**Nonlinear Processes
in Geophysics**

Trend assessment: applications for hydrology and climate research

M. Kallache, H. W. Rust, and J. Kropp

Potsdam Institute for Climate Impact Research, P.O. Box 60 12 03, 14412 Potsdam, Germany

Abstract. The assessment of trends in climatology and hydrology still is a matter of debate. Capturing typical properties of time series, like trends, is highly relevant for the discussion of potential impacts of global warming or flood occurrences. It provides indicators for the separation of anthropogenic signals and natural forcing factors by distinguishing between deterministic trends and stochastic variability. In this contribution river run-off data from gauges in Southern Germany are analysed regarding their trend behaviour by combining a deterministic trend component and a stochastic model part in a semi-parametric approach. In this way the trade-off between trend and autocorrelation structure can be considered explicitly. A test for a significant trend is introduced via three steps: First, a stochastic fractional ARIMA model, which is able to reproduce short-term as well as long-term correlations, is fitted to the empirical data. In a second step, wavelet analysis is used to separate the variability of small and large time-scales assuming that the trend component is part of the latter. Finally, a comparison of the overall variability to that restricted to small scales results in a test for a trend. The extraction of the large-scale behaviour by wavelet analysis provides a clue concerning the shape of the trend.

1 Introduction

Within the climate change debate it is common sense that the hydrological cycle would intensify as a consequence of an enhanced greenhouse effect. This has led to the accompanying assumption that also extreme events, such as heavy rain, droughts, or floods , will increase in frequency and/or magnitude (IPCC, 2001). However, an improved knowledge of the impacts of climate change on river discharges is urgently needed, since floods are among the costliest natural hazards in both economic and human terms (Mileti, 1999; DFO, 2004). Recent examples, such as the river Elbe/Danube flood in Europe in summer 2002 ($\approx$20 billion EUR losses and 55 deaths) or South Asia's worst monsoon flood for almost a decade in summer 2004 (Ganges/Bramaphutra/Irrawady; India, Bangladesh, Burma) with $\approx$1600 causalties and millions of displaced people, again bring the hazardousness of floods to public attention. Taking into account that climate change already induces slight impacts on the hydrological cycle, the actual pattern is ambiguous. Kundzewicz et al. (2004), for instance, could not find evidence of a general growth in flood flows when analysing 195 long time series of maximum annual river flows located world-wide. Similar results are published by Zhang et al. (2001) who found no increasing trend for Canadian rivers, but in some cases a decreasing development. For U.S. rivers Douglas et al. (2000) found no trends in flood flows, but evidence for an upward trend in low flows. Analysing streamflow data in Switzerland, the results of Bîrsan et al. (2003) indicate an increase of run-off in winter and a decrease in summer.

This situation shows the need for methods which are able to provide an improved knowledge on the mechanisms that produces a series, their seasonal variations, internal correlations, and possible trends. The most common technique used to assess the significance of trends in hydro-meteorological times series data is the rank-based Mann-Kendall test (Mann, 1945; Kendall, 1975). It accepts or rejects the null hypothesis of randomness against the alternative of a monotonic trend. The advantage of the Mann-Kendall test is that it relies only on a few assumptions: the potential trend may be either linear or nonlinear, and no assumptions are made regarding the underlying statistical distribution. Nevertheless, it is well recognised that the Mann-Kendall test is not robust against autocorrelation in the sense that false positive trend identifications get more likely (Fleming and Clarke, 2002). This effect depends on the sample size as well as on the magnitude of the trend to be identified. Prewhitening techniques introduced to remove effects induced by autocorrelation may

---

Correspondence to: M. Kallache
(malaak.kallache@pik-potsdam.de)

also bias the Mann-Kendall test result (Yue and Wang, 2002). Therefore we propose a semi-parametric approach to analyse correlated data.

Different potential driving forces have to be considered when analysing hydrological or meteorological time series, for example, anthropogenic influences and natural variability. The distinction of such components is still a matter of research. Our approach is based on the rather simple assumption that anthropogenic influences can be represented by a deterministic trend component, whereas natural variability is modelled by a stochastic process. It is desirable to distinguish between these components, because the former is expected to continue in the future whereas the latter changes direction in a more or less unpredictable manner.

However, smoothly varying trends are difficult to distinguish from natural variability on large scales caused e.g. by autocorrelation. This exchangeability implies that the evaluation of a possible deterministic trend inherently needs assumptions about the natural variability: trends of small magnitudes will less often be considered significant in a system which is known to be highly variable anyway. This fact is reflected by the selection of the stochastic model capturing the natural variability. The choice has an important influence on the sigma interval estimated for the trend parameters and therefore on the significance of the trend estimate. This is taken into account when an adequate goodness-of-fit test and model selection criteria are employed to choose suitable models.

Since long-term correlation might cause long excursions from the mean as well as local trends (Beran, 1994) the issues described above are particularly challenging when dealing with long-term correlated data. Several authors report long-term memory being present in river run-off records (e.g. Lawrence and Kottegoda, 1977; Montanari et al., 1997).

Aspects of the approach we follow in this paper have been addressed previously. Bloomfield (1992), Woodward and Gray (1995) and others evaluate trends in global temperature by combining a trend and an ARIMA component in a parametric approach. Beran (1994) describes the characteristics of long-term correlated processes and their effect on trend estimation, whereas Koutsoyiannis (2003) adapts elements of hydrological statistics to the Hurst phenomenon. Smith (1993) utilises the standard least squares estimator to test for trends under consideration of long-term memory. Bhattacharya et al. (1983), Taqqu et al. (1995), Bunde et al. (2002) and others address pitfalls when discriminating between deterministic trends and long-range components and provide possible ways of distinction. Sibbertsen (1999) uses kernel estimators to assess a trend, whereas Percival and Walden (2000) and Craigmile et al. (2004) employ wavelets. They provide a semi-parametric approach to test for the significance of a trend under assumption of a FD(δ) or an AR(1) model.

In this paper we extend their approach by including more flexible fractional ARIMA models. To account for the presence of a trend, we also use filtered data for the parameter estimation. The trend is considered as a slowly varying de-

terministic component on large scales. By using wavelet filters the data are decomposed into variations on large scales (further referred to as trend estimate $\widehat{T}$) and variations on smaller scales ($\widehat{X}$). The latter are assumed to be modelled adequately by a linear stochastic process. To evaluate the performance of the models, a goodness-of-fit test is applied. Selection from among the different models which passed this test, is achieved by applying an Akaike-type model selection criterion. Finally, under assumption of the chosen stochastic model, the time series data are tested for an underlying deterministic trend.

We have organised the paper as follows: in Sect. 2 the applied methods, i.e. wavelet analysis, trend estimation, the Whittle estimator for FARIMA(p, δ, q) models and the trend test itself, are briefly discussed. Section 3 introduces the data, and in Sects. 4 and 5 the results are presented and discussed.

2 Methodological concept

To be able to assess trends in a data-specific manner and to account for the correlation structure of the data, flexible tools like wavelet analysis and stochastic modelling are chosen. Here wavelet analysis is used to estimate a deterministic trend component from data. Then FARIMA(p,δ,q) processes are used to model the structure of the remaining fluctuations whereby the parameters are estimated via Whittle's approximation to the Maximum-Likelihood Estimator. Finally, wavelet coefficients are used to establish a test for significance of the estimated trend. All three steps are linked and serve together serve as the trend assessment method.

2.1 Wavelet analysis

For the wavelet analysis, we use the discrete wavelet transform (DWT) and the maximum overlap DWT (MODWT). The DWT is an orthonormal transform. The time series can be reconstructed by a linear combination of wavelets, analogous to a reconstruction by sinusoids in Fourier analysis. Each wavelet is essentially non-zero only within a finite interval of time, which allows for time-scale analysis. In this manner information about the variability which is not only local in scale but also local in time is retrieved. The wavelet basis is dilated when processing larger scales. Thus, the problem of under- or over-localisation, which occurs if the windowed fourier transformation is used, can be minimised (see Kaiser, 1994). The DWT operates on dyadic time series and is only computed for scales which are powers of two. It is defined in terms of a wavelet filter $\{h_{j,l} : l=0, ..., L_j-1\}$ and an associated scaling filter $\{g_{j,l}\}$. Here j denotes the levels $j=1, ..., J$, which are associated to the scales 2^{j-1}, and $L \equiv L_1$ the width of the wavelet.

Let the time series $\{Y_t\}$ be a realisation of a stochastic process with random variables Y_t, $t=0, ..., N-1$; N being dyadic. By convolution of the filters with $\{Y_t\}$, the wavelet coefficients $W_{j,k}$ and scaling coefficients $V_{j,k}$ for level $j=0, ..., J$ and time $k=0, .., N_j-1$ are obtained (with

the number of coefficients $N_j = N2^{-j}$ and wavelet length $L_j = (2^j - 1)(L-1) + 1$ at each level j). The larger the scales, the wider the applied filters. Thus, using an orthogonal transform implies that for larger scales less convolutions are done, i.e. less wavelet coefficients are available:

$$W_{j,k} \equiv \sum_{l=0}^{L_j-1} h_{j,l} Y_{2^j(k+1)-1-l \bmod N} \quad ,$$

$$V_{j,k} \equiv \sum_{l=0}^{L_j-1} g_{j,l} Y_{2^j(k+1)-1-l \bmod N} \quad . \qquad (1)$$

The algorithm is implemented efficiently by means of a pyramid algorithm (Mallat, 1989). Formally, the DWT of $\{Y_t\}$ is defined via matrix multiplication that yields a vector $\mathbf{W}_j \equiv [W_{j,0}, W_{j,1}, \ldots, W_{j,N_j-1}]'$ and $\mathbf{V}_j \equiv [V_{j,0}, V_{j,1}, \ldots, V_{j,N_j-1}]'$. Let $\mathbf{W} = [\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_J, \mathbf{V}_J]'$. The time series is then reconstructed as $Y = \mathcal{W}^T \mathbf{W}$, where the DWT matrix $\mathcal{W}$ is a $N \times N$ matrix with rows constructed from the filters. For a more detailed description of the DWT we refer to Appendix A. The exact form of $\mathcal{W}$ is provided in Percival and Walden (2000).

The choice of the mother wavelet depends on the data to be analysed. Possible choices are an orthonormal or non-orthonormal, a real or complex basis (an overview about the noteworthy aspects is given by Torrence and Compo, 1998). For our analysis we chose a discrete, orthogonal basis – the Daubechies "least asymmetric" wavelet basis of width eight (LA(8)). The wavelet width comprises smoothing in time and frequency in an acceptable manner regarding the data analysed. When using the $LA(\cdot)$ basis one can also benefit from the advantages of approximately linear filtering. Applying the LA(8) wavelet filter to $\{Y_t\}$ essentially yields a difference between Y_t and values before and after Y_t. The LA(8) scaling filter results in a weighted average of a window of length 2^{j+1} on scale 2^j. These scaling filters are defined through their squared gain function $\mathcal{G}(\omega)$ operating on the Fourier frequencies ω, which they share with all other Daubechies filters (for further details we refer to Appendix B). By additionally choosing the transfer function $G(\omega) = [\mathcal{G}(\omega)]^{1/2} e^{i\theta(\omega)}$ in such a way that the phase function $\theta(\cdot)$ has the smallest maximum deviation from a linear phase filter, i.e. is as close as possible to that of a linear phase filter, the definition of a $LA(\cdot)$ filter is completed. For the corresponding wavelet filters the relationship $\mathcal{H}(\omega) = \mathcal{G}(\omega + \pi)$ holds.

The MODWT (maximum overlap DWT) is a modified version of the DWT. With the MODWT, time series of non-dyadic length can be analysed. In contrast to the orthonormal DWT, the MODWT is a highly redundant non-orthogonal transform. In this work we use both the DWT and the MODWT, according to the respective aims (for more details we refer to Percival and Walden, 2000).

2.2 Trend estimation

In the following, we assume that the observed time series $\{Y_t : t = 0, \ldots, N-1\}$ can be modelled additively by a deter-

ministic trend component $\{T_t\}$ plus a realisation of a stochastic process $\{X_t\}$: $Y_t = T_t + X_t$. Using the DWT, the data vector Y is decomposed into a component $\widehat{T}$, representing the variability on large scales, and a component $\widehat{X}$ for small scales:

$$Y = \mathcal{W}^T \mathbf{W} = \mathcal{W}^T \mathbf{A} \mathbf{W} + \mathcal{W}^T (\mathbf{I}_N - \mathbf{A}) \mathbf{W} = \widehat{T} + \widehat{X} \qquad (2)$$

Let $\mathbf{I}_N$ be the $N \times N$ identity matrix. $\mathbf{A}$ contains all non-boundary wavelet coefficients, whereas $(\mathbf{I}_N - \mathbf{A})$ contains all scaling and boundary coefficients. Daubechies wavelet filter of length L zero out a polynomial trend of order K in the nonboundary wavelet coefficients if $K \leq L/2 - 1$. Since a reconstruction $Y = \mathcal{W}^T \mathbf{W}$ gives back the original time series, the trend $\{T_t\}$ is captured by the boundary wavelet coefficients and the scaling coefficients. The inverse wavelet transform of these coefficients $\widehat{T}$ can be used as an estimate of the trend. $\widehat{T}$ not only contains the deterministic trend component but also the stochastic variability on large scales and in the following is referred to as the trend estimate. The bias in estimating T by $\widehat{T}$ is zero for $K \leq L/2 - 1$ (cf. Craigmile et al., 2004, for further details).

2.2.1 Separating scale

Most reliable results are obtained by selecting the separating scale l_s between $\widehat{T}$ and $\widehat{X}$ as the largest scale with a non-zero number of wavelet coefficients (see Craigmile et al., 2004). As pointed out in Sect. 2.1, applying a Daubechies scaling filter is a kind of weighted smoothing procedure. The choice of l_s determines the amount of variation of the time series present in the trend estimate. In the following, all investigations are carried out with $l_s = 6$, which corresponds to a separating scale band of 32 to 64 months. So the estimated trends reflect variations of the time series which are larger than 5 years.

2.2.2 Boundary conditions

Wavelet coefficients are always a certain aggregation of the variations of the time series around a time point t_0. The larger the analysed scale, the more points become aggregated, i.e. the more wavelet coefficients are affected by the boundary conditions. To be able to calculate these wavelet coefficients, some kind of assumption for the elongation of the time series at both edges has to be made. One alternative is to assume periodicity, i.e. the end of the time series is identified with the beginning. Other ways would be to expect reflecting boundary conditions, where the time series is reflected at both edges, or to pad the series at the edges with the mean value (see Torrence and Compo, 1998, for more details). In the following we use the latter alternative. As one goes to larger scales, padding decreases the amplitude of the wavelet and scaling coefficients as more zero values enter the analysis. The trend estimation and the calculation of its sigma intervals is done on this elongated, padded series, whereas the test for trend is applied to the original empirical series

and its fluctuation estimate without padded parts. When using wavelet filters with width $L>2$ as basis, the boundary coefficients cause a higher variability of the trend estimate at the endpoints of the time series and accordingly wider sigma intervals there.

2.3 Parameter estimation

The natural variability present in the records is modelled using fractional ARIMA (FARIMA) models (Granger and Joyeux, 1980). These models are an extention of the well studied autoregressive integrated moving average (ARIMA) models (Box and Jenkins, 1976) incorporating long-range dependence. A process has long-range dependence or long-range correlation if the autocorrelation function $\rho(k)$ decays algebraically in the limit of large time lags k:

$$\lim_{k \to \infty} \frac{\rho(k)}{ck^{-\beta}} = 1, \qquad (3)$$

with c being a finite constant.

A FARIMA(p, δ, q) process is defined by

$$\phi(B)(1 - B)^{\delta} Y_t = \psi(B)\eta_t \qquad (4)$$

with B denoting the back shift operator $BY_t = Y_{t-1}$, $\eta_t \sim WN(0, \sigma)$, $\delta \in \mathbb{R}$ the fractional difference parameter and

$$\phi(z) = 1 - \sum_{i=1}^{p} \phi_i z^i$$

$$\psi(z) = 1 + \sum_{j=1}^{q} \psi_j z^j \qquad (5)$$

are the autoregressive (AR) and moving average (MA) polynomials of order p and q, respectively. A FARIMA(p, δ, q) process is stationary if $\delta<0.5$ and all solutions of $\phi(z)=0$ are outside the unit circle. It exhibits long-memory for $0<\delta<0.5$. FARIMA(p, δ, q) models with $\delta<0$ are called intermediate memory or "overdifferenced". In practice, this case is rarely encountered. For more details, see Beran (1994) and Ooms and Doornik (1999).

Three of the most simple models out of the FARIMA(p, δ, q) class are:

1. FARIMA(1,0,0) or AR(1)

$$(1 - \phi B) Y_t = \eta_t, \qquad (6)$$

2. FARIMA(0,δ,0) or FD(δ)

$$(1 - B)^{\delta} Y_t = \eta_t, \qquad (7)$$

3. FARIMA(1,δ,0)

$$(1 - \phi B)(1 - B)^{\delta} Y_t = \eta_t. \qquad (8)$$

The first model consists of only a short-range correlated component with parameter $\phi \equiv \phi_1$, the second model of a long-range correlated component only with long-memory parameter δ. The third model combines the previous two in the sense that for $\delta=0$ or $\phi=0$ it recovers the first or the second model, respectively.

The model parameters are estimated using Whittle's approximation to the Maximum-Likelihood Estimator (MLE). The latter is based on minimising

$$Q(\mathbf{\Theta}) = \sum_j \frac{\mathcal{P}(\omega_j)}{\mathcal{S}(\mathbf{\Theta}; \omega_j)}, \qquad (9)$$

where $\mathcal{P}(\omega_j)$ denotes the periodogram of the data at the Fourier frequencies $\omega_j = 2\pi j/N$ and $\mathcal{S}(\mathbf{\Theta}; \omega_j)$ the spectral density of the FARIMA process with parameter vector $\mathbf{\Theta}$ comprising $(\phi_1, ..., \phi_p, \delta, \psi_1, ..., \psi_q)$. At the minimum of $Q(\widehat{\mathbf{\Theta}})$, $\widehat{\mathbf{\Theta}}$ is an estimate for the model parameters. The standard deviation provided for the parameters are based on the assumption of a Gaussian process and are calculated from the asymptotic distribution of parameters in the limit of long records. As such they are only a first guess for the variability of the parameters.

An extensive discussion on modelling long-range dependence also with respect to FARIMA(p, δ, q) models and the Whittle estimator can be found in Beran (1994).

2.3.1 Trend influence and filtering algorithm

The influence of a deterministic trend component will usually bias the parameter estimation, because parameter estimation on $\mathbf{Y}$ implies assuming a possible trend to be negligible.

To account for the trend component, we estimate the parameters on $\widehat{\mathbf{X}}$, where the trend estimate $\widehat{\mathbf{T}}$ has been subtracted from the data. Although $E(\widehat{\mathbf{T}} - \mathbf{T}) = \mathbf{0}_N$ holds, the parameter estimation on $\widehat{\mathbf{X}}$ also is likely to be biased because $\widehat{\mathbf{T}}$ contains the deterministic trend component as well as stochastic variations on large scales and we have to use $\widehat{\mathbf{T}}$ instead of $E(\widehat{\mathbf{T}})$.

We studied this influence using low-order ($p, q \leq 1$) FARIMA(p, δ, q) models and trends with various strengths (0 to 3 times the standard deviation of the process) and shapes. The AR parameter ϕ as well as the fractional difference parameter δ are systematically estimated too low on the filtered series $\widehat{\mathbf{X}}$, where the bias decreases for an increasing trend. They are estimated too high on the original series $\mathbf{Y}$. This effect increases with increasing trend (see Fig. 1). The bias caused by a trend or the detrending procedure is negligible for the MA parameter ψ. The comparison of the trend test result obtained by estimating the parameters on $\mathbf{Y}$ and on $\widehat{\mathbf{X}}$ can constructively be used to draw further conclusions (see Sect. 4 for more details).

2.3.2 Goodness-of-fit and model selection

The model choice for the stochastic variability is motivated by a goodness-of-fit test and an Akaike-type model selection criterion.

A goodness-of-fit test for FARIMA(p, δ, q) models formulated in the spectral domain has been proposed by Beran (1992). The test statistic

$$T_B = \frac{A}{B^2} \qquad (10)$$

with

$$A = \frac{4\pi}{N} \sum_{j=1} \left(\frac{\mathcal{P}(\omega_j)}{\mathcal{S}(\omega_j; \boldsymbol{\Theta})} \right)^2 \quad \text{and}$$

$$B = \frac{4\pi}{N} \sum_{j=1} \frac{\mathcal{P}(\omega_j)}{\mathcal{S}(\omega_j; \boldsymbol{\Theta})} \qquad (11)$$

is asymptotically normal with mean $\mu_{T_B} = \pi^{-1}$ and variance $\sigma_{T_B}^2 = 2\pi^{-2} N^{-1}$. It has been shown that this test is equivalent to the Portmanteau test (see Percival et al., 2001) for uncorrelated residuals in the time domain. The smallest significance level for which the null hypothesis "H_0: the empirical data is compatible with being a realisation of the model" is rejected falsely, is denoted by

$$\alpha_{crit} = 1 - \Phi(T_B') \quad \text{for} \quad T_B' = \frac{T_B - \mu_{T_B}}{\sigma_{T_B}}, \qquad (12)$$

where $\Phi(\cdot)$ is the cumulative distribution function for a standard normal random variable.

The dynamics of the system, recorded in finite time series, cannot be reconstructed unambiguously, which corresponds to the possibility of fitting different stochastic models reasonably well to a time series.

As criteria for the model selection we employ different versions of the Akaike Information Criterion, defined by

$$\text{AIC}_f(\kappa) = -2 \log \mathcal{L}(\boldsymbol{\Theta}; Y) + f\kappa, \qquad (13)$$

where $\mathcal{L}(\boldsymbol{\Theta}; Y)$ denotes the likelihood of the parameter vector $\boldsymbol{\Theta}$. The factor f weights the influence of the penalty term due to the number of parameters κ. If f is set to 2 we get the Akaike Information Criterion (AIC) (Akaike, 1973, 1979), for $f = \log N$ the Bayesian Information Criterion (BIC) (Schwarz, 1978) and for $f = 2c \log \log N$ the Hannan-Quinn Information Criterion (HIC), in which we use $c = 1.0001$. The model to select is that with the smallest AIC$_f$. For FARIMA($p, \delta, 0$) models BIC and HIC are shown to yield consistent estimators for the autoregressive order p (Beran et al., 1998), for some low order FARIMA(p, δ, q) models Bisaglia (2002) performed simulation studies and recommends BIC and HIC as selection criteria. In the following the HIC was used.

2.4 Test for significance of trend

A test for trend is provided by comparing the variability in $\boldsymbol{Y}$ and $\widehat{\boldsymbol{X}}$. Let

$$p_c = \frac{\|\boldsymbol{Y}\|^2}{\|\widehat{\boldsymbol{X}}\|^2} \qquad (14)$$

be the test statistic, where $\|.\|$ denotes the Euclidian norm. Without loss of generality, we assume $\boldsymbol{Y}$ has zero mean.
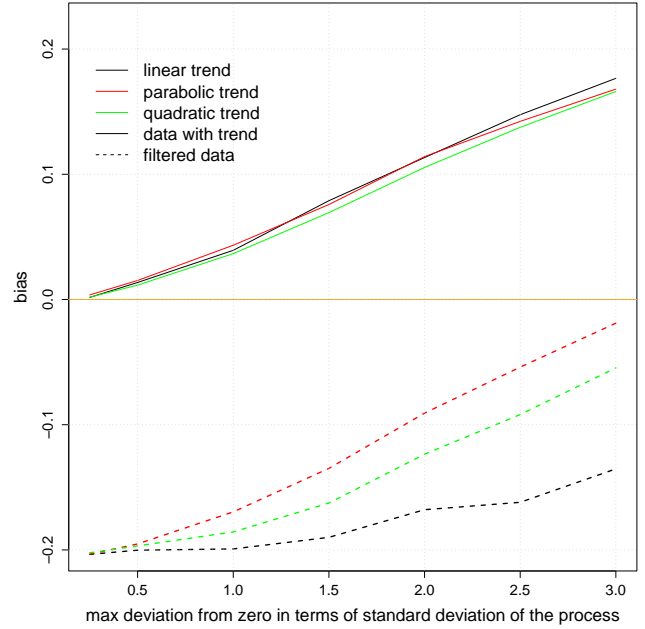

Fig. 1. Results of a Monte Carlo simulation to assess the bias due to a trend or to filtering when estimating the long-term parameter δ of a FARIMA(1,δ,1) model.

H_0 is now: $T_t = 0 \ \forall t = 0, ..., N-1$ versus H_1: not H_0. For $\{T_t\} \neq \{0\}$, p_c should be large. H_0 is rejected at a level of significance α if $p_{c(\text{empirical})}$ of $\boldsymbol{Y}$ exceeds $p_c(\alpha)$, whereby $p_c(\alpha)$ is the upper $100\alpha\%$-quantile of simulated p_c values. The distribution of the test statistic is estimated via Monte Carlo simulation with 4000 realisations of the chosen stochastic model with optimised parameters and without trend component.

2.4.1 Power of the trend test

To determine the power of the trend test, a Monte Carlo study is performed on realisations of an AR(1) and a FD(δ) model with diverse parameter choices and deterministic trends superimposed. The amplitude of the deterministic trend has been increased in small steps from 0 to 0.4 times of the standard deviation of the process. The following results were obtained:

(a) Analysing linear trends, where the trend starts at $t = 0$ and the endpoint $t = N-1$ is varied from 0 to 0.4σ, the power of the trend test is slightly lower than the power of a standard linear regression test (see also Craigmile et al., 2004).

(b) The power of the trend test does not weaken significantly when the linear trend starts in a later part of the time series. This was tested for a starting point as late as $t = (5/6)N$.

(c) The trend test itself is robust against changes in variance of the time series, which cancels out in Eq. (14).

(d) It is vulnerable against jumps in the data. This was tested by including a jump point in the middle of the series with amplitude 0 to 0.4σ. So the occurrence of
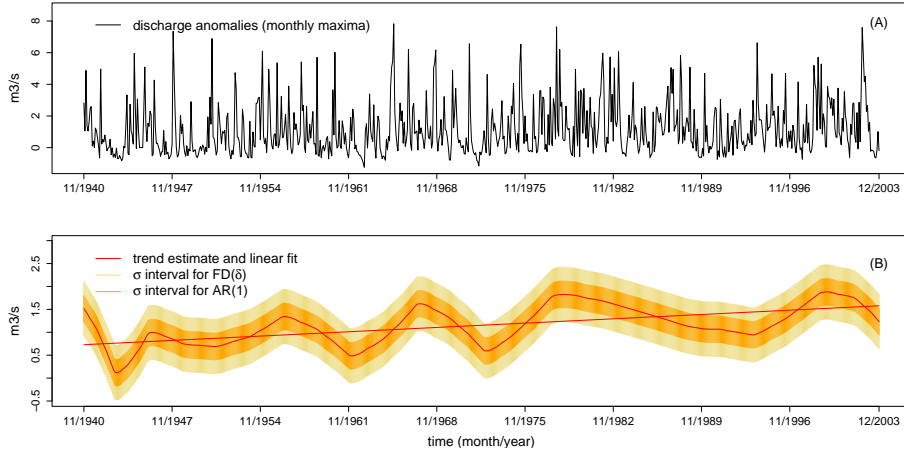
Fig. 2. (A) River discharge anomalies (monthly maxima) for the river Erms at Riederich and **(B)** trend estimate and linear fit with its standard deviation intervals for a FD(δ) model (δ=0.252, σ^2=2.465) and an AR(1) model (ϕ=0.328, σ^2=2.443).

jump points should be excluded. To deal with this problem Craigmile et al. (2004) proposed an extension to the methodology: by adopting the wavelet coefficients to the jump points the trend test can still be applied.

(e) The power does not only depend on the choice of the stochastic model, but also on the magnitude of the parameters. For an FD(δ) model, e.g. it reaches a power of one faster for small δ.

2.5 Variance of the trend

For a selected separating scale l_s the choice of the model and the magnitude of its parameters do not affect the shape of the trend estimate. However, reorganising Eq. (2) reveals the connection between $\widehat{T}$ and X, namely

$$\widehat{T} = \mathcal{W}^T \mathbf{A} \mathcal{W} Y = \mathbf{R}T + \mathbf{R}X \qquad (15)$$

with $\mathbf{R} = \mathcal{W}^T \mathbf{A} \mathcal{W}$.
Therefore holds $\text{cov}(\widehat{T})=\mathbf{R}\text{cov}(X)\mathbf{R}^T$, i.e. the variance of the trend estimate involves the covariance of X. So if X is a realisation of a stationary stochastic process, the autocovariance sequence $\rho_{\Theta}(k)$ of X can be used to calculate the variance of the trend

$$\text{var}(\widehat{T}_t) = \sum_{k=-(N-1)}^{N-1} \rho_{\Theta}(k)r_{t,k} \qquad (16)$$

with $r_{t,k}=\sum_{i=0}^{N-1-|k|} R_{t,i} R_{t,i+|k|}$. $R_{t,i}$ denotes the (t, i) element of $\mathbf{R}$ (for further details see Craigmile et al., 2004).

This relation implies that the choice of the stochastic model as well as its estimated parameters have an impact on the width of the sigma interval of the trend estimate. For an example see Fig. 2. Here fluctuations in the sigma interval due to the choice of the wavelet basis have been averaged out and the sigma intervals have been calculated for the padded series. The range of influence of the boundary conditions can be estimated analytically (see Percival and Walden, 2000).

Examining the character of the influence of the chosen model on the trend we find that the sigma intervals for the trend estimate obtained under an assumption of long-range dependence can be considerably larger than those obtained under a short-range correlated model. This is in accordance with Smith (1993) and Beran (1994).

The chosen model also influences the distribution of the test statistic given in Eq. (14) and may lead to an acceptance of the hypothesis that no trend exists in the case of long-term correlations, whereas this hypothesis is rejected in the case of short-term correlations. Thus, correctly identifying the correlation structure is crucial for a reliable trend test.

3 Data

For the analysis, river discharge records from several catchments near the river Neckar in Southern Germany were investigated. The series were selected according to their length and completeness. The run-off data was analysed for the jointly covered time period from November 1940 to December 2003. All data sets used are affected by a strong periodic component due to the annual cycle. This is approximately removed by calculating the average and variance of each day of the year over all years (missing values are replaced by the average for the specific day). The daily run-off anomalies are obtained by subtracting the average and dividing by the standard deviation. The latter removes the seasonal periodicity in the variance, see e.g. Hipel and McLeod (1994). FARIMA(p, δ, q) models do not reflect heteroscedasticity, therefore the results for the series normalised in this way are expected to be more reliable. Finally the daily measurements have been aggregated to monthly average and monthly maximum values. Records containing April to September and October to March data are used to study the summer and winter seasons separately and in the following are referred to as summer and winter data.
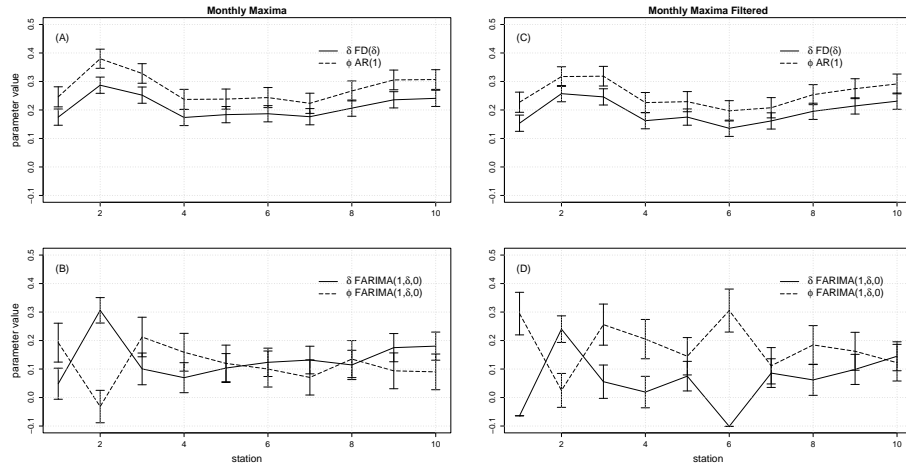
Fig. 3. (A) Fitted parameters for the FD(δ) and the AR(1) model and **(B)** the FARIMA(1,δ,0) model for the monthly maximum data and the monthly maximum filtered data, respectively (**C and D**). River Gauges: 1 Josbach/Hölzlebruck, 2 Ammer/Pfäffingen, 3 Erms/Riederich, 4 Glatt/Hopfau, 5 Wiese/Zell, 6 Dreisam/Ebnet, 7 Rems/Neustadt, 8 Fils/Plochingen, 9 Neckar/Horb, 10 Neckar/Plochingen.

4   Results and discussion

The examined time series are assumed to be composed of a deterministic trend component and a stochastic part (see Sect. 2.2).

The obtained trend estimates help to assess linear trend behaviour in the data. As shown e.g. in Fig. 2, the run-off data of the river Erms at Riederich (monthly maxima) is characterised by an increase. Nevertheless a general tendency of increase or decrease cannot be found for the investigated river run-off data sets.

In spite of these results the shape of the trend estimate and therefore the trend test result is influenced by the used wavelet filters (see Sect. 2.1) and the choice of the separating scale l_s (see Sect. 2.2.1). The latter may be compared to the selection of a window size when smoothing data to get a trend estimate.

The parameters for the stochastic models representing the natural variability of the time series are estimated using the Whittle estimator (see Sect. 2.3). We restrict our analysis to FARIMA(p, δ, q) models with $0 \leq \delta < 0.5$ and $p, q \leq 2$ ($q \leq 1$ for the winter and summer data due to the shortness of the time series). Filtering may produce data which is represented best by a model with $\delta < 0$, because not only the trend component but also the stochastic variations on large scales have been removed there. Such occurrences suggest that the long-term parameter δ is redundant. The counterpart model with $\delta = 0$ also is considered in the analysis. Therefore we exclude models with an estimated $\delta < 0$. All models which performed best in terms of the HIC turned out to be stationary, e.g. do have $\delta \leq 0.5$ and all solutions of $\phi(z) = 0$ outside the unit circle.

Obtained parameter values δ and ϕ for the FD(δ), AR(1) and FARIMA(1,δ,0) models are shown in Fig. 3 for the monthly maxima (original data and filtered data, the stations are ordered according to the increase in catchment size). The

estimates for the two one-parametric models do not change qualitatively. The magnitude of ϕ and δ is lower for the filtered data. For the more complex FARIMA(1,δ,0) model, filtering causes the short-term correlation parameter ϕ to get larger, whereas the long-term correlation parameter δ gets smaller; in some cases it becomes compatible with zero.

As discussed in Sect. 2.3.1 a possible trend biases the parameter estimation $\widehat{\Theta}'$ on Y, i.e. under assumption of no trend. To account for the bias the parameter estimation is performed under consideration of a deterministic trend component, represented by the trend estimate $\widehat{T}$, i.e. $\widehat{\Theta}''$ is estimated on the filtered series $\widehat{X}$. $\widehat{T}$ not only contains the deterministic trend component but also stochastic variations on large scales. Therefore, $\widehat{\Theta}''$ is likely to be biased in the other direction than $\widehat{\Theta}'$. Using both estimates for the analysis helps to interpret trend test results. A more reliable parameter set $\widetilde{\Theta}$ could be obtained by extracting a best trend guess $\widetilde{T}$ out of $\widehat{T}$, e.g. by fitting a polynomial to the trend estimate. $\widetilde{\Theta}$ is then estimated on the detrended series, where $\widetilde{T}$ has been substracted. Another possibility is to use the magnitude of $\widetilde{T}$ in terms of the standard deviation of the process itself to estimate the bias of $\widehat{\Theta}'$ and $\widehat{\Theta}''$ by Monte Carlo simulation (see Fig. 1). These issues will be the subject of further research.

To evaluate the performance of the models with optimised parameters, a goodness-of-fit test is applied and an Akaike-type model criterion is used to select among different models (see Sect. 2.3.2). In Table 1 the best performing models in terms of the HIC have been listed for the monthly averaged data and the monthly maxima, respectively. The apparently best HIC models coincide with a goodness-of-fit test failure in some cases. This either means that no appropriate model has been found to represent the data reasonably well within the FARIMA(p, δ, q) model class, or that the presence of a fairly weak trend is indicated, which biases the parameter estimation on Y as well as on $\widehat{X}$. In most cases a goodness-of-fit test failure occurs for the original data as well as for the

Table 1. FARIMA(p,δ,q) models performing best in terms of the HIC for (A) original/filtered monthly averaged data and for (B) original/filtered monthly maxima data. By filtering, variations on large scales have been eliminated to exclude a bias due to a possible trend. Cases where a significant trend has been detected have been underlined. Parameters written in italics refer to models which have been rejected by the goodness-of-fit test at a 95% level.

	(A) monthly averages best FARIMA(p, δ, q) model (original data/filtered data)			(B) monthly maxima best FARIMA(p, δ, q) model (original data/filtered data)		
	Total	Summer	Winter	Total	Summer	Winter
Josbach/Hölzlebruck	*(1,0,0)/(1,0,0)*	(1,0,0)/(1,0,0)	*(1,0,0)/(1,0,0)*	*(1,0,0)/(1,0,0)*	(1,0,0)/(1,0,0)	(1,0,0)/(0,0,1)
Ammer/Pfäffingen	(2,0,1)/(1,0,0)	(1,0,0)/(1,0,0)	(0,δ,1)/(1,0,0)	(0,δ,0)/(0,δ,0)	(0,δ,0)/(0,δ,0)	(0,δ,0)/(1,0,0)
Erms/Riederich	(1,0,0)/(1,0,0)	(1,0,0)/*(1,0,0)*	(1,0,0)/(1,0,0)	(1,0,0)/(1,0,0)	*(1,0,0)/(0,0,1)*	(1,0,0)/(1,0,0)
Glatt/Hopfau	*(1,δ,0)/(1,0,0)*	(0,δ,1)/(1,0,0)	*(1,0,0)/(1,0,0)*	*(1,0,0)/(1,0,0)*	(1,0,0)/(1,0,0)	(1,0,0)/(1,0,0)
Wiese/Zell	(1,0,0)/(1,0,0)	(1,0,0)/(1,0,0)	*(1,0,0)/(1,0,0)*	(0,δ,0)/*(1,0,0)*	(1,0,0)/(1,0,0)	(0,δ,0)/(1,0,0)
Dreisam/Ebnet	(1,0,0)/(1,0,0)	(1,0,0)/(1,0,0)	(1,0,0)/(1,0,0)	(0,δ,0)/*(1,0,0)*	(0,δ,0)/(1,0,0)	(1,0,0)/(1,0,0)
Rems/Neustadt	(1,δ,0)/(1,δ,0)	(1,0,0)/(1,0,0)	(1,0,1)/(1,0,1)	(0,δ,0)/(1,0,0)	(1,0,0)/(1,0,0)	(0,δ,0)/(1,0,0)
Fils/Plochingen	(1,0,0)/(1,0,0)	(1,0,0)/(1,0,0)	(1,0,0)/(1,0,0)	(1,δ,0)/(1,0,0)	*(1,0,0)/(1,0,0)*	(1,0,0)/(1,0,0)
Neckar/Horb	(1,δ,0)/(1,0,1)	(1,0,0)/(1,0,0)	(1,0,0)/(1,0,0)	(0,δ,0)/(1,0,0)	(1,0,0)/(1,0,0)	(0,δ,0)/(1,0,0)
Neckar/Plochingen	(1,0,1)/(1,0,1)	(1,0,0)/(1,0,0)	(1,0,0)/(1,0,0)	(0,δ,0)/(1,0,1)	(1,0,0)/(1,0,0)	(0,δ,0)/(1,0,0)

filtered data. For the river Erms at Riederich (averaged summer data) the best HIC model for the filtered data is rejected by the goodness-of-fit test, whereas the best HIC model is accepted for the original data. So the trend test result of not detecting a significant trend is supported here. In this way goodness-of-fit results may be utilised to check HIC choices and trend test results.

The three relatively simple models presented in Sect. 2.3 perform best in terms of the HIC in the majority of cases, which is presented in Table 1. By studying these model fits conclusions about the short and long-term behaviour of the time series can be drawn. In particular the short-term correlated AR(1) model suffices to reproduce the dynamics of the fluctuations of the data best in a lot of cases. However, the long-term parameter δ is required more often for the monthly maxima data than for the monthly averages. The monthly maxima winter data is even represented best by a FD(δ) model in half of the cases, whereas a long-term parameter is nearly never required for the monthly average winter data. This may suggest that long-term correlation plays an important role when examining extreme values.

Regarding the trend behaviour a definitively significant trend has been found for Dreisam/Ebnet (complete data) and Ammer/Pfäffingen (winter data) for the monthly averages and Erms/Riederich (complete data) and Wiese/Zell (summer data) for the monthly maxima. In those cases a significant trend has been found for parameter estimations on Y, as presented in Table 1.

Where a possible trend component is found to be significant only under assumption of the trend estimate, i.e. the parameter values estimated on the filtered data, it cannot be excluded that this "trend" may be explained as well by stochastic variability. Such a spurious case, e.g. is the monthly averaged complete data for Ammer/Pfäffingen. Part of the variability of the data may be explained by a larger AR parameter or by a deterministic trend component. The goodness-of-fit

values also do not provide a distinction here. Regarding the monthly maxima a classical constellation occurs in several cases: part of the variability in the data may either be explained by a long-term correlated model or by a short-term model in combination with a deterministic trend component (see Giraitis et al., 2001). In those cases a refined parameter estimation by using a trend guess $\widetilde{T}$, as described above, may lead to further conclusions.

In the majority of cases a significant trend was found only under assumption of an AR(1) model (results not shown). This emphasises the influence of the model choice on the trend test result, which has been described in Sect. 2.5.

5 Conclusions

A semi-parametric trend test has been applied to discharge anomalies from Southern Germany. We considered FARIMA(p, δ, q) models optimised by a Whittle estimator with p, $q \leq 2$ (see Sect. 2.3). For each station the complete data as well as data restricted to the summer (April-September) and winter (October-March) months have been investigated separately, both for monthly averaged and maximum values.

It is evidenced that simple models seem to represent the dynamics of the data quite well in terms of the HIC. The short-term correlated AR(1) model is chosen as best fit in most cases. However, a long-term parameter $\delta > 0$ is frequently required, especially for the monthly maxima data. Therefore, a trend analysis should take the possibility of long-term correlated data into account.

The separation of deterministic trend and natural variability is of high interest to water management authorities. A way towards a distinction is the extraction of a trend estimate and the evaluation of the significance of the trend, as presented here. As described in Sect. 2.4, the acceptance or

rejection of a trend depends on the chosen stochastic model to represent the natural variability as well as on the magnitude of its parameter values. By considering the best performing models in terms of the HIC and by comparing the trend test results for the original and the filtered data (under consideration of the goodness-of-fit test outcome) reliable trend test results can be achieved in a series of cases, as listed up in Sect. 4. For the area under investigation, no general trend tendency for river run-off could be found. This is in line with other results (cf. Sect. 1), but nevertheless does not imply that climate change has no impact on the hydrological cycle. Actually one has to assume that climate change will be observable with a time delay and with complex patterns in hydrology.

By refining the trend estimation procedure (as also described in Sect. 4) and therefore reducing the bias in the parameter estimation, the trend test may be improved to handle not yet determinable cases. This issue will be subject to further work.

Summing up, we feel that the results and the applied technique open a promising road towards more insight in the dynamics of river run-off series. A test for trend is ameliorated by considering the correlation structure of the data. This might also be relevant for the investigation of extreme values like floods and droughts.

Appendix A: The discrete wavelet transforms

The DWT operates on dyadic time series and is defined in terms of a wavelet filter and an associated filter known as the scaling filter. Formally, a wavelet filter $\{h_{j,l} : l=0, ..., L_j - 1\}$ ($j=1, ..., J$ running over scale and l over time) is a sequence that sums to zero, has unit energy, and is orthogonal to its even shifts:

$$\sum_{l=-\infty}^{\infty} h_{j,l} = 0 \quad ,$$

$$\sum_{l=-\infty}^{\infty} h_{j,l} h_{j,l+2n} = \begin{cases} 1 \text{ if } n = 0 \\ 0 \text{ otherwise} . \end{cases} \quad (A1)$$

Let $h_{1,l} \equiv h_l$ and $g_{1,l} \equiv g_l$. The most practical wavelet filters have a finite width L, which means that $h_l=0$ for $l<0$ and $l>L-1$, while $h_l \neq 0$ for $0 \leq l < L-1$. The associated scaling filter g_l is constructed as follows: $g_l \equiv (-1)^{l+1} h_{L-1-l}$.

Appendix B: The Daubechies wavelet filters

The squared gain functions (i.e. the squared modulus of the transfer function) for all Daubechies wavelet filters are given by the form

$$\mathcal{H}_{1,L}(\omega) = \left| \sum_{l=0}^{L-1} h_l e^{-i2\pi\omega l} \right|^2$$

$$= 2 \sin^L \left(\frac{\pi\omega}{2} \right) \sum_{l=0}^{L/2-1} \binom{L/2-1+l}{l} \cos^{2l} \left(\frac{\pi\omega}{2} \right) \quad (A2)$$

on the first scale and $\mathcal{G}_{1,L}(\omega) = \mathcal{H}_{1,L}(\pi-\omega)$ holds, respectively. The transfer function for the filters is given by the polar representation

$$G(\omega) = \left| \mathcal{G}_{1,L}(\omega) \right|^{1/2} e^{i\theta(\omega)} \quad (A3)$$

and $H(\omega)=G(\omega+\pi)$, respectively. For different phase functions $\theta(\omega)$ distinct Daubechies wavelet filter families are defined. The least asymmetric (LA) family of Daubechies wavelet filters has a phase function such that these filters get as close as possible to a linear phase filter.

Edited by: M. Thiel
Reviewed by: one referee

References

Akaike, H.: Information theory and an extension of the maximum likelihood principle, in 2nd International Symposium on Information Theory, edited by Petrov, B. and Csaki, F., Akademie Kiade, Budapest, 267–281, 1973.

Akaike, H.: A Bayesian extension of the minimum AIC procedure of autoregressive model fitting, Biometrika, 26, 237–242, 1979.

Beran, J.: A goodness-of-fit test for time series with long range dependence, Journal of the Royal Statistical Society B, 54, 749–760, 1992.

Beran, J.: Statistics for Long-Memory Processes, Chapman & Hall, 1994.

Beran, J., Bhansali, R., and Ocker, D.: On unified model selection for stationary and nonstationary short- and long-memory autoregressive processes, Biometrika, 85, 921–934, 1998.

Bhattacharya, R., Gupta, V., and Waymire, E.: The hurst effect under trends, Journal of Applied Probability, 20, 649–662, 1983.

Bîrsan, M.-V., Molnar, P., and Burlando, P.: Streamflow trends in Switzerland, in: Paleofloods, Historical Floods and Climatic Variability: Applications in Flood Risk, edited by: Thorndycraft, V. R., Benito, G., Barriendos, M., and Llasat, M. C., Barcelona, 2003.

Bisaglia, L.: Model selection for long-memory models, Quaderni di Statistica, 4, 2002.

Bloomfield, P.: Trends in global temperature, Climatic Change, 1–16, 1992.

Box, G. and Jenkins, G.: Time Series Analysis: forecasting and control, Prentice Hall, New Jersey, 1976.

Bunde, A., Havlin, S., Koscielny-Bunde, E., and Schellnhuber, H.: Atmospheric Persistence Analysis: Novel Approaches and Applications, 2002.

Craigmile, P., Guttorp, P., and Percival, D.: Trend assessment in a long memory dependence model using the Discrete Wavelet Transform, Environmetrics, 15, 313–335, 2004.

DFO: Dartmouth Flood Observatory, Tech. rep., data available at: http://www.dartmouth.edu/~floods/, 2004.

Douglas, E. M., Vogel, R. M., and Kroll, C. N.: Trends in floods and low floods in the United States: impact of spatial correlation, Journal of Hydrology, 240, 90–105, 2000.

Fleming, S. and Clarke, G.: Autoregressive noise, deserialization, and trend detection and quantification in annual river discharge time series, Canadian Water Resources Journal, 27, 335–354, 2002.

Giraitis, L., Kokoszka, P., and Leipus, R.: Testing for long memory in the presence of a general trend, Journal of Applied Probability, 38, 1033–1054, 2001.

Granger, C. and Joyeux, R.: An introduction to long-memory time series models and fractional differences, J. Time Series Anal., 1, 15–29, 1980.

Hipel, K. and McLeod, A.: Developement in Water Science, Elsevier, Amsterdam, 1994.

IPCC: Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, 2001.

Kaiser, G.: A Friendly Guide to Wavelets, Birkhäuser, Boston, 1994.

Kendall, M.: Rank Correlation Methods, Griffin, London, 1975.

Koutsoyiannis, D.: Climate change, the Hurst phenomenon, and hydrological statistics, Hydrological Sciences Journal, 48, 2003.

Kundzewicz, Z., Graczyk, D., Przymusinska, I., Maurer, T., Radziejewski, M., Svensson, C., and Szwed, M.: Detection of change in world-wide hydrological time series of maxium annual flow, Tech. rep., Global Runoff Data Centre (GRCD), 2004.

Lawrence, A. and Kottegoda, N.: Stochastic modelling of river-flow time series, J. R. Stat. Soc., 140, 1–47, 1977.

Mallat, S.: A theory for multiresolution signal decomposition: the wavelet representation, IEEE Transaction on Pattern Analysis and Machine Intelligence, 11, 674–693, 1989.

Mann, H.: Nonparametric tests against trend, Econometrica, 13, 245–259, 1945.

Mileti, D. S. (Ed.): Disasters by Design: A Re-Assessment of Natural Hazards ind the United States, Joseph Henry Press, Washington, 1999.

Montanari, A., Rosso, R., and Taqqu, M.: Fractionally differenced ARIMA models applied to hydrologic time series: Identification, estimation and simulation, Water Resour. Res., 33, 1035, 1997.

Ooms, M. and Doornik, J.: Inference and forecasting for Fractional Autoregressive Integrated Moving Average models, with an application to US and UK inflation, econometric Institute Report 9947/A, 1999.

Percival, D. and Walden, A.: Wavelet Methods for Time Series Analysis, Cambridge University Press, 2000.

Percival, D., Overland, J., and Mofjeld, H.: Interpretation of North Pacific Variability as a short- and long-memory process, J. Clim., 14, 4545–4559, 2001.

Schwarz, G.: Estimating the dimension of a model, Ann. Statist., 6, 461–464, 1978.

Sibbertsen, P.: Robuste Parameterschätzung im linearen Regressionsmodell bei Fehlertermen mit langem Gedächtnis, Verlag für Wissenschaft und Forschung, 1999.

Smith, R.: Long-range dependence and global warming, Statistics for the Environment, 1993.

Taqqu, M. S., Teverovsky, V., and Willinger, W.: Estimators for long-range dependence: an empirical study, Fractals, 4, 785–798, 1995.

Torrence, C. and Compo, P.: A practical guide to wavelet analysis, Bulletin of the American Meteorological Society, 79, 61–78, 1998.

Woodward, W. and Gray, H.: Selecting a model for detecting the presence of a trend, J. Clim., 8, 1929–1937, 1995.

Yue, S. and Wang, C.: Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test, Water Resources Research, 38, 2002.

Zhang, X., Harvey, K. D., Hogg, W. D., and Yuzyk, T. R.: Trends in Canadian streamflow, Water Resources Research, 37, 987–998, 2001.