Streamflow Dynamics in the Western United States: A Nonlinear Dynamical Analysis

B. Sivakumar[a]

[a] *Department of Land, Air & Water Resources, University of California, Davis, CA 95616, USA* (sbellie@ucdavis.edu)

Abstract: An attempt is made in the present study to understand the dynamics of streamflow in the western United States from a nonlinear dynamical perspective. Monthly streamflow data observed at 79 stations across 11 States in the western United States are analyzed. A nonlinear prediction method with a local approximation approach is employed. The method uses the concept of phase-space reconstruction to represent the underlying dynamics, i.e. reconstruction of the single-dimensional (or variable) streamflow series in a multi-dimensional phase-space. The analysis is carried out by categorizing the 79 stations into three groups on the basis of the magnitude of streamflow, as: (1) low-flow stations, having mean streamflow values less than 100 ft^3/s (2.382 m^3/s); (2) high-flow stations, having mean streamflow values more than 1000 ft^3/s (23.82 m^3/s); and (3) medium-flow stations, having mean streamflow values between 100 ft^3/s (2.382 m^3/s) and 1000 ft^3/s (23.82 m^3/s). The local approximation prediction approach yields, in general, reasonably good results for all these three categories, but the predictions for the low-flow stations are found to be relatively better than that obtained for the high-flow and medium-flow stations.

Keywords: Streamflow; Western United States; Nonlinear dynamics; Phase-space reconstruction; Local approximation prediction

1. INTRODUCTION

Understanding streamflow dynamics constitutes one of the most important problems in hydrology and water resources, as the topic is of relevance for the proper management of our water resources. Streamflow dynamics is governed by various physical mechanisms acting on a wide range of temporal and spatial scales. For instance, streamflow depends not only on the distribution of rainfall in time and space, but also on the type and the state of the basin, which, in turn, depend on climatic condition, vegetation-state, etc. Also, almost all mechanisms involved in the streamflow process present some degree of nonlinearity. All these factors make the problem of streamflow modeling non-trivial.

During the past few decades, a great deal of research has been devoted to the formulation and development of approaches and models to understand streamflow dynamics, and significant progress has been made. For several decades, a common belief among hydrologists has been that the tremendous variability observed in streamflow (and other hydrological phenomena) in time and space is due to the influence of a large number of dominant variables. Consequently, a majority of the investigations until now has employed the concept of a stochastic process for modeling the dynamics of streamflow (and other hydrological phenomena) [e.g. Thomas and Fiering, 1962; Matalas, 1967; Carlson et al., 1970; Delleur et al., 1976; Obeysekara and Salas, 1986]. However, the last decade, with the advent of deterministic chaos theory and other nonlinear concepts, has seen a significant shift in our approach to modeling the dynamics of streamflow (and other hydrological phenomena), as it has become increasingly realized that apparently irregular behavior could be the outcome of simple deterministic systems influenced by a few nonlinear interdependent variables. Applications of the concept of chaos theory towards improving our understanding of dynamics of streamflow (and other hydrological phenomena) have been on the rise [e.g. Rodriguez-Iturbe et al., 1989; Jayawardena and

Lai, 1994; Porporato and Ridolfi, 1997; Liu et al., 1998; Sivakumar et al., 1999, 2001a, b; Sivakumar, 2000]. The outcomes of such studies are encouraging, as reasonably good predictions are achieved for hydrological phenomena, in particular streamflow [e.g. Porporato and Ridolfi, 1997; Liu et al., 1998; Sivakumar et al., 2001b].

It is relevant to note, at this point, that both the stochastic and chaotic approaches possess certain important limitations that may hamper our ability to accurately model the streamflow dynamics. In view of such limitations, the appropriate approach for modeling the tremendously variable streamflow in time and space is system (or data) dependent. Consequently, a general assessment regarding whether one approach is better than the other or vice-versa is difficult to provide. It must be emphasized, however, that studies that have employed both approaches have revealed that the chaotic approach was better than the stochastic approach for the streamflow series analyzed [e.g. Jayawardena and Lai, 1994; Jayawardena and Gurung, 2000]. The author is not aware of any studies that have reported the opposite situation.

Having said the above, this study is not intended to participate in the continuing debate on the (identification of) appropriate approach for streamflow modeling. Rather, it attempts to investigate in detail the streamflow data at hand, with the premise that either of these two approaches could be appropriate for the problem, though it employs the chaotic approach. The streamflow studied is the monthly streamflow observed at 79 stations spread throughout 11 States in the western United States. Though one cannot deny that modeling streamflow dynamics at high resolutions, such as daily, is crucial for purposes like flood forecasting, the reasons for considering the monthly streamflow data in the present study are as follows: (1) The assessment of the quantity (and quality) of water availability in a basin and, hence, its management for efficient water supply is generally made over longer periods of time, such as months and years, rather than at daily scale; and (2) Monthly and annual streamflow series are more related to long-term climate as compared to that of the daily series and, therefore, streamflow dynamics at these scales may be used to establish connections between climate and streamflow.

2. STUDY AREA AND DATA

In this study, monthly streamflow dynamics in the western United States is studied using data collected over a period of 62 years (1932-1993) at 79 stations. These stations are located over 11 States, as follows: Arizona (AZ) – 1 station, Arizona/New Mexico (AZ/NM) - 1, California (CA) - 21, Colorado (CO) - 2, Colorado/Wyoming (CO/WY) - 1, Idaho (ID) - 10, Montana (MT) - 5, New Mexico (NM) - 9, Nevada (NV) - 2, Oregon (OR) - 9, Utah (UT) - 5, Washington (WA) - 12, and Wyoming (WY) - 1. Table 1 presents the streamflow stations State wise (numbers arranged starting from 01).

Table 1. State wise streamflow stations.

State	Station numbers
AZ	19
AZ/ NM	17
CA	24, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 46, 47
CO	05, 06
CO/WY	04
ID	53, 54, 59, 64, 65, 66, 68, 69, 70, 71
MT	01, 02, 03, 55, 56
NM	07, 08, 09, 10, 11, 12, 13, 16, 18,
NV	25, 26
OR	45, 67, 72, 73, 74, 76, 77, 78, 79
UT	15, 20, 21, 22, 23
WA	48, 49, 50, 51, 52, 57, 58, 60, 61, 62, 63, 75
WY	14

Streamflow data in the US is commonly expressed in 'water years,' which commence in October. The records used herein start in October

1931 and end in September 1993 and are average monthly streamflow (see Slack and Landwehr [1992] and Pichota et al. [1997] for details).

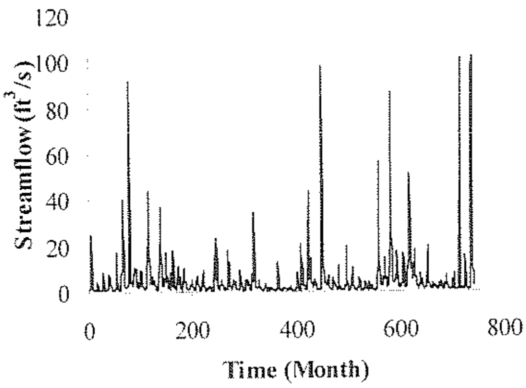
The magnitude of the streamflow varies greatly among the stations, due to both the different climatic regions in the western US and the different drainage basin characteristics associated with each streamflow station. Most of the drainage basins are medium to small sized (< 1000 km^2) and are located in middle to high elevations (> 500 m). For the present analysis, the 79 stations are grouped under three categories based on the magnitude of streamflow, as follows: (1) low-flow stations, having mean streamflow values less than 100 ft^3/s (2.382 m^3/s); (2) high-flow stations, having mean values more than 1000 ft^3/s (23.82 m^3/s); and (3) medium-flow stations, with mean values between the above two. The number of stations in each State falling under the three categories is as follows: Low-flow: CA (8), CO (2), ID (1), NV (1), NM (6), and UT (2); Medium-flow: AZ (1), AZ/NM (1), CA (10), CO/WY (1), ID (4), MT (1), NV (1), NM (3), OR (4), UT (3), WA (1), WY (1); and High-flow: CA (3), ID (5), MO (4), OR (5), WA (11). Table 2 presents the station numbers and the total number of stations falling in each of the above three categories.

Table 2. Streamflow stations categorized based on mean flow value.

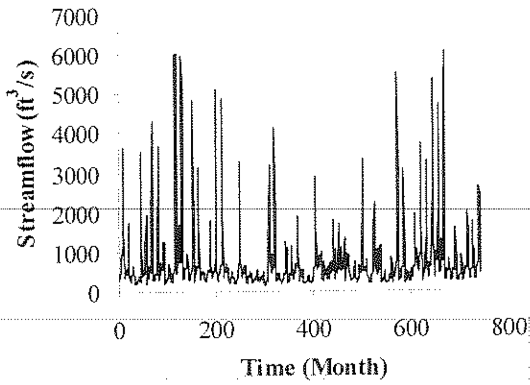
Streamflow	Stations
Low-flow Regime (Mean < 100 ft^3/s)*	05, 06, 07, 08, 09, 10, 12, 13, 22, 23, 24, 26, 27, 28, 29, 32, 33, 36, 37, 65 (Total: 20 stations)
Medium-flow Regime (100 $ft^3/s \leq$ Mean $\leq$ 1000 ft^3/s)*	02, 04, 11, 14, 15, 16, 17, 18, 19, 20, 21, 25, 30, 31, 34, 35, 38, 39, 40, 41, 42, 43, 45, 48, 53, 54, 64, 67, 68, 78, 79 (Total: 31 stations)
High-flow Regime (Mean > 1000 ft^3/s)*	01, 03, 44, 46, 47, 49, 50, 51, 52, 55, 56, 57, 58, 59, 60, 61, 62, 63, 66, 69, 70, 71, 72, 73, 74, 75, 76, 77 (Total: 28 stations)

* 1 ft3/s = 0.0238 m^3/s.

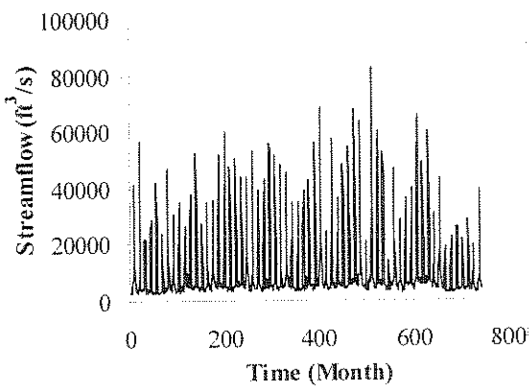
Figures 1(a) to 1(c) show the variation of streamflow in three stations (nos. 28, 43, and 69) representing the low, medium and high flow categories, respectively. These are located in California, New Mexico and Idaho, respectively.



(a)



(b)



(c)

Figure 1. Streamflow series in the western US representing (a) low-flow (sta. 28); (b) medium-flow (sta. 43); and (c) high-flow (sta. 69). (Units: 1 ft^3/s = 0.0238 m^3/s).

3. NONLINEAR LOCAL APROXIMATION PREDICTION

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

$$Y_j = (X_j, X_{j+\tau}, X_{j+2\tau}, ..., X_{j+(m-1)\tau}) \quad (1)$$

where $j = 1, 2, ..., N-(m-1)\tau$, m is the dimension of the vector Y_j, called as embedding dimension; and τ is a delay time [Packard *et al.*, 1980; 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,

$$Y_{j+T} = f_T(Y_j) \quad (2)$$

where Y_j and 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. In this study, a local approximation approach [e.g. Farmer and Sidorowich, 1987] is employed. 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 system dynamics is 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 Y_j from which the forecast is initiated, identifying neighbors Y_j^p, $p = 1, 2, ..., k$, with $j^p < j$, nearest to Y_j, which constitute the set corresponding to Y_j. The local functions can then be built, which take each point in the neighborhood to the next neighborhood: Y_j^p to Y_{j+1}^p. The local map F_T, which does this, is determined by a least squares fit minimizing

$$\sum_{p=1}^{k} \| Y_{j+1}^{p} - F_T Y_j^{p} \|^2 \quad (3)$$

In this study, the local maps are learned in the form of local polynomials [e.g. Abarbanel, 1996], and the predictions are made forward from a new point Z_0 using these local maps. For the new point Z_0, the nearest neighbor in the learning or training set is found, which is denoted as Y_q. Then the evolution of Z_0 is found, which is denoted as Z_1 and is given by

$$Z_1 = F_q(Z_0) \quad (4)$$

The nearest neighbor to Z_1 is then found, and the procedure is repeated to predict the subsequent values. The prediction accuracy is evaluated using the correlation coefficient (ρ). The time series plots are also used to choose the best prediction results among a large combination of results achieved with different embedding dimensions.

4. RESULTS AND DISCUSSION

The nonlinear local approximation method is now employed for predicting the monthly streamflow dynamics at the above 79 stations. The first 500 values in the series are used in the phase-space reconstruction for predicting the next 200 values. One-step ahead predictions are made for phase-spaces reconstructed with embedding dimensions from 1 to 9.

Figures 2(a) to 2(c), for instance, present a comparison, using time series plots, of the predicted and the observed streamflow values for stations 28, 43, and 69 (representing the low, medium and high flow stations, respectively). The plots shown correspond to the best predictions obtained (i.e. when the time series is reconstructed in a three-dimensional phase space). As can be seen, the predicted values are in reasonable agreement with the observed ones, except when the observed values are high. A closer look at the observed and predicted time series reveals that the local polynomial prediction approach captures reasonably not only the major trends but also the minor fluctuations in the streamflow. Similar results are achieved also for the other stations (Figures not shown).

Table 3 presents a summary of the streamflow prediction results (correlation coefficient, ρ) achieved for the 79 stations. As can be seen, all the stations, irrespective of the flow-regime, have ρ values higher than 0.70. Out of 20 stations in the low-flow category, 11 stations have ρ above 0.90, 8 stations have ρ between 0.80 and 0.90, and only one station has ρ less than 0.80. Out of 31 stations in the medium-flow category, 18 stations have ρ above 0.90, 10 stations have ρ between 0.80 and 0.90, and only 3 stations have ρ less than 0.80. Of the 28 stations in the high-flow category, 5 stations have ρ above 0.90, 21 stations have ρ between 0.80 and 0.90, and only 2 stations have ρ less than 0.80.

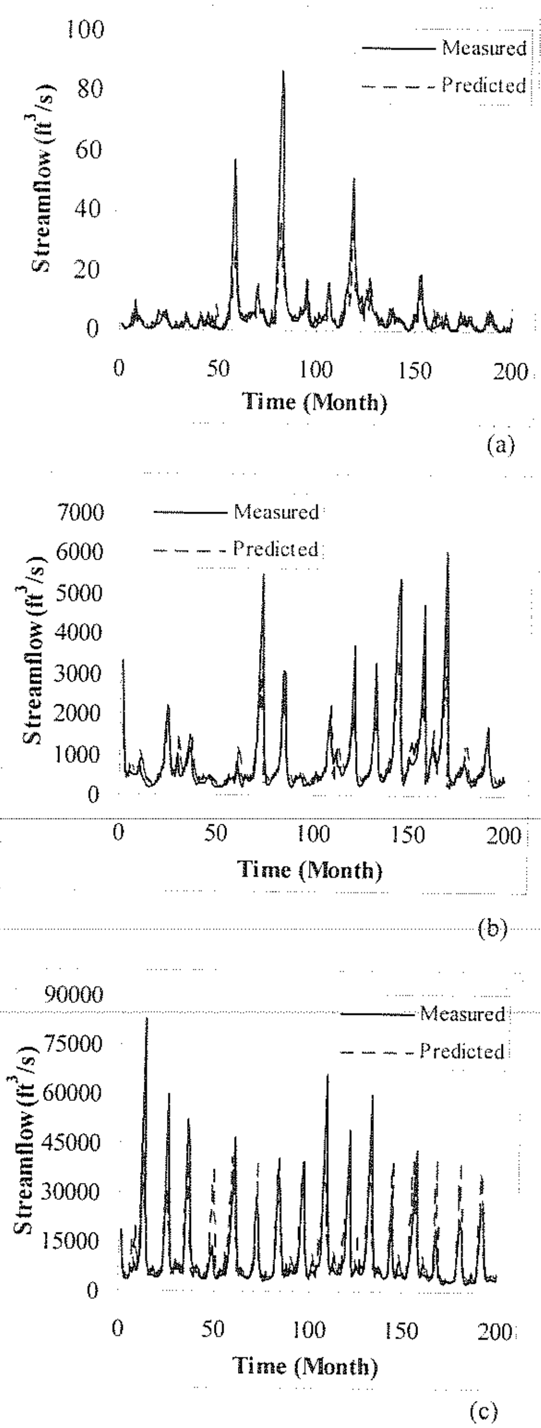


Figure 2. Examples of comparison between observed and predicted streamflow representing (a) low-flow; (b) medium-flow; and (c) high-flow stations.
(Units: 1 $ft^3/s = 0.0238\ m^3/s$).

The high ρ values and also the good agreement in the time series plots between the observed and predicted values indicate the suitability of the nonlinear local approximation method for predicting the streamflow dynamics. The reconstruction of the single variable (streamflow) series in multi-dimensional phase-space is, therefore, found to be capable of capturing the important features of the underlying dynamics.

Table 3. Results of nonlinear prediction analysis.

Low-flow stations		Medium-flow stations		High-flow stations	
Sta. No.	ρ	Sta. No.	ρ	Sta. No.	ρ
05	0.912	02	0.802	01	0.824
06	0.832	04	0.921	03	0.773
07	0.862	11	0.885	44	0.867
08	0.956	14	0.892	46	0.873
09	0.927	15	0.856	47	0.836
10	0.877	16	0.881	49	0.889
12	0.936	17	0.921	50	0.894
13	0.889	18	0.948	51	0.827
22	0.888	19	0.939	52	0.904
23	0.913	20	0.972	55	0.851
24	0.895	21	0.701	56	0.902
26	0.947	25	0.956	57	0.801
27	0.952	30	0.940	58	0.834
28	0.951	31	0.964	59	0.837
29	0.967	34	0.853	60	0.861
32	0.905	35	0.922	61	0.856
33	0.841	38	0.924	62	0.888
36	0.939	39	0.976	63	0.885
37	0.876	40	0.932	66	0.931
65	0.764	41	0.909	69	0.795
		42	0.940	70	0.845
		43	0.923	71	0.875
		45	0.949	72	0.928
		48	0.878	73	0.875
		53	0.810	74	0.906
		54	0.746	75	0.852
		64	0.959	76	0.894
		67	0.866	77	0.857
		68	0.757		
		78	0.890		
		79	0.912		

5. SUMMARY AND CONCLUSIONS

A nonlinear dynamical approach was employed to predict the streamflow dynamics in the western United States. The approach was based on the concept of phase-space reconstruction, i.e. reconstruction of single-variable series in a multi-dimensional phase-space to represent the underlying dynamics, and a local approximation method was used for making predictions. Monthly streamflow data observed at 79 stations spread over 11 States in the western US were analyzed, by grouping the stations into three categories, namely low, medium and high flow. The local approximation method yielded good results for streamflow from all the 79 stations, but was found to perform better for the low-flow stations. The method captured both the major trends and the minor fluctuations. The results also revealed that, in most cases, the best results were achieved when the series was reconstructed in a three-dimensional phase-space, implying that the streamflow dynamics could be understood from a low dimensional chaotic perspective.

6. REFERENCES

Abarbanel, H.D.I., *Analysis of observed chaotic data*, Springer-Verlag, New York, USA, 1996.

Carlson, R.F., J.A. McCormick, and D.G. Watts, Application of linear random models to four annual stream-flows series, *Water Resources Research*, 6, 1070-1078, 1970.

Delleur, J.W., P.C. Tao, and M.L. Kavvas, An evaluation of the practicality and complexity of some rainfall and runoff time series models, *Water Resources Research*, 12, 953-979, 1976.

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

Jayawardena, A.W., and F. Lai, Analysis and prediction of chaos in rainfall and stream flow time series, *Journal of Hydrology*, 153, 23-52, 1994.

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.

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.

Matalas, N.C., Mathematical assessment of synthetic hydrology, *Water Resources Research*, 3, 937-947, 1967.

Obeysekera, J.T.B., and J.D. Salas, Modeling of aggregated hydrologic time series, *Journal of Hydrology*, 86, 197-219, 1986.

Packard, N.H., J.P. Crutchfield, J.D. Farmer, and R.S. Shaw, Geometry from a time series, *Physical Review Letters*, 45(9), 712-716, 1980.

Pichota, T.C., J.A. Dracup, and R.C. Fovell, Western US streamflow and atmospheric circulation patterns during El Nino-Southern Oscillation, *Journal of Hydrology*, 201, 249-271, 1997.

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

Rodriguez-Iturbe, I., F.B. De Power, M.B. Sharifi, and K.P. Georgakakos, Chaos in rainfall, *Water Resources Research*, 25(7), 1667-1675, 1989.

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

Sivakumar, B., S.Y. Liong, C.Y. Liaw, and K.K. Phoon, Singapore rainfall behavior: chaotic? *Journal of Hydrologic Engineering, ASCE*, 4(1), 38-48, 1999.

Sivakumar, B., S. Sorooshian, H.V. Gupta, and X. Gao, A chaotic approach to rainfall disaggregation, *Water Resources Research*, 37(1), 61-72, 2001a.

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

Slack, J.R., and V.M. Landwehr, Hydro-Climatic Data Network (HCDN): A U. S. Geological Survey Streamflow Data Set for the United States for the Study of climate variations, 1847-1988, U. S. Geological Survey Open File Report 92-129, 1992.

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.

Thomas, H. A., and B. M. Fiering, Mathematical synthesis of streamflow sequences for the analysis of river basins by simulation, in *Design of Water-Resource Systems*, A. Mass et al. (eds), pp. 459-493, Harvard Univ. Press, Cambridge, Mass., 1962.