Reconstructing Complex System Dynamics from Time Series: A Method Comparison

Forough Hassanibesheli, Niklas Boers, and Jurgen Kurths

6 . May . 2020

Outline

Inverse Modelling Methods

- Langevin Equation (LE)
- Generalized Langevin Equation (GLE)
- Empirical model Reduction (EMR)

Results

- Reconstructing the dynamics and statistics of the underlying systems using Inverse Modelling Methods

Summary

Complex system :

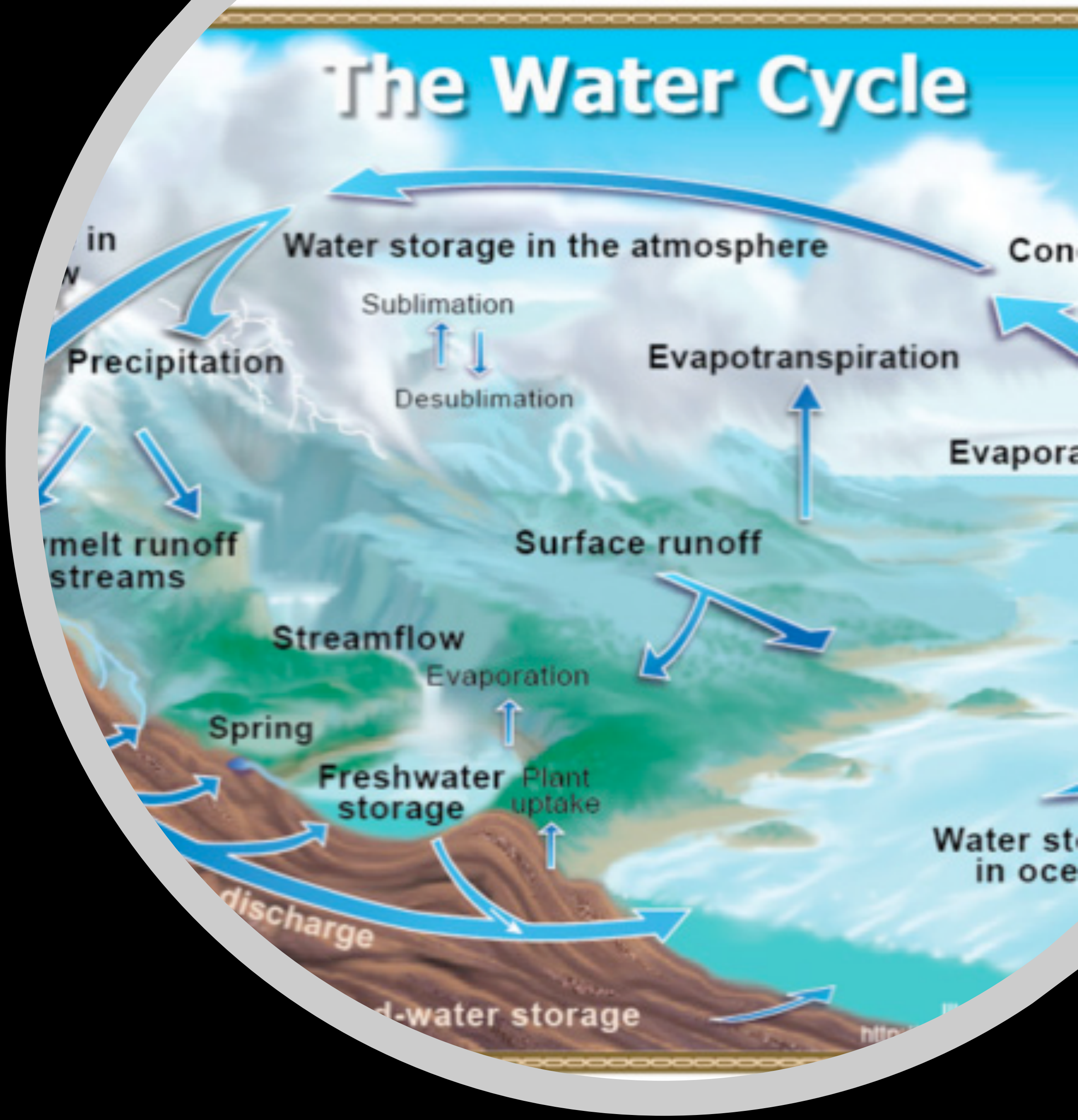
A system composed of highly interconnected components in which the collective property of an underlying system cannot be described by dynamical behavior of the individual parts alone.

A typical approach to study the behavior of a complex system is:

Focusing on the comparably few observed, macroscopic variables, assuming that they determine the key dynamics of the system, while the remaining ones are represented by noise

Using

Stochastic differential equations (SDEs)



Langevin Equation (LE) :

Consider a system for which the evolution of the macroscopic states x(t) obeys the following equation of motion:

$$\frac{dx(t)}{dt} = F(x,t) + G(x,t)\eta(t),$$

The drift F(x,t) and the diffusion G(x,t) can be directly estimated from the measured data without any prior knowledge about the internal dynamics of a system using **Kramers-Moyal (KM)** coefficients.

$$D^n(x,t) = \frac{1}{n!}\lim_{dt\to 0}\frac{1}{dt} < (x(t+dt) - x(t))^n|_{x(t)=x} >$$

The numerical discretization of the LE, in Itˆo's interpretation of stochastic integration, is as follows:

$$dx(t) = D^1(x,t)dt + \sqrt{D^2(x,t)}dw(t)$$

Generalized Langevin Equation (GLE) :

It has been proposed to account for long-range correlations and memory effects of complex systems that do not exhibit strong time scale separation. One pathway to derive a GLE is by means of the **Mori–Zwanzig formalism (MZ)** :

$$\frac{dx(t)}{dt} = \Omega x(t) - \int_{t_0}^{t} K(t-t')x(t')dt' + R(t,t_0)$$

The first, local term defines the self interactions of the macroscopic variables, the second, non-local term describes memory dependencies of the macroscopic variables, and the last term stands for the residual force associated with fast variables.

The deterministic term can be estimated using KM coefficients and the memory function K(t) can be derived from the autocorrelation function C(t) of the system:

$$\frac{dC(t)}{dt} = -\int_{t_0}^{t} K(t-t')C(t')dt'$$

Empirical Model Reduction (EMR) :

The general form of the EMR approach is as follows:

First, model the increments dx of an observed variable x as quadratic function of x plus a residual dr^0 that is typically obtained from a least-squares minimization yielding the parameters a, b^0; and c^0:

$$dx = (ax^2 + b^0 x + c^0)dt + dr^0$$

In the main step, the stochastic residual r^0 can be estimated through an iterative linear regression on the state variable x and residuals in previous steps. This recursive procedure terminates when the state- and time-correlations of the residual of the nth levels r^{n+1} vanish.

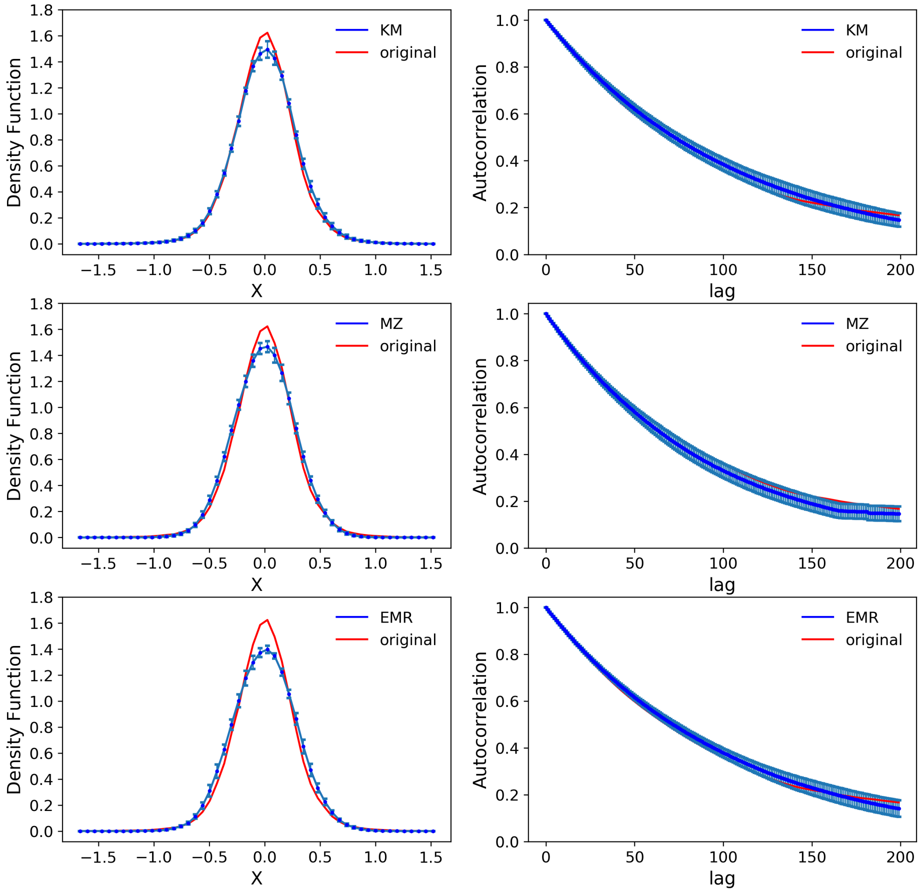
$$dr^0 = b^1[x, r^0]dt + r^1 dt$$

.....

$$dr^n = b^n[x, r^0, ..., r^n]dt + r^{n+1} dt$$

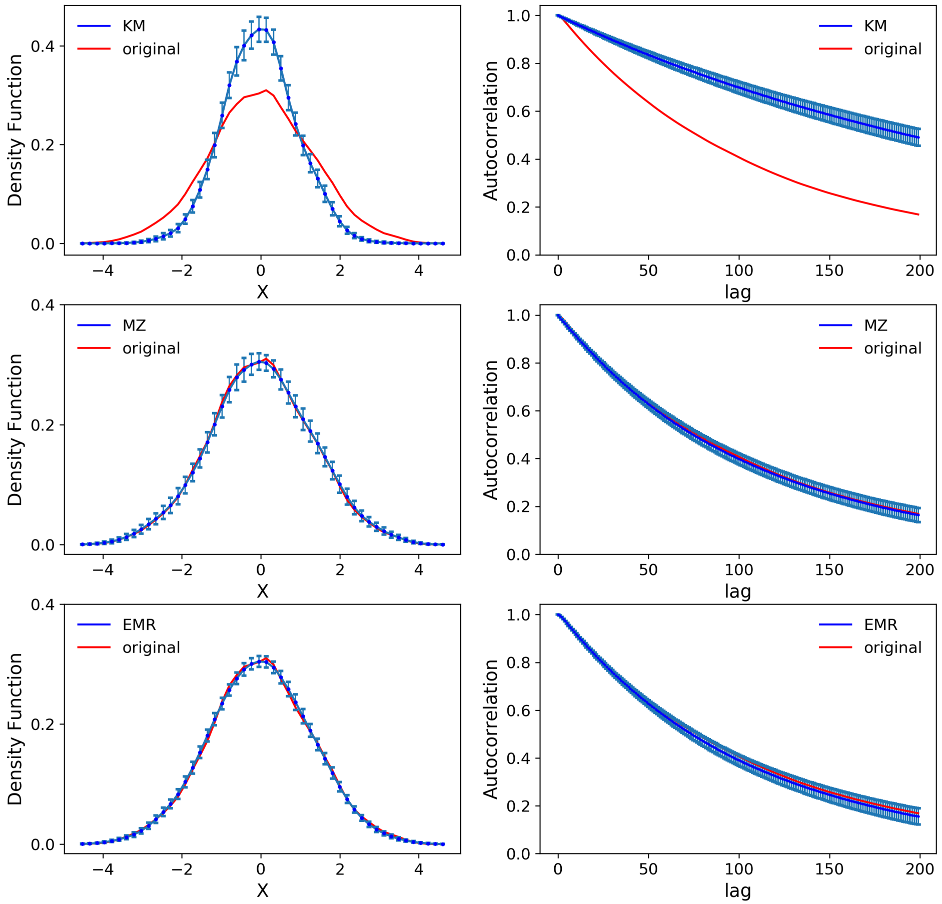
Ornstein–Uhlenbeck process with multiplicative noise:

$$\frac{dx(t)}{dt} = \theta(\mu - x(t)) + (1 + x^2)\eta(t)$$



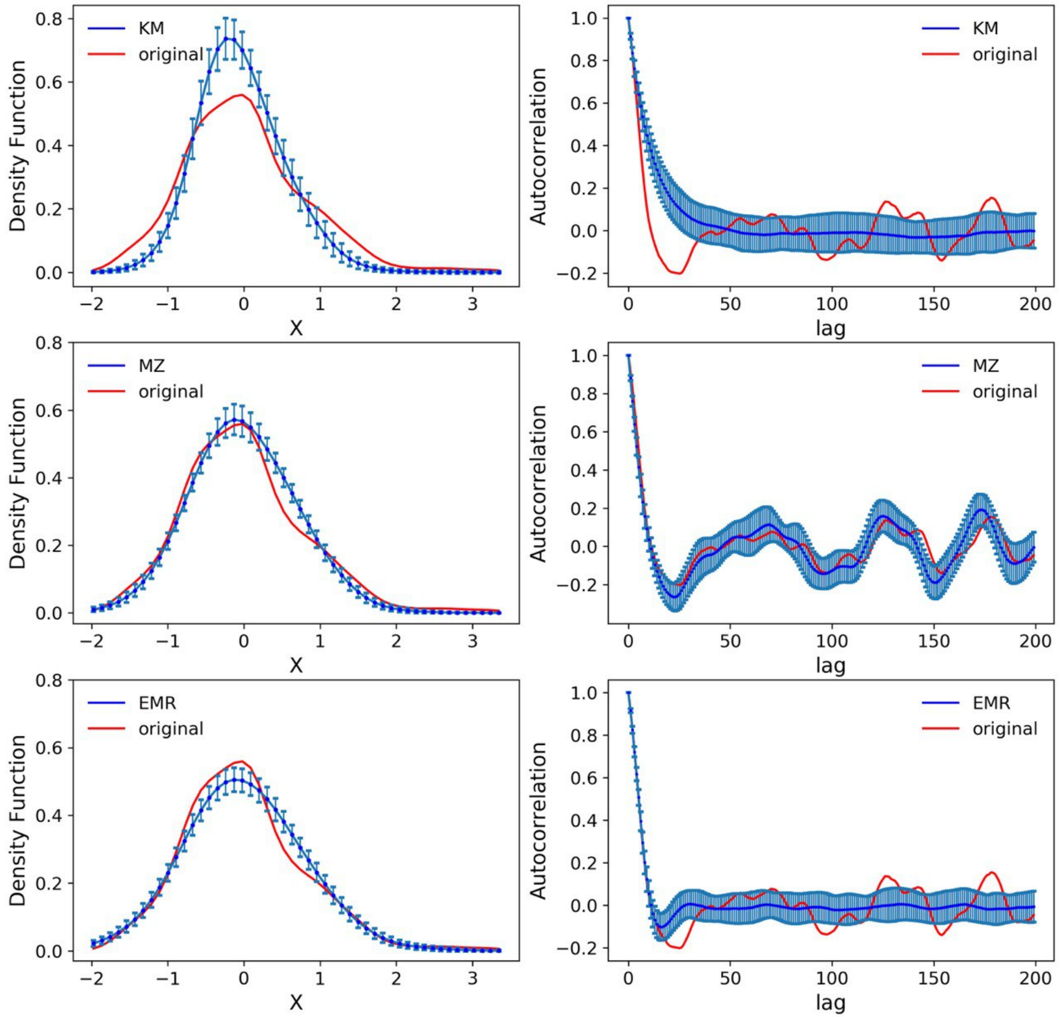
Ornstein–Uhlenbeck process with colored noise:

Here we substitute the stochastic term of the equation with a first order autoregressive process.

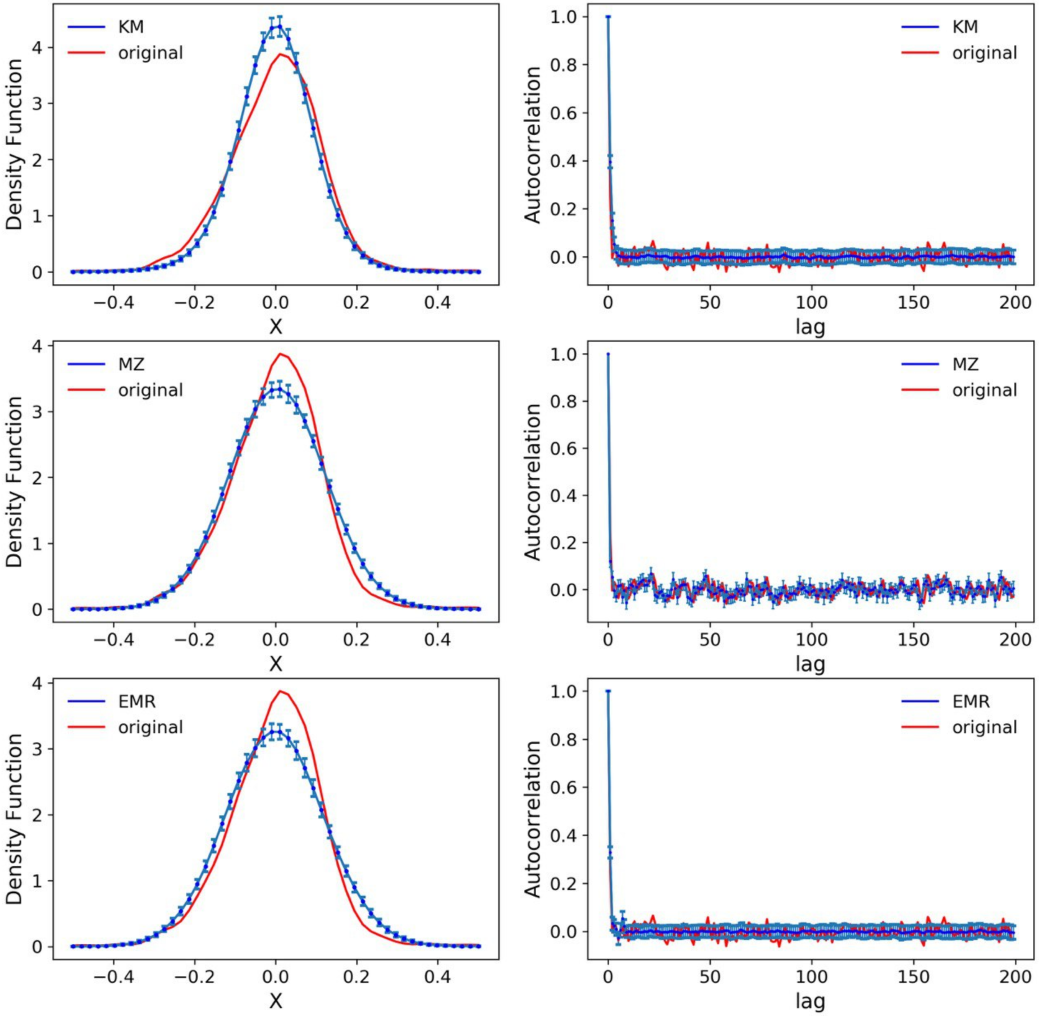


Summary statistics (PDFs in the left column and ACFs in the right column) of original and simulated time series derived from 1000 sample time series reconstructed by the three stochastic models (KM, MZ, and EMR), from top to bottom as indicated in the legend.

Nino-3 monthly SST indices averages from (1891 to 2015)



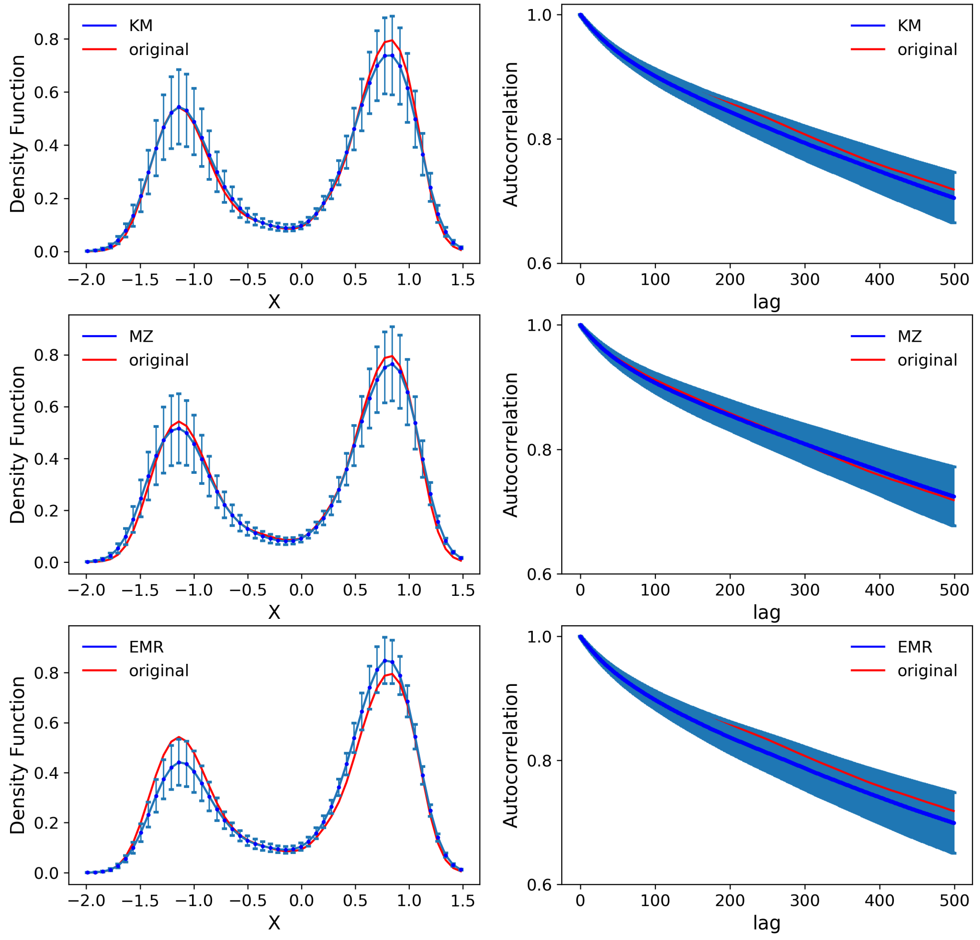
Weekly S&P500 stock index for the time span of 35 years (1950-1985)



Summary statistics (PDFs in the left column and ACFs in the right column) of original and simulated time series derived from 1000 sample time series reconstructed by the three stochastic models (KM, MZ, and EMR), from top to bottom as indicated in the legend.

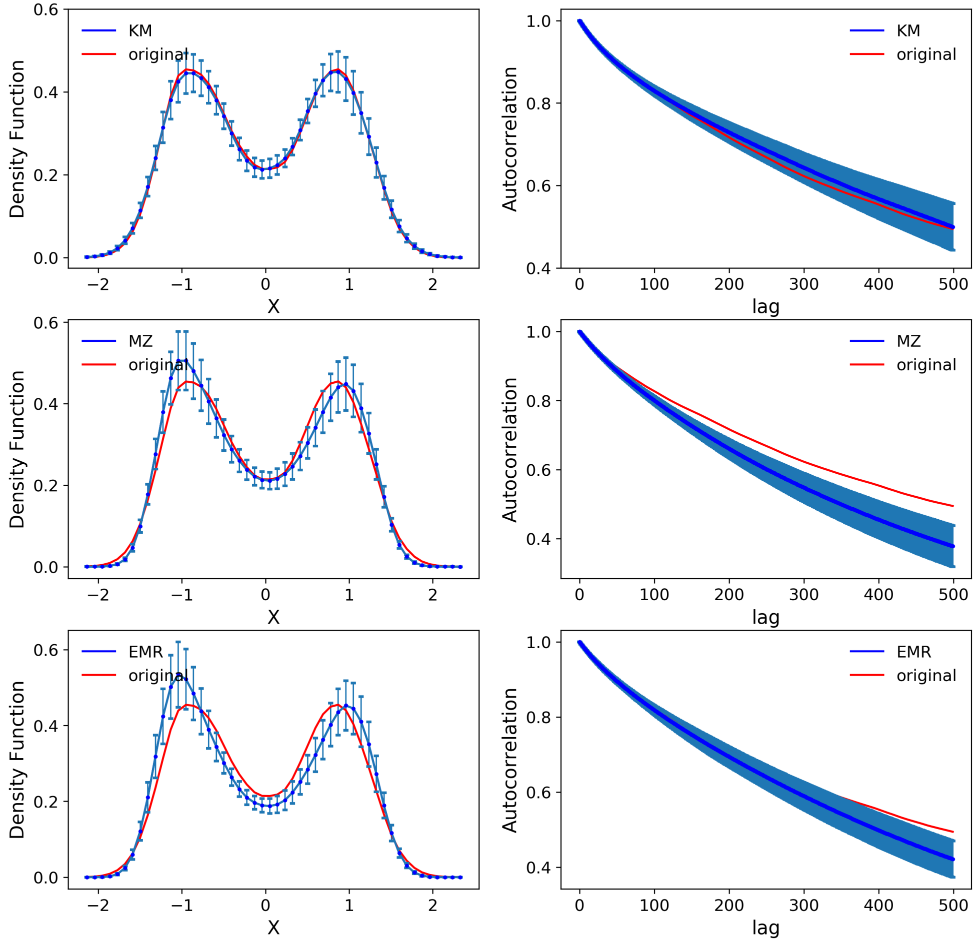
A particle moves in a double-well potential subjected to an additive noise

$$\frac{dX(t)}{dt} = \theta(x(t) - x^3(t)) + \sigma\eta(t)$$

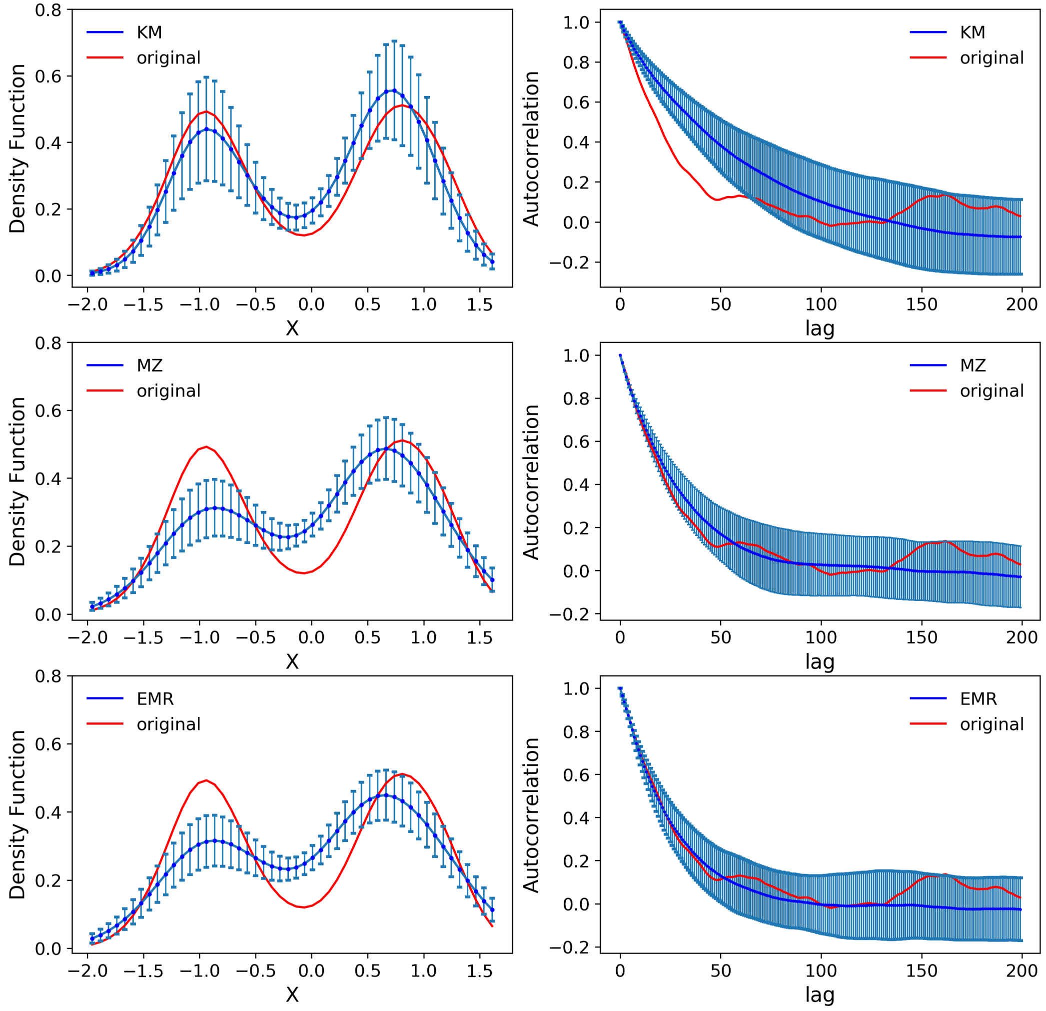


A particle moves in a double-well potential subjected to a multiplicative noise

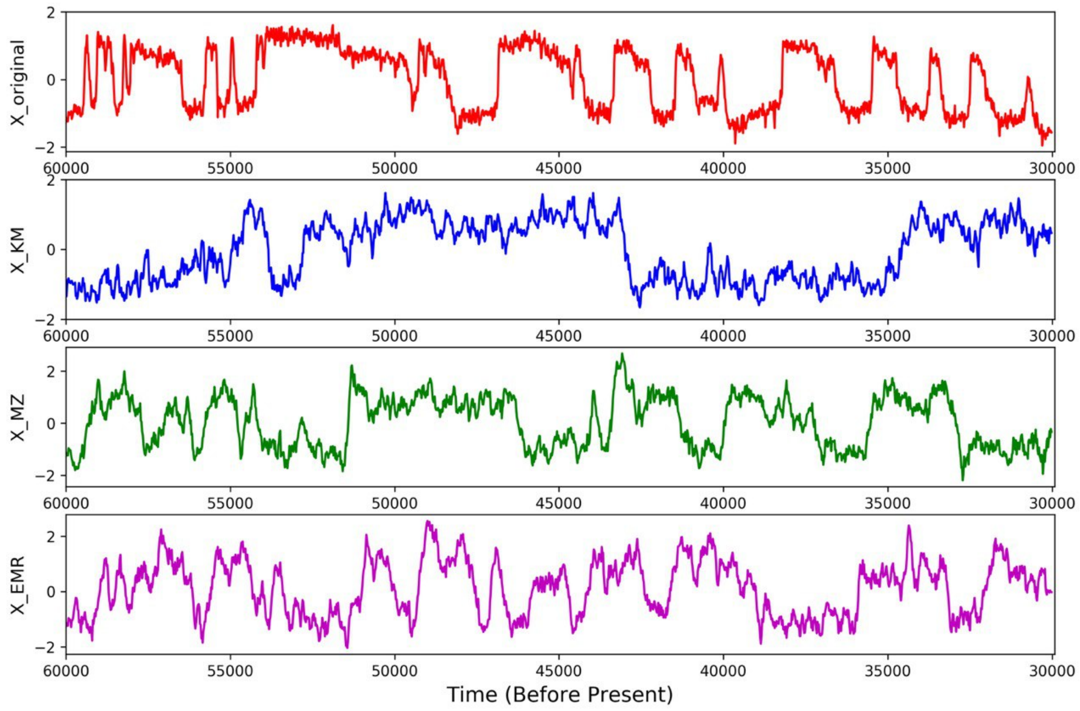
$$\frac{dX(t)}{dt} = \theta(x(t) - x^3(t)) + \sqrt{1 + x^2}\eta(t)$$



The high-resolution (20yr-average) Ca^{2+} (interpreted as a proxy for atmospheric circulation patterns), collected from the NGRIP ice core on the GICC05 time scale.



Because of the substantially better signal-to-noise ratio, we focus here on the Ca^{2+} time series between 60ka and 30ka b2k.



Original (red) and randomly chosen simulated time series based on KM, MZ, and EMR methods (from top to bottom respectively).

Summary

- In this work we attempt to reconstruct the dynamical equations of motion of both synthetical and real-world processes, thoroughly comparing (LE,GLE,EMR) in terms of their capability to reconstruct the dynamics and statistics of the underlying systems.

- Though we generally observe an appropriate performance of all approaches for unimodal systems, our results have revealed that the MZ method exhibits a better performance especially in the case of modeling ENSO dynamics.

- We propose that due to the memory contribution in real-world systems, LEs (as derived by the KM approach) cannot fully grasp the underlying behavior and it is essential to take into account the non-Markovian closure terms