1 Supplementary Material Part A. Combinations of general circulation model (GCM) and regional climate model (RCM)

RCMs	GCMs					
	CNRM-CM5	HadGEM2-ES	MPI-ESM-LR	EC-EARTH	IPSL-CM5A-MR	NorESM1-M
RCA4	✓	✓	✓	✓	✓	
CCLM4-8-17	✓	✓	✓	✓		
RACMO22E	✓		✓	✓		
ALADIN53	✓					
ALADIN63	✓					
REMO2009		✓				
RegCM4-6			✓			
WRF331F					✓	
WRF381P					✓	
HIRHAM5						✓
REMO2015						✓

Table S1. 20 GCM-RCM pairs considered. We always rely on the version r1i1p1 except for EC-EARTH where we use r12i1p1.

2 Supplementary Material Part B. Results for the elevation 900 m and 2100 m

We apply the same methodology at the elevation 900 m and 2100 m.

In Figure S1, we present the selected parameterization of the GEV distribution. In Figure S1 **(a)**, we observe that at 900 m many massif in the South of the Alps are discarded in the first step of the three-step selection method. This is simply due to the fact that more years without accumulated snow, i.e. annual maxima equal to zero, happen at 900 m than at 1500 m. Besides, a parameterization with zero adjustment coefficients is selected for most remaining massifs. In Figure S1 **(b)**, we observe that at 2100 m parameterizations with adjustment coefficients are selected for a majority of massifs.

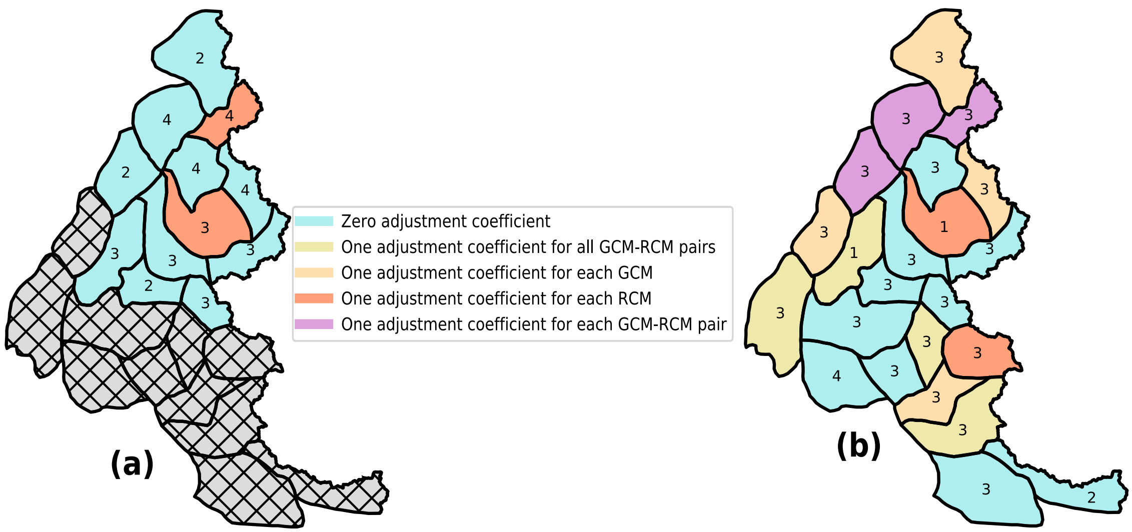

Figure S1. Map of the selected parameterization of the GEV distribution at **(a)** 900 m and **(b)** 2100 m. The selected parameterizations of the adjustment coefficients are illustrated with colors, while the selected numbers of linear pieces are written on the map.

In Figure S2 and Figure S3, we illustrate the relative change of return levels for +2, +3, and +4 degrees of global warming w.r.t. +1 degrees at 900 m and 2100 m of elevation. On one hand, for 900 m we find that the relative decrease is stronger than at 1500 m, as we observe that the average change of return level is equal to -0.4 kN m^{-2} (-15%), -1 kN m^{-2} (-33%), -1.5 kN m^{-2} (-48%) for +2, +3, and +4 degrees, respectively. On the other hand, for 2100 m, we find that the relative decrease is weaker than at 1500 m, as we observe that the average change of return level is equal to -0.9 kN m^{-2} (-9%), -2.1 kN m^{-2} (-22%), -3.4 kN m^{-2} (-36%) for +2, +3, and +4 degrees, respectively.

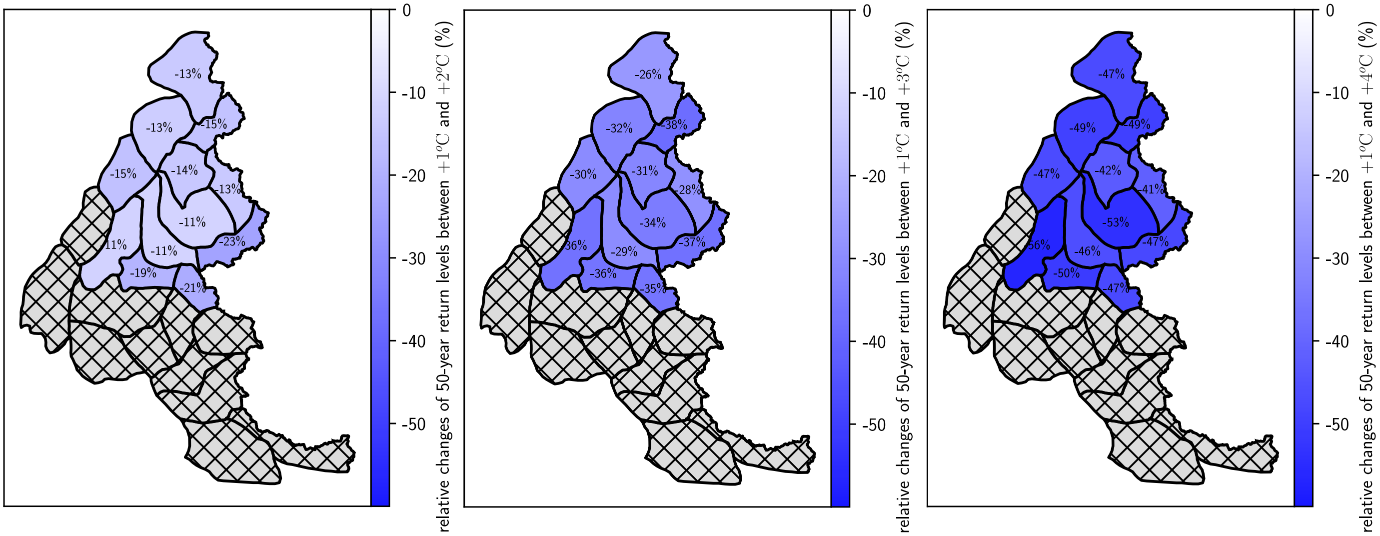

Figure S2. Relative changes in 50-year return levels (RL50) of snow load at 900 m for +2, +3, and +4 degrees of global warming under the scenario RCP8.5 w.r.t +1 degrees of global warming.

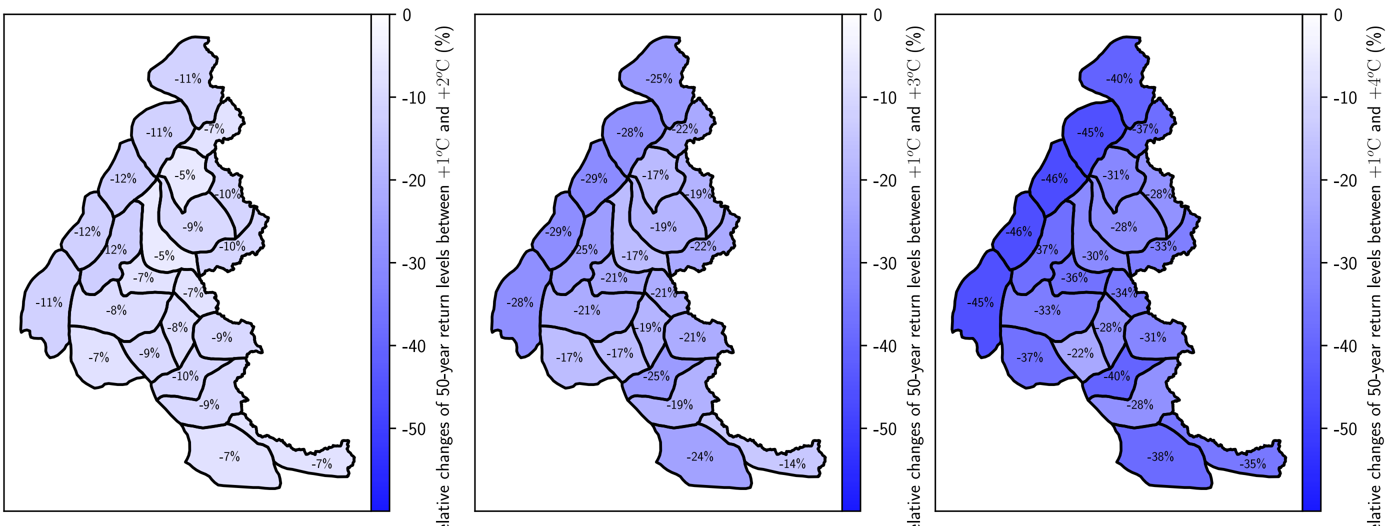

Figure S3. Relative changes in 50-year return levels (RL50) of snow load at 2100 m for +2, +3, and +4 degrees of global warming under the scenario RCP8.5 w.r.t +1 degrees of global warming.

3 Supplementary Material Part C. Detailed results of the evaluation experiments for elevation 1500 m

In Table S2, we detail the mean logarithmic score for the model-as-truth experiment at 1500 m. This score is averaged on the pseudo-observation of the validation set (2020-2100) for the 12 GCM-RCM pairs set as pseudo-observations. We find that a parameterization with zero linear pieces, which corresponds to a stationary GEV distribution, always performs worse than non-stationary approaches with at least one linear piece. A parameterization with one linear piece, which is the most used parameterization in the literature (Tab. 1), is either on par with approaches using more linear pieces (e.g. for the Haute-Tarentaise massif, and for the Vanoise massif) or worse than such approaches (e.g. for the Aravis massif, and the for the Bauges massif).

Massif name	Mean logarithmic score ($\overline{\text{LS}}$) for the model-as-truth experiment with zero adjustment coefficients.			
	1 linear piece	2 linear pieces	3 linear pieces	4 linear pieces
Aravis	1.73	1.71	**1.71**	1.71
Bauges	1.6	1.58	**1.58**	1.58
Beaufortain	1.63	1.63	**1.62**	1.62
Belledonne	1.5	1.48	**1.48**	1.48
Chablais	1.62	1.6	**1.6**	1.6
Champsaur	1.07	1.05	**1.04**	1.05
Chartreuse	1.65	1.64	**1.64**	1.64
Devoluy	0.89	0.87	**0.86**	0.87
Grandes-Rousses	1.23	1.22	**1.22**	1.22
Haute-Maurienne	0.84	0.83	**0.83**	0.83
Haute-Tarentaise	1.47	1.47	1.47	**1.47**
Maurienne	1.18	1.18	**1.17**	1.18
Mont-Blanc	1.7	1.69	**1.69**	1.69
Oisans	1.11	1.1	1.1	**1.1**
Parpaillon	0.83	0.81	**0.81**	0.81
Pelvoux	1.12	1.11	**1.11**	1.11
Queyras	0.5	0.49	**0.48**	0.48
Thabor	1.0	1.0	**0.99**	0.99
Ubaye	0.29	0.27	**0.26**	0.27
Vanoise	1.44	**1.44**	1.44	1.44
Vercors	1.29	1.27	**1.27**	1.27

Table S2. Mean logarithmic score for the model-as-truth experiment for the elevation 1500 m. The mean logarithmic score is averaged on the held-out pseudo-observations (2020-2100) for the 12 GCM-RCM pairs set as pseudo-observations for each massif.

In Table S3, we detail the mean logarithmic score for the split-sample experiment at 1500 m. The mean logarithmic score is averaged on three split-sample experiments, where the calibration set corresponds to $60\%, 70\%$, and 80% of the observations.

We find that a parameterization with zero adjustment coefficients performs better than with adjustment coefficients on many massifs, e.g. for the Oisans and Beaufortain massifs. For the massifs where a parameterization with adjustment coefficients is selected, we observe that the four parameterizations with adjustment coefficients (one for all GCM-RCM pairs, one for each GCM, one for each RCM, one for each GCM-RCM pair) perform almost similarly.

Massif name	Mean logarithmic score ($\overline{\text{LS}}$) for the split-sample experiment. For each massif, we rely on the selected number of linear pieces selected with the model-as-truth experiment.				
	Zero	One for all GCM-RCM pairs	One for each GCM	One for each RCM	One for each GCM-RCM pair
Aravis	**1.84**	1.88	1.88	1.88	1.88
Bauges	**1.7**	1.75	1.74	1.75	1.74
Beaufortain	**1.77**	1.83	1.83	1.83	1.83
Belledonne	**1.62**	1.67	1.67	1.67	1.67
Chablais	**1.77**	1.8	1.79	1.79	1.79
Champsaur	1.18	1.16	**1.15**	1.15	1.15
Chartreuse	1.76	1.75	1.74	1.74	**1.74**
Devoluy	1.18	**1.16**	1.17	1.17	1.17
Grandes-Rousses	**1.3**	1.44	1.43	1.43	1.43
Haute-Maurienne	**1.13**	1.14	1.14	1.14	1.14
Haute-Tarentaise	**1.52**	1.55	1.54	1.55	1.55
Maurienne	**1.3**	1.38	1.37	1.37	1.37
Mont-Blanc	1.71	1.71	1.71	1.71	**1.71**
Oisans	**1.11**	1.28	1.27	1.27	1.27
Parpaillon	**1.06**	1.08	1.08	1.08	1.07
Pelvoux	1.39	1.37	**1.37**	1.37	1.37
Queyras	1.06	0.96	0.96	0.96	**0.96**
Thabor	**1.1**	1.14	1.13	1.13	1.13
Ubaye	1.08	**1.05**	1.05	1.06	1.05
Vanoise	**1.42**	1.43	1.42	1.42	1.42
Vercors	1.46	1.42	1.41	**1.41**	1.41

Table S3. Mean logarithmic score for the split-sample experiment for the elevation 1500 m. The mean logarithmic score is averaged on three split-sample experiments , where the calibration set correspond to $60\%, 70\%$, and 80% of the observations.