Hydrology and Earth System Sciences

Analysis of projected hydrological behavior of catchments based on signature indices

M. C. Casper[1], G. Grigoryan[1], O. Gronz[1], O. Gutjahr[2], G. Heinemann[2], R. Ley[1], and A. Rock[1]

[1]University of Trier, Physical Geography, 54286 Trier, Germany
[2]University of Trier, Environmental Meteorology, 54286 Trier, Germany

Correspondence to: M. C. Casper (casper@uni-trier.de)

Abstract. To precisely map the changes in hydrologic response of catchments (e.g. water balance, reactivity or extremes), we need sensitive and interpretable indicators. In this study we defined nine hydrologically meaningful signature indices: five indices were sampled on the flow duration curve, four indices were closely linked to the distribution of event runoff coefficients. We applied these signature indices to the output from a hydrologic catchment model for three different catchments located in the Nahe basin (Western Germany) to detect differences in runoff behavior resulting from different meteorological input data. The models were driven by measured and simulated (COSMO-CLM) meteorological data. It could be shown that the application of signature indices is a very sensitive tool to assess differences in simulated runoff behavior resulting from climatic data sets of different sources. Specifically, the selected signature indices allow assessing changes in water balance, vertical water distribution, reactivity, seasonality and runoff generation. These indices showed that the hydrological model is very sensitive to biases in mean and spatio-temporal distribution of precipitation and temperature because it acts as a filter for the meteorological input. Besides model calibration and model structural deficits, we found that bias correction of temperature fields and further adjustment of bias correction of precipitation fields is absolutely essential. We conclude that signature indices can act as indirect "efficiency measures" or "similarity measures" for output from regional or local climate models.

1 Introduction

The world is presently facing rapid changes to the climate. The understanding and prediction of related hydrologic changes is one main question that hydrologists face today (Blöschl and Montanari, 2010; Schaefli et al., 2011). It is therefore essential that we precisely map the changes in hydrologic response of catchments (e.g. water balance, reactivity or extremes). In this context, hydrological models are applied to detect the impact of a changing climate on the hydrology of large basins (Mahmoud et al., 2009). Often, the output of global climate models is fed into the hydrological impact model (Taye et al., 2011). Several studies highlight the effects of different spatial resolutions of meteorological forcing on hydrologic simulations (Arnell, 2011; Casper et al., 2009; Segond et al., 2007; Tramblay et al., 2011). Sangati et al. (2009) evaluated the effects of different spatial resolutions in meteorological forcing on modelling of flash floods. They concluded that in their study a correct estimate of rainfall volume is not sufficient for accurate reproduction of flash flood events characterised by large spatial rainfall variability. Increasing the aggregation length may result in a significant distortion of rainfall field geometry and a deformation of the hydrograph. Therefore, on smaller catchment scales, the output of regional climate models (RCM) and local climate models (LCM) is used as forcing data of hydrological models (Teutschbein and Seibert, 2010; Marke et al., 2011). However, Piani et al. (2010) conclude that precipitation and temperature fields simulated by regional climate models are defective and therefore introduce significant

errors into hydrological models when used as forcing data. Errors in the climate models affect the spatial and temporal distribution of rainfall and temperature (Sennikovs and Bethers, 2009). Wood et al. (2004) state that hydrological models are sensitive to biases in the basin mean and spatial distribution of precipitation and temperature. In this context, Benoit et al. (2000) and Fowler et al. (2007) strongly recommend the use of hydrological models for evaluation of downscaled climate instead of using only precipitation observations. However, this raises the question how to quantify differences in hydrological behavior of hydrological catchment models. Van Werkhoven et al. (2009) argue that statistical metrics only quantify the distance between measured and simulated runoff based on assumptions about the statistical characteristics of the model residuals. They do not indicate how well the hydrological function of the real system is maintained by the model. This makes it necessary to include hydrologically-based metrics in model evaluation. These metrics provide dynamic aspects of the watershed system or hydrological model (Yadav et al., 2007; Zhang et al., 2008) and may therefore allow for a quantitative evaluation of hydrological behavior (Gupta et al., 2008). Examples for hydrologically-based metrics are signature indices derived from flow duration curves (Yilmaz et al., 2008) or from distributions of runoff event coefficients (Merz et al., 2006; Ley et al., 2011).

The flow duration curve (FDC) allows indication and classification of watershed functioning. The FDC summarizes a catchment's ability to produce discharge values of different magnitudes, and is therefore strongly sensitive to the vertical distribution of soil moisture within a basin (Yilmaz et al., 2008). Additionally, 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).

The analysis of event runoff coefficients is another diagnostic tool. The main controls on event runoff coefficients are the climate and the runoff regime through the seasonal water balance. They are especially sensitive to seasonal changes in precipitation patterns (Merz et al., 2006). Blume et al. (2007) summarize that event runoff coefficients are useful to understand how different landscapes "filter" rainfall into event based runoff and to explain the observed differences between catchments. They offer information on watershed response including changes from event to event, or from season to season.

In order to combine the strengths of both approaches, we define in this study nine hydrologically meaningful signature indices: five indices are sampled on the FDC (similar to Yilmaz et al., 2008), four indices are closely linked to the probability distribution of event runoff coefficients (Ley et al., 2011). We apply this signature index concept to the output from a hydrologic catchment model run for 3 subcatchments located in the Nahe basin (Western Germany) to detect differences in runoff behavior resulting from
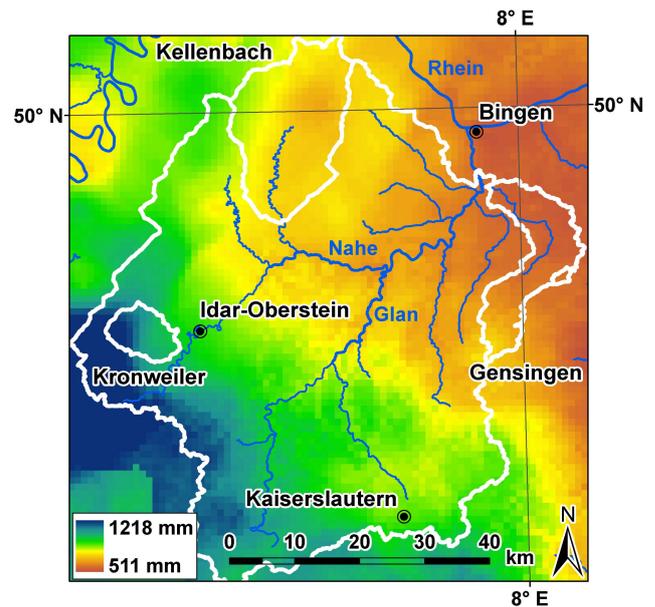
Fig. 1. Areal distribution of annual precipitation and the outlines of the catchments.

different meteorological data sets. The model is driven by measured and simulated (by a local climate model) meteorological data. We demonstrate the discriminating power of the selected signature indices by pairwise comparison of our simulation results.

2   Study area

The study area consists of three lower mesoscale gaged catchment areas in the low mountain ranges of the Nahe basin (Fig. 1), Germany: Kronweiler (64 km^2), Kellenbach (362 km^2) and Gensingen (197 km^2). Geology is characterized by Devonian schist, greywacke and quartzite in Kellenbach and most parts of Kronweiler. The south part of Kronweiler consists of Permian sedimentary and volcanic rocks. Tertiary clay and Pleistocene loess characterizes the geology of Gensingen. Mean annual precipitation reaches 930 mm in Kronweiler, followed by Kellenbach (675 mm) and Gensingen (545 mm). Mean annual potential evaporation reaches 615 mm in Gensingen and about 540 mm in Kronweiler and Kellenbach. Field capacity in Gensingen is much higher than in Kronweiler and Kellenbach. About 75 % of the area of Gensingen is used agriculturally, with 20 % vineyards and orchards. In Kellenbach and Kronweiler about half the area is forested. All watersheds are rural with little build-up area (all less than 6% of the area). The mean slope gradient of Kronweiler is 8.6°, much higher than for Kellenbach and Gensingen (about 4.5°). The runoff response behavior of the three catchments is quite different: Kronweiler shows high discharges, high reactivity and high runoff coefficients all year round. In contrast to this, Gensingen has low reactivity, low

Table 1. Catchment properties.

Catchment property		Kronweiler (64 km^2)	Kellenbach (362 km^2)	Gensingen (197 km^2)
Mean annual precipitation (mm yr^{-1})		930	675	545
Potential evaporation (mm yr^{-1})		535	540	615
Long term runoff rate (l s^{-1} km^2)	*year*	14.0	7.2	2.3
	winter	23.1	10.9	2.8
	summer	5.2	3.6	1.7
Mean runoff coefficient (93–08)	*year*	0.23	0.17	0.04
	winter	0.41	0.28	0.08
	summer	0.09	0.07	0.03
Land use		43 % arable land 3 % built-up area 54 % forest	58 % arable land 3 % built-up area 39 % forest	75 % arable land, orchards, vineyards 5 % built-up area 20 % forest
Soils (FAO85)		gleyic and humic podzols with cambisols	gleyic podzols with few cambisols	humic podzols and cambisols at upper reaches; luvisols, gleysols and regosols at lower reaches

discharges and low runoff coefficients with a high variability in winter. Runoff behavior of Kellenbach lies between the two other catchments (Table 1). Because of the large gradient in physical properties, we expect these three catchments to respond differently on changed climate input.

3 Methods

3.1 Hydrological model and input data

The water balance model LARSIM (Large Area Runoff Simulation Model) allows a continuous process- and area-detailed simulation of the mesoscale mainland water cycle (Ludwig and Bremicker, 2006). In simplistic terms, the watershed is subdivided into 1-D elements linked by a flood routing scheme. To account for sub-grid variability, interception, snow accumulation and melt, evapotranspiration and soil water movement (including runoff generation) are simulated separately for 16 distinct land-use classes within each element. The land-use specific soil column is simulated by a modified form of the Xinanjiang approach (Todini, 1996). Runoff concentration within each element is simulated by three parallel linear reservoirs: one for groundwater discharge, one for interflow, and one for direct runoff. Direct runoff comprises fast subsurface runoff and overland flow. Flood routing within the river sections is performed with a kinematic wave approach. The temporal resolution of the water balance calculation is one hour. The model needs as meteorological input spatial fields of precipitation, temperature, air pressure, wind speed, global radiation and relative

humidity. In this study, we run LARSIM with time series of climatic data from different sources. Each of data sets used has a length of ten years (see also Table 2):

1. Measured meteorological data from 56 stations of the German Meteorological Service (DWD), period 1993–2003 (1993 was for warm-up; simulation period was 1994–2003)

2. COSMO-CLM control run (scenario C20_1)

3. COSMO-CLM projection, period 2015–2024 (scenario A1B_1)

Measured meteorological data has been interpolated using INTERMET software (Gerlach, 2006). Here, a stratified kriging interpolation with external drift is implemented, which mainly bases on topographical information and automatic classification of weather conditions and altitudinal temperature gradient for each time step.

COSMO-CLM (CCLM) data originates from a run of version COSMO4.2-CLM3 on 5 km grid resolution within the LandCaRe 2020 project (Berg et al., 2008; Köstner et al., 2008; Rockel et al., 2006). Bias correction has been applied only to precipitation. Each data set has been bilinearly interpolated on a 1 km grid. Measured runoff at the three gaging stations Kronweiler, Kellenbach and Gensingen covers the period from 1990 to 2003.

3.2 Model calibration

We implemented a MATLAB version of the official water balance model of the State Office for Environment,

Table 2. Statistics of rainfall fields: Mean annual precipitation in mm.

			Kronweiler	Kellenbach	Gensingen	Nahe
1994–2003	measured	mean	937.6	690.9	563.9	754.6
		std	177.9	128.9	116.6	155.6
control run	CCLM, bias correction	mean	813.0	702.3	557.5	714.4
		std	97.9	96.2	74.2	85.5
2015–2024	CCLM, no bias correction	mean	790.3	728.8	677.7	754.7
		std	76.2	69.0	57.9	53.2
2015–2024	CCLM, bias correction	mean	819.1	673.2	546.1	705.5
		std	82.6	70.2	53.8	52.3

Water Management and Trade Control (LUWG) Rhineland-Palatinate (Mainz/Germany). For better comparison of results, we use in our study the official parameter sets (Elpers et al., 2008). This model has been calibrated manually in four iterative steps on catchment scale. Calibration has been mainly based on visual comparison of simulated and measured discharge. The following four steps have been executed: (1) calibration of base flow storage parameters for low flow periods, (2) calibration of interflow storage parameters for mean flow periods, (3) calibration of parameters influencing surface runoff components by focusing on flood peaks, (4) calibration of channel roughness to compensate for a visible lag in flood peaks. Steps (1) to (4) have been iterated until a reasonable compromise for all 3 runoff components has been reached. No spatial differentiation has been made: only one parameter set has been assigned to all elements of the catchment under investigation. Vegetation parameters have not been calibrated. Calibration period for the model was 1997 to 1999; validation period was 2000 to 2003. The year 1996 was for warm-up. Model validation has been done based on Nash-Sutcliffe-Efficiencies (NSE). For the catchments Kronweiler and Kellenbach NSEs are similar for both, calibration and validation period (Table 3). For the catchment Gensingen model efficiency is very low for the calibration period (NSE = 0.64). This can be at least partly explained by the fact, that the water balance could not be closed within reasonable parameter bounds. Inconsistencies in the rating curve after the year 1999 (e.g. increase of mean annual runoff) resulted in a significantly higher model efficiency for the validation period. Despite these inconsistencies, we decided to include this catchment in our investigation. It is by far the driest catchment in our area and we expected the highest sensitivity to climate change among all catchments.

3.3 Bias correction of precipitation

For bias correction of the aggregated CCLM 5 km daily precipitation fields we chose the quantile matching method (Maraun et al., 2010; Michelangeli et al., 2009; Piani et al., 2010; Sennikovs and Bethers, 2009). The quantile match-

ing is based upon the cumulative distribution function (CDF), defined as:

$$F(x) = P(X \leq x), \tag{1}$$

and the inverse of the CDF, defined as the quantile function:

$$F^{-1}(P) = x(F). \tag{2}$$

The CDF is either a parametric or non-parametric (i.e. empirical) function. Parametric functions for precipitation intensities are usually gamma or exponential functions (Piani et al., 2010). To account for correcting the probabilities for no precipitation (dry day) together with the probabilities of a wet day ($x > 0$), we chose an empirical CDF $F(x) = i/n$, with i the rank and n the sample size.

Let x_c be the daily precipitation intensities of a time series from CCLM and x_s a time series from a precipitation station, then the quantile matching sets:

$$F_s(x_s) = F_c(x_c). \tag{3}$$

By rearranging Eq. (3) using the quantile function it is possible to calculate a corrected time series for the CCLM from the quantiles of x_s with the probabilities $F_c(x_c)$:

$$x_{cor} = F_s^{-1}(F_c(x_x)). \tag{4}$$

On the left side of Eq. (4) stands the new time series and on the right a transfer function:

$$T(x_c) = F_s^{-1}(F_c(x_c)). \tag{5}$$

This quantile matching corrects the whole intensity distribution of the modeled precipitation and therefore preserves all moments (Sennikovs and Bethers, 2009). The gained transfer functions can be applied to the future scenario if assumed that the model error is the same for the control run and scenario run (Van Roosmalen et al., 2011) and the transfer functions do not change with time (stationarity) (Maraun et al., 2010). Provided that the bias correction is optimal in the control period and the model error is removed from the control run as well from the scenario run, the remaining signal is only due to climate change (Van Roosmalen et al., 2011).

Table 3. Nash-Sutcliffe-Efficiencies.

		Kronweiler	Kellenbach	Gensingen
1997–1999	Calibration period	0.85	0.83	0.64
2000–2003	Validation period	0.80	0.87	0.82

Because x_c is a discrete time series, $T(x_c)$ has to be interpolated to become a continuous function. This has been done with a linear approach (Gutjahr et al., 2011). A spatial interpolation of the transfer functions is carried out for all grid boxes containing no gaging station by an inverse distance weighting method, including the three nearest precipitation stations:

$$\hat{x}_{\text{cor}} = \frac{\sum_{i=1}^{n} \frac{1}{d_i^P} T_i(x_c)}{\sum_{i=1}^{n} \frac{1}{d_i^P}}, \tag{6}$$

with $n = 3$, the number of stations, d the distance of the station to the CCLM grid box center and P the power parameter. Finally, LARSIM needs hourly input data. Therefore, the bias corrected daily precipitation fields from CCLM are disaggregated to hourly fields $H_{i,k}^{\text{cor}}$ by:

$$H_{i,k}^{\text{cor}} = H_{i,k}^{\text{uncor}} \cdot \frac{D_{i,k}^{\text{cor}}}{D_{i,k}^{\text{uncor}}}, \tag{7}$$

with $H_{i,k}^{\text{uncor}}$ the original CCLM precipitation fields, $D_{i,k}^{\text{uncor}}$ the original uncorrected aggregated daily CCLM precipitation fields and $D_{i,k}^{\text{cor}}$ the resulting daily fields after the bias correction. Index i denotes the hours and index k denotes the days.

3.4 Flow Duration Curves

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 equaled or exceeded. This provides a probabilistic description of stream flow at a given location (Fig. 2).

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.

3.5 Calculation of runoff coefficients

Event runoff coefficients specify the percentage of precipitation that appears as significant runoff above base flow following directly the corresponding rainfall. This study uses the direct approach of event-based runoff coefficient (Eq. 8) as described by Merz et al. (2006) and Norbiato et al. (2009).

$$\text{ERC} = \frac{\sum Q_d}{A_{\text{eo}} \cdot \sum \text{prec} \cdot 1000} \cdot 3.6 \tag{8}$$

with: ERC = Event Runoff Coefficient, Q_d = direct event runoff [m^3 h^{-1}], A_{eo} = catchment area [km^2] and prec = areal event precipitation [mm h^{-1}].

The semi-automatic method to calculate event-based runoff coefficients was developed for Austria by Merz at al. (2006). We adapted this method for catchments in Rhineland-Palatinate by alteration of program parameters and verification of calculated runoff coefficients with manually calculated runoff coefficients. The same set of adapted criteria is used for all catchments in this study.

The calculation of runoff coefficients follows a four-step approach: First, observed runoff is separated into baseflow and direct flow using the digital filter proposed by Chapman and Maxwell (1996). Second, to define events, each runoff time series was screened from the largest peak flow to the second largest peak flow and so forth as specified below. A peak flow was assumed to be the peak flow of a potential event, if the ratio of direct runoff to baseflow at time of the peak was larger than 2 and there was no larger flow in the previous and following 12 h. For each peak flow, the start and end of an event was searched within a given time period by finding the time where the direct runoff becomes lower than a given threshold, which depends on the direct runoff at the time of the peak flow. If no starting point was found, the search was repeated by gradually increasing the time period and the threshold. With this iterative approach, the direct runoff at the beginning and end of an event is as small as possible (Merz et al., 2006; Norbiato et al., 2009). Next, direct event runoff and event rainfall volume were calculated and event runoff coefficients were estimated following Eq. (8). Last, to improve data quality we eliminated very small events, events caused by snow melt, events with insufficient data and events with poor event separation.

3.6 Signature indices

Signature indices are used to quantify features resulting from the comparison of FDCs or Empirical Cumulative Distribution Functions (ECDF) of event runoff coefficients. As a set, these features are a characteristic fingerprint of the differences in hydrological behavior.

We use five indices derived from FDCs, following the definitions given by Yilmaz et al. (2008). For illustration purposes, Fig. 2 shows two strongly different FDCs: FDC$_1$ and FDC$_2$.
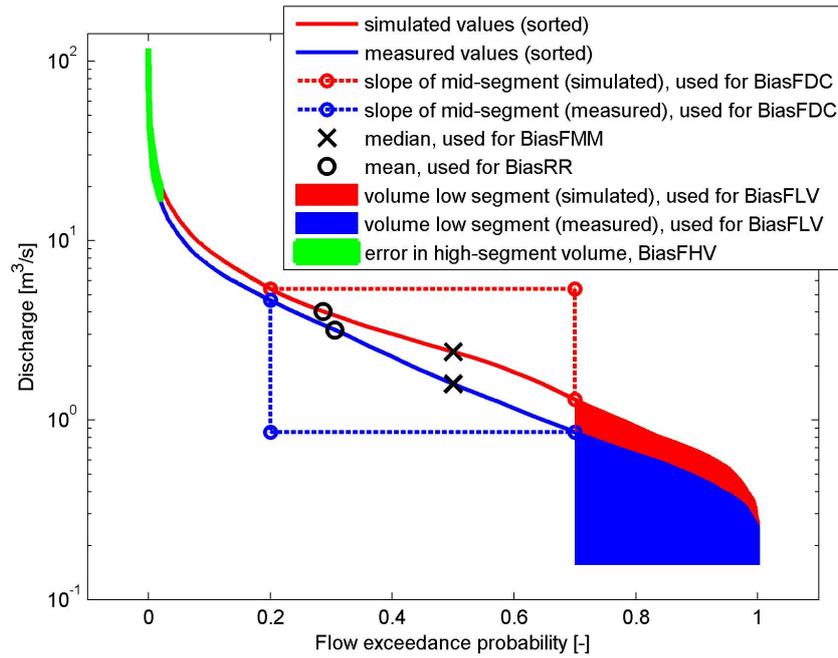
Fig. 2. Two different flow duration curves with highlighted features that are used to determine signature indices.

1. BiasRR: percent bias in the mean values:

$$\text{BiasRR} = \frac{\text{mean}(\text{FDC}_1) - \text{mean}(\text{FDC}_2)}{\text{mean}(\text{FDC}_2)} \cdot 100. \quad (9)$$

BiasRR, which is highlighted by circles (Fig. 2), quantifies the differences in balance.

2. BiasFDCmidslope: percent bias in slope of the mid-segment:

BiasFDCmidslope

$$= \frac{\left(\log(\text{FDC}_{1,0.2}) - \log(\text{FDC}_{1,0.7})\right) - \left(\log(\text{FDC}_{2,0.2}) - \log(\text{FDC}_{2,0.7})\right)}{\left(\log(\text{FDC}_{2,0.2}) - \log(\text{FDC}_{2,0.7})\right)} \cdot 100, \quad (10)$$

where $\text{FDC}_{i,p}$ is the runoff with exceedance probability p of FDC number i (red and blue triangles in Fig. 2). It quantifies the flashiness of flows.

3. BiasFHV: percent bias in high-segment volumes:

$$\text{BiasFHV} = \frac{\int_0^{0.02} \text{FDC}_{1,p}\,\mathrm{d}p - \int_0^{0.02} \text{FDC}_{2,p}\,\mathrm{d}p}{\int_0^{0.02} \text{FDC}_{2,p}\,\mathrm{d}p} \cdot 100, \quad (11)$$

this corresponds to the green area in Fig. 2 and compares the peak discharges.

4. BiasFLV: differences in long-term baseflow:

BiasFLV

$$= \frac{\int_{0.7}^{1}\left(\log(\text{FDC}_{1,p}) - \log(Q_{\min})\right)\mathrm{d}p - \int_{0.7}^{1}\left(\log(\text{FDC}_{2,p}) - \log(Q_{\min})\right)\mathrm{d}p}{\int_{0.7}^{1}\left(\log(\text{FDC}_{2,p}) - \log(Q_{\min})\right)\mathrm{d}p} \cdot 100 \quad (12)$$

where $Q_{\min}$ is the minimum value of $\text{FDC}_{1,1}$ and $\text{FDC}_{2,1}$, i.e. the lowest runoff at all. The two compared areas are highlighted in red and blue (Fig. 2).

5. BiasFMM: percent bias in mid range flow levels:

$$\text{BiasFMM} = \frac{\text{median}(\text{FDC}_1) - \text{median}(\text{FDC}_2)}{\text{median}(\text{FDC}_2)} \cdot 100. \quad (13)$$

It is highlighted in Fig. 2 by crosses. We defined BiasRR, BiasFLV and BiasFMM differently compared to Yilmaz et al. (2008). The other four indices use ECDFs of event runoff coefficients. ECDFs estimate the true underlying distribution function of the points of a sample by empirical measures of the sample. From the ECDFs of event runoff coefficients (ERCs) we derive four indices (Fig. 3):

6. BiasERC: percent bias of mean runoff coefficients.

$$\text{BiasERC} = \left(\frac{1}{m_1}\sum_{j_1=1}^{m_1}\text{ERC}_{j_1} - \frac{1}{m_2}\sum_{j_2=1}^{m_2}\text{ERC}_{j_2}\right) \cdot 100, \quad (14)$$

where m_1 and m_2 are the number of events from datasets 1 and 2

7. BiasERCcv: percent bias of coefficients of variation, describing variability of runoff coefficients of

one catchment.

$$\mathrm{BiasERCcv} = \left(\frac{\sqrt{\frac{1}{m_1-1}\sum_{j_1=1}^{m_1}(\mathrm{ERC}_{j1}-\mathrm{ERCmean}_1)^2}}{\mathrm{ERCmean}_1} \right.$$

$$\left. - \frac{\sqrt{\frac{1}{m_2-1}\sum_{j_2=1}^{m_2}(\mathrm{ERC}_{j2}-\mathrm{ERCmean}_2)^2}}{\mathrm{ERCmean}_2} \right) \cdot 100 \quad (15)$$

8. BiasERCSummer: percent bias of mean runoff coefficient in summer (May to October). Calculation like BiasERC (Eq. 14) for events between May and October.

9. BiasERCWinter: percent bias of mean runoff coefficient in winter (November to April). Calculation like BiasERC (Eq. 14) for events between November and April.

3.7 Sensitivity analysis of signature indices

The relatively short length (10 yr) of our simulation periods gives rise to the question how sensitive are our signature indices to the selection of the corresponding observation period. Fortunately, we have 14 yr of observation data at our disposal. This enables us to move the starting point of the 10-yr observation period needed for index calculation in daily steps by a maximum of 4 yr. For each shift of one day, we recalculate our signature indices. The resulting values are then plotted against the shift in days. This gives a first impression about the degree of uncertainty induced by the selection of the observation period for index calculation.

4 Results

To demonstrate the discriminating power of the nine signature indices, we apply the methodology for four different cases: (1) Assessment of model error, (2) Assessment of bias correction, (3) Assessment of CCLM control run and (4) Detection of climate change signal in the CCLM data set. In addition, we investigate the sensitivity of signature indices by shifting the observation period (5).

4.1 Assessment of model error

To assess the error of the hydrological model, deviations between simulated discharge and measured discharge are calculated (Fig. 4). For all catchments, the model shows higher mean event runoff coefficients in summer (positive values of BiasERCSummer). This indicates a wrong simulation of pre-event conditions during summer. For the catchments

Kronweiler and Kellenbach the other deviations are reasonably low. In contrast, simulated runoff for Gensingen is much higher than the measured one resulting in a large positive bias for 8 of the 9 indices. This can be explained by incorrect model calibration as well as by incorrect runoff measurements at the gaging station Gensingen.

4.2 Assessment of bias correction for precipitation

Bias correction affects the spatial precipitation pattern. For the catchment Kronweiler, bias correction increases the mean annual precipitation only by $29\,\mathrm{mm\,a^{-1}}$ ($+3.6\,\%$) (Table 2). Consequently, the hydrological behavior remains very similar to the uncorrected data set (Fig. 5). For the catchment Kellenbach, bias correction decreases mean annual precipitation by $55.6\,\mathrm{mm\,a^{-1}}$ ($-7.6\,\%$), resulting in a visible but moderate decrease of 7 index values. For the catchment Gensingen, bias correction decreases the mean annual precipitation by $131.6\,\mathrm{mm\,a^{-1}}$ ($-19.4\,\%$), whereas the simulated runoff decreases by $66\,\%$ (BiasRR) compared to the uncorrected dataset (Fig. 5). Much dryer pre-event conditions result in lower BiasERC with higher variance (BiasERCcv). While for the other two catchments bias correction of precipitation only causes moderate changes in hydrologic response, the example of Gensingen highlights the non-linear relationship between runoff and precipitation when mean annual precipitation falls below potential evaporation (Table 1). In addition, we have to note that bias correction for the Kronweiler catchment is far too low as we can conclude from the control run (Table 2). Here, the difference between bias corrected CCLM control run and measured rainfall is $-124.6\,\mathrm{mm\,a^{-1}}$ ($-13.3\,\%$).

4.3 Assessment of CCLM control run (bias corrected for precipitation)

Deviation between bias corrected CCLM control run and the measured climatic input is clearly visible (Fig. 6). For Kronweiler a small decrease in overall runoff (BiasRR), reactivity (BiasFDCmidslope) and peak flows (BiasFHV) can be detected. This can be explained by the lower yearly mean precipitation in the CCLM control run (Table 2). In contrast, a much higher mean event runoff coefficient in summer (BiasERCSummer) can be observed. This is caused by the lower mean annual temperatures in the CCLM control run. A temperature bias of approx. $-1.6\,°\mathrm{C}$ (Table 4) causes lower evaporation rates resulting in higher pre-event water content and thus higher event runoff coefficients.

The other two gaging stations show a clear increase for most of the signature indices, though there is no difference in yearly mean precipitation between the two datasets (Table 2). This increase is only seemingly contradictory: the difference can again be explained by the lower mean annual temperatures in the CCLM control run (bias of -1.2 and $-1.5\,°\mathrm{C}$).
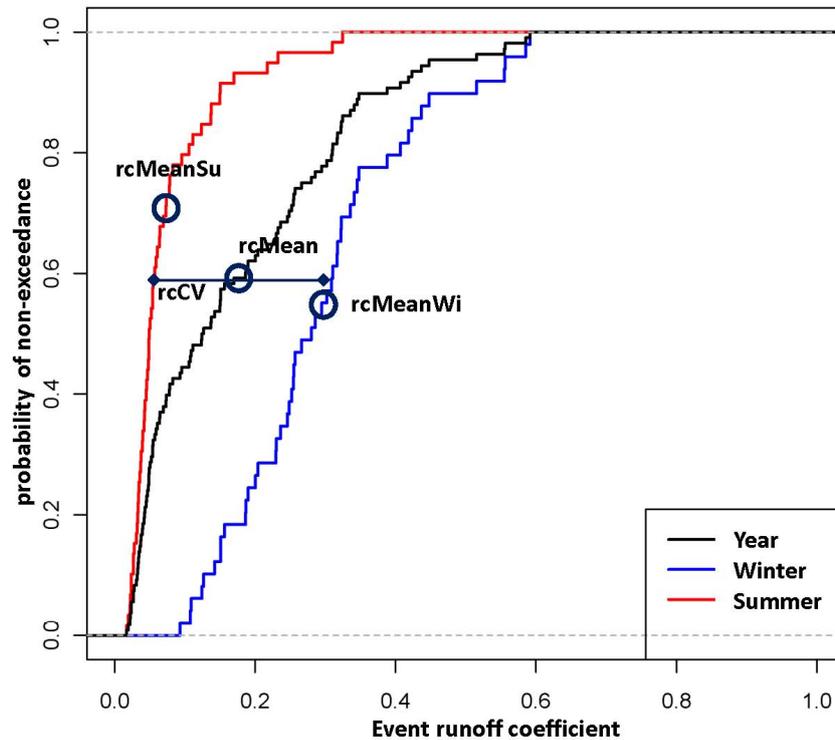
Fig. 3. Empirical Cumulative Distribution Function of event runoff coefficients (ERC) for catchment Kellenbach and derived signature indices.
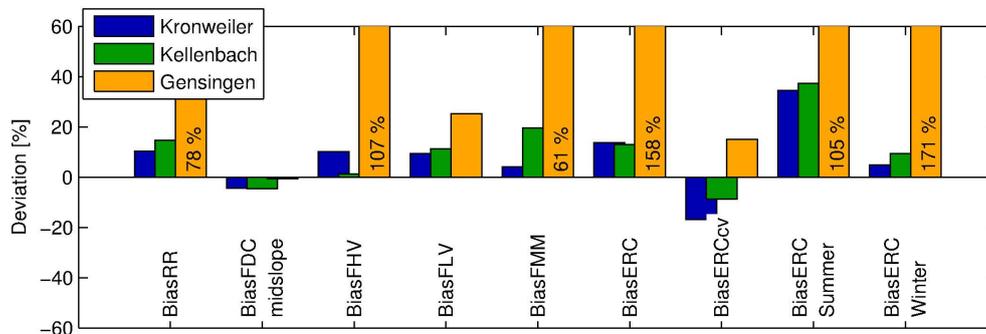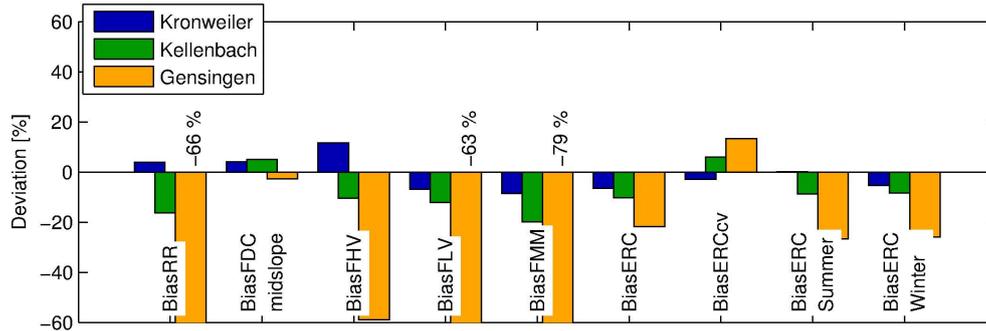

Fig. 4. Signature indices resulting from comparison of (1) the measured discharge time series and (2) the simulated discharge time series using measured meteorological input data, 1994–2003.

4.4 Detection of climate change signal in CCLM data

A climate change signal can be detected by analyzing the differences between the control run and the future projection of climate: For the catchments Kellenbach and Gensingen, a small decrease in annual precipitation (Table 2) and a clear increase in temperature (Table 4) cause a decrease in high flows (negative index BiasFHV) and event runoff coefficients (Fig. 7). Also the water balance (negative index BiasRR) and the reactivity of the catchment (negative index

BiasFDCmidslope) decrease. Partly contrasting, the catchment Kronweiler shows a small increase in annual precipitation, which seems to be compensated by the higher evaporation losses, resulting in index values close to zero. Only the high flow volume decreases (negative index BiasFHV). This decrease can probably be explained by a slightly different temporal distribution of rainfall, because the indices for the event runoff coefficients remain unchanged compared to the control run.

Fig. 5. Signature indices resulting from comparison of simulated discharge time series using (1) original CCLM-data and (2) bias corrected CCLM-data, 2015–2024.
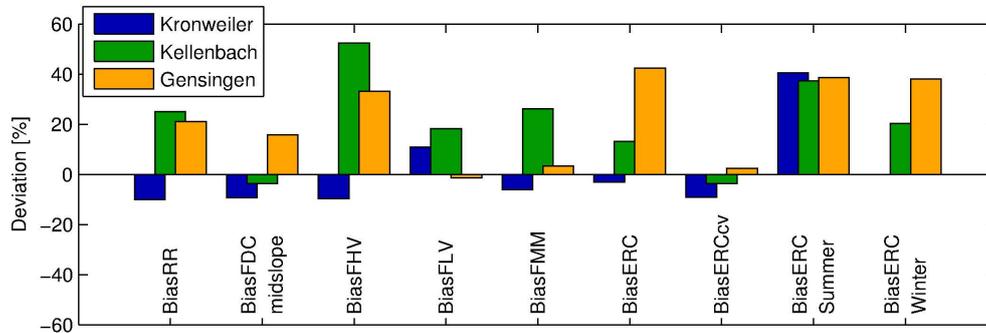
Fig. 6. Signature indices resulting from comparison of simulated discharge time series using (1) measured meteorological input data and (2) bias corrected CCLM-data, control period.

4.5 Sensitivity of signature indices to shift in observation period

A shift in the observation period results in visible fluctuations of signature indices (Fig. 8). Derivations from FDC are more sensitive than signature indices based on ERC. Inclusion of very dry or wet years or inclusion of extreme events (in our example the 100-yr flood event in 1993) mainly affects BiasRR, BiasFMM and BiasFHV. Signature Indices based on ERCs show a largely stable mean. BiasERCWinter remains almost constant. This is a specific property of the catchment Kronweiler; in winter, we continuously observe high event runoff coefficients due to high soil moisture conditions (see also Table 1). BiasRR and BiasFMM are highly correlated, but show – in our example – a considerable difference in level. This is not the case for the catchment Kellenbach; here, level and variation of both indices are identical (not shown).

5 Discussion

Our study revealed differences between simulated discharge (using measured meteorological data as input) and measured discharge for all gaging stations (Fig. 4). These differences are partly due to calibration errors of the model (Kronweiler, Kellenbach), but also clearly indicate a larger balance error at the gaging station Gensingen. Here, the large positive bias for 8 of 9 indices could be explained by incorrect model calibration as well as by incorrect runoff measurements at the gaging station Gensingen. Both presumptions are explainable by visible inconsistencies in the rating curve of this gaging station. It looks as though we have higher mean discharges after the year 1999. For the catchments Kronweiler and Kellenbach, we recommend the use of signature indices for multi-criteria model calibration leading to more behavioral model parameterizations (Van Werkhoven et al., 2008; Herbst et al., 2009a, b). In the case of Gensingen, model calibration on signature indices would lead to an invalid model parameterization due to inconsistencies in the rating curve of the gaging station. Only after a readjustment of the rating curve, model calibration can be successful.

We expect the highest impact of climate change on hydrologic systems when annual precipitation comes close to or even below potential evaporation. Due to the highly nonlinear behavior of our hydrological models, the sensitivity of the proposed signature indices increases in such cases (Fig. 5, see catchment Gensingen).

The large differences between the CCLM control period and the measured climatic data (Fig. 6) lead to the conclusion that bias correction of temperature fields and further

Table 4. Statistics of temperature fields: mean annual temperature in °C.

			Kronweiler	Kellenbach	Gensingen	Nahe
1994–2003	measured	mean	8.63	8.86	10.06	9.35
		std	0.74	0.80	0.67	0.76
control run	CCLM	mean	7.04	7.32	8.81	7.87
		std	0.65	0.69	0.67	0.67
2015–2024	CCLM	mean	7.74	7.97	9.48	8.63
		std	0.63	0.61	0.61	0.62
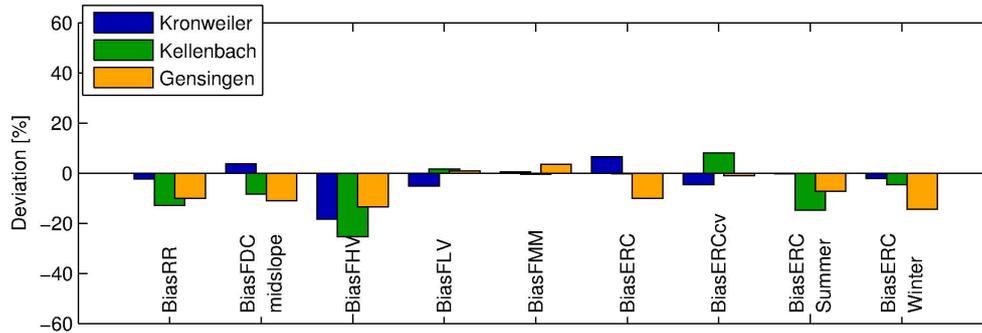
Fig. 7. Signature indices resulting from comparison of simulated discharge time series using bias corrected CCLM-data, 1 km resolution of the periods (1) control period and (2) 2015–2024.

adjustment of bias correction of precipitation fields, especially in the small mountainous ranges, is indispensable. Evaluation of the different approaches for bias correction of precipitation and temperature should be based on subsequent hydrologic simulation and calculation of the proposed signature indices. This is especially important when calculated for representative sub-catchments in a larger basin area, where only a sparse network of observation points is available for bias correction (Benoit et al., 2000). In this case, signature indices act as indirect "efficiency measures" or "similarity measures" for the control period of the simulation. Therefore, application of signature indices for the control period facilitates the decision on the suitability of the bias corrected data for future impact studies (Fowler et al., 2007).

The impact of projected climate change (2015–2024) on the hydrology of our catchments is very small (Fig. 7), particularly if compared to other sources of error (model structure, model calibration, bias correction). Nevertheless, all catchments show a tendency, which clearly corresponds to the expectations derived from the properties of CCLM data sets (see Tables 2 and 4).

Our study is based on relatively short time series of 10 yr length. This fact prohibits application of conventional statistics (e.g. derivation of frequency distributions). Fortunately, most of the proposed signature indices show only moderate sensitivity to a shift of the observation period (Fig. 8). BiasFHV evaluates the upper 2 % of all values and therefore

focuses on rainfall driven periods with high discharges without being too sensitive on single extreme peaks. However, BiasFHV is quite sensitive to extended flood periods (like the 100-yr event in winter 1993, as shown in Fig. 8). BiasRR and BiasFMM clearly react on fluctuations of the long-term water balance. To avoid misinterpretation of results, we recommend carefully selecting a representative observation period, especially with regard to extremes. In our case, the selected observation period (1994–2003) shows much higher standard deviation for annual precipitation than the CCLM control period (Table 2). This is mainly caused by 2 extreme values included in the data set: a 50-yr flood (1995) and the driest summer of the total observation period (2003).

In some special cases, the selection of mean values from the empirical distributions of event runoff coefficients (BiasERC, BiasERCSummer, BiasERCWinter) may not sufficiently distinguish the three distributions (Fig. 3). To circumvent this, these signature indices may be based on the slope of a particular segment of the distribution function (Ley et al., 2011) or the distribution function may be parameterized and classified as proposed by Merz et al. (2006).

6 Conclusions and outlook

It could be shown that application of signature indices is a valuable tool to assess differences in simulated runoff behavior resulting from climatic data sets of different source
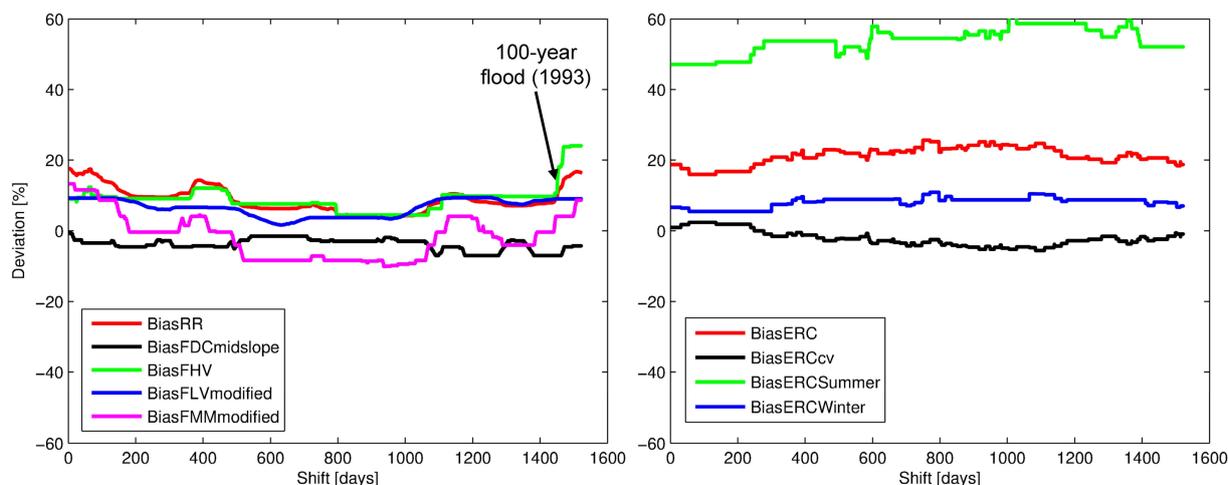
Fig. 8. Sensitivity of signature indices to selection of observation period (catchment Kronweiler).

and/or time reference. The hydrological model acts as a filter for the meteorological input and is therefore sensitive to biases in mean and spatial distribution of precipitation and temperature (Wood et al., 2004; Sangati et al., 2009).

Our study underlines the applicability of the proposed method; it supports on the one hand the elimination of the main sources of error (model calibration, bias correction), and it allows on the other a clear assessment of important dynamical aspects of runoff behaviour caused by a changing climate. Evaluation of dynamic system behavior remains applicable when direct comparison based on conventional statistical metrics is not possible. In any case, we could show that the proposed method clearly visualizes hydrological changes and allows for a straightforward interpretation of results. In detail, the selected signature indices allow fast assessment of changes in water balance (BiasRR, BiasFMM), vertical water distribution (BiasFHV, BiasFLV), reactivity (BiasFDCmidslope), runoff generation (BiasERC, BiasERCcv) and seasonality (BiasERCSummer, BiasERCWinter).

Fowler et al. (2007) claim with reference to Wood et al. (2004) that the minimum standard for any useful downscaling procedure for hydrological applications is that the "historic (observed) condition must be reproducible". To reach this goal, our future work will focus on improving bias correction for CCLM data sets. The assumptions for the bias correction method used in this study are stationarity of the transfer functions with time and that all possible extreme values occurred in the control period, since there is no extrapolation for future extremes implemented. Shifting the distribution to an extreme value distribution at the tail with a dynamic mixture model could be a feasible solution (Frigessi et al., 2003; Vrac and Naveau, 2007). Another shortcoming affects the physical consistency if one variable of the climate model is corrected with no respect to any covariance with other variables. This causes internal inconsistency if corrected variables are used together with uncorrected variables

(Knutti, 2008). In order to find the most suitable bias correction technique, evaluation of bias corrected data sets should always be based on methods which strongly rely on the interpretation of the hydrologic reaction of a given climate model output. The proposed signature index method represents a step forward in that direction.

Here, it should be mentioned that this study is not intended to draw scientifically sound conclusions on hydrological impact of climate change for our study area. For this purpose an ensemble approach would have been necessary (Knutti, 2008; Teutschbein and Seibert, 2010) as well as an improved bias correction method, which also considers the extreme value problem (Boé et al., 2007). Actually, on the selected scale of 5 km^2, ensemble runs of nested CCLM models are not yet available.

References

Arnell, N. W.: Uncertainty in the relationship between climate forcing and hydrological response in UK catchments, Hydrol. Earth Syst. Sci., 15, 897–912, doi:10.5194/hess-15-897-2011, 2011.

Benoit, R., Pellerin, P., Kouwen, N., Ritchie, H., Donaldson N., Joe, P., and Soulis, E.: Towards the use of coupled atmospheric and hydrologic models at regional scale, Mon. Weather Rev., 128, 1681–1706, 2000.

Berg, M., Wieland, R., Mirschel, W., and Wenkel, K.-O.: LandCaRe 2020 – ein Entscheidungsunterstützungssystem zur Vorhersage und Beurteilung der Potentiale ländlicher Gebiete unter dem Einfluss regionalen Klimawandels, in: Modellierung und Simulation von Ökosystemen, edited by: Gnauk, A., Shaker Verlag Aachen, 214–229, 2008.

Blöschl, G. and Montanari, A.: Climate change impacts – throwing the dice?, Hydrol. Process., 24, 374–381, doi:10.1002/hyp.7574, 2010.

Blume, T., Zehe, E., and Bronstert, A.: Rainfall-runoff response, event-based runoff coefficients and hydrograph separation, Hydrol. Sci. J., 52, 843–862, doi:10.1623/hysj.52.5.843, 2007.

Boé, J., Terray, I., Habets, F., and Martin, E.: Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. Int. J. Climatol., 27, 1643–1655, doi:10.1002/joc.1602, 2007.

Casper, M. C., Herbst, M., Grundmann, J., Buchholz, O., and Bliefernicht, J.: Influence of rainfall variability on the simulation of extreme runoff in small catchments, Hydrol. Wasserbewirts., 53, 132–137, 2009.

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.

Elpers, C., Hohenrainer, J., Großkinsky, B., Vollmer, S., and Richter, K. G.: Aufstellung von Wasserhaushaltsmodellen fuer die Landesflaeche von Rheinland-Pfalz und fuer das Moselgebiet, Teil B: Wasserhaushaltsmodelle Sieg und Rheinland-Pfalz, internal project report (IBL, Karlsruhe), 49 pp., unpublished, 2008.

Fowler, H. J., Blenkinsop, S., and Tebaldi, C.: Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling, Int. J. Climatol. 27, 1547–1578. doi:10.1002/joc.1556, 2007.

Frigessi, A., Haug, O., and Rue, H.: A dynamic mixture model for unsupervised tail estimation without threshold selection, Extremes, 5, 219–235, 2003.

Gerlach, N.: INTERMET – Interpolation meteorologischer Größen, in: Niederschlags-Abfluss-Modellierung zur Verlängerung des Vorhersagezeitraumes operationeller Wasserstands-Abflussvorhersagen, edited by: Bundesanstalt für Gewaesserkunde, Reihe BfG Veranstaltungen, 3/2006, 5–14, 2006.

Gutjahr, O., Heinemann, G., Casper, M. C., and Rock, A.: Statistical bias correction for daily precipitation fields from COSMO-CLM over the Nahe catchment area. European Geosciences Union (EGU), General Assemby, Vienna, Austria, Geophys. Res. Abstracts, Vol. 13, EGU2011-3074, 2011.

Liu, Y., Gupta, H. V., Springer, E., and Wagener, T.: Linking science with environmental decision making: Experiences from an integrated modeling approach to support sustainable water resources management, Environm. Modell. Softw., 23, 846–858, doi:10-1016/j.envsoft.2007.10-007, 2008.

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, 2009a.

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, 2009b.

Jacob, D. and Podzun, R.: Sensitivity studies with the regional climate model REMO, Meteor. Atmos. Phys, 63, 119–129, 1997.

Knutti, R.: Should we believe model predictions of future climate change?, Phil. Trans. Roy. Soc. A 366, 4647–4664, doi:10.1098/rsta.2008.0169, 2008.

Köstner, B., Berg, M., Bernhofer, Ch., Franke, J., Gömann, H., Kersebaum, K. C., Kuhnert, M., Lindau, R., Manderscheid, T., Mengelkamp, H.-T., Mirschel, W., Nendel, C., Nozinski, E., Pätzold, A., Simmer, S., Stonner, R., Weigel, H.-J., Wenkel, K. O., and Wieland, R.: Land, Climate and Resources (LandCaRe) 2020 – Foresight and Potentials in Rural Areas under Regional Climate Change. Ital J Agron/Riv Agron, 3, Suppl., 743–744, 2008.

Ley, R., Casper, M. C., Hellebrand, H., and Merz, R.: Catchment classification by runoff behaviour with self-organizing maps (SOM), Hydrol. Earth Syst. Sci., 15, 2947–2962, doi:10.5194/hess-15-2947-2011, 2011.

Ludwig, K. and Bremicker, M.: The Water Balance Model LARSIM – Design, Content and Applications, Freiburger Schriften zur Hydrologie, Bd. 22, 2006.

Mahmoud, M., Liu, Y., Hartmann, H., Stewart, S., Wagener, T., Semmens, D., Stewart, R., Gupta, H. V., Dominguez, D., Dominguez, F., Hulse, D., Letcher, R., Rashleigh, B., Smith, C., Street, R., Ticehurst, J., Twery, M., van Delden, H., Waldick, R., White, D., and Winter, L.: A formal framework for scenario development in support of environmental decision-making, Environ. Modell. Softw., 24, 798–808, doi:10.1016/j.envsoft.2008.11.010, 2009.

Maraun, D., Wetterhall, F., Ireson, A., Chandler, R., Kendon, E., Widmann, M., Brienen, S., Rust, H., Sauter, T., Themeßl, M., Venema, V., Chun, K., Goodess, C., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change. Recent developments to bridge the gap between dynamical models and the end user, Rev. Geophys., 48, RG3003, doi:10.1029/2009RG000314, 2010.

Marke, T., Mauser, W., Pfeiffer, A., and Zängl, G.: A pragmatic approach for the downscaling and bias correction of regional climate simulations: evaluation in hydrological modeling, Geosci. Model Dev., 4, 759–770, doi:10.5194/gmd-4-759-2011, 2011.

Merz, R., Blöschl, G., and Parajka, J.: Spatio-temporal variability of event runoff coefficients, J. Hydrol., 331, 591–604, doi:10.1016/j.jhydrol.2006.06.008, 2006.

Michelangeli, P.-A., Vrac, M., and Loukos, H.: Probabilistic downscaling approaches: Application to wind cumulative distributian function, Geophys. Res. Lett., 36, L11708, doi:10.1029/2009GL038401, 2009.

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, doi:10.1016/j.jhydrol.2009.06.044, 2009.

Piani, C., Haerter, J., and Coppola, E.: Statistical bias correction for daily precipitation in regional climate models over Europe, Theor. Appl. Climatol., 99, 187–192, doi10.1007/S00704-009-0134-9, 2010.

Rockel, B., Will, A., and Hense, A.: The Regional Climate Model COSMO-CLM (CCLM), Meteorologische Z., 17/4, 347–348, 2008.

Sangati, M., Borga, M., Rabuffetti, D., and Bechini, R.: Influence of rainfall and soil properties spatial aggregation on extreme flash

flood response modelling: An evaluation based on the Sesia river basin, North Western Italy, Adv. Water Resour. 32, 1090–1106. doi:10.1016/j.advwatres.2008.12.007, 2009.

Schaefli, B., Harman, C. J., Sivapalan, M., and Schymanski, S. J.: HESS Opinions: Hydrologic predictions in a changing environment: behavioral modeling, Hydrol. Earth Syst. Sci., 15, 635–646, doi:10.5194/hess-15-635-2011, 2011.

Segond, M.-L., Wheater, H. S., and Onof, C.: The significance of spatial rainfall representation for flood runoff estimation: A numerical evaluation based on the Lee catchment, UK, J. Hydrol., 347, 116–131. doi:10.1016/j.jhydrol.2007.09.040, 2007.

Sennikovs, J. and Bethers, U.: Statistical downscaling method of regional climate model results for hydrological modeling. 18th World IMACS/MODSIM Congress, 2009.

Taye, M. T., Ntegeka, V., Ogiramoi, N. P., and Willems, P.: Assessment of climate change impact on hydrological extremes in two source regions of the Nile River Basin, Hydrol. Earth Syst. Sci., 15, 209–222, doi:10.5194/hess-15-209-2011, 2011.

Teutschbein, C. and Seibert, J.: Regional Climate Models for Hydrological Impact Studies at the Catchment-Scale: A Review of Recent Modeling Strategies, Geography Compass, 4/7 (2010), 834–860, doi:10.1111/j.1749-8198.2010.00357.x, 2010.

Todini, E.: The ARNO rainfall-runoff model, J. Hydrol., 175, 339–385, 1996.

Tramblay, Y., Bouvier, C., Ayral, P.-A., and Marchandise, A.: Impact of rainfall spatial distribution on rainfall-runoff modelling efficiency and initial soil moisture conditions estimation, Nat. Hazards Earth Syst. Sci., 11, 157–170, doi:10.5194/nhess-11-157-2011, 2011.

Van Roosmalen, L., Sonnenborg, T. O., and Christensen, J. H.: Comparison of hydrological simulations of climate change using perturbation of observation and distribution-based scaling, Vadose Zone J., 10, 136–150, doi:10.2136/vzj2010.0112, 2011.

Van Werkhoven, K., Wagener, T., Reed, P., and Tang, Y.: Sensitivity-guided reduction of parametric dimensionality for multi-objective calibration of watershed models, Adv. Water Resour., 32, 1154–1169, doi:10.1016/j.advwatres.2009.03.002, 2009.

Vogel, R. M. and Fennesey, N. M.: Flow Duration Curves I: New Interpretation and Confidence Intervals, J. Water Res. Pl., 120, 485–504, 1994.

Vrac, M. and Naveau, P.: Stochastic downscaling of precipitation: from dry events to heavy rainfall, Water Resour. Res., 43, W07402, doi:10.1029/2006WR005308, 2007.

Wood, A. W., Leung, L. R., Sridhar, V., and Lettenmaier, D. P.: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, Climatic Change, 62, 189–216, 2004.

Wilby, R. L.: Evaluating climate model outputs for hydrological applications – Opinion. Hydrol. Sci. J., 55, 1090–1093, doi:10.1080/02626667.2010.513212, 2010.

Xu, H., Taylor, R. G., and Xu, Y.: Quantifying uncertainty in the impacts of climate change on river discharge in sub-catchments of the Yangtze and Yellow River Basins, China, Hydrol. Earth Syst. Sci., 15, 333–344, doi:10.5194/hess-15-333-2011, 2011.

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, doi:10.1016/j.advwatres.2007.01.005, 2007.

Yang, W., Andréasson, J., Graham, L. P., Olsson, J., Rosberg, J., and Wetterhall, F.: Distribution-based-scaling to improve usability of regional climate model projections for hydrological climate change impacts studies. Hydrol. Res., 41, 211–229, doi:10.2166/nh.2010.004, 2010.

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.

Zhang, Z., Wagener, T., Reed, P., and Bushan, R.: Ensemble streamflow predictions in ungauged basins combining hydrologic indices regionalization and multiobjective optimization. Water Resour. Res., 44, W00B04, doi:10.1029/2008WR006833, 2008.