Hydrol. Earth Syst. Sci. Discuss., 3, S1598–S1601, 2006
www.hydrol-earth-syst-sci-discuss.net/3/S1598/2006/

Hydrology and Earth System Sciences Discussions

HESSD

3, S1598–S1601, 2006

Interactive Comment

Interactive comment on "Detecting long-memory: Monte Carlo simulations and application to daily streamflow processes" *by* W. Wang et al.

W. Wang et al.

Received and published: 4 December 2006

We would like to thank the comments from both reviewers for improving the manuscript, especially the critical comments from referee #1. The reply to the general comments of two referees (Hydrol. Earth Syst. Sci. Discuss., 3, S765-S766, 2006; Hydrol. Earth Syst. Sci. Discuss., 3, S599-S603, 2006) is given in the discussion paper (Hydrol. Earth Syst. Sci. Discuss., 3, S805-S810, 2006). Many major revisions are made according to the comments of the reviewers, including the methods and some conclusions.

Here is the reply to the second interactive comment of referee #1 (Hydrol. Earth Syst. Sci. Discuss., 3, S811-S812, 2006):

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

(1) Indeed, it would be better to use several estimators so as to increase the reliability of the estimation due to the uncertainty in the detection of long-memory, especially for the analysis of daily streamflow observations. This is a limitation of the present study which has been put in the conclusion part. But in any research, we have to make a choice among various available methods. As the three methods we used here in the first manuscript are very commonly used in literature, even used in isolation sometimes, and the results given by the three methods are generally acceptable, therefore, we have not considered applying more methods. In addition, some other methods may work well in many cases, but may not be suitable in Monte Carlo simulations. For instance, with the aggregated variance method which is recommended by Koutsoyiannis (2002), we usually plot the logarithm of the variance of an aggregated (averaged) process versus the logarithm of the aggregation level, and then fit a least-squares line to the data, and the slope provides an estimate of d. But this is a heuristic method. In practice, we should exclude the points at the very low and high ends of the plots, because the short-range effects can distort the estimates of d if the low end of the plot is used, and at the very high end of the plot there are too few blocks to get reliable estimates of the variance. That makes it less suitable to be employed in the Monte Carlo simulations.

(2) Following the suggestion by the reviewer, we have added the results with the Whittle's estimator as an additional method to detect the existence of long-memory. The results with the S-MLE method are still preserved so as to make an additional comparison between the two MLE methods, and they are re-calculated for 500 simulations. Our Monte Carlo simulations show that the S-MLE can give reasonable results despite of its disadvantages mentioned by the reviewer over the Whittle's estimator. In fact, the algorithm of Haslett and Raftery (1989) is no less popular than the Whittle's method in literature. It has been cited over 1000 times since its publication. Its popularity can also been seen in S+FinMetrics, which is a module of S-Plus focusing on financial time series analysis. In the development of S+FinMetrics, Jan Beran who developed the code of Whittle's method in his monograph on long-memory processes (Beran, 1994) wrote many of the long memory functions while acting as a consultant to Insightful Corpora-

HESSD

3, S1598–S1601, 2006

Interactive Comment

tion (Zivot and Wang, 2003, Preface). However, the S+FinMetrics function FARIMA() which is the major function for fitting the ARFIMA model to a univariate time series is still implemented based on S-Plus function arima.fracdiff() (Zivot and Wang, 2003, Chapter 8.5), namely, the S-MLE method, whereas the Whittle's method is used for estimating the fractional integration parameter only for the case of the ARFIMA(0, d, 0) process (Zivot and Wang, 2003, Chapter 8.5).

(3) As pointed out by the reviewer, the minimal sample size depends on the behaviour of the time series, and it may be not possible to provide a general rule, but it would be helpful if we can find some rules for some specific cases. Our simulation results indicate that the four methods we used here may work well or poorly depending on different cases with different intensity of long memory, different intensity of short memory, and different data sizes. We hope that the results could be a helpful supplement to theoretical results, and could give at least a rough view when applying those methods to real-world applications. One example of this point is about the use of the GPH test. Agiakloglou et al. (1993) found that GPH estimators performed poorly for AR(1) processes with phi = 0.9 for sample size of 100 to 900. The simulation results of Hurvich and Beltrao (1993) also showed the poor performance of the GPH estimator when phi = 0.9 for not only AR(1) processes but also ARFIMA(1,d,0) processes. In our simulation study, it is shown that, on one hand, the power of GPH test does decrease with the increase of the autoregressive coefficient; on the other hand, its power increases with the increase of sample size. If we use a sample size of larger than 104 points, GPH test still performs very well for AR(1) processes with phi = 0.9. But the use of GPH test is helpless when f is larger than 0.95, even with a data size of larger than 10ˆ4. Another example is the result with the Whittle's estimator which is strongly recommended by the reviewer. As far as we know, it has not been reported in previous literature that Whittle's estimates would be negatively biased when there is a weak AR component in the ARFIMA(1,d,0) process (phi $<=$ 0.5), whereas positively biased when there is a strong AR component (phi = 0.9) for comparatively small data size ($<$ 3x10ˆ3). We wish that these results would more or less help when using these methods for real-world time

HESSD

3, S1598–S1601, 2006

Interactive Comment

series.

Finally, the reply to the technical comments of both reviewers:

1. To Reviewer #1: (1) Spelling errors are corrected as far as we found. (2) Acronyms are defined before their first use now. (3) The linkage between d and H is given in Section 1. (4) Sample data size is restricted because usually we cannot find data series in the real-world of unlimited size. Observed time series are of limited size because of several criteria I mentioned in Section 4.

2. To Reviewer #2: (1) In Eq.(1), Gamma function is defined, but without illustration so as to save space. (2) An introduction is given to the ARFIMA model in Section 1. (3) H is defined in Section 1. (4) By saying that the data should be recorded as early as possible, we mean that we choose the period of the river flow series as early as possible in the 20th century or even 19th century, assuming (maybe not a good assumption) that human intervention is generally less stronger in early period than in later period. (5) The deseasonalization procedure is described in more detail. (6) All the typing errors pointed out by the reviewer are corrected.

References

Agiakloglou, C., Newbold, P., Wohar, M.: Bias in an estimator of the fractional difference parameter. J. Time Ser. Anal., 14, 235-246, 1993. Beran J.: Statistics for Long-Memory Processes. Chapman & Hall, New York, 1994. Hurvich, C.M. and Bellrao, K.: Asymptotics for the low-frequency ordinates of the periodogram of a long-memory time series. Journal of Time Series Analysis, 14, 455-472, 1993. Koutsoyiannis, D.: The Hurst phenomenon and fractional Gaussian noise made easy, Hydrol. Sci. J., 47(4), 573-595, 2002. Zivot, E., Wang, J.: Modeling Financial Time Series with S-Plus. New York: Springer Verlag, 2003.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 3, 1603, 2006.

HESSD

3, S1598–S1601, 2006

Interactive Comment