Hydrol. Earth Syst. Sci. Discuss., 10, C7292–C7298, 2014
www.hydrol-earth-syst-sci-discuss.net/10/C7292/2014/

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Downstream prediction using a nonlinear prediction method" *by* N. H. Adenan and M. S. M. Noorani

Anonymous Referee #1

Received and published: 15 January 2014

General Comments:

In this manuscript, the authors applied nonlinear prediction method for analysis of four years of daily river flow data for the Langat River at Kajang, Malaysia which is known for regularly suffers from flooding. The authors developed two models name as, Model I and Model II. Both models involve same modeling procedure except determining embedding dimension for phase-space reconstruction. The former one uses the correlation dimension method while the latter one uses the false nearest neighbor approach for determining the embedding dimension which is essence of nonlinear prediction method. The authors argued that both models could give a good prediction for the river flow downstream. However, in my opinion the analyses are not sufficiently rigorous to conclude that the above conclusion. In the following, I have included several

comments.

Specific Comments:

1. The determination of the preliminary parameter pair (tau, m) plays a significant role in forecasting performance of nonlinear prediction method. As the authors mentioned, selection of the appropriate tau is important during the reconstruction of the phase-space. If tau is too small, the phase-space coordinates will not be independent enough to produce new information about the evolution of the system whereas if it is too large then all the relevant information is lost because of diverging trajectories (Islam and Sivakumar, 2002).

The authors proposed that "when a condition of tau=1 is used in phase space reconstruction, the results gave good predictions (Sivakumar, 2002; Sivakumar, 2003). Thus, in this study tau=1 is used". Based on the Takens theorem (Takens, 1981), tau could be randomly chosen, in the absence of noise and infinitely long time series. However, in the real life situation, discharge time series could never be free from noise, e.g. measurement error. Therefore, subtle analyses are required before selection of delay time.

In the literature, the most used methods for determining the delay time (tau) are the autocorrelation function (ACF) and mutual information function (MI). Generally, tau can be determined where the ACF attains the value of zero (Holzfuss and Mayer-Kress, 1986) or below a pre-determined value such as 0.5 (Schuster, 1988) and 0.1 (Tsonis and Elsner, 1988). In the phase-space reconstruction analysis a good choice of the time delay is required for geometrical and numerical analysis of attractor. Selection of proper delay time mainly based on system dynamics to unfold the attractor in d-dimensional plot, for instance Porporato and Ridolfi (1997) found essential delay time as 1, 7 days in Islam and Sivakumar (2002), 10 days in Elshorbagy et al. (2002), 14 days in Ng et al. (2007) and 146 days in Pasternack (1999). These differences mainly rise from the differences in autocorrelation structure, mutual information functions and/or underlying system dynamics. For instance, a surrogate time series that was generated from a stochastic autoregressive process with AR(1) where AR(1) coefficient is 0.80, which have similar statistics with the original time series that used in the study can be seen below.

Figure 1. A stochastically generated time series

As can be seen, a time series that generated from a stochastic process could yield a similar phase-space plot to a chaotic one. Therefore, a stand-alone phase-space plot with delay time tau=1 is not an indicator of chaotic dynamics against the authors proposed in page 14341 “the trajectories of the attractor are clearly shown in two phase diagrams. Thus, the data involved in this analysis are chaotic”. Finally, the effect of tau on phase-space plot is well explained in Islam and Sivakumar (2002). The authors investigated the effect of various delay times, i.e. 1, 7 and 200 on phase-space plot and concluded that tau=1 is highly dependent and tau=200 is highly independent. tau=7 seems to be provide some kind of compromise when compared to the above two. Therefore, selection of ïĄť should have been more rigorous.

2. The accurate estimation of correlation integrals largely depends on properly selected delay time, because, the correlation integrals are calculated from the reconstructed time series (Khatibi et al., 2011; Ng et al., 2007; Wu and Chau, 2010). Therefore, without properly selected delay time, the calculated correlation integrals and resulted embedding dimension which leads to Model I would be spurious.

3. As Wu and Chau (2010) mentioned “The false nearest neighbor (FNN) technique is not concerned with a dynamic system being deterministic or not. Thus, the correlation integral method can be more reliable for unfolding dynamical system”. However, in this study, the FNN gave higher embedding dimension than the correlation integral method and there is not any discussion about this phenomenon (e.g. is this resulted from the system dynamics?). Also, Model II that employs higher embedding dimension than the Model I gave better results than Model I. However, Khatibi et al. (2012) showed that the

performance of nonlinear prediction method is based on the selection of embedding dimension and the RMSE vs. embedding dimension plot shows that the best obtained results with k is in harmony with the embedding dimension with the correlation dimension not FNN. Generally, higher embedding dimension (especially higher embedding dimensions higher than that obtained with the correlation integral analysis) would give poor results than the lower ones as can be seen in (Sivakumar, 2003; Sivakumar, 2007; Sivakumar et al., 2002). This result can be originated from the under-estimated embedding dimension with the correlation integral analysis or over-estimated embedding dimension with the FNN analysis.

4. Another important parameter in the nonlinear prediction method is the number of nearest neighbors. In the study, there is not any explanation about how and how many nearest neighbors are selected?

5. The authors argued that both methods gave a good prediction performance for the river flow downstream. However, the obtained correlation coefficients for Model I and Model II are 0.6103 and 0.6360 which are essentially Rsquare=0.372 and Rsquare=0.404, respectively. I think these Rsquare values are rather poor and it is evident that there is no reason to use these models in downstream prediction.

6. Finally, the study employs two different approaches to determine the embedding dimension for nonlinear prediction method. In the study, there is not any discussion about system dynamics and I think there is no need for chaotic analysis for determining embedding dimension in this study. A trial-and-error process would be quite efficient in determining embedding dimension for more practical use.

References

Elshorbagy, A., Simonovic, S.P. and Panu, U.S., 2002. Estimation of missing streamflow data using principles of chaos theory. Journal of Hydrology, 255(1–4): 123-133.

Holzfuss, J. and Mayer-Kress, G., 1986. An Approach to Error-Estimation in the Application of Dimension Algorithms. In: G. Mayer-Kress (Editor), Dimensions and Entropies in Chaotic Systems. Springer Series in Synergetics. Springer Berlin Heidelberg, pp. 114-122.

Islam, M.N. and Sivakumar, B., 2002. Characterization and prediction of runoff dynamics: a nonlinear dynamical view. Advances in Water Resources, 25(2): 179-190.

Khatibi, R. et al., 2011. Dynamics of hourly sea level at Hillarys Boat Harbour, Western Australia: a chaos theory perspective. Ocean Dynamics, 61(11): 1797-1807.

Khatibi, R. et al., 2012. Investigating chaos in river stage and discharge time series. Journal of Hydrology, 414–415(0): 108-117.

Ng, W.W., Panu, U.S. and Lennox, W.C., 2007. Chaos based analytical techniques for daily extreme hydrological observations. Journal of Hydrology, 342: 17-41.

Pasternack, G.B., 1999. Does the river run wild? Assessing chaos in hydrological systems. Advances in Water Resources, 23(3): 253-260.

Porporato, A. and Ridolfi, L., 1997. Nonlinear analysis of river flow time sequences. Water Resources Research, 33(6): 1353-1367.

Schuster, H.G., 1988. Deterministic Chaos. VCH, Weinheim.

Sivakumar, B., 2002. A phase-space reconstruction approach to prediction of suspended sediment concentration in rivers. Journal of Hydrology, 258(1–4): 149-162.

Sivakumar, B., 2003. Forecasting monthly streamflow dynamics in the western United States: a nonlinear dynamical approach. Environmental Modelling & Software, 18(8-9): 721-728.

Sivakumar, B., 2007. Nonlinear determinism in river flow: prediction as a possible indicator. Earth Surface Processes and Landforms(32): 969-979.

Sivakumar, B., Jayawardena, A.W. and Fernando, T.M.K.G., 2002. River flow fore-

casting: use of phase-space reconstruction and artificial neural networks approaches. Journal of Hydrology, 265(1-4): 225-245.

Takens, F., 1981. Detecting strange attractors in turbulence. In: D.A. Rand, Jung, L.S. (Editor), Dynamical Systems and Turbulence, Lecture Notes in Mathematics. Springer-Verlag, Berlin, pp. 366-381.

Tsonis, A.A. and Elsner, J.B., 1988. The weather attractor over very short timescales. Nature, 333: 545–547.

Wu, C.L. and Chau, K.W., 2010. Data-driven models for monthly streamflow time series prediction. Engineering Applications of Artificial Intelligence, 23: 1350-1367.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 14331, 2013.

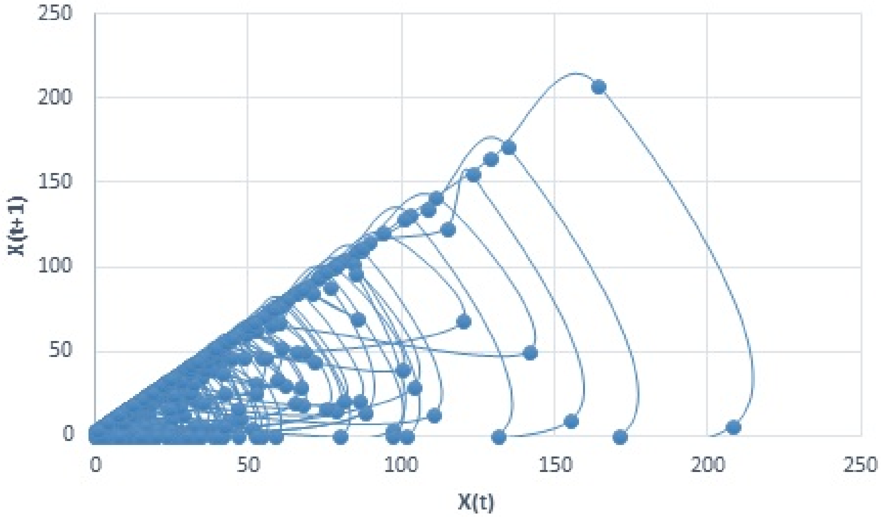

Fig. 1. A stochastically generated time series