Dear Dr.Toth,

We appreciate the helpful and constructive comments and our responses to the comments are as follows:

The paper presents a comparison of procedures (based either on clustering or on calendar) for calibrating and validating a rainfall-runoff model with parameter sets that depend of the period of the year.

Since the analysis of the seasonal variability of the dominating hydrological processes is a crucial topic, and the importance of keeping such variability into account when calibrating a rainfall-runoff model is often neglected in both research and hydrological practice, the addressed theme is of broad interest for the HESS readers.

The application to only one case study (even if with adequately long time-series) is certainly a serious limitation of the work, as highlighted also by the Editor, Ralf Merz, in addition a number of clarifications on the work are needed.

Reply: Agreed. We will apply the proposed methodology to additional two catchments (with different geophysical characteristics) to explore its generalisation capabilities.

The main concerns I have are:

(1) The final objective of the procedure is not clear.

The method is in fact not supposed to be used for choosing a different parameter set in real-time (like some examples in the cited literature, where the hydro-meteorological conditions PRECEDING the forecast instant are used for the classification/clustering and for the choice of the most adequate parametrization to be used for real-time forecasting), but it has to be used off-line, since the hydro-meteorological similarity is identified a posteriori in the clustering technique here applied. In fact, in order to identify the cluster to which a specific time instant belongs, the future rainfall (and temperature) values are needed in input.

The sub-annual calibration scheme based on FCM is therefore applicable only a posteriori; such drawback, in addition to its complexity, is not justified by the results, since the calendar grouping performs equally well than the best FCM.

Reply: We believe that the proposed methodology can be used in both off-line and real-time conditions. During a real-time situation, both temperature and precipitation for the event to be modelled could be obtained through real-time observation and weather nowcasting. However, we apologise for this misunderstanding due to the unclear description about the operational use of the proposed methodology, and will amend this in the revised manuscript.

Maybe the authors should elaborate more on the differences in the simulation results in comparison to the traditional approach, possibly in order to improve the model structure?

Reply: The simulation results of the proposed methodology have been compared with that of the traditional approach (Figure 7), which shows great improvements. This study aims to use more appropriate calibration schemes to compensate the deficiency of the model structure and invariant model parameters when considering seasonal variations of the catchment. How to improve the model structure may be explored in the future research.

The final aim of the study should be clearly stated in the introduction and a deeper interpretation of the results is needed in the result/conclusion sections.

Reply: The final operational aim of the study is to build a more appropriate hydrological model for water resource management (e.g., river flow extension by rainfall runoff modelling) or real-time flood forecasting (via data assimilation). Because hydrometeorological features/indices are considered, this study may also be useful for future rainfall-runoff modelling under climate change. We will add this in the revised manuscript as well as a deeper interpretation of the results.

(2) The reasons for the choice of the variables used in the clustering technique are not clear: of course many other hydro-meteorological features may be needed to appropriately identify the peculiarity of each subperiod/ 'season'.

Reply: Agreed. We will explore a wider range of hydrometeorological features and justify their use in the revised manuscript.

(3) It is important to underline that an important problem in using the model with changing parameter sets is the fact that the model is a continuously-simulating conceptual one, that needs all the previous simulation values (depending on the specific parameter set) to update the state variables. This also implies that when switching from one subperiod to the following one (e.g. at the end of the month and beginning of the new one) there may be some discontinuities in the simulated streamflow, due to the change. Such aspects are one of the main issues in the use of time-varying parameters in rainfall-runoff modelling and it's not very clear in the presentation.

Reply: Many thanks for highlighting this important discontinuity issue in changing model parameters. This is caused by the parallel running of models with different parameter sets. If the models are run in series, there will be no such discontinuity problem because the subsequent flow is mainly derived from the antecedent flow and part of the new effective rainfall. We will provide a better clarification in the revised manuscript.

Specific comments:

p. 1, 32-35: please be careful with the use and meaning of the terms 'stationarity' and 'climate change': see, Lins' note (2012) on the WMO website: http://www.whycos.org/chy14/download/file.php?id=13 in particular, in this case, the study does not address climate change, but interannual variations, so I don't think such digression (especially being a very complex and debated issue) is needed.

Reply: Many thanks for highlighting this important issue. We will pay more attention to this problem to avoid the digression in the revised manuscript.

Section 3.2: a flowchart or a diagram explaining the splitting and use of the different time periods would be very useful to understand the proposed approaches.

Reply: Agreed. We will add this diagram.

p.4: ll 30-31: explain how the months are merged: the 6-months periods are only the Jan to June one and the July to December one, or other 6 consecutive months periods have been analysed?

Reply: The 6-month periods are only the Jan to June one and the July to December one. This will be further explained in the revised manuscript.

p. 4, ll 36-37: specify that the clustering technique was applied for all the time-scales (1 month, 2 months, 4 months and 6 months).

Reply: Agreed. We will specify it.

p. 4, ll 37-41: please add more information on the selection of the input variables: which other variables have been considered, how you have chosen such five ones, etc; (see point 2) above)

Reply: Agreed. We will explore a wider range of hydrometeorological features and justify their use in the revised manuscript.

p. 5, ll 60-62: add that the description of the steps for identifying Kopt is reported in Section 4.1.

Reply: Agreed. We will add this description.

p.5, ll 68-69: as said above, explain that the problem in using the model with changing parameters is the fact that the model is a continuously-simulating conceptual one, that needs all the previous simulation values (depending on the specific parameter set) to update the state variables: for this reason the model has to be run for the entire observation period and not only for the analysed sub-period.

Reply: Agreed. The current calibration process is based on the parallel model run, and we will replace it with the series run for better modelling continuity. This will be clarified in the revised manuscript.

p. 5, ll. 77-79: more information on the optimization algorithm are needed and in particular either add the definition and meaning of 'nlminb', or remove such detail.

Reply: Agreed. We will provide more information on the optimization algorithm.

p. 5, l.84- 85: explain better how dealing with discontinuities in the simulated streamflow values when going from one period to the following one (see comments above).

Reply: This issue is caused by the parallel running of models with different parameter sets. If the models are run in series, there will be no such discontinuity problem because the subsequent flow is mainly derived from the antecedent flow and part of the new effective rainfall. We will replace the parallel run with the series run for better modelling continuity and provide a better clarification in the revised manuscript.

All section 4.1 must be thoroughly revised and reworded since it's very confusing and the utility of using the cluster validity index is far from demonstrated (in the only information referring to it, Fig. 4, the values of V_{XB} seem to fluctuate randomly):

- from ll. 90-93 and ll. 100-06 it is not clear how, eventually, the optimal number of clusters is identified, considering both the simulation results and the validity index;

Reply: The purpose of this section is to verify the utility of the index V_{XB} in identifying the optimal number of clusters for FCM algorithm through comparing with the value of NSE under different numbers of clusters. We apologise for this misunderstanding due to the unclear description and will amend this in the revised manuscript.

- l. 92: with 'according' you mean 'also considering'?

Reply: Yes, 'according' here means 'also considering'.

- ll.96-97: comment also on the results for the validation period.

Reply: Agreed. We will comment on the results for the validation period.

- Overall, the text does not report the final chosen value for Kopt at monthly time-scale, and most importantly, nor the text nor Table 2 report the final number of clusters chosen for each of the other time-scales: in table 2, both possible values of Kopt are shown for each time-scale. And the paper does not provide any information on the reasons for the choice of the number of clusters for all the other time-scales (2-, 4-, 6—months periods), since Figure 4 (in addition to reporting difficult to interpret results) refers only to the monthly time-scale (even if this is not stated in the caption).

Reply: Agreed. Although line 204 to 207 (Therefore, the cluster validity index V_{XB} in identifying the optimal number of clusters has a satisfactory performance and we used the cluster validity index V_{XB} to recognize the optimal number of clusters in this study) reports the cluster validity index V_{XB} is used to recognize the optimal number of clusters, we did not report the final chosen value of Kopt for different time scales. We will provide more information on the choice of the number of clusters for all the time scales and revised this section. We will also consider other better ways to define the number of clusters for all time scales, since the results of the cluster validity index are not good enough in this study.

p. 6, l. 11: why Fig. 5 refers only to the years 1990-1995? The calibration period is 1990-2000.

Reply: The calibration period (1960-2000) is too long to exhibit in the figure, so we only choose the period 1990-1995 to clarify the difference between the distribution of groups classified by the CBG method and clusters from the FCM algorithm for different time scales.

Fig. 5 does not provide any information on the relation between clusters and seasons: you should find a way to show this information and to analyse it deeper.

Reply: Agreed. We will improve Figure 5 to better present the information on the relation between clusters and seasons and analyse it deeper.

p. 6, l. 12: with 'sub-periods in one group' you mean the 'sub-periods in the same calendar position'?

Reply: Yes. We will state it more clearly.

p.7, l. 25-26: this result is hardly surprising: the first FCM is based on four over five input variables that depend only on rainfall, whereas temperature has a clear annual cycle, well-reproduced by a calendar method. And in fact Fig. 6 is not very useful, since all the plots show the same pattern in the rainfall variables. Probably more/different climatic variables would provide more insights in the hydrological behavior of the catchment during the year (see point 2)

Reply: Agreed. We will explore a wider range of hydrometeorological features and justify their use in the revised manuscript.

Section 4.3 (ll 36-49): adding a second FCM classification procedure based on different climatic variables as a sort of 'second thought' experiment in the results section makes the overall work difficult to follow: please introduce also such FCM algorithm earlier, in section 3, together with the other two techniques, and not here when discussing the results.

Reply: Agreed. We will reorganize this paper, in which all the calibration schemes are introduced and discussed together.

End of section 4.3: please add considerations also on the possible effect of snow (guided by temperature) in the study basin.

Reply: Snow is a fairly rare occurrence in the studied catchment, so we did not consider the effect of snow.

p. 8, l. 65: why Fig. 11 shows the simulation for the period 2005-2008? The validation period is 2001 to 2011.

Reply: The data of the validation period is too many to show in the figure, so we only choose the period 2005-2008 to represent the validation period. However, this figure is not very clear to describe the results, and an improved figure will be provided in the revised manuscript.

p. 8, ll.76-82: also this paragraph suffers in clarity from the 'late' addition of a second FCM scheme:

rephrase referring to both FCM algorithms as developed at the same time and with the same ‘dignity’.

Reply: Agreed. We will reorganize the paper, in which all the calibration schemes are introduced and discussed together.

p. 8, ll. 87-88: actually, given the confusion and lack of information in section 4.1, this conclusion (utility of cluster validity index for choosing Kopt) is not supported by what is presented in the current version of the manuscript.

Reply: We will further clarity the utility of cluster validity index for choosing Kopt in the revised manuscript.

p. 8, ll. 89-90 (and section 4.4): please elaborate more on such result (bi-monthly sub-periods as the best performing partition), trying to explain it, if possible.

Reply: Agreed. We will add more elaboration and explanation on this result.

We hope our responses to the comments are satisfactory and look forward to more suggestions.

Best regards,

Binru Zhao, the corresponding author