Markov switching autoregressive models and Bayesian hydrology

Roberta Paroli[1] and Luigi Spezia[2]

1. *(1) Dipartimento di Scienze Statistiche, Università Cattolica SC, Milano, Italy*
2. *(2) Biomathematics & Statistics Scotland, Aberdeen, UK*

Hydrological time series are realisations of complex stochastic systems. A few issues need to be taken into account by the modellers: non-Normality, non-linearity, non-stationarity, and long memory. Non-Normality is observed when the data density is multimodal or asymmetric or kurtic and the data cannot be considered as realisations from a Gaussian process. Non-linearity is assumed when the whole series does not show the same statistical peculiarities over all the observations, but they can be classified into a few homogeneous groups, each one with specific characteristics (e.g., different means and/or different variances). Non-linearity can also be assumed when the series exhibits asymmetries, e.g., when peaks are sharper (or more rounded) than the troughs, and/or when the cycles increase at a different rate from which they decrease. Weak non-stationarity is caused by generating processes having time-varying means and autocovariances (possibly due to periodic components and/or covariates). Finally, when the series shows high autocorrelations at the higher lags, with a slow decay, the observations are realisations from a long memory process. Because of these issues, hydrological time series can be analysed by Markov switching autoregressive models (MSARMs). MSARMs are pairs of discrete-time stochastic processes, one observed and one latent, or hidden. The hidden process is a finite-state Markov chain, whereas the observed process, given the Markov chain, is conditionally autoregressive. The dynamics of the observed process is driven by the dynamics of the latent one, so that each observation depends on the contemporary state of the Markov chain. By this theoretical structure, MSARMs allow: *i*) modelling non-linear and non-Normal time series by assuming that different autoregressions, each one depending on a hidden state, alternate according to the Markovian regime switching; *ii*) classifying the observations into a small number of homogeneous groups, labelled as the regimes of the Markov chain. We propose MSARMs within the Bayesian framework: Bayesian inference, model choice, and variable selection are performed numerically by Markov chain Monte Carlo (MCMC) algorithms. Our applications to isotope signatures, turbidity measurements, and river temperatures provide new clear examples of the suitability of the MSARMs in hydrological time series analysis in particular and environmental sciences in general. We hope our work can motivate other scientists to approach MSARMs and give their highly structured time series a valuable interpretation.