We thank the anonymous reviewer for the supportive and helpful comments. Below we address each comment and explain how we revised the manuscript in response:

1. All the results are based on simulated soil moisture patterns. While the authors clearly state this assumption, they provide no support that the simulated patterns are realistic. They calibrate the discharge that is produced by the model to the observed discharge, but I don't think they have any observations of soil moisture against which to compare. How do we know that the simulated patterns have any resemblance to the real patterns? This is a critical limitation on the analysis because everything the authors conclude relies on the validity of those patterns.

   Here, the reviewer raises an important point. Simulated soil moisture depends on (i) climate and soil data as well as (ii) model parameterization and structure. Applying a Monte Carlo uncertainty analysis, we accounted for the uncertainty related to model parameterization. Nevertheless, the influence of climate and soil data as well as the representation of processes was so far not addressed. A verification of the simulated soil moisture with soil moisture point measurements (e.g. gravimetric, TDR) is not feasible as these are highly variable over short distances. On the other hand, satellite based and hydrological simulated soil moisture estimates are spatially integrated values. For this reason, we validated the temporal progression of simulated soil moisture against the remotely sensed soil water index (SWI) (Wagner et al., 1999, http://www.ipf.tuwien.ac.at/radar/ers-scat/home.htm) by calculating the Pearson correlation coefficient between the standardized simulated soil moisture (SMI) and the SWI. Figure A shows the Pearson correlation coefficient (significant at 5 % level) between the median simulated SMI and the remotely sensed SWI. Except for a few local spots, basin wide high correlations are observed. The median correlation is 0.57, the difference between the 25th and 75th quantile is 0.2.

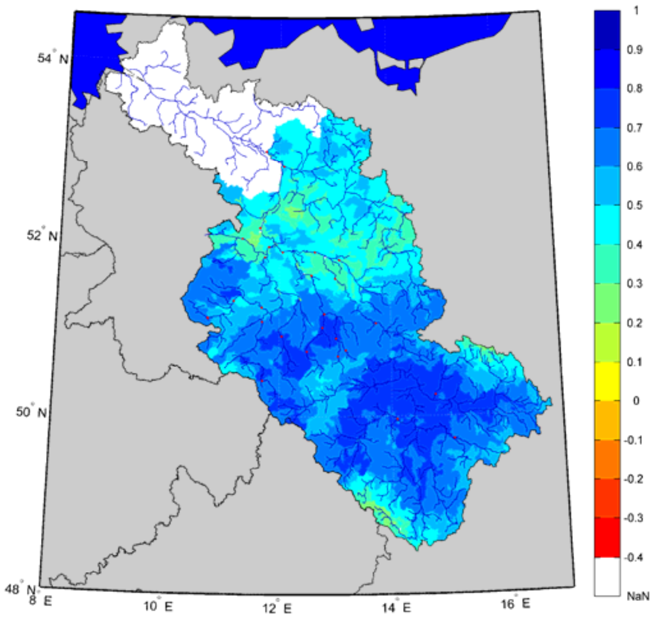

Figure A. Pearson correlation coefficient (significant at 5 % level) between the median simulated SMI and the remotely sensed SWI.

The individual examination of the Monte Carlo parameterizations leads to the same findings. Hence, we demonstrated that the simulated temporal progression of the SMI is well represented in the hydrological model.

We included this validation in the revised manuscript. In the 'Data' section, we added a brief description of the retrieval of the soil water index and data availability. The Pearson correlation coefficient as well as the assignment of the SWI to the subbasins is explained in section 'Methodology: Model calibration and validation'. The results including Fig. A are presented in 'Results: Hydrological modeling' and discussed in the subsequent section.

2. The methods in this paper are not adequately described. Numerous important details missing, but more importantly, the overall rationale for the approach is not given. I am very familiar with PCA, but I could not determine exactly what the authors have done from their description. There are many ways that PCA can be performed with a given dataset (for example, see the details of the papers that the authors cite), and these various approaches ultimately lead to different principal components (PCs) and different interpretation of those PCs. In particular, the authors say they performed PCA on the "spatial linear Pearson correlation matrix." If I understand them correctly, this is an unusual way to do the analysis. Why was this matrix selected? What is the rationale behind this unusual choice? What are the implications for the PCs and eigenvectors (i.e. how should the reader interpret those)? The authors also choose to include multiple simulated soil moisture patterns in the PCA rather than performing different PCAs for the different simulated patterns. Why was that approach selected, and what are the implications? I am less familiar with cluster analysis and their description of this topic is much shorter. I could not understand what was done.

    The sentence, "PCA is performed on the spatial linear Pearson correlation matrix" (page 10062, line 22) might have caused a misunderstanding. First, we calculated the spatial linear Pearson correlation matrix including multiple simulated soil moisture patterns. From this correlation matrix we extracted the eigenvectors (Eq. 3 and Eq. 4). Finally, the principal components are obtained by projecting the standardized soil moisture data onto the eigenvectors (Eq. 6). Different PCAs are performed for the different simulated patterns. We have removed the misleading sentence cited above as our approach follows the one proposed by the reviewer.

    We extended the description of the cluster analysis and added a reference of the applied cluster algorithm: "The hierarchical Ward cluster algorithm (Ward Jr., 1963) was implemented on the leading PCs to identify days of similar soil moisture patterns. At the beginning, every day D represents a single cluster. At each analysis step, the union of all possible cluster pairs is considered and the cluster pair t that offers the smallest increase in variance V is merged.

    $$V_t = \sum_D \sum_k |PC_{tDk} - \overline{PC}_{t\cdot k}|^2 \qquad (A)$$

    where PC_{tDk} denotes the value of the k-th PC at day D belonging to cluster t. The algorithm merges the days progressively until all days are united in a single cluster."

3. The implications of the results seem rather limited. The authors mainly focus on the direct results of the analysis such as the variance that is explained by the different PCs and so forth. The deeper implications of those results are not well explored. In the end, the main conclusion of the paper is that "these results underline the importance of catchment state for flood initiation and severity." This conclusion is relatively unsatisfying when compared to the worthy objectives of the paper. Can the authors provide more precise and impactful conclusions?

In the revised manuscript, we restructured the conclusions to better highlight the outcome of the analysis and the discussion. The following conclusions can be drawn: (1) Different soil moisture patterns are not equally associated with flood occurrence. (2) Patterns with catchment wide high soil moisture are associated with the majority of flood start days and are related to the most severe flood events. (3) Flood initiating soil moisture patterns vary seasonally. In winter, floods are initiated by overall high soil moisture content whereas in summer the flood initiating soil moisture patterns are diverse. (4) Seasonality of flooding triggered by a given soil moisture pattern, doesn't necessarily follow the same seasonal variation as the seasonality of the pattern. (5) Occurrence of a soil moisture pattern doesn't necessarily lead to flood initiation, but the probability of occurrence of a large-scale flood may be increased.
While the results underline the importance of catchment state for flood initiation and severity, (4) and (5) indicate that beside soil moisture, other patterns are relevant for flood initiation. Therefore, future work will extend the soil moisture classification to circulation patterns and snow. A combination of hydro-meteorological pattern types would enable to quantify the interaction of patterns of hydrological catchment conditions and meteorological conditions on flood initiation and magnitude.

Specific comments

a. Figure captions need to be more descriptive.

On the basis of comment 1) of the 2nd reviewer, we decided to remove Figure 3, 5 and 8. In the revised version, we changed the captions the following:

Figure 1: *Topographic map of the Elbe river basin. Yellow dots: Discharge gauges applied in flood identification. Red dots: Discharge gauges applied in flood identification as well as model calibration and validation. Crosses: Location of pixel centroids of the scatterometer data.*

Figure 4: *Loadings (left) and their corresponding PCs (right) of the leading four PCs. PCs are displayed for the sub-period 1982-1991. Minimum and maximum values correspond to the parameter uncertainty introduced by the rainfall-runoff model.*

Figure 6: *Median probability of cluster membership p_t of different PC-cluster combinations (left). PC-cluster combinations with a small median p_t are strongly influenced by model parameterization.*
Distribution of p_t for different numbers of clusters when clustering the leading four PCs (right).

Figure 7: *Soil moisture index (SMI) patterns of cluster centroids.*

b. Section 2: Study area Page 10057, lines 15-17: Climate is described both as temperate and as transitioning from maritime to continental. Are these descriptions consistent or contradictory?

We agree that those descriptions are contradictory. Therefore, we deleted "The climate is temperate" (page 10057, lines 15-16) and change the subsequent sentence in "situated in a transition zone between temperate (lower Elbe) and continental climate (Middle and Upper Elbe)" in the revised manuscript.

c. Section 3: Hydrological modeling Page 10060, lines 24-25: "accumulated soil transpiration" over what duration? Page 10060, line 25: Should "potential soil transpiration is reduced" be "actual soil transpiration: : :"? Also, wording is incorrect with ": : :in dependence of the number of days: : :"

We reformulated the statement: "As long as the soil evaporation accumulated since the last rainfall event is below 6 mm, actual soil evaporation equals potential soil evaporation. If the soil evaporation exceeds 6 mm, actual evaporation is assumed to decay exponentially."

d. Page 10061, line 9: cite chapter of reference, not entire reference

We followed the reviewers' advice in the revised manuscript:
(Maidment, 1993, chapter 10.2.3)

e. Page 10061: The model used is calibrated for discharge, not soil moisture patterns and as far as I can tell there no actual soil moisture values are used in the paper for validation of the modeled moisture patterns. The lack of soil moisture measurements to verify to any degree the modeled patterns means that the results presented in the manuscript remain speculative. The authors did note the possibility of verification of the patterns by remotely-sensed products.

Please refer to our reply on comment 1).

f. Page 10061, line 15: clarify what "behavioral model performances" means

We expanded the statement in the revised manuscript: "A Monte Carlo uncertainty analysis is carried out to identify an ensemble of parameter sets that lead to behavioral model performances. A model performance is considered as behavioral, if an a-priori set threshold of goodness of fit is exceeded." page 10061, line 14.

g. Page 10061, line 16: How were the nine sensitive parameters identified?

Preliminary sensitivity analysis was performed by randomly perturbing different parameters individually.

h. Page 10062, line 11: Page 10060, lines 8-9 indicate that multiple soil layers are used in the model, so need to explain what constitutes a daily soil moisture value. Is it an arithmetic average? Or a layer-depth weighted average value?

Daily soil moisture values of the soil profile are layer-depth weighted average values. We added this information in the revised manuscript.

i. Page 10063, lines 13-15: This is a valuable sentence explaining the relationship between eigenvectors and PCs. That relationship could be further clarified by explicitly stating that PCs are time-series and the eigenvectors are spatial patterns.

We followed the reviewers' advice in the revised manuscript.

j. Page 10064, line 13: climPCk and PCk are time-series, why is the correlation being done with a median value? What is a "decomposed" PC? Is it the same as the decomposition on page 10065, line 12? If so, then explain the decomposition at this point and how the median is calculated.

The "decomposed" PC is it the same as the decomposition on page 10063, line 15-17. However, on the basis of comment 1) of the 2nd reviewer, we decided to remove this section from the manuscript.

k. Page 10064, line 14: The formula for the Spearman's rank correlation coefficient is unintelligible.

   On the basis of comment 1) of the 2nd reviewer, we decided to remove this section from the manuscript.

l. Page 10064, 3.3 Cluster analysis: Explain more thoroughly, but briefly, cluster analysis in general. Many readers will not be familiar with that technique. Page 10064, line 18: Why up to 15 PCs used? Page 10064, line 19: What is the weight that is being assigned to a PC? Page 10064, line 22-23: Why is a distance metric needed and what are you measuring the distance between? What is the hierarchical Ward algorithm and provide a reference for it. Page 10065, lines 11-17: How are the "leading PCs" determined? How many are there? Also, what are the results of the cluster analysis? It seems that each PC, a time-series, is divided into clusters consisting of days that are somehow similar to each other. In the second cluster analysis, what is a cluster centroid? An SMI value? One should not have to assume that the numbers of clusters in two analyses are the same but should be able to control that number. Page 10065, lines 18-26: Is cluster enumeration somehow ordered? Otherwise, cluster 1 in one analysis may be equivalent to cluster 8, for example, in another analysis. But in order to assign a single cluster number to a day, the enumeration should have some meaning or order to it.

   The cluster analysis identifies days of similar soil moisture patterns and summarizes them in one cluster. We extended the description of the cluster analysis and added a reference of the applied cluster algorithm. Please refer to our reply on comment 2).
   The cluster centroid is the point with coordinates equal to the median of the variables of all observations in that cluster (page 10065, line 16). As the cluster centroid therefore characterizes the cluster, it can be applied to compare the results of various cluster analysis independent of cluster enumeration.
   In the Ward cluster algorithm, the user has to set the cluster number, as the algorithm merges the days progressively until all days are united in a single cluster. To determine the optimum number of clusters, cluster validation methods can be applied (page 10065, line 2-8).

   The individual PCs are not standardized i.e. the quantile difference (75 % and 25 %) of the first PC is 64, of the second PC is 18 and of the 15th PC is 3. Therefore, the first PC has the highest weight when applying the variance criteria of Ward (Eq. A) and the influence of a PC on the clustering declines the lower its explained variance. In the revised manuscript, we reformulated our statement: "Applying the variance criteria of Ward, the weight assigned to a particular PC depends on its explained variance as the individual PCs are not standardized i.e. the first PC has the highest weight."
   The use of up to 15 PCs is an arbitrary decision justified by the small amount of explained variance by the 15th PC (2.8 %) and the even lower explained variance for the subsequent PCs. In total, the leading 15 PCs explain more than 90 % of total variability. Please refer also to our reply on comment w).
   The leading first PC refers to the eigenvector with the highest variance. The leading second PC refers to the eigenvector with the second highest variance and so forth. In total there are 1945 'leading PCs', although only the first view explain a high portion of total variance.

"For the comparability of the simulated profile soil moisture of different soil types, the values are standardized by the field capacity of each soil type. This standardized profile soil moisture content is termed soil moisture index (SMI)" (Page 10062, line 16). In the revised manuscript, we put the SMI definition in a separate paragraph.

m. Move section 3.4 Flood event identification to the beginning of section 3 so the ordering agrees with section 4.

   We followed the reviewers' advice in the revised manuscript.

n. Page 10066, lines 12-14: Provide a brief justification why the flood identification method used was chosen from the five methods listed.

   We added a brief justification in the revised manuscript (page 10066, line 12): "We identified large-scale flood events in the Elbe river basin using an approach proposed by Uhlemann et al. (2010) as the method is non-restrictive to a certain return period taking the simultaneous or time shifted occurrence of peak discharges at many sites into account."

o. Section 4: Results Page 10067, lines 20-22: Are the validation gauges a completely separate set from the calibration gauges? The validation gauges need to be identified on Figure 1, along with specifying what 'identification' gauges are. This is an example of a paragraph that is too short and needs to be expanded or combined with the subsequent paragraph.

   The validation gauges are identical to the calibration gauges excluding gauge Laucha where discharge time series weren't available for the considered time period. We added this information in the text as well as in Figure 1. 'Identification' gauges are the gauges used in the flood event identification method. We clarified this in the figure and captions.

p. Page 10068, lines 12-19: Selection, justification and use of a consistent number of PCs throughout the manuscript will aid in reader comprehension. For example, currently forty-three PCs are significant according to this paragraph but twenty are shown in Figure 3, five are shown in Figure 4, and up to fifteen are used in the cluster analysis.

   We agree with the reviewer. In the revised manuscript, we use a consistent number of PCs. In the end, we apply cluster analysis on the leading 4 PCs (Please refer also to our reply on comment w). Therefore, we limited Figure 4 to 4 PCs and integrated the information provided by Figure 3 for these PCs in the text.
   On the basis of comment 1) of the 2nd reviewer, we decided to remove Figure 5 as well as Figure 8 from the manuscript.

q. Page 10068, line 20 – Page 10069, line 12: Ensure correct reference to eigenvectors and PCs throughout this range of text. For example, the sentence beginning with "The second PC: : :" indicates that the PC shows a north-south partition; however, the PC in this analysis is a time-series and the eigenvector is the spatial as shown in Figure 4. Accurately referencing the different products of the PCA is required to avoid confusion about the manner in which the PCA is performed.

   We carefully rechecked this in the revised version.

r. Page 10070: Could you note the divergent results of the different cluster validation methods but pick one to use in the analysis? Page 10069-10072: I am unclear as to exactly how the cluster analysis is performed and what comprises a cluster. Is a cluster a range of PC anomaly values? I need a better description of cluster centroids and how they lead to SMI patterns. In general, the description of the cluster methodology needs a good deal of work to clarify the techniques, products and results. I am lost in varying numbers of clusters, probabilities and patterns. Without a better explanation of the technique and the analyses performed and their purpose, I am unable to assess the validity of the results.

   Please refer to our reply on comment 2), l) and w).

s. Figure 1 Legend needs better formatting Identify validation gauges

   Please refer to our reply on comment o).

t. Figure 3 The number of PCs shown on the figure seems arbitrary. Twenty is already less than the total number of significant PCs but the variance explained could be considered negligible after only a few PCs.

   We agree with the reviewer and integrated the information provided by Figure 3 in the text.

u. Figure 4 This is a good figure for illuminating the form of eigenvectors (spatial patterns) and PCs (time-series). Center "Principal Components" title text over figures

v. Figure 5 A table containing this data would be easier to understand.

   On the basis of comment 1) of the 2nd reviewer, we decided to remove this graph from the manuscript.

w. Figure 6 These figures are too complicated. The number of PCs shown is arbitrary and not consistent with other portions of the manuscript. The numbers of clusters shown seem arbitrary and random and there are too many different numbers. I have difficulty determining the message that the figures are attempting to convey.

   As the number of PCs used in the cluster analysis as well as the number of clusters is an arbitrary decision, we conducted a sensitivity analysis varying the number of PCs between one and 15 and the number of clusters between one and 100. Based on the probability of cluster membership (Eq. 10), we selected the optimum PC-cluster combination: The leading 4 PCs are separated into 10 clusters.
   In the revised manuscript, we restructured the section 'Methodology: Cluster analysis' to better highlight our main purpose.

x. Figure 7 I am unclear as to what a cluster centroid is and how it leads to a spatial moisture pattern.

   Please refer to our reply on comment l).

y. Figure 8 I believe this figure shows the distribution of the values in PC1 and PC2 among the different clusters (i.e., each box and whisker indicates the distribution of anomaly values from PC1 or PC2 within each numbered cluster). If so, the caption should be amended to make that more clear. Also, describe what the features of the boxes and whiskers represent (e.g., median or mean, interquartile range or 1 standard deviation, etc.)

Yes, this figure shows the distribution of the values in PC1 and PC2 among the different clusters. On the basis of comment 1) of the 2nd reviewer, we decided to remove this graph from the manuscript.

References:

Maidment, D. R.: Handbook of Hydrology, McGraw-Hill, Inc. New York, 1993.

Uhlemann, S., Thieken, A. H., and Merz, B.: A consistent set of trans-basin floods in Germany between 1952 - 2002, Hydrol. Earth Syst. Sci., 14, 1277-1295, 2010.

Wagner, W., Lemoine, G., and Rott, H.: A method for estimating soil moisture from ERS scatterometer and soil data, Remote Sens. Environ., 70, 191-207, 1999.

Ward Jr., J. H.: Hierarchical grouping to optimize an objective function, J. Am. Stat. Assoc., 58, 236-244, 1963.