Author comments on

Interactive comment on
"Catchment classification by runoff behaviour with self-organizing maps (SOM)" by R. Ley et al.

U. Ehret (Referee)

The authors thank Mr. Ehret for his helpful, detailed comments, suggestions and editorial corrections, which will help to improve the paper!

Here, comments of Mr. Ehret are printed in grey, answers of the authors in black.

Scientific quality

The scientific approach and the carrying out of the work is mainly clearly explained, self-contained and valid. However, I want to raise some issues here that apply to catchment classification in general and also to proposed work:

• **What shall be classified and for what purpose?** The catchment as a physiogeograpic unit described by e.g. slope, soil, vegetation, land use or the combination of the catchment and its hydroclimatic conditions? Of course the two are linked as the catchment evolved under the hydroclimatic conditions, but it makes a strong difference on the applicability of the classification: If we classify solely based on catchment properties, we may group catchments that show different response behavior as they may be exposed to different hydroclimate. If we classify based on the combination, the two aspects become inseparable. Which of the two approaches is appropriate depends on the intended use of the classification system. The authors apply the second approach in the paper, which makes sense as the intention is to compare it to catchment response behavior. On the other hand, it would be the wrong approach if the purpose of the classification is model parameter transfer (e.g. soil hydraulic parameters), as they are (should be) mainly related to the catchments physiogeographic characteristics.

− **In summary:** The issue of what shall be classified and for what purpose should be clearly stated in the paper. All further decisions (choice of the metrics etc.) in the work should be discussed in the light of this purpose.

The study contributes to the discussion for the HESS special issue "Catchment classification and PUB". The aim of the paper is to classify catchments by their response behaviour with a view to regionalisation and to apply self-organizing maps (SOMs) for catchment classification.
Catchment response behaviour as a result of a combination of physiogeographic and hydroclimatic conditions is the basis of building hydrologically homogenous regions or groups of catchments with similar response behaviour. Therefore, in our first step, we cluster catchments only by their response behaviour on base of indices "independent" of physiogeographic and hydroclimatic conditions. In a next step, we cluster catchments by the most important physiogeographic and hydroclimatic conditions to relate them to special catchment response behaviour and to analyse possible improvements and limits of the method.
Single parameter transfer is in our approach not foreseen.
We will present this aim of the paper in a more clearly way in the revised version.

The influence of channel on the response behaviour may be one catchment property responsible for certain response behaviour. We considered this by calculation of the maximum and mean flow length: the maximum flow length for all catchments vary between 7 and 91 km, with a mean of 31 km. The catchment with the largest flow length of 91 km is the fifth largest one with 681 km². Mean and maximum flow lengths are highly correlated, therefore mean flow length was no further considered. Maximum flow length is correlated with runoff indices describing FDCs up to 0.56 for a single index, and show nearly no correlation to indices describing runoff coefficients. Therefor flow length was not considered as property for clustering by physical catchment properties (see also • 3064/13: choice of the indices).

Rainfall extend is typically in the range of most of the catchment sizes. Only in summer some mostly convective rainfall events affect only parts of catchments.

To consider slope heterogeneity we calculated the standard deviation of mean slope and percentage of area of certain slopes for each catchment. All these measures are highly correlated with mean slope (spearman's rank coefficient > 0.8) and therefore excluded from analysis (see also • 3064/13: choice of the indices).

For the chosen indices describing runoff behaviour catchment size effects are not observable.
A correlation analysis of catchment size and runoff indices showed the largest correlation (spearman's rank coefficient) between Aeo and MWL of 0.47 and HS -0.38, both indices are derived from the flow duration curve. Correlation coefficients between Aeo and the other indices, especially indices from runoff coefficients are considerably lower. Also within the 46 catchments (87 %) smaller than 370 km², correlation coefficients show no significant correlation between catchment size and chosen runoff indices.
Also Merz and Blöschl (2009) found no correlation of mean event runoff coefficients and their variability to catchment size for catchments in Austria. Yadav et al. (2007) suggest that catchment area isn't the most important physical characteristic describing response behaviour for English catchments.

In order to classify catchments only by their response behaviour, indices with a high influence of catchment properties like catchment size or precipitation are not suitable for our study. An influence of physical catchment property like MAP would make a comparison of the both pools of clusters behaviour and physiogeographic and hydroclimatic conditions impossible.
FDCs normalized by catchment area/runoff ratio shows high influences of mean annual precipitation, FDCs normalized by mean runoff are influenced by extreme values, considered in special indices. Normalization by median runoff shows no influence of MAP and good differentiation of FDCs.

We will change the text in page 3061 lines 17-20 as follow:
"To compare the FDCs of different catchments by their response behaviour we have to normalize them in a way that influences of physiogeographic and hydroclimatic conditions are mostly excluded. Normalisation by catchment area shows high influences of mean annual precipitation and is therefore not suitable for this study. Normalization by mean results in FDCs not well distinguishable for low exceedance probabilities and shows a high dependence of extreme values, covered by other indices. Therefore, runoff for FDCs is normalized by its median value (Q50). By doing so …"

The parameter of 12 h in 3055/18 is used only for identification of peak flows as the highest flow in a period of 24 hours. The duration of an event is calculated by minimizing the direct flow before and after a peak flow (start and end of an event) and considers different travel time ranges of the catchments. In the revised version of the paper we will describe this more clearly.

From the event runoff coefficients and FDCs we calculated a huge amount of indices, which describe, seen by themselves, important aspects of runoff behaviour. A correlation analysis showed very high correlations between many of these indices. High correlated indices (spearman's rank coefficient > 0.8) don't bring new insights in the analysis and therefore were excluded from analysis. On the other hand we considered an even distribution of indices with respect to season, high and low flow and the importance of indices.
An explanation of the selection process will be added to the revised version.

Presentation quality

This information will be added in 3053/4:" InterMet takes into account data from about 200 rain gauges, meteorological data ..."

Both papers describe classifications with SOM.
Toth (2009) classifies hydro-meteorological catchment conditions, i.e. parts of time series by SOM for streamflow forecasting for one catchment. It is a good example for using SOM in hydrology and show the wide range of applications. It will be added to the examples on page 3051.
Di Prinizo et al. (2011) in the same special issue cluster catchment by response behaviour and properties too, but with a different focus on selection of indices.
Our approach was -to our knowledge at time of writing- new. The text will be rewritten to "... have been used also in this special issue by Di Prinzio et al., but with a contrasting size of study area and variables and with a different focus."

Single events show differences between manually and automatically calculated coefficients between 0.01 and 0.15, with only a few outliers. Differences between mean ERC und ERCman for one catchment is from 0.01 to 0.10, the mean difference to all ERCman of about 0.05 (Fig. 1). Because of different methods for baseflow-separation automatically calculated runoff coefficients are often slightly higher than manually calculated (linear baseflow-separation). Manual calculation of event runoff coefficients was done for the period 1996 to 2003 with a focus on events with high runoff. The overlap of events between manually and automatically chosen events for this period is, depending on catchment, between 50 and 87 %.
In the revised version of the paper the text of p. 3055 line10 will be: "....indicates a good fit with a mean difference to all ERCman of about 0.05.".
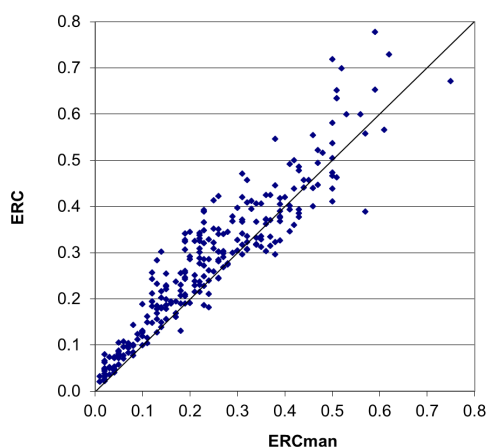

Fig. 1: Comparison of event based runoff coefficients manually (ERCman) and automatically (ERC) calculated for 7 catchments with different qualities of fit.

Up to this part of the paper, no physical catchment properties are discussed. A discussion why the classification is as it is can be made after clustering by physical properties in following section 3.3.
A reference to section 3.4 (Comparison) will be added at line 14: "... values of all indices. A discussion and interpretation of these clusters, related to physical catchment properties, will be done in section 3.4 (Comparison)."

We will change the last sentence of chapter 3.2 into:
"From the cluster by response behaviour we can identify three borderline cases (catchments 1, 26 and 43). They belong to neurons at the edge of a cluster with medium distance to a neighbouring neuron of another cluster. In their first 6 BMUs there is a frequent change between both clusters, with slightly increasing quantisation errors. Therefore we assign them to both clusters: catchment 1 to B and C, catchment 26 to A and B and catchment 43 to D and A. "

The selection and weighting process was done similar to the process to choose indices of runoff behaviour:
First we calculated a large number of indices describing physical catchment properties like geology, soils, land use, catchment size, flow length, climate etc. All of them have an impact on runoff, some more, some less. Second, to reduce this amount of indices, we eliminated all indices with high correlation coefficients (spearman's rank coefficient > 0.8) to other indices. And third, we run a correlation analysis with indices based on runoff behaviour and indices of physical catchment properties to identify indices with a high impact on runoff. Fourth, we chose indices with high correlation to single indices of runoff behaviour considering physical indices identified in other studies as most important for runoff behaviour.

An explanation of the selection process will be added to the revised version.

The text will be changed to:
"- mean long-term potential evaporation (ET), calculated from raster data of the mean annual potential grass reference evapotranspiration from the "Hydrological Atlas of Germany" for the period from 1961 to 1990."

Figure 1 base on a DEM with a resolution of 20 m. We will add a map of mean annual precipitation to Figure 1 in the revised version of the paper.

We will add a new figure 8 to the revised paper with a dendrogram, u-matrix with cluster and component planes of clustering by physical catchment properties as described in section 3.3.

There are borderline catchments in both pools of clusters: 3 catchments for runoff behaviour and 5 catchments for physical catchment properties. To improve the overlap we count these catchments to their second cluster. Result of this are 38 than 30 of 45 catchments in the overlap.

The Text will be changed to:
"…from 67 to about 80 %. There are 3 borderline cases for clusters based on runoff behaviour and 5 catchments in clusters based on physical catchment properties. To improve the overlap we count these 8 catchments near the overlap to their second cluster. Thus …"

We will add an example to the revised text at 3068/13:
"For example catchment 28 (Planig, 171 km²) fits well to clusters C and **II**, like most catchments from cluster C; Catchment 39 (Schulmühle, 145 km²) belongs to cluster C too but fits to cluster **IV**. Both catchment aren't borderline cases and show very low to medium quantization errors to their clusters. The indices of both catchments are inside the range of indices of cluster C, they can be assumed as similar behaving. On the u-matrix (Fig. 8) clusters II and IV are neighbouring clusters but well separated by a row of neurons indicating high differences, shows a clear separation of both clusters. This case of process equifinality points to different processes but similar runoff behaviour and can be figured out by comparing both pools of clusters."

All the corrections mentioned by Mr. Ehret as 5. Minor comments will be modified in the revised version of the paper.

Yes, it is the 1.0 minus 0.8 quantile: will be changed.

This is because of bad data quality and many missing data for very low flow of some catchments.