Monte-Carlo Testing for Preservation of Historical Statistics in Synthetic Time Series

Ian McLeod

University of Western Ontario, Canada

January 4, 2012

Please see:
http://www.stats.uwo.ca/faculty/aim/TIES2012/
for pdf copies of this presentation and papers referred to therein.

Introduction

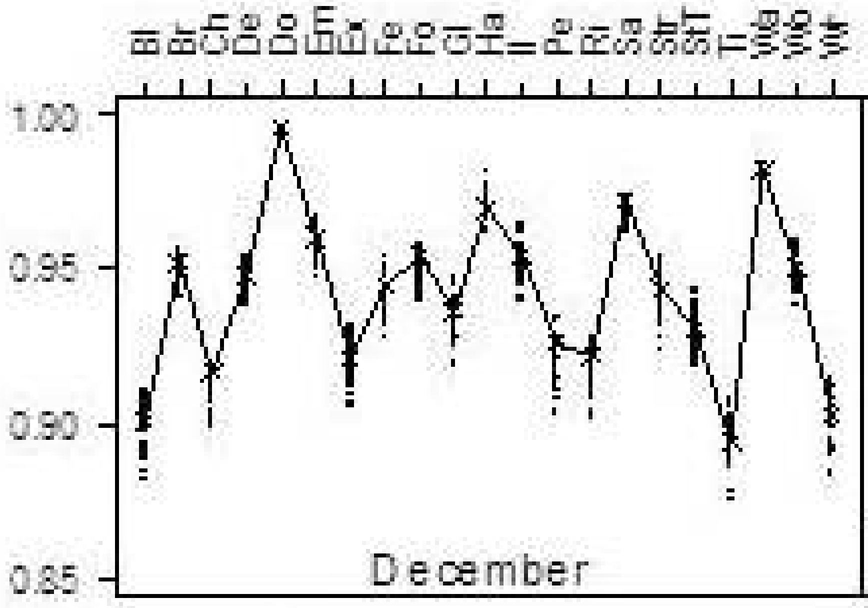
- Simulated or synthetic weather and hydrological time series used for planning
- Does the model adequately capture important characteristics of the series: seasonal means and variances, long-range dependence, critical drought, etc.
- informal and formal testing

Weather Generator, Downscaling, Climate Change[1]

- Empirical nearest neighbour model for weather simulation.
- Daily time series for 22 weather stations in southwestern Ontario (Blyth,Brantford,Dorchester,Embro,Exeter,Foldens,Glen Allan,Ilderton,London A, etc.)
- 7 weather variables (precipitation, max temperature, min temperature, humidity, wind speed (eastward, northward), pressure.
- Long time series from January 1, 1979 to December 31, 2005 (length 9862).
- Are temporal and spatial correlations preserved as well as seasonal means, variances, extremes, etc.

[1] with Leanna King and Dr. Simonovic

Spatial Correlation with London A



Formal Significance Test for Preservation

- given data, $\mathcal{D}$
- statistic of interest, $K_{\text{obs}} = K(\mathcal{D})$
- model fitted to data, $\hat{\mathcal{M}}$
- $\mathcal{H}_0 : K_{\text{obs}} \sim K(\hat{\mathcal{M}})$, where $K(\hat{\mathcal{M}})$ denotes the distribution of the statistic K if the fitted model is correct.

Similarity with Diagnostic Checks

It is standard practice to check the adequacy of the a fitted statistical model using a significance test.

These tests can be tricky!

For example if we fit $z_t = \phi z_{t-1} + a_t, t = 1, \ldots, n$, where $a_t \sim \mathrm{NID}(0, 1)$, we may use the lag one residual autocorrelation, $\hat{r}_{\hat{a}}(1)$ is the test statistic.

But,

$$\hat{r}_{\hat{a}}(1) = \phi^2/n.$$

Not as might be expected since

$$\hat{r}_a(1) = 1/n.$$

Residual Autocorrelation Test (con't)

If we knew the model then we could obtain $a_t, t = 1, \ldots, n$. In this case $\hat{r}_a(1)$ is normal with mean zero and variance $1/n$ but it turns out that $\hat{r}_{\hat{a}}(1)$ is normal mean zero and variance ϕ^2/n. If wrong variance is used, test is much too conservative.

This occurs because $\hat{\phi}$ and $\hat{r}_{\hat{a}}(1)$ may be highly correlated asymptotically!

This phenomenon is very general and occurs for other statistics such as Hurst's K and the critical period statistics.

Monte-Carlo Test

- Fit model, $\hat{\mathcal{M}}$. Compute observed test statistic, K_{obs}.
- Set $N \longleftarrow 1000$, number of MC iterations. Repeat next step N times.
- Simulate $\hat{\mathcal{M}}$ and compute value of K with new data (fitting model again if needed).
- Compute two-sided p-value,

$$p = (N+1)^{-1}(1 + \sum I(|K| > |K_{\mathrm{obs}}|)).$$

Informal Monte-Carlo Test

Just using a few simulations and comparing the null distribution of the test statistic with the observed value is often adequate as in the spatial correlation with the weather generator.

Monte-Carlo Test Theorem

Dufour (2006, Proposition 2.6) establishes that under regularity conditions the usual Monte-Carlo test procedure provides a valid test in the presence of a nuisance parameter provides that consistent estimators are used.

Dufour, J.-M. (2006). Monte Carlo Tests with Nuisance Parameters: A General Approach to Finite-Sample Inference and Non-Standard Asymptotics. *Journal of Econometrics*, 113/2, 443-477.

Monte-Carlo Test Theorem: Applications in Diagnostic Checking

Mahdi & McLeod (2011) and references therein discuss many applications of this theorem to portmanteau residual diagnostic checks in time series.

Mahdi, E. and McLeod, A.I. (Online, 05/09/2011). Improved Multivariate Portmanteau Diagnostic Test. Journal of Time Series Analysis. DOI: 10.1111/j.1467-9892.2011.00752.x.

Hurst Phenomenon & Long Memory

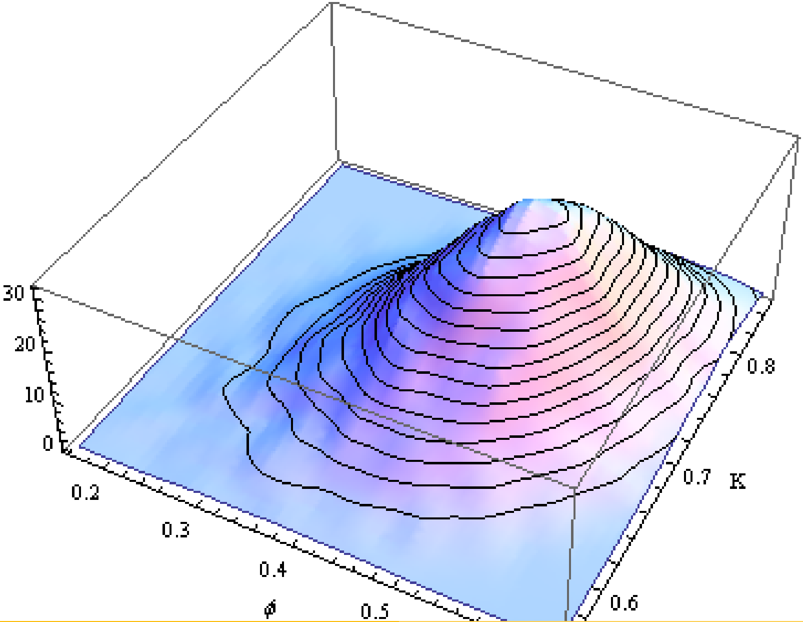
- Hurst (1951,1957) fit the equation $R_n = (n/2)^K$, where R_n is the rescaled cumulative range to 690 geophysical time series.
- Obtained $K = 0.73$ with $\sigma_K = 0.09$ based on these 690 series.
- W. Feller & others noted $K \asymp 0.5$ for all ARMA-like time series.
- Long-memory time series can be used. More complicated and computationally difficult models.

Hurst, H.E. (1957). A Suggested Statistical Model of some Time Series which occur in Nature. *Nature* 180, 494-494 doi:10.1038/180494a0

Hurst Phenomenon & AR(1)

- Simulated 25,000 replications of an AR(1) with $\phi = 0.5$ and $n = 100$.
- Obtained joint distribution for $\hat{\phi}$ and K
- Mean and standard deviation of K are 0.73 and 0.06 respectively.

Empirical Distribution $\hat{\phi}$ and K



Preservation of Hurst's K & Critical Period Statistics

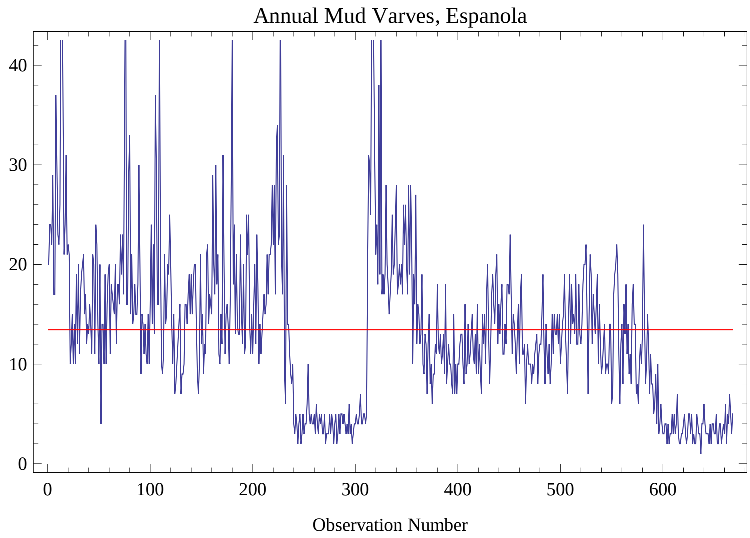
Simple ARMA models were fit to 22 annual geophysical time series and the MC test for preservation of Hurst's K suggested no evidence of lack of fit.

Similarly we showed that PAR models preserve the critical period statistics in monthly riverflows.

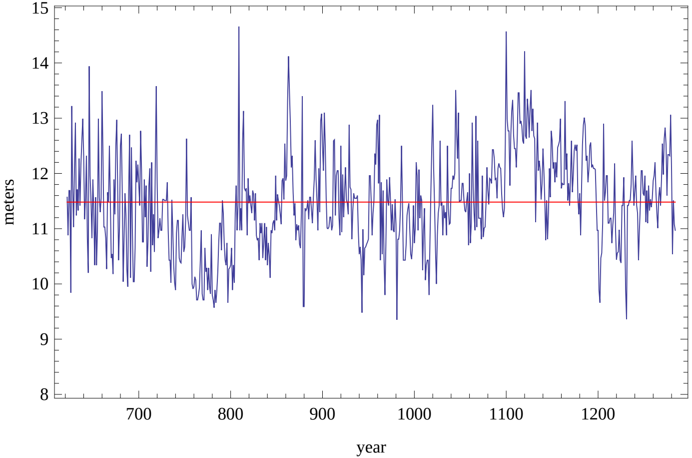
Hipel, K.W. & McLeod, A.I. (1994), *Time Series Modelling of Water Resources and Environmental Systems*, Elsevier.
`http://www.stats.uwo.ca/faculty/aim/1994Book/`

Espanola Mud Varves, Annual, $K = 0.84$

Out[68]=



Nile River Minima, 622 to 1284



Best Model, Nile River Minima

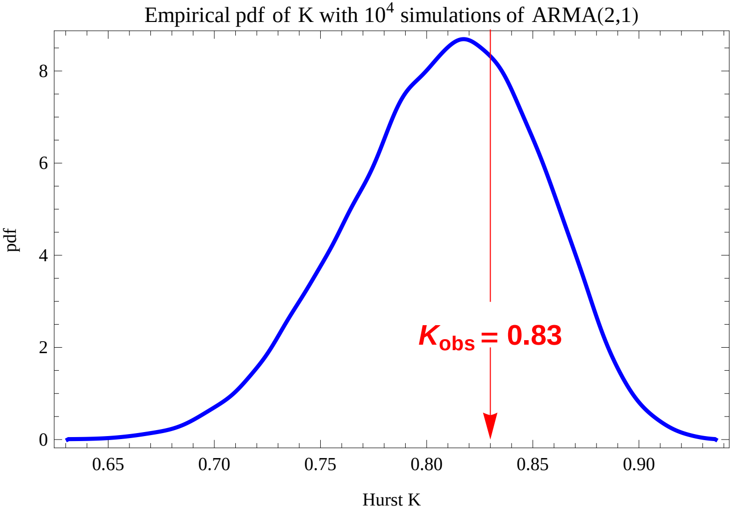
	Loglikelihood	AIC	BIC
ARMA(2,1)	237.61	-469.22	-455.73
FGN(mle)	236.52	-471.04	-466.54

Table: AIC/BIC for best models

The series length is $n = 663$ so the BIC penalty for parameters is $6.5k$ as opposed to $2k$ for the AIC.

Distribution of Hurst K in Fitted ARMA(2,1)

Out[116]=



Dynamic Visualizations

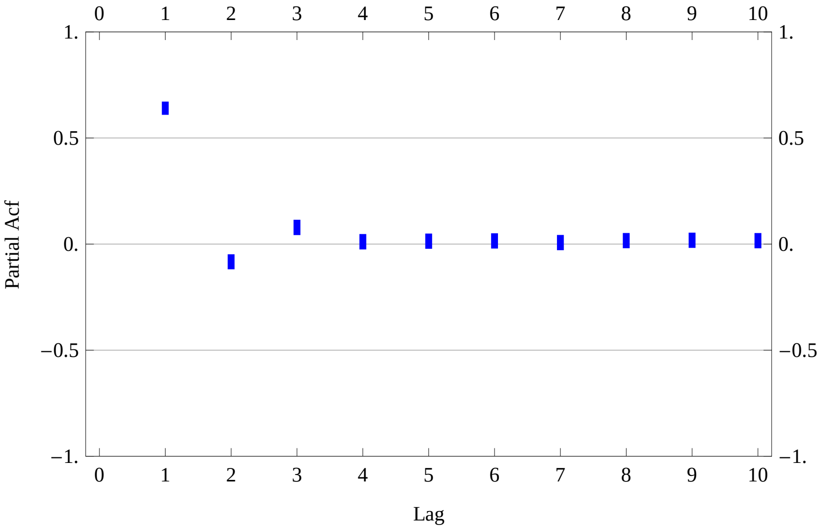
Available using *Mathematica's* cdf format and in the Wolfram Demonstrations.

- Visual Comparison of FGN and ARMA(2,1) for Nile Minima Series
- Visualizing Fractional Gaussian Noise

Daily Max Temperature Series

- Long time series from January 1, 1979 to December 31, 2005 (length 9862).
- Detrended by fitting harmonic regression with period 365.25.
- New PACF plot presented next.

New PACF Plot



References for New PACF Plot

McLeod, A. I., Ying Zhang & Changjiang Xu (2011), FitAR: Subset AR Model Fitting. R package version 1.92. `http://CRAN.R-project.org/package=FitAR`

McLeod, A.I. & Zhang, Ying (2008). Improved Subset Autoregression: With R Package. *Journal of Statistical Software* 28/2, 1-28. http://www.jstatsoft.org/v28/i02.

McLeod, A.I. & Zhang, Y. (2006). Partial Autocorrelation Parameterization for Subset Autoregression. *Journal of Time Series Analysis*, 27, 599-612.

Long-memory in financial time series?

Couillard, M. and Davison, M. (2005). A comment on measuring the Hurst exponent of financial time series. *Physica A* 3 48, 404-418.