Authors response to comment on hess-2022-21

Anonymous Referee #2

Author comment on **A gridded multi-site precipitation generator for complex terrain: An evaluation in the Austrian Alps** by Hetal Dabhi et al., Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2022-21-RC2, 2022

The paper entitled 'A gridded multi-site precipitation generator for complex terrain: An evaluation in the Austrian Alps' by Hetal Dabhi and co-authors describes an extension of the daily stochastic spatio-temporal precipitation generator of Kleiber et al. (2012) to mountainous areas with complex topography, and illustrates the proposed framework in the Austrian Alps based on a network of 29 meteorological stations.

The paper is well written and structured, and the topic is relevant for the journal Hydrology and Earth System Sciences. The selected case study seems appropriate to test the proposed precipitation generator in an area with complex topography.

Despite the above qualities, the manuscript also contains major shortcomings, which, in my opinion, must be corrected before the paper can be considered for publication:

We are thankful to Reviewer-2 for the time and constructive feedback on our manuscript. In the following we provide our response comment by comment.

Major comments:

1) It is claimed that the paper provides an extension of the model of Kleiber et al. (2012) to areas with complex topography, but the proposed extensions (Kriging with external drift (KED) of model parameters and altitude dependent covariance function) are not very convincing. According to assessment results, these extensions do not offer significant improvements compared to the original model despite the addition of a lot of model complexity. In my opinion, the authors should therefore test the original model in their case study, and propose extensions only if the added complexity generates clear improvements in simulation results. If it is not the case, I would recommend to stick to the simplest possible model, and introduce the current paper as a case study testing the performance of the original model in the presence of complex topography, which I believe is already an interesting contribution.

Response: Concerning 'complexity of the model', it should be noted that we have proposed only two changes to the original model, which do not add much complexity to the model: i) Allow elevation dependence in the covariance structure, ii) KED instead of OK for interpolation of the parameters. Elevation dependence in the covariance structure is the natural assumption for the complex topography. Such elevation dependence has already been incorporated in the correlation structure by Wilks (1999) and Wilks (2009). Due to the inclusion of the elevation difference in the covariance structure, only one parameter increases in each of the two models (i.e., the occurrence and amount models), which is the range parameter in the vertical direction. Also, we have explicitly stated in the results and conclusions that our model adds value due to the KED and not much due to the inclusion of the anisotropy. We also have implemented the original model (Iso-OK) in the same study domain but showed results only in Figure 19 and 20. We therefore do not see the need to make the original model the centre of our study. Actually, Figures 19 and 20 clearly suggest that the original model is indeed not a good choice especially for the amount model. We started with the hypothesis that anisotropy is required in the complex terrain and proved that it is not the case in the European Alps in such a small region. We have explicitly stated on L515 that Iso-KED is superior to the fully anisotropic case. This is also a choice of how the results are presented. One can start with the hypothesis that the original model is sufficient in the complex terrain and if not then propose the necessary changes. Or one can start with the hypothesis that the anisotropy is needed in the complex terrain which is a natural assumption and assess whether the hypothesis is correct or not. We have chosen the latter approach. However, we will mention the information on the complexity of the model in the revised manuscript.

2) Almost no information is provided about the spatial interpolation of model parameters (using different versions of Kriging, in particular KED), which is however a critical step of the model set-up. I think that it is inevitable to give more details about the Kriging step, including: (i) mention if a nugget term is used, and if yes what is the nugget contribution to the total variance, (ii) mention which variogram model is used, which method is used to fit or infer variograms, and maybe show some examples of adjusted variograms, (iii) prove by some data analysis than model parameters are (linearly) dependent of altitude to justify the use of KED, and finally (iv) display maps of kriged model parameters.

Response: We agree that the information provided on the spatial interpolation of the parameters may have been on the low side — this was due to the fact that there are quite a few parameters (and their number grows with the number of data points and is not fixed as in a parametric model). Based on the reviewer's comment we will add the following information to the revised manuscript.

i) will be mentioned in revised manuscript.

ii) An appropriate information will be added in the revised manuscript.

iii) Please see the plots below in Figure R1. The use of the KED is natural in the complex terrain. It has been used in the literature for precipitation interpolation in the mountains. We will add this information to the revised manuscript.

iv) Please see the plots for shape and scale parameters in Figure R2 and Figure R3. We will include this information in the revised manuscript.

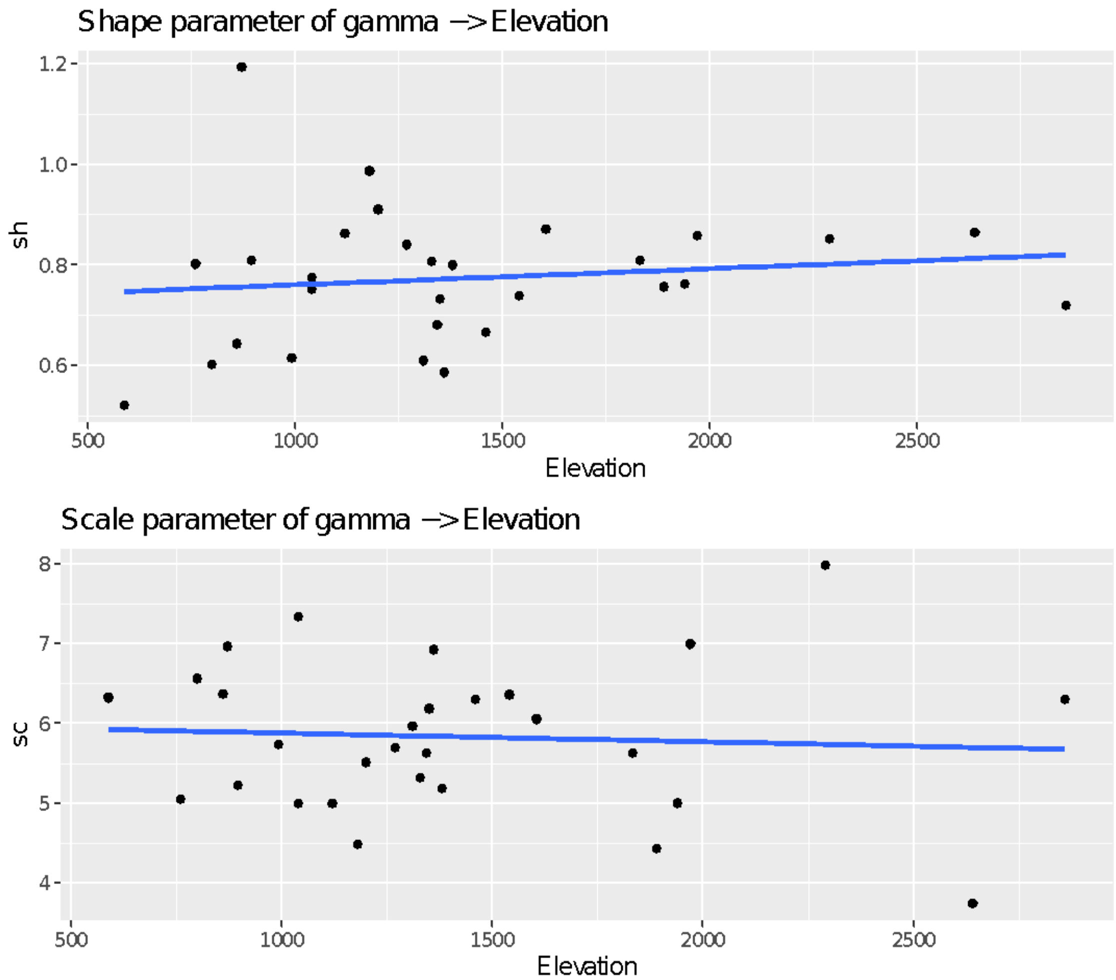

Figure R1: The shape parameters (top) and scale parameters (bottom) of 29 sites plotted against their elevation

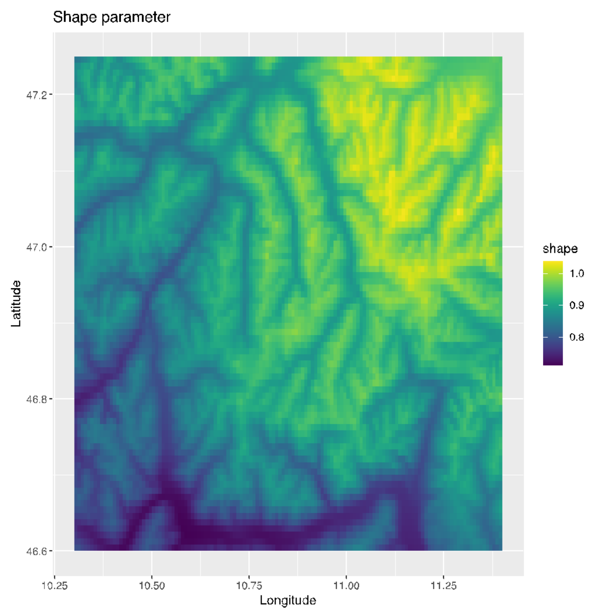

Figure R2: Interpolated shape parameter

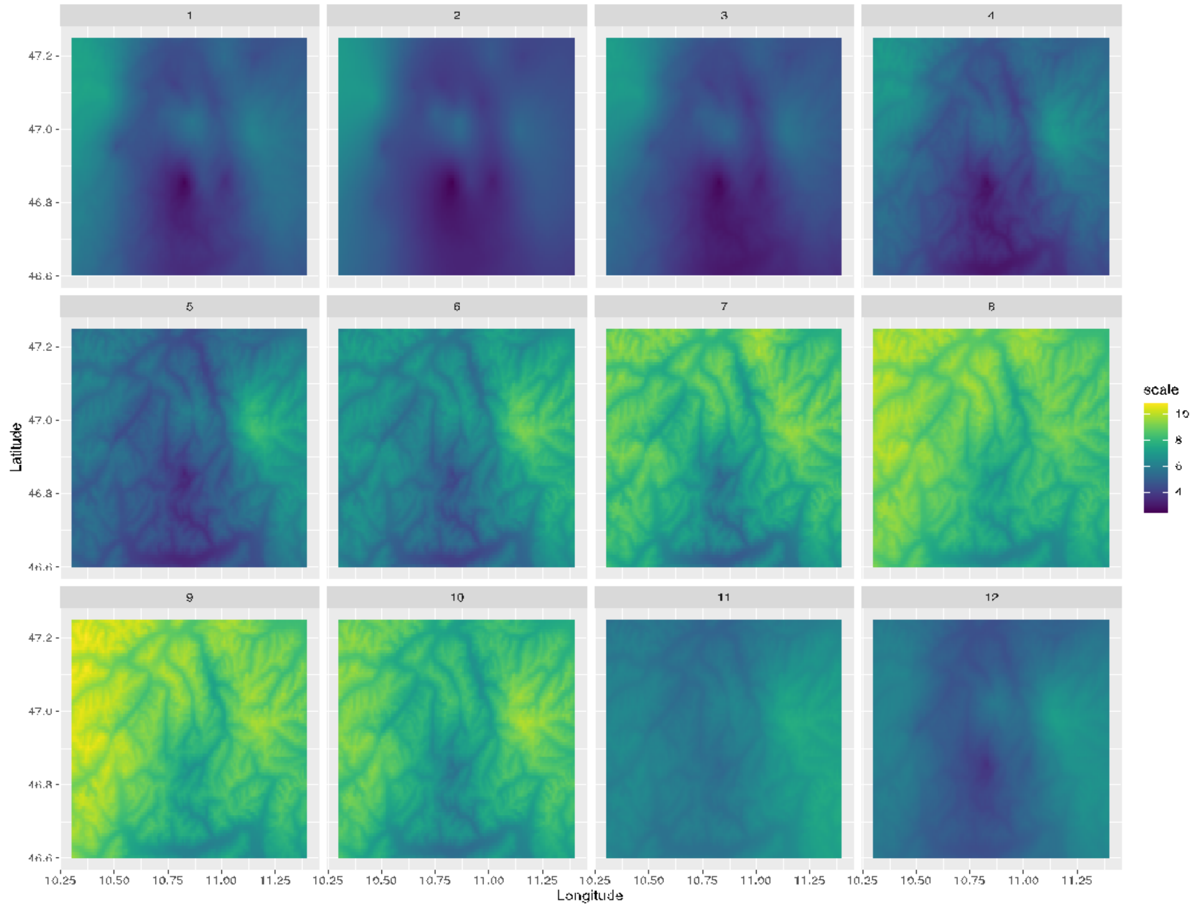

Figure R3: Interpolated scale parameter for each month

3) A cross-validation is missing to evaluate the performance of the interpolation of model parameters by Kriging, and to assess the density of stations required for model calibration.

Response: We thank the reviewer for the suggestion to add a cross-validation. We will include cross-validation in the revised manuscript. Following the 'logic' of our approach (i.e., to show exemplary results, such as in Figures. 2, 3, 4 where a 'good', a 'topographically exposed' and a 'bad' station are shown, we have decided to perform the cross-validation by holding those three stations out.

However, the other suggestion by the reviewer, i.e. the assessment of the density of the stations required for model calibration is out of the scope of this study. We decided to include as many sites as possible and with any (high-quality) data, the more the better. Moreover, the observed data display that the spatio-temporal structure of precipitation is highly variable in mountainous regions, and in that regard, we believe that such an assessment would not be feasible.

4) The proposed model fails to reproduce precipitation at one of the three stations selected for illustration (i.e. Prutz), and according to Fig 9 also performs poorly for Dresdner Hütte, Kühtai, Nauders and Sankt Leonhard 2 (i.e. 17% of stations in total), and the reasons why the model fails at these locations are not

investigated in enough details. The only explanation given for Prutz is that this station exhibits 'precipitation characteristics quite distinct from those of neighbouring (or more distant) stations', but nothing is said about why a model calibrated at station locations is unable to reproduce precipitation at the exact same locations. For me it is very probable that the Kriging of model parameters introduces errors when clustered stations exhibit distinct statistics (which is the case for Prutz, that forms a cluster with Ried in Oberinntal and Fendels) - and possibly also where sharp gradients of precipitation occur (Nauders or St Leonhard im Pitztal 2) as well as at the edges of the study area (Marienberg, St Martin or Kühtai) - but one cannot see it in the present manuscript because the Kriging step is completely overlooked.

Response: We thank the reviewer for the quite detailed comment and substantiation with individual sites. We think, however, to have provided quite an extensive discussion on the problematic sites in Section 3.1 – and it is that in the complex terrain precipitation is highly variable in space and time and our data displays this behaviour. L588-596 explain the reasons for the bad performance at those sites. We have supported our argument by showing correlations in the observed and synthetic data at the 29 sites, which is shown in the Figure S4 in the Supplement. On L324–L328 (and also L588–L596) we explain the reasons, i.e. that the use of the same set of covariates at all the locations is the likely reason that the seasonality at those distinct sites is not well captured.

The 'probable reason' as pointed out by the reviewer has also been mentioned – see the explanation on L427–L431, which states that the kriging interpolation of the model parameters is the reason for the bad performance of the model [in certain parts of the domain]. We even explicitly discuss the case of the station St. Martin (L427–L429). We haven't discussed every station's performance individually in the manuscript, but collectively throughout the manuscript, we have mentioned the reasons for the bad performance at those sites.

5) Regarding the problem at Prutz station, I would be interested in seeing (in supplementary material for instance) the raw time series for Prutz, Ried in Oberinntal, Fendels and Ladis to be sure that the modeling problem does not simply originate from instrumental errors at Prutz station... the fact that this station has so much differences with Ried in Oberinntal which is located 1km apart, at the exact same altitude, and with similar neighboring topography is very surprising to me.

Response: Indeed, this was surprising for us too, but the data are from different data providers so the time of data recording, the method of data collections etc. may differ which results in data being different. We have carefully selected only those stations, which have gone through proper data quality check. Before selecting the stations for our study, we contacted each individual service provider and after getting an approval from them, we allowed only the high-quality stations in our study. Also, regarding Prutz and other stations showing very different characteristics than the surrounding stations, we again contacted the data providers to know the reason for the data being different as we also suspected that it could be due to instrumental errors or any other problem with the data. However, it was confirmed by the data providers that the data provided by them are of highest quality — and it should be not so surprising to see such behaviour in the mountainous region. To this end, we had the choice to either exclude those stations from the study or keep them as they are. We believe that 'throwing away' such stations from the study just for the reason that the data are not pleasant to the model is not a good scientific practice. Hence, we allowed those stations — and decided to show them as 'problematic'.

Minor comments:

In addition to these major concerns, I also have some minor comments detailed hereafter. Line and figure numbers refer to: https://hess.copernicus.org/preprints/hess-2022-21/hess-2022-21.pdf.

1) that the value of 0.1mm corresponds to the lowest resolution of the rain gauges used in the case study. But I was not able to find this information. Please mention the resolution of the rain gauges in section 3.1.

   Response: We have adopted the threshold of 0.1 mm, which is the threshold typically used by the data providers and is also a common practice in the field. As per we know, a rain gauge does not have a 'resolution' as a thermometer.

2) L 125-128 and Eq4: Making the covariance of the occurrence altitude dependent seems a relatively arbitrary choice, and leads to a complex (and hardly tractable) model. This must be supported by some preliminary data analysis, showing that such altitude dependent covariance of precipitation occurrence actually exists in your study dataset. In addition, I wonder if including station altitude as a covariate in the vector Xo would not be a more convenient modeling choice. This may lead to a simpler model (also easier to 'validate' using the AIC/BIC model selection procedure introduced L260).

   Response: The altitude dependence of the parameters has already been raised in major comment (2). Please refer to the discussion there.

   Station altitude as a covariate is indeed an 'obvious choice' as a covariate — at first sight. We originally also have considered including altitude as a covariate in the model, but at each location, this is just one number, i.e. a constant, while all other covariates are time-dependent, and the altitude doesn't change with time, so how does that one constant value influence the model? Or can it even be considered as a covariate if it is just a constant? Even if we allow altitude as a covariate, adding one covariate in the model would add one parameter at each location, so in total 29 parameters which after interpolation would increase to as many parameters as the number of grid points. In our view, the elevation dependence suggested by us makes the model simpler compared to considering elevation as a covariate.

3) Eq 5 (L156): It is not clear to me how this equation derives from Eq 2 and Eq 4 considered at a single site, and using previous day's occurrence as a covariate. Could you give more details (maybe in supplementary or a reference)?

   Response: Eq. 5 is not derived from Eq. 2 and Eq. 4, but it is stated that at individual location the model reduces to logistic regression. The idea is to use the logistic regression at individual stations with the selected covariates. One of the covariates is previous day's occurrence. This covariate is obtained from the observed data. Hence, when the logistics model is fitted at a station, the only location dependent value is the regression parameter (β). Once the regression parameters are obtained at the 29 stations, a Gaussian process with the mean function as in Eq. 2. and the covariance structure as in Eq. 4 is implemented over the region. Such models, where previous day's occurrence is considered as a covariate already exist in literature. We already have given appropriate references on L148–L149 in the manuscript.

4) L166: Using KED with altitude as drift to interpolate regression parameters means that these parameters are all (linearly) correlated with altitude. This should be shown by a data analysis. In addition this leads to a complex model, and I wonder (as in my comment about Eq 5) if it would not be easier and equally effective to include altitude in Xo, and test if altitude is a relevant covariate.

   Response: Please see the answer for major comment (2) and minor comment (2).

5) Eq 6 and paragraph L179-188: the mean function of the latent process \mu_{A}, and the parameters of the Gamma distribution are space and time dependent, and in addition are interpolated by KED. This is a lot of parameters! You should quantify and acknowledge the complexity of your model. I'm not sure that so much model complexity (and degree of freedom) is necessary, but I'm ready to be convinced by a careful data analysis showing that all these dependencies are indeed present in your dataset. If i'm not mistaken, in Kleiber et al (2012), the mean function of the latent process Wa used to model precipitation amount is fixed to zero (and not regressed on covariates with regression parameters additionally interpolated by KED), which makes the model of precipitation amount way simpler. Such addition of complexity must be supported by data.

   Response: Indeed, the model has a large number of parameters because a Gaussian Process is a non-parametric method. As the number of data increases, the number of parameters also increases, which is different to the case of parametric models where the number of parameters remains fixed with increasing the data size (this is also the reason why the 'mean' and the 'covariance' are referred to as functions in Gaussian process modelling). We think that our model is equally as complex as that of Kleiber et al (2012). It has only two parameters more than that of Kleiber et. al (2012) — which are the two range parameters in the vertical direction due to the inclusion of the elevation dependence. Kleiber et al. (2012) presented the ***general framework*** for the multi-site gridded model but tested the model only for **multi-site** data generation and not for **multi-site gridded** data, i.e. it was tested at stations with observations and the interpolation part was not carried out and, in that sense, it was indeed less complex than our model. Also, please refer to our response to major comment (2) for the comment on dependencies.

6) Figure 1: I think station 22 (Pitztaler Gletscher) should be in red instead of station 24 (Obergurgl).

   Response: We thank the reviewer for pointing out this mistake. Earlier, we selected Obergurgl as a representative station for high mountain stations but later we changed to Pitztal Glacier and mistakenly updated the wrong figure.

7) L218-220: Could this extreme value be an outlier? Prutz is surrounded by very nearby stations (few kilometer apart) in all directions, and none of them measure more than 35mm this day (compared to

156mm in Prutz). I agree that summer convective rains can be very localized, but I'm still surprised by this observation, and I think this requires more investigation. And this also rises concerns about the quality of data at Prutz.

Response: Please refer to our response to major comment (5).

8) L223-224: different precipitation features at St Leohard 2: more details are needed to ensure that this station operates properly (same comment for Prutz).

   Response: This station is operated by the Austrian Weather Service (ZAMG) and is properly quality controlled. Also, please see our response to major comment (5).

9) L229-232: I do not understand this paragraph. How this 7-days window increases the amount of data? And how simulation will add robustness to the observations? This is very unclear. Maybe because I do not understand what you name data.

   Response: Since have only 30 years of observed data, there are only 30 values for each day (for example, January 1 has only 30 values). There are many days, for which all the 30 observations were dry, for example, so that the determined daily probability of dry days was 1. This would also then imply that every simulated realisation of those dates is a dry day, which is not realistic. There is always some — albeit possibly small — probability of precipitation on a particular day for the climate in the Alps. Similarly, for transition probabilities also some days (i.e., dates) were displaying the probabilities either 1 or 0. Hence, by allowing a 7-days window centred at the day of interest, we were able to escape the 1/0 values of the observed probabilities. We will explain this more precisely in the revised manuscript.

10) Eq 7 and Eq 8: A lot of temporal covariates are tested, but only one covariate linked to atmospheric circulation. Why such unbalance? I agree that NAOI can influence precipitation in the Austrian Alps, but it is definitely not the only covariate one can think about. In my opinion you should test other climate covariates, or justify why you think NAOI is enough.

    Response: We agree that other covariates should have been incorporated in the model, however, at the time when we designed the experiment, we considered only NAOI as we wanted to see how NAOI alone is able to simulate the precipitation, especially during the months when NAOI is said to be linked with the precipitation. In a future study we will assess the impact of other covariates related to atmospheric circulation.

11) L288-291: More information about the APGD dataset is required to allow the reader to understand the main features of this reference dataset. In addition, it should be mentioned somewhere that APGD is not a perfect reference. Finally, the use of a 5km spatial resolution reference does not make it possible to assess the fine scale patterns generated by your 1km resolution model. Hence, all fine scale patterns seen in Fig 11 and Fig 14 may only be artifacts of using KED driven by altitude. This must be mentioned somewhere in the paper, or Fig 11 and Fig 14 should be aggregated at 5km resolution to avoid over-interpretation of the results.

Response: This is indeed an important suggestion. We will add more information about the APGD data in the revised manuscript. As per the figures showing 1 km resolution of synthetic data, we believe to show the generated data as they are. We respectfully disagree with the comment that Figure 11 and 14 should be aggregated to 5 km because this information would then be lost.

12) Fig 5 and Fig 6: Frequency -> Frequency (%).

Response: We thank the reviewer for pointing this out. However, it is not in %, but it should be mentioned that it is frequency 'per year'. It has been mentioned in Figure 20, but needs to be mentioned in other figures, too. We will make the necessary changes in the revised manuscript accordingly.

13) Fig 9: Very useful figure. It could be improved by: (1) using the same range of values for abscissa and ordinates, and for all stations. (2) Add station Id in addition to station names, and maybe order stations according to their Id (as in Table 1). (3) Mention in caption which quantiles are used (percentiles I presume).

Response: We thank the reviewer for this comment. This will indeed improve the readability of the figure. Necessary changes will be made in the revised manuscript.

14) Section 4.2: It would also be interesting to display q-q plots of areal daily precipitation amount.

Response: We included the q-q plot of the areal daily precipitation amount initially, however given the length of the article we had to exclude it. Also Figure 13 in the manuscript gives the flavour of how the plot looks like.

15) L420: The influence of topography on precipitation occurrence (and also amount) may be an artifact of the model. I don't say that it is the case, but just that you do not prove it in this paper.

Response: Indeed, when modelling environmental data, one should always consider the possibility that the result could be a model artifact. In case of precipitation in mountainous terrain, however, it is pretty well established that orography does not only receive precipitation — but it also triggers precipitation (cf. 'orographic precipitation' (e.g. Rotunno and Houze, (2007)) which, in a mountain range like the Alps, is responsible for the vast majority of precipitation amount). Even summer precipitation is triggered (convective initiation) by processes that are related to orography (slope, exposition). If a model then returns an orography influence, we think that this rather reflects a model quality than an artifact. Proving this, however is not overly simple.

16) Section 4.3: Results in this section prove that the anisotropic model is not relevant in the present case. They also show very few improvement when using KED instead of OK. When considering how much model complexity and arbitrary hypotheses about orographic precipitation enhancement are added with KED I wonder if a direct application of the model of Kleiber et al, (2012), with maybe altitude as a covariate (to be selected by BIC/AIC) would not be a better option.

Response: We agree that the anisotropic model has only a negligible/neutral impact on model performance (it does have a slightly positive impact on the frequency of dry spells, though), while the use of KED instead of OK has a small but consistent positive impact. We respectfully disagree, however, about the added complexity (which is nil) and the arbitrariness of the assumptions (see our responses to major comment (1) and minor comment (15) above). In our model comparison (Figure 19) we have also included the original model of Kleiber et al. (2012) — thus allowing to assess the impact. In this sense, we have demonstrated that using the original model would not be a better option (even if the difference is not overwhelming).

Furthermore, *even if the model had zero impact*, we think that demonstrating that an obvious extension for complex mountainous terrain, i.e. to include terrain information in one way or the other (i.e., the vertical dimension), does not improve model performance, would be a scientifically valuable contribution. Hence, other researchers could spare their time to 'try out' this obvious extension and focus on other possible improvements. If the impact is positive but not very large (as in the present case), we hope to pave the ground for further studies which can make this impact even larger.

17) L532-533: It would be interesting to compare the original and the extended model in a schematic, including the number of parameters involved.

Response: Please refer to our response to major comment (1) and minor comment (5).

18) L678-827: many typos in the doi of the references.

Response: We thank the reviewer for pointing out the typos. They will be corrected in the revised manuscript.

References:

1. *Simultaneous stochastic simulation of daily precipitation, temperature and solar radiation at multiple sites in complex terrain DS Wilks, Agricultural and Forest Meteorology 96 (1-3), 85-101.*

2. *Wilks, D. S.: A gridded multisite weather generator and synchronization to observed weather data, Water Resources Research, 45, https://doi.org/10.1029/2009WR007902, 2009.*

3. *Rotunno, R. and Houze, R. A.: Lessons on orographic precipitation from the Mesoscale Alpine Programme, Q. J. Roy. Meteor. Soc., 133, 811–830, https://doi.org/10.1002/qj.67, 2007.*