Comments from Referee #3

We thank Reviewer #3 for these comments and this in-depth review of our manuscript. We appreciate this expertise although we think that some aspects of our work have been misunderstood, probably due to a lack of explanations in the current manuscript. Please find below our responses to these comments.

Main comments:

RC3#1. My feeling is that this paper is not providing a good solution to a real problem. They consider several GCM-RCM model pairs ('model' in the following). In the Introduction the authors point out that previous studies evaluated extreme value statistics (EVS) for individual models and in some cases they took then the ensemble mean of return levels. In this regard the authors are concerned that the estimates for separate models are not so reliable because of data scarcity. However, they seem to do this themselves.

Thank you for this comment. Indeed one of our main motivations for the proposed method is to improve the estimates using a joint estimation approach. Indeed, fitting individual GEV to each climate simulation leads to a set of very uncertain estimates. Taking the mean of very uncertain estimates of a return level is not equivalent to having one estimate of the mean return level obtained from the whole set of climate simulations, as is done in this work. This will be clarified in the revised version of the manuscript.

RC3#2. They introduce the concept of "adjustment coefficients", which is really just a difference of an estimate of the GEV parameter for a particular model (or subset of the data) from the estimate upon lumping all the data. I think we really don't need a name for this, beside the problem that they do what they criticised.

It must be clear that these “adjustement coefficients” are not direct differences between estimates of GEV parameters. They are estimated by maximizing the likelihood function (4) and represent a compromise between fitting past observations and future maxima as obtained from the different climate models. The name for these parameters can of course be debated but they represent the adjustment of a systematic shift between maxima obtained by the different climate models (see Figure 3 above in response to the comment RC2#2). According to the properties of the climate models, this shift can be assumed to be the same for each GCM, each RCM. We also test the possibility of estimating these parameters for all GCM/RCM pairs. In this application, these parameters do not often lead to the best model, as shown in Figure 3c of the submitted manuscript. However, we believe that this possibility can be interesting in other applications where annual maxima from climate models largely differ from observed maxima in the historical period. As indicated in our response to comment RC2#2, a simple sketch of the concept around these parameters will be added in the revised version.

RC3#3. On the other hand, lumping all the data together, in order to have seemingly more robust estimates, is also problematic. As i pointed out in some recent publications of mine, a model ensemble (or multi-model ensemble, MME, as it's most often called) does not represent an objective probability distribution. As such, fitting a GEV to MME data is flawed methodology. It has no meaningful probabilistic interpretation. On the contrary, doing this for (converged) initial condition ensembles is fine. It is actually a profound scientific challenge on how to use MME data in a meaningful way. I don't mean to discourage anyone from trying, though, and hope that real progress can be made. The authors promised a constraining of the estimates/projections using observational data. Emergent constraints is now a popular concept, but it appears to me that the authors did nothing like that. They simply threw the observational data in the mix. However, the information from it is diluted by the large amount of model data.

Thank you for this comment. We entirely agree that fitting probabilistic models to samples obtained from a multi-model ensemble is highly questionable, as it is questionable to apply a statistical test to assess the statistical significance of projected averages obtained from a MME (see Storch and Zwiers, 2013). However, MMEs are widely used to assess the impact of climate change because it seems difficult to trust more one particular model (or one GCM-RCM-impact model chain) than the others. Emerging constraints (e.g. Ribes et al., 2021) try to answer this challenge but we agree that the limitation of our approach must be clearly acknowledged. This will be done in the revised version.

In this work, it is also true that we put the same weight on the observed maxima and the maxima from the climate simulations. As a consequence, it is true that observed maxima have a limited influence on the estimated parameters. Some attempts to put more weights on the observations did not lead to satisfying results because of the limited period covered by the observations, which does not provide much information about the evolution of the GEV parameters. This limitation is also discussed in our response to comment RC1#2 and will be discussed in more depth in the revised manuscript.

RC3#4. Obs data is rather used for bias correction. I'm not sure this was done. Or, if it was done, then it seems to have even less use to throw the obs data into the mix for doing EVS.

A quantile mapping method (ADAMONT) is a bias-correct method that was applied to climate simulations. This method shows some limitations to correct the tail of the distribution since, by definition, extreme values are rarely observed. This explains why some biases can still be expected in the bias-corrected climate simulations.

RC3#5. I share my detailed comments on the manuscript with the authors in an annotated pdf. Hopefully it is useful one way or another. I'm sorry that i cannot be more positive this time. If i misunderstand something, i'm happy to learn from the author's response.

We appreciate these comments and are also happy to discuss them and defend our position.

Specific comments in the annotated pdf

RC3#6. l.3: chain of the MME?

As indicated in our response to comment RC2#4, this term is sometimes used to refer to a particular climate simulation chain (e.g. scenario/GCM/RCM). This term will be removed in the revised manuscript.

RC3#7. l.7: flexible?

“Flexible” was used in the sense that the proposed nonstationary GEV model can be adapted to a great variety of nonlinear evolution thanks to the piecewise functions. This term will be removed since it does not seem to be informative.

RC3#8. l.8 “functions of time”?

As indicated in section 3.2, parameters are functions of anomalies of global mean temperature. This will be added.

RC3#9. l.8. “adjustment coefficients”?

These adjustement coefficients are additional parameters (see our response to the comment RC3#2).

RC3#10. l.17. Maybe this should go in the end of the sentence.

Thank you for this suggestion, this will be done.

RC3#11. l.19: I cannot think anything else.

“One of” will be removed in the revised version.

RC3#12. l.19: Even the most basic climatic means are obtained/projected by simulations.

We agree with the reviewer.

RC3#13. l.23: I would say that quantiles and tail probabilities are more often used by climate scientists than EVT.

It is true that return levels are often obtained using different empirical formulas (e.g. Gringorten) or various distributions. EVT seems to be dominant in the past ten years (see, e.g., Serinaldi and Kilsby, 2014, concerning rainfall extremes). We propose to replace “usually” by “often” to indicate that the EVT approach is undoubtedly popular for the assessment of climate extremes.

RC3#14. l.24. Why robustly? If we are interested in too high return levels, surely the estimates can be very poor.

This comment was also made by reviewer #2 (see comment RC2#7) and we agree that “robustly” should be removed here.

RC3#15. l.29. I don't know what you mean by chain. I guess you mean time series. If so, "chain" is not used in English in this context.

The term “chain” will be removed (see also our response to the comment RC2#4).

RC3#16. l.32. 30 years is not really a slice, is it?

“slices” will be replaced by “periods”.

RC3#17. l.44 Do we really need this reference for nonstat EVT? This seems to unnecessarily inflate the bibliography. There is Coles already cited.

This reference will be removed.

RC3#18. l.47 chain

This will be replaced by climate simulation.

RC3#19. l.51 This is poor English. Otherwise, this concept is unfortunately flawed. We can only fit a probability distribution if we have reason to believe that there is an objective physical measure from which our data is drawn from. However, regarding a MME there isn't. It is an ensemble of opportunity (or chance). There isn't a probability space from which Earth System Models are samples. This is what we were articulating in another context of forced response. The right approach is considering Initial condition large ensembles because in that case there is an objective physical probability measure generated by the one climate model that represents some hypothetical physics.

This part of the sentence will be rephrased and replaced by “on a temporal non-stationary GEV model fitted to all ensemble members”. See our response to comment RC3#3 concerning the rest of this remark.

RC3#20. l.53 why this reference?

This reference will be removed.

RC3#21. l.56 This should not be assumed but rather checked. What one would check is whether the ensemble still depends on its initialisation or already converged to the snapshot attractor. The check can proceed by checking the stat significance of the distributions (concerning a particular scalar observable) bw TWO ensembles that were initialised differently. (Drotos et al. J Clim 2015).

We agree but it does not seem that convergence has been checked in the papers cited in Table 1 (Kharin and Zwiers (2004), Wang et al. (2004), Aalbers et al. (2018), Fix et al. (2018)).

RC3#22. l.57. On l 51-52 the text reads such that you follow this approach.

Thank you for this comment. This sentence was misleading and will be modified.

RC3#23. l.60 How does this solve the described problem?

Using past observations and adjustment coefficients, this nonstationary GEV model assumes that different climate simulations can be used to project future return levels. This was perhaps unclear from this version of the manuscript and we hope that the proposed amendments will motivate this approach (see also the response to comment RC3#2).

RC3#24. Why do you exclude the shape parameter from this treatment?

This is explained later at l. 137: “because it sometimes leads to prediction failures.” for example when a negative shape parameter leads to predictive GEV distribution with an upper bound and when some maxima exceeds this upper bound.

RC3#25. l.63. By this do you mean "the various GEV distributions of all GCM-RCM pairs"? Variability of climate trajectories does not sound anything like that.

Here, we mean that a MME obtained with many different climate models can lead to a large variation of projected changes. This is true for average changes (see, e.g. Evin et al., 2021a for a EURO-CORDEX RCM ensemble) and for extremes (Rajczak and Schär, 2017).

RC3#26. l.64. “jointly fitted” What does this refer to.

This refers to the joint estimation over all simulations of the MME but “jointly” is not really needed here and will be removed.

RC3#27. l.65 Why do you say that past observations represent the most likely climate trajectory? This is so strange language to me. And i'm not convinced that it makes deep sense.

Thanks for this comment. We agree that “to represent the most likely climate trajectory” was not clear, it will be removed.

RC3#28. l.65 By this do you mean that you fit output data from individual models whereby you have different e.g. location parameter estimates for the different models? If so, there seems to be no difference from what you describe on l46. Rather, you stop short of proceeding with obtaining some ensemble statistics, e.g. mean of the return levels.

Thank you for this comment. As indicated above, the purpose of these adjustment coefficients is to account for systematic shifts in climate simulations, either for the location or the scale parameter. However, we consider different options for these coefficients (one by GCM, one by RCM, etc.) and this will be illustrated by a simple sketch. However, we insist on the fact that these adjustement coefficients are part of the joint inference and are **not** obtained by first fitting individual GEV models to the different climate simulations.

RC3#29. l.68. Does the data look piecewise linear? If not, why not using some polynomial?

Polynomial functions have been tested but can lead to poor models for some types of evolutions. In particular, they are not very well suited to changes of trends (e.g. an increase followed by a decrease). Piecewise linear functions provide a certain flexibility in this regard.

RC3#30. l.76 What does this mean. I suppose you might need to provide a reference here.

This is explained at l.93-94. “Statistically adjusted” will be removed from this sentence.

RC3#31. l.78 resort?

Thank you for this comment, this will be corrected.

RC3#32. It would be good to have a short explanation what it is so that one doesn't need to look it up.

We will add the following definition “the pressure exerted by the snow”.

RC3#33. l.85. It would be helpful to say that snow load has the physical unit of pressure, that is Pascal.

Snow load usually has the units of kN/m2. This will be added.

RC3#34. l.85. Otherwise, one has to wonder if this study is very useful in case it is very common to have different snow depths just meters apart, say, from house to house or street to street.

This is a relevant comment and it is true that snow depth is known to vary a lot in space. However, aside from these measurement issues, there is a clear need to assess snow load hazards as it is indicated in the abstract (for the design of roofs) and in our response to the comment RC2#6.

RC3#35. l.85. If i wanted to see what a particular model has in a particular model, i could not find it. Perhaps do not connect the lines.

We agree that it is difficult to look at a particular climate simulation due to the many trajectories. Here the figure mainly aims at showing the overall variability from one year to the other and between the different members of the ensemble.

RC3#36. l.92. “The” is missing.

“The” will be added.

RC3#37. l.92. software package Crocus?

Thank you for this comment. Crocus is actually the part of the S2M numerical chain that produces the evolution of the snowpack. This precision will be added to the manuscript.

RC3#38. l.99. “The” missing”.

This will be added.

RC3#39. l.99. “Observational”?

Thanks, this will be corrected.

RC3#39. l.101. Is it because this way you have a much better "stat agreement" bw. the models?

It is true that there is a better agreement between the statistical properties of the snowload maxima produced by the different climate models when the covariate is the global temperature than when it is time. Indeed, the climate sensitivity of each GCM can produce large differences of warming for a particular time horizon (e.g. 2050). Taking a warming horizon (e.g. +2°C) removes this “climate sensitivity effect” and makes the different projections more consistent (see also Verfaillie et al., 2018).

RC3#40. l.103. “does not depend on the internal variability of GMST” I don't understand what you mean.

We mean that the “smoothed GMST” is not affected by the internal variability simulated by the GCM (see the fluctuations of the raw GMST at the decadal scale in Fig. 2).

RC3#41. l.109-111: “Indeed theoretically, as the central limit theorem motivates asymptotically sample means modeling with the normal distribution, the Fisher–Tippett–Gnedenko theorem (Fisher and Tippett, 1928; Gnedenko, 1943) encourages asymptotically sample maxima modeling with the GEV distribution.” Please tidy up the English here.

Thank you, this part will be rephrased as follows: “Indeed, similarly to the central limit theorem that motivates to model means obtained from different samples using a normal distribution, the Fisher–Tippett–Gnedenko theorem (Fisher and Tippett, 1928; Gnedenko, 1943) encourages to model maxima using the GEV distribution.”

RC3#42. l.111: “Y” Realised values of a random variable Y are usually denoted by a lower case letter y_i; l.112 This is meant to denote the rand var indeed.

Thank you for this comment. We will modify this sentence as follows: “if Y is the random variable representing annual maxima, we can assume…”.

RC3#43. “u+ denotes max(u,0)” Where is this coming from?

This notation is sometimes used to indicate that the cdf is defined only for positive values of $1+\xi(y-\mu)/\sigma$, when ξ is not equal to 0 (see, e.g., Eq. 1 in Coles and Perrichi, 2003).

RC3#44. l.116 “Due to these theoretical justifications, the GEV distribution enables a robust estimation of return levels” This is a dogma at best. The variance and bias of estimates depends on the amount of data and the tail probability 1/T associated with the estimated return level.

This comment was also raised by the reviewer RC2 (see comment RC2#7). We agree that we do not provide strong results showing a particular robustness of the return level estimates. We propose to remove it when it appears in the manuscript.

RC3#45. l.119 “for the design working life of building” please make sure the English is correct.

Thank you for this comment, this part of the sentence will be replaced by “for building design”.

RC3#46. l.124. “the smoothed anomaly of GMST.” to be removed.

Ok, this will be removed.

RC3#47. l.124. “log link function” Provide reference where the link function is defined.

It is defined in Eq. 2 where we indicate that the log transform of the scale parameter is a piecewise linear function of T. To clarify this point, we will move this sentence after the equation.

RC3#48. l.125 “numerical optimization.” What does it refer to? I suppose the maximum search for MLE.

Yes, this is correct, we will precise this point with a reference to section 3.3.

RC3#49. l/126. “T_t^obs” I don't think you need to write t here.

The subscript *t* indicates that the global temperature is obtained for the year *t*. It is actually missing at l.121 above and will be added.

RC3#50. l.126. For linear piece l, the slope is sum_i=1^l mu_i, right?

For a linear piece l, the slope is mu_l, and starts from the end of the previous linear piece (l-1).

RC3#51. l.127. “θ is vector of coefficients” What coefficients? Also, the realised values don't depend on anything. Please be careful with the notation.

Thank you for this comment. θ is actually the vector of parameters corresponding to the linear pieces, i.e. $\{\mu_i, \sigma_i, \xi_i, i=1,...,L\}$, this will be clarified.

RC3#52. l.129. “κ_2” You also have kappa_1 in eq. (2).

κ_1 corresponds to T_{min}, and is the starting point of the piecewise linear function. In this sentence, we simply explain where the slope of the piecewise linear function is changing. However, the subscript for κ_{L-1} was incorrect and will be replaced by κ_L.

RC3#53. l.131. “Let” This sentence could be put more concisely.

We agree, this sentence will be rephrased.

RC3#54. l.131. “Maximum” -> maxima.

Here, we refer to a single maximum.

RC3#55. l.131. “pair” you can loose this, the reason why you would introduce k is to denote a pair.

Thank you for this suggestion, this will be done.

RC3#56. l.133-143. “$GCM^{pair\ k}$” k is not a co-variate, it's an ARBITRARY label to identify a model. “aim at adjusting the distribution of GCM-RCM pairs w.r.t. the distribution of the past observations.” I cannot see where in eq. (3) the observations are involved. “adjustment coefficient” Up to here i still don't know what is an "adjustment coeff". “assume that these adjustment coefficients are constant, i.e. the same for historical and future climates.” I guess this means that you have one set of coeff, not one set per piece of the piece-wise lin fun. However, why not? I didn't read (Brown et al. 2014) to see if they justify this assumption. “μ_{GCMi}” okay, but what is the functional form, i.e., the co-variate dependence. What is the co-variate to start with? Is it the integer i?

This paragraph will be strongly modified in order to clarify how these adjustement coefficients are introduced and to explain the rationale behind this use. The adjustement coefficients are additional parameters that assume that the different members of the ensemble share the same nonstationary GEV model up to a constant shift for the location and scale parameters (see proposed illustration in Fig. 3, in response to the comment RC2#2 above). This shift can be shared by all members (“One for all GCM-RCM pairs”). In this case, there is a single parameter that represents a common shift for all members of the ensemble with respect to the observations. We also consider the possibility of having different shifts for each GCM or for each RCM (“One for each GCM”,”One for each RCM”), and for all members (“One for each GCM-RCM pair”). Obviously, adding many parameters to our model is only desirable if it leads to an important improvement of our predictive skills, evaluated using the split-sample experiment. As it is not often the case in our application (see Fig. 3 of the manuscript), we did not test more complex forms for the adjustement coefficients. This argument will thus replace the reference to Brown et al., 2014 in the revised manuscript. Again, it must be clear that the adjustement coefficients are directly estimated by maximising the likelihood (4) and **not** by first fitting individual GEV models to each climate simulation.

RC3#57. l.137. “because it sometimes leads to prediction failures.” This is a very opaque explanation.

See response to comment RC3#24 above.

RC3#58. l.146 “all annual maxima of a given massif, i.e. annual maxima from the observations and from the 20 GCM-RCM pairs” This is the first time you give indication how you mean to use the obs and model data together.

Otherwise, i'm not at all convinced if this is any good solution to the problem that models are different from each other and from reality. You just throw reality into the mix. But if you have more and more model data, it matters less and less that you have a seed of the truth.

Thank you for this comment which is also related to the comment RC1#2 above. We agree that the combination of the observations and of the different GCM-RCM pairs in Eq. (4) is simplistic and could be improved since the information provided by the observations is diluted by the large size of the ensemble (number of members times the length of the projections). We will discuss these limitations more clearly and in more depth in the revised paper. Different tests have been carried out in order to put more weights on the observations but the formulation of these weights was problematic. However, we see our contribution as a first step in that direction.

RC3#59. l.149. Eq. 4. We really don't need this formula. It's enough to say that you perform MLE for all the data, obs and model. Or, it would be enough, had it been worth to do, as i commented above. Furthermore, you forget to mention that you also do the fitting to individual model pairs too, or various subsets of your data.

We prefer to keep this equation that helps the reader to understand how observed maxima and simulated maxima from the projections are combined (despite the limitations of this approach). It is true that the fitting is also done for various subsets of the data for the evaluation experiments, and this will be added. Nonstationary GEV models are fitted to each climate simulation only for the sake of illustration in Figure 4 and it would be confusing to add this explanation at this stage of the paper in our opinion.

RC3#60. l.153-154. “the calibration of the non-stationary GEV distribution.” no idea what this is.

We agree that this part of this sentence was unclear and not necessary, it will be removed.

RC3#61. l.169. Unfortunately, i do not understand what you mean by evaluating predictive performance here.

As indicated in Gneiting and Raftery (2007), the log-score is used to assign “a numerical score based on the predictive distribution and on the event or value that materializes”, in our case, the predictive distribution is obtained from the GEV distribution defined in Eq. (2) with an estimated parameter vector $\hat{\theta}$and the value used to evaluate the predictive performance (the future data of the pseudo-observations in the model-as-truth experiment, and the observations which have been discarded from the dataset fitting in the split-sample experiment). Additional explanations and the reference to Gneiting and Raftery (2007) will be added to the revised manuscript.

RC3#62. l.171. ”we select one parameterization” Rather, the selection determines the set of data you will fit.

The final GEV model is always obtained using the observations dataset and the ensemble of climate simulations composed of 20 GCM-RCM pairs. Maybe this was unclear in the original version of the manuscript and we hope that the revised manuscript will clarify this point.

RC3#63. l.171-172. “we select the number of linear pieces with a model-as-truth experiment using zero adjustment coefficients for the GEV parameters” not sure how this would be done

The assumption is that the model-as-truth experiment gives a first indication on the global evolution of the GEV parameters, based on the predictive performance of the long climate runs (1950-2100), while past observations are often limited to assess these evolutions. For this reason, most of the nonstationary GEV models fitted on past observations assume linear trends for the GEV parameters (usually the location and the scale). Our study proposes an approach to assess more complex evolutions of the GEV parameters based on a climate ensemble, even if some limitations must be acknowledged.

RC3#64. l.179. “uncertainty interval” confidence interval it's called. However, for the said reason i'm not optimisitic it's any meaningful .

Thank you for this comment, we will replace “uncertainty interval” by “confidence interval” in the revised manuscript.

RC3#65. l.181. Eq 5. On line 118 you already have this eq. This is a rather unnecessary duplication.

We agree that this equation can appear unnecessary if the reader is familiar with nonstationary GEV models. However, on l.118, the return level is provided for a stationary GEV distribution while the return level depends on T in Eq. 5. Furthermore, it seems important to stress that these return levels are not based on the adjustment coefficients.

RC3#66. l.226. Seeing the discontinuity of the slopes for many individual model pairs i wonder if the piece-wise lin model is really good. Perhaps the chi-squared test should really be done. Coles notes what to do in nonstationary EVS, which is what i also followed: https://nhess.copernicus.org/preprints/nhess-2020-117/nhess-2020-117.pdf

Thank you for this comment. First, from the comments above, it seems that there was a misunderstanding about these individual fittings which are used solely for the sake of illustration. We never fit GEV models to individual GCM-RCM pairs in the rest of the paper. The gray curves in Fig. 4 showing the return levels obtained from these individual GEV models are not very smoothed indeed, maybe due to the lack of robustness of these fittings. It is in fact one important motivation for the joint inference of the likelihood (4). The goodness-of-fit of the final GEV models (colored curves in Fig. 4) have been carried out using the Anderson-Darling tests (see comment RC1#1) and these results will be discussed in the revised manuscript. Finally, we agree that an interesting approach for nonstationary GEV models is to express the GEV parameters as a linear combination of additional covariates. This is an approach that we have also considered in other applications (see, e.g. Evin et al., 2021b, for an application to extreme avalanche cycles). When there is a clear relationship between the statistical properties of the extreme variables and some climate descriptors (as is the case in the aforementioned paper between extreme cold temperatures in Europe and the arctic oscillation index), this approach is particularly powerful. However, this is not the case in our application to extreme snow loads in the French Alps where the relationships between climate indices and climate extremes related to precipitation events are often very weak (see Belkhiri and Kim, 2021). A possible idea would be to use the statistical properties of the climate simulations as covariates. These possible avenues will be discussed in the revised version of the manuscript.

RC3#67. l.266. “100 maxima” That does not sound very many.

Our point was that 20 x 100 maxima is a large amount of information compared to the 61 observed maxima.

RC3#68. l.287 “The 90% uncertainty intervals” Why not the usual 95% but a 90% CI?

This 90% CI was preferred to rely on the 5% and 95% quantiles instead of 2.5% and 97.5% quantiles which are estimated with more uncertainties.

References

Belkhiri, L., and T.-J. Kim. 2021. “Individual Influence of Climate Variability Indices on Annual Maximum Precipitation Across the Global Scale.” *Water Resources Management* 35 (9): 2987–3003. https://doi.org/10.1007/s11269-021-02882-8.

Coles, S., and L. Pericchi. 2003. “Anticipating Catastrophes through Extreme Value Modelling.” *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 52 (4): 405–16. https://doi.org/10.1111/1467-9876.00413.

Drótos, G., T. Bódai, and T. Tél. 2015. “Probabilistic Concepts in a Changing Climate: A Snapshot Attractor Picture.” *Journal of Climate* 28 (8): 3275–88. https://doi.org/10.1175/JCLI-D-14-00459.1.

Evin, G., S. Somot, et B. Hingray. 2021a « Balanced Estimate and Uncertainty Assessment of European Climate Change Using the Large EURO-CORDEX Regional Climate Model Ensemble ». *Earth System Dynamics* 12, n° 4: 1543- 69. https://doi.org/10.5194/esd-12-1543-2021.

Evin, G., P. D. Sielenou, N. Eckert, P. Naveau, P. Hagenmuller, and S. Morin. 2021b. “Extreme Avalanche Cycles: Return Levels and Probability Distributions Depending on Snow and Meteorological Conditions.” *Weather and Climate Extremes*, July, 100344. https://doi.org/10.1016/j.wace.2021.100344.

Gneiting, T., and A. E Raftery. 2007. “Strictly Proper Scoring Rules, Prediction, and Estimation.” *Journal of the American Statistical Association* 102 (477): 359–78. https://doi.org/10.1198/016214506000001437.

Hu, G., T. Bódai, and V. Lucarini. 2019. “Effects of Stochastic Parametrization on Extreme Value Statistics.” *Chaos: An Interdisciplinary Journal of Nonlinear Science* 29 (8): 083102. https://doi.org/10.1063/1.5095756.

June-Yi Lee, Tamás Bódai. (2021) Indian summer Monsoon Variability 1st Edition, El Niño-teleconnections and beyond: Chapter 20, Future Changes of the ENSO-Indian Summer Monsoon Teleconnection, Elsevier, pp. 393-412.

Rajczak, J., and C. Schär. « Projections of Future Precipitation Extremes Over Europe: A Multimodel Assessment of Climate Simulations ». *Journal of Geophysical Research: Atmospheres* 122, n° 20 (2017): 10,773-10,800. https://doi.org/10.1002/2017JD027176.

Ribes, A., S. Qasmi, and N. P. Gillett. 2021. “Making Climate Projections Conditional on Historical Observations.” *Science Advances* 7 (4): eabc0671. https://doi.org/10.1126/sciadv.abc0671.

Serinaldi, F., and C. G. Kilsby. « Rainfall Extremes: Toward Reconciliation after the Battle of Distributions ». *Water Resources Research* 50, n° 1 (1 janvier 2014): 336- 52. https://doi.org/10.1002/2013WR014211.

Storch, Hans von, and Francis Zwiers. 2013. “Testing Ensembles of Climate Change Scenarios for ‘Statistical Significance.’” *Climatic Change* 117 (1): 1–9. https://doi.org/10.1007/s10584-012-0551-0.

Tél, T., Bódai, T., Drótos, G., Haszpra, T., Herein, M., Kaszás, B., Vincze, M. (2020) The Theory of Parallel Climate Realizations – A New Framework of Ensemble Methods in a

Changing Climate: An Overview, Journal of Statistical Physics, http://doi.org/10.1007/s10955-019-02445-7

Verfaillie, D., M. Lafaysse, M. Déqué, N. Eckert, Y. Lejeune, et S. Morin. « Multi-component ensembles of future meteorological and natural snow conditions for 1500m altitude in the Chartreuse mountain range, Northern French Alps ». *The Cryosphere* 12, n° 4 (10 avril 2018): 1249- 71. https://doi.org/10.5194/tc-12-1249-2018.

Vionnet, V., E. Brun, S. Morin, A. Boone, S. Faroux, P. Le Moigne, E. Martin, et J.-M. Willemet. « The detailed snowpack scheme Crocus and its implementation in SURFEX v7.2 ». *Geosci. Model Dev.* 5, n° 3 (24 mai 2012): 773- 91. https://doi.org/10.5194/gmd-5-773-2012.