Hydrol. Earth Syst. Sci. Discuss., 12, 11223–11268, 2015
www.hydrol-earth-syst-sci-discuss.net/12/11223/2015/
doi:10.5194/hessd-12-11223-2015


Hydrology and Earth System Sciences Discussions
Open Access



# Simultaneous calibration of hydrological models in geographical space

**A. Bárdossy[1], Y. Huang[1], and T. Wagener[2]**

[1]Institute for Modelling Hydraulic and Environmental Engineering, University of Stuttgart, Stuttgart, Germany
[2]Department of Civil Engineering, Queen's School of Engineering, University of Bristol, Bristol, UK

Received: 3 October 2015 – Accepted: 5 October 2015 – Published: 30 October 2015

Correspondence to: A. Bárdossy (andras.bardossy@iws.uni-stuttgart.de)

Published by Copernicus Publications on behalf of the European Geosciences Union.

## Abstract

Hydrological models are usually calibrated for selected catchments individually using specific performance criteria. This procedure assumes that the catchments show individual behavior. As a consequence, the transfer of model parameters to other ungauged catchments is problematic. In this paper, the possibility of transferring part of the model parameters was investigated. Three different conceptual hydrological models were considered. The models were restructured by introducing a new parameter $\eta$ which exclusively controls water balances. This parameter was considered as individual to each catchment. All other parameters, which mainly control the dynamics of the discharge (dynamical parameters), were considered for spatial transfer. Three hydrological models combined with three different performance measures were used in four different numerical experiments to investigate this transferability. The first numerical experiment, individual calibration of the models for 15 selected MOPEX catchments, showed that it is difficult to identify which catchments share common dynamical parameters. Parameters of one catchment might be good for another catchment but not reversed. In the second numerical experiment, a common spatial calibration strategy was used. It was explicitly assumed that the catchments share common dynamical parameters. This strategy leads to parameters which perform well on all catchments. A leave one out common calibration showed that in this case a good parameter transfer to ungauged catchments can be achieved. In the third numerical experiment, the common calibration methodology was applied for 96 catchments. Another set of 96 catchments were used to test the transfer of common dynamical parameters. The results show that even a large number of catchments share similar dynamical parameters. The performance is worse than those obtained by individual calibration, but the transfer to ungauged catchments remains possible. The performance of the common parameters in the second experiment was better than in the third, indicating that the selection of the catchments for common calibration is important. In the fourth numerical experiment, the common parameters obtained

from the 96 US catchments were used to model two selected German catchments. The results indicate that the dynamical model parameters have skill even under very different conditions.

## 1 Introduction

Hydrological models are widely used to describe catchment behavior, and for subsequent use for water management, flood forecasting and other purposes. Hydrological modeling is usually done for catchments with observed precipitation and discharge data. The unknown (and partly not measurable) parameters of a conceptual or to some extent physics-based model are adjusted in a calibration procedure to reproduce the measured discharge from the observed weather and catchment properties. Due to the high variability of catchment properties and hydrological behavior (Beven, 2000), this modeling procedure is usually performed individually for each catchment. Different catchments are often modeled using different models. This great variety of models and catchments makes a generalization of the description of the hydrological processes very challenging (Sivapalan, 2003). Additionally, even for a selected model applied for a specific catchment, the parameter identification is not unique, a great number of parameter vectors might lead to the very similar performance (Beven and Freer, 2001).

Hydrological modeling of catchments without measured discharge is big challenge, i.e. the so called predictions in ungauged basins problem (Sivapalan, 2003). Instead of model calibration, parameters have to be estimated on the basis of other information. A great number of interesting methods have been developed for this purpose. An excellent summary can be found in Blöschl et al. (2013).

One possibility of dealing with this problem is through catchment classification. It is attempted to identify groups of catchments which behave similarly. Based on catchment properties, it is assumed that catchments form groups whose members can be modeled in a similar way (Seibert, 1999; Oudin et al., 2008). There have been





HESSD
12, 11223–11268, 2015



Title Page
Abstract Introduction
Conclusions References
Tables Figures
◄ ►
◄ ►
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

many attempts to develop classification schemes (Grigg, 1965; Sawicz et al., 2011). However, the task is of great importance, McDonnell and Woods (2004) discuss the need for a widely accepted classification system. Among others, a good classification would help to model the rainfall–runoff process for ungauged catchments (Wagener et al., 2007).

Catchment similarity can be described by comparing their corresponding discharge series. Correlations (Archfield and Vogel, 2010) or copulas (Samaniego et al., 2010) can be used for this purpose. Much of the variability in discharge time series is controlled by weather patterns. Therefore, it is likely that similarity in discharge is higher for catchments with well correlated weather, which often requires geographical closeness (Archfield and Vogel, 2010). However, discharge series produced by catchments can be very different under different meteorological conditions. Even the same catchment behaves differently in a dry and in a wet year. Due to the different weather forcing, the above methods would consider the same catchment in one time period as dissimilar to itself in another time period.

One can also define catchment similarity using hydrological models (McIntyre et al., 2005; Oudin et al., 2010). Catchments are similar if they can be modeled reasonably well by the same model using the same model parameters. Due to observation errors and specific features in the calibration period, the adjustment of the model can be very specific of the observation period leading to an overcalibration (Andréassian et al., 2012).

The focus of this paper is to investigate if the transformation of precipitation to discharge possible independently of the weather. For this purpose, the hydrological model parameters are separated into two groups:

- Parameters describing the water balances, which are strongly related to climate.
- Parameters describing the dynamics of the runoff triggered by weather.

The second group of parameters is supposed to be weather independent and represent the focus of this paper. To simplify the problem, a single new parameter





HESSD
12, 11223–11268, 2015

$\eta$ was introduced to describe water balance. This parameter is conditional on the other model parameters and adjusts the long term water balances.

The purpose of this paper is to investigate to what extent do different catchments share a similar dynamical rainfall–runoff behavior and can be modeled using the same model parameters with exception of the newly introduced individualized water balance parameter $\eta$.

Hydrological models are usually judged according to the degree of reproducing discharge dynamics and water balances. While water balances are mainly driven by weather in terms of precipitation, temperature, radiation and wind. Dynamics is controlled by catchment properties in terms of size, terrain, slopes, soils etc. Landscapes are formed during long time through climate, and are thus in a kind of quasi equilibrium. The hypothesis of this paper is that this equilibrium is mirrored in a similar dynamic behavior. Thus, a large number of catchments can be modelled by using the same dynamic parameters.

Three simple conceptual hydrological models combined with three different performance measures are used to describe the rainfall–runoff behavior on the daily time scale for a large number of catchments.

The following four different numerical experiments, including calibration and validation procedures, are carried out for different sets of selected catchments:

1. The usual catchment by catchment calibration is carried out. In order to test if dynamical model parameters are shared, the parameters are directly transferred to all of other catchments.

2. Instead of the traditional catchment by catchment calibration, it is assumed that the model parameters are similar for a set of catchments in a close geometrical setting. Thus a simultaneous calibration of the models is carried out and tested both in a gauged and an ungauged version.





HESSD
12, 11223–11268, 2015

3. The geographical extent of the catchments used for simultaneous calibration is expanded. A great number of assumed ungauged catchments are used for testing the hypothesis.

4. Finally, the transferability of the model parameters to catchments under very different climatic and geographical conditions is tested.

The hypothesis is that the rainfall–runoff process can be described using the same dynamical hydrological model parameters for a number of catchments. The very different climatic conditions and water balances of the catchments are considered by the newly introduced specific parameter $\eta$ controlling the long term water balance of each catchment individually. The other model parameters control the discharge dynamics on both short and long time scales. These dynamical parameters are supposed to be shared despite the great heterogeneity of the catchments. This procedure simplifies the hydrological model parameter estimation for ungauged catchments, namely the procedure is reduced to the estimation of a single parameter $\eta$, which can be related to long term water balances.

The paper is structured as follows: after the introduction, the investigation area is described. This is followed by a description of the three conceptual hydrological models and the three performance criteria used for calibration and validation. In section four, the new model parameter $\eta$ controlling the water balance is introduced. In section five to eight, four numerical experiments are described and the results are presented, starting with the individual calibration of the models and ending with a transfer of the model parameters to randomly selected catchments. The paper concludes with a discussion of the results.

## 2 Investigation area and available data

The study area is the eastern United States. Locations of the 196 catchments used in this study are shown in Fig. 1. The catchments for a subset used for the



Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

12, 11223–11268, 2015

Simultaneous calibration

A. Bárdossy et al.

international Model Parameter Estimation Experiment (MOPEX) project. Catchments range in size from 134 to 9889 $km^2$ and exhibit aridity indices (long-term potential evapotranspiration to precipitation rates) between 0.41 and 3.3, hence representing a heterogeneous dataset. Time-series data of daily streamflow, precipitation, and temperature for all catchments were provided by the MOPEX project (Duan et al., 2006). Catchments within this dataset are minimally impacted by human influences. Streamflow information within this dataset was originally provided by the United States Geological Survey (USGS) gauges, while precipitation and temperature was supplied by the National Climate Data Center (NCDC). The MOPEX dataset has been used widely for hydrological model comparison studies (see references in Duan et al., 2006).

## 3 Hydrological models and performance criteria

Three simple conceptual hydrological models were applied in this study. The reason for this is that the great number of calibration and validation experiments (totally $> 10^{10}$-year discharge calculations) could only be performed with relatively simple model structures. We considered it is important to see if the results are similar for different models and performance measures. In a subsequent study, spatially distributed models will be considered.

### 3.1 HYMOD

The HYMOD (Boyle et al., 2001) is a conceptual rainfall–runoff model derived from the Probability Distributed Model (Moore, 1985). The soil moisture accounting module of HYMOD utilizes a Pareto distribution function of storage elements of varying sizes. The storage elements of the catchment are distributed according to a probability density function defined by the maximum soil moisture storage CMAX and the distribution of soil moisture stores *b* (Wagener et al., 2001). Evaporation from the soil moisture store occurs at the rate of the potential evaporation estimates using the Hamon approach.





HESSD
12, 11223–11268, 2015

Following evaporation, the remaining rainfall and snowmelt is used to fill the soil moisture stores. A routing module divides the excess rainfall using a split parameter $\alpha$ which separates fluxes amongst two parallel conceptual linear reservoirs meant to simulate the quick and slow flow response of the system (defined by residence times $k_q$ and $k_s$).

## 3.2 HBV

The HBV model is a conceptual model and was originally developed at the Swedish Meteorological and Hydrological Institute (SMHI) (Bergström and Forsman, 1973). Snow accumulation and melt, actual soil moisture and runoff generation are calculated using conceptual routines. The snow accumulation and melt is based on the degree-day approach. Actual soil moisture is calculated by considering precipitation and evapotranspiration. Runoff generation is estimated by a non-linear function of actual soil moisture and precipitation. The dynamics of the different flow components at the subcatchment scale are conceptually represented by two linear reservoirs. The upper reservoir simulates the near surface and interflow in the sub-surface layer, while the lower reservoir represents the base flow. They are connected through a linear percolation rate. Finally, there is a transformation function consisting of a triangular weighting function with one free parameter for smoothing the generated flow.

## 3.3 Xinanjiang model (XAJ)

The XAJ model was established in the early 1970s in China. This conceptual rainfall–runoff model has been applied to a large number of basins in the humid and semi-humid regions in China. The lumped version of XAJ model consisted of four main components (Zhao, 1995). The evapotranspiration is represented by a 3-layer soil moisture module which differentiates upper, lower and deeper soil layers. Runoff production is calculated based on rainfall and soil storage deficit, tension water capacity curve is introduced to provide for a non-uniform distribution of tension water capacity throughout the whole catchment. The runoff separation module separates the determined runoff into three parts, namely surface runoff, interflow and groundwater. The flow routing module transfers the local runoff to the outlet of the basin. In order to account for the precipitation that is contributed from snowmelt, the degree-day snowmelt approach is added in this model. In this study, the model has 16 parameters which can be adjusted using calibration.

### 3.4 Performance criteria

Model calibration depends strongly on the performance criteria used. In order to obtain reasonably general results, three different criteria were selected to evaluate model performance.

The Nash–Sutcliffe Efficiency (Nash and Sutcliffe, 1970) between the observed and modeled flow is most frequently taken as the first evaluation criterion:

$$O^{(1)} : \mathrm{NS} = 1 - \frac{\sum_{T=1}^{T}(Q_\mathrm{o}(t) - Q_\mathrm{m}(t))^2}{\sum_{T=1}^{T}\left(Q_\mathrm{o}(t) - \overline{Q_\mathrm{o}}\right)^2} \tag{1}$$

Here $Q_\mathrm{o}(t)$ is the observed discharge and $Q_\mathrm{m}(t)$ is the modeled discharge on a given day $t$. The abbreviation NS is used subsequently for this performance measure.

The NS model performance criterion was often criticized (for example in Schaefli and Gupta, 2007), and several modifications and other criteria were suggested. One interesting suggestion was published in Gupta et al. (2009), the authors suggest using a performance measure which accounts for the water balances and the correlation of the observed and modeled time series separately. Their approach was slightly modified and the following performance criterion was introduced:

$$O^{(2)} : \mathrm{GK} = 1 - \beta\left(\frac{\sum_{T=1}^{T}(Q_\mathrm{o}(t) - Q_\mathrm{m}(t))}{\sum_{T=1}^{T}Q_\mathrm{o}(t)}\right)^2 - (1 - r(Q_\mathrm{o}, Q_\mathrm{m}))^2 \tag{2}$$





HESSD
12, 11223–11268, 2015
Simultaneous calibration
A. Bárdossy et al.

Here $r(Q_o, Q_m)$ is the correlation coefficient between the observed and modeled time series of discharge. $\beta$ is a weight to express the importance of the water balance. In our study, $\beta = 5$ was selected. The reason for selecting this version of the coefficient is that a model should produce good water balances and appropriate discharge dynamics simultaneously. The quadratic form in Eq. (2) assures that both aspects are considered, and the worse of them is dominating. The abbreviation GK is used subsequently for this performance measure.

The Nash–Sutcliffe coefficient of the logarithm of the discharges is focusing on the low flow conditions more than the traditional NS coefficient:

$$L_{NS} = 1 - \frac{\sum_{T=1}^{T} (\log(Q_o(t)) - \log(Q_m(t)))^2}{\sum_{T=1}^{T} \left(\log(Q_o(t)) - \overline{\log(Q_o)}\right)^2} \tag{3}$$

To equally concentrate on high and low flows, a combination of the original NS and the logarithmic NS is used as a third measure:

$$O^{(3)} : \mathrm{NS} + \mathrm{LNS} = \frac{\mathrm{NS} + L_{NS}}{2} \tag{4}$$

The abbreviation NS + LNS is used subsequently for this performance measure.

The three performance criteria were modified, hence the higher the value the better the model. Further the best value for the criteria is 1.

## 4 Method

### 4.1 Model parameter to control water balance

Climatic conditions are of central importance for water balances. The relationship of potential to actual evapotranspiration can differ strongly due to water or energy limitations. This suggests that catchments might have similar dynamical behavior but





HESSD
12, 11223–11268, 2015

with different water balances. In order to account for this, the model parameters could be separated to form two groups, one group with parameters controlling the water balances and another controlling the discharge dynamics. This separation of existing model parameters is difficult, as they often influence simultaneously both components. Instead of an artificial model specific separation, a new parameter $\eta$ was introduced to all three models. This parameter controls the ratio between daily potential and actual evapotranspiration depending on the available water and depends on the long term water balance only. This parameter $\eta$ gives:

$$E_{ta} = \begin{cases} E_{tp} & \text{if } \frac{\text{SM}}{\text{CMAX}} > \eta \\ \min\left(\frac{\text{SM}}{\eta \cdot \text{CMAX}} E_{tp}, \text{SM}\right) & \text{else} \end{cases} \quad (5)$$

Here SM is the actual soil water available for evapotranspiration. CMAX is the maximum possible soil moisture. $E_{tp}$ stands for the potential and $E_{ta}$ for the actual evapotranspiration, respectively.

The parameter $\eta$ regulates the water balances in accordance with the dynamical parameters. It can be calculated directly for each parameter vector $\boldsymbol{\theta}$. This is necessary as it is thought to establish correct water balances. Thus it is a catchment and parameter vector dependent parameter. $f(\eta) = V_{iM}(\eta, \theta)$ is monotonically decreasing function of $\eta$. If the model can provide correct long term water balances then:

$$V_{iM}(1,\theta) < V_{iO} < V_{iM}(0,\theta) \quad (6)$$

As $f(\eta) = V_{iM}(\eta, \theta)$ is continuous, there is a unique $\eta(\theta)$ for which:

$$V_{iM}(\eta(\theta),\theta) = V_{iO} \quad (7)$$

If Eq. (6) is not fulfilled, then the parameter vector $\boldsymbol{\theta}$ is not appropriate for the model.

The parameter $\eta$ is fitted individually for each $\boldsymbol{\theta}$, in this way a correct water balance is assured for the calibration period.





HESSD
12, 11223–11268, 2015

## 4.2 Experimental design

In this study, the ROPE algorithm (Bárdossy and Singh, 2008) was applied for model parameter optimization. Each calibration yielded 10 000 convex sets of good parameters. Four numerical experiments on a large number of catchments were carried out to investigate the transferability of the model parameters under different calibration strategies. For a clear explanation and understanding of the methods, the procedure and results for these four experiments are presented in the following four sections.

# 5 Numerical experiment 1: individual calibration and parameter transfer

The first experiment is thought to test the transferability of the model parameters under the usual individual calibration for each catchment.

As a first step, 15 catchments with reliable data and slightly varying catchment properties in the eastern United States were selected.Locations of the selected gauges are marked as red plus on Fig. 1. Table 1 lists the basic catchment properties and Table 2 summarizes the meteorological conditions for the selected 15 catchments, respectively (Falcone et al., 2010). The tables show that despite their geographical proximity, these catchments have quite different weather and hydrographic properties.

For the 15 selected catchments, an individual calibration was performed using all three models and all three performance measures. Data series from year 1951 to 2000 were split up into 5 sub-periods. This leads to 45 calibrations for each catchment. Each calibration yielded convex sets $\mathcal{G}_i$ of good parameters for each catchment $i$. 10 000 parameter vectors from each of these sets were generated. (Note that the corresponding parameter $\eta$ was estimated for each element of the parameter set separately.)

Let $O_i^{(j)}(\theta)$ denote the value of the objective function $j$ for a parameter vector $\boldsymbol{\theta}$ in catchment $i$. The best objective function value for each individual catchment is denoted with $O_i^{(j)*}$.





HESSD
12, 11223–11268, 2015

The left part of Fig. 2 shows the mean values of the objective function NS for the 10 000 parameter vectors for the calibration period 1971–1980 for the three selected models (denoted as individual calibration). As expected, the models perform differently in different catchments. The reasons for this are observation errors both in input and output as well as a possible inability of the model to reasonably well represent the main hydrological processes.

The ranges of the model parameters are relatively large. As a first step, we checked if the catchments have common parameter vectors. For each pair of catchments $(i, j)$, for the same performance measure and time period, the intersection of the convex hull of the good parameter sets $\mathcal{G}_i \cap \mathcal{G}_j$ is empty showing that there are no common best parameters. From the result, seemingly none of the catchments are similar.

As a next step, the 10 000 generated best dynamical parameter vectors for a given time period and hydrological model obtained for catchment $i$ were applied to model all other catchments using the same hydrological model and time period. Note that the value of $\eta$ is not transferred but adjusted to the true long term water balance. Figure 3 show the color coded matrices for the mean NS performance and GK performance of the three hydrological models using transferred parameters for all 15 catchments for a calibration period (1971–1980).

The performance of the transferred parameter vectors displays a strongly varying picture. While in some cases the catchments seem to share parameter vectors with reasonably good performance, in other cases the transfer lead to weak performances. A further surprising fact is that none of the matrices is symmetrical. One can see that some catchments are good donors as their parameters are good for nearly all catchments, while others have parameters which are hardly transferable.

The asymmetry of the parameter transition matrices cannot be explained by catchment properties. Two different catchments seem to share well performing parameters if calibrated on one catchment and no common good parameters if calibrated on the other one. Take the catchments 1 and 12 with the NS performance as an example. For all three models, parameters calibrated for catchment 1 are not





HESSD
12, 11223–11268, 2015

suitable for catchment 12, but parameters of catchment 12 perform reasonably well for catchment 1. The matrices for NS show different performances with different models. In general, HBV model performs the best. The average value of the matrix is 0.62 for HBV, 0.55 for HYMOD and 0.54 for XAJ model. Furthermore, the correlation of transferred model performance between different models are all greater than 0.7. From the viewpoint of parameter transferability, the three models perform similarly, if a parameter transfer is reasonable from catchment $i$ to $j$ for one model then it is also reasonable for the other models. The results for the GK performance differ from those of the NS performance. Here the XAJ model seems to give the generally best transferable parameters. Model parameters from other catchments are almost useless for catchment 15 for all three models.

The difference of the transferability for these two performance measures could be explained by different focuses – while NS is mainly focusing on the squared difference between the observed and modeled discharge, GK focuses on water balances and good timing and NS+LNS is strongly influenced by low flow events. It is interesting to observe that catchment 12 is a very bad receiver for model parameters for NS, while it is an excellent receiver for GK. This means that different events have different influence on the performance. A possible explanation for the asymmetry is the fact that the catchments have different weather forcing in the calibration period. It could be that runoff events which are most important for a performance measure occur in the calibration period frequently in one catchment leading to good transferability, and seldom in the other causing weak transferability of the parameters from one catchment to another.

The transferability of the model parameters was also tested for an independent validation period between 1991 and 2000. Figure 4 shows the corresponding color coded results for NS as performance measure. The matrices are similar to those obtained for calibration. Catchment 12 remained a bad receiver but a good donor indicating that the bad performance is unlikely to be caused by observation errors.





HESSD
12, 11223–11268, 2015

Further, for some columns the off diagonal elements are larger than the diagonal ones which is a sign of a possible overcalibration of models.

To investigate the influence of weather on calibration, the hydrological models calibrated for different time periods using the same model and performance measure were compared. As the different time periods represent different weather conditions, the calibrations lead to different parameter sets. As a comparison, the differences in calibrated model parameters using the same model and performance measure for different catchments were compared. As an example, Fig. 5 shows two calibrated parameters of the HYMOD model for catchment 13 on three different 10 years time periods. Figure 6 shows the same parameters obtained by calibration for three different catchments 7, 8 and 13 during time period 1951–1960. The structural similarity of the two figures suggests that the difference between the different catchments is comparable to the difference between the different time periods. In hydrological modeling, it is usually assumed that model parameters are constant over time assuming no significant change in climate or other characteristics. The results however show the assumption that parameters are the same over space is not completely unrealistic. The figures even suggest that there might be parameter vectors which perform reasonably well for all 15 catchments. As a next step, an experiment to test this assumption was devised.

## 6 Numerical experiment 2: simultaneous calibration

Since for many pairs of catchments, the parameter transfer worked reasonably well. As a next step, we investigated if there are parameters which perform reasonably well for all catchments. As seen in the previous section, none of the catchments share optimal parameters. Therefore common sub-optimal parameters have to be found.

In order to identify parameter vectors which perform simultaneously well for each catchment, the hydrological models were calibrated for all 15 catchments simultaneously. The simultaneous calibration of the model for all catchments is a multi-





HESSD
12, 11223–11268, 2015

objective optimization problem. The goal is to find parameter vectors which are almost equally good for all catchments with no exception. As the models perform differently for the different catchments due to data quality and catchment particularities, the performance was measured through the loss in performance compared to the usual individual calibration. Thus the objective function was formulated using the formulation of the compromise programming method (Zeleny, 1981):

$$R^{(j)}(\theta) = \sum_{i=1}^{n} \left( O_i^{(j)*} - O_i^{(j)}(\theta) \right)^p \tag{8}$$

Here index $j$ indicates the type of the individual performance measure specified in Eqs. (1), (2) and (4). The goal in this objective function is to minimize $R^{(j)}$. Here $p$ is the so called balancing factor, the larger $p$ is the more the biggest loss in performance contributes to the common performance. In order to obtain parameters which are good for all catchments, a relatively high $p = 4$ was selected for all three performance measures.

As same as individual calibration, the ROPE algorithm was used for the simultaneous calibration. The optimized parameter sets $\mathcal{H}^{(j)}$ are simultaneously well performed for each model and time period. The left part of Fig. 2 compares the performance of the individually calibrated and the common calibration for the 15 selected catchments using NS as performance criterion. As expected, the results show that the individual calibrations lead to better performances, but the joint parameter vectors perform reasonably well for all catchments.

As the goal of modeling is not the reconstruction of already observed data, the performances on a different validation period (1991–2000) were also compared. The right part of Fig. 2 shows the mean model performances for the 15 individually calibrated and the common calibrated datasets. Result shows the use of the parameters obtained from the common calibration for each catchment are sometimes even better than those obtained by using the individually calibrated parameters.

These results indicate that instead of transferring model parameters from a single catchment, a parameter transfer might perform better if the parameters obtained through common calibration on all other catchments are used. In order to test this kind of parameter transfer, a set of simple "leave one out" calibrations were performed. This means that for a catchment *i*, the hydrological models were simultaneously calibrated the remaining 14 catchments. Each time another catchment *i* was not considered for calibration, leading to 15 simultaneous calibrations. These common model parameters were then applied for the catchment which was left out. The performance of the models on these catchments in the calibration period is reasonably good for all catchments. Figure 7 shows the result of HBV and HYMOD using the NS performance measure. It compares the performance of the parameters obtained via individual calibrations (red x-mark), parameter transfers from other catchments individually (blue plus) and the transfer of the common parameters obtained by leave one out procedure (green diamond). The performance of common parameters is obviously weaker than that of the individual calibration, but better than many parameter transfer obtained using individual parameter transfer. To test the effective potential of the transferability of the common parameters, a validation period was used. Figure 8 shows the results for the validation time period 1991–2000. In this case, the common calibration performs very well. For HYMOD, it outperforms the model obtained by individual calibration for 6 out of the 15 catchments. For the other catchments, the loss in performance is relatively small. Note that this good performance of the common models was obtained without using any information of the target catchment. The transfer of parameters obtained from individual calibrations on other catchments shows a highly inhomogeneous picture as described in experiment 1. The transferred common calibration is better than most of these performances. Further note that the results of experiment 1 show that there is no explanation why certain transfer work well and others do not. Thus for the transfer of model parameters to ungauged catchments, common calibration seems to be a reasonable method.





HESSD
12, 11223–11268, 2015

In order to illustrate how model parameters of the leave one out common calibration perform in validation, two hydrographs are presented. Figures 9 and 10 show a part of the observed, the modeled and the common calibration transferred hydrographs for a randomly selected parameter set obtained by individual calibration and leave one out common calibration of HBV for catchments 5 and 14. While for catchment 5, the common calibration leads to a hydrograph which is slightly better than that obtained by individual calibration, in the second case for catchment 14 the performance is reversed. However, in both cases the common parameters, which were obtained without using any observations of the catchment perform surprisingly well.

## 7 Numerical experiment 3: extension to other catchments

The results of the previous experiment suggest that even more catchments might share parameters which perform well on all. The 15 catchments used in experiments 1 and 2 are however to some extent similar and can thus not necessarily be considered as representative for a great number of other catchments. Thus, for the third experiment, 192 catchments of the MOPEX dataset were considered. 96 of them were randomly selected for common calibration (marked as blue circle on Fig. 1), the other 96 catchments were used as receivers to test the performance of the common parameters (marked as green triangle on Fig. 1). HBV model using three selected performance measures were considered in this experiment.

For each of the 192 catchments, an individual model calibration was carried out using 1971–1980 as calibration period. Common calibration was performed for the selected 96 catchments the same way as in experiment 2, for HBV model using all performance measures.

As a first step, the model performances for the individual and common calibration were compared. As expected and already seen in experiment 2, the performance for the common calibration is lower than the individual one for HBV using all performance measures. For example, the mean performance NS over all 96 catchments drops from 0.69 to 0.50. When one applies the models for the validation period 1991–2000, the individually calibrated model mean performance is 0.65, while for the common calibration the mean increases to 0.51. Figure 11 shows the histograms of the performance NS for the calibration and validation periods for the individual and the common calibrations. Results indicate the robustness of the common calibration. The transfer to the 96 assumed ungauged catchments shows very similar performance for the common parameters as for the catchments selected for common calibration. Figure 12 shows the histograms of the performance NS for the individual calibration and the transfer for the assumed ungauged catchments. It can be seen clearly from the histogram that there is very little difference between the performance for the gauged and the ungauged catchments. In 90 % of catchments, the common calibration works reasonably well even for the ungauged cases. The common parameters describing runoff dynamics of all 192 catchments indicate that there is a high degree of similarity of these catchments.

Comparing the results of the common calibration using the 96 catchments to that obtained using the 15 catchments, one can observe that the increase of catchments considered for the common calibration lead to a decrease of the performance. This is as expected that there is less common behavior of a large set of catchments as for a few. Thus the parameters obtained through common calibration can be regarded to describe the common dynamical behavior of many very different catchments over a large geographical area. If one is interested to find model parameters for a specific ungauged catchment, the common calibration using a more careful selection of the donor set of catchments is likely to lead to good parameter transfers.

The water balances of the 192 catchments are very different leading to very different $\eta$ parameters. Figure 13 shows the distribution of $\eta$ values for three randomly selected common good parameter sets for HBV model using NS as performance measure for the calibration time period. It can be seen clearly from the curve that for the same catchment, $\eta$ is specific for different dynamical parameter sets. And due to the differences in water balance, different catchments requires very different $\eta$-s to control actual evapotranspiration. Furthermore, for all 192 catchments, parameter $\eta$ present very similar tendency for different dynamical parameter sets. Figure 14 plots the mean $\eta$ value against the ratio of the long term actual evapotranspiration to potential evapotranspiration ($E_{ta}/E_{tp}$) for each catchment. It shows strong negative correlation (−0.72) between $\eta$ and $E_{ta}/E_{tp}$.

## 8 Numerical experiment 4: application to catchments in other geographical regions

The 96 catchments used for experiment 3 represent a large variety of hydrological and meteorological conditions. Their use for other catchments showed a quite good performance for all performance criteria. Thus the question whether theses parameters describe a general hydrological behavior independently of the location arises. In experiment 4, the common parameters obtained by common calibration in experiment 3 were used to model two German catchments with appropriate data availability.

The selected Rottweil catchment and Fils catchment at Süssein are located in Southwest of Germany with a drainage size of 455 and 345 $km^2$, respectively. The hydrometeorological data from 1971 to 1980 were used to test the transferability of common parameter sets calibrated for the 96 US catchments. The simulation performances for these two catchments are listed in Table 3. Result shows all dynamical model parameters obtained through simultaneous calibration for the 96 US catchments worked well for both catchments. But for the Rottweil catchment, model performance is worse than for the Fils catchment. It indicates that there is some skill in the transferred parameters, but the differences are substantial. Figures 15 and 16 show part of the observed and the modeled hydrographs using the NS performance measure. We can see the transfer is reasonable and the dynamics of the discharge are similar to the US case. This experiment demonstrated that even very distant and different catchments may behave similarly.





HESSD
12, 11223–11268, 2015

## 9 Discussion

### 9.1 Robust parameters

The four experiments were carried out in way that a set of parameters (usually represented by 10 000 individual parameter sets) was used. This leads to a considerable fluctuation of the results. Modelers often prefer to use single parameter vector. If a single parameter vector is desired, then according to Bárdossy and Singh (2008), the deepest parameter set is the most likely candidate to be robust. This study also indicates the deepest parameter set perform slightly better than the mean of the parameter sets considered.

### 9.2 Variability of $\eta$

As described in Sect. 4.1, the water balance related parameter $\eta$ is specific for each catchment and each model parameter vector. For each individual catchment, we can have a large variation in $\eta$. And for the same set of good parameters that matching different water balances, different catchments always require very different $\eta$-s to control actual evapotranspiration.

### 9.3 Estimation of $\eta$ for ungauged catchments

In the numerical experiments, in order to estimate water balance parameter $\eta$, the long term discharge volumes were treated as known variables for both gauged and ungauged catchments. Therefore for practical application, the long term discharge volumes have to be estimated for ungauged catchments. This problem is not explicitly treated in this paper. For the study area, the discharge coefficients which relate discharge volumes to (known) precipitation show a quite smooth spatial behavior as shown on Fig. 17. Thus the regionalization of this parameter seems to be a not extremely complicated task. According to the previous analysis of $\eta$, for each common dynamical parameter set, one can have a possible estimator of $\eta$ for a certain catchment based on the regionalization of discharge coefficients.

## 10 Conclusions

In this paper, the transfer of the dynamical parameters of hydrological models was investigated. In order to cope with the clear differences in water balances due to water or energy limitations, a new model parameter $\eta$ controlling the actual evapotranspiration was introduced. This parameter is determined for each vector of other model parameters by adjusting the long term water balances. This parameter was not transferred, only the other parameters controlling flow dynamics and short term water balances were assumed to be shared by many catchments. In order to assure the generality of the results, three different simple lumped hydrological models were used in combination with three different performance measures in four numerical experiments on a large number of catchments. The following conclusions can be drawn:

- Hydrological models are often overfitted during calibration. The parameters are sometimes more specific for the calibration time period and their relation to catchment properties seems to be unclear. This makes parameter transfers or parameter regionalization based on individual calibration difficult.

- In the second experiment, a common calibration strategy was introduced and tested on a small number (15) of catchments. For the common calibration, an overall objective function which considers all catchments simultaneously and allows little compensation is required. Compromise programming offers a good possibility for this purpose. This methodology was able to identify parameter sets which work reasonably for all catchments. Testing the parameters on an independent time period shows that common parameters perform comparably well as those obtained using individual calibration. The transfer of the common parameters to model ungauged catchments works well. Note that





HESSD
12, 11223–11268, 2015

the water balance parameters have to be estimated individually for the ungauged catchments.

- Extending the number of catchments covering a very large spatial scale (continental) still allows a reasonable common calibration and transfer of the dynamical parameters. The performance on this scale is weaker than on the smaller scales, but a transfer to other ungauged catchments is still possible.

- Parameters obtained via continental scale calibration are transferable to model catchments on other continents. This shows that there is a partly common behavior of most catchments. However note that the performance of these common parameters is significantly worse than what can be obtained using individual calibration.

- The fact that many catchments share common parameters which describe their dynamical behavior does not mean that they have the same dynamical behavior. The model output highly depends on the parameter $\eta$ which varies from catchment to catchment and also as a function of the other model parameters describing dynamical behavior.

- The results of the experiments were similar for all three hydrological models applied independently of the choice of the performance measures. Note however that the common parameters corresponding to the different performance measures differ considerably. Common behavior is dependent how one evaluates the performance of the models.

- The performance of the common parameters depends strongly on the selection of the catchments used to assess them. The optimal choice of the appropriate catchments was not investigated in the framework of this research. The second experiment suggests that a reasonable geographic proximity of the catchments might be a good choice for common calibration.





HESSD
12, 11223–11268, 2015

Common parameters offer a good possibility for the prediction of ungauged catchments. One single parameter $\eta$ which controls the long term water balances has to be estimated individually. This however can be done using other modelling approaches including regionalization methods.

The results show that on the daily time scale many catchments behave similarly and the same dynamical parameter sets can be used for all of them independently from the selected lumped model and the performance function selected. This means that hydrological behavior on the daily scale is dominated by precipitation characteristics and actual evapotranspiration. Differences in catchment properties play a secondary role and cannot be captured well by parameters of simple lumped models within our study area.

*Acknowledgements.* The research of the second author (Yingchun Huang) was supported by China Scholarship Council.

## References

Andréassian, V., Le Moine, N., Perrin, C., Ramos, M.-H., Oudin, L., Mathevet, T., Lerat, J., and Berthet, L.: All that glitters is not gold: the case of calibrating hydrological models, Hydrol. Process., 26, 2206–2210, 2012. 11226

Archfield, S. A. and Vogel, R. M.: Map correlation method: Selection of a reference streamgage to estimate daily streamflow at ungaged catchments, Water Resour. Res., 46, W10513, doi:10.1029/2009WR008481, 2010. 11226

Bárdossy, A. and Singh, S. K.: Robust estimation of hydrological model parameters, Hydrol. Earth Syst. Sci., 12, 1273–1283, doi:10.5194/hess-12-1273-2008, 2008. 11234, 11243

Bergström, S. and Forsman, A.: Development of a conceptual deterministic rainfall-runoff model, Nord. Hydrol., 4, 174–190, 1973. 11230

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. 11225





HESSD
12, 11223–11268, 2015

Beven, K. and Freer, J.: Equifinality, data assimilation, and data uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, J. Hydrol., 249, 11–29, 2001. 11225

Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., and Savenije, H. e.: Runoff Prediction in Ungauged Basins: Synthesis across Processes, Places and Scales, Cambridge University Press, Cambridge, 2013. 11225

Boyle, D. P., Gupta, H. V., Sorooshian, S., Koren, V., Zhang, Z., and Smith, M.: Toward improved streamflow forecasts: value of semidistributed modeling, Water Resour. Res., 37, 2749–2759, 2001. 11229

Duan, Q., Schaake, J., Andreassian, V., Franks, S., Goteti, G., Gupta, H., Gusev, Y., Habets, F., Hall, A., Hay, L., Hogue, T., Huang, M., Leavesley, G., Liang, X., Nasonova, O., Noilhan, J., Oudin, L., Sorooshian, S., Wagener, T., and Wood, E.: Model Parameter Estimation Experiment (MOPEX): an overview of science strategy and major results from the second and third workshops, J. Hydrol., 320, 3–17, 2006. 11229

Falcone, J. A., Carlisle, D. M., Wolock, D. M., and Meador, M. R.: GAGES: a stream gage database for evaluating natural and altered flow conditions in the conterminous United States: Ecological Archives E091-045, Ecology, 91, 621–621, 2010. 11234

Grigg, D.: The logic of regional systems 1, Ann. Assoc. Am. Geogr., 55, 465–491, 1965. 11226

Gupta, H., Kling, H., Yilmaz, K., and Martinez, G.: Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling, J. Hydrol., 377, 80–91, 2009. 11231

McDonnell, J. and Woods, R.: On the need for catchment classification, J. Hydrol., 299, 2–3, 2004. 11226

McIntyre, N., Lee, H., Wheater, H., Young, A., and Wagener, T.: Ensemble predictions of runoff in ungauged catchments, Water Resour. Res., 41, W12434, doi:10.1029/2005WR004289, 2005. 11226

Moore, R. J.: The probability-distributed principle and runoff production at point and basin scales, Hydrol. Sci. J., 30, 273–297, 1985. 11229

Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models. 1. A discussion of principles, J. Hydrol., 10, 282–290, 1970. 11231

Oudin, L., Andréassian, V., Perrin, C., Michel, C., and Le Moine, N.: Spatial proximity, physical similarity, regression and ungaged catchments: a comparison of regionalization approaches based on 913 French catchments, Water Resour. Res., 44, W03413, doi:10.1029/2007WR006240, 2008. 11225

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. 11226

Samaniego, L., Bárdossy, A., and Kumar, R.: Streamflow prediction in ungauged catchments using copula-based dissimilarity measures, Water Resour. Res., 46, W02506, doi:10.1029/2010WR009290, 2010. 11226

Sawicz, K., Wagener, T., Sivapalan, M., Troch, P. A., and Carrillo, G.: Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA, Hydrol. Earth Syst. Sci., 15, 2895–2911, doi:10.5194/hess-15-2895-2011, 2011. 11226

Schaefli, B. and Gupta, H.: Do Nash values have value?, Hydrol. Process., 21, 2075–2080, 2007. 11231

Seibert, J.: Regionalisation of parameters for a conceptual rainfall-runoff model, Agric. Forest Meteorol., 98, 279–293, 1999. 11225

Sivapalan, M.: Prediction in ungauged basins: a grand challenge for theoretical hydrology, Hydrol. Process., 17, 3163–3170, 2003. 11225

Wagener, T., Boyle, D. P., Lees, M. J., Wheater, H. S., Gupta, H. V., and Sorooshian, S.: A framework for development and application of hydrological models, Hydrol. Earth Syst. Sci., 5, 13–26, doi:10.5194/hess-5-13-2001, 2001. 11229

Wagener, T., Sivapalan, M., Troch, P., and Woods, R.: Catchment classification and hydrologic similarity, Geogr. Compass, 1, 901–931, 2007. 11226

Zeleny, M.: Multiple Criteria Decision Making, McGraw-Hill, New York, USA, 1981. 11238

Zhao, R. J. and Liu, X.: The Xinanjiang model, in: Computer Models of Watershed Hydrology, Water Resources Publications, Littleton, Colorado, USA, 215–232, 1995. 11230





HESSD

12, 11223–11268, 2015

**Table 1.** Catchment properties for the selected 15 catchments

| Streamgauge ID | Streamgauge name | Drainage area ($km^2$) | Shape factor | Field capacity | Percentage of relative humidity | Average porosity | Base flow index | Snow proportion (%) |
|---|---|---|---|---|---|---|---|---|
| 01548500 | Pine Creek at Cedar Run, PA | 1564 | 0.14 | 0.32 | 67.1 | 0.42 | 0.44 | 26.6 |
| 01606500 | So. Branch Potomac River near Petersburg, WA | 1663 | 0.15 | 0.31 | 74.9 | 0.28 | 0.45 | 19.5 |
| 01611500 | Cacapon River near Great Cacapon, WV | 1753 | 0.17 | 0.269 | 68.4 | 0.27 | 0.41 | 15.6 |
| 01663500 | Hazel River at Rixeyville at Rixeyville, VA | 743 | 0.16 | 0.30 | 66.9 | 0.39 | 0.51 | 12.1 |
| 01664000 | Pappahannock River at Remington, VA | 1606 | 0.11 | 0.294 | 67.1 | 0.40 | 0.50 | 11.8 |
| 01667500 | Rapidan River near Culpeper, VA | 1222 | 0.13 | 0.32 | 67.9 | 0.40 | 0.51 | 10.6 |
| 02016000 | Cowpasture River near Clifton Forge, VA | 1194 | 0.18 | 0.28 | 69.8 | 0.27 | 0.43 | 16.0 |
| 02018000 | Craig Creek at Parr, VA | 852 | 0.24 | 0.27 | 67.5 | 0.30 | 0.44 | 11.3 |
| 02030500 | Slate River near Arvonia, VA | 585 | 0.20 | 0.30 | 66.4 | 0.46 | 0.48 | 8.5 |
| 03114500 | Middle Island Creek at Little, WV | 1186 | 0.14 | 0.36 | 69.4 | 0.27 | 0.21 | 15.6 |
| 03155500 | Hughes River at Cisco, WV | 1171 | 0.14 | 0.36 | 68.7 | 0.27 | 0.22 | 14.9 |
| 03164000 | New River near Galax, VA | 2929 | 0.09 | 0.29 | 71.5 | 0.43 | 0.64 | 13.3 |
| 03173000 | Walker Creek at Bane, VA | 790 | 0.24 | 0.32 | 71.9 | 0.37 | 0.46 | 13.5 |
| 03180500 | Greenbrier River at Durbin, WV | 344 | 0.26 | 0.36 | 77.4 | 0.27 | 0.37 | 25.3 |
| 03186500 | Williams River at Dyer, WV | 332 | 0.33 | 0.36 | 73.8 | 0.28 | 0.36 | 24.3 |

**Table 2.** Climate variables of the 15 selected catchments.

| No | Streamgauge ID | Annual precipitation (mm) | Average temperature (°C) | Annual potential evapotranspiration (mm) | Annual runoff (mm) |
|---|---|---|---|---|---|
| 1 | 01548500 | 951.7 | 7.2 | 727.0 | 495.1 |
| 2 | 01606500 | 948.6 | 10.3 | 716.3 | 378.3 |
| 3 | 01611500 | 905.6 | 10.8 | 800.0 | 310.5 |
| 4 | 01663500 | 1049.9 | 11.7 | 897.2 | 402.6 |
| 5 | 01664000 | 1027.7 | 12.0 | 906.1 | 367.5 |
| 6 | 01667500 | 1087.4 | 12.3 | 915.2 | 380.4 |
| 7 | 02016000 | 1029.5 | 11.0 | 746.0 | 402.9 |
| 8 | 02018000 | 1010.6 | 11.4 | 764.6 | 406.3 |
| 9 | 02030500 | 1075.9 | 13.5 | 918.2 | 350.3 |
| 10 | 03114500 | 1089.7 | 11.4 | 737.4 | 483.9 |
| 11 | 03155500 | 1057.8 | 11.6 | 740.0 | 443.7 |
| 12 | 03164000 | 1247.9 | 10.6 | 807.4 | 593.3 |
| 13 | 03173000 | 958.6 | 11.1 | 762.7 | 371.9 |
| 14 | 03180500 | 1224.2 | 8.3 | 710.9 | 543.2 |
| 15 | 03186500 | 1401.5 | 9.1 | 710.9 | 945.0 |





HESSD
12, 11223–11268, 2015

**Table 3.** HBV model performance of the Germany catchments.

| Catchment | NS | GK | NS + LNS |
|---|---|---|---|
| Rottweil | 0.47 | 0.90 | 0.52 |
| Fils | 0.58 | 0.97 | 0.68 |





HESSD
12, 11223–11268, 2015

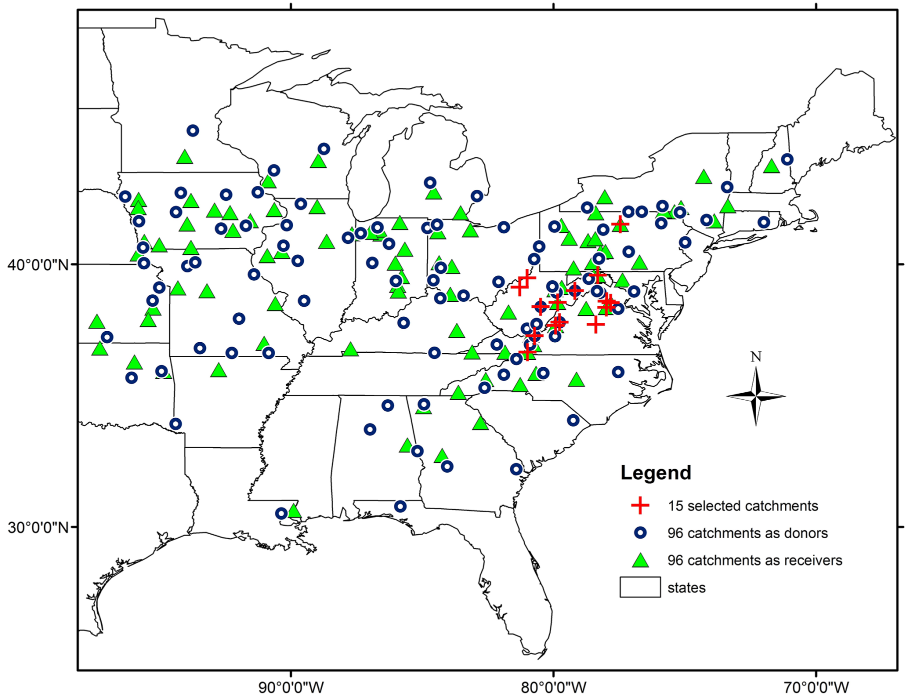


**Figure 1.** Location of the catchments selected for the experiments.

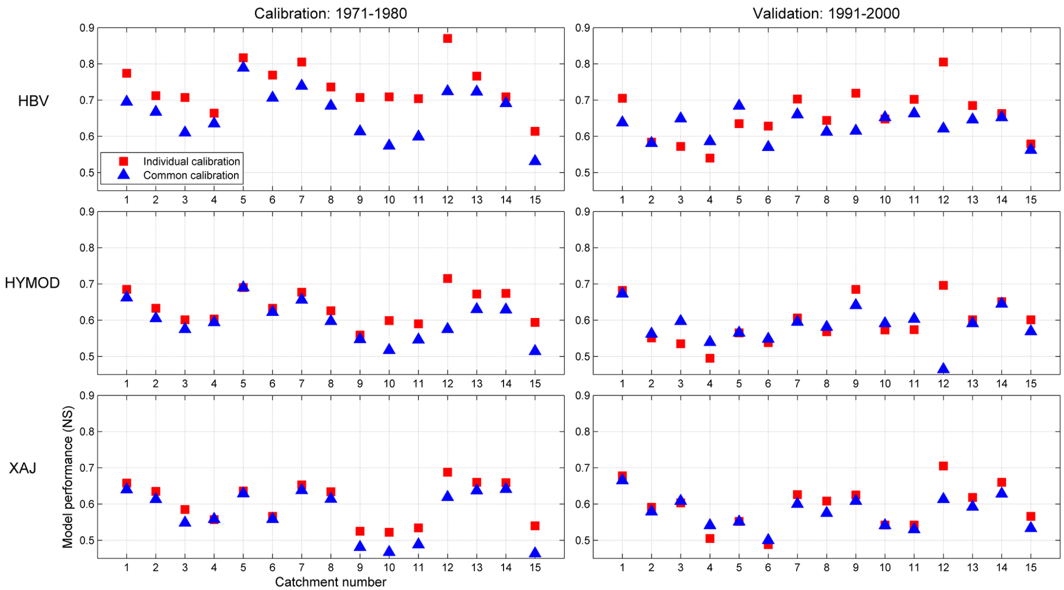


**Figure 2.** Performance of the individually calibrated and the common calibrated models using NS as performance criterion.

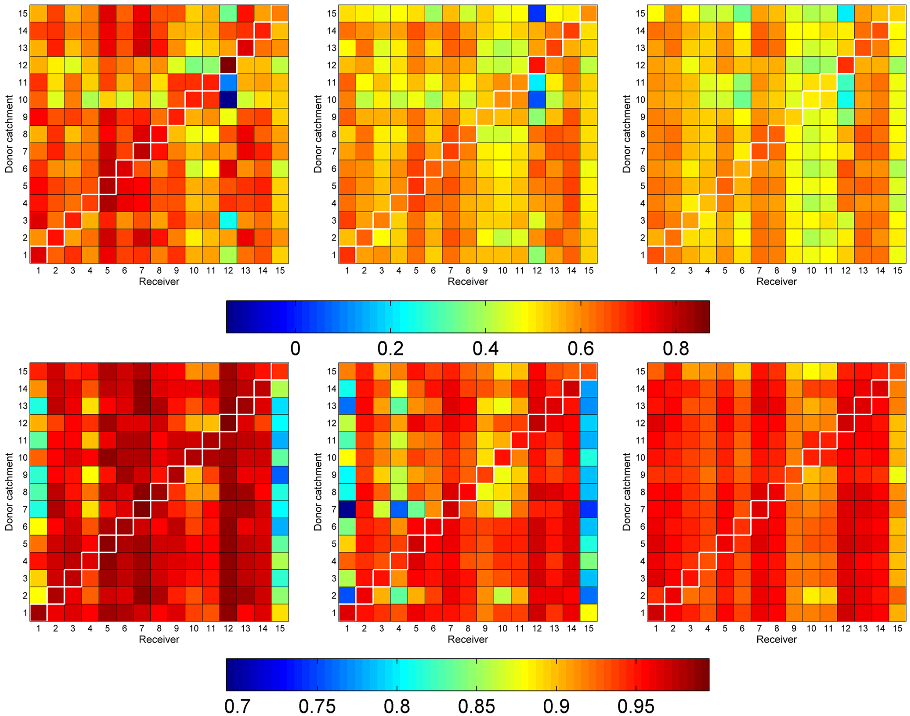


**Figure 3.** Color coded matrices for the mean model performance of the parameter transfer for the selected 15 catchments. The upper panel used NS as performance measure, the lower panel used GK as performance measure.





HESSD
12, 11223–11268, 2015

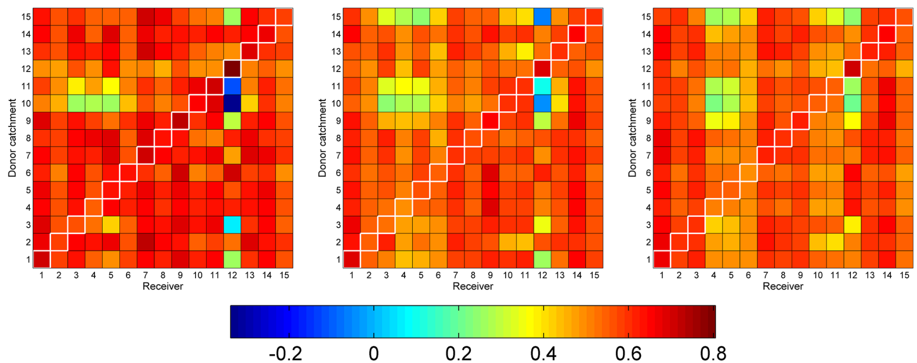


**Figure 4.** Color coded matrices for the mean NS model performance of the parameter transfer for the validation period for the selected 15 catchments.

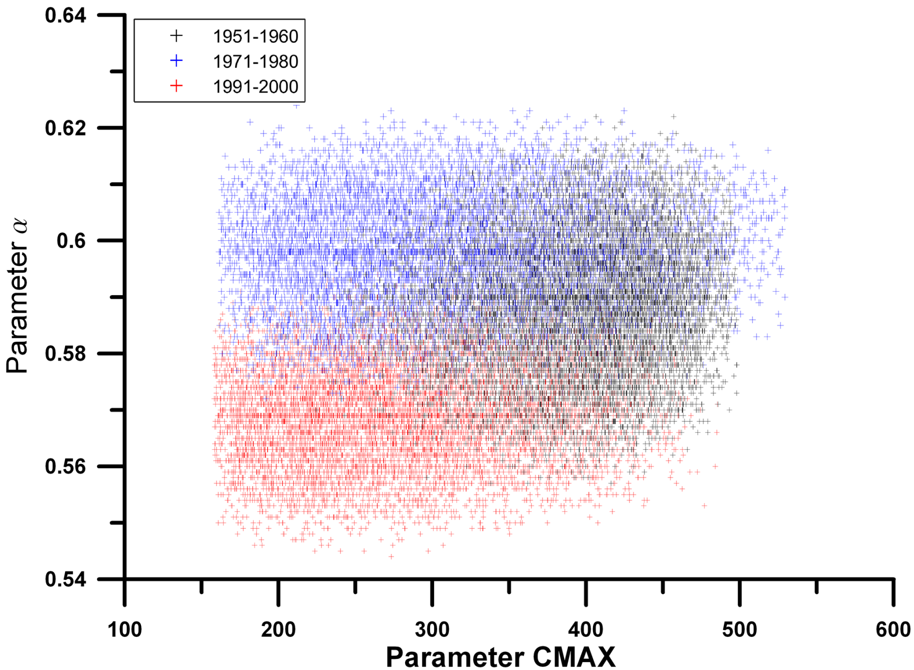


**Figure 5.** Scatterplots for two selected HYMOD parameters CMAX and $\alpha$ obtained via model calibration using NS as performance measures for catchment 13 (black: 1951–1960, blue: 1971–1980 and red: 1991–2000).





HESSD
12, 11223–11268, 2015



Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

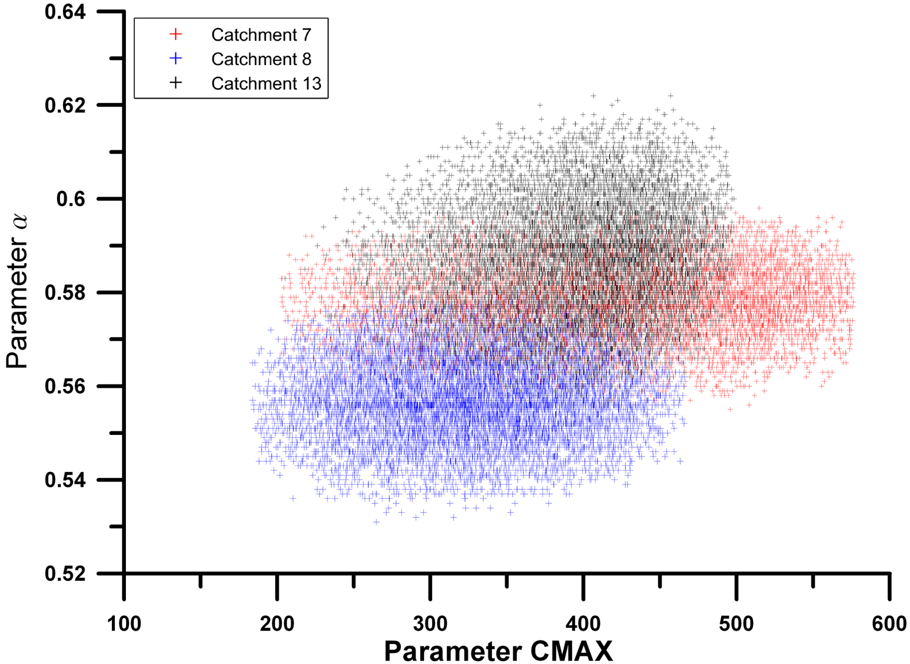


**Figure 6.** Scatterplots for two selected HYMOD parameters CMAX and $\alpha$ obtained via model calibration using NS as performance measures for catchments 7 (red), 8 (blue) and 13 (black) for 1951–1960.





HESSD
12, 11223–11268, 2015

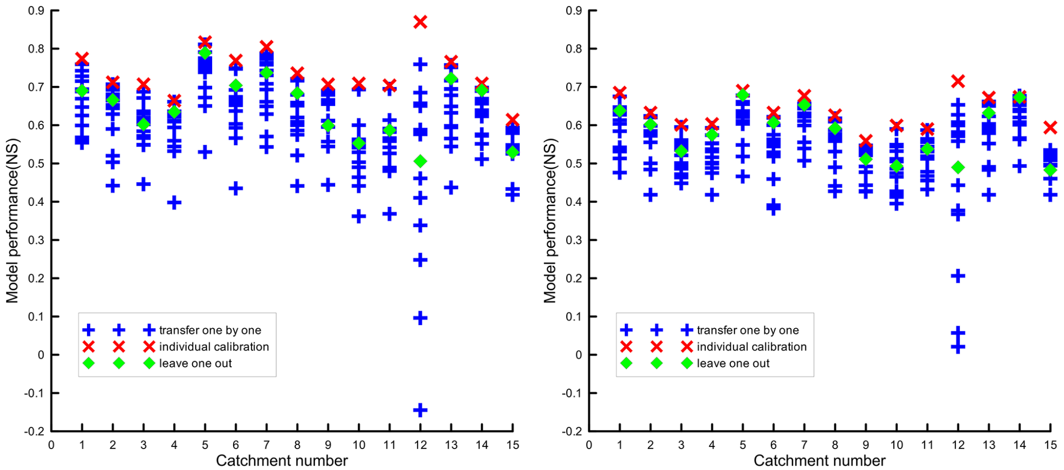


**Figure 7.** Mean NS model performance of the calibration, individual parameter transfer and for the leave one out transfer for the selected 15 catchments for the calibration time period 1971–1980. Left panel: HBV, right panel: HYMOD.

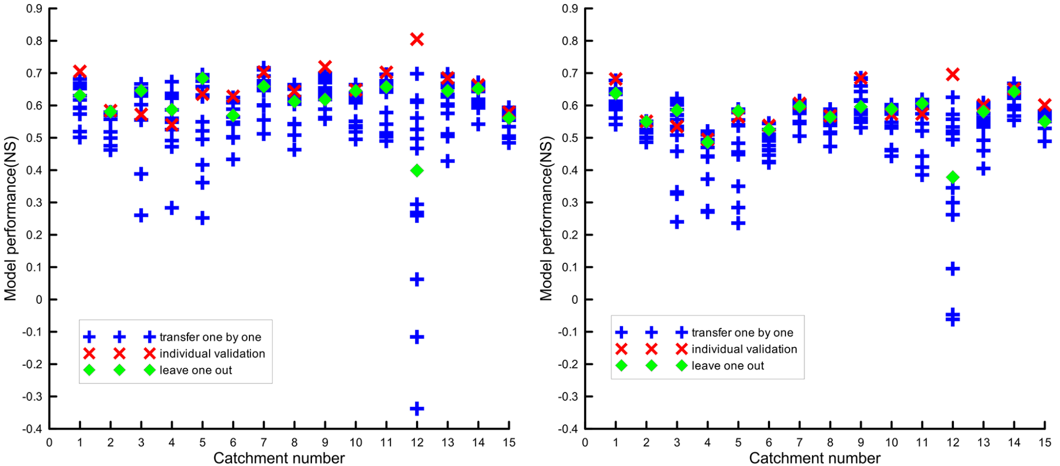


**Figure 8.** Mean NS model performance of the calibration, individual parameter transfer and for the leave one out transfer for the selected 15 catchments for the validation time period 1991–2000. Left panel: HBV, right panel: HYMOD.

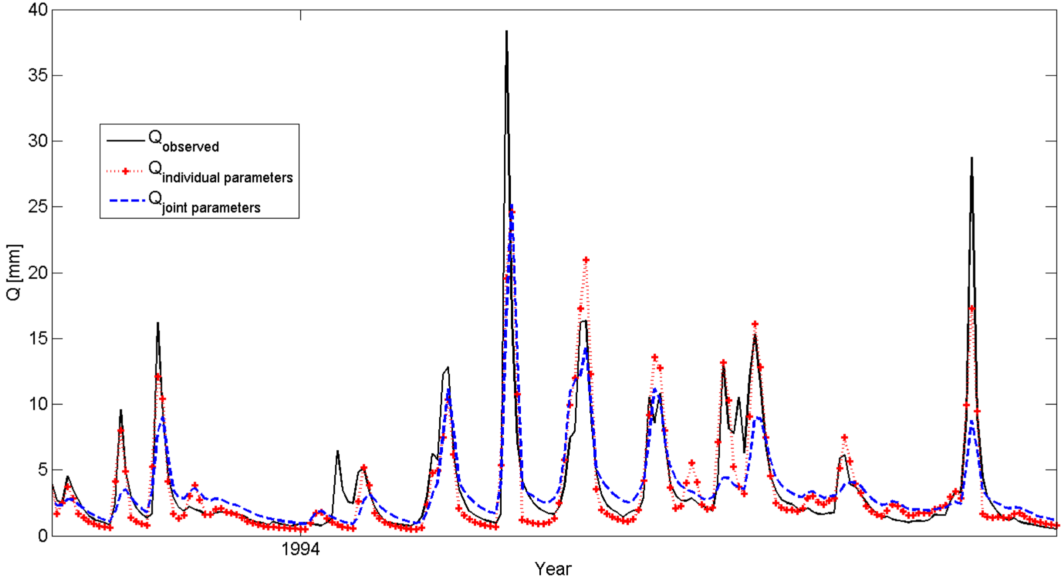


**Figure 9.** Runoff hydrographs for catchment 14 obtained using individual and leave one out common calibrations of HBV using the GK performance measure.





HESSD
12, 11223–11268, 2015

Simultaneous calibration

A. Bárdossy et al.

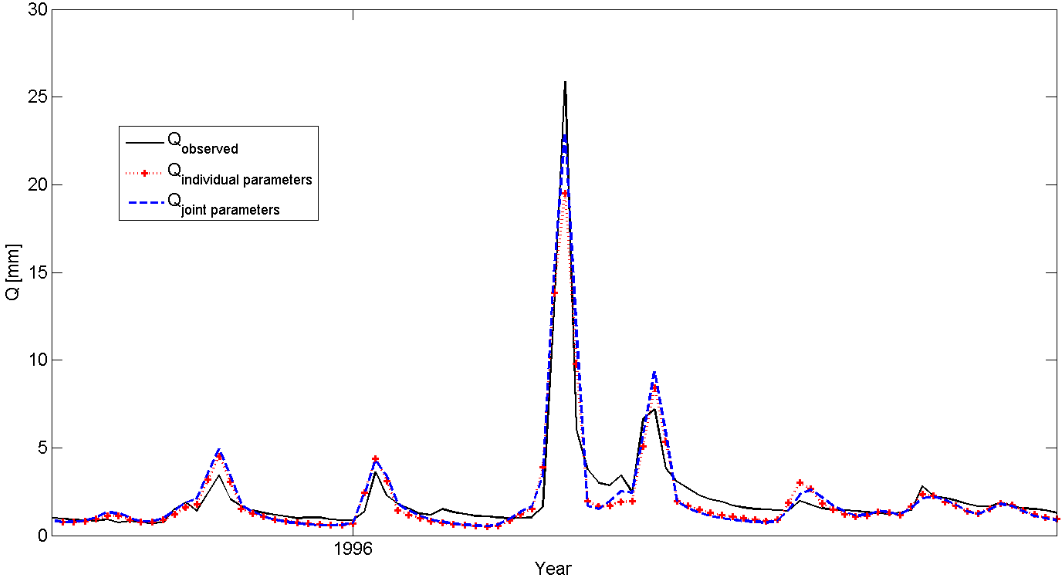


**Figure 10.** Runoff hydrographs for catchment 5 obtained using individual and leave one out common calibrations of HBV using the NS performance measure.

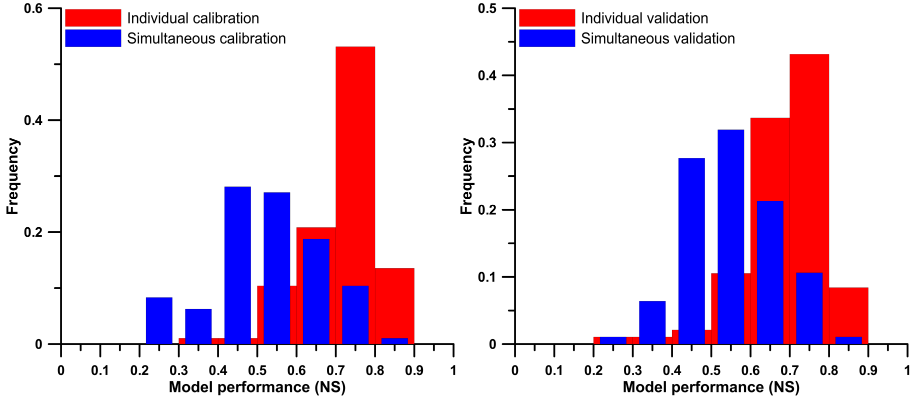


**Figure 11.** Histograms of the NS model performance of HBV for the 96 selected (donor) catchments. Left: calibration period (1971–1980), right: validation period (1991–2000).





HESSD
12, 11223–11268, 2015

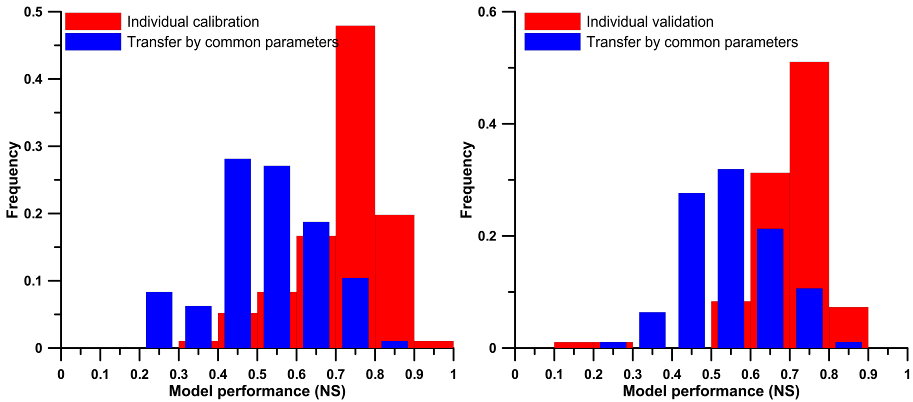


**Figure 12.** Histograms of the NS model performance of HBV for the 96 test (ungauged) catchments. Left: calibration period (1971–1980), right: validation period (1991–2000).





HESSD
12, 11223–11268, 2015

Simultaneous calibration

A. Bárdossy et al.

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

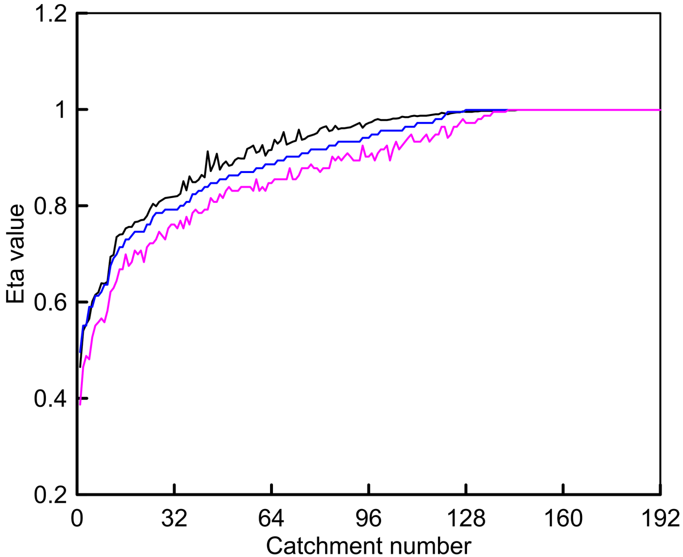


**Figure 13.** Distribution of water balance parameter $\eta$ for three randomly selected common parameter vectors obtained via HBV using the NS performance measure for 192 selected catchments.





HESSD
12, 11223–11268, 2015

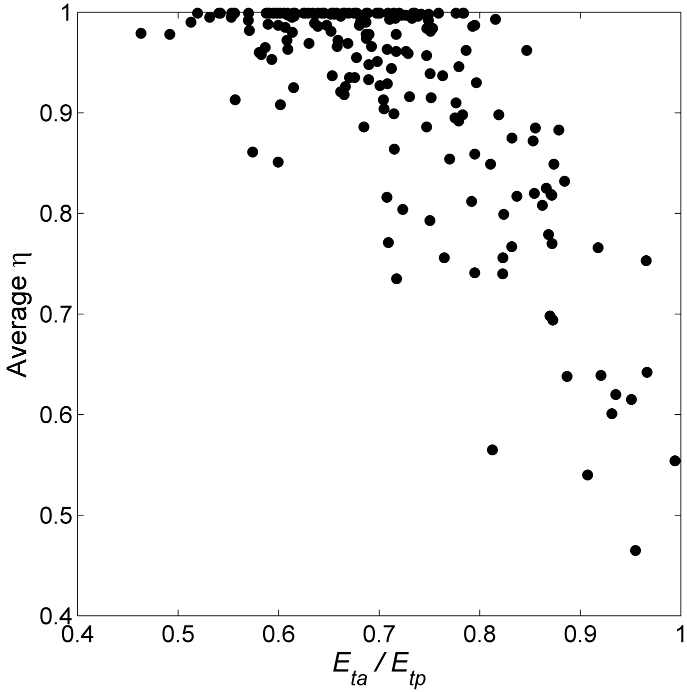


**Figure 14.** Scatterplots of mean $\eta$ value and ratio of actual evapotranspiration to potential evapotranspiration for 192 selected catchments.

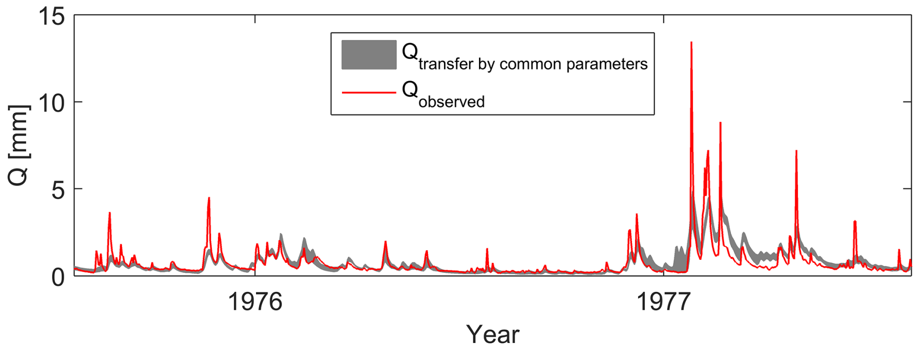


**Figure 15.** Observed and modeled discharges of the Rottweil catchment. Modelling was performed using the common parameters of the 96 US catchments obtained by calibration using HBV for NS.

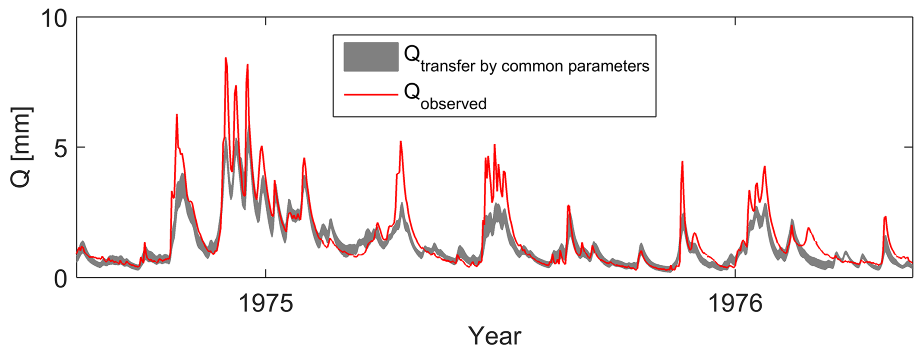


**Figure 16.** Observed and modeled discharges of the Fils at Süssen. Modelling was performed using the common parameters of the 96 US catchments obtained by calibration using HBV for NS.





HESSD
12, 11223–11268, 2015

Simultaneous calibration

A. Bárdossy et al.

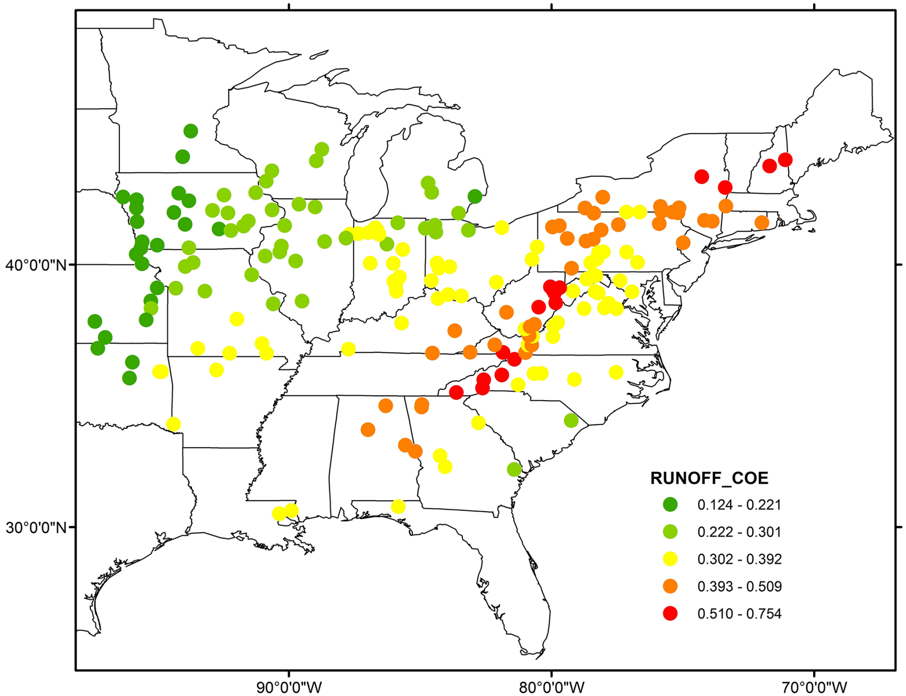


**Figure 17.** The discharge coefficient of the catchments selected for the experiments.





HESSD
12, 11223–11268, 2015