



Hydrologic system complexity and nonlinear dynamic concepts for a catchment classification framework

B. Sivakumar^{1,2} and V. P. Singh³

¹School of Civil and Environmental Engineering, The University of New South Wales, Sydney, Australia

²Department of Land, Air and Water Resources, University of California, Davis, USA

³Department of Biological and Agricultural Engineering & Department of Civil and Environmental Engineering, Texas A & M University, College Station, USA

Correspondence to: B. Sivakumar (s.bellie@unsw.edu.au)

Received: 12 April 2011 – Published in Hydrol. Earth Syst. Sci. Discuss.: 2 May 2011

Revised: 5 October 2012 – Accepted: 16 October 2012 – Published: 8 November 2012

Abstract. The absence of a generic modeling framework in hydrology has long been recognized. With our current practice of developing more and more complex models for specific individual situations, there is an increasing emphasis and urgency on this issue. There have been some attempts to provide guidelines for a catchment classification framework, but research in this area is still in a state of infancy. To move forward on this classification framework, identification of an appropriate basis and development of a suitable methodology for its representation are vital. The present study argues that hydrologic system complexity is an appropriate basis for this classification framework and nonlinear dynamic concepts constitute a suitable methodology. The study employs a popular nonlinear dynamic method for identification of the level of complexity of streamflow and for its classification. The correlation dimension method, which has its base on data reconstruction and nearest neighbor concepts, is applied to monthly streamflow time series from a large network of 117 gaging stations across 11 states in the western United States (US). The dimensionality of the time series forms the basis for identification of system complexity and, accordingly, streamflows are classified into four major categories: low-dimensional, medium-dimensional, high-dimensional, and unidentifiable. The dimension estimates show some “homogeneity” in flow complexity within certain regions of the western US, but there are also strong exceptions.

1 Introduction

As in most other fields of science and engineering, growth in the field of hydrology during the past century has been unprecedented, largely driven by the invention of powerful computers, measurement devices, remote sensors, geographic information systems (GIS), digital elevation models (DEM), and networking facilities. This growth may be viewed in terms of: (1) the various sub-fields that have been created essentially to “break down” hydrology into specific components for more focused and detailed studies (e.g. surface hydrology, subsurface hydrology, groundwater hydrology, forest hydrology, mountain hydrology, urban hydrology, isotope hydrology, snow and glacier hydrology, ecohydrology); and (2) the numerous scientific theories and mathematical techniques that have been developed/applied for modeling and prediction of hydrologic systems and the associated processes (e.g. deterministic techniques, stochastic methods, scaling and fractal theories, artificial neural networks, chaos theory, wavelets, entropy theory, evolutionary computing).

Despite this growth, there remain many grand challenges in performing good hydrologic teaching, research, and practice. Among others, two major concerns are dominating discussions and debates on current hydrologic studies: (1) hydrologic models being developed are often more complex, having too many parameters and requiring too much data, than perhaps needed; and (2) models are often developed for specific situations, and their extensions and generalizations to other situations are rather difficult. In addition, our general lack of emphasis in studying the crucial connections

between the (model) theories and the actual system properties (e.g. data), our increasing emphasis in applying specific (and often pre-selected) mathematical techniques independently as opposed to the integration of techniques for modeling hydrologic systems, and our focus mainly on local-scale hydrologic problems rather than global-scale hydrologic issues have also come under severe scrutiny (e.g. Beven, 2002; Sivakumar, 2008). With growing concerns on the occurrence of global climate change and its potential impacts on water resources and the environment (including more frequent and greater magnitudes of extreme events, such as floods and droughts), the limitations of the “confines of traditional hydrology” and the need to go beyond and perform cross-disciplinary research integrating hydrology with atmospheric science, geomorphology, geochemistry, ecology, and other areas have also been increasingly recognized (see, for example, Paola et al., 2006, for some details).

In view of these concerns, many studies during the past decade or so have emphasized the need for simplification in modeling wherever possible as well as a common framework in hydrology (e.g. Grayson and Blöschl, 2000; McDonnell and Woods, 2004). Within this context, some attempts have also been made towards a catchment classification framework (e.g. Snelder et al., 2005; Sivakumar et al., 2007; see also the other articles in the current special issue “Catchment Classification and PUB” for some latest studies), with an aim to streamline catchments into different groups and sub-groups on the basis of their salient characteristics (e.g. data and process complexity) and to provide directions to model developers on the level of model complexity to invoke. Nevertheless, these attempts are only preliminary and research in this direction is still in a state of infancy. Indeed, there are even questions on the basic form of the classification framework and on the components to be included (e.g. Wagener et al., 2007). Therefore, identification of an appropriate basis for the classification framework and development of a suitable methodology are crucial for moving forward in hydrology.

The present study attempts to offer some workable guidelines for an appropriate basis and a suitable methodology towards a classification framework in hydrology. The study argues, through highlighting the relevance of complexity and nonlinearity in hydrologic systems, that system complexity is an appropriate basis for the classification framework and nonlinear dynamic concepts constitute a suitable methodology for assessing system complexity. With this, it examines the usefulness of a nonlinear dynamic method for streamflow classification. This is done by employing the correlation dimension method (e.g. Grassberger and Procaccia, 1983a) to streamflow from a large network of gaging stations in the western United States. Monthly streamflow data observed over a period of 52 yr from 117 gaging stations across 11 states are considered for analysis. The identification of the level of complexity and the subsequent classification are

made based on the dimensionality of the streamflow time series.

The rest of this paper is organized as follows. Section 2 presents a brief account of major attempts on classification in hydrology. Section 3 highlights the role of complexity and nonlinearity in hydrologic systems. Section 4 describes the correlation dimension method. Section 5 presents the details of streamflow data from the western United States and results of their analysis. Conclusions and directions for further research are presented in Sect. 6.

2 Classification in hydrology: a brief history and scope

The realization of the need for a classification framework in hydrology is not entirely new. It had indeed been discussed some time ago, and since then several studies have also attempted to advance the idea. These studies have investigated different ways for developing such a framework and their implications, including river morphology (e.g. Rosgen, 1994; Poff et al., 2006), river/flow regimes (e.g. Beckinsale, 1969; Haines et al., 1988), landscape and land use parameters (e.g. Merz and Blöschl, 2004; Wardrop et al., 2005), similarity indices (e.g. Olden and Poff, 2003; Ali et al., 2012), eco-hydrologic factors (e.g. Harris et al., 2000; Olden et al., 2011), geostatistical properties (e.g. Vormoor et al., 2011), entropy (e.g. Krasovkaia, 1997), nonlinear and chaotic dynamic properties (e.g. Sivakumar et al., 2007), and other relevant characteristics/methods (e.g. Chapman, 1989; Isik and Singh, 2008). Extensive details of these studies are available both in the traditional hydrologic literature and in related fields (e.g. geomorphology, ecohydrology, and freshwater biology); for some very latest accounts, see Ali et al. (2012) and also the articles in the current special issue “Catchment Classification and PUB.”

Although useful in their own ways, these studies are largely inadequate for a generic classification framework. In addition to the limitations that exist in each of the different forms, a coherent effort to bring these disparate forms together for a workable classification framework is also missing. The urgency to formulate a generic classification framework in hydrology is increasingly realized now, especially with our current practice of employing more and more sophisticated mathematical techniques and developing more and more complex models for each and every individual hydrologic system/situation, rather than the emphasis needed for addressing broader-scale hydrologic issues (e.g. Sivakumar, 2008).

The fundamental idea behind a classification framework in hydrology is to streamline hydrologic systems into groups and sub-groups to recognize salient characteristics that are emblematic and to develop suitable methods/models. This classification also serves as a middle-ground to the following two extremes: (1) treatment of all hydrologic systems in the same way, regardless of the differences among them; and

(2) treatment of each and every individual hydrologic system in its own way, regardless of the similarities among them. Either of these approaches has enormous implications for modeling, including complexity of the models, data and computer requirement, accuracy of results, and overall understanding of the systems. The classification framework, therefore, is aimed at providing an optimum way of studying hydrologic systems, taking into account both minimization of costs and maximization of benefits. In the end, it should help modelers identify suitable catchments to apply their models to and also users to identify suitable models for their catchments.

For its usefulness to be realized both at the global and at the regional/local levels, the classification framework should be able to accommodate important general as well as specific characteristics of hydrologic systems/processes. The framework must also be simple enough and commonly agreeable to provide a “universal” language for communication and discussion in hydrology and water resources. The crucial questions now are: (1) What form should the classification framework assume? (2) What components need to be included? (3) What is the appropriate methodology for its formulation? and (4) How to effectively verify such a classification framework? A few studies have attempted to address these questions and relevant issues, such as the examples below.

Wagener et al. (2007) reviewed the existing approaches to define hydrologic similarity, which has often been invoked for classification purposes, and offered some general guidelines for catchment classification that include the use of catchment structure, hydro-climatic region, and catchment functional response, among others. They also identified the following requirements for a classification framework: (1) mapping catchment form/hydro-climatic conditions on catchment function across spatial and temporal scales; (2) including partition, storage, and release of water in catchment functions; (3) consideration of uncertainty in the metrics/variables used; and (4) basing on functions characterized by streamflow to start with and subsequently expanding to other more complex functions.

Using the Shannon entropy, Krasovskaia (1995, 1997) developed a quantitative methodology for studying river flow regimes and their classification. The entropy-based methodology involves: (1) classification of mean monthly flows into different types; (2) identification of discriminating periods for different classes; (3) specification of instability index; (4) computation of instability index value for each regime type; and (5) computation of instability index for all flow series. Another method for grouping river regimes, developed by Krasovskaia (1997), employs minimization of an entropy-based objective function. This function uses a concept of information loss resulting from flow aggregation and determining the difference between the series aggregated into one group.

Sivakumar et al. (2007) explored the utility of a simple nonlinear data reconstruction approach, called phase space

reconstruction, for assessing the complexity of hydrologic systems and, thus, for their classification. They used the “region of attraction of trajectories” in the phase space to identify data as exhibiting “simple” or “intermediate” or “complex” behavior and, correspondingly, classify the system as potentially low-, medium-, or high-dimensional. The utility of this reconstruction concept was first demonstrated on two artificial time series possessing significantly different characteristics and levels of complexity (purely random and low-dimensional deterministic), and then tested on a host of river-related data representing different geographic regions, climatic conditions, basin sizes, processes, and scales. The ability of the phase space to reflect the river basin characteristics and the associated mechanisms, such as basin size, smoothing, and scaling, was also observed. The “dimensionality” and “complexity” ideas used by Sivakumar et al. (2007) were along the lines of the dominant processes concept (DPC), which was originally introduced in the context of hydrologic model simplification (Grayson and Blöschl, 2000) and subsequently suggested as a potential means for formulation of a classification framework (e.g. Woods, 2002; Sivakumar, 2004a).

Following up on the preliminary ideas by Sivakumar et al. (2007) based on just a few example cases, we attempt here to advance the studies on nonlinear dynamic concepts for identifying complexity of hydrologic systems and for their classification. To this end, we particularly consider that the extent of “complexity” of the system is reflected by the “variability” of the representative (observed) data (i.e. streamflow in the present case), which, in turn, is assessed by its “dimensionality”. We apply the correlation dimension method for studying data dimensionality and system complexity, and use such information for classification purposes.

3 Complexity and nonlinearity in hydrologic systems

3.1 Complexity in hydrologic systems

Although words “complex” and “complexity” are widely used both in scientific theory and in common practice, there is no general consensus on the definition. Nevertheless, one workable definition may be this: “consisting of interconnected or interwoven parts”. Qualitatively, to understand the behavior of a complex system, we must understand not only the behavior of the parts but also how they act together to form the behavior of the whole. This is because: (1) we cannot describe the whole without describing each part; and (2) each part must be described also in relation to other parts. For a quantitative description, the central issue again is defining quantitatively what “complexity” means. In the specific context of classification of systems, such as the one addressed in this study, it may perhaps be even more useful to ask: (1) What do we mean when we say that one system is more complex than another? and (2) Is there a way

to identify the complexity of one system and to compare it with the complexity of another system? To develop a quantitative understanding of complexity, a variety of tools can be used. These may include: statistical (e.g. coefficient of variation), nonlinear dynamic (e.g. dimension), information theoretic (e.g. entropy), or some other measure. In this study, we discuss the nonlinear dynamic tools, which allow identification of complexity of different systems and interpretations and distinctions on “more complex” and “less complex” systems. In particular, we attempt to assess the complexity of the system in terms of variability of the data through dimension estimation.

Hydrologic phenomena arise as a result of interactions between climate inputs and landscape characteristics that occur over a wide range of space and time scales. Due to the tremendous heterogeneities in climatic inputs and landscape properties, such phenomena may be highly variable and “complex” at all scales. Consequently, they are not fully understood. In the absence of perfect knowledge, a simplified way to represent them may be through the concept of “system”. There are many different definitions of a system, but perhaps the simplest may be: “a system is a set of connected parts that form a whole”. Chow (1964) defined a system as an aggregate or assemblage of parts, being either objects or concepts, united by some form of regular interaction or inter-dependence. Dooge (1967a), however, defined a system as: “any structure, device, scheme, or procedure, real or abstract, that inter-relates in a given time reference, an input, cause, or stimulus, of matter, energy, or information and an output, effect, or response of information, energy, or matter”. This definition by Dooge is much more comprehensive and instructive.

With this system concept, the entire hydrologic cycle may be regarded as a hydrologic system, whose components might include precipitation, interception, evaporation, transpiration, infiltration, detention storage or retention storage, surface runoff, interflow, and groundwater flow, and perhaps other phases of the hydrologic cycle. Each component may be treated as a sub-system of the overall cycle, if it satisfies the characteristics of a system set out in its definition. Thus, the various components of the hydrologic system can be regarded as hydrologic sub-systems. To analyze the total system, the simpler sub-systems can be treated separately and the results combined according to the interactions between the sub-systems (especially with the assumption of linearity). Whether a particular component is to be treated as a system or sub-system depends on the “objective of the inquiry” (Singh, 1988).

In this “objective of the inquiry” context, Sivakumar (2008) suggests that hydrologic systems may be viewed from three different, but related, angles: process, scale, and purpose of interest. Depending upon the angle at which they are viewed, hydrologic systems may be either simple or complex; for example, the rainfall occurrence in a desert may be treated as an extremely simple process since there may be

no rainfall at all, while the runoff process in a large river basin may be highly complex due to the basin complexities and heterogeneities, in addition to rainfall variability. Consequently, hydrologic modeling must also be viewed from these three angles; in other words, the appropriate model to represent a given hydrologic system may also be either simple or complex. The obvious question, however, is: how simple or how complex should the models be? This issue is addressed in this study, since the basic purpose behind formulation of a catchment classification framework is the identification of the most appropriate model (type and complexity) for a given catchment.

Since complexity is a fundamental and central characteristic of hydrologic systems, and is also a representation of their generality and specificity, it should form the basis for a classification framework. The study by Sivakumar et al. (2007), for example, offers some clues as to the use of complexity (defined in terms of extent of data variability) as a viable means for a classification framework.

3.2 Nonlinearity in hydrologic systems

Much of the research in hydrologic systems, at least until the 1990s, has been based on the assumption of “linearity”; i.e., the relation between cause (e.g. input) and effect (e.g. output) is linear or proportional. One of the important factors that contributed to, or necessitated, this linear approach was the lack of computational power to develop the (perhaps more complex) nonlinear mathematical methods. However, the “nonlinear” behavior of hydrologic systems had been known for a long time (e.g. Izzard, 1966; Dooge, 1967b).

The nonlinear behavior of hydrologic systems is evident in various ways and at almost all spatial and temporal scales. The hydrologic cycle itself is an example of a system exhibiting nonlinear behavior, with almost all of the individual components themselves exhibiting nonlinear behavior as well. The climatic inputs and landscape characteristics are changing in a highly nonlinear fashion, and so are the outputs, often in unknown ways. The rainfall-runoff process is nonlinear, almost regardless of the basin area, land uses, rainfall intensity, and other influencing factors. In fact, the effects of nonlinearity can be tremendous, especially when the system is sensitively dependent on initial conditions. This means, even small changes in the inputs may result in large changes in the outputs (and large changes in the inputs may turn out to cause only small changes in the outputs), a situation popularly termed as “chaos” in the nonlinear science literature (e.g. Lorenz, 1963).

With significant developments in computational power during the past three decades or so, and also with major advances in measurement technology and mathematical concepts, studies on the nonlinearity and related properties of hydrologic systems have started to gain attention. Nonlinear stochastic methods (e.g. Kavvas, 2003), artificial neural networks (e.g. Govindaraju, 2000), data-based mechanistic

models (e.g. Young and Beven, 1994), and nonlinear dynamics and chaos (e.g. Sivakumar, 2000) are some of the popular nonlinear techniques that have found extensive applications in hydrology. This study discusses the utility of nonlinear dynamic techniques as a suitable methodology for studying the complexity of hydrologic systems and, thus, for formulation of a catchment classification framework. In particular, we apply a popular nonlinear dynamic method, the correlation dimension method, to streamflow time series for classification purposes.

4 Correlation dimension method

During the past three decades or so, significant advances have been made in the field of nonlinear sciences to study complex systems. Numerous methods have been developed and applied in various fields, including physics, chemistry, biology, earth sciences, ecology, economics, engineering, medicine, and psychology. Extensive details of the applications of nonlinear dynamics and chaos concepts in hydrology are found in Sivakumar (2000) and in the broader field of geophysics in Sivakumar (2004b).

Popular among the methods developed within the context of nonlinear dynamic and chaos theories are correlation dimension, Lyapunov exponent, false nearest neighbors, nonlinear prediction, surrogate data, and redundancy methods. Almost all of these methods involve data embedding and nearest neighbor search, identifying different yet related properties of the underlying system dynamics. In this study, we employ the correlation dimension method for complexity determination of time series.

Correlation dimension is a measure of the extent to which the presence of a data point affects the position of the other points lying on the attractor in (a multi-dimensional) phase space or coordinate system. The correlation dimension method uses the correlation integral (or function) for determining the dimension of the attractor in the phase space and, hence, for distinguishing, broadly, between low-dimensional and high-dimensional systems. The concept of the correlation integral is that a time series arising from deterministic dynamics will have a limited number of degrees of freedom equal to the smallest number of first-order differential equations that capture the most important features of the dynamics. Thus, when one constructs phase spaces of increasing dimension, a point will be reached where the dimension equals the number of degrees of freedom, beyond which increasing the phase space dimension will not have any significant effect on correlation dimension.

Many algorithms have been formulated for the estimation of the correlation dimension of a time series. Among these, the Grassberger–Procaccia algorithm (Grassberger and Procaccia, 1983a) has been and continues to be the most widely used one, especially in hydrologic studies. The algorithm uses the concept of phase space reconstruction (e.g.

Packard et al., 1980) for representing the dynamics of the system from an available single-variable time series. Given a single-variable series, X_i , where $i = 1, 2, \dots, N$, a multi-dimensional phase space can be reconstructed as (Takens, 1981):

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

where $j = 1, 2, \dots, N - (m - 1)\tau$; m is the dimension of the vector Y_j , called embedding dimension; and τ is an appropriate delay time, which is an integer multiple of sampling time. It must be noted that if time series of multiple variables are available (e.g. relevant climate and hydrologic variables influencing streamflow dynamics, such as rainfall, temperature, and infiltration), then such can be directly used for reconstruction, which will be a more realistic representation of the system dynamics and, thus, will yield more reliable results.

A correct phase space reconstruction in a dimension m generally allows interpretation of the system dynamics (if the variable chosen to represent the system is appropriate) in the form of an m -dimensional map f_T , given by

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

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.

For an m -dimensional phase space, the correlation function $C(r)$ is given by

$$C(r) = \lim_{N \rightarrow \infty} \frac{2}{N(N-1)} \sum_{\substack{i,j \\ (i \leq i < j \leq N)}} H(r - \|Y_i - Y_j\|), \quad (3)$$

where H is the Heaviside step function, with $H(u) = 1$ for $u > 0$, and $H(u) = 0$ for $u \leq 0$, where $u = r - \|Y_i - Y_j\|$, r is the vector norm (radius of sphere) centered on Y_i or Y_j . If the time series is characterized by an attractor, then $C(r)$ and r are related according to

$$C(r) \approx \alpha r^\nu, \quad (4)$$

$r \rightarrow 0$
 $N \rightarrow \infty$

where α is a constant and ν is the correlation exponent or the slope of the $\text{Log } C(r)$ versus $\text{Log } r$ plot. The slope is generally estimated by a least square fit of a straight line over a certain range of r (i.e. scaling regime) or through estimation of local slopes between r -values.

The dimensionality of the time series is determined by checking if there is a saturation of ν with increasing m ; the saturation value of ν is defined as the correlation dimension (d) of the attractor. In general, a low saturation value of ν is considered as an indication of a low-dimensional system, while a high (or no) saturation value is considered as an indication of a high-dimensional system. The nearest integer above this saturation value is generally an indication

of the number of variables dominantly governing the system dynamics. Although the correlation dimension method is widely used for distinguishing low-dimensional systems and high-dimensional systems, additional categories of system dimensionality (e.g. medium) can also be formed based on correlation dimension values. This is attempted in the present study to achieve better grouping of streamflow time series.

It is relevant to note, at this point, that the reliability of the Grassberger–Procaccia algorithm (or any other algorithm for that matter) for correlation dimension estimation of real time series (e.g. streamflow observations) has been under considerable debate, in view of the potential limitations that may exist with the method and/or the data. Some of the relevant issues are data size (e.g. Havstad and Ehlers, 1989), data noise (e.g. Schreiber and Kantz, 1996), presence of zeros (e.g. Tsonis et al., 1994), temporal correlations and delay time selection for phase space reconstruction (e.g. Fraser and Swinney, 1986), even stochastic processes yielding low correlation dimensions (e.g. Osborne and Provenzale, 1989), and others. As most of these issues are also highly relevant to hydrologic time series, there have been criticisms on the correlation dimension estimates reported for hydrologic time series as well (e.g. Schertzer et al., 2002; Koutsoyiannis, 2006).

Numerous studies have addressed these issues and allayed the concerns on the reliability of correlation dimension estimates of hydrologic time series. Indeed, some studies have pointed out that many of the criticisms on dimension estimates are often unreliable and unfounded; see Sivakumar et al. (2002a) in response to Schertzer et al. (2002) regarding the issue of data size. These issues and concerns as well as clarifications and interpretations regarding correlation dimension estimates of hydrologic time series have already been extensively discussed in the literature (e.g. Sivakumar, 2000; Sivakumar et al., 2002b). Therefore, further details are not reported herein, and the interested reader is directed to such studies. However, as the issue of data size is particularly relevant for the 117 streamflow time series analyzed in this study (with “only” 624 values in each series), we will briefly discuss the reliability of our correlation dimension estimates in Sect. 5.3. We will also briefly explain our selection of the delay time for phase space reconstruction and its implications.

5 Data, analysis, and results

5.1 Data

In this study, monthly streamflows from the western United States (US) are studied, with data collected over an extensive network of 117 gaging stations (see Fig. 1). The stations are spread over 11 states in the western US: Arizona (AZ), California (CA), Colorado (CO), Idaho (ID), Montana (MT), Nevada (NV), New Mexico (NM), Oregon (OR), Utah (UT), Washington (WA), and Wyoming (WY). The

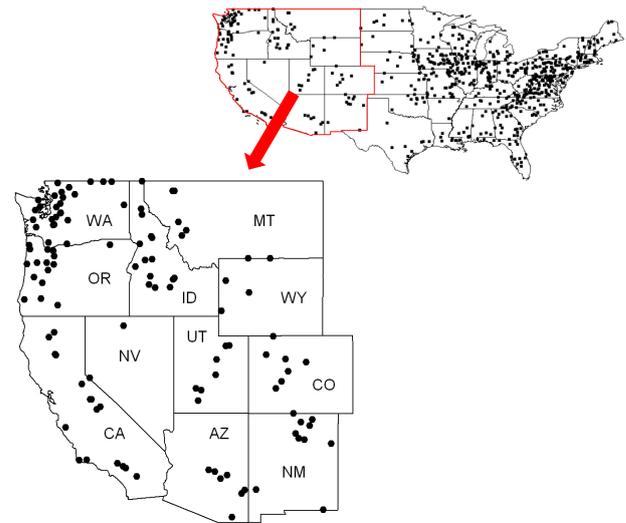


Fig. 1. Map of the western United States and locations of 117 streamflow gaging stations. AZ – Arizona; CA – California; CO – Colorado; ID – Idaho; MT – Montana; NM – New Mexico; NV – Nevada; OR – Oregon; UT – Utah; WA – Washington; WY – Wyoming.

drainage areas range from as small as 22.79 km² (8.8 mi²) (Station #11058500 in California) to as large as 35 094 km² (13 550 mi²) (Station #13317000 in Idaho); as many as two-thirds of the catchments are small- to medium-sized, i.e. less than 1000 km² (or approximately 400 mi²).

Streamflow data in the US are commonly expressed in “water years”, which commence in October. The records used in this study are those observed over a period of 52 yr, starting in October 1951 and ending in September 2003, and are average monthly streamflow values. The magnitude of streamflow varies greatly among the 117 stations (e.g. even during the same period) as well as within a station (e.g. at different periods). Notable observations of the flow variations (during the 52-yr period of 1951–2002) are as follows:

- the mean flows range from as low as 0.06 m³ s⁻¹ (1.97 ft³ s⁻¹) at Station #11063500 in CA to as high as 322 m³ s⁻¹ (11 550 ft³ s⁻¹) at Station #13317000 in ID;
- the standard deviation values range from as low as 0.11 m³ s⁻¹ (3.92 ft³ s⁻¹) at Station #11063500 to as high as 373.5 m³ s⁻¹ (13 193 ft³ s⁻¹) at Station #13317000;
- the coefficient of variation (CV) values (defined as the standard deviation divided by the mean) range from as low as 0.295 at Station #11367500 in CA to as high as 4.324 at Station #10258500 in CA;
- the maximum flow observed was 2339 m³ s⁻¹ (82 600 ft³ s⁻¹) at Station #13317000 (the minimum flow at this station was 64 m³ s⁻¹ (2257 ft³ s⁻¹)), while the flow was zero in 15 stations at one time or another;

- flows over $300 \text{ m}^3 \text{ s}^{-1}$ (or about $10\,000 \text{ ft}^3 \text{ s}^{-1}$) were observed at 18 stations and flows less than $0.3 \text{ m}^3 \text{ s}^{-1}$ (or about $10 \text{ ft}^3 \text{ s}^{-1}$) were observed at 38 stations (23 stations if those having zero flows are excluded); and
- five stations (Station #9448500, #9498500, #9508500, #12401500, and #14301000) had a maximum flow of over $300 \text{ m}^3 \text{ s}^{-1}$ (about $10\,000 \text{ ft}^3 \text{ s}^{-1}$) and also a minimum flow of less than $3 \text{ m}^3 \text{ s}^{-1}$ (about $100 \text{ ft}^3 \text{ s}^{-1}$).

All these observations clearly reflect the extreme variability in streamflow among the 117 stations. The variability in streamflow is due to, among others: (1) the different climatic regions in the western US; (2) the different drainage basin characteristics associated with the streamflow stations; and (3) the variations in hydroclimatic factors and land-use changes over a period of time at any of these stations.

Further details on these 117 streamflow stations in the western US (as well as the numerous other ones in the conterminous US), including streamflow data retrieval, are available at: <http://nwis.waterdata.usgs.gov/nwis>. The reader is also directed to Sivakumar (2003) and Tootle and Piechota (2006) for some of the studies relevant to streamflow at these stations.

5.2 Analysis and results

The correlation dimension analysis is performed on each of the above 117 streamflow time series. The phase space diagrams and the correlation exponent plots (i.e. local slope versus $\log r$) are carefully interpreted to achieve appropriate grouping of these time series.

Both phase space diagrams and correlation dimension plots show varying degrees of results among the 117 time series. The phase space diagrams exhibit attractors ranging from reasonably well-structured ones (i.e. in a well-defined region in the phase space) to totally “shapeless” ones (i.e. difficult to identify any kind of structure), and others in between these two extremes. Similarly, the correlation exponent plots show dimensionalities ranging from very low values of saturation of ν at one extreme (say less than 3) to unidentifiable ones at the other, and others in between.

Based on careful examination of phase space diagrams and correlation dimension results of all 117 streamflow series, we are able to identify four reasonably distinct groups. This identification is made based on the dimensionality of the attractor (d , i.e. saturation value of ν) as the primary criterion, since the dimensionality results allow a slightly better interpretation (qualitatively and quantitatively) compared to phase space diagrams. However, we also place particular emphasis on the consistency between dimensionality and attractor shape (phase space diagram) for each group, for a more reliable grouping. The four groups and the associated dimensionalities are as follows: (1) low-dimensional, with $d \leq 3.0$; (2) medium-dimensional, with $3.0 < d \leq 6.0$; (3) high-dimensional, with $d > 6.0$; and (4) unidentifiable.

The selection of the number of groups and the range of dimension values for each group is somewhat arbitrary. Nevertheless, they are certainly reasonable, especially in the context of the number of stations studied in the present study, since too many groups (with only minor differences among them) or just two groups (e.g. high-dimensional and low-dimensional) do not really serve the purpose of classification of 117 time series. Further, the above grouping according to correlation dimensions is also reasonable in the context of process/model complexity, since the influence of more than six dominant governing variables (i.e. $d > 6.0$) often leads to high complexity in dynamics (requiring “complex” models), whereas that of 3 or less variables can confidently be considered to lead to simpler dynamics (requiring “simple” models), with other in between (medium-complexity dynamics, requiring medium-complexity models).

For discussion here, we present the results for two time series from each of these four groups. The stations representing these time series are as follows: (1) low-dimensional – Station #10032000 (WY) and Station #13317000 (ID); (2) medium-dimensional – Station #11315000 (CA) and Station #11381500 (CA); (3) high-dimensional – Station #12093500 (WA) and Station #14185000 (OR); and (4) unidentifiable – Station #8408500 (NM) and Station #11124500 (CA).

Figure 2a–h presents the phase space diagrams for streamflow series from the above eight stations. The diagrams correspond to the reconstruction in two dimensions ($m = 2$) with delay time $\tau = 1$, i.e. the projection of the attractor on the plane $\{X_i, X_{i+1}\}$. The following general observations may be made: (1) the plots on the first row exhibit reasonably well-structured attractors in the phase space, suggesting that the systems are likely less complex and low-dimensional; (2) the second row plots indicate slightly wider scattering of the attractor, suggesting systems of medium complexity and medium dimension; (3) the plots on the third row exhibit much wider scattering (especially with one or a few outliers), suggesting highly complex and high-dimensional systems; and (4) the last two plots do not show any identifiable patterns, thus making it hard to include them in any of the above three groups.

Figure 3a–h presents the correlation dimension results for the above eight streamflow series; the plots show the local slopes (i.e. correlation exponent, ν) as a function of radius, r , for embedding dimensions, m , from 1 to 20 (bottom to top curves). These plots allow an even better interpretation in regards to the dimensionality and complexity of the underlying systems: (1) the top row plots reveal saturation of ν at a value less than 3 (shown using a thick horizontal line; see below for further details about identification of this saturation), suggesting low-dimensional and less complex systems; (2) the second row plots yield slightly higher dimensions (but less than 6), suggesting medium-dimensional and slightly more complex systems; (3) the plots on the third row do not indicate any saturation of ν , suggesting high-dimensional and

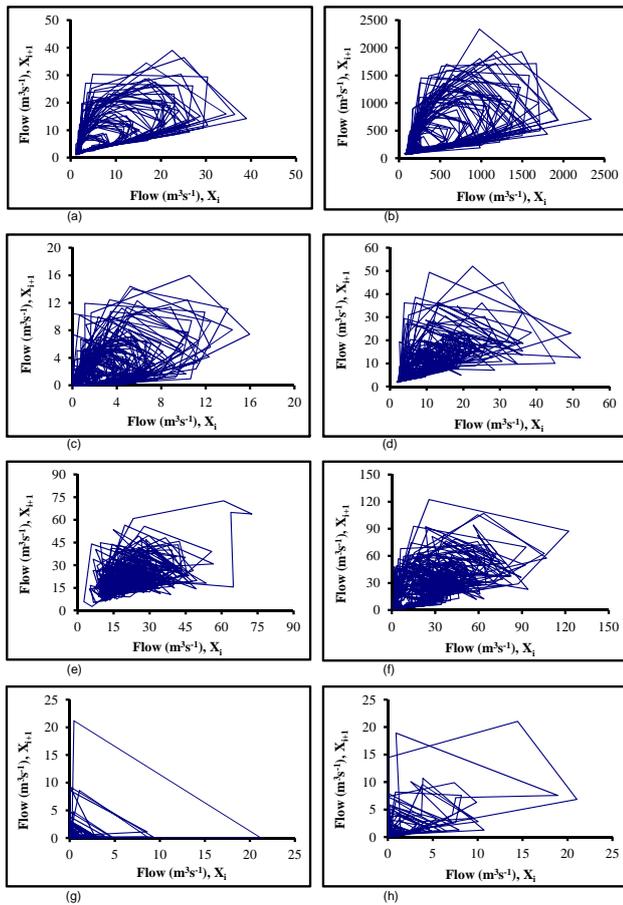


Fig. 2. Phase space diagram: (a) Station #10032000; (b) Station #13317000; (c) Station #11315000; (d) Station #11381500; (e) Station #12093500; (f) Station #14185000; (g) Station #8408500; and (h) Station #11124500.

highly complex systems; and (4) the results for the last two series do not show any clear indication regarding the dimension value or group (as they show neither saturation of ν nor high-dimensionality) and, therefore, are considered “unidentifiable”.

At this point, a few remarks about the identification of the scaling region and estimation of the correlation exponent are in order. As mentioned earlier, the scaling region can be identified in the following ways: (1) identifying the long “straight line” portion in the $\text{Log } C(r)$ versus $\text{Log } r$ plot (i.e. correlation function versus radius); and (2) the “horizontal line” in the local slope versus $\text{Log } r$ plot. It is important to note that a “perfect straight line” or a “perfect horizontal line” in these plots may be found when the data are completely clean, but is often very hard to find when the data are noisy, as is the case with streamflow (and other hydrologic) data; the higher the embedding dimension (or attractor dimension), the harder it is to find the scaling region. Also, when the data are noisy, the slopes are hard to find at small r -values, and there is normally a shift in the r -values that yield the best results;

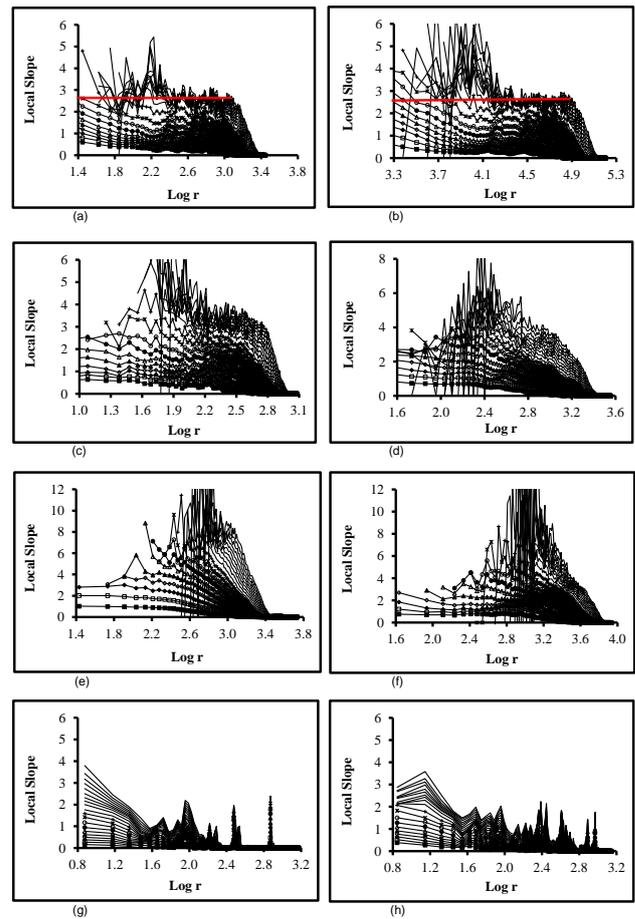


Fig. 3. Correlation dimension – Local slopes: (a) Station #10032000; (b) Station #13317000; (c) Station #11315000; (d) Station #11381500; (e) Station #12093500; (f) Station #14185000; (g) Station #8408500; and (h) Station #11124500.

again, the difficulty increases at higher embedding dimensions (and higher attractor dimensions). Therefore, it is often helpful, and necessary, to use as many ways as possible to be more confident of the scaling region identification and correlation exponent estimation. Further details on the effects of noise on the correlation dimension estimate, in particular reference to hydrologic data (rainfall), are presented in, for example, Sivakumar et al. (1999b), and the interested reader is directed to such.

In view of the above, we use not only the local slope versus $\text{Log } r$ for identification and estimation (shown in Fig. 3) but also the $\text{Log } C(r)$ versus $\text{Log } r$ plot (figures not shown) and the changes in the individual values of the calculated slopes against changing r . Since the (local) slopes may sometimes change dramatically between successive values of r , especially at small r -values (see Fig. 3), we also estimate the slopes averaged over a range of values of r (5 values at a time), in a moving average manner. The dimensions we arrive at are based on looking at all these combinations and making the best estimate.

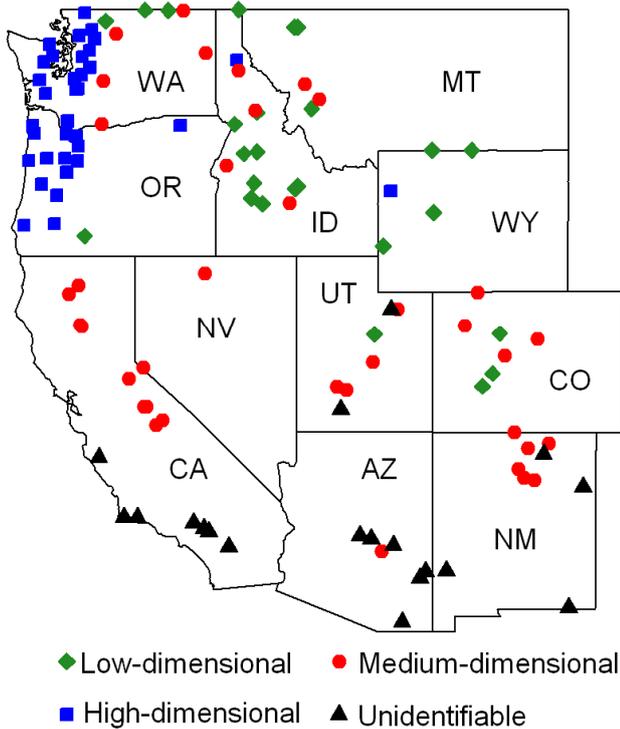


Fig. 4. Grouping of streamflow stations according to correlation dimension (d) estimates: low-dimensional ($d \leq 3.0$); medium-dimensional ($3.0 < d \leq 6.0$); high-dimensional ($d > 6.0$); and unidentifiable (d not identifiable).

Figure 4 presents the grouping of the 117 streamflow time series in the western US, according to the above dimensionality (and phase space) criterion. The grouping shows some kind of “homogeneity” in the dimensionality and complexity of streamflow dynamics within certain regions. For instance: (1) streamflow dynamics in the far northwest (i.e. western parts of WA and OR) are generally high-dimensional; (2) the dimensionality of streamflows in the far south and southwest (southern CA, southern AZ, southern NM) is generally unidentifiable; (3) the complexity of streamflow dynamics in the west (northern CA and NV) is generally medium-dimensional; and (4) low-dimensional complexity is generally observed for streamflows in Wyoming. However, this “homogeneity” is not true for every region, and there are indeed strong exceptions. For example: (1) both low-dimensional and medium-dimensional complexity of streamflow dynamics are observed in some other regions, especially in the east and north (including CO, ID, MT, and some parts of WA); and (2) streamflow dynamic complexity in some regions is rather very mixed, ranging from low-dimensional to medium-dimensional to unidentifiable (UT and, to some extent, northern NM).

5.3 Discussion

The above classification of streamflow based on complexity and nonlinear dynamic concepts, with dimensionality (and other relevant properties) as a criterion, is both useful and interesting. In particular, the dimension estimates and the grouping of streamflow time series (Fig. 4) clearly show that: (1) the dimensionality concept captures the complexity of streamflow dynamics at individual stations independently and then allows classification regardless of the proximity of catchments, without resorting to a “regionalization” approach and the assumptions involved therein; and (2) a “regionalization” approach, even for monthly streamflows, is not necessarily the right way to classification, despite the close proximity of some catchments. In other words, the dimension estimates reflect that “near” does not mean “similar” and, consequently, that extrapolation (and interpolation) may not always work even when using data from nearby catchments. This observation has important implications for predictions in ungaged basins (PUBs), especially when they involve extrapolation/interpolation schemes.

Notwithstanding that the dimensionality concept and the proposed classification are useful, it is still somewhat premature to offer definitive conclusions and guidelines. Some reasons for this and also possible ways to address them are as follows. We are currently studying these issues, and will report the details in the future.

- Despite the consideration of a study area as large as the western United States and streamflow time series from as many as 117 stations, the extent of area covered and number of time series analyzed are still considerably smaller when compared to the numerous combinations that may be encountered with respect to catchments (e.g. climatic conditions, catchment properties, streamflow characteristics). Therefore, it is important to study a significantly large number of catchments and streamflow time series. In the specific context of the western United States, it would be important to study many more catchments, especially in the following parts: western and southern Arizona, western California, eastern Colorado, eastern and southern Idaho, almost entire Montana, almost entire Nevada, western and southern New Mexico, eastern Oregon, northwest and southeast Utah, eastern Washington, and eastern Wyoming.
- In the present study, only monthly streamflow time series are analyzed. Since the dimensionality and complexity of streamflow (and other hydrologic) processes could change with respect to temporal scale (e.g. Regonda et al., 2004; see also Sivakumar et al., 2001), it is crucial to study streamflow data observed at least at a few other scales (e.g. daily, annual) to verify the dimension estimates and classification. However, as Sivakumar (2008) suggests, and as mentioned earlier, “scale” is a vital component in the definition of a “system”.

Other vital components are “process” and “purpose of interest”. For instance, one often requires different models for average events and extreme events (e.g. droughts and floods); see Sivakumar (2005b) for a discussion on this, especially on the role of thresholds. In most cases, study of monthly streamflow dynamics is more appropriate for medium-term to long-term water planning and management (including environmental flow requirements), rather than flood forecasting, which requires data at daily and even much finer timescales. Therefore, a classification framework may (or may not) be limited by how a system is defined.

- The correlation dimension method is only one among a number of nonlinear dynamic-based methods available for estimating dimensionality and assessing complexity of systems, despite the fact that it has been the most widely used. Two other methods are the false nearest neighbor algorithm (e.g. Kennel et al., 1992) and the Kolmogorov entropy method (e.g. Grassberger and Procaccia, 1983b). Therefore, it would be particularly useful to employ these methods to verify, and possibly confirm, the correlation dimension estimates. As linear approaches and nonlinear approaches often complement each other, and the fact that streamflow (and other hydrologic) processes often exhibit both linear and nonlinear properties (depending upon catchments, scales, etc.), it would also be helpful to apply linear techniques to study the complexity and perhaps find better ways to classify the streamflow time series. In this regard, coupling/integration of nonlinear and linear techniques may also be possible.

At this point, it is also important to discuss the reliability of the correlation dimension estimates obtained for the 117 streamflow time series analyzed in this study. As mentioned earlier, there have been criticisms on the dimension estimates reported for hydrologic time series, especially in light of the potential limitations that may exist with the method/data (e.g. data size, data noise, presence of zeros, temporal correlation). Here, we address two issues that are particularly relevant to the streamflow time series analyzed and methodology used in this study: data size (“only” 624 values) and temporal correlation (delay time $\tau = 1$ for phase space reconstruction).

One of the most common criticisms on the use of correlation dimension method (especially the Grassberger–Procaccia algorithm) for hydrologic (and other real) time series is that it significantly underestimates the dimension when the data size is small (e.g. Nerenberg and Essex, 1990; Schertzer et al., 2002). Many studies have already addressed this issue through various means (e.g. Lorenz, 1991; Sivakumar et al., 2002a). These studies essentially point out that: (1) the data size is not a function of embedding (or attractor) dimension; and (2) it is not appropriate to simply look at the data length alone (in terms of the sheer number of values) and that it is far more important to assess if the time

series is long and representative enough (in terms of period of coverage and sampling time) to capture the essential dynamics of the system evolution. For instance, studies have shown that even a few hundred data (about 300 or so) would be sufficient for dimension estimate (e.g. Sivakumar, 2005a) if the period of coverage is long enough for the sampling time studied (e.g. Sivakumar et al., 2002b). The dimension estimates obtained for the 117 streamflow time series in the present study only offer further support to this. With “only” 624 values in each streamflow series, the correlation dimension method still yields dimension values ranging from very low to very high (including non-saturation of ν), clearly reflecting the variability of the data and complexity of the underlying dynamics and also defying the widely-perceived relationship between data size and embedding dimension. The primary reason for this is that the streamflow data studied are long enough (52 yr at monthly scale) to adequately represent the dynamic changes that occur in the respective catchments.

There are questions regarding the selection of an appropriate delay time (τ) for phase space reconstruction and correlation dimension estimation, since a small τ may result in temporal correlations between the values in the reconstructed vector while a large τ may result in completely independent ones. Various methods/guidelines have been proposed for τ selection to have the best separation of neighboring trajectories, including autocorrelation function (e.g. Holzfuss and Mayer-Kress, 1986), mutual information (e.g. Fraser and Swinney, 1986), and correlation integral (Liebert and Schuster, 1989). Regardless of the method used and the value of τ obtained, it is also not clear how such a value is actually relevant to the dynamics that take place in the underlying system; see Sivakumar et al. (2007) for relevant issues in using the autocorrelation function method for τ selection, even for the case of a well-known artificial low-dimensional chaotic system, the Henon map (Henon, 1963). For instance, use of the autocorrelation function method and selection of the lag time at which the autocorrelation function first crosses the zero line yield τ -values ranging from 2 to 40 among all the 117 streamflow series. These τ -values do not seem to indicate any consistent relevance to seasonality or other catchment dynamics; for instance, $\tau = 40$ indicates a delay time of over three years ($\tau = 3$ is obtained in a few cases). Similar problems have also been encountered in dealing with other hydrologic data, whether at the monthly scale or at other temporal scales; see Sivakumar et al. (2006) for some details in regards to rainfall data from California. Considering these issues and associated complications, in the present study, τ is chosen equal to the sampling (i.e. $\tau = 1$ month) for phase space reconstruction of each of the 117 streamflow time series. We believe such a selection is reasonable, especially since there are no significant correlations at lag time equal to 1; in most cases, the correlation at lag time one is about 0.4–0.6, which is relatively small for streamflow (it is likely that significant correlations occur for daily flows, especially for large catchments). Extensive details on the issue of

τ selection are already available in the hydrologic literature (e.g. Sangoyomi et al., 1996; Sivakumar et al., 1999a) and, therefore, are not reported herein.

6 Conclusions and further research

Hydrologic models play a crucial role in the assessment of water resources availability and decisions on water planning and management. Consequently, hydrologic modeling has become an important research endeavor, particularly facilitated by recent technological and methodological advances. Although numerous hydrologic models have been developed (often with increasing structural complexity and mathematical sophistication), identifying which model is appropriate for which catchment remains a fundamental problem. To this end, the need for a classification framework that streamlines catchments into different groups and sub-groups for a more effective and efficient model selection is increasingly realized. However, an appropriate basis and a suitable methodology for such a framework are still elusive.

This study offers one possible way to view the classification problem in hydrology through an inverse approach; i.e., going backward from system outputs. It argues that hydrologic system complexity forms an appropriate basis for the classification framework and nonlinear dynamic concepts constitute a suitable methodology for assessing system complexity. Discussing the relevance of complexity and nonlinearity in hydrologic systems and also the utility of nonlinear dynamic tools for complexity determination and system identification, the study employs a nonlinear dynamic method for classification of streamflow in the western United States. Applying the correlation dimension method (a dimensionality-based method having its basis in data reconstruction and nearest neighbor concepts) to monthly streamflow time series from 117 stations in the western US, the study classifies these time series into four distinct groups: low-dimensional, medium-dimensional, high-dimensional, and unidentifiable. The dimension estimates for the 117 streamflow time series show some “homogeneity” in the complexity of streamflow dynamics within certain regions of the western US. However, there are also strong exceptions to this within some other regions. These results not only indicate the utility of the dimensionality concept for classification but also suggest that a “regionalization” approach may not always be the right way to classification. As “regionalization” is arguably one of the most important aspects of extrapolation/interpolation of hydrologic data and, hence, for predictions in ungauged basins (PUBs), the present results have important implications to advance our studies on PUBs.

Since dimensionality of a time series is a representation of the level of complexity of the underlying system dynamics (and number of dominant governing variables), the above nonlinear dynamic- and dimensionality-based classification

certainly helps in identifying the appropriate structure and complexity of models. It is important to further verify, and confirm, the present results through other methods (both nonlinear and linear) that can be supplementary and complementary. Verification also needs to be done through: (a) establishing relationships between the data patterns/complexity and the actual catchment/process properties; and (b) studying the outputs simulated from existing hydrologic models and varying their complexities. The effectiveness of any such classification also needs to be tested on a wide variety of catchments and hydrologic data representing different climatic conditions, catchment characteristics, land use properties, and types of data, among others. Detailed studies in these directions are underway, and the results will be reported in future publications.

Finally, it is important to remember that classification of catchments is not the “be-all and end-all” of research on catchments, but rather only a means towards achieving broader goals of planning and management of our water resources, environment, ecosystems, and other relevant earth systems and resources. Nevertheless, catchment classification certainly allows us to study catchments more effectively and efficiently and develop more appropriate strategies, in terms of simplification in models/model development, generalization in our modeling approach, and improvement in communication both within the hydrologic community and across disciplines, as much as possible. Needless to say, catchment classification needs to be tuned towards the broader goals, which are carefully identified and properly defined, in order for us to assess whether catchment classification is necessary and to evaluate whether a proposed classification framework is successful. The present study has highlighted some of the issues associated with these, including the need to define a “system” with the necessary angles to view it from (e.g. process, scale, purpose). The study of monthly streamflow dynamics in the present study is tuned towards identification of models for medium-term to long-term water planning and management (including environmental flow requirements), rather than flood forecasting, which requires data at daily and even much finer timescales. Although an accurate assessment of the classification proposed in this study still requires some good distance to travel, the dimensionality concept certainly has potential, including in identifying where a “regionalization” approach is more effective, where it is not, and where and why the transitions occur. We hope that future studies will further help realize the true potential of the correlation dimension concept, and other nonlinear dynamic concepts, for formulation of a catchment classification framework.

Acknowledgements. Support for this work was provided by the Australian Research Council (ARC). Bellie Sivakumar acknowledges the financial support from ARC through the Future Fellowship grant (FT110100328). We thank B. Selle and two anonymous referees for their constructive comments and useful

suggestions, and M. Sivapalan for his Short Comment on an earlier version of this manuscript. We also thank Attilio Castellarin for inviting us to contribute to the Special Issue on “Catchment Classification and PUB” and for his constructive suggestions and recommendation on our manuscript.

Edited by: A. Castellarin

References

- Ali, G., Tetzlaff, D., Soulsby, C., McDonnell, J. J., and Capell, R.: A comparison of similarity indices for catchment classification using a cross-regional dataset, *Adv. Water Resour.*, 40, 11–22, 2012.
- Beckinsale, R. P.: River regimes, edited by: Chorley, R. J., Water, Earth, and Man. Methuen, London, 455–471, 1969.
- Beven, K. J.: Uncertainty and the detection of structural change in models of environmental systems, edited by: Beck, M. B., Environmental Foresight and Models: A Manifesto, Elsevier, The Netherlands, 227–250, 2002.
- Chapman, T.: Classification of regions, edited by: Falkenmark, M. and Chapman, T., Comparative Hydrology: An Ecological Approach to Land and Water Resources, Paris, UNESCO, 67–74, 1989.
- Chow, V. T.: Handbook of Applied Hydrology, McGraw-Hill, New York, USA, 1964.
- Dooge, J. C. I.: The hydrologic cycle as a closed system, *Int. Assoc. Sci. Hydrol. Bull.*, 13, 58–68, 1967a.
- Dooge, J. C. I.: A new approach to nonlinear problems in surface water hydrology: hydrologic systems with uniform nonlinearity, *Int. Assoc. Sci. Hydrol. Publ.*, 76, 409–413, 1967b.
- Fraser, A. M. and Swinney, H. L.: Independent coordinates for strange attractors from mutual information, *Phys. Rev. A*, 33, 1134–1140, 1986.
- Govindaraju, R. S.: Artificial neural networks in hydrology. II: Hydrological applications, *ASCE J. Hydrol. Eng.*, 5, 124–137, 2000.
- Grassberger, P. and Procaccia, I.: Measuring the strangeness of strange attractors, *Physica D*, 9, 189–208, 1983a.
- Grassberger, P. and Procaccia, I.: Estimation of the Kolmogorov entropy from a chaotic signal, *Phys. Rev. A*, 28, 2591–2593, 1983b.
- Grayson, R. B. and Blöschl, G.: Spatial Patterns in Catchment Hydrology: Observations and Modeling. Cambridge University Press, Cambridge, UK, 2000.
- Haines, A. T., Finlayson, B. L., and McMahon, T. A.: A global classification of river regimes, *Appl. Geogr.*, 8, 255–272, 1988.
- Harris, N. M., Gurnell, A. M., Hannah, D. M., and Petts, G. E.: Classification of river regimes: a context for hydroecology, *Hydrol. Process.*, 14, 2831–2848, 2000.
- Havstad, J. W. and Ehlers, C. L.: Attractor dimension of nonstationary dynamical systems from small data sets, *Phys. Rev. A*, 39, 845–853, 1989.
- Henon, M.: A two-dimensional mapping with a strange attractor, *Commun. Math. Phys.*, 50, 69–77, 1963.
- Holzfuß, J. and Mayer-Kress, G.: An approach to error-estimation in the application of dimension algorithms, in: Dimensions and Entropies in Chaotic Systems, edited by: Mayer-Kress, G., Springer, New York, 114–122, 1986.
- Isik, S. and Singh, V. P.: Hydrologic regionalization of watersheds in Turkey, *ASCE J. Hydrol. Eng.*, 13, 824–834, 2008.
- Izzard, C. F.: A mathematical model for nonlinear hydrologic systems, *J. Geophys. Res.*, 71, 4811–4824, 1966.
- Kavvas, M. L.: Nonlinear hydrologic processes: conservation equations for determining their means and probability distributions, *ASCE J. Hydrol. Eng.*, 8, 44–53, 2003.
- Kennel, M. B., Brown, R., and Abarbanel, H. D. I.: Determining embedding dimension for phase space reconstruction using a geometric method, *Phys. Rev. A*, 45, 3403–3411, 1992.
- Koutsoyiannis, D.: On the quest for chaotic attractors in hydrological processes, *Hydrolog. Sci. J.*, 51, 1065–1091, 2006.
- Krasovskaia, I.: Quantification of the stability of river flow regimes, *Hydrolog. Sci. J.*, 40, 587–598, 1995.
- Krasovskaia, I.: Entropy-based grouping of river flow regimes, *J. Hydrol.*, 202, 173–191, 1997.
- Liebert, W. and Schuster, H. G.: Proper choice of the time delay for the analysis of chaotic time series, *Phys. Lett. A*, 141, 386–390, 1989.
- Lorenz, E. N.: Deterministic nonperiodic flow, *J. Atmos. Sci.*, 20, 130–141, 1963.
- Lorenz, E. N.: Dimension of weather and climate attractors, *Nature*, 353, 241–244, 1991.
- McDonnell, J. J. and Woods, R. A.: On the need for catchment classification, *J. Hydrol.*, 299, 2–3, 2004.
- Merz, B. and Blöschl, G.: Regionalization of catchment model parameters, *J. Hydrol.*, 287, 95–123, 2004.
- Nerenberg, M. A. H. and Essex, C.: Correlation dimension and systematic geometric effects, *Phys. Rev. A*, 42, 7065–7074, 1990.
- Olden, J. D. and Poff, N. L.: Redundancy and the choice of hydrologic indices for characterizing streamflow regimes, *River Res. Appl.*, 19, 101–121, 2003.
- Olden, J. D., Kennard, M. J., and Pusey, B. J.: A framework for hydrologic classification with a review of methodologies and applications in ecohydrology, *Ecohydrology*, 5, 503–518, doi:10.1002/eco.251, 2011.
- Osborne, A. R. and Provenzale, A.: Finite correlation dimension for stochastic systems with power-law spectra, *Physica D*, 35, 357–381, 1989.
- Packard, N. H., Crutchfield, J. P., Farmer, J. D., and Shaw, R. S.: Geometry from a time series, *Phys. Rev. Lett.*, 45, 712–716, 1980.
- Paola, C., Fofoula-Georgiou, E., Dietrich, W. E., Hondzo, M., Mohrig, D., Parker, G., Power, M. E., Rodriguez-Iturbe, I., Voller, V., and Wilcock, P.: Toward a unified science of the Earth’s surface: opportunities for synthesis among hydrology, geomorphology, geochemistry, and ecology, *Water Resour. Res.*, 42, W03S10, doi:10.1029/2005WR004336, 2006.
- Poff, N. L., Olden, J. D., Pepin, D. M., and Bledsoe, B. P.: Placing global stream flow variability in geographic and geomorphic contexts, *River Res. Appl.*, 22, 149–166, 2006.
- Regonda, S., Sivakumar, B., and Jain, A.: Temporal scaling in river flow: can it be chaotic?, *Hydrolog. Sci. J.*, 49, 373–385, 2004.
- Rosgen, D. L.: A classification of natural rivers, *Catena*, 22, 169–199, 1994.
- Sangoyomi, T. B., Lall, U., and Abarbanel, H. D. I.: Nonlinear dynamics of the Great Salt Lake: dimension estimation, *Water Resour. Res.*, 32, 149–159, 1996.
- Schertzer, D., Tchiguirinskaia, I., Lovejoy, S., Hubert, P., and Bendjoudi, H.: Which chaos in the rainfall-runoff process? A discus-

- sion on "Evidence of chaos in the rainfall-runoff process" by Sivakumar et al., *Hydrolog. Sci. J.*, 47, 139–147, 2002.
- Schreiber, T. and Kantz, H.: Observing and predicting chaotic signals: is 2% noise too much?, in: *Predictability of Complex Dynamical Systems*, edited by: Kravtsov, Yu. A. and Kadtko, J. B., Springer Series in Synergetics, Springer, Berlin, 43–65, 1996.
- Singh, V. P.: *Hydrologic Systems: Volume 1, Rainfall-Runoff Modeling*, Prentice Hall, New Jersey, USA, 1988.
- Sivakumar, B.: Chaos theory in hydrology: important issues and interpretations, *J. Hydrol.*, 227, 1–20, 2000.
- Sivakumar, B.: Forecasting monthly streamflow dynamics in the western United States: a nonlinear dynamical approach, *Environ. Modell. Softw.*, 18, 721–728, 2003.
- Sivakumar, B.: Dominant processes concept in hydrology: moving forward, *Hydrol. Process.*, 18, 2349–2353, 2004a.
- Sivakumar, B.: Chaos theory in geophysics: past, present and future, *Chaos Soliton. Fract.*, 19, 441–462, 2004b.
- Sivakumar, B.: Correlation dimension estimation of hydrologic series and data size requirement: myth and reality, *Hydrolog. Sci. J.*, 50, 591–604, 2005a.
- Sivakumar, B.: Hydrologic modeling and forecasting: role of thresholds, *Environ. Modell. Softw.*, 20, 515–519, 2005b.
- Sivakumar, B.: Dominant processes concept, model simplification and classification framework in catchment hydrology, *Stoch. Env. Res. Risk A.*, 22, 737–748, 2008.
- Sivakumar, B., Liang, S. Y., Liaw, C. Y., and Phoon, K. K.: Singapore rainfall behavior: chaotic?, *ASCE J. Hydrol. Eng.* 4, 38–48, 1999a.
- Sivakumar, B., Phoon, K. K., Liang, S. Y., and Liaw, C. Y.: A systematic approach to noise reduction in chaotic hydrological time series, *J. Hydrol.*, 219, 103–135, 1999b.
- Sivakumar, B., Sorooshian, S., Gupta, H. V., and Gao, X.: A chaotic approach to rainfall disaggregation, *Water Resour. Res.*, 37, 61–72, 2001.
- Sivakumar, B., Berndtsson, R., Olsson, J., and Jinno, K.: Reply to "Which chaos in the rainfall-runoff process?" by Schertzer et al., *Hydrolog. Sci. J.*, 47, 149–158, 2002a.
- Sivakumar, B., Persson, M., Berndtsson, R., and Uvo, C. B.: Is correlation dimension a reliable indicator of low-dimensional chaos in short hydrological time series?, *Water Resour. Res.*, 38, 1011, doi:10.1029/2001WR000333, 2002b.
- Sivakumar, B., Wallender, W. W., Horwath, W. R., Mitchell, J. P., Prentice, S. E., and Joyce, B. A.: Nonlinear analysis of rainfall dynamics in California's Sacramento Valley, *Hydrol. Process.*, 20, 1723–1736, 2006.
- Sivakumar, B., Jayawardena, A. W., and Li, W. K.: Hydrologic complexity and classification: a simple data reconstruction approach, *Hydrol. Process.*, 21, 2713–2728, 2007.
- Snelder, T. H., Biggs, B. J. F., and Woods, R. A.: Improved eco-hydrological classification of rivers, *River Res. Appl.*, 21, 609–628, 2005.
- Takens, F.: Detecting strange attractors in turbulence, in: *Dynamical Systems and Turbulence*, edited by: Rand, D. A. and Young, L. S., Lecture Notes in Mathematics 898, Springer-Verlag, 366–381, 1981.
- Tootle, G. A. and Piechota, T. C.: Relationships between Pacific and Atlantic ocean sea surface temperatures and U.S. streamflow variability, *Water Resour. Res.*, 42, W07411, doi:10.1029/2005WR004184, 2006.
- Tsonis, A. A., Triantafyllou, G. N., Elsner, J. B., Holdzkom II, J. J., and Kirwan Jr., A. D.: An investigation on the ability of nonlinear methods to infer dynamics from observables, *B. Am. Meteorol. Soc.*, 75, 1623–1633, 1994.
- Vormoor, K., Skaugen, T., Langsholt, E., Dieckrüger, B., and Skjøien, J. O.: Geostatistical regionalization of daily runoff forecasts in Norway, *Int. J. River Basin Management*, 9, 3–15, 2011.
- Wagener, T., Sivapalan, M., Troch, P., and Woods, R. A.: Catchment classification and hydrologic similarity, *Geog. Compass*, 1, 901–931, 2007.
- Wardrop, D. H., Bishop, J. A., Easterling, M., Hychka, K., Myers, W. L., Patil, G. P., and Taille, C.: Use of landscape and land use parameters for classification of watersheds in the mid-Atlantic across five physiographic provinces, *Environ. Ecol. Stat.*, 12, 209–223, 2005.
- Woods, R. A.: Seeing catchments with new eyes, *Hydrol. Process.*, 16, 1111–1113, 2002.
- Young, P. C. and Beven, K. J.: Data-based mechanistic modeling and rainfall-flow non-linearity, *Environmetrics*, 5, 335–363, 1994.