

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

Anonymous Referee #2

Received and published: 7 June 2011

The aim of this paper is to build a classification using self-organizing maps. This approach has not been applied before. It is thus the first application.

The paper is well documented and well presented. I have only few comments that could make this paper more didactic. This article should be accepted after minor revision.

We may expect more technical advices for hydrologists who would like to apply SOM. There are several parameters and we do not know how they were fixed. Drawbacks and advantages of SOM could be also listed (it is not clear why SOM should be used instead of one classical clustering approach; in particular there is still one question about the number of clusters to be considered).

C2031

P 3051: The authors should justify the choice of the signature indices and, in connection to these variables, the reasons for looking for clusters. Indeed the relevance of a classification depends on the final objective and on the use of the analysis. In this paper we may suspect that SOM are established to study spatial variability of flood characteristics at regional scale.

P 3055: Why was the method suggested by Merz et al. (2006) modified? Parameters for the extraction was fixed (peak flow with direct runoff $> 2 \times$ baseflow and a maximal time window of 24 hours, threshold fixed to 5 mm to identify rainfall event). How were these parameters defined? For large basin ($> 1000 \text{ km}^2$) the duration may exceed 24 hours. Is there any bias in the selection (underestimation of the related runoff coefficients)? In addition there is certainly a time lag between the rainfall event and the answer of the basin. The time windows are not identical for runoff and rainfall. How are they defined to ensure consistency?

Excluding events due to snow melt is relevant to study link between rainfall-runoff relations. The selection does not consider as well events with low runoff coefficients in late summer (due to drought conditions). Why is it important to exclude them since these events describe the basin behavior and the dependency to actual conditions (objective mentioned P 3056)?

P 3058: Results are presented. The number of neurons has been chosen arbitrary. The authors should develop this part: on which basis was this number fixed? What do the authors mean by “the best choice”? What are the consequences of choosing a high number of neurons?

P 3060 and P 3080: Labels for Fig.4b and Fig.4c are inverted.

P 3062: I am not familiar with SOM. Fig 4a shows 30 neurons located in a matrix with 11×9 units ; Fig 4c shows 23 neurons located in a matrix with 6×5 units. I do not understand why the authors consider 25 neurons and the links between a and c since the sizes are different.

C2032

The range of Fig 5 is not the same as the range of Fig 4a.

P 3064: Is it possible to locate the three borderline catchments on the map (Fig 6)? Please give more details on these basins (e.g. are they the most highly urbanized basins?).

P 3066-3067: In Table 3 the correspondence is established between clusters based on the number of shared basins. Is this criterion sufficient especially when the number is similar for different pairs of clusters?

There is an overlap of 67% which corresponds to 30 basins among the 45. The overlap reaches 80% (36 basins among the 45) considering these catchments. The increase of 13% is not consistent with the number of borderline catchments (= 3 , P 3064).

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 8, 3047, 2011.