Dear Dr. Leijnse,

We are very grateful for your constructive review of our paper. We will try to take advantage of your advice for improving the manuscript. We respond below to your comments item by item:

- The introduction should be clarified. It should be made clear how this paper contributes to our understanding of the use of geostatistical methods for rainfall interpolation, and what sets it apart from other work that has been done in this field. I suggest that the authors move the discussion of findings in the literature in Section 4 to the introduction. In this way it can be made clear what results have already been presented, and what the contribution of this paper is. The second question that the authors state they will attempt to answer in the introduction ("With the raingage in which position?", p. 7387, lines 9-10) is not tackled in the remainder of the paper. This should also be clarified.

This paper contributes to the understanding of the use of geostatistical methods for daily rainfall interpolation whereas other previous studies focused on monthly and annual time step. Difference in time step can contribute to different result of the interpolation methods.

The revised manuscript will include your recommendation and give more clearly what results have already been presented. For better clarity, the second question in the objective will be removed from the revised manuscript because it not tackled in the remainder of the paper. It will be replaced by "how does the raingages' density affect the interpolation results at daily time step?"

- Using many different variogram models to estimate the optimal variogram will in principle yield better results than when only one type of variogram model is used. However, it also makes the interpretation of the results presented in this study more difficult. It would be interesting to know how the optimal variogram models and their parameters change in time. Is there a single variogram model that is often the optimal one, or are they all used equally frequently? And how large is the improvement over using only, say, the spherical model? While these are all interesting questions, I feel that the paper would gain much in clarity if only one variogram model were to be used, especially given the main objective of the paper. Looking at the example of Fig. 2, the spread of the points around the different fitted variograms is much larger than the difference among the variogram models. This leads me to believe that using these seven different variogram models will not greatly improve results.

We agree that may not improve the result for rainfall based on Fig.2. We will clearly state this in the revised manuscript. However, using these seven variogram models can resolve the negative result of the ordinary kriging as raised in the following comments. In the revised manuscript, we can report how the frequency of these seven models is over the 30-year daily rainfall.

A discussion on how the optimal variogram models and their parameters change in time and how the result for single model is seems to be out of the scope and would not give more clarity, in our opinion.

- Zero rainfall is not mentioned. This deserves some attention, as it is one of the points that make rainfall modelling and interpolation difficult.

For some days where zero rainfall occurred at all stations, we supposed that there is no rain in the whole catchments, and thus no interpolation for these days. We will try to make more detail about zero rainfall in the revised manuscript.

- On p. 7393, lines 5-6, it is stated that "kriging can lead to negative estimates. Thus, the variogram model was changed to another one until the kriging estimates were all positives." The kriging estimates can only become negative for some types of kriging (for example, it is not possible with ordinary kriging, where all values are between the minimum and maximum data values). And it is entirely possible that all variogram models will yield negative rainfall at some point in space for for instance kriging with external drift. Was the variogram model changed if one of the kriging methods yielded negative results somewhere, or was the model only changed for the kriging methods that yield negative values? And what happens if all variogram models yield negative results with a particular kriging method? The "C+10000" statement in Fig. 1 does not help me understand how this is accomplished.

It is possible that ordinary kriging leads to negative estimates (Deutsch, 1996). Negative weights in ordinary kriging occur when data close to the location being estimated display outlying data. Depending on the variogram and the data values, the negative weights can be significant. Also, negative weights when applied to high data values may lead to negative and nonphysical estimates. For this study, the kriging methods are computed separately. If a variogram model yielded negative rainfall somewhere, the variogram model changed only for that kriging method. Using these seven variogram models in this study allowed us to avoid negative estimates for ordinary kriging. However, all variogram models yield negative result with UNK and KED at some days. We modified the Kriging system to disallow negatives weights, based on the a posteriori correction described by Deutsch (1996). We will clearly state this in the revised manuscript.

The “C+10000” statement is just to change the variogram model. We will clarify in the figure.

- I find Figs 5 and 6 impossible to interpret. It would be much better if the DEM would not be displayed in these figures (it can be seen in Fig. 3), and that the rainfall fields would be presented as color maps (such as the ones for the Thiessen interpolation).

We will present the rainfall as colour maps in the revised manuscript

- Results for which interpolation technique are presented in Fig. 7?

Fig. 7 shows the average RMSE of all methods. We will separate the graphic for each method in the revised manuscript and explain more clearly the legend of the figure.

- The fact that the Robertville (misspelled as “Rotbertville” on p. 7400, line 17) rain gauge produces such a large RMSE in Fig. 7 compared to the other gauges is striking. Can this be explained in some way? What can be concluded from this regarding the robustness of RMSE of 7 gauges as a measure of the quality of the interpolation technique? As stated in the conoclusions (p. 7404, lines 11-14), a fuller assessment of the strengths and weaknesses can be given using a more thorough cross-validation. Given the fact that this is central to this paper, I think that the authors should consider including such analyses in this paper.

The Robertville raingage produces largest RMSE compared to the other gages. This can be explained by the effect of the high elevation and the position of that raingage located at the extremity of the zone. The least errors occur over flat plains and the largest over mountainous area (Basistha et al., 2008).

As stated in section 2.3, the evaluation of such a comparison of different interpolators was usually made by cross validation which involves temporarily discarding data from the sample data set; the value at the same location is then estimated using the remaining samples. Most of work use cross validation technique with monthly or annual time steps. The sample size from the cross validation is the number of sample data (number of existing raingages). In our case, a complete cross validation for the daily time steps of 30-year precipitation was not realistic. Therefore, seven raingages in the study area were randomly selected to be used for validation, in view of the fact that the existing observed daily rainfall series of these seven raingages provided a large enough sample size. Moreover, these sevens raingages are spread over the catchment area and they also cover the whole elevations’ range of the catchment as seen in Fig.3.

Minor comments:

- In the abstract, what is meant by “the majority of existing geostatistical algorithms are available only for single-moment data”, and how does this relate to the rest of the paper? There are ar least two open-source packages (gstat and geoR; both based on R) that offer a wide range of geostatistical methods.

That statement means that the existing algorithms are available only for single moment data like elevation. We need to add or develop complement algorithms to automate the computation for time series data. We agree this statement is not much important and does not much relate to the rest of the paper. We will remove this sentence from the abstract.

- On p. 7387, lines 1-2, it is stated that “a daily time step is optimal for an understanding of the soil-plant-water relationship and long-term catchment management simulation”. I don’t think that this statement is true in general. It may hold for certain aspects of the soil-plant-water relationships and catchment management simulation on certain spatial and temporal scales, but it is certainly not generally true. I therefore strongly suggest removing this statement, or providing ample evidence that supports it.

We will remove this statement from the revised manuscript.

- In Eq. (1) (p. 7388), it should be made clear that λ_i is a function of g.

λ_i is not a function of g. g is the interpolated point (grid).

- On p. 7390, it should be clearly stated that statistical homogeneity is assumed.

We will add the statement in the revised manuscript.

- On p. 7390, lines 8-10, the authors mention that the distances are binned. How large are these bins? And does the bin size depend on the number of gauges used?

The bin size depends on the number of gauges used. The bin size is 5km when we use all available raingages. We will make detail in the revised manuscript.

- The distance between two points h is a vector in this paper. For simplicity, the authors could consider using a scalar instead, because they use the assumptions of statistical stationarity and isotropy.

We will use scalar instead of vector in the revised manuscript.

- The logarithmic variogram is one I have never heard of. This may be because it is minus infinity close to h = 0, which is impossible to interpret in terms of variances. I therefore strongly suggest removing this variogram model from the analyses.

This model is often used for describing variables regularized by a sampling support, which implies that the infinite variance problem disappears (Chilès and Delfiner, 1999; Diggle and Ribeiro Jr, 2007). In reality, the logarithmic model is negative for small values of h, which occurred rarely in our study. We are sorry that we can not remove this model from the analyses that are already done.

- For the power model, why is the parameter θ_2 constrained to be smaller than 2?

The parameter θ_2 is restricted between 0 and 2 in order that the model is a conditional positive definite function and then a permissible variogram model (Pardo-Igúzquiza, 1998).

- For the spherical and penta-spherical models, the last two conditions overlap. The last condition should be $h > \theta_2$, and not $h \neq 0$.

We will correct accordingly.

- On p. 7392, lines 5-7, it is stated that the most common method for fitting semivariogram models is by eye. Can the authors provide some references in which this has indeed been done? In most of the work that I'm aware of automatic fitting has been used.

The usual practice is to fit the models manually to experimental variograms". This is done in the SGeMS (Stanford Geostatistical Modeling Software)

- On p. 7397, line 4, it should be mentioned that Pearson's coefficient is computed from the correlation between rainfall amount and elevation.

We will modify accordingly.

- In Section 3.1, the authors could consider using a table to summarize the annual mean rainfall at the highest and lowest points of, and the average over the catchmant. Other values that could be added to this would be the minimum and maximum values over the catchment (which should be similar for e.g. inverse distance weighting and ordinary kriging).

We will make a table accordingly.

- In the analysis effect of the number of gauges, the authors could include a discussion of the results presented by Berne et al. (2004, J. Hydrol., 299, 166-179).

We will include accordingly.

- Are the points in Fig. 8 the means of the seven gauges? and are the error bars plus and minus the standard deviation of these seven gauges?

Yes, we will clarify the legend of that figure.

References:

Basistha, A., Arya, D.S. and Goel, N.K.: Spatial Distribution of Rainfall in Indian Himalayas – A case study of Uttarakhand Region. Water Resour. Manag. 22, 1325-1346, 2008.

Chilès, J. P., and Delfiner, P.: Geostatistics: Modelling Spatial Uncertainty. John Wiley & Sons, New York, 1999.

Deutsch, C.V.: Correcting for negative weights in ordinary kriging. Comput Geosci, 22, 765-773, 1996.

Diggle, P. J., and Ribeiro Jr, P. J.: Model-based Geostatistics. Springer Series in Statistics, Springer Science +Business Media, LLC, 228 pp., 2007.

Pardo-Igúzquiza, E.: MLREML4: A Program for the inference of the power variogram model by maximum likelihood and restricted maximum likelihood. Comput Geosci, 24, 537-543, 1998.