ORIGINAL PAPER

Seasonality of low flows and dominant processes in the Rhine River

Hakan Tongal · Mehmet C. Demirel ·
Martijn J. Booij

Abstract Low flow forecasting is crucial for sustainable cooling water supply and planning of river navigation in the Rhine River. The first step in reliable low flow forecasting is to understand the characteristics of low flow. In this study, several methods are applied to understand the low flow characteristics of Rhine River basin. In 108 catchments of the Rhine River, winter and summer low flow regions are determined with the seasonality ratio (SR) index. To understand whether different numbers of processes are acting in generating different low flow regimes in seven major sub-basins (namely, East Alpine, West Alpine, Middle Rhine, Neckar, Main, Mosel and Lower Rhine) aggregated from the 108 catchments, the dominant variable concept is adopted from chaos theory. The number of dominant processes within the seven major sub-basins is determined with the correlation dimension analysis. Results of the correlation dimension analysis show that the minimum and maximum required number of variables to represent the low flow dynamics of the seven major sub-basins, except the Middle Rhine and Mosel, is 4 and 9, respectively. For the Mosel and Middle Rhine, the required minimum number of variables is 2 and 6, and the maximum number of variables is 5 and 13, respectively. These results show that the low flow processes of the major sub-basins of the Rhine could be considered as non-stochastic or chaotic processes. To confirm this conclusion, the rescaled range analysis is applied to verify persistency (i.e. non-randomness) in the processes. The estimated rescaled range statistics (i.e. Hurst exponents) are all above 0.5, indicating that persistent long-term memory characteristics exist in the runoff processes. Finally, the mean values of SR indices are compared with the nonlinear analyses results to find significant relationships. The results show that the minimum and maximum numbers of required variables (i.e. processes) to model the dynamic characteristics for five out of the seven major sub-basins are the same, but the observed low flow regimes are different (winter low flow regime and summer low flow regime). These results support the conclusion that a few interrelated nonlinear variables could yield completely different behaviour (i.e. dominant low flow regime).

Keywords Rhine river · Low flows · Seasonality ratio · Correlation dimension · Mutual information · Rescaled range analysis

1 Introduction

Hydrological events such as high flows and low flows are controlled by climatic surpluses and deficits that propagate through the river basin in different seasons. In this paper, we are interested in low flows and we focus on the identification of the number of dominant processes (i.e. predictors) for low flows by using nonlinear methods and the seasonality properties of low flows by analysing the persistence of low flow occurrence days. Detecting the seasonality will help to understand the dominant processes and identifying the number of dominant processes will help to link the processes to the identified numbers. The resultant predictors then can be incorporated in a low flow forecast model.

H. Tongal (✉)
Department of Civil Engineering, Engineering Faculty,
Süleyman Demirel University, 32260 Isparta, Turkey
e-mail: hakantongal@sdu.edu.tr

M. C. Demirel · M. J. Booij
Water Engineering and Management, Faculty of Engineering
Technology, University of Twente, P.O. Box 217, 7500 AE
Enschede, The Netherlands

Low flow is defined as an integral phase of large rivers (Smakhtin 2001). Different processes and storages are involved in the occurrence of low flows. Identifying the number of dominant processes is the first step prior to an elaborated analysis of which dominant processes and which spatial and temporal scales are relevant for the occurrence of low flows. In this paper we particularly address the number of dominant processes that are minimally required for an appropriate model which includes dominant processes in the basin (Booij 2003). The dominant processes concept has been introduced in the last decade (Woods 2002; Sivapalan 2003) to avoid over-parameterization problems that occur when processes that are not important are represented in models (Grayson and Blöschl 2001; Sivakumar 2008). Here, we define a dominant process as a major hydrological driver or a storage component sustaining low flows in a prolonged dry period.

A reliable forecast of low flows several weeks or months in advance using dominant processes can benefit, generally, the management of freshwater resources, and particularly in Europe, the freight shipment and cooling water supply for the energy sector as heavily industrialised cities are located along the rivers. Other river functions such as drinking water supply, agricultural water supply and ecological and recreational activities are also vulnerable to low flows and can profit from reliable forecasts as well.

Many different methods have been applied to reveal dominant processes in river basins such as chaos theory (Sivakumar 2001; Regonda et al. 2004; Sivakumar 2008), wavelet theory (Markovic and Koch 2005; Schaefli et al. 2007), circular statistics (Laaha and Blöschl 2006) and time series analysis techniques (Zaidman et al. 2001; Kavvas et al. 1977).

Laaha and Blöschl (2006) identified winter and summer low flow regions in Austrian river basins using the circular statistics as a part of the seasonality analysis of low flows. Schaefli et al. (2007) identified dominant processes for high flows for the Alpine catchments in Switzerland by analysing the cross correlations between precipitation, temperature and discharge through application of the wavelet coherence method. It was shown that the processes, leading to high flows, can be detected from the extraordinary co-variations of the relevant hydrological processes at various temporal scales. They found a significant coherence between temperature and precipitation during the floods in autumn 1993 and 1994 showing that these floods where the result of exceptional meteorological situations. Further, they also found significant coherence between discharge and temperature at scales between a few days and a month showing that the Alpine catchments do not seem to significantly react to day to day variations of temperature. This is from the fact that only very small areas start contributing to the discharge in contrast to lowland flat

catchments, where a short temperature variation affects large areas simultaneously (Schaefli et al. 2007). Wilcox et al. (1991) analysed daily runoff series to verify if snowmelt runoff processes exhibit chaotic dynamics. The chaotic behaviour of a hydrological process could reveal the determinism in the system which consequently helps identifying the principal drivers of the process. However, Wilcox et al. (1991) found no evidence of chaotic behaviour in the runoff processes. Lisi and Villi (2001) explored chaotic behaviour in the dynamics of river flows using the correlation dimension method. They used the identified chaotic feature (i.e. determinism) in predictions. The results showed that nonlinear river flow modelling is a more effective method to improve the discharge predictions in comparison to a classical linear model. Eng and Milly (2007) analysed the base flow recession time component using regional regression models of low flow characteristics. They found that a base flow recession time component in addition to drainage area can be used as predictor in these low flow regression models. The estimation error was significantly reduced (i.e. halved) in their linear approach including the time component compared to the conventional low flow regression model using only drainage area as predictor.

However, to our knowledge, nonlinear analysis methods have not been applied to low flows. The objective of this study is to identify the number of dominant processes that are acting in generating summer and winter low flows in the Rhine basin. As an essential part of nonlinear theory, the chaos theory offers an appropriate tool to illustrate the characteristics of dynamical systems as well as defining the number of dominant processes (Yu et al. 2011). The correlation dimension method is one of the most frequently employed methods to identify chaotic and non-chaotic behaviour of a system (Sivakumar 2004a). Long term memory properties could affect low flow dynamics. The rescaled range analysis provides an appropriate framework to reveal the long term memory properties.

The Rhine River is the fourth largest river in Europe and heterogeneous in terms of the dominant processes and we, therefore, sub-divided the Rhine basin into seven sub-basins based on previous studies (e.g. Belz and Frauenfelber-Kääb 2007; Hurkmans et al. 2008).

The river discharges at the outlet of the sub-basins were analysed using circular statistics (i.e. seasonality ratio index) and nonlinear techniques such as the mutual information analysis, correlation integral analysis and rescaled range analysis. In this study, low flows are defined as flows below a threshold equal to the 75 % exceedence probability (hereafter Q75) whereas e.g. Laaha and Blöschl (2006) used Q95 as a low flow threshold.

The organization of the paper is as follows. The next section describes the study area and discharge data, the
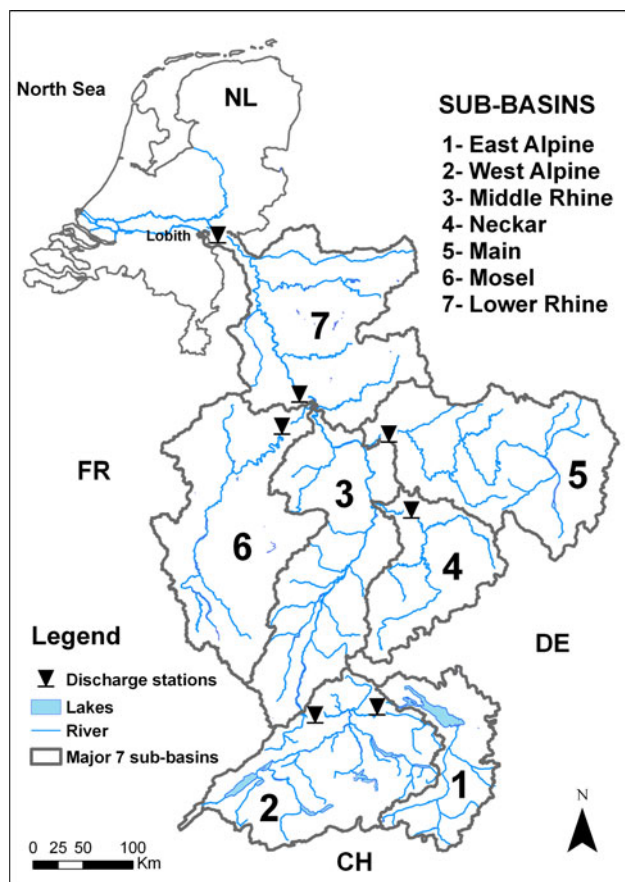
Fig. 1 Pre-defined 7 sub-basins in the Rhine River numbered as (*1*) East Alpine (*2*) West Alpine (*3*) Middle Rhine (*4*) Neckar (*5*) Main (*6*) Mosel (*7*) Lower Rhine

Table 1 Spatial characteristics and generated discharges in the seven sub-basins

Sub-basin	Area (km^2)	Altitude range (m)	Annual generated discharge (mm)
East Alpine (1)	16,051	143–3,270	890
West Alpine (2)	17,679	252–3,980	1,021
Middle Rhine (3)	37,908	67–1,340	344
Neckar (4)	12,616	90–970	363
Main (5)	24,833	83–939	244
Mosel (6)	27,262	59–1,326	410
Lower Rhine (7)	23,738	5–779	273

methods we applied to estimate the number of dominant processes and the seasonality ratio index are explained in the third section, the results are discussed in the fourth section and finally conclusions are drawn in the fifth section.

2 Study area and data

2.1 Study area

The Rhine River flows through a 1,233 km long course (Fig. 1) and its basin covers an area of 185,300 km^2. The Rhine basin covers part of Switzerland, Germany, Luxembourg, France and the Netherlands and about 60 million people live in the basin (Huisman et al. 2000). Besides that, more than 60 percent of the Netherlands' fresh surface waters comes from the Rhine (Middelkoop and Van Haselen 1999). The altitudes in the river basin vary from 4,000 m in the Alps to 6 m below sea level in the Netherlands (Table 1).

The average discharge in the upstream area before Lake Constance is around 1,000 m^3 s^{-1}. This discharge level increases up to 2,300 m^3 s^{-1} at the Dutch border (Lobith) (Fig. 2). In general, the discharge regime of the Rhine River is mainly dominated by rainfall and melt-water coming from snow and Alpine glaciers. About 70 % of the summer flow at Lobith comes from the Alps, whereas only around 30 % of the winter flow comes from the Alps as winter precipitation is stored as snow in winter until it melts in late spring (Fig. 3).

In this study, the Rhine sub-basins are organised at two spatial scales; at the scale of 134 catchments and at the scale of 7 sub-basins. It should be noted that the 134 catchments, shown in Fig. 4, have already been identified for previous hydrological modelling purposes in the Rhine River (e.g. Renner et al. 2009; Te Linde et al. 2010; Te Linde et al. 2008; Reggiani et al. 2009). The scale of 7 sub-basins is chosen to explore the large scale signals of dominant processes such as snow melt and groundwater discharges. The upper Rhine is sub-divided into the West Alpine (WA) and East Alpine (EA) Rhine to analysis the impact of Lake Constance on low flows, in addition to snow storage in the Alps. The middle Rhine is sub-divided into four sub-basins: Neckar, Main and Mosel tributaries and the main channel between Basel and Koblenz, which is called the Middle Rhine (MR). The remaining part of the Rhine River from Koblenz to Lobith is called the Lower Rhine (LR) as shown in Fig. 1.

2.2 Discharge data

Daily discharge (Q) series have been provided by the Global Runoff Data Centre (GRDC), Koblenz (Germany). Discharge data of 170 stations for 108 out of the 134 catchments are used to estimate the seasonality ratio indices. There are no discharge stations in 26 sub-basins. Besides, 7 stations at the outlets of the 7 major sub-basins are included in the nonlinear analyses. The daily discharge data series span from 1989 to 2008. The net generated annual discharge for each sub-basin is given in Fig. 3. For the West Alpine, Middle Rhine and Lower Rhine

Fig. 2 Long term average daily discharge for the seven major sub-basins
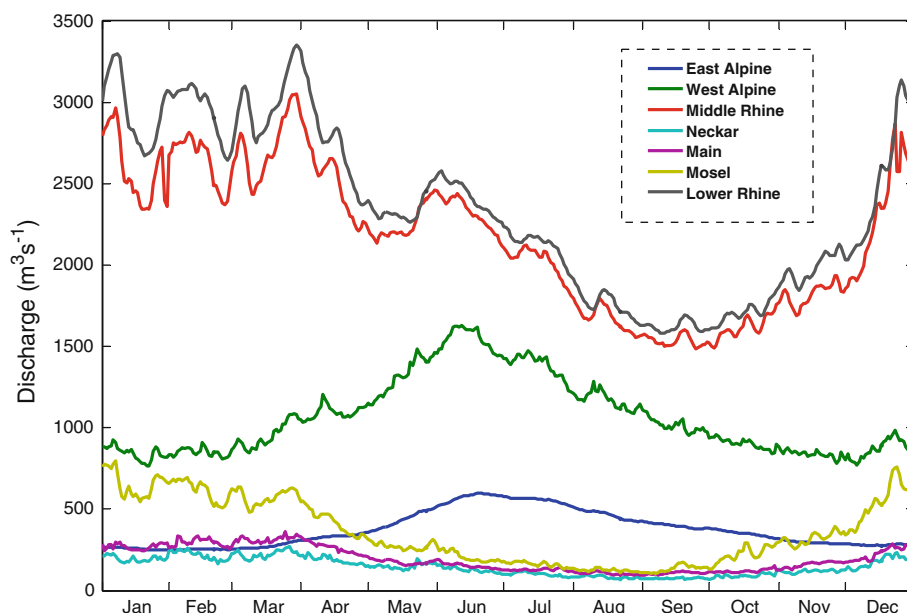
**Fig. 3** Net generated annual discharge (mm) in each sub-basin
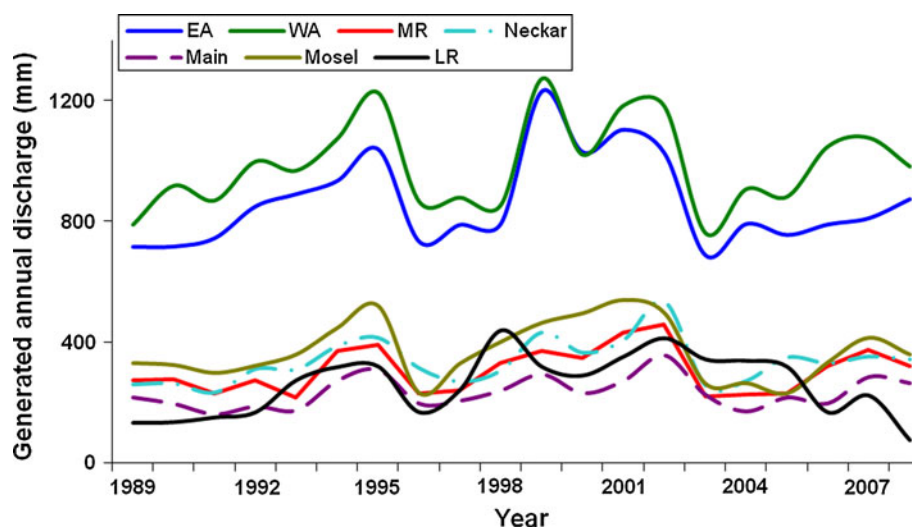
sub-basins, the annual discharge generation is calculated as a net discharge, i.e. the inlet discharge subtracted from the outlet discharge over the sub-basin area. The discharge generation is highest in the WA sub-basin whereas it is lowest in the Main sub-basin (Table 1). Dry years such as 1985 and 2003 can be identified in Fig. 3, since the discharge generation is significantly less in the upstream basins.

3 Methodology

We applied several steps to analyse the low flows and to identify the number of dominant processes in the Rhine basin. These steps are shown in a general framework (Fig. 4). The first step is the seasonality analysis of

discharge series from 108 catchments to render the general picture of low flow seasonality in the Rhine basin. In the second step, we defined 7 major sub-basins by spatially aggregating the 134 catchments. The number of 7 sub-basins is selected to distinguish between large tributaries and associated different processes like snow melt processes in the Alpine sub-basins and groundwater flows in the rainfed sub-basins. In the third step, we applied the correlation dimension method to identify the number of dominant processes in each major sub-basin. The delay times required for the correlation dimension analysis were calculated with the mutual information method. In the fourth step, the validities of the estimated correlation dimensions were examined with the rescaled range analysis. In the fifth and final step, the spatial variability of correlation dimension values was investigated by comparing them with the
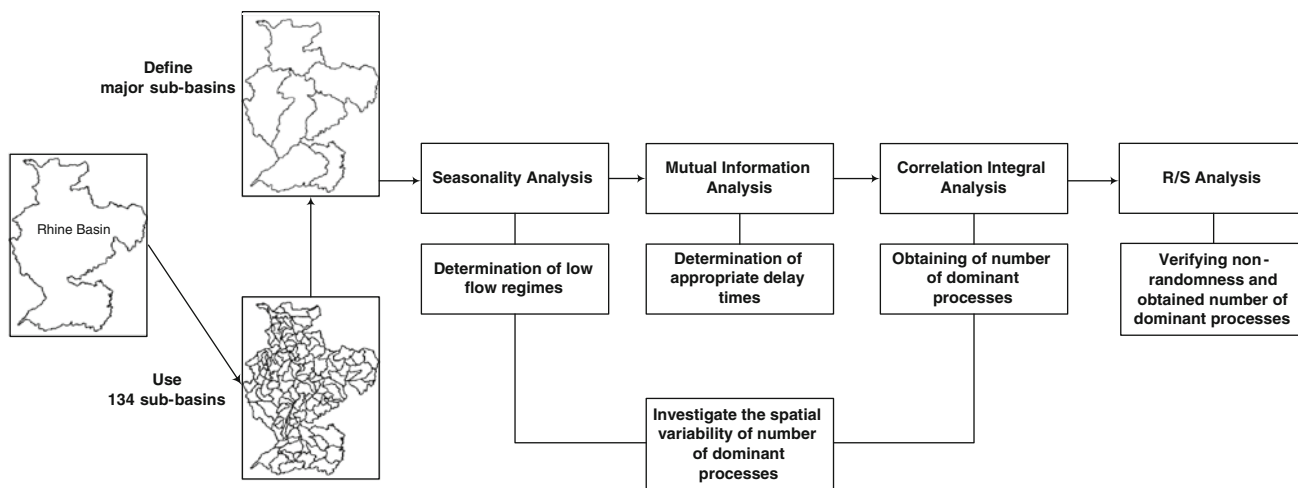
Fig. 4 General framework for identifying the number of dominant winter and summer processes in the Rhine River

mean seasonality ratio indices of the seven major sub-basins.

3.1 Seasonality ratio index

Low flows in summer and winter have different spatial characteristics caused by the variability of catchment properties and climatology. The seasonality ratio (SR) index reflects these low flow characteristics and underlying hydrological processes (Laaha and Blöschl 2006). Daily discharge data with a minimum length of 20 years from 108 of the 134 catchments were used to calculate the seasonality ratio for each sub-basin. The series were divided into summer and winter discharge series. The threshold level, Q75 (flows extracted from flow duration curve and exceeded 75 % of the time) was selected for the definition of low flows in the Rhine River for summer low flows (Q_{75s}) and winter low flows (Q_{75w}). The SR index is calculated as shown in Eq. (1).

$$SR = Q_{75s}/Q_{75w} \qquad (1)$$

The calculated SR values will be presented in a map. Values of SR > 1 indicate the presence of a winter low flow regime and values of SR < 1 indicate the presence of a summer low flow regime (Laaha and Blöschl 2007).

3.2 Correlation dimension method

In the literature, several methods are available to identify the existence of chaos in a time series (e.g. the Lyapunov exponent, the Kolmogorov entropy, the continuous power spectrum methods). Among these methods, the correlation dimension method has been widely employed in attempts to understand streamflow dynamics from a chaotic perspective (Wang et al. 2006; Sivakumar 2007; Lisi and Villi 2001; Ng et al. 2007). The advantage of the correlation dimension method is that it represents the essential features of a system using time series of only a single variable of the entire system. The resulting correlation dimension indicates not only whether the system is chaotic but also the number of dominant processes governing the system (Sivakumar 2004b).

The most commonly used algorithm for computing the correlation dimension is the Grassberger–Procaccia algorithm (Grassberger and Procaccia 1983). In the Grassberger–Procaccia algorithm, dynamics of the considered system is tried to be captured by constructing an appropriate series of state vectors in the m-dimensional phase space. The phase-space can be reconstructed using the method of delays (Takens 1981; Packard et al. 1980; Ruelle 1981) which tries to capture the underlying dynamics by using single-variable series. The idea behind this method is that a nonlinear system is characterized by self-interaction and a single variable can carry the information of the dynamics of the entire multi-variable system (Sivakumar 2008). For a scalar time series (e.g. streamflow) X_i, where $i = 1, 2, \ldots, N$, N is the number of data points, the phase space can be reconstructed according to:

$$Y_j = (X_j, X_{j+\tau}, X_{j+2\tau}, \ldots, X_{j+(m-1)\tau}) \qquad (2)$$

where $j = 1, 2, \ldots, N - (m-1)\tau$, m is the embedding dimension of the Y_j vector and τ is a delay time. The long term behaviour of a system in the phase space is characterized by a geometric object which is called a d-dimensional attractor. According to the embedding theorem of Takens (1981), to characterize a dynamic system with an attractor dimension d, a ($m = 2d + 1$)-dimensional phase space is required. However, Abarbanel et al. (1990) suggested that $m > d$ would be sufficient.

To construct a representative phase space by the time delay method, a proper choice of τ is critical. The autocorrelation function and the mutual information (MI)

methods are the two commonly preferred methods in determining the delay time (Williams 1997). However, Frazer and Swinney (1986) proposed that since the auto-correlation function measures the linear dependence between successive points, this method may not be appropriate for analyzing nonlinear dynamics. They suggested the use of the local minimum of the mutual information which measures the general dependence between successive points based on probability distributions.

For a discrete time series with X_i and $X_{i-\tau}$ as successive values, the mutual information function MI, is computed according to

$$MI = \sum_{i}^{N} \sum_{i-\tau}^{N} P(X_i, X_{i-\tau}) \log_2 \left[\frac{P(X_i, X_{i-\tau})}{P(X_i)P(X_{i-\tau})} \right] \quad (3)$$

where, N is the number of data points, $P(X_i)$ and $P(X_{i-\tau})$ are the individual probabilities of X_i and $X_{i-\tau}$, respectively and $P(X_i, X_{i-\tau})$ is the joint probability (Sahoo et al. 2009).

The appropriate delay times were selected from the first local minimum of the mutual information function (Frazer and Swinney 1986). To find local minimum of the mutual information function the following formula can be used

$$\text{Relative Change} = \left[\frac{MI_{next} - MI_{current}}{MI_{current}} \right] \times 100 \quad (4)$$

where, $MI_{current}$ is the mutual information value of which relative change will be calculated, MI_{next} is the successive value of $MI_{current}$. Where the relative change of mutual information as a function of delay time, starts to be constant with lag time could be taken as first local minimum. With a properly selected time delay, the considered time series can be reconstructed in the m-dimensional phase space. The correlation integral is then calculated as follows:

$$C(r) = \frac{2}{N(N-1)} \sum_{j=1}^{N} \sum_{i=1}^{N} H(r - \|Y_i - Y_j\|) \quad (5)$$

where H is the Heaviside step function, with $H(u) = 1$ for $u > 0$, and $H(u) = 0$ for $u \leq 0$ where $u = r - \|Y_i - Y_j\|$, r is the radius of a sphere centred on Y_i or Y_j, and N is the number of data points.

The norm, $\|Y_i - Y_j\|$ may be any of the three usual norms, the maximum norm (maximum absolute difference between the Y_i and Y_j), the diamond norm (sum of all the absolute differences) or the standard Euclidean norm) (Jayawardena and Lai 1994). In this study the Euclidean norm is adopted.

$C(r)$ gives the probability of two randomly selected vectors that lie within a certain distance (Ng et al. 2007). The exponent d of the space is then related to the correlation integral as:

$$C(r) \propto r^d \quad (6)$$

and the correlation exponent d can be estimated by taking the logarithm of both sides of Eq. (6). The slope is estimated by a least square fit of a straight line over a certain range of radius (r) (Sivakumar 2005). To detect if the underlying dynamics of the system are chaotic, the correlation exponent (d) must be plotted as a function of the embedding dimension (m). If the correlation exponent increases as the embedding dimension increases without reaching a saturation value, then the system is considered as purely random (e.g. white noise) or non-chaotic (Khokhlov et al. 2008; Wang et al. 2006; Xu et al. 2010).

In a deterministic system, the correlation exponent value reaches a saturation value. The saturation value of the correlation exponent is called the correlation dimension of the system (Xu et al. 2010). The correlation dimension of the system gives a clue about the essential number of independent variables (or dominant variables) to characterize the motion of the system (Sivakumar and Jayawardena 2002). It is generally assumed that the correlation dimension equals the number of degrees of freedom of the system and a larger number of embedding dimensions than the correlation dimension causes redundancy (Ng et al. 2007).

3.3 Rescaled range analysis

In this study, in addition to the correlation dimension analysis, the rescaled range analysis, also known as R/S analysis, is employed to assess long term memory properties of the discharge series. Using this method, persistency and non-persistency properties of the processes can be identified. The results can be used as an additional evidence to confirm or to reject the results obtained from correlation dimension analysis. Long-term persistency could also give a clue about the non-stochasticity of the processes or, in other words, chaotic properties of the processes.

Let $x(t)$ be a time series and for any positive integer $n \geq 1$, the mean value series is defined as:

$$\{x\}_n = \frac{1}{n} \sum_{t=1}^{n} x(t) \quad n = 1, 2, \ldots \quad (7)$$

The cumulative deviation and the extreme deviation is, respectively:

$$X(t, n) = \sum_{u=1}^{t} (x(u) - \{x\}_n) \quad 1 \leq t \leq n \quad (8)$$

$$R(n) = \max_{1 \leq t \leq n} X(t, n) - \min_{1 \leq t \leq n} X(t, n) \quad n = 1, 2, \ldots \quad (9)$$

The standard deviation is:

$$S(n) = \left[\frac{1}{n} \sum_{t=1}^{n} \left(x(t) - \{x\}_n\right)^2\right]^{0.5} \quad n = 1, 2, \dots \quad (10)$$

Hurst proposed a proportional relationship between the ratio between R and S and the positive integer, n, as follows:

$$R/S \propto \left(\frac{n}{2}\right)^{H} \quad (11)$$

where; H is the Hurst coefficient which lies between 0 and 1 (Hurst 1951). The logarithm of both sides of the Eq. (11) will give a linear relationship between the R/S statistic and $(n/2)$ where the slope (H) can easily be found by least square estimation. A Hurst coefficient close to 0.5 indicates a random process where the observations are independent and there is no correlation between one data value and any future value (i.e. has no long memory). A Hurst coefficient between 0.5 and 1 indicates that the future trend of the time series will be consistent with the past. This consistency increases as H approaches to 1. When the H value is between 0 and 0.5, the process has an anti-persistent character and the future trend of the series will be contrary to the past (Xu et al. 2010). This anti-persistency increases as H approaches to 0. However, the studies which do not provide statistical significance test to judge if the Hurst coefficients are statistically different from the one characterizing an independent process may be premature to conclude anything about the presence of long term memory. Therefore, we applied the statistical significance test which is particularly developed for the Hurst coefficient by Couillard and Davison (2005), where the detailed explanation can be found.

4 Results and discussion

4.1 Seasonality ratio index

The SR indices for the 108 of the 134 catchments are shown in Fig. 5. The map shows the seasonal characteristics of low flows. Values of SR > 1 demonstrate the presence of a winter low flow regime and values of SR < 1 indicate the presence of a summer low flow regime. It is clear that the Alpine regions show a winter low flow regime caused by snow storage and frost processes. In the Middle Rhine, some parts show a winter low flow regime and others show a summer low flow regime. This could result from being an extension of the Alpine region (winter low flows) and being between the Neckar and Mosel sub-basins (summer low flows). The remaining sub-basins are dominated by summer low flows. The map with SR values gives a clue that there could be a different number of processes acting in generating different low flow regimes
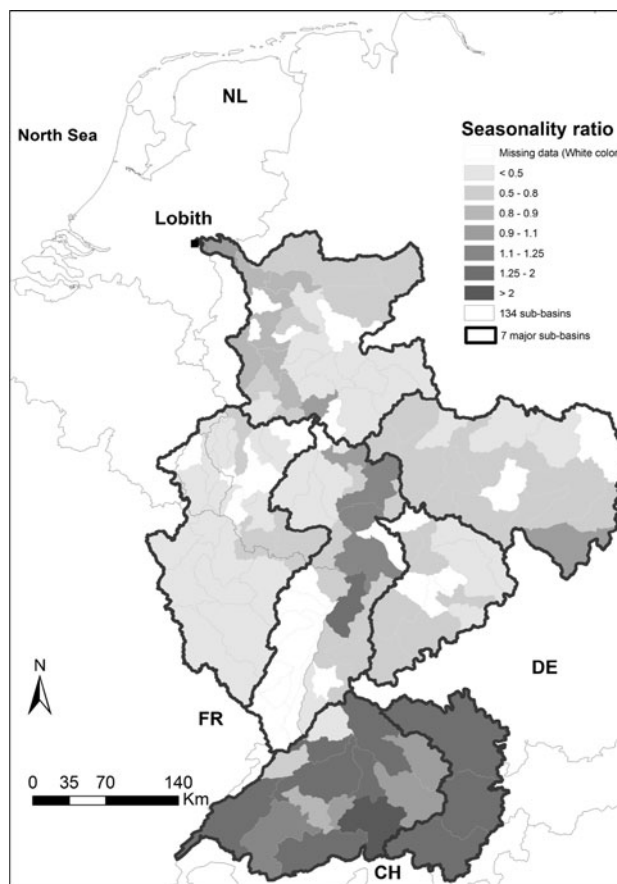
Fig. 5 Seasonality ratio index for 108 catchments in the Rhine basin

(Fig. 5). To investigate this hypothesis, nonlinear analysis methods were applied.

4.2 Correlation dimension

The daily discharges of seven major sub-basins were analysed with the correlation dimension method to reveal the dynamic characteristics. The data set for each sub-basin containing a total of 7,305 daily discharge values is assumed to be enough for chaotic analysis. Sivakumar (2005) showed that 1,000 data points for correlation dimension estimation by applying the Grassberger–Procaccia algorithm were sufficient. To calculate the correlation integral, the delay time (τ) was computed using the mutual information method.

First local minimums were selected as where the relative change in the mutual information function is stable and fluctuates around a constant value. If the delay time is chosen as the first minimum of the mutual information function, then the reconstructed state vector would consist of components that possess minimal mutual information between them (Sivakumar 2000). The mutual information functions against lag time with their relative changes are

shown in Fig. 6 and optimal delay times are given in Table 2.

The first local minimums of mutual information functions were taken as the delay times where the relative changes started to become stable. In Fig. 6, the time series of the West Alpine sub-basin is the most correlated one since its mutual information function value for lag zero is the highest. However, the time series of the West Alpine region is de-correlated after 8 days which could be a clue about higher dependency on the short term that allows short term predictions. The time series of the Neckar has the lowest auto-correlation and modelling and predicting might be difficult. From Table 2, the highest necessary lag time to construct a representative phase space is obtained for the Main and Lower Rhine and the smallest lag time is obtained for the West Alpine and Mosel. After the determination of the delay times, the correlation integrals were computed using the Grassberger–Procaccia algorithm for embedding dimensions (m) varying from 1 to 25 (Fig. 7) for all discharge series. These plots show the relationship between ln $C(r)$ and ln (r) for the daily discharge series for different embedding dimensions. For a small radius, hardly any points will be captured and for a too large radius nearly all the available points will be included. For the former situation, the plotted relation might curve away from the straight line and for the latter situation the correlation sum becomes 1 constantly, since the radius captures all the available data points and the number of the captured data points remain constant at the maximum available in the data and therefore, the curve depart away from the straight line. Thus, there is a middle segment where the correlation integral starts to deviate from linearity (i.e. scaling region). From the scaling regions, the slopes (i.e. according to Eq. (5), the slopes are equal to the correlation exponents, d) were calculated by least square estimation and the results are given in Fig. 8. As an example, for two different embedding dimensions, ($m = 4$ and $m = 25$) in the Lower Rhine, the scaling regions and estimated correlation exponent values are given in Fig. 9. In all sub-basins, the correlation exponent values increase with an embedding dimension up to a certain value and remain constant for higher values. The saturation of the correlation exponents beyond certain embedding dimension values could be an indication of the existence of deterministic chaos (Table 3). The obtained correlation dimension for Mosel is depicted in Fig. 9c.

The nearest integer above the correlation dimension (i.e. the saturation value of the correlation exponents) provides the minimum embedding dimension for reconstructing the phase-space that is the number of variables (i.e. the number of dominant variables) necessary to model the dynamics of the system (Sivakumar and Jayawardena 2002; Khokhlov et al. 2008). Therefore, the results from this analysis

indicate that the required minimum number of variables to model the system dynamics is equal to 2 for the Mosel, 6 for the Middle Rhine and 4 for the other sub-basins. Since the required minimum numbers of variables to model the system dynamics are relatively low, accurate modelling of the system is possible by considering only a few variables. The results are comparable with those presented in a recent work on low flow indicators in the Rhine basin (Demirel et al. 2011). In that study, the potential evapotranspiration and groundwater levels have been revealed as the most important indicators for the Mosel sub-basin. Further, the correlation between low flows and the precipitation was high for the rainfed sub-basins i.e. Neckar and Main. However, due to the mixed discharge regime, the correlations between low flows and the pre-selected indicators were not high in the MR and LR sub-basins except for the potential evapotranspiration which revealed significant correlations with low flows.

The value of the embedding dimension at which the correlation exponent attains saturation (Fig. 8) is considered to provide the maximum number of variables or upper bound on the dimension of the phase-space sufficient to describe the dynamics of the system (Table 3). However, this embedding dimension can also be estimated by the equation $m \geq 2d + 1$ (Takens 1981). From Table 3, we can conclude that the upper bound for the Mosel is about 5 variables, for the Middle Rhine is about 13 variables and for the other sub-basins it is about 9 variables. It is obvious that different processes are important in these sub-basins leading to low flows. The difference in the minimum and maximum number of variables of the Mosel and Middle Rhine sub-basins from the other sub-basins could be caused from these different processes and the catchment geology. All these local differences can change the basin response and discharge seasonality. Therefore, the discharge signal at the outlet of the basin reveals different number of minimum and maximum variables. In several studies, different results from similar basins were attributed not only to the sub-basins' different inter-dynamic characteristics (i.e. reciprocal relationships) but also to the noise effect-outliers (Sivakumar 2005; Ng et al. 2007). To understand whether over or underestimation exist, rescaled range analysis was applied.

4.3 Rescaled range analysis

In this study, the R/S analysis was applied to make inference about the dynamics of the processes. By using this method, persistency and/or anti-persistency characteristics of the runoff process can be revealed. Also, the R/S analysis allows one to interpret the validity of the correlation dimension results by comparing these results with the Hurst coefficients for the considered sub-basins. The Hurst

coefficients were computed for all discharge series for the period from 1989 to 2008 (Table 4).

The results show that all daily discharge series have values of H greater than 0.5, indicating the presence of long term persistence or long-memory. This result supports the finding about the possible presence of deterministic chaos in the considered sub-basins by rejecting the null hypothesis that the processes are stochastic ($H \neq 0.5$). The strongest long-memory is obtained in the East Alpine sub-

basin and the lowest one in the Mosel sub-basin. The results of the Hurst coefficients as persistence indicators are in line with the size of the arrows showing the strength of the seasonality in the sub-basins (Demirel et al. 2011).

The p-values in parentheses show that at a 95 % confidence interval the null hypothesis of randomness for all of the series can be rejected. In other words, the Hurst coefficients are statistically significant that the long term persistency in the processes cannot be rejected at a 95 %

Fig. 6 Mutual information (MI) functions and their relative changes for discharge series of major sub-basins
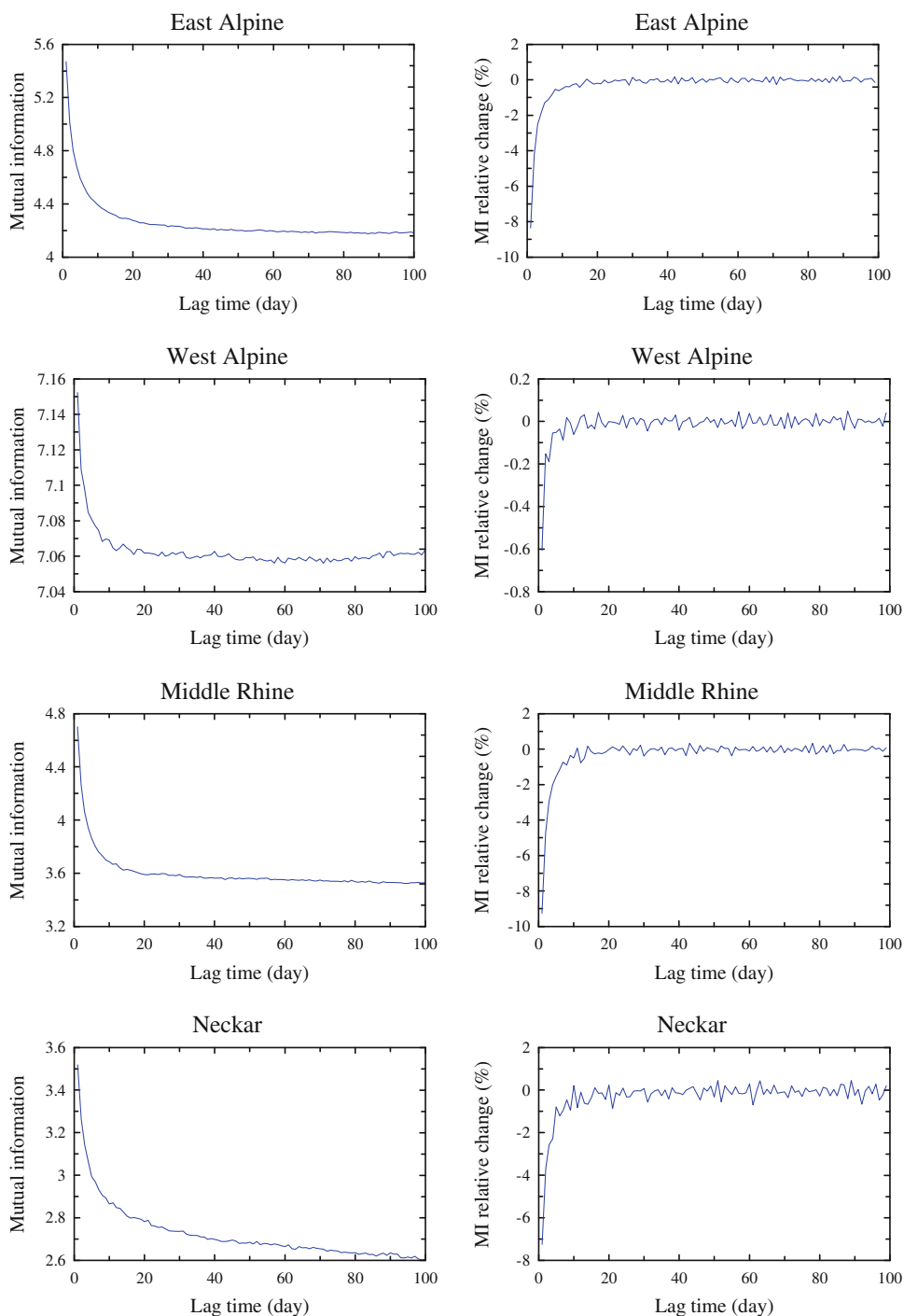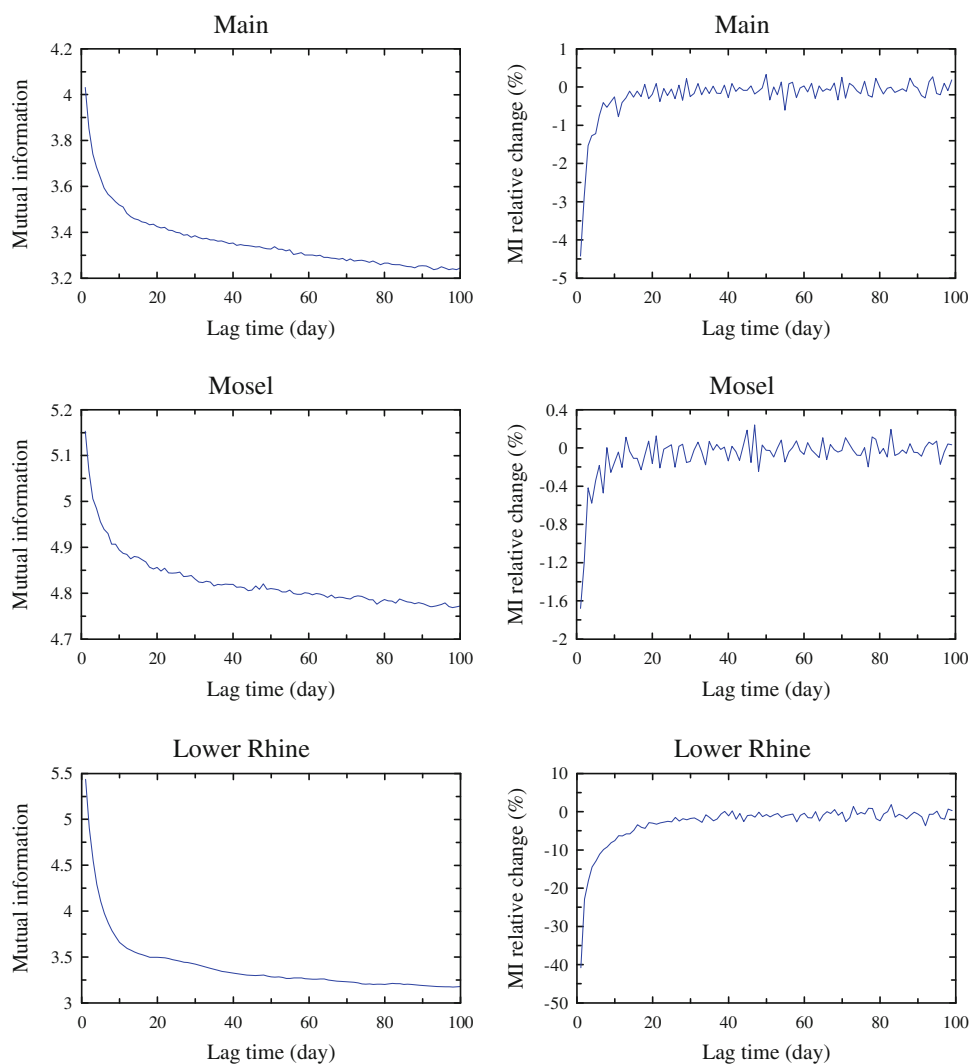
Fig. 6 continued

Table 2 First local minima of mutual information functions of seven major sub-basins

| Sub-basin	East Alpine (I)	West Alpine (II)	Middle Rhine (III)	Neckar (IV)	Main (V)	Mosel (VI)	Lower Rhine (VII)
First local minimum (day)	17	8	11	10	18	8	18

confidence level. The Hurst coefficients are also in agreement with the obtained correlation dimension values except for the Mosel and Middle Rhine sub-basins. This finding could be a clue for underestimation or over estimation. The Hurst coefficient of the Mosel sub-basin is close to the Hurst coefficient values of Neckar and Main sub-basins. Also, the correlation dimension values of the Neckar and Main sub-basins are close to each other. Therefore, the correlation dimension of the Mosel sub-basin should be larger than 1.30 and should be close to the correlation dimension values of Neckar and Main sub-basins. Thus, it can be concluded the correlation dimension value for the Mosel is underestimated. For the Middle Rhine sub-basin, the correlation dimension value should be smaller than

5.95, because its Hurst exponent (0.8396) is close to the Hurst exponent of the East Alpine, the West Alpine and the Lower Rhine sub-basins.

The reason of underestimation and overestimation could be the data quality, noise effects (i.e. errors arising from measurement), outliers, improperly chosen delay time-scaling regions or sub-basins inter-dynamics (intermittency, long-term correlations, seasonality etc.). However, all the data were obtained from the same source (GRDC) and the same procedure was followed in choosing delay time and scaling regions.

The main factor for the underestimation of the MR sub-basin must be from the fact that the annual discharge generation is very low and snow melt water is mixed with

Fig. 7 Correlation integrals as a function of different embedding dimensions (*m*) for daily discharge series
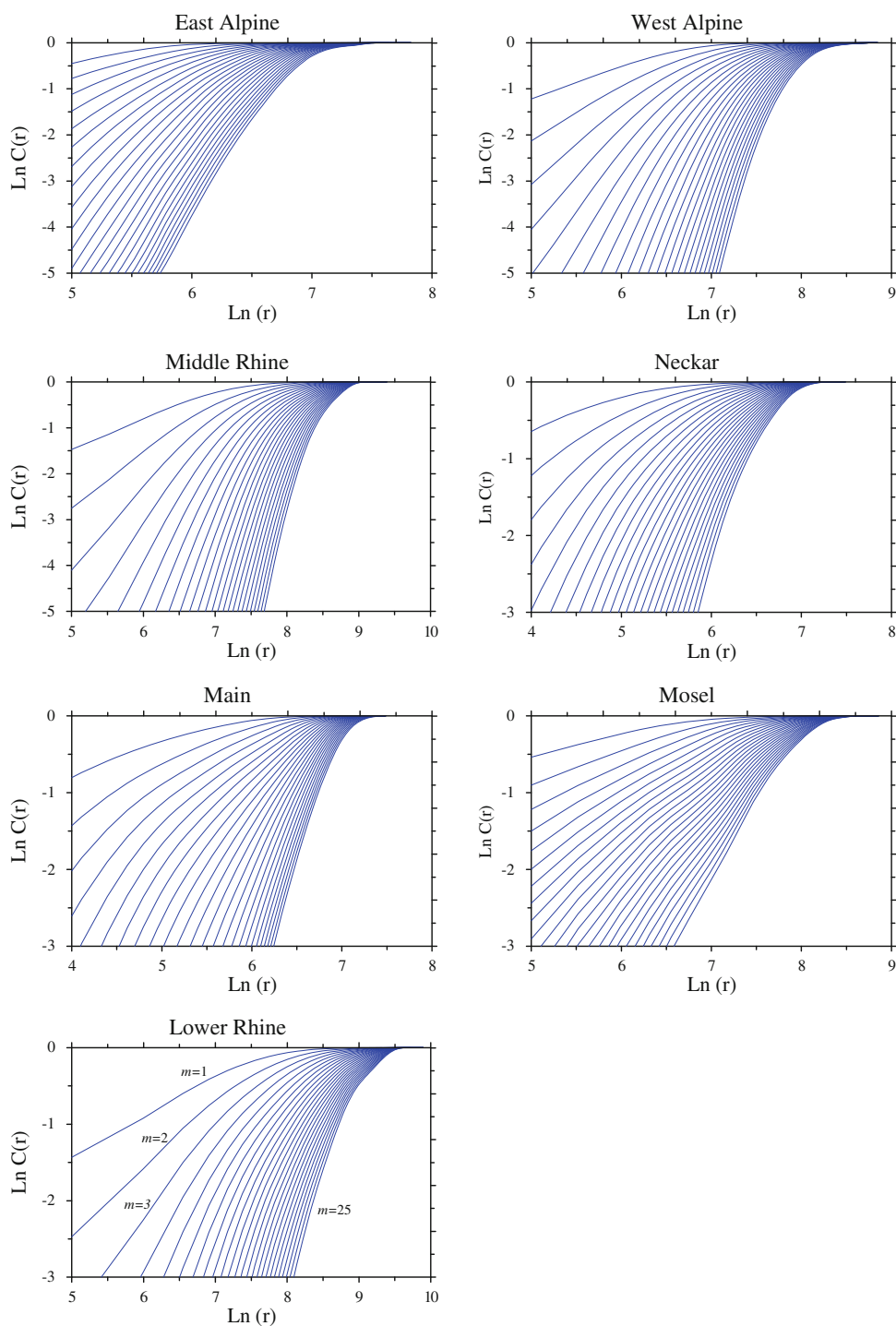
rainfall water in this sub-basin. As a result, the high and low flows are affected by these interacting processes. Secondly, human interventions in the Rhine may cause noise effects hindering the sub-basins flow dynamics. The reader can refer to Sivakumar (2000), Elshorbagy et al. (2002) and Ng et al. (2007) for a detailed discussion of the noise effects on the correlation dimension estimations.

By disregarding underestimated and overestimated dimensions, a minimum number of 4 variables is needed to capture the sub-basins dynamics and a maximum number of 9 variables to fully describe the systems' dynamics. These dominant variable numbers are in agreement with those of (Xu et al. 2010; Sivakumar and Jayawardena 2002; Elshorbagy et al. 2001; Islam and Sivakumar 2002), where
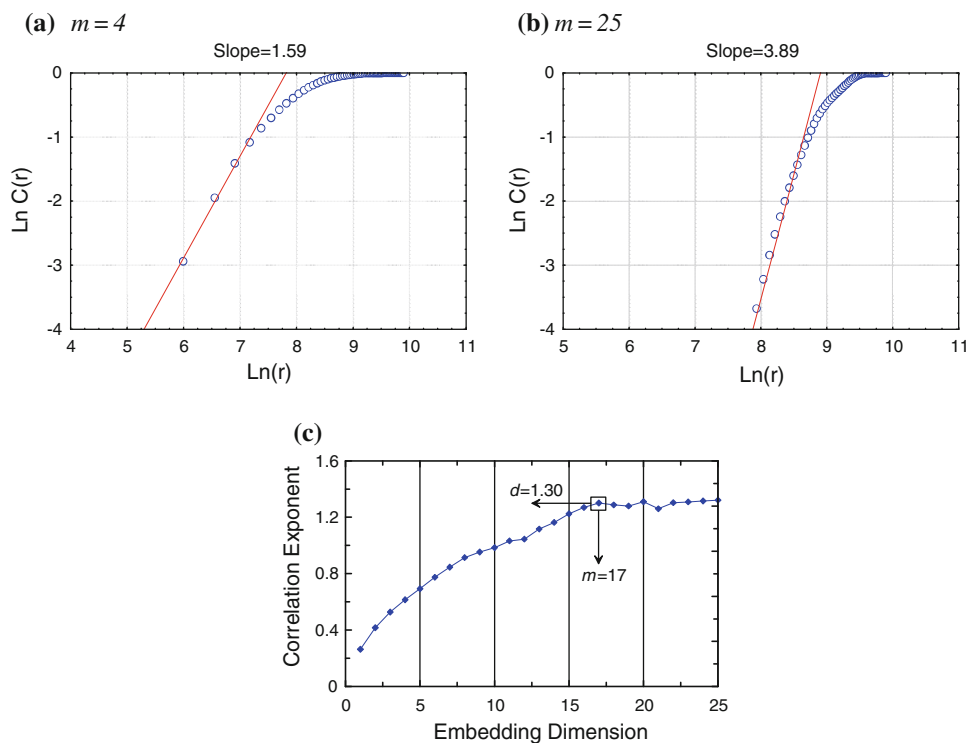
Fig. 8 Correlation exponent versus embedding dimension for the daily discharge series
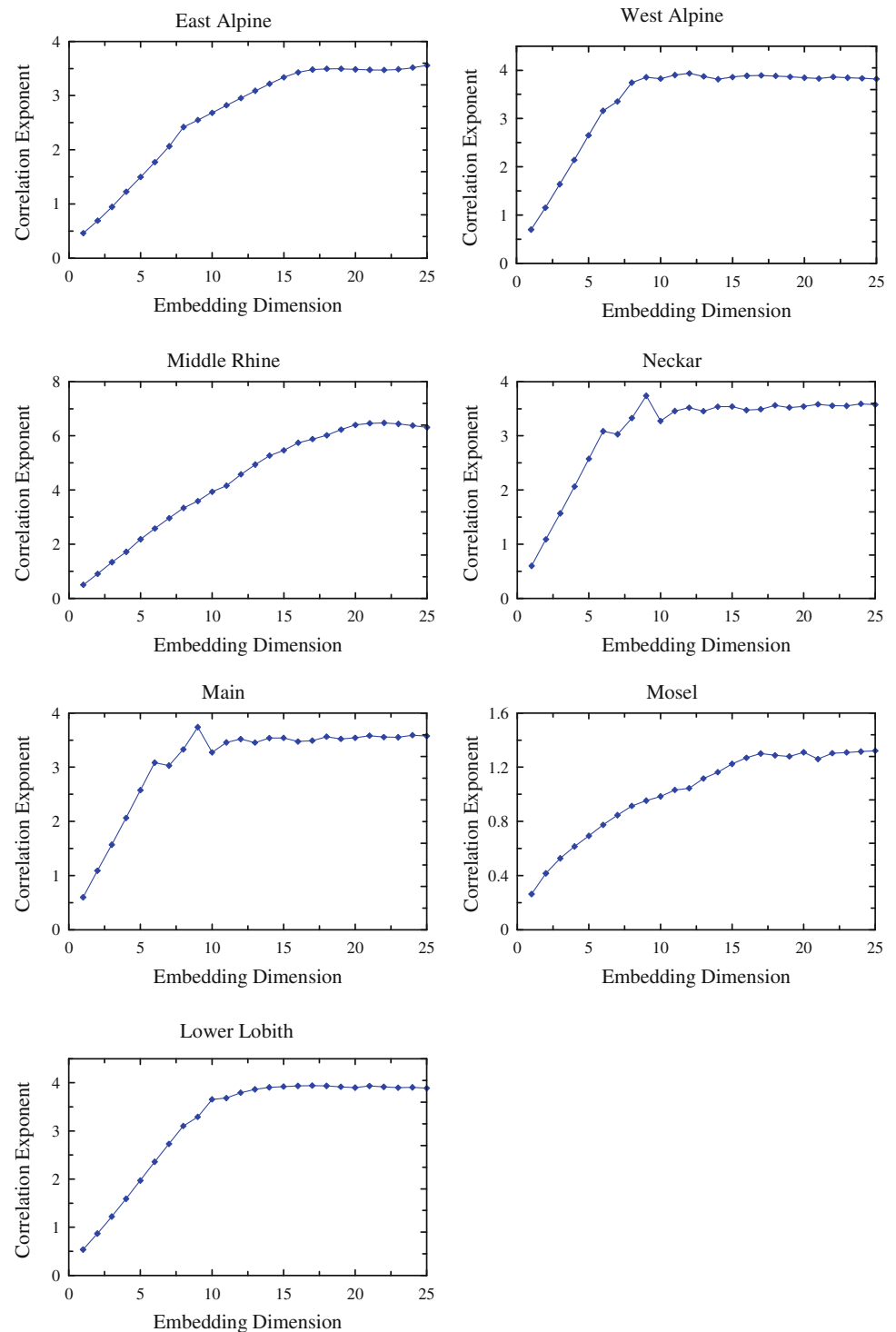
the minimum number of dominant variables of examined streamflows were 2, 4, 3 and 4, and the maximum number of dominant variables were 5, 9, 7 and 9, respectively.

In order to answer the question whether the correlation dimensions vary spatially, and whether summer or winter low flow regimes have impacts on correlation dimension, SR values of the 7 major sub-basins are needed. Thus, the

mean values were estimated from 108 catchments and the results can be seen in The East Alpine and West Alpine regions show a winter low flow regime and the others show summer low flow regimes (Table 5). We could not find any meaningful relationship between the mean SR values and the delay times, correlation dimensions, embedding dimensions and Hurst coefficients. Therefore, it can be

Fig. 9 Illustration of correlation exponent estimation for Lower Rhine for two embedding dimensions, **a** $m = 4$ and **b** $m = 25$, and **c** correlation dimension estimation for Mosel

Table 3 The correlation dimensions for the daily discharges of seven major sub-basins

Sub-basin	East Alpine (I)	West Alpine (II)	Middle Rhine (III)	Neckar (IV)	Main (V)	Mosel (VI)	Lower Rhine (VII)
Embedding dimension from Fig. 8	18	9	18	11	12	17	19
Correlation dimension	3.92	3.86	5.95	3.54	3.52	1.30	3.92
Embedding dimension from $m \geq 2d + 1$	9	9	13	9	9	5	9

Table 4 Hurst exponents and correlation dimensions for the daily discharge series in the seven major sub-basins of the Rhine River

Sub-basin	East Alpine (I)	West Alpine (II)	Middle Rhine (III)	Neckar (IV)	Main (V)	Mosel (VI)	Lower Rhine (VII)
H	0.8546 ($p = 0.000$)	0.8270 ($p = 0.000$)	0.8396 ($p = 0.000$)	0.7865 ($p = 0.000$)	0.7781 ($p = 0.000$)	0.7671 ($p = 0.000$)	0.8245 ($p = 0.000$)
Correlation dimension	3.92	3.86	5.95	3.54	3.52	1.30	3.92

concluded that the correlation dimension values are independent from the dominant low flow regime. This is reasonable since the governing dynamics (i.e. attractors) may show a highly different behaviour because of a high dependency on hydrometeorologic and/or geophysical conditions. These conditions could be thought as initial conditions for the runoff systems that will evolve through time. For instance, for the East Alpine and West Alpine sub-basins there are four dominant variables which govern the system dynamics and in these regions a winter low flow regime is dominant. Also, for the Main, Neckar and Lower Rhine sub-basins there are four dominant variables but the dominant low flow regimes are summer low flow regimes. Therefore, with an equal number of dominant variables, quite different low flow regimes can be found which are highly dependent on basin location. This result also supports the conclusion that a random hydrological phenomenon does not necessarily need to be the outcome of a large

Table 5 The mean SR values of 108 catchments

Sub-basin	East Alpine (I)	West Alpine (II)	Middle Rhine (III)	Neckar (IV)	Main (V)	Mosel (VI)	Lower Rhine (VII)
Mean SR Values	1.8628	1.2539	0.6967	0.5067	0.6007	0.3806	0.6139
Correlation dimension	3.92	3.86	5.95	3.54	3.52	1.30	3.92

number of variables. With the same minimum number of dominant variables (i.e. 4 dominant variables) completely different flow regimes are observed in the East Alpine, West Alpine sub-basins (winter low flow regime) and the Neckar, Main and Lower Rhine sub-basins (summer low flow regime) which could be the result of nonlinear interactions of only a few variables that are interdependent (Sivakumar 2005).

One of the benefits of the present results is that to the authors' knowledge this paper is the first that addresses the potential relationship between the correlation dimension and its spatial variation. Additional research is needed to provide a deeper understanding of this phenomenon.

5 Conclusions

Hydrologic systems are complex and dynamic in nature as their current and future states depend on numerous variables. Dominant low flow regimes in the Rhine basin were investigated using seasonality ratio (SR) indices. Daily discharge data were obtained from 172 stations over 108 catchments. While the Alpine regions are dominated by winter low flows, the other sub-basins are dominated by summer low flows. However, the Middle Rhine is dominated by both summer and winter low flow because of its particular location.

To analyse whether different numbers of variables are acting in generating different low flow regimes, nonlinear analyses were applied. For the correlation dimension analysis, proper delay times were determined with the mutual information method. The correlation integrals were computed with the Grassberger–Procaccia algorithm for embedding dimensions, varying from 1 to 25, and the correlation exponents were determined from the scaling regions of these plots. Except for the Middle Rhine and Mosel, the minimum and maximum required numbers of variables were found to be 4 and 9, respectively to properly capture the dynamics of the considered sub-basins. For the Mosel and Middle Rhine sub-basins, the required minimum number of variables is 2 and 6, and the maximum number of variables is 5 and 13, respectively. By applying the rescaled range analysis (R/S analysis), which is a powerful tool to analyse long-term dependence, the validities of the estimated correlation dimensions were investigated. The estimated R/S statistics are all above 0.5, indicating that persistent long-term memory characteristics (i.e. non-randomness) are present in the runoff processes.

The estimated correlation dimension values and the Hurst exponents were compared with each other and it was found that the correlation dimension value of the Middle Rhine sub-basin was overestimated and the estimated correlation dimension value of the Mosel sub-basin was underestimated. The spatial variability of the correlation dimensions was compared with the mean values of the SR index for the seven sub-basins, however no meaningful relationship was obtained. The minimum numbers of required variables of the East Alpine, West Alpine, Neckar, Main and Lower Rhine sub-basins to capture the underlying dynamics are all equal to 4 but the observed low flow regimes are different (i.e. winter low flow and summer low flow regimes) within the considered sub-basins. This result supports the conclusion that a few interrelated nonlinear variables could yield completely different behaviour (i.e. dominant low flow regime). In future research, the authors will investigate the dominant variables themselves which are interacting in generating low flow regimes.

References

Abarbanel HDI, Brown R, Kadtke JB (1990) Prediction in chaotic nonlinear systems: methods for time series with broadband Fourier spectra. Phys Rev A 41:1782–1807

Belz JU, Frauenfelber-Kääb R (2007) Das Abflussregime des Rheins und seiner Nebenflüsse im 20. Jahrhundert: Analyse, Veränderungen, Trends (in German). vol 9789070980337. KHR/CHR Lelystad

Booij MJ (2003) Determination and integration of appropriate spatial scales for river basin modelling. Hydrol Processes 17(13):2581–2598

Couillard M, Davison M (2005) A comment on measuring the Hurst exponent of financial time series. Phys A 348:404–418

Demirel MC, Booij MJ, Hoekstra AY (2011) Identification of appropriate lags and temporal resolutions of low flow indicators in the Rhine River in forecasting low flows with different lead times. Hydrol Processes. doi:10.1002/hyp.9402

Elshorbagy A, Panu US, Simonovic SP (2001) Analysis of cross-correlated chaotic streamflows. Hydrol Sci J 46(5):781–793

Elshorbagy A, Simonovic SP, Panu US (2002) Noise reduction in chaotic hydrological time series: facts and doubts. J Hydrol 256:147–165

Eng K, Milly PCD (2007) Relating low-flow characteristics to the base flow recession time constant at partial record stream gauges. Water Resour Res 43(1):W01201. doi:10.1029/2006wr005293

Frazer AM, Swinney HL (1986) Independent coordinates for strange attractors from mutual information. Phys Rev A 33(2): 1134–1140

Grassberger P, Procaccia I (1983) Characterization of strange attractors. Phys Rev Lett 50(5):346–349

Grayson R, Blöschl G (2001) Summary of pattern comparison and concluding remarks. In: Grayson R, Blöschl G (eds) Spatial patterns in catchment hydrology: observations and modeling. Cambridge University Press, Cambridge, pp 355–396

Huisman P, De Jong J, Wieriks K (2000) Transboundary cooperation in shared river basins: experiences from the Rhine, Meuse and North Sea. Water Policy 2(1):83–97

Hurkmans R, De Moel H, Aerts J, Troch PA (2008) Water balance versus land surface model in the simulation of Rhine river discharges. Water Resour Res 44(1):W01418. doi:10.1029/2007wr006168

Hurst HE (1951) Long-term storage capacity of reservoirs. Trans Am Soc Civil Eng 116:770–808

Islam MN, Sivakumar B (2002) Characterization and prediction of runoff dynamics: a nonlinear dynamical view. Adv Water Resour 25(2):179–190

Jayawardena AW, Lai F (1994) Analysis and prediction of chaos in rainfall and stream flow time series. J Hydrol 153(1–4):23–52

Kavvas ML, Cote LJ, Delleur JW (1977) Time resolution of the hydrologic time-series models. J Hydrol 32(3–4):347–361

Khokhlov V, Glushkov A, Loboda N, Serbov N, Zhurbenko K (2008) Signatures of low-dimensional chaos in hourly water level measurements at coastal site of Mariupol, Ukraine. Stoch Environ Res Risk Assess 22(6):777–787

Laaha G, Blöschl G (2006) Seasonality indices for regionalizing low flows. Hydrol Processes 20(18):3851–3878. doi:10.1002/hyp.6161

Laaha G, Blöschl G (2007) A national low flow estimation procedure for Austria. Hydrol Sci J 52(4):625–644

Lisi F, Villi V (2001) Chaotic forecasting of discharge time series: a case study. J Am Water Resour Assoc 37(2):271–279

Markovic D, Koch M (2005) Wavelet and scaling analysis of monthly precipitation extremes in Germany in the 20th century: interannual to interdecadal oscillations and the North Atlantic Oscillation influence. Water Resour Res 41. doi:10.1029/2004wr003843

Middelkoop H, Van Haselen COG (1999) Twice a river. Rhine and Meuse in The Netherlands. RIZA report 99.003. RIZA, Arnhem

Ng WW, Panu US, Lennox WC (2007) Chaos based analytical techniques for daily extreme hydrological observations. J Hydrol 342:17–41

Packard NH, Crutchfield JP, Farmer JD, Shaw RS (1980) Geometry from a time series. Phys Rev Lett 45(9):712–716

Reggiani P, Renner M, Weerts AH, van Gelder P (2009) Uncertainty assessment via Bayesian revision of ensemble streamflow predictions in the operational river Rhine forecasting system. Water Resour Res 45(2):W02428. doi:10.1029/2007WR006758

Regonda SK, Sivakumar B, Jain A (2004) Temporal scaling in river flow: can it be chaotic? Hydrol Sci J 49(3):373–385

Renner M, Werner MGF, Rademacher S, Sprokkereef E (2009) Verification of ensemble flow forecasts for the River Rhine. J Hydrol 376(3–4):463–475

Ruelle D (1981) Chemical kinetics and differentiable dynamical systems. In: Nonlinear phenomena in chemical dynamics. Springer, Berlin

Sahoo GB, Schladow SG, Reuter JE (2009) Forecasting stream water temperature using regression analysis, artificial neural network, and chaotic non-linear dynamic models. J Hydrol 378:325–342

Schaefli B, Maraun D, Holschneider M (2007) What drives high flow events in the Swiss Alps? Recent developments in wavelet spectral analysis and their application to hydrology. Adv Water Resour 30(12):2511–2525

Sivakumar B (2000) Chaos theory in hydrology: important issues and interpretations. J Hydrol 227(1–4):1–20

Sivakumar B (2001) Rainfall dynamics at different temporal scales: a chaotic perspective. Hydrol Earth Syst Sci 5(4):645–652

Sivakumar B (2004a) Chaos theory in geophysics: past, present and future. Chaos Solitons Fract 19:441–462

Sivakumar B (2004b) Dominant processes concept in hydrology: moving forward. Hydrol Process 18:2349–2353

Sivakumar B (2005) Correlation dimension estimation of hydrological series and data size requirement: myth and reality. Hydrol Sci J 50(4):591–603

Sivakumar B (2007) Nonlinear determinism in river flow: prediction as a possible indicator. Earth Surf Proc Land 32:969–979. doi:10.1002/esp.1462

Sivakumar B (2008) Dominant processes concept, model simplification and classification framework in catchment hydrology. Stoch Environ Res Risk Assess 22(6):737–748

Sivakumar B, Jayawardena AW (2002) An investigation of the presence of low-dimensional chaotic behaviour in the sediment transport phenomenon. Hydrol Sci J 47(3):405–416

Sivapalan M (2003) Process complexity at hillslope scale, process simplicity at the watershed scale: is there a connection? Hydrol Processes 17(5):1037–1041

Smakhtin VU (2001) Low flow hydrology: a review. J Hydrol 240(3–4):147–186

Takens F (1981) Detecting strange attractors in turbulence. In: Rand DA, Jung LS (eds) Dynamical systems and turbulence, lecture notes in mathematics, vol 898. Springer, Berlin, pp 366–381

Te Linde AH, Aerts J, Hurkmans R, Eberle M (2008) Comparing model performance of two rainfall-runoff models in the Rhine basin using different atmospheric forcing data sets. Hydrol Earth Syst Sci 12(3):943–957

Te Linde AH, Aerts JCJH, Bakker AMR, Kwadijk JCJ (2010) Simulating low-probability peak discharges for the Rhine basin using resampled climate modeling data. Water Resour Res 46(3):W03512. doi:10.1029/2009wr007707

Wang W, Vrijling JK, Van Gelder PHAJM, Ma J (2006) Testing for nonlinearity of streamflow processes at different timescales. J Hydrol 322:247–268

Wilcox BP, Seyfried MS, Matison TH (1991) Searching for chaotic dynamics in snowmelt runoff. Water Resour Res 27. doi:10.1029/91wr00225

Williams GP (1997) Chaos theory tamed. Joseph Henry Press, Washington

Woods R (2002) Seeing catchments with new eyes. Hydrol Processes 16:1111–1113

Xu J, Li W, Ji M, Lu F, Dong S (2010) A comprehensive approach to characterization of the nonlinearity of runoff in the headwaters of the Tarim River, western China. Hydrol Process 18:136–146. doi:10.1002/hyp.7484

Yu B, Huang C, Liu Z, Wang H, Wang L (2011) A chaotic analysis on air pollution index change over past 10 years in Lanzhou, northwest China. Stoch Env Res Risk Assess 25(5):643–653. doi:10.1007/s00477-011-0471-y

Zaidman MD, Rees HG, Young AR (2001) Spatio-temporal development of streamflow droughts in north-west Europe. Hydrol Earth Syst Sci 6(4):733–751