Geophysical Research Abstracts
Vol. 12, EGU2010-955-1, 2010
EGU General Assembly 2010

Building Chaotic Model From Incomplete Time Series

Michael Siek (1) and Dimitri Solomatine (1,2)

(1) UNESCO-IHE Institute for Water Education, Hydroinformatics and Knowledge Management, Delft, Netherlands (m.siek@unesco-ihe.org, +31-15-2122921), (2) Water Resources Section, Delft University of Technology, The Netherlands

This paper presents a number of novel techniques for building a predictive chaotic model from incomplete time series. A predictive chaotic model is built by reconstructing the time-delayed phase space from observed time series and the prediction is made by a global model or adaptive local models based on the dynamical neighbors found in the reconstructed phase space. In general, the building of any data-driven models depends on the completeness and quality of the data itself. However, the completeness of the data availability can not always be guaranteed since the measurement or data transmission is intermittently not working properly due to some reasons.

We propose two main solutions dealing with incomplete time series: using imputing and non-imputing methods. For imputing methods, we utilized the interpolation methods (weighted sum of linear interpolations, Bayesian principle component analysis and cubic spline interpolation) and predictive models (neural network, kernel machine, chaotic model) for estimating the missing values. After imputing the missing values, the phase space reconstruction and chaotic model prediction are executed as a standard procedure. For non-imputing methods, we reconstructed the time-delayed phase space from observed time series with missing values. This reconstruction results in non-continuous trajectories. However, the local model prediction can still be made from the other dynamical neighbors reconstructed from non-missing values.

We implemented and tested these methods to construct a chaotic model for predicting storm surges at Hoek van Holland as the entrance of Rotterdam Port. The hourly surge time series is available for duration of 1990-1996. For measuring the performance of the proposed methods, a synthetic time series with missing values generated by a particular random variable to the original (complete) time series is utilized. There exist two main performance measures used in this work: (1) error measures between the actual and missing value estimates; and (2) the accuracy of the built chaotic model predictions. The model results indicate that the proposed methods are able to build a chaotic model from incomplete time series and to provide reliable and accurate predictions.