

Stochastic Simulation Model for non-stationary time series using Wavelet AutoRegressive Model

Moon, Young-Il* and Kwon, Hyun-Han**

Abstracts

Many hydroclimatic time series are marked by interannual and longer quasi-period features that are associated with narrow band oscillatory climate modes. A time series modeling approach that directly considers such structures is developed and presented. The essence of the approach is to first develop a wavelet decomposition of the time series that retains only the statistically significant wavelet components, and to then model each such component and the residual time series as univariate autoregressive processes. The efficacy of this approach is demonstrated through the simulation of observed and paleo reconstructions of climate indices related to ENSO and AMO, tree ring and rainfall time series. Long ensemble simulations that preserve the spectral attributes of the time series in each ensemble member can be generated. The usual low order statistics are preserved by the proposed model, and its long memory performance is superior to the direct application of an autoregressive model.

Keyword: Wavelet decomposition, nonstationary, nonlinear, time series simulation, autoregressive model

1. Introduction

Stochastic hydrologic methods have been very useful for a variety of water resources problems where temporal uncertainty needs to be quantified. The time series models that were developed extensively since the 1960s have typically assumed that the series modeled comes from a stationary or cyclostationary process. Thus the literature has developed around autoregressive moving average models and their extensions to consider seasonality through periodic terms. Multi-site models and space-time disaggregation approaches have also been considered.

However, as record lengths have increased, hydrologists have become aware of the low frequency structure of climate and associated hydrologic time series. Traditionally, an ARMA model is considered for such a time series. Such a model is capable of generating linear oscillations, even with relatively low values of p and q . However, applications of such models often do not reproduce the spectral signature of the time series, specifically the amplitude-frequency modulation over time that is seen in moving window spectra or wavelet spectra (Kwon et al., 2007).

An objective of this study is to explore the use of autoregressive models with wavelet

* Member-University of Seoul, Professor-E-mail : ymoon@uos.ac.kr

* Member-Columbia University Associate Research Scientist-E-mail : hk2273@columbia.edu

decomposition as a time series simulator for systems with nonstationary time series without a priori specifying any of these modeling structures. The continuous wavelet transform is applied to decompose a univariate time series into several statistically significant components and then a linear AR model is employed to simulate each component extracted from wavelet transform analysis, as well as the residual "noise" term.

2. Wavelet Autoregressive Model

Consider a time series x_t , $t=1,\dots,N$, recorded at monthly, that exhibits low frequency variations at intraseasonal, interannual and longer time scales, as seen in many hydroclimatic time series. Consider the decomposition of this series into S component series RC_{st} that represent "signal" and a residual term ϵ_t .

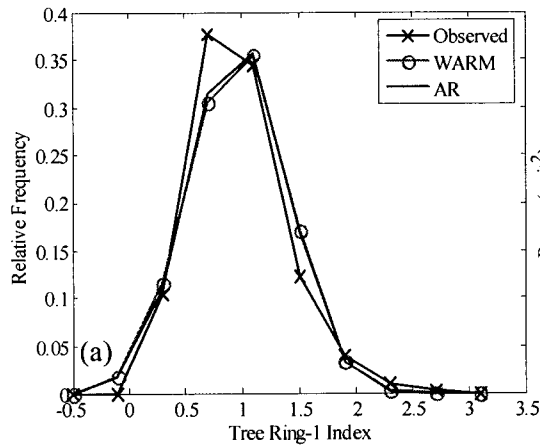
$$x_t = \sum_{k=1}^S RC_{st} + \epsilon_t \quad (1)$$

The decomposition in (1) considers that there are S orthogonal or independent series that carry the low frequency information, and the residual, ϵ_t , is a stochastic process. The notion is that the dynamics of each of these terms (RC_{st} and ϵ_t) is simpler to model using an autoregressive model than an autoregressive model for the composite dynamics of all the components. In general, each could be modelled using an appropriate time series technique. Here, we consider a linear autoregressive model for each term, leading to the following model structure:

3. Applications

Applications to 2 real world time series are presented the 384year record of Treering and the 110 year record of the NINO4 index of ENSO. The record of tree ring data used for simulation extends over 384 years. We note that the entire 10 year, 20 year, 40 year 100 year band has a GWP level higher than the significance level. Both the AR model and WARM model generally preserve the marginal distribution (see Figure 1-a). The WARM simulations are considerably better at reproducing the decadal and multidecadal spectral signatures in the original observations (see Figure 1-b). The AR, simulated spectra typically do have low frequency components, but are generally broad band, instead of being sharp as in the WARM case.

a) Relative Frequency



b) Global Wavelet Power Spectrum

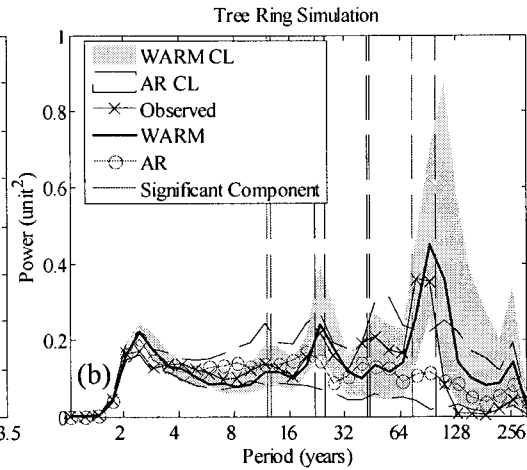
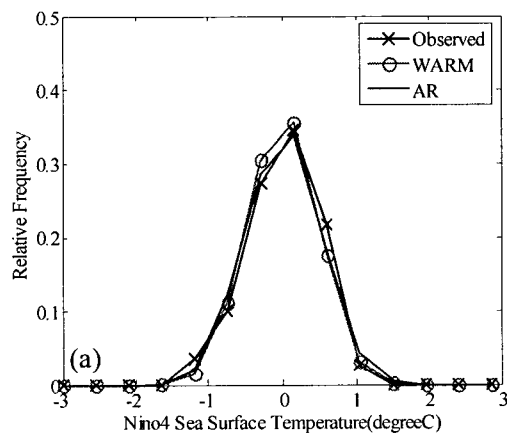


Figure 1. Comparison of average relative frequency distribution (a) and Global wavelet power spectrum (b) of 1000 simulations for the Treering in FL using the AR, ARMA and WARM models.

110 years Observed data are used. We note that 3 year to 5 year band has a GWP level higher than the significance level. At this stage one could conclusively conclude that WARM usually reproduces the observed spectrum over the frequency band of fitting, while the AR model does not (see Figure 2). The mean AR simulated spectrum is often outside the WARM spectrum uncertainty bounds, suggesting that there is a statistically significant difference between simulations from these two models.

a) Relative Frequency



b) Global Wavelet Power Spectrum

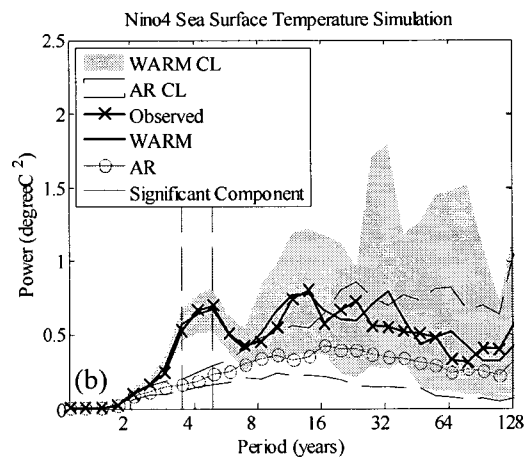


Figure 2. Comparison of average relative frequency distribution (a) and Global wavelet power spectrum (b) of 1000 simulations for the Nino4 Sea Surface Temperature using the AR, ARMA and WARM models.

4. Conclusion

Many hydroclimatic time series are marked by interannual and longer quasi-period features that are associated with narrow band oscillatory climate modes. Main question in simulating hydroclimatic time series is that how the non-normality and low frequency memory observed in hydroclimatic time series can be effectively modeled by an approach that decomposes the time series into significant spectral components and noise and then models each such process using a traditional AR process? To address this question empirically, we considered two examples, and compared the application of AR model with WARM. The essence of the approach is to first develop a wavelet decomposition of the time series that retains only the statistically significant wavelet components, and to then model each such component and the residual time series as univariate autoregressive processes. Long ensemble simulations that preserve the spectral attributes of the time series in each ensemble member can be generated. The usual low order statistics are preserved by the proposed model, and its long memory performance is superior to the direct application of an autoregressive model.

References

1. Kwon, H.-H., U. Lall, and A. F. Khalil (2007), Stochastic simulation model for nonstationary time series using an autoregressive wavelet decomposition: Applications to rainfall and temperature, *Water Resour. Res.*, 43, doi:10.1029/2006WR005258.