strengthen the global response to this threat (Masson-Delmotte et al., 2018). We choose to focus on the scenario RCP8.5 to have the broadest spectrum of potential
270   changes for the 50-year return level of snow load. Also, to obtain Eq. 4 we assume that all annual maxima are conditionally independent given the vector of parameters θ which is a classical hypothesis. Following the principle of parsimony, we assume that the adjustment coefficients are constant, i.e. the same for historical and future climates. Besides, as mentioned in Section 3.2, we did not consider adjustment coefficients for the shape parameter because it sometimes leads to prediction failure, i.e. the predictive distribution gives a null probability to some future annual maxima. This illustrates the trade-off between i) improving
275   adjustments on the historical period, i.e. that the adjusted distribution better match the distribution of the observations and ii) having assumptions that help to constrain the predictive distribution on the future period.

For the two-step selection method, we first rely on a model-as-truth experiment to select the number of linear pieces. It assesses the optimal number of linear pieces to predict annual maxima of the pseudo-observations for the evaluation set (2020-2100), i.e. to find a good trade-off between underfitting and overfitting for the calibration set. In this first step, adjustment coefficients are not considered, such that this experiment does not depend on a specific parameterization. Then, the best parameterization of the adjustment coefficients is selected with a split-sample experiment. It assesses whether applying adjustment coefficients helps to predict observations of the evaluation set, i.e. whether it is reasonable to assume that the observations do not follow the same distribution as the GCM-RCM pairs. The evaluation score is average for three split-sample experiments where the evaluation set corresponds to the last $40\%, 30\%$, and 20% of the observations (Sect. 4.1). Thus, evaluation sets of the three split-sample experiments contain 24, 17, 12 annual maxima, respectively, which is a limited amount to robustly select the best parameterization of the adjustment coefficients.

The 90% uncertainty intervals of return levels (Fig. 4) account both for the sampling uncertainty (Appendix A) and the climate model uncertainty (distributions are fitted together from the past observations and all GCM-RCM pairs). In contrast, approaches that estimate return levels separately for each ensemble member usually do not account for the sampling uncertainty, i.e. the sampling uncertainty of return levels estimated on each ensemble, even if this uncertainty can be large because return levels are estimated with only one ensemble member. One limitation of our approach is that, contrary to the climatological expectations, the width of uncertainty intervals does not increase with global warming (Fig. 4). This is presumably a consequence of assuming constant adjustment coefficients.

5.3 Related works

First, our methodology based on adjustment coefficients can be seen as an extension of Brown et al. (2014), which estimates non-stationary GEV distribution simultaneously with both observations and a single GCM-RCM pair, and introduces constant bias terms for each GEV parameter. There also exists some links with a debiasing method proposed for annual maxima from GCM-RCM projections (Fontolan et al., 2019). For the location parameter we consider additive adjustment coefficients that can be seen as bias terms, while the adjustment coefficients of the scale parameter that are multiplicative (due to the log link function) can be viewed as bias correction factors (Hosseinzadehtalaei et al., 2021). In this paper, we choose the name "adjustment coefficients" because we introduce them to improve the statistical adjustments. Our idea to add adjustment coefficients for each GCM/RCM or GCM-RCM pairs into the non-stationary extreme value distributions (Tab. 2) comes from the ANOVA framework, which can be applied to partition the uncertainty of GCM-RCM projections by identifying GCM/RCM main effects, or GCM/RCM interactions (Hawkins and Sutton, 2009; Evin et al., 2019).

Then, our approach based on piecewise linear functions for the non-stationarity of the GEV parameters can be viewed as using linear splines. In the literature, there exists many extreme value theory approaches using splines. For instance, linear splines have been applied to model the temporal non-stationarity (Wilcox et al., 2018), while cubic splines are often considered to model spatial extremes (Chavez-Demoulin and Davison, 2005; Gaume et al., 2013).

6 Conclusions and outlook

310 Following the recent trend of statistical methods that constrain climate projections using past observations (Brunner et al., 2020), we propose a novel non-stationary extreme value approach for GCM-RCM ensembles that estimates a non-stationary GEV distribution from both the past observations and all GCM-RCM pairs together. Specifically, we rely on a flexible non-stationary generalized extreme value (GEV) distribution with i) piecewise linear functions to model the changes in the three GEV parameters ii) adjustment coefficients for the location and scale parameters to adjust the GEV distributions of the GCM-

315 RCM pairs with respect to the GEV distribution of the past observations. In order to select one parameterization of the GEV distribution (number of linear pieces, parameterization of the adjustment coefficients) we design a two-step selection procedure based on two evaluation experiments for GCM-RCM ensembles: a model-as-truth experiment and a split-sample experiment.

In this article, as a case study, the proposed approach is applied to snow load in the French Alps at 1500 m of elevation, using 20 GCM-RCM pairs statistically adjusted from the EURO-CORDEX experiment under the scenario RCP8.5. More generally,

320 the proposed approach could also be applied to other scenarios, climate variables, and climate projection ensembles.

Many extensions of this work could be considered. First, our parameterization of the GEV distribution cannot account for different changes of distributions among the GCM-RCM pairs. Indeed, in the case where a parameterization of the GEV distribution without adjustment coefficients is selected, the GCM-RCM pairs are considered as independent and identically distributed. By contrast, in the case where a parameterization with adjustment coefficients is selected, GCM-RCM pairs can

325 have different distributions. However, these distributions are still constrained to have the same changes with global warming because adjustment coefficients are constant. In future works, to better account for different changes of distributions among the GCM-RCM pairs, we could imagine adjustment coefficients that vary with global warming. A second potential extension of this work could be to improve the parameterization of the GEV distribution by adding weights for each GCM-RCM pair. In our methodology, GCM-RCM pairs are currently considered as equally plausible even though it is known that for each application

330 some of them can have a better agreement with the past observations. Following the intuition of weighting schemes for climate ensemble (Knutti et al., 2017), we could design a parameterization of the GEV distribution that assigns more weights, i.e. more confidence, to climate models that agree more with the observation.

Appendix A: Uncertainty estimation

We estimate the uncertainties resulting from in-sample variability with a semi-parametric bootstrap resampling method adapted

335 to non-stationary extreme distributions (Efron and Tibshirani, 1993; Kharin and Zwiers, 2004). This method relies on a transformation $f_{\text{GEV} \rightarrow \text{Standard Gumbel}}$ to the standard Gumbel distribution. Indeed, if $Y_x \sim \text{GEV}(\mu(x), \sigma(x), \xi(x))$, then $f_{\text{GEV} \rightarrow \text{Standard Gumbel}}(Y_x) = \frac{1}{\xi(x)} \log(1 + \xi(x) \frac{Y_x - \mu(x)}{\sigma(x)}) \sim \text{Gumbel}(0, 1)$. Let $\boldsymbol{y} = (y_1, ..., y_S)$ denote a vector of annual maxima, with S the size of the vector. The transformed observations, a.k.a. residuals, are computed as $\epsilon_m = f_{\text{GEV} \rightarrow \text{Standard Gumbel}}(y_m)$, using $\widehat{\boldsymbol{\theta}}$ for $\mu(x), \sigma(x), \xi(x)$.

We generate $B = 1000$ bootstrap samples with a four steps procedure. First, we compute the residual vector $\boldsymbol{\epsilon} = (\epsilon_1, ... \epsilon_S)$.

340 Then, for each bootstrap sample i, from these residuals we draw with replacement a sample of size S denoted as $\tilde{\epsilon}_1^{(i)}, ..., \tilde{\epsilon}_S^{(i)}$. Further, we transform these bootstrapped residuals into bootstrapped annual maxima as follows: $\forall m, \tilde{y}_m^{(i)} = f_{\text{GEV} \rightarrow \text{Standard Gumbel}}^{-1}(\tilde{\epsilon}_m^{(i)})$.

Finally, we estimate the GEV parameter $\widehat{\boldsymbol{\theta}}^{(i)}$ with the bootstrapped annual maxima $\tilde{y}_1^i, ..., \tilde{y}_S^i$. To sum up, this bootstrap procedure provides a set $\{\widehat{\boldsymbol{\theta}}^{(1)}, ..., \widehat{\boldsymbol{\theta}}^{(i)}, ..., \widehat{\boldsymbol{\theta}}^{(B)}\}$ of B GEV parameters that represents the in-sample variability.

Author contributions. ELR, GE and NE designed the research. ELR performed the analysis and drafted the first version of the manuscript.
345   All authors discussed the results and edited the manuscript.

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

Data availability. The full S2M reanalysis on which this study grounds is freely avalaible on AERIS (Vernay et al., 2019). For each GCM, the global mean surface temperature can be computed from https://climexp.knmi.nl/CMIP5/Tglobal/. For the observation, the global mean surface temperature from HadCRUT5 can be downloaded from the following webpage https://crudata.uea.ac.uk/cru/data/temperature/
350   HadCRUT5.0Analysis_gl.txt.

Acknowledgements. ELR holds a PhD grant from INRAE. We are grateful to Ben Youngman for his "evgam" R package. Inrae, CNRM and IGE are members of Labex OSUG. We are indebted to Raphaëlle Samacoïts from Météo France for providing us the latest version of the climate projection data.

References

355 Aalbers, E. E., Lenderink, G., van Meijgaard, E., and van den Hurk, B. J.: Local-scale changes in mean and heavy precipitation in Western Europe, climate change or internal variability?, Climate Dynamics, 50, 4745–4766, https://doi.org/10.1007/s00382-017-3901-9, 2018.

Abramowitz, G., Herger, N., Gutmann, E., Hammerling, D., Knutti, R., Leduc, M., Lorenz, R., Pincus, R., and Schmidt, G. A.: Model dependence in multi-model climate ensembles: weighting, sub-selection and out-of-sample testing, Earth System Dynamics, 6, 1–20, https://doi.org/10.5194/esd-2018-51, 2019.

360 AghaKouchak, A., Easterling, D., Hsu, K., Schubert, S., and Sorooshian, S.: Extremes in a changing climate: detection, analysis and uncertainty, vol. 65, Springer Science & Business Media, https://doi.org/10.1007/978-94-007-4479-0, 2012.

Beniston, M., Stephenson, D. B., Christensen, O. B., Ferro, C. A., Frei, C., Goyette, S., Halsnaes, K., Holt, T., Jylhä, K., Koffi, B., Palutikof, J., Schöll, R., Semmler, T., and Woth, K.: Future extreme events in European climate: An exploration of regional climate model projections, Climatic Change, 81, 71–95, https://doi.org/10.1007/s10584-006-9226-z, 2007.

365 Brown, S. J., Murphy, J. M., Sexton, D. M., and Harris, G. R.: Climate projections of future extreme events accounting for modelling uncertainties and historical simulation biases, Climate Dynamics, 43, 2681–2705, https://doi.org/10.1007/s00382-014-2080-1, 2014.

Brunner, L., McSweeney, C., Ballinger, A. P., Befort, D. J., Benassi, M., Booth, B., Coppola, E., de Vries, H., Harris, G., Hegerl, G. C., Knutti, R., Lenderink, G., Lowe, J., Nogherotto, R., O'Reilly, C., Qasmi, S., Ribes, A., Stocchi, P., and Undorf, S.: Comparing Methods to Constrain Future European Climate Projections Using a Consistent Framework, Journal of Climate, 33, 8671–8692,
370 https://doi.org/10.1175/jcli-d-19-0953.1, 2020.

Cabrera, A. T., Heras, M. D., Cabrera, C., and Heras, A. M. D.: The Time Variable in the Calculation of Building Structures . How to extend the working life until the 100 years ?, in: 2nd International Conference on Construction and Building Research, pp. 1–6, http://oa.upm.es/22914/1/INVE_MEM_2012_152534.pdf, 2012.

Caires, S., Swail, V. R., and Wang, X. L.: Projection and analysis of extreme wave climate, Journal of Climate, 19, 5581–5605,
375 https://doi.org/10.1175/JCLI3918.1, 2006.

Castebrunet, H., Eckert, N., Giraud, G., Durand, Y., and Morin, S.: Projected changes of snow conditions and avalanche activity in a warming climate: The French Alps over the 2020-2050 and 2070-2100 periods, Cryosphere, 8, 1673–1697, https://doi.org/10.5194/tc-8-1673-2014, 2014.

Chavez-Demoulin, V. and Davison, A. C.: Generalized additive modelling of sample extremes, Journal of the Royal Statistical Society. Series
380 C: Applied Statistics, 54, 207–222, https://doi.org/10.1111/j.1467-9876.2005.00479.x, 2005.

Coles, S. G.: An introduction to Statistical Modeling of Extreme Values, vol. 208, Springer, London, https://doi.org/10.1007/978-1-4471-3675-0, 2001.

Cooley, D.: Return Periods and Return Levels Under Climate Change, in: Extremes in a Changing Climate - Detection, Analysis & Uncertainty, pp. 97–114, Springer Science & Business Media, https://doi.org/10.1007/978-94-007-4479-0, 2012.

385 Croce, P., Formichi, P., Landi, F., and Marsili, F.: Climate change: Impact on snow loads on structures, Cold Regions Science and Technology, 150, 35–50, https://doi.org/10.1016/J.COLDREGIONS.2017.10.009, https://www.sciencedirect.com/science/article/pii/S0165232X17304834, 2018.

Croce, P., Formichi, P., Landi, F., and Marsili, F.: Harmonized European ground snow load map: Analysis and comparison of national provisions, Cold Regions Science and Technology, 168, 102 875, https://doi.org/10.1016/j.coldregions.2019.102875, https://doi.org/10.
390 1016/j.coldregions.2019.102875, 2019.

Dkengne Sielenou, P., Viallon-Galinier, L., Hagenmuller, P., Naveau, P., Morin, S., Dumont, M., Verfaillie, D., and Eckert, N.: Combining random forests and class-balancing to discriminate between three classes of avalanche activity in the French Alps, Cold Regions Science and Technology, 187, 103 276, https://doi.org/10.1016/j.coldregions.2021.103276, 2021.

Durand, Y., Laternser, M., Giraud, G., Etchevers, P., Lesaffre, B., and Mérindol, L.: Reanalysis of 44 yr of climate in the French Alps (1958-2002): Methodology, model validation, climatology, and trends for air temperature and precipitation, Journal of Applied Meteorology and Climatology, 48, 429–449, https://doi.org/10.1175/2008JAMC1808.1, 2009.

Eckert, N., Parent, E., Naaim, M., and Richard, D.: Bayesian stochastic modelling for avalanche predetermination: From a general system framework to return period computations, Stochastic Environmental Research and Risk Assessment, 22, 185–206, https://doi.org/10.1007/s00477-007-0107-4, 2008.

Eckert, N., Keylock, C. J., Castebrunet, H., Lavigne, A., and Naaim, M.: Temporal trends in avalanche activity in the French Alps and subregions: From occurrences and runout altitudes to unsteady return periods, Journal of Glaciology, 59, 93–114, https://doi.org/10.3189/2013JoG12J091, 2013.

Efron, B. and Tibshirani, R. J.: An introduction to the bootstrap, Chapman and Hall, https://doi.org/10.1201/9780203217252.ch1, 1993.

Evin, G., Curt, T., and Eckert, N.: Has fire policy decreased the return period of the largest wildfire events in France? A Bayesian assessment based on extreme value theory, Natural Hazards and Earth System Sciences, 18, 2641–2651, https://doi.org/10.5194/nhess-18-2641-2018, 2018.

Evin, G., Hingray, B., Blanchet, J., Eckert, N., Morin, S., and Verfaillie, D.: Partitioning uncertainty components of an incomplete ensemble of climate projections using data augmentation, Journal of Climate, pp. 2423–2440, https://doi.org/10.1175/jcli-d-18-0606.1, 2019.

Favier, P., Eckert, N., Faug, T., Bertrand, D., and Naaim, M.: Avalanche risk evaluation and protective dam optimal design using extreme value statistics, Journal of Glaciology, 62, 725–749, https://doi.org/10.1017/jog.2016.64, 2016.

Fisher, R. A. and Tippett, L. H. C.: Limiting forms of the frequency distribution of the largest or smallest member of a sample, Mathematical Proceedings of the Cambridge Philosophical Society, 24, 180–190, https://doi.org/10.1017/S0305004100015681, 1928.

Fix, M. J., Cooley, D., Sain, S. R., and Tebaldi, C.: A comparison of U.S. precipitation extremes under RCP8.5 and RCP4.5 with an application of pattern scaling, Climatic Change, 146, 335–347, https://doi.org/10.1007/s10584-016-1656-7, 2018.

Fontolan, M., Xavier, A. C. F., Pereira, H. R., and Blain, G. C.: Using climate change models to assess the probability of weather extremes events: A local scale study based on the generalized extreme value distribution, Bragantia, 78, 146–157, https://doi.org/10.1590/1678-4499.2018144, 2019.

Fowler, H. J., Ekström, M., Blenkinsop, S., and Smith, A. P.: Estimating change in extreme European precipitation using a multimodel ensemble, Journal of Geophysical Research Atmospheres, 112, https://doi.org/10.1029/2007JD008619, 2007.

Fowler, H. J., Cooley, D., Sain, S. R., and Thurston, M.: Detecting change in UK extreme precipitation using results from the climateprediction.net BBC climate change experiment, Extremes, 13, 241–267, https://doi.org/10.1007/s10687-010-0101-y, 2010.

Gaume, J., Eckert, N., Chambon, G., Naaim, M., and Bel, L.: Mapping extreme snowfalls in the French Alps using max-stable processes, Water Resources Research, 49, 1079–1098, http://doi.wiley.com/10.1002/wrcr.20083, 2013.

Gnedenko, B.: Sur la distribution limite du terme maximum d'une série aléatoire, The Annals of Mathematics, 44, 423–453, https://doi.org/10.2307/1968974, 1943.

Hanel, M. and Buishand, T. A.: Analysis of precipitation extremes in an ensemble of transient regional climate model simulations for the Rhine basin, Climate Dynamics, 36, 1135–1153, https://doi.org/10.1007/s00382-010-0822-2, 2011.

Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in regional climate predictions, Bulletin of the American Meteorological Society, 90, 1095–1107, https://doi.org/10.1175/2009BAMS2607.1, 2009.

430   Hosseinzadehtalaei, P., Ishadi, N. K., Tabari, H., and Willems, P.: Climate change impact assessment on pluvial flooding using a distribution-based bias correction of regional climate model simulations, Journal of Hydrology, 598, 126 239, https://doi.org/10.1016/j.jhydrol.2021.126239, https://doi.org/10.1016/j.jhydrol.2021.126239, 2021.

IPCC: Special Report: The Ocean and Cryosphere in a Changing Climate, In press, https://www.ipcc.ch/report/srocc/, 2019.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., Geor-

435   gopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J. F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., and Yiou, P.: EURO-CORDEX: New high-resolution climate change projections for European impact research, Regional Environmental Change, 14, 563–578, https://doi.org/10.1007/s10113-013-0499-2, 2014.

440   Kharin, V. V. and Zwiers, F. W.: Estimating extremes in transient climate change simulations, Journal of Climate, 18, 1156–1173, https://doi.org/10.1175/JCLI3320.1, 2004.

Kharin, V. V., Zwiers, F. W., Zhang, X., and Hegerl, G. C.: Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations, Journal of Climate, 20, 1419–1444, https://doi.org/10.1175/JCLI4066.1, 2007.

Kharin, V. V., Zwiers, F. W., Zhang, X., and Wehner, M.: Changes in temperature and precipitation extremes in the CMIP5 ensemble, Climatic

445   Change, 119, 345–357, https://doi.org/10.1007/s10584-013-0705-8, 2013.

Knutti, R., Sedláček, J., Sanderson, B. M., Lorenz, R., Fischer, E. M., and Eyring, V.: A climate model projection weighting scheme accounting for performance and interdependence, Geophysical Research Letters, 44, 1909–1918, https://doi.org/10.1002/2016GL072012, 2017.

Kyselý, J., Picek, J., and Beranová, R.: Estimating extremes in climate change simulations using the peaks-over-threshold method with a

450   non-stationary threshold, Global and Planetary Change, 72, 55–68, https://doi.org/10.1016/j.gloplacha.2010.03.006, 2010.

Le Roux, E., Evin, G., Eckert, N., Blanchet, J., and Morin, S.: Non-stationary extreme value analysis of ground snow loads in the French Alps: a comparison with building standards, Natural Hazards and Earth System Sciences, 20, 2961–2977, https://doi.org/10.5194/nhess-2020-81, 2020.

Marty, C., Tilg, A.-M., Jonas, T., Marty, C., Tilg, A.-M., and Jonas, T.: Recent Evidence of Large-Scale Receding Snow Water Equivalents

455   in the European Alps, Journal of Hydrometeorology, 18, 1021–1031, https://doi.org/10.1175/JHM-D-16-0188.1, 2017.

Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D., Skea, J., Shukla, P. R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors, S., Matthews, J. B. R., Chen, Y., Zhou, X., Gomis, M. I., Lonnoy, E., Maycock, T., Tignor, M., and Waterfield, T.: An IPCC Special Report on the impacts of global warming of 1.5, https://doi.org/10.1017/CBO9781107415324, 2018.

Montanari, A. and Koutsoyiannis, D.: Modeling and mitigating natural hazards: Stationarity is immortal!, Water Resources Research, 50,

460   9748–9756, https://doi.org/10.1002/2014WR016092, 2014.

Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., Dunn, R. J., Osborn, T. J., Jones, P. D., and Simpson, I. R.: An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set, Journal of Geophysical Research: Atmospheres, 126, https://doi.org/10.1029/2019JD032361, 2021.

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., Van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram,

465   T., Meehl, G. A., Mitchell, J. F., Nakicenovic, N., Riahi, K., Smith, S. J., Stouffer, R. J., Thomson, A. M., Weyant, J. P., and Wilbanks, T. J.:

The next generation of scenarios for climate change research and assessment, Nature, 463, 747–756, https://doi.org/10.1038/nature08823, 2010.

O'Gorman, P. A.: Contrasting responses of mean and extreme snowfall to climate change, Nature, 512, 416–418, https://doi.org/10.1038/nature13625, http://www.nature.com/articles/nature13625, 2014.

470   Quéno, L., Vionnet, V., Dombrowski-Etchevers, I., Lafaysse, M., Dumont, M., and Karbou, F.: Snowpack modelling in the Pyrenees driven by kilometric-resolution meteorological forecasts, Cryosphere, 10, 1571–1589, https://doi.org/10.5194/tc-10-1571-2016, 2016.

Rao, A. R. and Hamed, K. H.: Flood Frequency Analysis, CRC Press, https://doi.org/10.1201/9780429128813, 2000.

Revuelto, J., Lecourt, G., Lafaysse, M., Zin, I., Charrois, L., Vionnet, V., Dumont, M., Rabatel, A., Six, D., Condom, T., Morin, S., Viani, A., and Sirguey, P.: Multi-criteria evaluation of snowpack simulations in complex alpine terrain using satellite and in situ observations,

475   Remote Sensing, 10, 1–32, https://doi.org/10.3390/rs10081171, 2018.

Roth, M., Buishand, T. A., Jongbloed, G., Klein Tank, A. M., and van Zanten, J. H.: Projections of precipitation extremes based on a regional, non-stationary peaks-over-threshold approach: A case study for the Netherlands and north-western Germany, Weather and Climate Extremes, 4, 1–10, https://doi.org/10.1016/j.wace.2014.01.001, 2014.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bulletin of the American Meteorological

480   Society, 93, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1, 2012.

Tramblay, Y. and Somot, S.: Future evolution of extreme precipitation in the Mediterranean, Climatic Change, 151, 289–302, https://doi.org/10.1007/s10584-018-2300-5, 2018.

Um, M. J., Kim, Y., Markus, M., and Wuebbles, D. J.: Modeling nonstationary extreme value distributions with nonlinear functions: An application using multiple precipitation projections for U.S. cities, Journal of Hydrology, 552, 396–406,

485   https://doi.org/10.1016/j.jhydrol.2017.07.007, 2017.

Verfaillie, D., Déqué, M., Morin, S., and Lafaysse, M.: The method ADAMONT v1.0 for statistical adjustment of climate projections applicable to energy balance land surface models, Geoscientific Model Development, 10, 4257–4283, https://doi.org/10.5194/gmd-10-4257-2017, 2017.

Verfaillie, D., Lafaysse, M., Déqué, M., Eckert, N., Lejeune, Y., and Morin, S.: Multi-component ensembles of future meteorological and

490   natural snow conditions for 1500 m altitude in the Chartreuse mountain range, Northern French Alps, Cryosphere, 12, 1249–1271, https://doi.org/10.5194/tc-12-1249-2018, 2018.

Vernay, M., Lafaysse, M., Hagenmuller, P., Nheili, R., Verfaillie, D., and Morin, S.: The S2M meteorological and snow cover reanalysis in the French mountainous areas (1958 - present), [Data set]. AERIS, https://doi.org/10.25326/37, 2019.

Vernay, M., Lafaysse, M., Monteiro, D., Hagenmuller, P., Nheili, R., Samacoits, R., Verfaillie, D., and Morin, S.: The S2M meteorological and

495   snow cover reanalysis over the French mountainous areas, description and evaluation 1958 - 2020, Earth System Science Data Discussions, https://doi.org/10.5194/essd-2021-249, 2021.

Vionnet, V., Dombrowski-Etchevers, I., Lafaysse, M., Quéno, L., Seity, Y., and Bazile, E.: Numerical weather forecasts at kilometer scale in the French Alps: Evaluation and application for snowpack modeling, Journal of Hydrometeorology, 17, 2591–2614, https://doi.org/10.1175/JHM-D-15-0241.1, 2016.

500   Vionnet, V., Six, D., Auger, L., Dumont, M., Lafaysse, M., Quéno, L., Réveillet, M., Dombrowski-Etchevers, I., Thibert, E., and Vincent, C.: Sub-kilometer Precipitation Datasets for Snowpack and Glacier Modeling in Alpine Terrain, Frontiers in Earth Science, 7, 1–21, https://doi.org/10.3389/feart.2019.00182, 2019.

Wang, X. L., Zwiers, F. W., and Swail, V. R.: North Atlantic ocean wave climate change scenarios for the twenty-first century, Journal of Climate, 17, 2368–2383, https://doi.org/10.1175/1520-0442(2004)017<2368:NAOWCC>2.0.CO;2, 2004.

505   Wehner, M. F.: Characterization of long period return values of extreme daily temperature and precipitation in the CMIP6 models: Part 2, projections of future change, Weather and Climate Extremes, 30, 100 284, https://doi.org/10.1016/j.wace.2020.100284, 2020.

Wilcox, C., Vischel, T., Panthou, G., Bodian, A., Blanchet, J., Descroix, L., Quantin, G., Cassé, C., Tanimoun, B., and Kone, S.: Trends in hydrological extremes in the Senegal and Niger Rivers, Journal of Hydrology, 566, 531–545, https://doi.org/10.1016/J.JHYDROL.2018.07.063, https://www.sciencedirect.com/science/article/pii/S0022169418305766, 2018.

510   Winter, H. C., Brown, S. J., and Tawn, J. A.: Characterising the changing behaviour of heatwaves with climate change, Dynamics and Statistics of the Climate System, p. dzw006, https://doi.org/10.1093/climsys/dzw006, 2017.