Journal of Hydrology (2007) 336, 186–198

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jhydrol

Temporal scaling comparison of real hydrological data and model runoff records

V. Livina [a,*], Z. Kizner [b], P. Braun [c], T. Molnar [c], A. Bunde [d], S. Havlin [b]

[a] School of Environmental Sciences, University of East Anglia, Norwich, UK
[b] Department of Physics, Bar-Ilan University, Ramat-Gan 52900, Israel
[c] Bayerisches Landesamt für Wasserwirtschaft, Lazarettstrasse 67, D-80636 München, Germany
[d] Institut für Theoretische Physik III, Justus-Liebig-Universität Giessen, Heinrich-Buff-Ring 16, 35392 Giessen, Germany

Received 16 February 2005; received in revised form 21 December 2006; accepted 2 January 2007

KEYWORDS
ASGi model;
Time series analysis;
Detrended fluctuation analysis;
Multifractality;
Nonlinear volatility;
Statistical evaluation of models

Summary We show that the scaling properties of river runoff records represent a useful tool for evaluating precipitation-runoff models that are widely used in hydrology for assessment of the water balance in a given river catchment. In this respect, it is important that the model maps the processes that control the water balance. The main field of application is therefore water management in a given area over a long time scale (at least several years). Here, we compare the temporal scaling properties of the runoff of three Bavarian rivers (Naab, Regnitz, and Vils) with the corresponding ASGi model records. In the evaluation, we use: (i) detrended fluctuation analysis (DFA); (ii) multifractal analysis; (iii) periodic volatility analysis; and (iv) long-term volatility analysis. Our study generally shows close similarity between real and simulated data for the main statistical parameters (e.g., correlation and multifractal exponents). Therefore, the ASGi model output seems to adequately describe real basin processes and might be useful for hydrological purposes, such as a posteriori estimation of water balance in a river catchment.

Introduction

Over the last few decades, many hydrological models have been developed, allowing to simulate records of river flux for short- and long-term intervals (Beven and Moore, 1993; Beven, 1998; Dutta et al., 2000; Beven, 2004). The modelling was spurred in 1980–1990s, when computers enabled processing a large amount of geophysical data and solving numerically complex fluid dynamics equations.

There are different approaches in hydrological modelling, and their intercomparison has been discussed intensively (see Refsgaard and Knudsen, 1996; also Marshall et al., 2005). The linear stochastic modelling, where autoregressive processes are implemented (e.g., fractional ARIMA, Montanari et al., 2000), by definition, is unable to reproduce nonlinear features of river runoff series and will not be considered in this paper. Another approach is deterministic

* Corresponding author. Tel.: +44 7896330375.
E-mail address: v.livina@uea.ac.uk (V. Livina).

0022-1694/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2007.01.014

modelling. Distributed physically based models are developed on partial differential equations governing hydrological systems (TOPMODEL, MIKE SHE, WATFLOOD, IISDHM, and others; for review see Beven, 2004). Such models use field measurements and remote sensing records as input data and can be exploited for water management. Third, there are parametric conceptual models, where the systems of deterministic equations are supplemented by some empirical relations (Jones, 1997).

The distributed physically-based models incorporate streamflow, precipitations, evaporation, averaged stochastic groundwater flow and spatial properties of the catchment, e.g., elevation data, topographic indices, glacier grid, land use, and soil types. The spatial scaling parameters are different for different river subcatchments, and only in exceptional cases is it possible to verify them by means of explicit variations of the grid width in use. Another problem which is difficult to resolve is the exact initial and boundary conditions for solving partial differential equations, and those conditions are, to a great extent, unknown, being a stochastic component of the system. This explains why many models fail in reproducing and predicting the patterns of real time series (Wood et al., 1988; Blöschl, 2001). For instance, the models of the atmosphere/ocean system (AOGCM), which are based on the same partial differential equations of fluid dynamics, demonstrated lack of long-term memory in temperature records, compared to real time series, as has been recently shown using modern tools of statistical physics (Govindan et al., 2002; Vyushin et al., 2004).

A contemporary parametric conceptual model tunes a detereministic kernel by implementing some empirical relations, and this is supposed to bring the output records closer to real observations. Although the conceptual models require additional calibration to obtain proper values of the parameters, they allow to solve the problem of the lack of initialisation data in physically-based models. Also, the conceptual models provide a better adjustment to a particular catchment (Refsgaard and Knudsen, 1996). A typical example of a parametric model which combines advances of distributed models with empirical balance, is the conceptual grid-based ASGi model[1] (Braun et al., 1998; Becker and Braun, 1999).

The model data comparison is traditionally based on quantitative criteria of model accuracy (for instance, the Nash–Sutcliffe coefficient). However, there are recent attempts to apply more advanced techniques, like Bayesian approach (Marshall et al., 2005). Still, this analysis does not take into account nonlinear characteristics of the data which are important for understanding of the process dynamics.

Here, we perform novel validation of a hydrological model by applying techniques that have been recently developed in statistical physics, such as Detrended Fluctuation Analysis (the major paper by (Peng et al., 1994), with multiple applications in many fields of the modern physics in the last 10 years — see e.g., Santhanam et al., 2006; Kantelhardt et al., 2006), nonlinear volatility analysis (Ashkenazy et al., 2001; Livina et al., 2003a), and multifractal analysis (Muzy et al., 1991; Arneodo et al., 1995; Ivanov et al., 1999; Struzik, 2000; Kantelhardt et al., 2002; Enescu et al., 2006). These techniques allow to quantify essential statistical features (correlations, scaling and nonlinearity) in time series and thus provide understanding of the underlying dynamical processes in the system.

The ASGi model and the rivers studied

We study the daily discharge time series generated by the ASGi model for three Bavarian rivers. Implementing empirical hydrological relations, the model, in contrast with 'pure' PDE-based models, does not map the physical processes by means of exact conservation laws, but establishes an empirical balance between sinks and sources in the grid elements. It is based on the water balance simulation model, CH (WaSiM-ETH, Schulla, 2000; Jasper et al., 2002; Gurtz et al., 2003; Verbunt et al., 2003).

The WaSiM-ETH model has been used in more than twenty institutions in Switzerland, Germany, Austria, Iceland and Slovakia. A more specialised ASGi model has been used in three German institutions. Due to the model layout which is well adapted to basins of European rivers, the model can be useful for many European users.

The model consists of the following main components (Fig. 1): (i) cell interpolation and adjustment of input meteorological data; (ii) estimation of evapotranspiration and, if necessary, irrigation; (iii) snow-glacier submodel; (iv) interception submodel; (v) infiltration submodel; (vi) soil submodel; (vii) groundwater submodel; and (viii) discharge routing. Coupling different parts of the model and taking into account losses due to evapotranspiration, the main sources of river discharge (surface runoff, interflow and baseflow) are integrated. The baseflow is generated for the entire subbasin as an average value. Interflow is generated for each grid cell separately and then averaged over the area. The surface runoff is routed to subbasin outlet, using a subdivision of the basin into flow time zones, which are zones of equal flow times for surface runoff to reach the subbasin outlet. Because of the local soil type, the ASGi model uses the first version of WaSiM, with TOPMODEL groundwater modelling, instead of the second version, implementing Richard's equation for groundwater. Sufficient adequacy of the model can be illustrated directly by means of visual comparison of real and simulated fluxes in hydrographs (Fig. 2). The calibration of the model was carried on by trial and error estimation using Nash–Sutcliffe efficiency criterion, and the flow was found to be reasonably well predicted (see 'Discussion').

The Naab river basin area (about 5500 km^2) is located mainly in the Oberpfalz region in Bavaria, and partly in the Czech Republic. The Naab is a left tributary of the Danube river, which it joins near Regensburg, in the Danube's upper basin. The catchment is placed in the Fränkische Alb (Eastern continuation of the Schwäbische Alb in Baden-Württemberg), a mountain chain up to 1000 m high (500 m at average) in Bavaria. At its origin, the Naab is a mountain stream, and in the lower part it is a slow plain stream. Therefore, the catchment is highly spatially variable and difficult for hydrological modelling and forecasting. Geologically, the basin is a compound of sand, karst, and slate grit.

[1] Kontinuierlicher Abfluss und Stofftransport-Integrierte Modellierung unter Nutzung von Geoinformationssystemen, Germany.

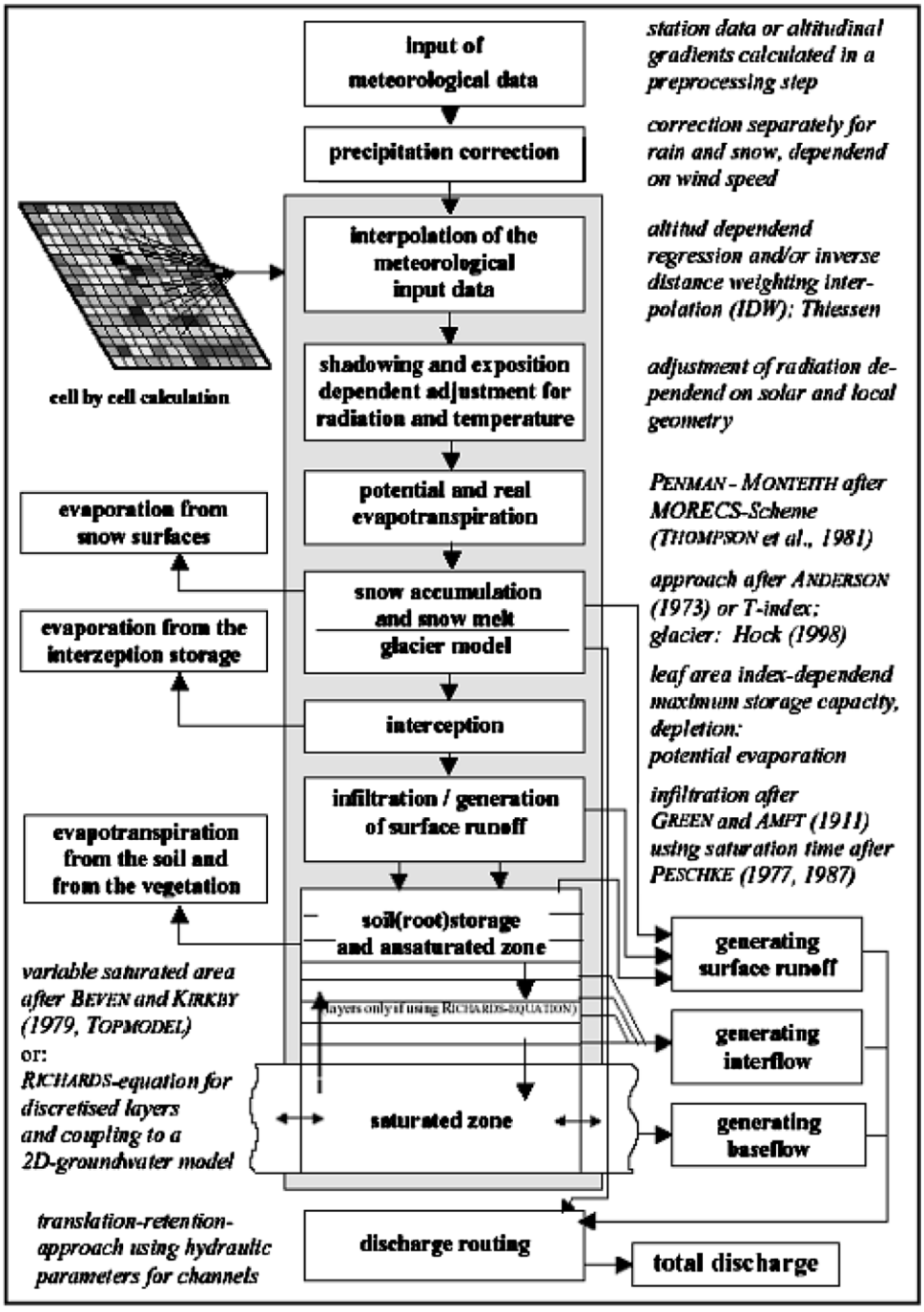

Figure 1 Schematic of the WaSiM-ETH model.

The catchment of the Danube's tributary Vils lies in the northern part of Bavaria, the Oberpfalz. This area is characterized by complex geological conditions (Jura). The main difficulty in this context is existence of karst areas, which shows some unusual water storage properties. In order to consider this specific behaviour, it was necessary to introduce some empirical corrections in the model.

The Regnitz river is a left tributary of the Main river in the Lower Bavaria. It is formed by the confluence of the Regnitz and Pegnitz and flows into the Main below Bamberg. The catchment area is about 7000 km^2 and varies in height between 240 and 650 m. The geology is comprised of slate, clay, and malm-soil types which are of different permeability and storage capacity.

The maps of precipitation, evaporation and runoff of the part of Bavaria where the considered river basin are placed are shown in Figs. 3–5. The hydrological characteristics of the basins are summarised in Table 1.

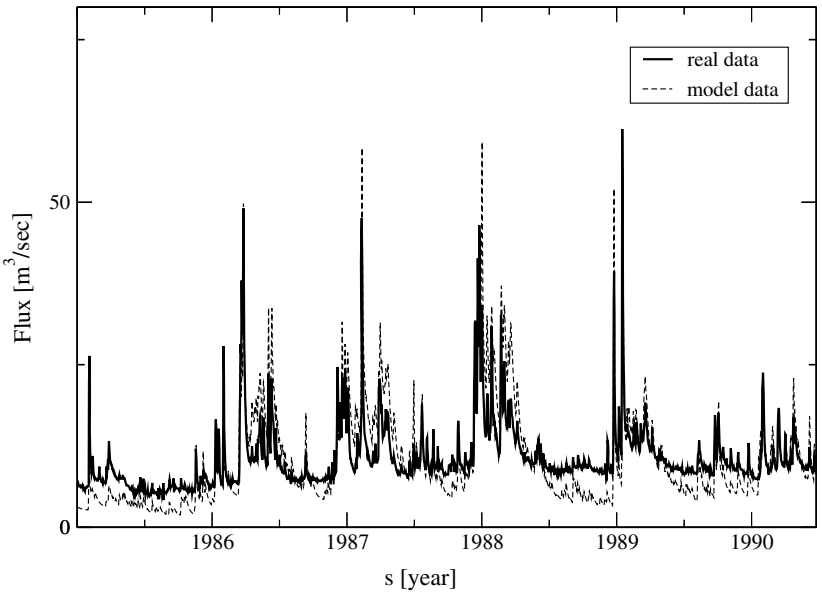

Figure 2 Five-year cycles of real and simulated flux records for Vils river; model data demonstrate similar pattern with slightly different magnitudes of the fluctuations.

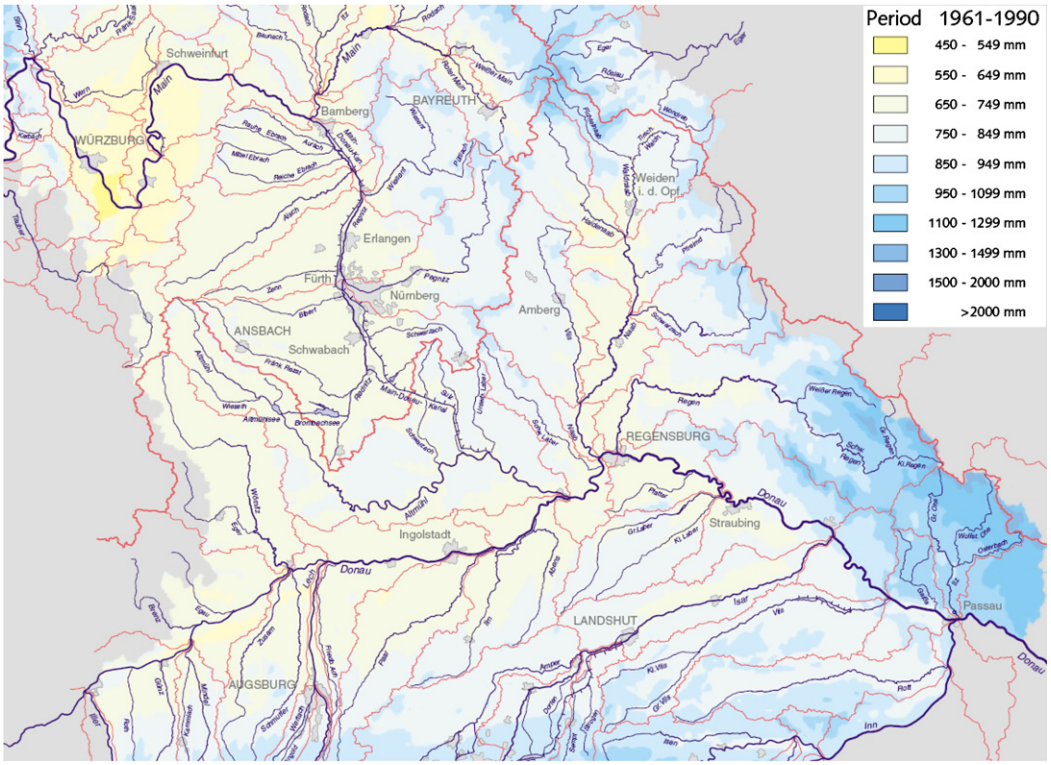

Figure 3 Yearly averaged precipitation map of the part of Bavaria where the basins are placed.

Detrended fluctuation analysis

Methodology

In recent years, the detrended fluctuation analysis (DFA) method has become a widely used tool for the study of statistical scaling properties of nonstationary time series (Peng et al., 1994). It has been applied successfully, e.g., to DNA sequences (Buldyrev et al., 1995), heart-rate dynamics (Peng et al., 1995; Bunde et al., 2000; Ashkenazy et al., 2001), to econometric time series (Mantegna and Stanley, 1999; Matia et al., 2003), climate dynamics (Koscielny-Bunde et al., 1998; Govindan et al., 2002; Kantelhardt et al., 2006), and other fields of the modern physics. The

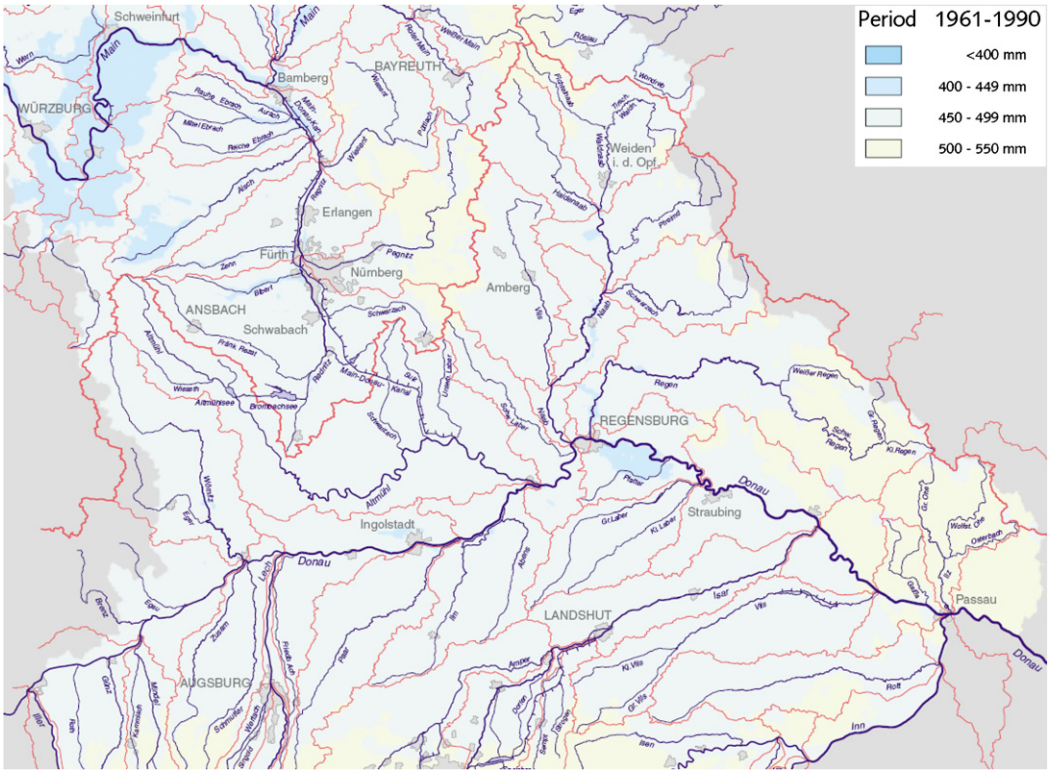

Figure 4 Yearly averaged evaporation map of the part of Bavaria where the basins are placed.

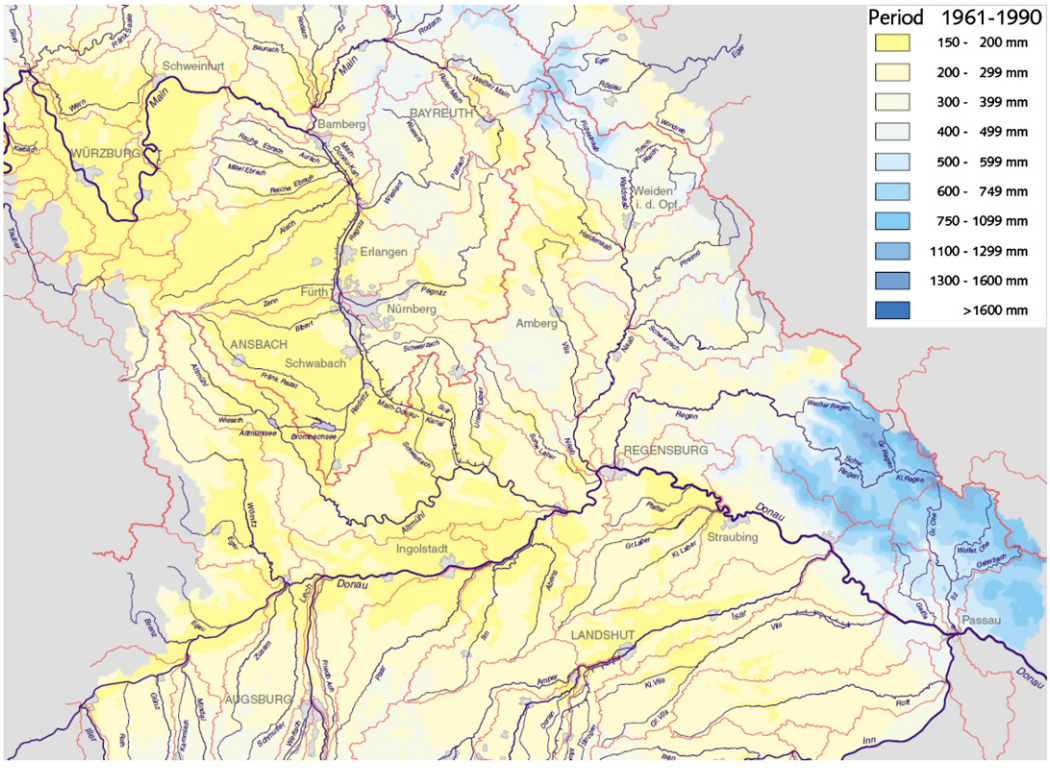

Figure 5 Yearly averaged runoff map of the part of Bavaria where the basins are placed.

essence of the method is studying the properties of the fluctuations in the data after removing the seasonal trend and other nonstationarities, and this is achieved by multiple averaging of the mean-root square variance over windows of variable length in the integral of the series. The method provides robust results on two-point correlations, as com-

Table 1 Hydrological characteristics of the three rivers (1961–1990)

River	Station	Elevation		Area (km^2)	Precipitation (mm/year)	Evaporation (mm/year)	Runoff		
		Minimum (m)	Maximum (m)				(1/(s km^2))	Average high (m^3/s)	Average low (m^3/s)
Naab	Heitzenhofen	350	1000	5246	700–900	400–700	2.18	280	18.9
Vils	Schmidmühlen	350	620	756	550–900	450–570	–	130	–
Regniz	Pettstadt	240	645	6999	650–1000	450–600	4.8–12.7	287	20.7

pared to conventional power-spectrum and auto-correlation function analyses. Moreover, in case of highly nonstationary data (in particular, hydrological), the auto-correlation function is incapable to reveal the correlations properly; in the power spectrum, the analysis is often deviated by the noisiness of the power amplitude. The DFA is devoid of these weakness, and this makes it an efficient tool for studying the statistical properties of the time series of complex systems.

Since we are concerned here with studying flux fluctuations, before applying statistical scaling techniques we eliminate the seasonal periodicities in the data by removing the seasonal average (deseasoning). In the framework of the *k*-order DFA method (DFAk), we integrate the fluctuation series and divide the range of definition of the obtained profile function into windows of size *s*. Next, within each window, we calculate the best polynomial fit and evaluate the difference between the polynomial and profile function. We average the obtained values over all windows, and afterwards we repeat the procedure for different window scales *s*, deriving $F(s)$:

$$F^2(s) = \frac{1}{K}\sum_{v=1}^{K}\frac{1}{s}\sum_{(v-1)s+1}^{vs}[Y_v(i) - p_v^k(i)]^2,$$

where $Y_n = \sum_{i=1}^{n}\Delta X_i$ is the profile function of the deseasoned time series ΔX_i, $p_v^k(i)$ is the best polynomial fit of order k, $K = 2N/s$ is the number of non-overlapping windows of size s (indexed v), and N is the length of the time series.

When the series satisfies a scaling law, we observe a power-law behaviour of the fluctuation function:

$$F(s) \sim s^{\alpha},$$

where α is the scaling exponent. For uncorrelated records, $\alpha = 0.5$, while for long-term correlated (persistent) records, $\alpha > 0.5$, where the auto-correlation function decays as $C(s) \sim s^{-\gamma}$ with an increasing time lag s, and α is related to γ by $\alpha = 1 - \frac{\gamma}{2}$.

Results

We analysed daily runoff of the Naab (26 years), the Regnitz (30 years) and the Vils (26 years) rivers and corresponding

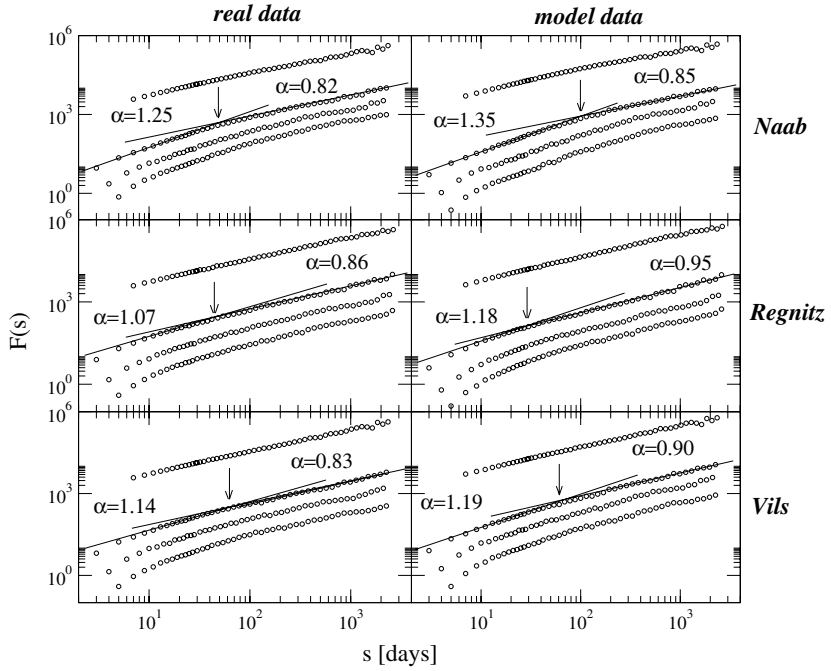

Figure 6 Detrended fluctuation analysis of orders 0 (top curve) and 1–3 (three curves below); arrows denote average points of crossovers, where the scaling exponents are changing; for all the three rivers, simulated data reproduce crossovers (slightly shifted for Naab river) and demonstrate close correlation exponents.

records simulated by the ASGi model for the same period of time. Applying DFA analysis to real and simulated series, we observed high correlations ($\alpha \sim 1.1$) for a period of about 30—70 days and a relatively smooth change to a smaller exponent ($\alpha \sim 0.8$) for longer periods of time. Fig. 6 shows the DFA0-3 results for real and simulated data. The model reproduces both the large correlation exponent in the short-term range and the lower exponent in the asymptotic time range. The crossover in the simulated Naab data is only slightly shifted towards larger time scales, and the short-term correlation exponent is slightly overestimated (1.35, as compared to 1.25 in the real data). For the Regniz and the Vils rivers, the model reproduces scaling behaviour in a nearly perfect manner.

Because of the DFA0 restriction on the exponent value (DFA0 exponent cannot be higher than one), in the short-term regime, DFA0 underestimates the correlation exponent α (as compared to DFA1-3 which provide values higher than one). In long-term regime, for all time series, DFA0 curves become straight and parallel to DFA1-3, with exponents close to those of DFA1. This indicates the absence of a linear trend in both real and simulated data, despite the generally claimed global warming trend. We suggest that this is because the data were recorded in the 1960—1990s under almost the same industrial conditions in the area, and, therefore, the records were not significantly influenced by the global warming trend.

Conventional hydrological statistics and its comparison with the scaling analysis

Methodology

In this section, three statistical criteria conventionally used in hydrology are described. We also discuss the advantages of the scaling analysis and its difference from the conventional methods. The criteria are the following (see, for instance, Hogue et al., 2006):

$$\text{Daily Root Mean Square} \quad \sqrt{1/n\sum_{i=1}^{N}(q_{\mathrm{m}}^{i}-q_{\mathrm{o}}^{i})^{2}},$$

$$\text{Nash–Sutcliffe efficiency} \quad 1-\left(\sum_{i=1}^{N}(q_{\mathrm{m}}^{i}-q_{\mathrm{o}}^{i})^{2}\Big/\sum_{i=1}^{N}(q_{\mathrm{o}}^{i}-\bar{q})^{2}\right),$$

$$\text{Percent bias(\%bias)} \quad \left[\sum_{i=1}^{N}(q_{\mathrm{m}}^{i}-q_{\mathrm{o}}^{i})\Big/\sum_{i=1}^{N}(q_{\mathrm{o}}^{i})\right]\times 100,$$

where q_{o} and q_{m} are the observed and modelled fluxes, respectively, $i = 1, \ldots, N$ is time index (daily flux), $\bar{q}$ is the mean value of the series q. It is easy to see that for two identical time series DRMS = 0, NS = 1, and %bias = 0.

Results

In Fig. 7, we plot the values of these traditional hydrological characteristics for the rivers under consideration. As was shown in hydrograph of the Vils river (Fig. 2), the patterns of observed and modelled data are similar, but the values of the traditional criteria do not show this. Since the criteria provide point-wise comparison of the series, their values are very fragile towards minimal changes in temporal organization of the data array.

This can be explained by a simple experiment. Consider, for example, the flux of the Naab river. By shift of the data three time steps back, we obtain new series with just the same pattern. However, the three criteria applied for comparison of the new series with the initial one provide estimates that differ dramatically from what could be expected for identical series: NS = 0.65 (instead of 1), DRMS = 4.52 (instead of 0), %bias = 0.03 (instead of 0). If we look at the values of the DFA correlation exponent α, it does not change at all, because the DFA algorithm takes into account the temporal organization of the data and, in this example, recognizes that the two series have the same scaling properties. In fact, here we compared a series with itself, but the conventional criteria failed. In another exper-

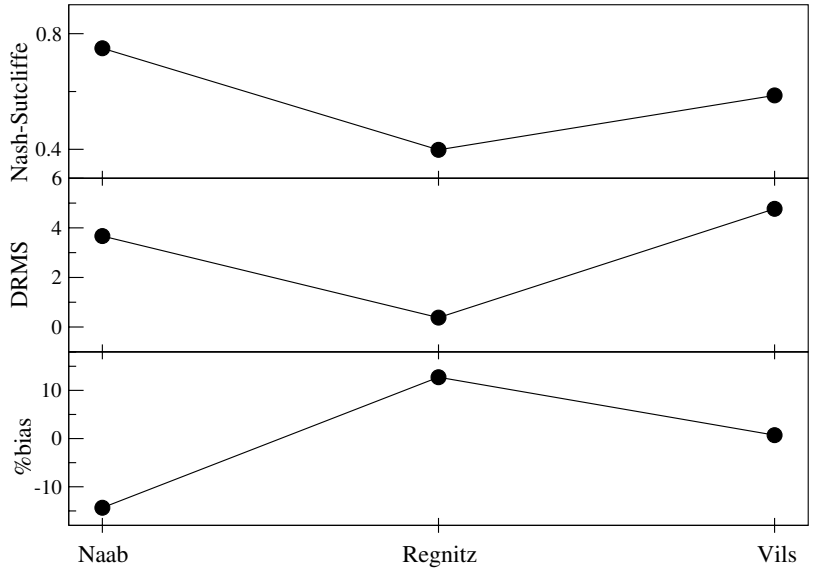

Figure 7 Conventional hydrological criteria for three rivers: Nash—Sutcliffe, daily root mean square, and %bias.

iment, introduction of just one spike in the series (which can be caused by a measurement error), changes NS significantly, and two practically identical series are qualified as incomparable. Thus, in practice, quite a successful model run might be accepted as an inappropriate when tested using conventional criteria, whereas it would require just a minimal adjustment (if any) when its temporal organization is analyzed.

The scaling exponent describes a very important property of a series, the temporal memory. Unlike NS and other conventional measures, it does not monitor the coincidence of values at each time step, but rather assesses the typical behaviour of the flow. Scaling properties of a river flux can be regarded as a 'fingerprint' of the latter, with specific shape DFA curve, values of α at different time scales, and position of crossover reflecting the change in the memory at particular time scale. Although the traditional criteria serve for comparison of observed and simulated flows point-by-point, they can actually be the same for different rivers, while telling nothing about the dynamics of a particular river system.

Therefore, the traditional analysis would flourish by the addition of the scaling exponent α. The point-wise criteria can be very useful at the final stage of the fine tuning of the model, when the scaling properties are reproduced well and are detected by the scaling exponent α.

Multifractal analysis

Methodology

Multifractality is a unique property of a nonlinear dynamic system. It characterizes a multiplicative mixture of contributing subprocesses and provides essential information about stable states of the system. Studying multifractal properties is crucial for revealing complexity of the dynamical process under consideration (Tessier et al., 1996; Pandey et al., 1998; Kantelhardt et al., 2002, 2003). Here, our tool for this study is the multifractal detrended fluctuation analysis (MFDFA) method (Koscielny-Bunde et al., 1998; Kantelhardt et al., 2002), which is a generalised DFA method. It has been recently shown (Kantelhardt et al., 2003) that this method demonstrates results equivalent to those obtained by the well-established wavelet transform modulus maxima (WTMM) technique (Arneodo et al., 1995; Muzy et al., 1991).

In the MFDFA procedure, the moments $F_q(s)$ are calculated by (i) integrating the initial series, (ii) splitting the series into non-overlapping segments of length s, (iii) calculating the mean-square deviations $F^2(v,s)$ from polynomial fits in each segment, (iv) averaging $[F^2(v,s)]^{q/2}$ over all segments, and (v) taking the qth root:

$$F_q(s) = \left\{ \frac{1}{2N_s} \sum_{v=1}^{2N_s} [F^2(v,s)]^{q/2} \right\}^{1/q}. \quad (1)$$

Afterwards, we determine the scaling behaviour of the fluctuation function by analysing the log–log plots of $F_q(s)$ versus s for each value of q. The series is said to possess a power-law scaling if

$$F_q(s) \sim s^{h(q)}.$$

For stationary time series, $h(2)$ is identical to the well-known Hurst exponent H (Feder, 1988), and $h(q)$ is termed a generalized Hurst exponent. Only if small and large fluctuations scale differently, will there be a significant dependence of $h(q)$ on q. For positive values of q, $h(q)$ describes the scaling behaviour of the segments with large fluctuations. For negative values of q, $h(q)$ describes the scaling behaviour of the segments with small fluctuations.

We use third-order polynomials in the fitting procedure of step (iii) (MFDFA3), thus eliminating quadratic trends in the data. We consider both positive and negative moments $F_q(s)$ and determine them for time scales s between $s = 5$ and $s = N/4$.

In the standard multifractal formalism, the scaling exponent $\tau(q)$ is determined via the partition function

$$Z_q(s) = \frac{1}{N_s} \sum_{v=1}^{N_s} |Y(vs) - Y((v-1)s)]|^q \sim s^{\tau(q)},$$

where Y is the profile of the series and q is a real parameter as in MFDFA. The exponent $h(q)$ can be related to the exponent $\tau(q)$ as follows:

$$\tau(q) = qh(q) - 1.$$

Another way to characterize a multifractal time series is to calculate the singularity spectrum $f(\alpha)$, which is related to $\tau(q)$ via the first-order Legendre transform (Feder, 1988):

$$\alpha = \tau'(q), \quad f(\alpha) = q\alpha - \tau(q),$$

where α is the Hölder exponent. Therefore, we can obtain the following relations:

$$\alpha = h(q) + qh'(q), \quad f(\alpha) = q[\alpha - h(q)] + 1.$$

The width of the spectrum $f(\alpha)$ characterizes the strength of multifractal effects in the data. For monofractal data, the spectrum $f(\alpha)$ collapses into a single point, and both functions $\tau(q)$ and $h(q)$ are linear.

Results

To reduce the effect of nonstationarities in the data, we apply the MFDFA to the deseasoned time series ΔX_i in the asymptotic scale (above 90 days). In Fig. 8 one can see the partition function $F_q(s)$ (Eq. (1)), for $q = -8, -6, -4, -2, -1, 1, 2, 4, 6, 8$. The simulated series demonstrate behaviour similar to that of the real data, though the corresponding curves for negative moments have slightly different exponents. For the Naab river at negative moments, the curves for simulated data have sharper crossover, as compared to real data, and, in consequence, different scaling exponents.

In Fig. 9, curves of the function $h(q)$ for real and simulated data (circles) are plotted. Dashed lines denote the confidence intervals of $h(q)$, which were calculated based on the ensemble of subsets of each record in the following manner. Using a sliding window of 6000 days in length, we obtain N−6000+1 subsamples, and then for each of them calculate $h(q)$. Next, we average the results over the ensemble of subsamples and arrive at the estimates of mathematical expectation and a confidence interval of $h(q)$. This is, in essence, one of the resampling method (Efron, 1982). It does not necessarily provide a fully reliable

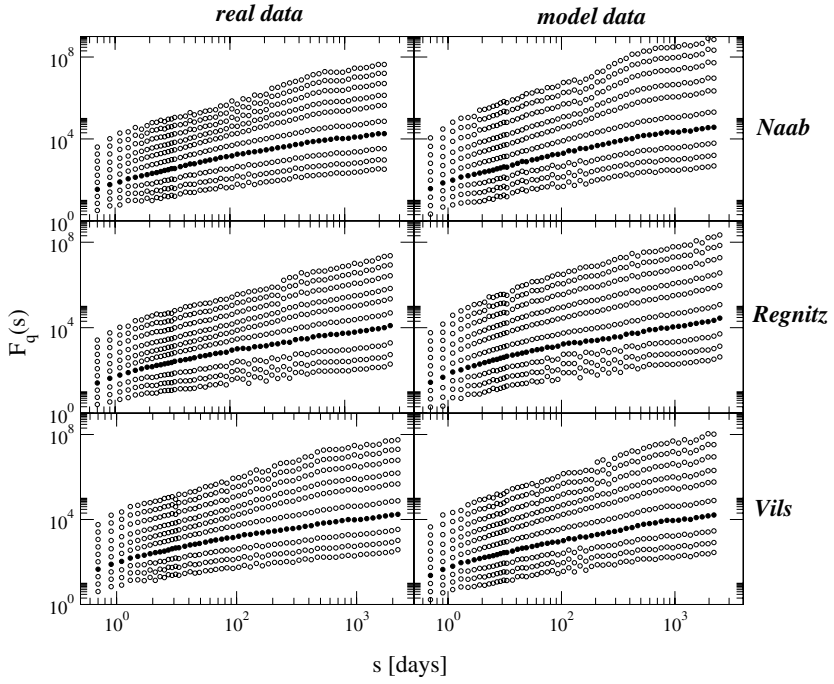

Figure 8 Multifractal fluctuation function $F_q(s)$ versus time scale s obtained with MFDFA3. The curves correspond to moment values $q = -8, -6, -4, -2, -1, 1, 2, 4, 6, 8$, from top to bottom (solid symbols for $q = 2$, conventional DFA3).

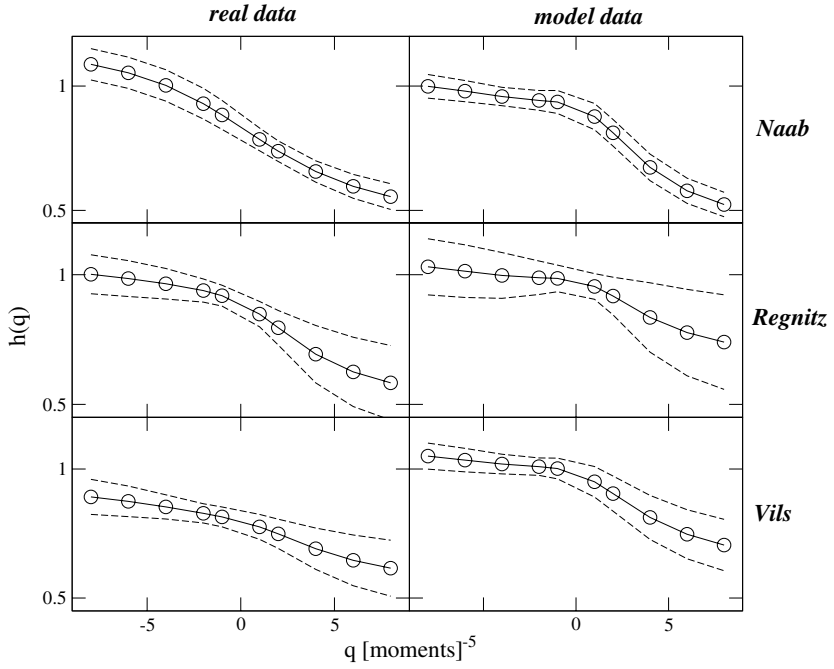

Figure 9 Generalized Hurst exponent measured in the scale above 90 days. Dashed lines represent confidence interval calculated for data subsets of length 6000 days.

estimate of the confidence interval, but can be used in our case for comparison of the results of the multifractal analysis for the real and simulated data.

In Fig. 9, the confidence intervals for the Regnitz, for both real and simulated data, are largest among the considered time series. This can be caused by nonstationarities in the data (subsets of length 6000 days). For simulated Naab data at negative q, the generalized Hurst exponent is underestimated, and for the Vils — overestimated. For Regnitz, the simulated data have overestimated exponents for positive q. In Fig. 10, again for the simulated Regnitz data, the discrepancy is highest when

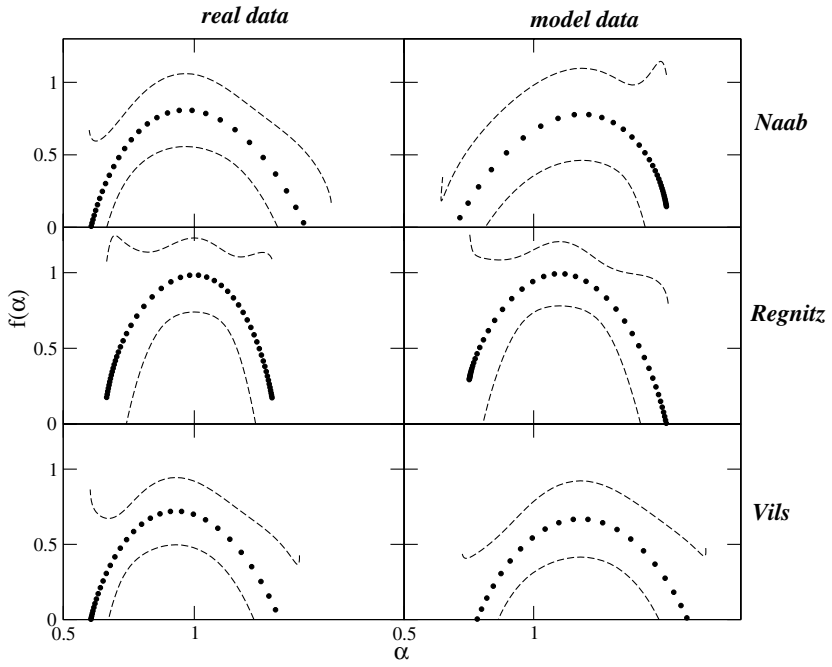

Figure 10 Multifractal spectra measured in time scale above 90 days. Filled circles correspond to data under consideration, dashed lines represent confidence interval calculated for data subsets of length 6000 days.

one compares the multifractal spectra calculated for the whole set and the confidence interval based on the ensemble of subsets. Note that the width of the spectra (or strength of the multifractal effect) is reproduced properly in the simulated data, but the exponent α is overestimated for all three rivers.

Nonlinear volatility analysis

Methodology

Recently, we have identified nonlinear long-term and periodic volatilities in river flow records (Livina et al., 2003a). We define nonlinearity with respect to the Fourier phases in the following way: if the statistical properties of a time series depend only on the power spectrum and the probability distribution, regardless of the Fourier phases, the series is considered to be *linear*. Otherwise, the series is defined as *nonlinear* (for more details see Livina et al., 2003a,b; Ashkenazy et al., 2001, 2003). We consider the absolute values of river flow increments (volatility), and in order to study nonlinearity of river flow data, we use a surrogate data test (Schreiber and Schmitz, 2000). The surrogate data have the same probability distribution and almost the same power spectrum as the original series, but with random Fourier phases. When the data are nonlinear in the above sense, the power spectrum of the volatility series has a seasonal peak (periodic volatility), which disappears after the phase randomization procedure (Livina et al., 2003a). After elimination of the nonlinearities by randomizing Fourier phases of the river flow increment series, the seasonal periodicity in the volatility is significantly weakened. Moreover, the volatility series shows long-term correlations, which are also eliminated after randomizing the Fourier phases of the river flow increment series. We use periodic and long-term volatilities as measures of nonlinearity, which should be reproduced in the simulated series.

Results

To apply the long-term volatility analysis, we consider the absolute values of the river flux increments (deseasoned), the volatility series $\widetilde{X}_k = |X_{k+1} - X_k|, k = 1, \ldots, N-1$. Here, to study correlations in the volatility data, DFA3 is used. In the time range 1 week $-$ 1 year, we obtain a correlation exponent $\alpha \sim 0.85$, and for window scales larger than 1 year, $\alpha \sim 0.59$ (see Fig. 11). After applying the surrogate data test for nonlinearity, the exponent decreases to $\alpha \sim 0.53$ for window scales larger than one week. The results of the surrogate data test, i.e., the decrease of the volatility exponent from a large value to uncorrelated value, indicates the nonlinearity of the river flow. The same effect is observed in the simulated data. For each data sample, we generated ten surrogate series and calculated mean value and standard deviation of their correlation exponents. The summary of the results for real and simulated data is given in Tables 2 and 3.

In Fig. 12, one can see the power spectra of the volatilities for the real and simulated data, before (thick curves) and after (thin curves) applying the surrogate data test. The simulated data demonstrate spectra similar to those of the real data. After randomizing the phases in both the real and simulated data, the periodic peaks vanish. Therefore, the ASGi model reproduces the nonlinear periodic volatility property in the river flow data.

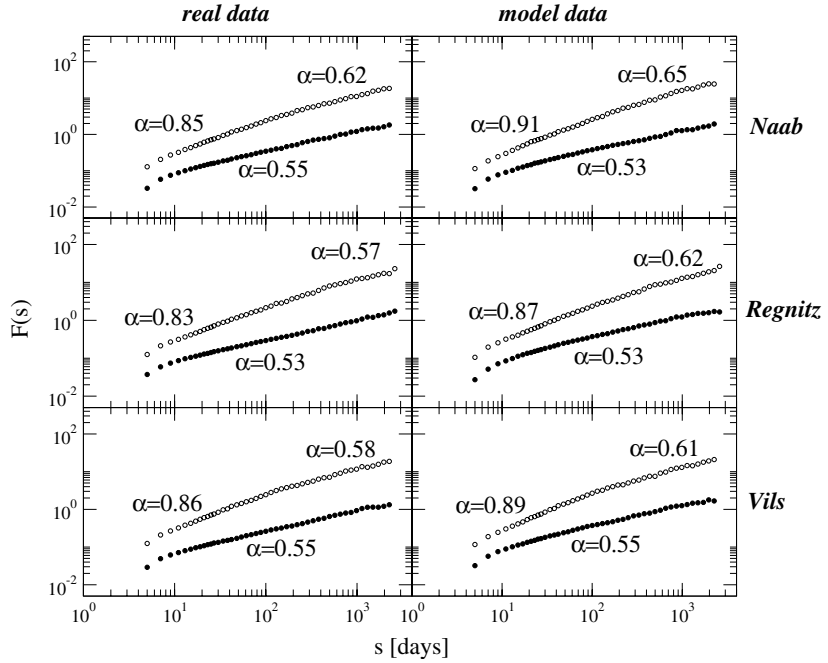

Figure 11 Long-term volatility effect (open symbols for initial data, filled symbols for phase randomized data); correlations are well reproduced by the model, both in short- and long-term ranges. After phase randomization, real and model data become uncorrelated, indicating the nonlinearity in the initial data.

Table 2 DFA3 exponents in long-term volatilities of real data

River	Data	Short-range	Long-range
Naab	Real	0.85	0.62
	Surrogate	0.54 ± 0.02	
Regnitz	Real	0.83	0.57
	Surrogate	0.51 ± 0.03	
Vils	Real	0.86	0.58
	Surrogate	0.53 ± 0.02	

Table 3 DFA3 exponents in long-term volatilities of model data

River	Data	Short-range	Long-range
Naab	Model	0.91	0.65
	Surrogate	0.53 ± 0.02	
Regnitz	Model	0.87	0.62
	Surrogate	0.52 ± 0.02	
Vils	Model	0.89	0.61
	Surrogate	0.53 ± 0.02	

Conclusion

We have applied the DFA method to compare the scaling properties of real and ASGi simulated river flux series. We found that flux records are highly short-term correlated and less correlated in the asymptotic range. We suggest that this type of memory in the river data is caused by the inertia of river basin aquifers (groundwater storage), which are the main sources for the river stream: the high water level is most likely followed by a still high level, and vice-versa. We applied multifractal analysis and found small differences between real and simulated flux records. To study correlation properties of absolute values of river flow increments (i.e., the volatility), DFA3 was used. The volatility series are correlated for time scales smaller than one year, and these correlations become much less pronounced for time scales larger than one year. Through the use of a surrogate data test which randomizes the Fourier phases of the series, we show that volatility correlations is an indication of nonlinearity, and that the nonlinearity decreases for time scales larger than one year. These effects have been observed in both real and simulated data, which confirms good performance of the ASGi model.

Our main conclusion is that the ASGi model does reproduce the main statistical properties of the river flux fluctuations: (i) values of scaling exponents and crossover in DFA curves; (ii) general multifractal behaviour of the data, though not precisely, since there is a difference in multifractal exponents between real and simulated data; (iii) scaling exponents of the volatility series for small and large time scales; and finally, (iv) the periodic volatility of the volatility series.

Our methods provide advanced statistical measures (scaling exponents, width of multifractal spectrum, peak height of normalised power spectrum, and correlation exponent of volatilities) for studying hydrological records. For instance, as we showed in ''Conventional hydrological statistics and its comparison with the scaling analysis'' section in a simple experiment, the scaling exponent properly detects similarity of series in cases when the conventional

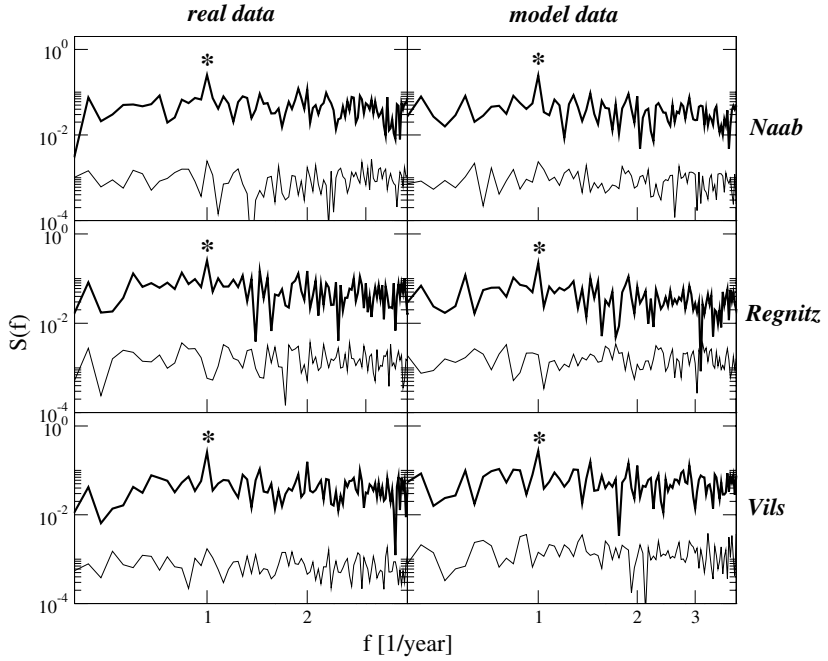

Figure 12 Periodic volatility effect: periodicity peak in power spectrum of real and model volatility series (thick lines) and results of phase randomization test (thin lines). Periodicity peaks denoted by stars become much less pronounced after randomization of the phases, both in real and simulated data, indicating existence of nonlinearity of the volatility series.

criteria fail. Moreover, conventional criteria tell nothing about nonlinearity and its strength, which is an important feature of the observed river data that should be reproduced by hydrological models. Different model sets can produce the same values of the Nash–Sutcliffe coefficient, and one of the series might be linear while the other might be not. Therefore, the modern techniques of nonlinear statistical analysis are to be employed in hydrological modelling to provide advanced comparison of real and model data and deeper evaluation of model performance.

To summarize, the ASGi model data generally demonstrate close similarity to real river flux. By means of our methods, some statistical characteristics have been estimated. We suggest that they can be used for evaluating hydrological models. The ensemble confidence interval can be used to quantify the reliability of multifractal estimates. We hypothesize that the difference in the width of multifractal spectra might be caused by noise artefacts in the simulated data (like improper seasonal pattern), and this should be taken into account in further development of the ASGi model. We suggest that closer similarity between the multifractal properties of real and simulated time series would improve the model's performance.

Acknowledgements

We thank Dr. Y. Ashkenazy for useful discussions.

References

Arneodo, A., Bacry, E., Graves, P.V., Muzy, J.F., 1995. Characterizing long-range correlations in DNA-sequences from wavelet analysis. Phys. Rev. Lett. 74, 3293–3296.

Ashkenazy, Y., Ivanov, P.Ch., Havlin, S., Peng, C.-K., Goldberger, A.L., Stanley, H.E., 2001. Magnitude and sign correlations in heartbeat fluctuations. Phys. Rev. Lett. 86, 1900–1903.

Ashkenazy, Y., Havlin, S., Ivanov, P.Ch., Peng, C.-K., Schulte-Frohlinde, V., Stanley, H.E., 2003. Magnitude and sign scaling in power-law correlated time series. Physica A 323, 19–41.

Becker, A., Braun, P., 1999. Disaggregation, aggregation and spatial scaling in hydrological modelling. J. Hydrol. 217, 239–252.

Beven, K.J. Ed., 1998. Distributed Hydrological Modelling: applications of the Topmodel Concept (Advances in Hydrological Processes). Wiley, 356pp.

Beven, K.J., 2004. Rainfall Runoff Modelling: The Primer. Wiley, 372pp.

Beven, K.J., Moore, I.D. (Eds.), 1993. Terrain Analysis and Distributed Modelling in Hydrology (Advances in Hydrological Processes). Wiley, p. 256.

Blöschl, G., 2001. Scaling in hydrology. Hydrol. Process. 15, 709–711.

Braun, P., Holle, F.-K., Becker, M., 1998. Die Fraktaltheorie in der HW-Statistik: Entwicklung Einer Neuen Multifraktalen Verteilungsfunktion. DGM 42 (5), 1–15.

Buldyrev, S.V., Goldberger, A.L., Havlin, S., Mantegna, R.N., Matsa, M., Peng, C.-K., Simons, S., Stanley, H.E., 1995. Long-range correlations properties of coding and noncoding DNA sequences: GenBank analysis. Phys. Rev. E 51, 5084–5087.

Bunde, A., Havlin, S., Kantelhardt, J.W., Penzel, T., Peter, J.H., Voigt, K., 2000. Correlated and uncorrelated regions in heartrate fluctuations during sleep. Phys. Rev. Lett. 85, 3736–3739.

Dutta, D., Herath, S., Misake, K., 2000. Flood inundation simulation in a river basin using a physically based distributed hydrologic model. Hydrol. Process. 14, 497–519.

Efron, B., 1982. The Jackknife, the Boostrap and other Resampling Plans. Society for Industrial and Applied Mathematics, Philadelphia (pp. 92).

Enescu, B., Ito, K., Struzik, Z.R., 2006. Wavelet-based multiscale resolution analysis of real and simulated time-series of earthquakes. Geophys. J. Int. 164 (1), 63.

Feder, J., 1988. Fractals. Plenum Press, New York, 283pp.

Govindan, R.B., Vjushin, D., Brenner, S., Bunde, A., Havlin, S., Schellnhuber, H.-J., 2002. Global climate models violate scaling of the observed atmospheric variability. Phys. Rev. Lett. 89, 028501.

Gurtz, J., Zappa, M., Jasper, K., Lang, H., Verbunt, M., Badoux, A., Vitvar, T., 2003. A comparative study in modelling runoff and its components in two mountainous catchments. Hydrol. Process. 17 (2), 297—311.

Hogue, T.S., Gupta, H., Sorooshian, S.A., 2006. 'User-Friendly' approach to parameter estimation in hydrological models. J. Hydrol. 320, 202—217.

Ivanov, P.C., Amaral, L.A.N., Goldberger, A.L., Havlin, S., Rosenblum, M.G., Struzik, Z.R., Stanley, H.E., 1999. Multifractality in human heartbeat dynamics. Nature 399 (6735), 461—465.

Jasper, K., Gurtz, J., Herbert, L., 2002. Advanced flood forecasting in Alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model. J. Hydrol. 267 (1-2), 40—52.

Jones, J.A.A., 1997. Global Hydrology: Processes, Resources and Environmental Management. Addison Wesley Longman, 399pp.

Kantelhardt, J.W., Zschiegner, S.A., Koscielny-Bunde, E., Havlin, S., Bunde, A., Stanley, H.E., 2002. Multifractal detrended fluctuation analysis of nonstationary time series. Physica A 316, 87—114.

Kantelhardt, J.W., Rybski, D., Zschiegner, S.A., Braun, P., Koscielny-Bunde, E., Livina, V.N., Havlin, S., Bunde, A., 2003. Multifractality of river runoff and precipitation: comparison of fluctuation analysis and wavelet methods. Physica A 330 (1-2), 240—245.

Kantelhardt, J.W., Koscielny-Bunde, E., Rybski, D., Braun, P., Bunde, A., Havlin, S., 2006. Long-term persistence and multifractality of precipitation and river runoff records. J. Geophys. Res. Atmosph. 111 (D1), D01106.

Koscielny-Bunde, E., Bunde, A., Havlin, S., Roman, H.E., Goldreich, Y., Schellnhuber, H.-J., 1998. Indication of a universal persistence law governing atmospheric variability. Phys. Rev. Lett. 81, 729—732.

Livina, V., Ashkenazy, Y., Braun, P., Monetti, R., Bunde, A., Havlin, S., 2003a. Nonlinear volatility of river flux fluctuations. Phys. Rev. E 67, 042101.

Livina, V.N., Ashkenazy, Y., Kizner, Z., Strygin, V., Bunde, A., Havlin, S., 2003b. A stochastic model of river discharge fluctuations. Physica A 330 (1-2), 283—290.

Mantegna, R.N., Stanley, H.E., 1999. An Introduction to Econophysics: Correlations and Complexity in Finance. Cambridge University Press, Cambridge, 158pp.

Marshall, L., Nott, D., Sharma, A., 2005. Hydrological model selection: a Bayesian alternative. Water Resources Research 41, W10422.

Matia, K., Ashkenazy, Y., Stanley, H.E., 2003. Multifractal properties of price fluctuations of stock and commodities. Europhysics Lett. 61 (3), 422—428.

Montanari, A., Rosso, R., Taqqu, M.S., 2000. A seasonal fractional ARIMA model applied to the Nile River monthly flows at Aswan. Water Res. Res. 36 (5), 1249—1259.

Muzy, J.F., Bacry, E., Arneodo, A., 1991. Wavelets and multifractal formalism for singular signals: Application to turbulence data. Phys. Rev. Lett. 67, 3515—3518.

Pandey, G., Lovejoy, S., Schertzer, D., 1998. Multifractal analysis of daily river flows including extremes for basins of five to two million square kilometres, one day to 75 years. J. Hydrol. 208 (1—2), 62—81.

Peng, C.-K., Buldyrev, S.V., Havlin, S., Simons, M., Stanley, H.E., Goldberger, A.L., 1994. Mosaic organization of DNA nucleotides. Phys. Rev. E 49, 1685—1689.

Peng, C.-K., Havlin, S., Stanley, H.E., Goldberger, A.L., 1995. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. Chaos 5, 82—87.

Refsgaard, J.C., Knudsen, J., 1996. Operational validation and intercomparison of different types of hydrological models. Water Res. Res. 32 (7), 2189.

Santhanam, M.S., Bandyopadhyay, J.N., Angom, D., 2006. Quantum spectrum as a time series: fluctuation measures. Phys. Rev. E 73 (1), 015201.

Schreiber, T., Schmitz, A., 2000. Surrogate time series. Physica D 142 (3—4), 346—382.

Schulla, J., 2000. Hydrologische Modellierung von Flussgebieten zur Abschätzung der Folgen von Klimaänderungen, Zürcher Geographische Schriften, Heft 69, Geographisches Institut der ETH Zürich; pp. 1—187.

Struzik, Z.R., 2000. Determining local singularity strengths and their spectra with the wavelet transform. Fractals 8 (2), 163—179.

Tessier, Y., Lovejoy, S., Hubert, P., Schertzer, D., Pecknold, S., 1996. Multifractal analysis and modeling of rainfall and river flows and scaling, causal transfer functions. J. Geophys. Res. 101 (D21), 26427—26440.

Verbunt, M., Gurtz, J., Jasper, K., Lang, H., Warmerdam, P., Zappa, M., 2003. The hydrological role of snow and glaciers in alpine river basins and their distributed modeling. J. Hydrol. 282 (1—4), 36—55.

Vyushin, D., Zhidkov, I., Havlin, S., Bunde, A., Brenner, S., 2004. Volcanic forcing improves atmosphere—ocean coupled general circulation model scaling performance. Geophys. Res. Lett. 31, L10206.

Wood, E.F., Sivapalan, M., Beven, K.J., Band, L., 1988. Effects of spatial variability and scale with implications to hydrological modelling, 1988. J. Hydrol. 102, 29—47.