Evidence of Low-dimensional Determinism in River Flow Dynamics: Examples from Predictions

B. Sivakumar

Department of Land, Air and Water Resources, University of California, Davis, CA 95616, USA
(sbellie@ucdavis.edu)

Abstract: Whether or not river flow dynamics exhibit low-dimensional determinism remains an unresolved question. While, on one hand, studies on the use of low-dimensional deterministic techniques for modeling and prediction of river flow dynamics are on the rise and the outcomes (regarding the presence of low-dimensional determinism) are encouraging, on the other hand, suspicions and criticisms on such studies continue to exist as well. An important reason for such suspicions and criticisms, however, is that the correlation dimension method, used as a determinism identification tool in most of those studies, possesses inherent limitations when applied to (real) river flow series, which are always finite (and often short) and contaminated with noise (e.g. measurement error). In view of this, the present study addresses the issue of low-dimensional determinism in river flow dynamics using prediction as an indicator. This is done by: (1) reviewing studies that have employed low-dimensional deterministic approaches (coupling phase-space reconstruction and local approximation techniques) for river flow predictions; and (2) identifying determinism (or distinguishing determinism and stochasticity) based on the level of prediction accuracy in general, and based on the prediction accuracy against the phase-space reconstruction parameters (also called as "inverse approach") in particular. The results not only provide possible evidence to the presence of low-dimensional determinism in the river flow series studied but also support, both qualitatively and quantitatively, the low correlation dimensions reported for such series. Therefore, low-dimensional deterministic techniques are a viable alternative for studying river flow dynamics, if only sufficient caution is exercised in their application and in interpreting the outcomes.

Keywords: *River flow; Stochasticity; Determinism; Phase-space reconstruction; Correlation dimension; Local approximation prediction; Inverse approach*

1. INTRODUCTION

Whether river flow dynamics are governed dominantly by a large number of variables or by only a very few variables has been and continues to be an unresolved question. The (seemingly) irregular behaviors of river flow phenomena and the (significant) variability they exhibit both in time and in space have led a majority of researchers to employ the concept of stochastic process for modeling and prediction of their dynamics. The ability of the stochastic models to fairly represent the important (statistical) characteristics of river flow series and the reasonably good predictions achieved on their evolutions have further strengthened our view on the usefulness of the stochastic process concept for river flow.

However, as river flow at some time/space scales are not as irregular and as complex as that at other time/space scales (suggesting possible simplicity in the former), the appropriateness of the stochastic process concept for every river flow phenomenon remains to be answered. The limitation that lies with the stochastic models may be explained with reference to the (deliberate) removal therein of the simple and easily predictable components of river flow (e.g. trend, seasonal and annual cycles), which is essentially done so that the time series under investigation suits the (random) assumption involved in the models. If this were the case, then any attempt to assess the usefulness and appropriateness of the stochastic models based on the river flow predictions achieved is misleading.

The necessity of the stochastic process concept for every river flow phenomenon comes into further question with our knowledge of chaos theory (Lorenz, 1963), according to which complex and irregular looking phenomena (such as river flow) might also be the outcome of simple deterministic systems with only a few nonlinear interdependent variables with sensitive dependence on initial conditions. Searching for the possible presence of low-dimensional determinism (also commonly called as deterministic chaos) in river flow dynamics and, upon identification of its presence, predicting their dynamics using nonlinear deterministic prediction techniques have been among the most exciting research activities in hydrology in recent times. Within this overall framework, studies have concentrated on only identification (e.g. Stehlik, 1999) or only prediction (e.g. Liu et al., 1998; Sivakumar et al., 2001) or both identification and subsequent prediction (e.g. Porporato and Ridolfi, 1997; Lambrakis et al., 2000; Islam and Sivakumar, 2002).

In studies investigating the presence of deterministic chaos in river flow dynamics, it is a common practice to employ the correlation dimension method (e.g. Grassberger and Procaccia, 1983) and assume the presence of low dimension either as a proof (e.g. Stehlik, 1999) or as a preliminary evidence (e.g. Porporato and Ridolfi, 1997; Islam and Sivakumar, 2002) of chaos. However, these studies and the reported results are often criticized, since the correlation dimension method has been found to possess certain important limitations, such as: (1) it is designed under the assumptions that the time series is infinite and noise-free; and (2) finite and low correlation dimensions might result also from linear stochastic processes (e.g. Osborne and Provenzale, 1989; Schertzer et al., 2002). Even though, the bases for such criticisms are often unfounded [see Sivakumar (2000) and Sivakumar et al. (2002a) for details] and the arguments put forth therein are unsupported [see Sivakumar et al. (2002b) for details], the potential limitations that exist in the use of the correlation dimension method for identifying chaos cannot be brushed aside completely.

In view of the above, there is a need to find an alternative way to (verify and) support the results reported by past studies regarding the possible presence of deterministic chaos in river flow dynamics. This issue is addressed in the present study, where an inverse approach (e.g. Casdagli, 1989, 1991; Sugihara and May, 1990) is used to identify deterministic chaos. As per this approach, the presence of deterministic chaos is identified using the river flow prediction results reported by such studies themselves. This is done by: (1) assessing the ability of the nonlinear deterministic local approximation approaches for predicting river flow dynamics; and (2) comparing the prediction results with respect to the parameter(s) used in the local approximation procedure. The inverse approach may be considered much more reliable than the correlation dimension method for chaos identification, since it is essentially based on predictions, which is the primary purpose behind characterizing a system in the first place.

The organization of the paper is as follows. Section 2 presents a brief account of the nonlinear prediction method and the inverse approach to identify chaos. Section 3 reviews the studies on river flow predictions using the nonlinear prediction method and also discusses the results in terms of the inverse approach. Conclusions are presented in Section 4.

2. NONLINEAR PREDICTION METHOD

In the nonlinear prediction method, the underlying dynamics of the system (e.g. time series) under investigation is represented by reconstructing the phase-space, i.e. embedding the single-dimensional (river flow) series, X_i, $i = 1, 2, \ldots, N$, in a multi-dimensional phase-space, according to:

$$\boldsymbol{Y}_j = (X_j, X_{j+\tau}, X_{j+2\tau}, \ldots, X_{j+(m-1)\tau}) \qquad (1)$$

where $j = 1, 2, \ldots, N-(m-1)\tau$, m is the dimension of the vector $\boldsymbol{Y}_j$, called as embedding dimension; and τ is a delay time or interval (e.g. Takens, 1981). A (correct) phase-space reconstruction in a dimension m allows one to interpret the underlying dynamics in the form of an m-dimensional map f_T, that is,

$$\boldsymbol{Y}_{j+T} = f_T(\boldsymbol{Y}_j) \qquad (2)$$

where $\boldsymbol{Y}_j$ and $\boldsymbol{Y}_{j+T}$ are vectors of dimension m, describing the state of the system at times j (current state) and $j+T$ (future state), respectively. The problem then is to find an appropriate expression for f_T (e.g. F_T).

There are several approaches for determining F_T. However, the local approximation approach (e.g. Farmer and Sidorowich, 1987) is widely employed (in river flow studies). In this approach, the f_T domain is subdivided into many subsets (neighborhoods), each of which identifies some approximations F_T, valid only in that subset and, hence, in this way, the underlying system dynamics are represented step by step locally in the phase-space.

The identification of the sets in which to subdivide the domain is done by fixing a metric || || and, given the starting point $\boldsymbol{Y}_j$ from which the forecast is initiated, identifying neighbors $\boldsymbol{Y}_j^p$, $p = 1, 2, \ldots, k$, with $j^p < j$, nearest to $\boldsymbol{Y}_j$, which constitute the set corresponding to $\boldsymbol{Y}_j$. The local functions can then be built, which take each point in the neighborhood to the next neighborhood: $\boldsymbol{Y}_j^p$ to $\boldsymbol{Y}_{j+1}{}^p$. The local maps may be learned in the form of local averaging (e.g. Farmer and Sidorowich, 1987) or local polynomials (e.g. Abarbanel, 1996). The local averaging procedure has an important advantage over the local polynomial technique, as it is computationally inexpensive.

The prediction accuracy may be evaluated using statistical evaluators, such as correlation coefficient (ρ), root mean square error (RMSE), and coefficient of efficiency (R^2). The time series plots and scatter diagrams may also used to choose the best prediction results among a large combination of results achieved with different embedding dimensions and number of neighbors.

The presence of chaos in the underlying dynamics may generally be assessed using the inverse approach, as follows:

1. By assessing the general performance of the nonlinear deterministic local approximation approach. A high prediction accuracy may be an indication that the underlying dynamics are nonlinear deterministic, whereas a low prediction accuracy is expected if the dynamics are stochastic; and
2. By checking the prediction accuracy against, for instance:
   (a) The embedding dimension (m): If the dynamics are chaotic, then the prediction accuracy would increase (to its best) with the increase in the embedding dimension up to a certain point, called the optimal embedding dimension (m_{opt}), and remain close to its best for embedding dimensions higher than m_{opt}. For stochastic time series, there would be no increase in the prediction accuracy with an increase in the embedding dimension and the accuracy would remain the same for any value of the embedding dimension (e.g. Casdagli, 1989);
   (b) The lead time (T): For a given embedding dimension, predictions in chaotic systems deteriorate considerably faster than in stochastic systems when the lead time is increased. This is due to the sensitivity of chaotic systems to initial conditions (e.g. Sugihara and May, 1990); and
   (c) The number of neighbors (k): Smaller number of neighbors would give the best predictions if the system dynamics are chaotic, whereas for stochastic systems, the best predictions are achieved when the number of neighbors is large (e.g. Casdagli, 1991).

3. RIVER FLOW PREDICTIONS USING NONLINEAR PREDICTION METHOD

3.1. Review

In recent years, several studies have attempted river flow (and spring discharge) predictions using nonlinear deterministic local approximation prediction methods. In these studies, river flow series observed at different temporal scales, different climatic regions, and different basin characteristics have been analyzed. As the presence of noise in the data could affect the predictions, some studies have even attempted predictions of noise reduced river flow series and compared with that of the raw series (e.g. Porporato and Ridolfi, 1997; Jayawardena and Gurung, 2000). Comparisons between the nonlinear prediction method and other techniques (such as stochastic methods and artificial neural networks) for river flow predictions have also been made by some studies (e.g. Jayawardena and Gurung, 2000; Lambrakis et al., 2000; Sivakumar et al., 2002a, c).

Table 1 presents the river flow prediction results in terms of correlation coefficient (ρ) and coefficient of efficiency (R^2) reported by some of the above studies [the correlation dimensions, D, obtained are also presented therein]. These studies are chosen in such a way that the associated river flow series cover wide temporal, climatic, and basin ranges, and that they may be considered reasonable representations of flow series for subsequent interpretations. As most of these studies have also employed other techniques (e.g. correlation dimension method) to identify chaos, the additional results could further facilitate the interpretations.

3.2. Discussion of Results: Inverse Approach

Prediction accuracy

As may be seen from Table 1, the predictions achieved for the various river flow series using the nonlinear local approximation method are extremely good. The correlation coefficient and the coefficient of efficiency are, in general, greater than 0.90. Comparisons of the observed and the predicted flow series through direct time series plots, such as the one shown in Figure 1 for the daily river flow from the Lindenborg catchment in Denmark (Islam and Sivakumar, 2002), and scatter diagrams [figure not shown herein; see Islam and Sivakumar (2002)] reveal excellent agreement between the two. Excellent agreement between observed and predicted river flow values have also been reported by other studies (e.g. Porporato and Ridolfi, 1997, 2001; Lisi and Villi, 2001; Sivakumar et al., 2001, 2002a, c).

Table 1. River flow predictions using nonlinear local approximation prediction method (lead time = 1 day or 1 month).

River flow data	ρ	R^2	D	Reference*
Daily				
Denmark	0.99	0.98	< 4	IS(02)
Greece	0.90		< 4	LAPGB(00)
Italy	0.95		< 4	PR(97)
	0.98	0.95		PR(01)
	0.88		< 3	LV(01)
Thailand	0.99		< 2	JG(00)
	0.99	0.96		SJF(02)
	0.79		< 3	JG(00)
USA	0.85		< 4	LIRL(98)
	0.99	0.99	< 3	SJ(02),S(03)
Monthly				
Brazil	0.89	0.94		SBP(01)
	0.89		< 4	SPBU(02)
Sweden	0.99	0.99	< 6	SBOJK(00)

*IS(02) – Islam and Sivakumar (2002); LAPGB(00) – Lambrakis et al. (2000); PR(97) – Porporato and Ridolfi (1997); PR(01) – Porporato and Ridolfi (2001); LV(01) – Lisi and Villi (2001); JG(00) – Jayawardena and Gurung (2000); SJF(02) – Sivakumar et al. (2002c); LIRL(98) – Liu et al. (1998); SJ(02) – Sivakumar and Jayawardena (2002); S(03) – Sivakumar (2003); SBP(01) – Sivakumar et al. (2001); SPBU(02) – Sivakumar et al. (2002a); SBOJK(00) – Sivakumar et al. (2000).

These observations clearly indicate the usefulness and suitability of the nonlinear local approximation method for river flow predictions, where the underlying river flow dynamics are captured in the phase-space step by step in local neighborhoods. As the local approximation approach is essentially a nonlinear deterministic procedure, a possible implication of the above extremely good river flow prediction results is that the underlying dynamics are deterministic. The low correlation dimensions (< 6) obtained for these series seem to support the above interpretation. Also, comparisons of predictions from local approximation methods with those from stochastic methods (e.g. Jayawardena and Gurung, 2000) and artificial neural networks (e.g. Lambrakis et al., 2000; Sivakumar et al., 2002a, c) indicate that the former perform better (sometimes significantly) or at least equally well. This result is additional evidence to the possible existence of deterministic behavior in the river flow dynamics.

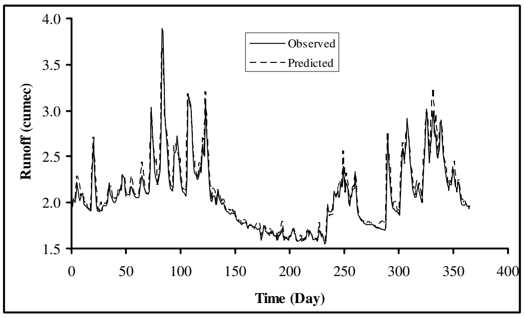

Figure 1. Comparison between observed and predicted values: daily river flow series from Lindenborg catchment in Denmark [from Islam and Sivakumar (2002)].

Prediction vs. embedding dimension

Figure 2, for instance, presents the relationship between the prediction accuracy (in terms of correlation coefficient) and phase-space (embedding) dimension when the nonlinear prediction method is employed to the daily flow series from Nakhon Sawan station at the Chao Phraya River basin in Thailand (Sivakumar et al., 2002c). As may be seen, the prediction accuracy increases with the increase in the embedding dimension up to a certain point (m = 3) and then decreases when the dimension is increased further. The presence of this low optimal embedding dimension, i.e. m_{opt} = 3, seems to indicate that the river flow dynamics exhibit low-dimensional chaotic behavior, dominantly governed by only a very few (in the order of 3) nonlinear interdependent variables. The decrease in the prediction accuracy (rather than an expected saturation) with the increase in the embedding dimension could be due to the presence of noise in the data [see, for instance, Porporato and Ridolfi (1997) and Sivakumar

(2000) for details]. Low optimal embedding dimensions have also been observed for several other river flow series studied for prediction purposes (e.g. Porporato and Ridolfi, 1997; Liu et al., 1998; Lambrakis et al., 2000; Sivakumar et al., 2001, 2002a; Islam and Sivakumar, 2002), indicating the possible low-dimensional deterministic nature of such river flow dynamics.

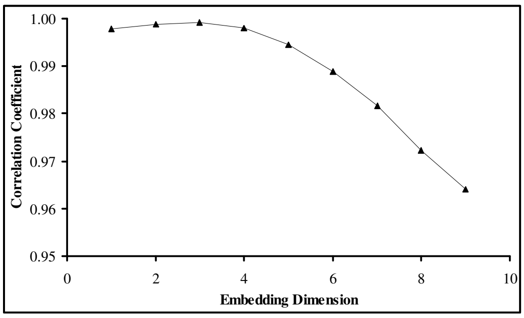

Figure 2. Prediction accuracy against embedding dimension: daily river flow series from Nakhon Sawan station at the Chao Phraya River basin in Thailand [from Sivakumar et al. (2002c)].

Prediction vs. lead time

For a chaotic system, due to its sensitivity to initial conditions, prediction accuracy decreases rapidly when the prediction lead time increases. Such a result is observed in the predictions obtained for the monthly river flow series from the Coaracy Nunes/Araguari River watershed in northern Brazil, shown in Figure 3 (Sivakumar et al., 2001), indicating the presence of chaotic behavior in the river flow dynamics. Similar results have also been reported for several other river flow series (e.g. Porporato and Ridolfi, 1997; Lambrakis et al., 2000; Lisi and Villi, 2001; Sivakumar et al., 2002c), revealing the deterministic nature of the underlying dynamics.

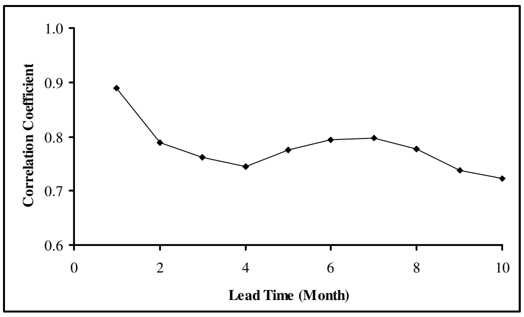

Figure 3. Prediction accuracy against lead time: monthly river flow series from Coaracy Nunes/Araguari River watershed in northern Brazil [from Sivakumar et al. (2001)].

Prediction vs. number of neighbors

Even though, extensive details of the effect of number of neighbors used in the nonlinear prediction method on the river flow predictions have not been made available [except the study by Jayawardena et al. (2002)], results from the trial-and-error procedures, as adopted in most of the studies, indicate that the best predictions are achieved only when the number of neighbors is small (e.g. Porporato and Ridolfi, 1997, 2001; Lisi and Villi, 2001; Sivakumar, 2003), i.e. typically less than 10% of the total number of neighbors. These results are an indication that the underlying dynamics are more of a deterministic nature, rather than of a stochastic nature.

4. CONCLUSIONS

Studies reporting possible presence of low-dimensional deterministic chaos in river flow dynamics are often criticized, due essentially to the potential limitations of the correlation dimension method, which is widely used to identify chaos. An attempt was made herein to address this issue by identifying chaos from river flow predictions (i.e. referred to as “inverse approach”). A review of studies on river flow predictions using nonlinear deterministic prediction methods provided convincing evidence to the possible low-dimensional deterministic nature of river flow dynamics. More importantly, the reported prediction results (in one way or another) also supported (qualitatively and quantitatively) the correlation dimension results reported by such studies. The present observations clearly reveal that nonlinear low-dimensional deterministic techniques are a viable alternative for studying river flow dynamics, if sufficient caution is exercised in the implementation of the available methods and also in the interpretation of the results.

5. REFERENCES

Abarbanel, H.D.I., *Analysis of Observed Chaotic Data*, Springer-Verlag, New York, 1996.

Casdagli, M., Nonlinear prediction of chaotic time series, *Physica D*, 35, 335-356, 1989.

Casdagli, M., Chaos and deterministic versus stochastic nonlinear modeling, *Journal of the Royal Statistical Society B*, 54(2), 303-328, 1991.

Farmer, D.J., and J.J. Sidorowich, Predicting chaotic time series, *Physical Review Letters*, 59, 845-848, 1987.

Grassberger, P., and I. Procaccia, Measuring the strangeness of strange attractors, *Physica D*, 9, 189-208, 1983.

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

Jayawardena, A.W., and A.B. Gurung, Noise reduction and prediction of hydrometeorological time series: dynamical systems approach vs. stochastic approach, *Journal of Hydrology,* 228, 242-264, 2000.

Jayawardena, A.W., W.K. Li, and P. Xu, Neighborhood selection for local modeling and prediction of hydrological time series, *Journal of Hydrology*, 258, 40-57, 2002.

Lambrakis, N., A.S. Andreou, P. Polydoropoulos, E. Georgopoulos, and T. Bountis, Nonlinear analysis and forecasting of a brackish karstic spring, *Water Resources Research*, 36(4), 875-884, 2000.

Lisi, F., and V. Villi, Chaotic forecasting of discharge time series: A case study, *Journal of the American Water Resources Association*, 37(2), 271-279, 2001.

Liu, Q., S. Islam, I. Rodriguez-Iturbe, and Y. Le, Phase-space analysis of daily streamflow: characterization and prediction, *Advances in Water Resources*, 21, 463-475, 1998.

Lorenz, E.N., Deterministic nonperiodic flow, *Journal of the Atmospheric Sciences*, 20, 130-141, 1963.

Osborne, A.R., and A. Provenzale, Finite correlation dimension for stochastic systems with power-law spectra, *Physica D*, 35, 357-381, 1989.

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

Porporato, A., and L. Ridolfi, Multivariate nonlinear prediction of river flows, *Journal of Hydrology*, 248(1-4), 109-122, 2001.

Schertzer, D., I. Tchiguirinskaia, S. Lovejoy, P. Hubert, and H. Bendjoudi, Which chaos in the rainfall-runoff process? A discussion on 'Evidence of chaos in the rainfall-runoff process' by Sivakumar et al., *Hydrological Sciences Journal*, 47(1), 139-147, 2002.

Sivakumar, B., Chaos theory in hydrology: Important issues and interpretations, *Journal of Hydrology*, 227(1-4), 1-20, 2000.

Sivakumar, B., Sediment transport dynamics: Assessing the suitability of low-dimensional deterministic approaches, *Advances in Water Resources* (in preparation), 2003.

Sivakumar, B., R. Berndtsson, J. Olsson, K. Jinno, and A. Kawamura, Dynamics of monthly rainfall-runoff process at the Göta basin: A search for chaos, *Hydrology and Earth System Sciences*, 4(3), 407-417, 2000.

Sivakumar, B., R. Berndtsson, and M. Persson, Monthly runoff prediction using phase-space reconstruction, *Hydrological Sciences Journal*, 46(3), 377-388, 2001.

Sivakumar, B., and A.W. Jayawardena, An investigation of the presence of low-dimensional chaotic behavior in the sediment transport phenomenon, *Hydrological Sciences Journal*, 47(3), 405-416, 2002.

Sivakumar, B., M. Persson, R. Berndtsson, and C.B. Uvo, Is correlation dimension a reliable indicator of low-dimensional chaos in short hydrological time series? *Water Resources Research*, 38(2), 10.1029/2001WR000333, 3-1 - 3-8, 2002a.

Sivakumar, B., R. Berndtsson, J. Olsson, and K. Jinno, Reply to "Which chaos in the rainfall-runoff process," *Hydrological Sciences Journal*, 47(1), 149-158, 2002b.

Sivakumar, B., A.W. Jayawardena, and T.M.G.H. Fernando, River flow forecasting: Use of phase-space reconstruction and artificial neural networks approaches, *Journal of Hydrology*, 265(1-4), 225-245, 2002c.

Stehlik, J., Deterministic chaos in runoff series, *Journal of Hydrology and Hydromechanics*, 47(4), 271-287, 1999.

Sugihara, G., and R.M. May, Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series, *Nature*, 344, 734-741, 1990.

Takens, F., Detecting strange attractors in turbulence, in *Dynamical Systems and Turbulence, Lecture Notes in Mathematics 898*, D.A. Rand and L.S. Young, 366-381 (eds), Springer-Verlag, Berlin, 1981.