Trend Assessment of Correlated Data

Malaak Kallache

Henning Rust

Jürgen Kropp

Trend Assessment of Correlated Data *

Malaak Kallache [a], Henning Rust[a] and Jürgen Kropp[a]

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

Abstract The assessment of trends in meteorology and/or hydrology still is a matter of debate. Capturing typical properties of time series, like trends, is highly relevant for the discussion of potential impacts (e.g. global warming or flood occurrence). In order to enhance capabilities of analytical strategies run-off data from river gauges in southern Germany are analysed systematically regarding their trend behaviour. The trend is assumed to be a slowly varying deterministic component caused e.g. by human impact like global warming. Its detection is difficult since it might be superimposed by natural variability also present on large time scales. In an innovative approach a polynomial trend component and a stochastic model part are combined. With the stochastic model long-term and short-term correlations in time series data are considered. A reliable test for a significant trend can be performed via three steps: First, a stochastic fractional ARIMA model is fitted to the empirical data. In a second step, wavelet analysis is applied to separate the variability of small and large time-scales, assuming that the trend component is part of the latter. A comparison of the overall variability to that restricted to small scales results in a test for a trend. For the analysed series no significant trend could be found under the assumption of the models presented. The extraction of the large scale behavior by wavelet analysis provides valuable hints concerning the shape of the trend.

Keywords: time series analysis; trend test; wavelets; stochastic model; FARIMA

1 Introduction

Trend assessment is an important problem in time series analysis. Analysing hydrological and/or meteorological time series, one has to cope with the relative shortness of measured data in comparison to the time span of potential driving forces of variability. Also, as discussed by Beran [1994], it is possible to fit different stochastic models reasonably well to finite time series. The question, which model is most suitable, cannot be answered unambiguously. Often time series data do not only consist of a stochastic component, but also of a deterministic instationarity like a trend. In the context of climate and hydrological research this distinction might be transmittable to the determination of anthropogenic influence and natural variability.

In order to contribute to trend analysis in time series we decompose the data into variations on larger scales (further referred to as trend estimate $\hat{\mathbf{T}}$) and variations on smaller scales ($\hat{\mathbf{X}}$), which are assumed to be represented adequately by a linear stochastic model. Under this assumption the data are tested for an underlying deterministic trend. Since long-term correlations might cause long excursions from the mean as well as local trends (Beran [1994]), the detection of a significant deterministic trend there is more challenging. By comparing the models by means of the Bayesian Information Criteria (BIC) we want to contribute to the problem of distinction of trend and short memory on the one hand and long memory on the other hand (see also Giraitis et al. [2001]). The selected model also has an important influence on the confidence interval estimated for the trend parameters. The view of a trend as smooth deterministic changes on large scales is for example shared by Craigmile et al. [2003], Percival and Walden [2000] or Bloomfield [1992].

We have organised this paper as follows: in section 2 the applied methods, i.e. wavelet analysis, the Whittle estimator for FARIMA models and the trend test itself, are briefly discussed. Section 3 introduces the data, and in section 4 and section 5 the results are presented and discussed.

*contact: malaak.kallache@pik-potsdam.de

2 METHODS

2.1 Wavelet Analysis

Wavelet analysis is carried out by applying the discrete wavelet transform (DWT) and the maximum overlap DWT (MODWT). The DWT is an orthonormal transform. The time series are reconstructed by a linear combination of wavelets, analogous to a reconstruction by sinusoids in Fourier analysis. Because each wavelet is essentially non-zero only within a finite interval of time, time-scale analysis can be achieved. In this manner information about the variations local in time and scale can be retrieved. Furthermore, the wavelet basis is dilated when processing larger scales. Thus the problem of under or overlocalisation, which occurs by using the windowed fourier transformation, is minimised (see Kaiser [1994]).

The selection of the mother wavelet is data dependent. Possible choices are an orthonormal, a non-orthonormal, a real, or a complex basis. An overview about the noteworthy aspects is given in Torrence and Compo [1998]. Here, we find a discrete, orthogonal basis – the Daubechies "least asymmetric" wavelet basis of width eight (LA(8)) – to be an appropriate choice. 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. Let a time series be a realization of a stochastic process $\mathbf{Y}_t = \{Y_t\}$ with random variables Y_t , $t = 0, ..., N-1$. Applying the LA(8) wavelet filter to $\mathbf{Y}_t$ essentially yields a difference between Y_t and values before and after Y_t. The LA(8) scaling filter yields a weighted average of length two on the unit scale.

The MODWT (maximum overlap DWT) is a modified version of the DWT. It is a highly redundant nonorthogonal transform and time series of non-dyadic length can be analysed with it. For more details refer to Percival and Walden [2000].

2.2 Parameter Estimation

In order to classify trends as significant, a model is fitted to empirical data representing natural variability. A canonical class of linear models is given by the autoregressive integrated moving average (ARIMA) models (see Box and Jenkins [1976]). This model category, however, is not suitable to model long-range correlations. Long-range correlation or long-term memory is present if the autocorrelation function $\rho(k)$ decays algebraically for large time lags k, i.e.

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

with c being a finite constant. This results in $\sum_{-\infty}^{\infty} \rho(k)$ diverging. On the other hand, for short-range correlations, the autocorrelations decay exponentially and are thus summable. Several authors report long-term memory being present in river runoff records, e.g. Lawrence and Kottegoda [1977]; Montanari et al. [1997]. To include the possibility of of long-range correlations, the model class is extended to fractional ARIMA (FARIMA) models (Granger and Joyeux [1980]) allowing a flexible description of short-range and long-range correlated data.

For the records analysed, we considered different fractional ARIMA models including one to three parameters. On the basis of the BIC and a goodness-of-fit test the following three models are found to be appropriate for a comparison of the data sets under investigation: An AR(1) model given by

$$x_t = \phi x_{t-1} + \eta_t \tag{2}$$

with $\eta_t \sim$WN(0,σ), an FD(δ) model

$$(1-\mathrm{B})^{\delta} x_t = \eta_t, \tag{3}$$

with $\mathrm{Bx_t} = \mathrm{x_{t-1}}$, and finally a FARIMA(1,δ,0) model

$$(1-\mathrm{B})^{\delta} x_t = \phi x_{t-1} + \eta_t. \tag{4}$$

For the estimation of the model parameters Whittle's approximation to the maximum-likelihood estimator is used. The latter is based on minimising

$$Q(\theta) = \sum_j \frac{I(\omega_j)}{f(\theta;\omega_j)}, \tag{5}$$

where $I(\omega_j)$ denotes the periodogram of the data at the Fourier frequencies $\omega_j = 2\pi j/N$ and $f(\theta;\omega_j)$ the spectral density of the FARIMA process. The vector $\theta = (\phi, \delta)$ contains the model parameters. In this formulation the variance of the residuals is rescaled to one. At the minimum of $Q(\hat{\theta})$, $\hat{\theta}$ is an estimate of the model parameters. An extensive discussion on FARIMA models, the Whittle estimator and an implementation of the algorithm can be found in Beran [1994].

The influence of a strong deterministic trend component will bias the parameter estimation. Considering this effect, the parameter estimation is performed also for the filtered set of data with the variations on large scales removed by wavelet filtering.

2.3 Trend Estimation and Test for Significance

The time series $\mathbf{Y}_t$, is regarded as being composed of a stochastic process $\mathbf{X}_t$ and a deterministic trend component $\mathbf{T}_t$: $\mathbf{Y}_t = \mathbf{T}_t + \mathbf{X}_t$. By using the DWT, the data vector $\mathbf{Y}$ is decomposed in a component $\hat{\mathbf{T}}$, representing the variability on large scales, and a component $\hat{\mathbf{X}}$ for small scales: $\mathbf{Y} = \hat{\mathbf{T}} + \hat{\mathbf{X}}$. $\hat{\mathbf{T}}$ captures the deterministic trend as well as the stochastic variability on large scales and is referred to as the trend estimate. The separating scale l_s between $\hat{\mathbf{T}}$ and $\hat{\mathbf{X}}$ is selected ensuring that enough data points are left unaffected by the boundaries, see Craigmile et al. [2003]. The choice of l_s influences the test result: shifting l_s for one scale might change result of a trend test (see Percival and Walden [2000]). In the following all investigations are carried out with $l_s = 5$, which corresponds to a time scale of 64 months. The boundary conditions are assumed to be periodic. To minimise the boundary effects, the series is padded at the end with the mean (see Torrence and Compo [1998]). Applying the trend test requires the d-th backwards difference of the stochastic component to be stationary.

For a selected l_s the model choice and the magnitude of the model parameters do not affect the shape of the trend. However, the parameters estimated for the stochastic model influence the confidence band of the trend (see Figure 2) and the variance of the trend estimate involves the covariance of $\mathbf{X}_t$. If $\mathbf{X}_t$ is a realization of a stationary stochastic process, the autocovariance sequence (ACVS) of $\mathbf{X}_t$ can be used to calculate the variance of the trend.

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

$$p_c = \frac{||\mathbf{Y}_t||^2}{||\hat{\mathbf{X}}_t||^2} \tag{6}$$

be the test statistic, where $||.||$ denotes the euclidian norm and $\mathbf{Y}_t$ has zero mean. H_0 is now: $T_t = 0$ $\forall t = 0, ..., N-1$ versus H_1: not H_0. For $\mathbf{T}_t \neq \{0\}$, p_c should be large. H_0 is rejected at a level of significance α if $p_{c(empirical)}$ of $\mathbf{Y}_t$ 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 simulations with 4000 runs of the stochastic model, which include the optimised parameters but no trend.

Examining the power of the trend test, the following results have been obtained: analysing linear trends the power of the trend test is slightly lower than the power of a standard linear regression test, as stated in Craigmile et al. [2003]. The power of the trend test does not weaken significantly, when the linear trend starts in a later part of the time series. It reaches a power of one faster for small δs than for large ones. The trend test is robust against changes in variance of the time series, but it is vulnerable against jumps in the data. So the occurrence of breakpoints should be excluded.

2.4 Goodness-of-fit and model selection

A goodness-of-fit test is used to test whether a model yields a valid description of the data. We applied a test proposed by Beran [1992], which compares the periodogram of the empirical data and the spectral density function of the fitted model. This test is basically a formulation of the portmanteau test for uncorrelated residuals. The smallest significance level for which the null hypothesis H_0: "the empirical data is compatible with being a realization of the model" is falsely rejected is denoted by α_{crit}.

As mentioned the Bayesian Information Criterion (BIC) is used to compare the performance of different models. If the true model is among the models explored, the BIC is minimal for the proper model. In a simulation study, Bisaglia [2002] has shown that also for the long-range correlated FARIMA models the BIC is a consistent selection criterion.

3 Data

For the analysis river discharge records from several catchments near the river Neckar in southern Germany have been investigated. The series were selected according to their length and completeness. The run-off data jointly covers the time period from November 1940 to October 1996. The presented analyses have been carried out for this period and the complete time period available for each single time series. All data sets used are affected by a strong periodic component due to the annual cycle. This is approximately removed by calculating the daily average and variance over all years , where missing values are replaced by the average for the specific day. Subsequently the anomaly is obtained by subtracting the average and dividing by the standard deviation from the measured data. Dividing by the standard deviation removes periodicity in the variance, see e.g. Hipel and McLeod [1994]. The daily measurements have been aggregated to monthly values.

4 RESULTS

The obtained parameter values δ and ϕ for the FD(δ), AR(1) and FARIMA(1,δ,0) models are shown in Figure 1 for the time span from 1940 to 1996. In comparison to other models including one to three parameters, they were selected as best fitting models by the Bayesian Information Criterion (BIC). Furthermore conclusions about the short and long-term behaviour of the time series can be drawn by studying these three model fits.

Considering the BIC, mostly the AR(1) model was determined as best fit. For Horb/Neckar, Plochingen/Neckar and Neustadt/Rems, the FD(δ) or the FARIMA(1,δ,0) was selected as best model. Analysing the complete time period covered reveals a shift towards long-term correlated models by the BIC best choice. As shown in Figure 1 B applying a FARIMA(1,δ,0) model results in smaller values of δ and ϕ, which in the case of Pfäffingen/Ammer leads from an instationary FD(δ) model to a stationary FARIMA(1,δ,0) model. Apparently, using the FARIMA(1,δ,0) model, also leads to an estimation of the long-range parameter δ near zero at 5 stations. This suggests that long-range correlation is not relevant in that case. The winter data (October to March, not shown) contains only 3 gauges with δ being significantly different from zero, which indicates that for the winter period long-range correlations are not a dominant characteristic.

The applied trend test did not reveal a significant trend for any of the time series, regardless whether the full year or only summer (April to September) or winter (October to March) are considered. The negative trend result is the same for both the jointly covered and the entire time period available for each single series.

Since the fitting algorithm is sensitive to the trend component, the results are compared to the trend test on the filtered time series, where the variations on large scales are removed. Here, not only the trend component is filtered out but also the stochastic variations on large scales. Thus parameter optimisation with respect to the long-range correlated components will be affected. Using parameter fits from the filtered series does also not reveal a significant trend. On the other hand, using parameter fits on the wavelet coefficients – which are unaffected by polynomial trends of order three – yields to systematically lower estimated parameters for the FD(δ) and AR(1) model. Here, a significant trend was found only under assumption of an AR(1) model for Pfäffingen/Ammer, Hopfau/Glatt

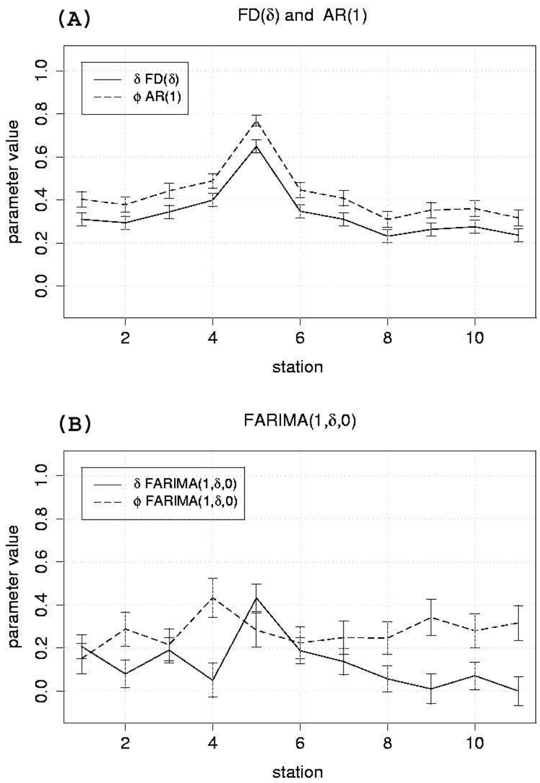

Figure 1: (A) Fitted parameters for the FD(δ) and the AR(1) model and (B) the FARIMA(1,δ,0) model. River Gauges: 1 Neustadt/Rems, 2 Plochingen/Fils, 3 Plochingen/Neckar, 4 Riederich/Erms, 5 Pfäffingen/Ammer, 6 Horb/Neckar, 7 Hopfau/Glatt, 8 Oberwolfach/Wolf, 9 Hölzlebruck/Josbach, 10 Ebnet/Dreisam, 11 Zell/Wiese.

and Oberwolfach/Wolf. This suits the finding that standard errors for the trend fit obtained under an assumption of long-range dependence can be considerably larger than those obtained under an short-range correlated autoregressive model (see Smith [1993]) The broadened confidence interval for estimated trend parameters can thus lead to an acceptance of the hypothesis that there is no trend in case of long-term correlations, whereas this hypothesis is rejected in case of short-term correlations (see Beran [1994]). In Figure 2 the different confidence bands for the trend estimation under different model assumptions are shown.

These results differ from KLIWA [2003], where a significant trend component was found at Plochingen/Fils, Riederich/Erms, Pfäffingen/Ammer and Hopfau/Glatt by using the Mann-Kendall test. Further work will focus on a better understanding of this detail.

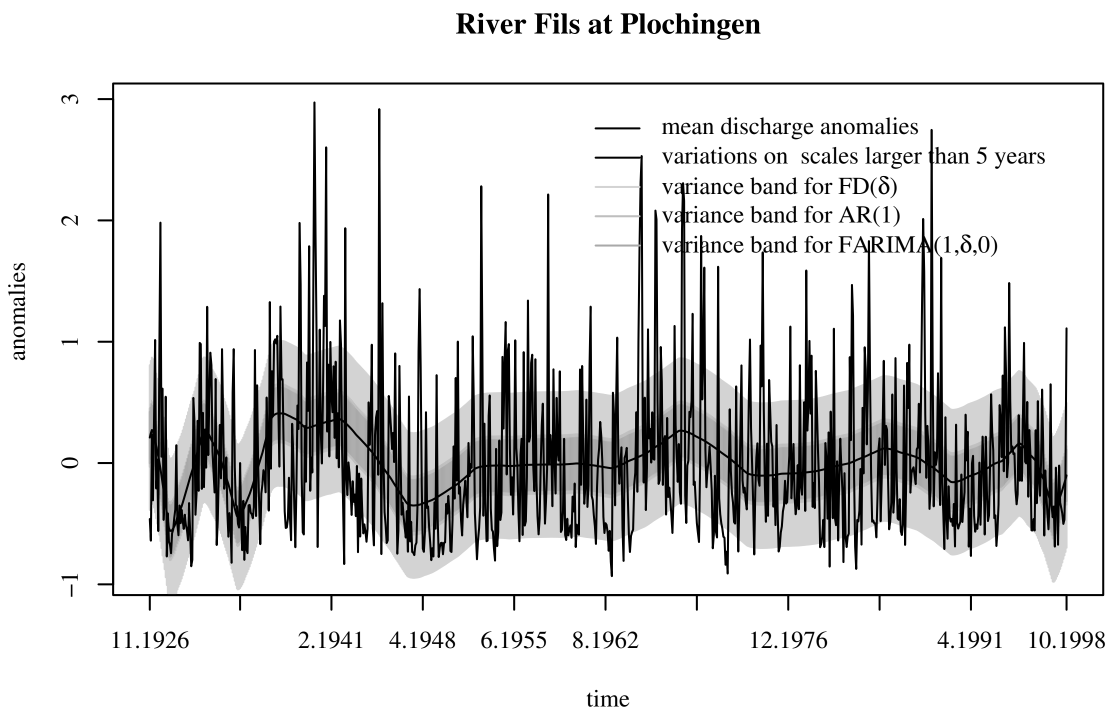

Figure 2: Normalised mean discharge anomalies for the river Fils at Plochingen and trend estimate. The 95% confidence band for the variance of the trend estimation is drawn under assumption of a FD(δ), an AR(1) and a FARIMA(1,δ,0) model.

In table (1) the goodness-of-fit test results $100\alpha_{crit}$, for both, the complete data and the filtered data, are listed. For $\alpha = 0.05$ the null hypothesis of the empirical data being a realization of the fitted models should be rejected when $100\alpha_{crit} \leq 5$. Apparently a failed goodness-of-fit test coincidences with the best BIC value in some cases. The relative closeness of α_{crit} for the full dataset and the filtered data indicates, that the filtering routines has to be ameliorated to enhance the fitting routine. This is subject of further work.

5 Conclusions

A model dependent trend test has been applied to mean discharge anomalies from southern Germany. We considered an AR(1), FD(δ) and FARIMA(1,δ,0) model with parameters optimised by a Whittle estimator. For each station the full time series as well as the summer (April-September) and winter (October-March) components have been investigated separately. For the time period jointly covered by all data sets (1940-1996) and for the full length of each record no significant trend could be found for the three models assumed. Since the Whittle estimator is biased for an underlying trend, the parameter estimation has been repeated for data with the variability on large scales removed by wavelet filtering. The result remained unaffected, no significant trend could be found. Preliminary investigations using parameter estimation on the wavelet coefficients yield smaller values for the AR(1) and FD(δ) parameter. In this case for Pfäffingen/Ammer, Hopfau/Glatt and Oberwolfach/Wolf a trend has been detected as being significant using an AR(1) model. This will be subject of further investigation, especially in the context of positive trend results obtained with a Mann-Kendall test reported in KLIWA [2003]. The separation of deterministic trend and natural variability is of high interest to water management authorities.

The shape of the estimated trends $\hat{\mathbf{T}}$ of the analysed hydrological data do not reveal a monotonic trend, but rather do contain segments of increase and decrease. This indicates, that for trend tests the analysed time period might be crucial.

The correlation structure – e.g. long-range or short-range correlations – of river-runoff has important consequences for the investigation of extreme val-

	FD(δ)			AR(1)			FARIMA(1,δ,0)		
	Total	Summer	Winter	Total	Summer	Winter	Total	Summer	Winter
Neustadt/Rems	**01**/01	84/85	**38**/34	09/**08**	**94/96**	32/**27**	07/07	94/95	37/29
Plochingen/Fils	05/05	77/78	19/20	**43/43**	**93/ 95**	**27/29**	40/42	93/94	27/29
Plochingen/Neckar	18/17	46/45	32/28	61/**55**	**78/79**	**34/26**	**58**/57	77/77	34/26
Riederich/Erms	17/18	41/43	37/34	**88/86**	**87/86**	**45/39**	87/86	87/86	45/39
Pfäffingen/Ammer	54/53	54/52	79/78	**46/40**	**80/76**	**68/61**	61/59	80/76	72/61
Horb/Neckar	21/20	24/24	28/25	**33/25**	**52/52**	**33/25**	**45**/41	51/51	33/25
Hopfau/Glatt	00/00	65/67	02/01	**08/12**	**51/61**	**14/13**	10/11	71/73	14/13
Oberwolfach/Wolf	03/02	38/38	01/00	**34/24**	**36/38**	**10/05**	27/24	35/35	10/05
Hölzlebruck/Josbach	00/00	64/63	00/00	**19/15**	**82/83**	**07/04**	19/15	83/83	07/04
Ebnet/Dreisam	07/04	10/08	01/00	**12/25**	**11/22**	**15/13**	20/25	15/22	15/13
Zell/Wiese	15/13	04/04	00/00	**47/43**	**22/28**	**00/00**	47/43	20/28	00/00

Table 1: $100\alpha_{crit}$ values from the goodness-of-fit test are listed. Values before the backslash denote results for all data analysed in the jointly covered time period from 1940 to 1996. Values after the backslash denote results for the filtered data, where variations on large scales have been eliminated before fitting the model parameters to exclude a bias due to a possible trend. Bold values indicate the model with best BIC results.

ues like floods and droughts. According to the model selection criterion (BIC) for most of the gauges considered an AR(1) process is suggested as being sufficient to describe the runoff-anomalies. This seems to contrast with Hurst coefficient larger 0.5 found frequently for run-off anomalies. Investigations in this direction with an expanded model canon and various selection criteria will be carried out in forthcoming works.

References

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.

Bisaglia, L. Model selection for lon-memory models. *Quaderni di Statistica*, 4, 2002.

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

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

Craigmile, P., P. Guttorp, and D. Percival. Trend assessment in a long memory dependence model using the discrete wavelet transform, 2003. in press.

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

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

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

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

KLIWA. Langzeitverhalten der hochwasserabflüsse in baden-württemberg und bayern. Technical report, 2003.

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

Montanari, A., R. Rosso, and M. Taqqu. Fractionally differenced arima models applied to hydrologic time series: Identification, estimation and simulation. *Water Resour. Res.*, 33(5):1035, 1997.

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

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

Torrence, C. and P. Compo. A practical guide to wavelet analysis. *Bulletion of the American Meteorological Society*, 79(1):61 – 78, 1998.