DO NOT CIRCULATE

STOCHASTIC HYDROLOGY

/ TOPIC 4 /

# Time series (Single and Multi-variate)

**Due**
25 Jun 2023

You will use the daily precipitation at Princeton provided in the EXCEL file ('Princeton_Precipitation_2002-2014') for this exercise.

**Task 0 –** First divide the data into the following seasons (Dec-Feb; Mar-May; Jun-Aug; Sep-Nov) by concatenating the seasons together from the 12 years.

**Task 1 –** Using a 2-state (rain, no-rain), first-order Markov chain model, determine the transition probability matrix for each season, the steady state probabilities and the lag-1 correlation. Discuss the seasonality of precipitation in Princeton. If you divide the record into two 6 year periods, do you see any change in the transition probabilities? Discuss your findings.

Also provided is an EXCEL file ('Assignment_time_series') that contains 5 time series, each 1096 long, coming from one of: AR(1), AR(2), MA(1), MA(2), ARMA(1,1). (There is nothing special about the length, it's long enough to get good statistical characterizations)

**Task 2 –** Calculate, for each series, the statistics of the 5 time series: mean, variance, and correlations and partial correlations up to lag-3, using all of the 1096 data points.

**Task 3 –** Try to determine which series came from which model, and estimate the parameter values using all 1096 data points.

There is a second folder listed as 'Multi-site' that contains times series for 5 sites that are correlated, coming from a multi-variate AR(1) model – again 1096 long.

**Task 4 –** We assume that the model is: $Z_t = AZ_{t-1} + B\epsilon_t$ ,where $Z$ is a 1×5 vector of the process at the 5 sites, and $\epsilon_t$ is a 1×5 vector of Gaussian(0,1) random noise terms. Calculate the parameter values in $A$ and $B$. For $B$ do the estimation:

- assuming $B$ is a lower triangular matrix
- using the eigenvector approach.