

Catchment classification by runoff behaviour with self-organizing maps (SOM)

R. Ley¹, M. C. Casper¹, H. Hellebrand², and R. Merz³

¹University of Trier, Forschungszentrum für Regional- und Umweltstatistik, Trier, Germany

²Hydrosol, Technologie Zentrum Trier, Trier, Germany

³Helmholtz Centre for Environmental Research, Department of Catchment Hydrology, Halle, Germany

Received: 22 March 2011 – Published in Hydrol. Earth Syst. Sci. Discuss.: 29 March 2011

Revised: 14 July 2011 – Accepted: 9 September 2011 – Published: 16 September 2011

Abstract. Catchments show a wide range of response behaviour, even if they are adjacent. For many purposes it is necessary to characterise and classify them, e.g. for regionalisation, prediction in ungauged catchments, model parameterisation.

In this study, we investigate hydrological similarity of catchments with respect to their response behaviour. We analyse more than 8200 event runoff coefficients (ERCs) and flow duration curves of 53 gauged catchments in Rhineland-Palatinate, Germany, for the period from 1993 to 2008, covering a huge variability of weather and runoff conditions. The spatio-temporal variability of event-runoff coefficients and flow duration curves are assumed to represent how different catchments “transform” rainfall into runoff. From the runoff coefficients and flow duration curves we derive 12 signature indices describing various aspects of catchment response behaviour to characterise each catchment.

Hydrological similarity of catchments is defined by high similarities of their indices. We identify, analyse and describe hydrologically similar catchments by cluster analysis using Self-Organizing Maps (SOM). As a result of the cluster analysis we get five clusters of similarly behaving catchments where each cluster represents one differentiated class of catchments.

As catchment response behaviour is supposed to be dependent on its physiographic and climatic characteristics, we compare groups of catchments clustered by response behaviour with clusters of catchments based on catchment properties. Results show an overlap of 67 % between these two pools of clustered catchments which can be improved using the topologic correctness of SOMs.

1 Introduction

An important task of science in any particular field is to “perpetually organize a body of knowledge gained by scientific inquiry” (Wagener et al., 2007). Classification groups together those systems that are similar, limiting the variability within classes (McDonnell and Woods, 2004). Thus, classification in itself can be a valuable first contribution in gaining understanding of systems.

In hydrology, a classification of catchments based on a rigorous analysis of patterns in observed data is almost non-existent (McDonnell and Woods, 2004; Woods, 2002). Classification is a task of learning a classification model that maps each attribute set x to one of the predefined class labels y (Tan et al., 2006). Prior to the use of a classification model we have to define classes of similar catchments. For example Burn (1997) used an agglomerative hierarchical clustering algorithm to define homogeneous regions for catchment regionalization. Basin similarity is expressed using seasonality measures derived from the mean date of occurrence of the annual maximum flood. Hall and Minns (1999) demonstrated that Representative Regional Catchments (RRC) whose characteristics are hydrologically more appealing than geographical proximity might define classes. They employed techniques like Kohonen networks and fuzzy c -means, which are straightforward in application and were found to identify broadly similar RRCs. Merz et al. (2006) described six climatic regions in Austria by Event Runoff Coefficients (ERCs). ERCs are highly correlated with mean annual precipitation, but poorly with soil type and land use. Oudin et al. (2010) defined similarity of catchments on the basis of model parameter transferability and compared them with a pool of apparently physically similar catchments.

With respect to this current state of basin classification Wagener et al. (2007) provide a review of catchment



Correspondence to: R. Ley
(leyrita@uni-trier.de)

classification and conclude that “there is a need for a classification system in which each catchment is another data point and would increase the understanding about how catchment functions are defined.” Furthermore, they define requirements for a classification framework which “should provide a mapping of landscape form and hydro-climatic conditions on catchment function ... while explicitly accounting for uncertainty and for variability at multiple temporal and spatial scales.” Thus, opposite to the uniqueness of place concept (Beven, 2000) and despite the fact that complexity and differences between catchments can be overwhelming, patterns and connections might be discernible according to Wagener et al. (2007). They therefore place catchment classification as a necessary step in the advancement of hydrological sciences.

Different classification procedures have been proposed in the literature. Traditionally, geographically coherent regions have been applied in regional flood analyses. In most practical applications (e.g. Stedinger et al., 1992), the regions are found by expert judgement, i.e. by a subjective assessment of which catchments one would expect to behave similar in terms of their runoff behaviour. These subjective considerations are usually based on a personal knowledge of the analyst of the catchment characteristics, climatic inputs and runoff response of the catchments. There have also been a number of attempts at identifying homogeneous regions by multivariate statistical methods that are used for grouping catchments according to their similarity in catchment attributes. Cluster analysis is one of the popular statistical methods for combining catchments into groups (e.g. Acreman and Sinclair, 1986; Burn, 1997). The idea of cluster analysis is to identify groups (regions) in such a way that the similarity of catchments within one region is maximized while similarity between regions is minimized. Other methods used to form groups are factor analysis, principal component analysis, artificial neural networks, fuzzy sets and canonical correlation analyses. A promising technique for classifying catchment in hydrology is SOM, which, to our knowledge, have not yet been used before in classifying catchment response behaviour. Di Prinzio et al. (2011), in this special issue, also used SOM to classify catchment response behaviour, but with a contrasting size of study area and variables and with a different focus.

A SOM consists of an unsupervised learning neural network algorithm that performs a non-linear mapping of the dominant structures present in a high-dimensional data field onto a lower-dimensional grid (Herbst et al., 2009b). Moreover, a SOM represents all input data sets in such a way that the distance and proximity relationships (i.e. the topology), are preserved as much as possible. Properties that distinguish SOM from other data mining tools are that it is numerical, non-parametric, and insensitive against a small portion of missing data. SOMs represent graded relationships, provide visualizations of structures in high dimensional data sets, need no assumptions about data distribution or cluster shapes and may find unexpected structures in the data (Kaski,

1997). There are only a few parameters to fix before training a SOM: the distance measure, a neighbourhood function, normalization and the size of the SOM. Once a SOM is trained, it can be used for many purposes e.g. detecting data structures, clustering, visualizing results via plots and describing clusters. In general, neural networks are good at solving problems such as classification, prediction, association and particularly clustering. SOMs have found diverse applications in various fields of data mining in business, engineering, industry, medicine and science (Maimon and Rokach, 2005). Hall and Minns (1999) used a small SOM for regionalization of catchments by physical catchment characteristics. Ramachandra Rao and Srinivas (2008) used SOMs for regionalization of watersheds in Indiana, USA, to identify plausible regions. Herbst et al. (2009a) used SOMs for hydrological model evaluation and optimization. Toth (2009) used SOM to classify hydro-meteorological catchment conditions for streamflow forecasting. Kalteh et al. (2008) gives an overview of various applications of SOM in hydrology e.g. analysis of rainfall-runoff processes, simulation of surface water quality, clustering ecologic items or remote sensing methods to cluster soil moisture. They conclude that “SOM is a promising technique suitable to investigate, model, and control many types of water resources processes and systems”.

The main objective of this study is to apply SOMs for catchment classification by their response behaviour. Clustering of catchments is the first step to build a classification scheme. Catchments with similar physical catchment properties may show different runoff behaviour because of certain combinations of physiographic and climatic properties. Therefore, we cluster catchments by response behaviour, independent of these properties. Additionally, using the same technique, the catchments are grouped independently of their response behaviour according to their physiographic and climatic catchment attributes, such as mean annual precipitation and mean slope. The comparison of the two independent grouping results sheds light on the main research questions of PUB (Sivapalan et al., 2003), i.e. to what extent do catchments with similar physiographic and climatic catchment attributes show similar response behaviour?

Catchment response behaviour is described by signature indices, derived from event-based runoff coefficients and flow duration curves. Event-based runoff coefficients offer information about changes between events and seasons and give a first idea of the hydrological functioning of the catchment under different conditions. They help to understand how different landscapes “filter” rainfall into runoff and to explain observed differences with catchment characteristics and related runoff mechanisms (Blume et al., 2007). They are used as a diagnostic variable to represent runoff generation in catchments, particularly if a range of catchments and a range of events are to be compared by a single indicator (Merz et al., 2006). The flow duration curve is commonly used to indicate and classify watershed functioning,

summarize the ability of a catchment to produce flow values of different magnitudes, and is therefore strongly sensitive to the vertical redistribution of soil moisture within a basin (Yilmaz et al., 2008).

Compared to studies characterizing catchments by response behaviour over large and diverse areas, e.g. in Austria (Merz et al., 2006), in France and England (Oudin et al., 2010), in the UK (Yadav et al., 2007) or in Australia (van Dijk, 2010) the current study area comprises 53 small to medium-sized catchments in the low mountain ranges of in Rhineland-Palatinate, Germany. This area was also part of the study area to describe catchment behaviour with winter storm flow coefficients in regression models by Hellebrand et al. (2007).

With respect to the main objective, a four-step approach is envisaged:

- Step (i) specifies information from hydrological catchment characteristics obtained from hydrograph analysis: event-based runoff coefficients and flow duration curves.
- Step (ii) uses Self-Organizing Maps (SOM) for clustering catchments with similar response behaviour and describes clusters as a basis for classification. Additionally we implement Hierarchical Clustering on the SOM to define the number of classes and cluster borders.
- Step (iii) clusters catchments by important physical catchment properties with the same method like (ii).
- Step (iv) compares both pools of clusters as a review of interpretability of the clustering by response behaviour. Furthermore it allows connecting catchment behaviour with physical catchment properties as a basis for catchment classification covering both kinds of catchment characteristics.

The implementation of these four steps pursues a scale-independent and region-independent basin classification method which lies in line with the classification system as proposed by Wagener et al. (2007). However, it should be noted that this study does not propose a universal classification system, but merely tries to uncover the potential of a basin classification method in a rather geographically restricted area.

2 Data and methodology

2.1 Study area

The study area consists of 53 small to medium-sized gauged catchment areas in Rhineland-Palatinate, Germany (Fig. 1a). The catchments are situated in the low mountain ranges of the Rheinisches Schiefergebirge, the Saar-Nahe-Bergland and the Rhine Valley. Most of the catchments belong to

the basins of Ahr, Wied, Lahn and Nahe, which all drain to the river Rhine. Five catchments drain directly to the Rhine. Among the 53 catchments there are 35 upstream catchments; three catchments are triple nested. The catchments areas vary from 9 to 1469 km², 46 catchments areas are less than 400 km² and 2 are larger than 1000 km². Elevation ranges from about 100 m a.s.l. in the Rhine valley up to 818 m a.s.l. in the Hunsrück, mean elevation is about 341 m a.s.l. Geology differs from schist, greywacke, and quartzite in the Rheinisches Schiefergebirge to sedimentary rock in the Saar-Nahe-Bergland and Rhine Valley. Many watersheds are characterised by tertiary and quaternary volcanism (basaltic rocks, pumice stone and tuff).

Mean annual precipitation ranges from 530 mm yr⁻¹ in the south-east up to 1108 mm yr⁻¹ in the west and north (study period 1993–2008, Fig. 1b). Almost all watersheds are rural with little urbanization except for four catchments which are moderately urbanized (11 % to 14 %). The main land use is agricultural, but varies between 7 % and 90 % for single catchments. Some watersheds, especially in the southeast, support viticulture and orchards.

2.2 Data

For the data-driven classification method, hourly runoff and areal precipitation data for the period from January 1993 to December 2008 are available. These time series cover a wide range of diverse annual or seasonal precipitation and runoff events: from years with high precipitation and exceptionally heavy floods like 1993 or 1995 to years with very dry summer periods like 2003.

Aerial precipitation was calculated with “InterMet” (Gerlach, 2006), which interpolates meteorological data using kriging technique. To calculate aerial precipitation for Rhineland-Palatinate and adjacent areas, InterMet takes into account data from about 200 rain gauges, meteorological data, prevailing atmospheric conditions, orography, and satellite and radar data. Typical rainfall fields extend in the range of most of the catchment sizes. In summer some mostly convective rainfall events affect only parts of catchments.

2.3 Methods

As described in the introduction, a four-step method is pursued for a catchment classification method. This section explains the indices describing physical catchment properties and the analytical methods.

2.3.1 Event-based runoff coefficients

This study uses the Event-based Runoff Coefficient (ERC) as described by Merz et al. (2006) and Norbiato et al. (2009). ERCs describe the rainfall amount that appears as a

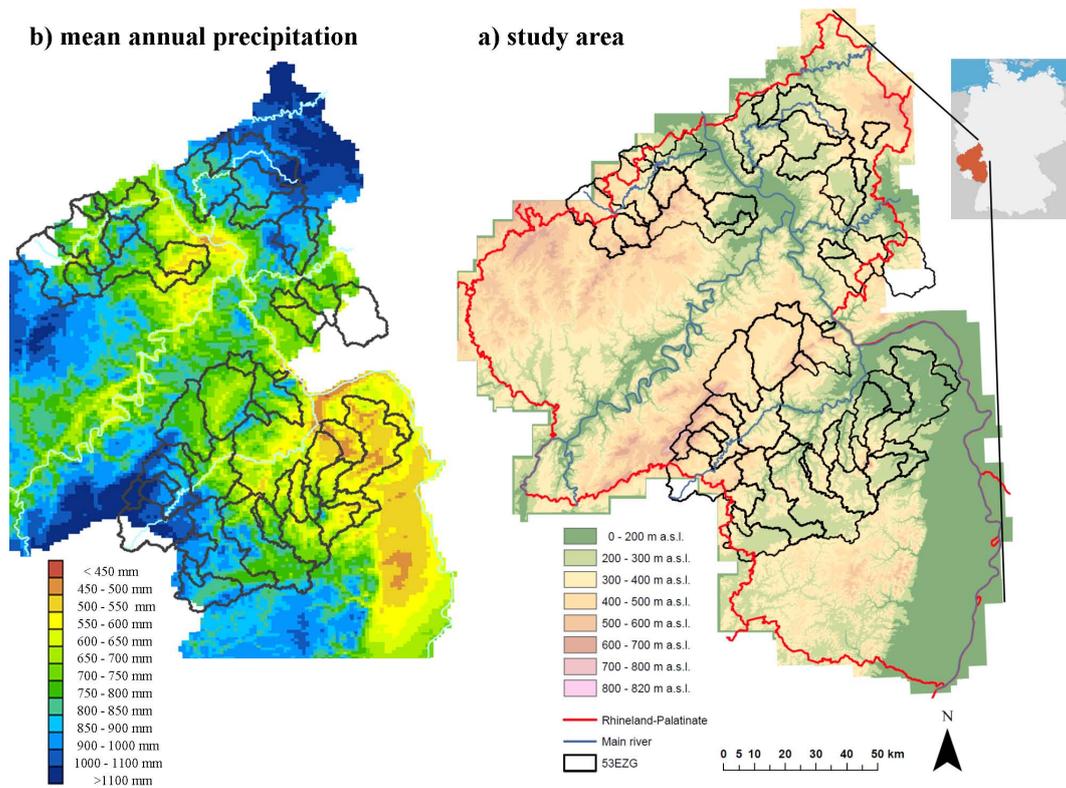


Fig. 1. Study area: (a) 53 catchments in Rhineland-Palatinate, Germany and (b) mean annual precipitation 1993 to 2008. Note that some catchments are nested, so their areas overlap. Smaller catchments are plotted on top of larger ones.

significant runoff above base flow, directly following the corresponding rainfall as given in Eq. (1).

$$\text{ERC} = \frac{\sum Q_d}{A_{eo} \cdot \sum \text{prec} \cdot 1000} \quad (1)$$

with:

ERC = Event Runoff Coefficient

Q_d = direct event runoff [$\text{m}^3 \text{h}^{-1}$]

A_{eo} = catchment area [km^2]

prec = areal event precipitation [mm h^{-1}].

The method of Merz et al. (2006) was developed for catchments in Austria. For this study we modified some parameters responsible for catchment separation to optimize the method for catchments in Rhineland-Palatinate.

Comparing the resulting ERCs with manually calculated ERCs for 17 catchments in Rhineland-Palatinate verifies them. The comparison between manually and automatically calculated coefficients indicates a good fit with a mean difference to all manually calculated coefficients of about 0.05. The same set of adapted criteria is used for all catchments in this study.

ERCs are calculated by a four-step approach:

- Separation of observed runoff into baseflow and direct flow using the digital filter proposed by Chapman and Maxwell (1996).

- Identification of peak flows, start and end of events and event-rainfall as described by Merz et al. (2006). A peak flow was identified with direct runoff twice as high as baseflow and with no larger flow 12 h before and after. To assist in event separation, a characteristic time scale of the runoff dynamics of each runoff peak was estimated. Start and end of an event were searched in an iterative process with the characteristic time scale and thresholds to find the time where the direct runoff at the beginning and end of an event is as small as possible (Merz et al., 2006). Event rainfall was defined as the amount of rainfall within a time period before start and end of an event depending on the characteristic time scale.

- Calculation of event runoff and rainfall volume and estimation of ERCs following Eq. (1).

- The calculated ERCs contain unsuitable events which are: very small events, events with insufficient data or poor event separation and events caused by snow melting.

We exclude very small events with a rainfall amount less than 5 mm or a maximum discharge less than the long-term mean discharge measured separately for summer and winter. Frequently, these events are not separated

accurately or display only low fluctuations of discharge. These exclusions affect events in summer and winter and are independent of resulting ERC.

Events caused by snow melt following a special dynamic of runoff and do not play a major role for the response behaviour in the study area.

Unsuitable events are eliminated from the data set to improve data quality. Eliminated events with very high discharge or precipitation are checked and in case of poor separation manually calculated and saved for analysis. The two events with the highest ERCs are verified by visual inspection and manual recalculation.

To represent all ERCs of one catchment, we use the Empirical Cumulative Distribution Function (ECDF). The ECDF is associated with the empirical measures of the sample, i.e. it estimates the true underlying distribution function of the points of a sample. Steep slopes of the ECDF indicate an independency of actual catchment conditions. Flat slopes of the ECDF indicate a high variability of ERCs influenced by actual catchment conditions.

2.3.2 Flow duration curves

For additional information about runoff behaviour concerning direct runoff data, we use Flow Duration Curves (FDC). The FDC is the complement of the cumulative distribution function of streamflow. In an FDC, discharge is plotted against exceedance probability and shows the percentage of time that a given flow rate is equalled or exceeded and provides a probabilistic description of stream flow at a given location. A steep slope of the FDC indicates flashiness of the stream flow response to precipitation input whereas a flatter curve indicates a relatively damped response and a higher storage (Yadav et al., 2007).

Normalization of runoff allows for a comparison of FDC and can be done by mean (e.g. Yadav et al., 2007), by median (e.g. Clausen and Biggs, 2000) or by catchment area. Opposite to common daily, monthly and annual FDCs (e.g. Vogel and Fennessey, 1994; Yadav et al., 2007), we use FDCs based on hourly discharge.

2.3.3 Indices

As homogenous regions may be found for almost any set of variables (Nathan and McMahon, 1990), the selection of the most appropriate set of indices constitutes an important step.

A single characteristic cannot describe all facets of catchment response behaviour. Therefore, we calculated a huge amount of indices which describe, seen by themselves, important aspects of runoff behaviour. Indices representing ERCs are statistically derived from data of all ERCs of one catchment, again separately for summer and winter. Indices from FDCs following the “signature measures” of Yilmaz et al. (2008) describe high, medium and low flow segments of

the FDCs. A correlation analysis shows very high correlations between many of these indices. High correlated indices (Spearman's $\rho > 0.8$; Spearman, 1904) do not bring new insights into the analysis and therefore were excluded. On the other hand, we consider an even distribution of indices with respect to season, high and low flow and the importance of indices. As a result of the weighting and selection processes 12 significant indices are identified (Table 1). Six indices describe mean and median ERCs, their variability and seasonality and two indices characterise the mid and high segment of the ECDFs. Four indices from the FDC describe catchment behaviour of high, intermediate and low flow. The values of these 12 indices characterise the unique runoff response behaviour of one catchment.

2.3.4 SOM

The 12 indices describing catchment response behaviour of one catchment (Table 1) can be seen as a vector that spans a high-dimensional data space. The data set of each catchment, which is called input vector, represents the response behaviour of a catchment. For exploratory data analysis of these high-dimensional data spaces of many catchments, we use a Self-Organizing Map (SOM) which is first described 1982 by Teuvo Kohonen (Kohonen, 1982). A SOM is an unsupervised learning algorithm based on artificial neural networks to produce a low-dimensional representation of a high-dimensional input data set. The goal of training a SOM is to present input vectors in an easily understandable form that preserves as much of the essential information in the data as possible. Kohonen (2001) provides a detailed description of the algorithm and its properties.

A SOM consists of centroids called neurons which are organized on a regular grid. Each neuron is represented by a prototype vector with the same dimension as the input vectors.

There are two primary quality properties of SOM: representing input data and data topology accuracy. The quality of data representation is usually measured using a quantisation error defined by Euclidian distance between input vectors and prototype vectors. The quantisation error should be as small as possible, but there is no universal reference for it. The prototype vector with the lowest quantisation error to a certain input vector is called the Best Matching Unit (BMU) and is the basis for assigning input vectors, and thus catchments, to neurons. Moreover, the quantisation error of the whole SOM measures the data representation by the average Euclidian distance between input vectors and their associated BMU. The topologic error measures the topologic representation: the percentage of two prototype vectors that are closest to a given input vector but are not neighbours on the map lattice.

The number and composition of neurons, which constructs the dimension of the SOM, can be determined automatically by means of a heuristic algorithm (Vesanto et al., 2000) or

Table 1. Indices describing catchment behaviour.

index	name	calculation	information
mean runoff coefficient	Mean	$\text{Mean} = \frac{1}{m} \sum_{j=1}^m \text{ERC}_j$ <p>ERC_j = event runoff coefficient of event <i>j</i> <i>m</i> = number of events</p>	overall feature of ERC, highly correlated to many other indices
coefficient of variation	CV	$\text{CV} = \frac{\sqrt{\frac{1}{m-1} \sum_{j=1}^m (\text{ERC}_j - \text{Mean})^2}}{\text{Mean}}$	variability of ERC
seasonality	Season	$\text{Season} = \frac{\text{mean Winter} - \text{mean Summer}}{\text{Mean}}$	differences between winter and summer
median Summer	MedianSm	Median ERC from Mai to October	central tendency of
median Winter	MedianWi	Median ERC from November to April	magnitude of ERC in summer or winter
coefficient of variation in Winter	CVWi	CV of ERC from November to April	variability of ERC in winter
slope of the ECDF, medium range	QSM	$\text{QSM} = \frac{0.8 \text{ quantile} - 0.2 \text{ quantile}}{\text{mean}}$	variability in medium range of runoff-coefficients
slope of the ECDF, high range	QSH	$\text{QSH} = \frac{1.0 \text{ quantile} - 0.8 \text{ quantile}}{\text{mean}}$	variability in high range of runoff-coefficients
high-flow segment volume of the FDC	MWH	$\text{MWH} = \frac{\sum_{h=1}^H Q_h}{H}$ <p><i>h</i> = 1, 2, ...<i>H</i> flows with exceedance probabilities <0.02</p>	watershed response to large precipitation events
low-flow segment volume of the FDC	MWL	$\text{MWL} = \frac{\sum_{l=1}^L Q_l}{H}$ <p><i>l</i> = 1, 2, ...<i>L</i> flows with exceedance probabilities 0.7–0.9</p>	long-term sustainability of flow (0.9 instead of 1.0 because of interpolated missing data 0.9 to 1.0)
mid segment slope of the FDC	MS	$\text{MS} = m2 - m7$ <p><i>m</i>2, <i>m</i>7: flow exceedance probability 0.2 and 0.7</p>	vertical redistribution of water
high segment slope of the FDC	HS	$\text{HS} = m0.005 - m2$ <p><i>m</i>0.005, <i>m</i>2: flow exceedance probability 0.005 and 0.2</p>	variability of watershed response to large precipitation events

manually. Small SOMs with few neurons often do not cover the complete variance of the input vectors and minimize interpretability; large maps perform a vague clustering.

If we want to cluster without a predefined number of clusters and get different levels of clustering that allow interpretations of cluster, we have to train a SOM with more neurons than the estimated number of clusters but small enough to define clusters. We choose the number of neurons after comparison of differently sized SOMs to get a SOM with a reasonable quantisation error, no topologic error, evident levels of clustering and a high interpretability.

Prior to the SOM training, each index has to be normalized to ensure that all indices have equal importance independent from their values. All indices are scaled linearly so that the variance of each is equal to one. Although indices can be weighted individually, we prefer to weight all indices equally. However, the choice of indices is a kind of weighting in itself.

The competitive and cooperative training algorithm compares the Euclidean distance of an input vector to all prototype vectors of the SOM. Next, the neuron with the most similar prototype vector and its neighbours are adjusted towards the input vector. At each training step of the SOM sequential training, one input vector is chosen randomly and processed to the SOM to update the closest and nearby neurons. In the course of the training, the prototype vectors and their neighbours are “tuned” to all input vectors. The neurons of the SOM gradually specialize to represent the input and become ordered on the map lattice with nearby neurons having similar data items. The final prototype vectors form a discrete approximation of the input data distribution (Herbst et al., 2009b).

As an alternative to the sequential approach, we use the batch training algorithm of the “SOM-Toolbox for Matlab 5”, described in Vesanto et al. (2000), Herbst and Casper (2008) and Herbst et al. (2009b). Instead of using a single data vector at a time, the whole data set is presented to the map before any adjustments are made. In each training step, the data set is partitioned according to the Voronoi regions of the prototype vectors, i.e. each input vector belongs to the data set of the neuron to which it is closest. After this, the prototype vectors are updated according to the weighted average of the input samples (Vesanto et al., 2000). The batch training algorithm speeds up the training process, does not need a learning rate factor and makes the SOM independent from a random choice of dataset at start of the training, which makes a SOM reproducible.

A SOM forms a semantic map where similar samples are mapped closer together and dissimilar apart and can be used for visualizing different features. Popular visualizations of SOMs, used in this paper are:

- a. U-matrix (Fig. 4a): visualizes distances between two neighbouring neurons as well as the median distance from each neuron to its neighbours, and thus shows the cluster structure of the map: high values indicate cluster borders; uniform areas of low values indicate clusters themselves.
- b. Component planes (Fig. 4b): show mean values of each index onto neurons and allows recognizing separating factors between clusters.
- c. Assignment of catchments to neurons (Fig. 4c): labels input vectors represented by catchment numbers to their corresponding BMU.
- d. Distribution of index properties for each neuron (Fig. 7) demonstrated by bar charts.

All visualizations are linked by position: in each figure, a certain position corresponds to the same neuron.

2.3.5 Hierarchical clustering

Hierarchical clustering is also an unsupervised method like a SOM. In its agglomerative approach it starts with single data points as individual clusters and at each step merges the closest pair of clusters until one cluster remains. This requires the definition of cluster proximity. In this study we use the unweighted group average distance which defines cluster proximity as the average pairwise proximity among all pairs of points in different clusters. Usually the result is displayed as a tree-like diagram called dendrogram which displays both the cluster-subcluster relationships and the order in which the clusters were merged (Tan et al., 2006). The lengths of the limbs of the dendrogram show the proximity of cluster or points. Data items can be clustered by cutting the dendrogram at a desired level.

3 Results

3.1 Runoff coefficients and flow duration curves

A total of 8259 Event Runoff Coefficients (ERCs) for 53 catchments were analysed, which corresponds to 100 to 200 events per watershed. There is a large variability in the ERCs. Figure 2 shows the distribution of the ERCs for 18 randomly selected catchments for a yearly (a) and seasonal (b and c) analysis. The 18 selected catchments represent a wide range of catchment behaviour and show differences of ERC between catchments. For lack of space, it is not possible to show all 53 boxplots in this figure.

Single runoff coefficients range from 0.003 to 0.89. The yearly mean runoff coefficients range from 0.03 to 0.36. Most catchments show a strong seasonal trend with higher ERCs in winter (Fig. 2b) and lower ERCs in summer (Fig. 2c). Mean runoff coefficients in summer range from 0.02 to 0.15, while in winter they range from 0.05 to 0.56. Although, ERCs tend to be low in summer,

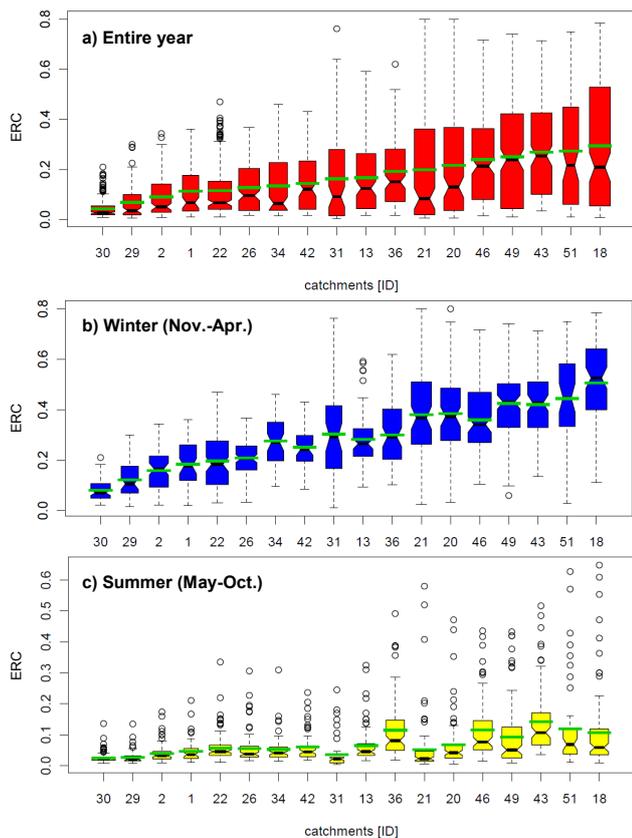


Fig. 2. Boxplots of ERC for 18 catchments for (a) the entire year, (b) winter and (c) summer. The boxplots show the 25th and 75th percentile, the median, values inside the interquartile range from the box and outliers of the event runoff coefficients. The green lines indicate the mean ERC. The widths of the boxes are proportional to the square-roots of the number of ERCs for this catchment.

some single summer events with larger ERCs occur in each catchment.

These outliers are caused by intensive convective precipitation events often in combination with heavy thunderstorms. Variability of ERCs tend to be larger in winter than in summer, but most catchments show less skewed distributions and relative short boxes which indicate the 25th and 75th percentiles in Fig. 2b. Thus, during winter, 50 % of the events of a catchment reach an analogous ERC, whereas the other 50 % display a much higher variability.

Event durations are between less than one day and many days.

The Empirical Distribution Functions (ECDFs) in Fig. 3a display the distribution of ERCs for each catchment. The lower part of the ECDFs is dominated by very low ERC that occur in summer. The ECDFs show characteristic shapes for each catchment.

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. Normalization 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 in this study. Therefore, runoff for FDCs is normalized by its median value ($Q_{.5}$) to eliminate most of the influences of mean annual precipitation, catchment area and extreme high or low flow and get well distinguishable FDCs. The FDC (Fig. 3b) show a wide range of runoff behaviour in all segments from flat curves to steep curves. Moreover, even in the geographically restricted region which includes nested catchments there are quite different response characteristics suitable for comparison and clustering.

3.2 Clustering by catchment response behaviour

In this paper two or more catchments are assumed to behave similar if similarity in the ECDFs and FDCs are given. Catchments with similar response behaviour are grouped by training a SOM and implementing hierarchical clustering on the SOM.

The dimension of the SOM is manually adjusted to 30 neurons on a 5×6 grid. This dimension has optimal explanatory power of the U-matrix (Fig. 4a) and a topological error of zero.

The SOM grid with the assignment of 53 available catchments to neurons (Fig. 4c) shows 25 neurons labelled with up to four catchment numbers. Catchments that are labelled together on one neuron can be considered as behaving similar and build the smallest unit of clustering. SOM reduces the variability of 53 catchments to 25 neurons. To further reduce the number of groups we apply a cluster analysis. The U-matrix (Fig. 4a) shows obviously different parts of the map: two blue-coloured regions in the upper part indicate units with a high degree of similarity, which can be seen as separate clusters. A row of red and orange colours separates the two bottom rows of neurons from the remaining map and builds a cluster border. Further visual clustering or setting of cluster borders is difficult. Therefore, we perform hierarchical clustering on the SOM. The dendrogram (Fig. 5) shows the arrangement of five reasonable clusters, which are named with capital letters A to E. The transfer of the cluster from the dendrogram to the U-matrix of the SOM is displayed in Fig. 5 as well. With the help of the assignment of catchments to neurons (Fig. 4c) we can relate catchments to the clustering. Despite overlapping ECDFs and FDCs (Fig. 6), the clusters are clearly visible in both diagrams. The spatial distribution of clustered catchments (Fig. 6) shows a grouping of catchments with similar response behaviour, but also some catchments which do not belong to a group.

Quantisation errors of input vectors to their Best Matching Unit (BMU) indicate the adequateness of representing catchment data by the BMU. Most of the input vectors have

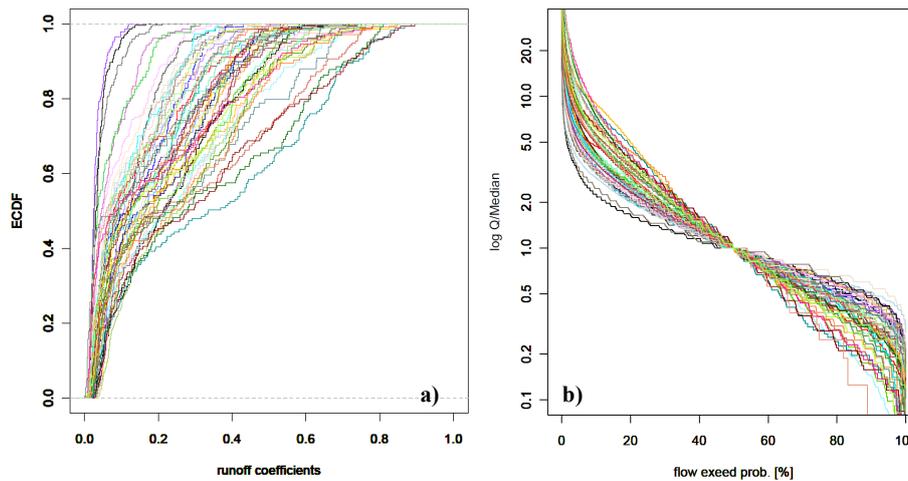


Fig. 3. Catchments response behaviour of 53 catchments for the entire year: (a) empirical cumulative distribution functions of event runoff coefficients and (b) flow duration curves, normalized by median.

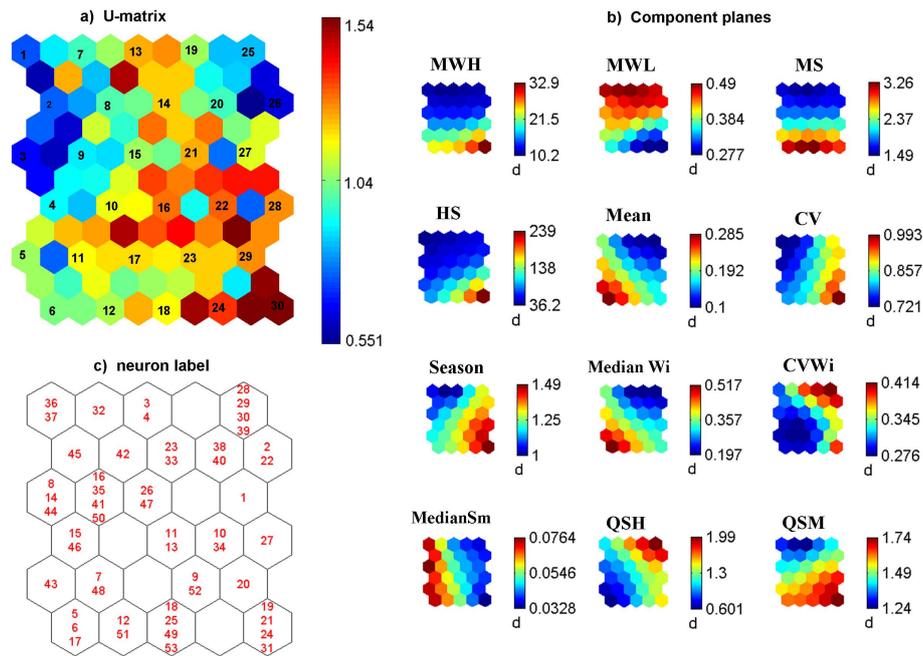


Fig. 4. Visualizations of a SOM: (a) unified distance matrix: neurons of the SOM are indicated by numbers. It shows structures by visualising distances between neighbouring neurons and, on additional hexagons between neurons, median distances between two neurons. (b) Component planes for each index display mean values of each index on the neurons of the SOM. (c) Assignment of catchments to neurons labels catchment IDs to their corresponding BMU1. All visualizations of SOMs are linked by position: in each figure, a certain position corresponds to the same neuron.

quantisation errors less than 2.7, 58% are below 1.6. For comparison: the general quantisation error is 1.64. But four catchment data sets have very high quantisation errors between 2.98 and 4.43. These catchments (3, 4, 21 and 31), labelled to neurons at the edge of the SOM, can be considered having special response behaviour, not very well covered by the SOM. They display extreme values for most of

the indices. Because every two of these catchments match to one neuron they can be considered forming two sub-clusters.

With the information from the component planes (Fig. 4b) and the distribution of index properties for each neuron (Fig. 7), we can characterise each cluster by a special combination of aspects of the response behaviour, listed in Table 2. All clusters except cluster B show either very high or

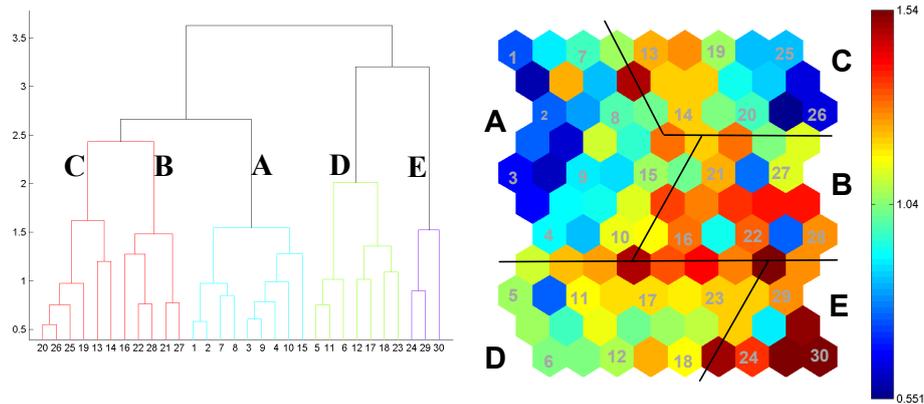


Fig. 5. Dendrogram on SOM and U-matrix with cluster. Numbers in the dendrogram correspond to neuron numbers on the U-matrix; capital letters indicate clusters.

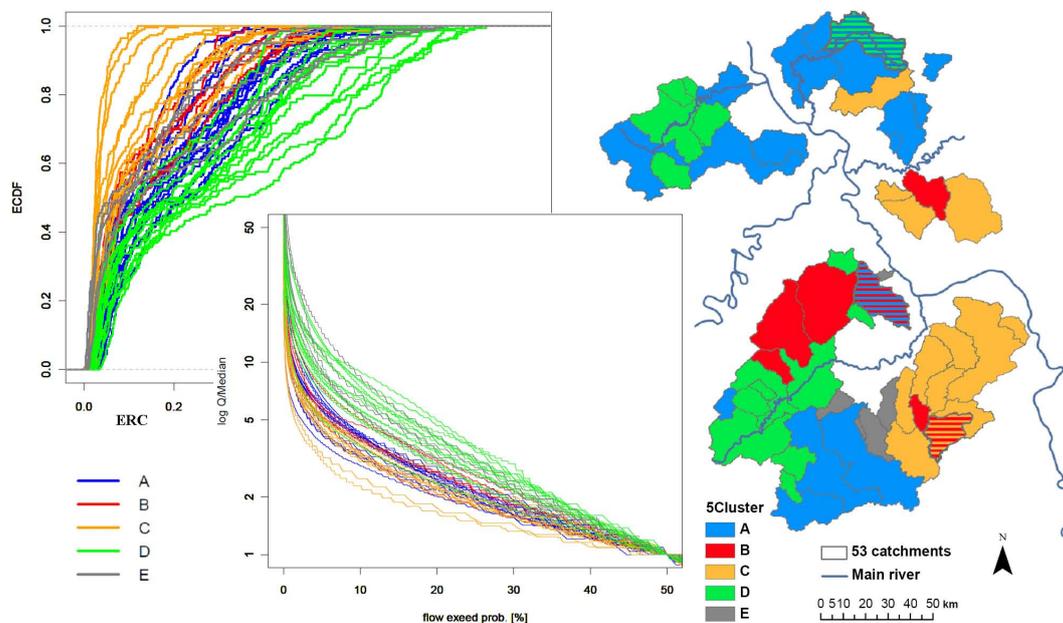


Fig. 6. Clusters by catchment response behaviour: empirical distribution function of event runoff coefficients, flow duration curve (limb < 50 %) and spatial distribution. The striped catchments in the map are borderline cases, shown in colours of their first and second best cluster. Note, some catchments are nested, so their areas overlap. Smaller catchments are plotted on top of larger ones.

low values of one or more characteristic aspect of runoff response, which makes them distinguishable from each other. This discernibility is a clear indication of the potential of the SOM to build classes of similar behaviour. Cluster B shows a medium response behaviour indicated by medium values of all indices. These clusters will be discussed and interpreted together with physical catchment properties in Sect. 3.4 (Comparison).

A possible disadvantage to our and many other methods of classification is the certainty of the allocation of catchments to a particular class. Classes of similar catchments by response behaviour or by physical catchment characteristics do not show sharp borders, especially for catchments in

a geographically restricted area. This is the case with most classifications in hydrology and other earth sciences where heterogeneity is involved (Merz and Blöschl, 2003). These cases may cause allocation of similar catchments to different clusters. To overcome this, Hall and Minns (1999) used fuzzy c-means to draw attention to “borderline” cases having significant membership levels of more than one class and raised the question how to proceed with borderline cases. The topological correctness of the SOM in combination with the quantisation errors addresses this problem. A borderline catchment is identified as catchment that is labelled to neighbouring neurons with a not exceeding medium distance and belonging to different clusters. As a decisive criterion for

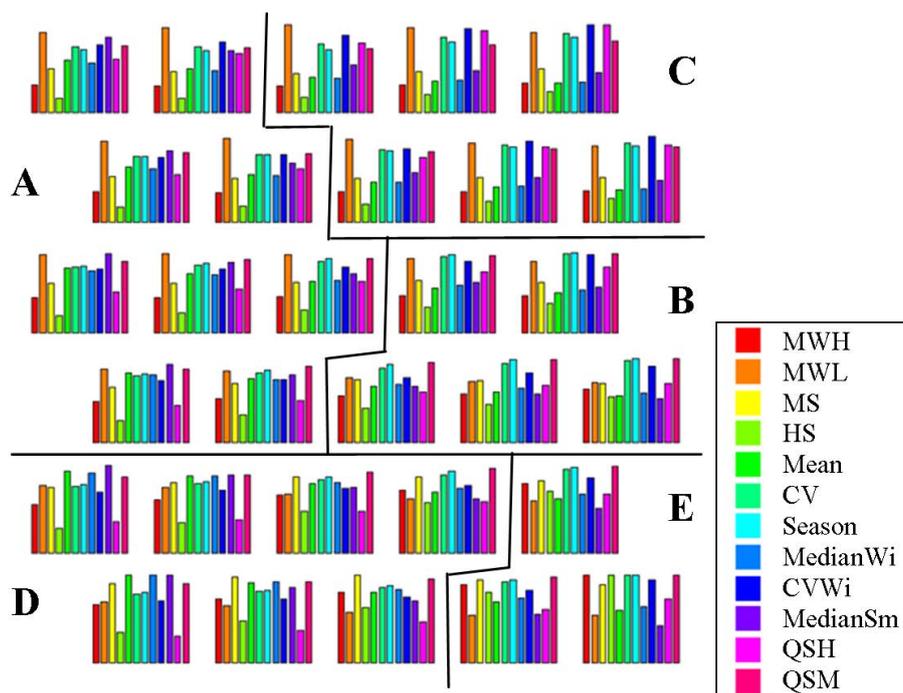


Fig. 7. Distribution of index properties for each neuron as bar charts. Lines indicate cluster border, letters cluster names.

Table 2. Characteristics of cluster A to E by catchment response behaviour.

cluster	characteristic			
	runoff coefficients	variability runoff coefficients	seasonality runoff coefficients	reactivity runoff
A	~ summer:+	-	-	-
B	--	+	+	--
C	-	~ winter++	--	-
D	++	--	--	+
E	+ summer:-	++	++	++

with: ++ very high, + high, ~ medium, -- low, - very low

borderline catchments we use their quantisation errors to the BMU and to the neuron of the neighbouring cluster. Catchments with a similar quantisation error to different neurons can be recognized as belonging to two neurons and thus may belong to two clusters. With this procedure, we can improve clustering and areal grouping of catchments (Fig. 6). We can identify three borderline cases (catchments 1, 26 and 43) from the clusters by response behaviour. 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.

3.3 Physical catchment properties

A list of physical catchment properties is compiled from a digital elevation model and different digital sources that describe catchment size, flow length, drainage density, topography, geology, soils, land use and climate of the catchments. All of them have more or less impact on runoff.

Yadav et al. (2007) recognised that using too many dynamical properties simultaneously often results in a rejection of all models, which is also our experience. Furthermore, different degrees of influence make it necessary to weight and choose certain groups of catchment characteristics. To reduce this amount of indices and to avoid redundancies, we eliminate all indices with high correlation coefficients (Spearman’s rho >0.8; Spearman 1904) to other indices. From the remaining indices, we choose indices with high correlation to indices of runoff behaviour considering physical indices identified in other studies as most important for runoff behaviour.

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 rho; Spearman, 1904) between size and MWL of 0.47 and HS -0.38, both indices are derived

from the flow duration curve. Correlation coefficients between catchment size and the other indices are considerably lower. Within the 46 catchments (87 %) smaller than 370 km², correlation coefficients show no correlation between catchment size and the 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.

The result of the selection process to find responsible catchment attributes for the study region are six meaningful catchment properties:

- Mean Annual Precipitation 1993–2008 (MAP),
- 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 (Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit, 2000),
- Wetness index: ratio of MAP to ET (Wet),
- mean slope derived from a digital elevation model with a resolution of 20 m (Slope),
- average field capacity derived from Bodenübersichtskarte 1:200 000 von Rheinland-Pfalz (BÜK 200) (FK),
- percentage of area with intense to medium ground water recharge, derived from Hydrologischer Atlas Rheinland-Pfalz and Geologische Übersichtskarte von Rheinland-Pfalz 1:300 000 (GUEK 300) (GW).

Climate and antecedent soil moisture are suggested as a major control in runoff generation (Sankarasubramanian and Vogel, 2002; Yadav et al., 2007; Merz and Blöschl, 2009). Therefore, half of the used indices describe climate. The wetness index as an expression of long-term water balance is highly correlated with MAP. Using MAP and the wetness index can also be seen as a weighting of MAP as the most important catchment property.

The spatial arrangement of clusters by behaviour (Fig. 6) follows a simplified gradient of increasing mean annual precipitation (MAP) from East to West and South to North. There is a high correlation between MAP and mean ERCs with a trend of increasing mean ERCs with increasing MAP. With increasing MAP, it becomes more likely that initial conditions are wet, thus enhancing runoff generation. The seasonal variability of ERCs (Fig. 2) reveals the changes in antecedent soil moisture and evapotranspiration. Climate has the largest influence on runoff behaviour not least by influencing drainage characteristics, controlling geomorphological structures, soils and vegetation (Sivapalan, 2005; Norbiato et al., 2009).

Pfister et al. (2002) and Hellebrand et al. (2007) found a strong relationship between winter storm flow coefficients and the permeability of the substratum for catchments in the Grand Duchy of Luxembourg and Rhineland-Palatinate which is represented here by the average field capacity and the ground water recharge.

Other catchment properties like flow length, catchment size, slope heterogeneity or land use show no important influence on the chosen indices of runoff behaviour or are eliminated because of high correlation to included indices.

Note, that because of not available data for all catchments, clustering of catchments by physical properties and comparison of catchments (Sect. 3.4) is based on 45 catchments.

Training a SOM with 4 × 7 neurons and performing hierarchical clustering on the SOM, the same method as used before, lead to 5 reasonable clusters of catchment properties (Fig. 8). The SOM grid with the labelled neurons (Fig. 8d) shows a clustering of the 45 catchments on 22 neurons. There are up to 4 catchments labelled on one neuron, building the smallest unit of clustering. The dendrogram (Fig. 8a) justifies the choice of five clusters, which are named with roman numerals I to V. With the help of the labelled neurons (Fig. 8d) and the component planes (Fig. 8c), we characterise these clusters as described in Sect. 3.4 (comparison).

Within the clusters by physical catchment properties we identify 5 borderline catchments: 1, B and C; 26, A and B; 27, B and E; 39, C and B; 43, D and A.

3.4 Comparison

The spatial variability of catchment response behaviour is supposed to be dependent on its physiographic and climatic characteristics. Following this a priori assumption, we search for corresponding clusters of catchments with similar physical catchment characteristics to illustrate the plausibility of the classification by response behaviour.

For each cluster of catchment behaviour, we counted the number of catchments matching a certain cluster of catchment properties (Table 3). Corresponding clusters show the highest overlap of catchments (Table 3, yellow). For most of the clusters, the number of shared basins allows a clear assignment. Only catchments from clusters B and IV cannot be clearly assigned to a corresponding cluster. Furthermore, clusters B and IV both represent medium response behaviour and medium catchment properties, i.e. they show reasonable corresponding catchment properties. Therefore, we assign cluster B to cluster IV. Attaching borderline catchments to their second best cluster (Table 3) supports the assignment B to IV.

The overlap between the two cluster sets is about 67 %. This result complies with Oudin et al. (2010) who found an overlap between hydrologically similar catchments and sets of physical catchment properties of 60 % in France and England. However, they used parameters of hydrological models and other physical catchment properties. Moreover, their

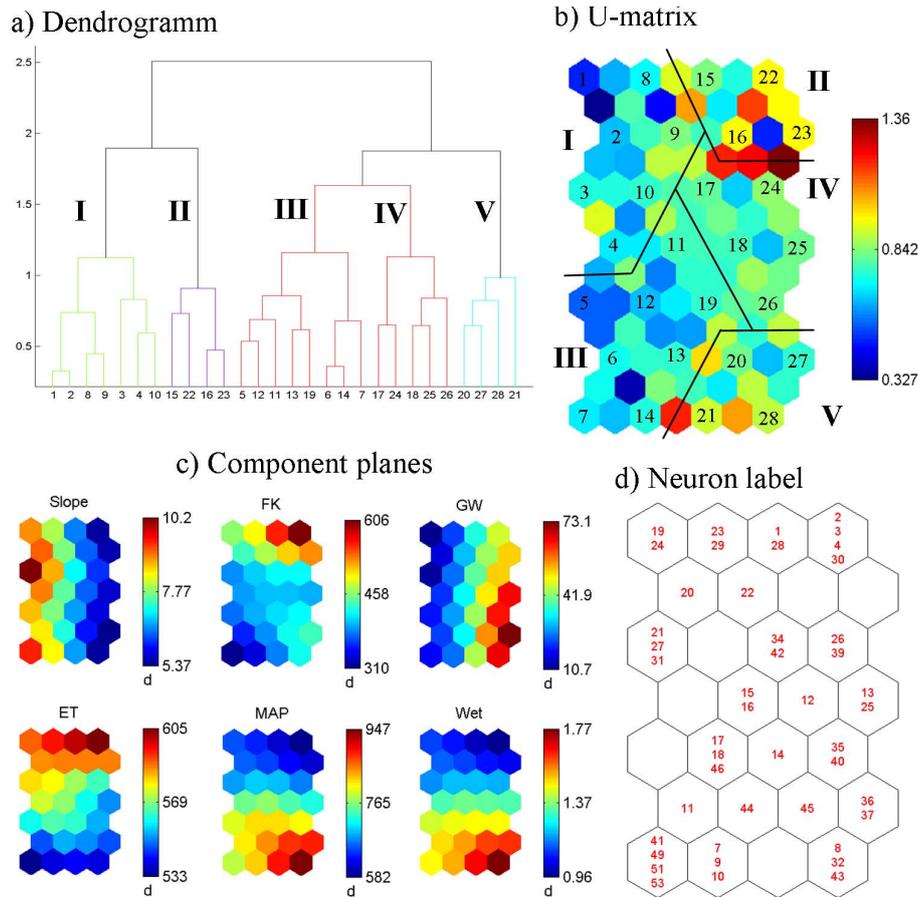


Fig. 8. SOM and clusters by physical catchment properties. **(a)** Dendrogram of SOM indicates cluster I to V, **(b)** unified distance matrix with neurons indicated by numbers and clusters; shows structures by visualising distances between neighbouring neurons and, on additional hexagons, median distances between two neurons. **(c)** Component planes for each index display mean values of the index onto the neurons of the SOM. **(d)** Assignment of catchments to neurons labels catchment IDs to their corresponding BMU1. In each figure, a certain position corresponds to the same neuron.

study area with 893 catchments is much larger than ours. This may indicate a possibility for classification and accompanying problems across scales.

Corresponding clusters show appropriate catchment properties (clusters I to V) and catchment response behaviour (clusters A to E):

- I: Very steep catchments with low water storage capacity, low MAP and high ET; it correspond to E with high ERCs and very high variability, seasonality and reactivity.
- II: Low slopes, high water storage capacity but dry with high ET; corresponds to C with very low ERCs and reactivity, low seasonality and medium variability but high variability in winter.
- III: High mean slopes, very low water storage capacities, high MAP and low to medium ET; it correspond to two quite different clusters: A and D. With these

catchment properties one would expect high ERCs and reactivity and low variability and seasonality – as in cluster D, not cluster A. In this case we can assume a lack of important physical characteristics in the analysis. Most of the catchments of cluster A are larger, shallower, have larger flow lengths, larger drainage density and have more agriculturally used areas and less forests than catchments of cluster D.

- IV: No exceptional values at all; it corresponds to B. Both clusters show medium values for all aspects of response behaviour as well as of physical catchment characteristics.
- V: Very flat catchments, very high ground water recharge and MAP, low ET; it corresponds to A with medium ERCs, very low variability, seasonality and reactivity.

Table 3. Overlap between clusters of response behaviour (A to E) and clusters of catchment properties (I to V). Numbers indicate the number of corresponding catchments. Numbers in brackets indicate the number of corresponding catchments considering borderline catchments assigned to their second best cluster. Corresponding clusters are marked in yellow.

catchment properties	response behaviour				
	A	B	C	D	E
I		1 (1)	3 (1)		5 (5)
II		1 (–)	5 (8)		
III	6 (7)	2 (2)		7 (8)	
IV	3 (–)	2 (3)	2 (2)	2 (1)	
V	5 (7)			1 (–)	

4 Discussion

Although the study area is comparatively small and homogeneous and there are many nested catchments, the indices describing catchment response behaviour show a high discriminating power and enable us to characterise catchments. Clustering of catchments by their response behaviour as well as by their physical properties leads to five clusters, each representing different facets of flow regimes or types of catchments. The comparison of the two pools of clusters show plausible classes of catchments. Each class of response behaviour represents different facets of flow regimes and can be characterised by high or low values of special aspects of response behaviour. Two classes are determined by steep flow duration curves and by the highest ERCs. This behaviour corresponds with steep catchment slopes and low water storage capacities. Different mean annual precipitation causes corresponding higher or lower runoff coefficients. Catchments with flatter flow duration curves display the lowest runoff coefficients. Moreover, due to their higher water storage capacities low seasonality and variability of ERCs are apparent. Within this group of catchments, differences in mean annual precipitation and potential evapotranspiration cause different runoff coefficients.

The comparison of clusters based on (i) response behaviour and (ii) physical catchment properties can be seen as a review of interpretability of the clustering by response behaviour. It allows connecting catchment behaviour with physical catchment properties as a basis for catchment classification covering both kinds of catchment characteristics. Most of the catchment characteristics, which are obtained by the classification method, can be interpreted well.

In Sect. 3.2, we define borderline catchments as catchments on the border of two clusters, near the overlap. We

can identify 3 borderline catchments in the clusters based on runoff behaviour and 5 borderline cases in the clusters based on physical catchment properties. If we assign these 8 borderline catchments to their second best cluster, we improve the overlap between the two pools of clusters from 67 % to 84 % (Table 3).

Thus, we divide the classified catchments into three groups:

1. catchments within the overlap: catchment behaviour is closely related to the considered physical catchment properties,
2. catchments near the overlap: catchment behaviour is related to the considered catchment properties and additional property(s),
3. catchments outside the overlap: individual physical catchment properties are responsible for catchment response behaviour.

The Classification of catchments in group (1) is clear. For the classification of catchments in group (2), we may define a differing class taking into account additional facts. Catchments of group (3) represent a subclass which is caused by special catchment properties. This subclass of catchments with similar behaviour but different physical catchment properties describes a so called “process equifinality” in the runoff signal (Hellebrand et al., 2011). Different processes producing the same values of discharge characteristics in different basins. 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. None of these catchments are borderline cases, both show very low to medium quantisation errors to their BMUs. The indices of both catchments are inside the range of indices of cluster C, they can be assumed to as behave similarly. On the U-matrix (Fig. 8b), clusters II and IV are neighbouring clusters but well separated by a row of neurons indicating high differences, showing 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.

The clusters in this study show geographic trends or groups, but no continuous regions. In the geographically restricted study area, the scale of diversity in landscape and in catchment characteristics is small. Therefore, the differences between catchments of different clusters are smaller than between clustered catchments in other landscapes, which are larger and more diverse. Classification of catchments has to consider scale and diversity of the respective region.

5 Conclusions

In this study, we define each five distinguishable classes of catchment behaviour and physical catchment properties by using SOMs. The comparison of the two pools of clusters, show a high overlap between both cluster sets. We take this as a sign that SOM has detected well structures inherent the catchment behaviour.

The SOM serves as a tool for unsupervised clustering, but also to analyse and visualize the clusters itself. In spite of using a default dimension of the SOM we train a SOM with 30 neurons, which is the optimal SOM dimension for the used data set. For other data sets, the optimal dimension of a SOM has to be found in an iterative process. Here, the large number of neurons makes hierarchical clustering on the SOM necessary to define cluster borders. On the other hand, a large number of neurons in combination with the topological correctness of SOMs offer a scope of interpretation of the clusters itself and their relationships, to find catchments with extreme response behaviour or to identify and discuss borderline cluster to improve clustering. Visualizations like the bar charts with index properties for each neuron in Fig. 7 or component planes in Fig. 4b describe clusters and identify indices causes clustering. Furthermore, SOMs allow to project any data set with the same dimensionality outside the training data set onto the map and thus to a certain cluster. Therefore a SOM can be used as a tool for classification of (new) data sets.

The presented approach allows the identification of clusters of similar catchment response behaviour and next, to classify catchments by their response behaviour with a view to physical catchment properties. A comparison with catchments clustered by physical properties shows a reasonable overlap which can be improved by analysis of SOMs. Still, a number of catchments show individual runoff response behaviour which cannot be explained by catchments properties considered here. The individuality of catchments limits the possibilities of simplifying runoff generation. To find a meaningful classification of catchments, both kinds of catchment characteristics have to be taken into account.

The classification presented in this study is valid for catchments in the study area in Rhineland-Palatinate. Catchments in other landscapes or at other scales may have other classification items because of influence, different processes have on the runoff response. However, the methodology presented in this paper is independent of landscape and scale.

Acknowledgements. This investigation was carried out within the Research Center for Regional and Environmental Statistics (Forum-Stat) at the University of Trier, financed by the Research-Initiative Rhineland-Palatinate. The authors wish to thank the State Office for Environment, Water Management and Trade Control (LUWG) Rhineland-Palatinate at Mainz for providing data of runoff and precipitation. We would like to thank Uwe Ehret, two anonymous reviewers and the editor Attilio Castellarin for their constructive comments that greatly helped improving this paper.

Edited by: A. Castellarin

References

- Acreman, M. C. and Sinclair, C. D.: Classification of drainage basins according to their physical characteristics: an application for flood frequency analysis in Scotland, *J. Hydrol.*, 84, 365–380, 1986.
- Beven, K. J.: Uniqueness of place and process representations in hydrological modelling, *Hydrol. Earth Syst. Sci.*, 4, 203–213, doi:10.5194/hess-4-203-2000, 2000.
- Blume, T., Zehe, E., and Bronstert, A.: Rainfall-runoff response, event-based runoff coefficients and hydrograph separation, *Hydrolog. Sci. J.*, 52, 843–862, 2007.
- Bundesministerium für Umwelt: Naturschutz und Reaktorsicherheit (Hrsg), *Hydrologischer Atlas von Deutschland*, 2000.
- Burn, D. H.: Catchment similarity for regional flood frequency analysis using seasonality measures, *J. Hydrol.*, 202, 212–230, 1997.
- Chapman, T. G. and Maxwell, A. I.: Baseflow Separation – Comparison of Numerical Methods with Tracer Experiments, I. E. Aust. Natl. conf. Publ. 96/05, 539–545, 1996.
- Clausen, B. and Biggs, B. J. F.: Flow variables for ecological studies in temperate streams: grouping based on covariance, *J. Hydrol.*, 237, 184–197, 2000.
- Di Prinzio, M., Castellarin, A., and Toth, E.: Data-driven catchment classification: application to the pub problem, *Hydrol. Earth Syst. Sci.*, 15, 1921–1935, doi:10.5194/hess-15-1921-2011, 2011.
- Gerlach, N.: INTERMET – Interpolation meteorologischer Größen, *Veranstaltungen 3/2006 – Niederschlag-Abfluss-Modellierung zur Verlängerung des Vorhersagezeitraumes operationeller Wasserstands- und Abflussvorhersagen, Kolloquium am 27 September 2005 in Koblenz, Bundesanstalt für Gewässerkunde*, 2006.
- Hall, M. and Minns, A.: The classification of hydrologically homogeneous regions, *Hydrolog. Sci. J.*, 44, 693–704, 1999.
- Hellebrand, H., Hoffmann, L., Juilleret, J., and Pfister, L.: Assessing winter storm flow generation by means of permeability of the lithology and dominating runoff production processes, *Hydrol. Earth Syst. Sci.*, 11, 1673–1682, doi:10.5194/hess-11-1673-2007, 2007.
- Hellebrand, H., Müller, C., Fencia, F., Matgen, P., and Savenije, H.: A process proof test for model concepts: modelling the meso-scale, *Phys. Chem. Earth*, 36, 42–53, 2011.
- Herbst, M. and Casper, M. C.: Towards model evaluation and identification using Self-Organizing Maps, *Hydrol. Earth Syst. Sci.*, 12, 657–667, doi:10.5194/hess-12-657-2008, 2008.
- Herbst, M., Gupta, H. V., and Casper, M. C.: Mapping model behaviour using Self-Organizing Maps, *Hydrol. Earth Syst. Sci.*, 13, 395–409, doi:10.5194/hess-13-395-2009, 2009a.
- Herbst, M., Casper, M. C., Grundmann, J., and Buchholz, O.: Comparative analysis of model behaviour for flood prediction purposes using Self-Organizing Maps, *Nat. Hazards Earth Syst. Sci.*, 9, 373–392, doi:10.5194/nhess-9-373-2009, 2009b.
- Kalteh, A. M., Hjorth, P., and Berndtsson, R.: Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application, *Environ. Modell Softw.*, 23, 835–845, 2008.

- Kaski, S.: Data Exploration Using Self-Organizing Maps, Dr. thesis, Department of Computer Science and Engineering, Helsinki University of Technology, Helsinki, 57 pp., 1997.
- Kohonen, T.: Self-Organize Formation of Topologically Correct Feature Maps, *Biol. Cybern.*, 43, 59–69, 1982.
- Kohonen, T.: Self-Organizing Maps, 3rd Edn., Information Sciences, Springer, Berlin, Heidelberg, New York, 501 pp., 2001.
- Maimon, O. and Rokach, L.: Data mining and knowledge discovery handbook, Springer, New York, 1383 pp., 2005.
- McDonnell, J. J. and Woods, R.: On the need for catchment classification, *J. Hydrol.*, 299, 2–3, 2004.
- Merz, R. and Blöschl, G.: A process typology of regional floods, *Water Resour. Res.*, 39, 1340, doi:10.1029/2002WR001952, 2003.
- Merz, R., Blöschl, G., and Parajka, J.: Spatio-temporal variability of event runoff coefficients, *J. Hydrol.*, 331, 591–604, 2006.
- Merz, R. and Blöschl, G.: A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria, *Water Resour. Res.*, 45, W01405, doi:10.1029/2008WR007163, 2009.
- Nathan, R. J. and McMahon, T. A.: Identification of homogenous regions for the purposes of regionalisation, *J. Hydrol.*, 121, 217–218, 1990.
- Norbiato, D., Borga, M., Merz, R., Blöschl, G., and Carton, A.: Controls on event runoff coefficients in the eastern Italian Alps, *J. Hydrol.*, 375, 312–325, 2009.
- Oudin, L., Kay, A., Andréassian, V., and Perrin, C.: Are seemingly physically similar catchments truly hydrologically similar?, *Water Resour. Res.*, 46, W11558, doi:10.1029/2009WR008887, 2010.
- Pfister, L., Iffly, J.-F., and Hoffmann, L.: Use of regionalized stormflow coefficients with a view to hydroclimatological hazard mapping, *Hydrolog. Sci. J.*, 47, 479–492, 2002.
- Ramachandra Rao, A. and Srinivas, V. V.: Regionalization of watersheds, An Approach based on Cluster Analysis, Springer Science + Business Media B. V., 2008.
- Sankarasubramanian, A. and Vogel, R. M.: Annual hydroclimatology of the United States, *Water Resour. Res.*, 38, 1–12, doi:10.1029/2001WR000619, 2002.
- Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendiondo, E. M., O'Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S., and Zehe, E.: IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences, *Hydrolog. Sci. J.*, 48, 857–880, doi:10.1623/hysj.48.6.857.51421, 2003.
- Sivapalan, M.: Pattern, Process and Function: Elements of a Unified Theory of Hydrology at the CatchmentScale, in: *Encyclopaedia of hydrological sciences*, edited by: Anderson, M., John Wiley, London, 193–219, 2005.
- Spearman, C.: The Proof and Measurement of Association between Two Things, *Am. J. Psychol.*, 15, 72–101, 1904.
- Stedinger, J. R., Vogel, R. M., and Foufoula-Georgiou, E.: Frequency analysis of extreme events, in: *Handbook of Hydrology*, edited by: Maidment, D. R., McGraw-Hill, New York, p.18.35, 1992.
- Tan, P.-N., Steinbach, M., and Kumar, V.: *Introduction to Data Mining*, Pearson Education, Addison Wesley, Boston, 769 pp., 2006.
- Toth, E.: Classification of hydro-meteorological conditions and multiple artificial neural networks for streamflow forecasting, *Hydrol. Earth Syst. Sci.*, 13, 1555–1566, doi:10.5194/hess-13-1555-2009, 2009.
- van Dijk, A. I. J. M.: Climate and terrain factors explaining streamflow response and recession in Australian catchments, *Hydrol. Earth Syst. Sci.*, 14, 159–169, doi:10.5194/hess-14-159-2010, 2010.
- Vesanto, J., Himberg, J., Alhoniemi, E., and Parhankangas, J.: *SOM Toolbox for Matlab 5*, Helsinki University of Technology, Report A59, Espoo, Finland, 60 pp., 2000.
- Vogel, R. M. and Fennesey, N. M.: Flow Duration Curves I: New Interpretation and Confidence Intervals, *J. Water Res. Pl.-ASCE*, 120, 485–504, 1994.
- Wagener, T., Sivapalan, M., Troch, P. and Woods R.: Catchment Classification and Hydrologic Similarity, *Geography Compass*, 1/4, 901–931, 2007.
- Woods, R.: Seeing catchments with new eyes, *Hydrol. Process.*, 16, 1111–1113, 2002.
- Yadav, M., Wagener, T., and Gupta, H.: Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins, *Adv. Water Resour.*, 30, 1765–1774, 2007.
- Yilmaz, K. K., Gupta, H. V., and Wagener, T.: A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model, *Water Resour. Res.*, 44, W09417, doi:10.1029/2007WR006716, 2008.