Detecting determinism and nonlinearity in river-flow time series

AMILCARE PORPORATO & LUCA RIDOLFI
Department of Civil and Environmental Engineering, Duke University, 123 Hudson Hall, Durham, North Carolina 27708-0287, USA
porporato@polito.it

Abstract The temporal evolution of river-flow dynamics is investigated using methods from nonlinear time series analysis. After a discussion on the possible sources of nonlinear deterministic components, different methods for the detection of nonlinear and deterministic components in river flow time series are illustrated and the application to three river-flow time series from basins of different dimensions and characteristics is presented. All the results of different nonlinear techniques show noticeable evidence of determinism and nonlinearity at high discharge values, while at low discharge values, and especially during recessions, the series seem to be the result of a mostly linear dynamics. The complexity of the dynamics and the degree of nonlinearity are also discussed in relation to the basin dimensions and the kind of hydroclimatic regime.

Key words rainfall–runoff transformation; river-flow forecasting; nonlinear time series analysis; recession curves; entropy

INTRODUCTION

The physical mechanisms governing river flow dynamics are many, and act on a huge range of temporal and spatial scales. Although not all components are complex in themselves, the vastness of the space–time domain, the number of the processes involved, and the fact that almost all of them present some degree of nonlinearity, make the problem of river-flow formation highly non-trivial. To model such a complex system, one hopes that only few of the various mechanisms become prevalent in different phases of the process, so that the system dynamics undergoes a simplification due to a reduction in the number of the effective degrees of freedom.
Researchers have frequently employed typical methods for the analysis of complex systems to study hydrological phenomena, aiming at separating the low-dimensional deterministic components from the high-dimensional ones. The importance of this research stems from the fact that it involves the issue of proper modelling of hydrological systems and the prediction of their behaviour for practical purposes. As effectively pointed out by Amorocho (1967, pp. 862, 868): “For while it is obviously essential to strive towards a separation of the manifestly deterministic elements of a system, it is futile to attempt to step beyond the limits of the physically possible by carrying the banner of strict determinism against all odds. [...] The nonlinearity of hydrologic systems has been recognized for many years. It is of interest, however, to attempt qualitative determination of the degree of linearity of particular catchments, because they permit an assessment of the range of approximation likely to be obtained from the application of either linear or low-order nonlinear methods for the postulation of inflow–outflow relationships”. More than 25 years later, the words by Pilgrim & Cordery (1993) still urge researchers to investigate the issue of nonlinearity in rainfall–runoff transformation: “The form and degree of nonlinearity, [...] which are often selected arbitrarily in practice, [...] can have a major effect on design estimates. Surprisingly little attention has been paid to the process involved in nonlinearity of response of drainage basins.” A reliable detection of nonlinearity can be an essential information in addressing modelling efforts (Judd & Mees, 1995).

Among the studies devoted to assessing the degree of nonlinearity and overall complexity of the system governing river flow, no general conclusion has been reached up to now. Some of the partial explanations which have been given are in disagreement or even in contradiction with each other and this explains why obtaining an objective indication from real data is extremely precious and has become an important goal in hydrology as well as in other scientific fields.

In this paper, some modern techniques of nonlinear time series analysis are employed (for an outline see Abarbanel et al., 1993; Kantz & Schreiber, 1997) with the aim of detecting and locating determinism and nonlinearity in the system governing the time behaviour of river flow. The present investigation carries on previous works by the authors (Porporato & Ridolfi, 1996, 1997, 2001), where some evidence of the presence of deterministic components and nonlinearity in river flow was reported. The analysis presented here is based on the univariate analysis of three river-flow time series of different characteristics employing a variety of methods. Other papers have also recently contributed to nonlinear analysis of hydrological time series (see e.g. the review by Sivakumar, 2000); those more akin to the present line of analysis are by Jayawardena & Lai (1994), Lall et al. (1996), Sangoyomi et al. (1996), Sivakumar et al. (1999, 2001a,b, 2002), Krasovskaia et al. (1999), Jayawardena & Gurung (2000), Elshorbagy et al. (2001), Islam & Sivakumar (2002), Phoon et al. (2002), Schertzer et al. (2002), Sivakumar (2002), Sivakumar & Jayawardena (2002) and Zhou et al. (2002).

ORIGIN OF DETERMINISM AND NONLINEARITY IN RIVER FLOW

It is useful to begin this analysis by qualitatively discussing why and where deterministic components or nonlinear dynamics are to be expected in river-flow time series.
The first source of determinism is related to climate dynamics that produces the input of the rainfall–runoff transformation and also determines in many ways the state of the basin with a sort of “parametric forcing” on vegetation cover, soil saturation, etc. Many studies have suggested the possible presence of low-dimensional deterministic components in the dynamics of the global climate and its external forcing (e.g. Matsumoto et al., 1995 and references therein). Others have conjectured that low-dimensional chaotic dynamics can originate from the nonlinear interaction of climate and large water bodies (e.g. Tsonis & Elsner, 1990; Tziperman et al., 1997). Recent studies have reported clues of low-dimensional dynamics in the response of large basins, where connections with the local climate are likely to be present (Lall et al., 1996; Sangoyomi et al., 1996). Moreover, for basins having hydrological regimes influenced also by snowmelt or by perennial glaciers, additional delays and feedbacks with the climate dynamics are certainly present. Even if the actual dynamics of atmosphere at a meteorological time scale is more likely of the kind of spatio-temporal chaos of a high dimension, the coupling with the basin with its various feedbacks could also give rise to recurrent low-dimensional components.

The strong low-pass filtering action of the basin, while smoothing out some of the space–time complexity of rainfall, could also make more evident the low-dimensional deterministic components originating from both climate and rainfall–runoff transformation. Low-dimensional components are introduced by the very action of the basin (e.g. Richards-Pecou, 2002): besides the influence of climatic and meteorological dynamics, discharge time series bear the fingerprint of the basin characteristics (topography, geology, channel network geometry, vegetation, human actions, etc.). For example, groundwater systems behave as a nonlinear dissipative system with few degrees of freedom (e.g. Duffy, 1996), and recession curves usually show a quite simple behaviour (e.g. Tallaksen, 1995; Wittenberg, 1999; Elshorbagy et al., 2001). Chiu & Huang (1970) successfully proposed a low-dimensional (second-order) nonlinear ordinary differential equation (ODE) for the falling limbs of the hydrograph, which closely reproduced measured data. This was evident also in the Poincaré sections presented in Porporato & Ridolfi (1997): this last aspect will be addressed later in the paper.

Regarding nonlinearity, there is no doubt that climatic and atmospheric dynamics are strongly nonlinear (e.g. Saltzman, 1983; Houghton, 1991; Lorenz, 1991), as nonlinearity certainly is nature of the majority of the feedbacks in the soil–atmosphere interaction (Brubaker & Entekhabi, 1996), rainfall formation, and rainfall–runoff transformation. The sources of nonlinearity in the rainfall–runoff transformation are many and they combine differently in basins of different sizes and hydrological regimes; examples include: the rainfall interception due to vegetation (a typical season-dependent, threshold-like mechanism), the unsaturated groundwater flow (where the hydraulic conductivity depends in a strongly nonlinear way on the degree of soil saturation), and open channel hydraulics. Hillslope-related nonlinearities seem to be predominant in small catchments where overland flow and unsaturated subsurface flow are predominant, whilst nonlinearity of channel-network response becomes predominant as the basin area increases (Robinson et al., 1995). Since the pioneering works by Amorocho (1963) and Jacobi (1966), many studies have increasingly given support to the better performance of nonlinear models in comparison to linear ones for the rainfall–runoff transformation (Amorocho & Brandstetter, 1971; Rao & Rao, 1984; Hsu et al., 1995).
Granted the presence of nonlinearity, its detection in data from natural systems is not a simple task. The interaction of the many degrees of freedom and the presence of external noise can mask its presence. This also explains the existence of somewhat contradictory results, or at least those of validity limited to the particular cases: some studies have found a higher degree of nonlinearity in small basins (Pilgrim, 1976; Wang et al., 1981), while others have drawn opposite conclusions (Robinson et al., 1995; Goodrich et al., 1997). Liu & Brutsaert (1978) detected a higher degree of nonlinearity during periods of high discharge, whereas Pilgrim (1976) found an opposite behaviour.

TESTS FOR DETERMINISM AND NONLINEARITY

The detection of nonlinearity and determinism in time series is one of the most intense research fields in nonlinear time series analysis. Due to the difficulty in obtaining reliable and clear-cut results from time series of natural phenomena, different methods have to be used in combination to assess the coherence of the results obtained (Abarbanel et al., 1993; Kantz & Schreiber, 1997). Different methods are discussed in the following with special attention to the application to river flow time series.

State-space reconstruction of river flow time series

The univariate approach to nonlinear time series analysis of river flow can be justified by the methods of state-space reconstruction. These are the basis of the modern nonlinear time series analysis (e.g. Abarbanel et al., 1993), as well as the starting point of all the methods that will be employed. According to Takens’ theorem (Takens, 1981; see also Sauer et al., 1991), as each variable contains information on the overall system dynamics, it is possible to reconstruct a space (the so-called embedding space) equivalent to the original phase space of the system from a single scalar time series. Referring to discharge time series, \( \{q_i\}_{i=1,...,N} \), the sequence of \( M = N - (m-1)\tau \) \( m \)-dimensional vectors of the form:

\[
Q_i = \{q_i, q_{i-\tau}, q_{i-2\tau}, ..., q_{i-(m-1)\tau}\}
\]

(1)

describes an object topologically equivalent to the attractor of the system. The dimension \( m \) of the space is called the embedding dimension and the parameter \( \tau \) is a suitable delay time. In the embedding space, the dynamics of the system can thus be described in the form of a discrete map:

\[
Q_{i+1} = f(Q_i)
\]

(2)

or, in the limit of a continuous time signal, as a flow. For a detailed discussion of the choice of optimal embedding dimension and delay time for river-flow time series, reference is made to Porporato & Ridolfi (1996, 1997, 2001).

Unfortunately, Takens’ theorem has limitations in its practical application (Casdagli et al., 1991; Porporato & Ridolfi, 2001). However, even in the case of partial reconstruction, careful applications of the embedding techniques can be very useful to study the dynamics underlying the time series (under particular conditions, meaningful
state-space reconstructions can also be achieved for systems with high-dimensional forcing (Richter & Schreiber, 1998)).

From the hydrological viewpoint, there are various reasons for information loss in the reconstruction. Firstly, when going from the governing variables to discharge, there is a projection error due to both noise in the measurement and spatial variability of driving variables (and more generally to their high dimensionality): as a consequence, a discharge sequence given by equation (1) may correspond to different states of the dynamical system. Secondly, because of their high-dimensional components, the climatic forcings (mainly rainfall and temperature) can be reconstructed in only a very small part from their past history. However, even with these limitations, a considerable amount of information about the system can be obtained by studying the univariate relationship (equation (2)) and, most important for the present investigation, if one finds evidence of determinism and/or nonlinearity in the reconstructed attractor, these necessarily have to be present in the original state space as well.

**Poincaré sections**

Once the state space has been reconstructed, Poincaré sections may be employed to reduce dimensionality. These were previously employed by Porporato & Ridolfi (1996, 1997) to evidence the low-dimensional character of some phases (in particular recession periods) of the dynamics of the Dora Baltea time series. Their use is further pursued here to analyse the issue of nonlinearity, especially during recessions.

Consider the three-dimensional embedding space \{q_i, q_{i-\tau}, q_{i-2\tau}\}, and perform a section of the reconstructed attractor with a plane orthogonal to the main diagonal. A new reference frame \{u, v, w\} can be defined, having the main diagonal of the embedding space as the first axis, and the third axis intersecting both the main diagonal and the original \(q_{i-2\tau}\) axis (Fig. 1). As will be shown in the applications, the

![Fig. 1 Reference frames used to study the Poincaré sections in the three-dimensional embedding space (see text for details).](image)
general shape of the Poincaré section may be sketched as in Fig. 1. One may notice that the points in the third quadrant, which correspond to discharge triplets during recessions, tend to collapse on a straight line. Only for the first parts of recessions, when discharge is still high (i.e. the farthest points from the origin on the third quadrant), will such curves start to deviate downward.

Using the variable defined in Fig. 1, the straight line in the third quadrant may be described by \( u < 0, \ v < 0 \) and \( w = av \), with \( a \) constant. The coordinate transformation between the new reference \( \{u, v, w\} \) and the embedding reference \( \{q_i, q_{i-\tau}, q_{i-2\tau}\} \) can be obtained by a simple rotation of the axes as:

\[
\begin{align*}
  u &= \frac{\sqrt{3}}{3} q_m \\
  v &= \frac{\sqrt{2}}{2} \left( -q_{i-2\tau} + q_{i-\tau} \right) \\
  w &= \frac{\sqrt{6}}{6} \left( -q_{i-2\tau} - q_{i-\tau} + 2q_i \right)
\end{align*}
\]

where \( q_m = (q_i + q_{i-\tau} + q_{i-2\tau})/3 \). Therefore, the conditions \( u < 0, \ v < 0 \) become \( q_{i-\tau} < q_{i-2\tau}, \ q_i < \frac{1}{2} \left( q_{i-\tau} + q_{i-2\tau} \right) \), simply stating that the second value of the point has to be lower than the first and that the third has to be lower than the arithmetic mean of the first two. More interesting is the third condition:

\[
q_i = \frac{1 + \sqrt{3}a}{2} q_{i-\tau} + \frac{1 - \sqrt{3}a}{2} q_{i-2\tau}
\]

Equation (4) is a second-order linear difference equation, whose solution is (Bender & Orszag, 1978):

\[
q_i = C_1 + C_2 \left( \frac{1 - \sqrt{3}a}{2} \right)^i
\]

where \( C_1 \) and \( C_2 \) are arbitrary constants. Expression (5) is often used in hydrology to model hydrograph recession curves in basins with permanent snow and glaciers (Tallaksen, 1995). The constant \( C_1 \) is related to the base flow, while \( C_2 \) and \( a \) bear the imprint of the geometric and hydraulic characteristics of the basin. Equation (5) allows a different mathematical interpretation of recession curves and proves that, for the time series under examination, at medium-low discharge values the recession curves are the result of a mostly stationary linear system. Only for higher discharge values, when the points start deviating downward from the straight line, will the dynamics arguably become significantly nonlinear.

Nonlinear redundancies

The use of redundancies to test for nonlinearity in time series was introduced by Palus et al. (1993; see also Palus, 1995), and then generalized by Prichard & Theiler (1995). Given \( m \) delay variables (equation (1)), redundancies are the \( m \)-dimensional extension
Detecting determinism and nonlinearity in river-flow time series

of the mutual information:
\[ R(m, \tau, \delta) = mH(m = 1, \tau = 0, \delta) - H(m, \tau, \delta) \]  

where \( H(m, \tau, \delta) \) the \( m \)-dimensional extension of the Shannon entropy of their joint probability distribution, \( p(q_1, ..., q_{j-(m-1)\tau}) \), and \( \delta \) is the partition of the embedding space used to estimate such probabilities. The marginal redundancy:
\[ R'(m, \tau, \delta) = R(m, \tau, \delta) - R(m - 1, \tau, \delta) \]

quantifies the amount of information regarding \( q_{j-(m-1)\tau} \) that is contained in \( \{ q_1, ..., q_{j-(m-1)\tau} \} \).

To increase the accuracy of the estimate, instead of using the \( m \)-dimensional entropies, Prichard & Theiler (1995) proposed the use of the second-order redundancies:
\[ R_2(m, \tau, r) \approx -m \log_2 \left[ C_2(m = 1, \tau = 0, r) \right] - \log_2 \left[ C_2(m, \tau, r) \right] \]  

and
\[ R'_2(m, \tau, r) \approx -\log_2 \left[ C_2(m = 1, \tau = 0, r) \right] + \log_2 \left[ \frac{C_2(m, \tau, r)}{C_2(m - 1, \tau, r)} \right] \]

which are related to the second-order correlation integral (Grassberger & Procaccia 1983):
\[ C_2(m, \tau, \delta) = \frac{1}{M} \sum_{\tau} \sum_{j=1}^{M} \delta \left( r - \| Q_i - Q_j \| \right) \]  

where \( \delta \) is the Heaviside step function (this function is zero, when the argument is lower than or equal to zero, and one otherwise), \( \| \cdot \| \) is the norm of a vector, and \( r \) is the radius of a region about the generic \( Q_i \).

Related to the (second-order) marginal redundancy is the (second-order) Kolmogorov-Sinai (KS) entropy:
\[ K_2 = \lim_{m \to \infty} \frac{R'_2(\tau = \tau_1) - R'_2(\tau = \tau_2)}{\tau_1 - \tau_2} \]  

which is the mean rate of information creation by the system. It is very important in characterizing the average predictability of a system of which it represents the sum of the positive Lyapunov exponents (i.e. long-term average exponential rates of divergence or convergence of nearby states in the phase space). KS entropy quantifies the average amount of new information on the system dynamics brought by the measurement of a new value of the time series. In this sense, KS entropy measures the rate of information produced by the system, being zero for periodic or quasi-periodic (i.e. completely predictable) time series, and infinite for white noise (i.e. unpredictable by definition).
Deterministic vs stochastic (DVS) test

The DVS plots (Casdagli, 1991; Casdagli & Weigend, 1993) are based on the nonlinear prediction method (Farmer & Sidorowich, 1987). These aim to assess the degree of determinism and nonlinearity by analysing the link between forecasting results and the extent of the portion of the attractor considered to model the dynamics. Briefly, once a suitable $m$-dimensional embedding space has been reconstructed using equation (1), a predictive model $\hat{q}_{i+T} = \varphi(Q_i)$ is searched for, which minimizes the square prediction errors ($T$ is the forecast interval). A fitting set $\{q_j\}_{j=1,...,N_f}$ is used to estimate the coefficients of $\varphi$, and a testing set $\{q_j\}_{j=1,...,N_t}$ is used to evaluate the model. For each point $Q_i$ of the testing set, an affine model:

$$q_{i+T}^{(f)} = \alpha_0 + \sum_{n=1}^{m} \alpha_n q_{j=n+T}^{(f)}$$

with coefficients $\alpha_0, \alpha_1, ..., \alpha_m$, is fitted to the $k$ nearest neighbours $Q_{j}^{(f)} (l = 1, 2, ..., k)$ of $Q_i$ on the attractor reconstructed from the fitting set. The model is then applied to $Q_i$ and the forecast $\hat{q}_{i+T}(k)$ is evaluated. Repeating this procedure for each point of the testing set, the mean absolute forecast error:

$$E_m(k) = \frac{\sum_{i=1}^{N_t-(m-1)\tau} |\hat{q}_{i+T}(k) - q_{i+T}|}{N_t - (m-1)\tau}$$

(13)

can thus be computed for given $m, \tau, k$. DVS diagrams are finally obtained by plotting the error $E_m(k)$ vs $k$ for different embedding dimensions.

Varying $k$ from small to high values is equivalent to increasing the dimension of the neighbourhood of the local model. Small values of $k$ correspond to a nonlinear, deterministic approach to modelling, while a large number of neighbours correspond to fitting a global, stochastic linear model. When the underlying nonlinear dynamics is low-dimensional, then DVS plots present a minimum (i.e. more accurate forecasts) at relatively small $k$ values. The opposite happens when the minimum is located towards the stochastic extreme, meaning that significant nonlinearities and deterministic behaviours are not evident in the reconstructed state space. The DVS tests have been previously used in hydrology by Sivakumar et al. (1999) to analyse rainfall data.

Wavering product

The last method employed is based on the so-called wavering product, which was introduced by Liebert et al. (1991) as a topological method to choose the optimum embedding parameters. A modification is proposed here to investigate some qualitative characteristics of the time series.

The wavering product is defined using the quantity:

$$P_\tau(m) = \left\{ \prod_{k=1}^{p} \frac{l_{m+1}^i[i,j(k,m)]}{l_{m+1}^i[i,j(k,m+1)]} \cdot \frac{l_{m}^i[i,j(k,m)]}{l_{m}^i[i,j(k,m+1)]} \right\}^{1/p}$$

(14)
Detecting determinism and nonlinearity in river-flow time series

where \( l_{m+1}^i[i, j(k, m)] \) is the distance, measured in \( \mathbb{R}^{m+1} \), between the generic reference point of the attractor \( Q_i \) and its \( k \)th nearest neighbour in \( \mathbb{R}^m \), whilst \( l_{m+1}^i[i, j(k, m+1)] \) is the distance, measured in \( \mathbb{R}^{m+1} \), between the same reference point and the point which is its \( k \)th nearest neighbour in \( \mathbb{R}^{m+1} \). Analogous definitions apply to \( l_{m+1}^i[i, j(k, m)] \) and \( l_{m+1}^i[i, j(k, m+1)] \) in \( \mathbb{R}^m \). \( P_i(m, \tau) \) attains value 1 only if the \( p \) nearest neighbours remain the same when going from \( m \) to \( m + 1 \). In such a case the embedding dimension \( m \) is sufficient to unfold the attractor, otherwise \( P_i(m, \tau) > 1 \). In order to eliminate the explicit \( \tau \)-dependence, and to underline small deviations in the reconstruction, the wavering product is defined as:

\[
W_i(m, \tau) = \frac{1}{\tau} \ln[P_i(m, \tau)]
\]

The wavering product can be used to recognize if the dynamics have different characteristics in different zones of the attractor. For the same embedding dimension, higher values of \( W_i \) indicate parts of attractor with a greater degree of complexity compared to those with \( W_i \) equal or close to zero for which the trajectories are better unfolded. In this way the wavering product is a measure of the degree of violation of the topological properties of the embedding and provides an indication of the local complexity of the attractor.

APPLICATION

The methods described above are now applied to daily-averaged discharge time series of three rivers in northern Italy. A map of the rivers is shown in Fig. 2, and the related characteristics are summarized in Table 1. The three rivers have been chosen because their basins have different sizes and, although they belong to the same main basin and

![Fig. 2 Map of the three rivers in northern Italy (the arrows indicate the measurement stations).](image-url)
Table 1 Main characteristics of the time series.

<table>
<thead>
<tr>
<th>River</th>
<th>Basin area (km²)</th>
<th>Mean elevation (m.a.s.l.)</th>
<th>Mean discharge (m³ s⁻¹)</th>
<th>Standard deviation (m³ s⁻¹)</th>
<th>Max. discharge (m³ s⁻¹)</th>
<th>Min. discharge (m³ s⁻¹)</th>
<th>Measurement period</th>
<th>Number of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dora Baltea</td>
<td>3313</td>
<td>2020</td>
<td>93</td>
<td>81</td>
<td>1260</td>
<td>18</td>
<td>01/01/1941–31/12/1979</td>
<td>14246</td>
</tr>
<tr>
<td>Tanaro</td>
<td>7985</td>
<td>663</td>
<td>139</td>
<td>191</td>
<td>2690</td>
<td>8</td>
<td>01/01/1953–31/12/1978</td>
<td>9496</td>
</tr>
<tr>
<td>Po</td>
<td>70091</td>
<td>200</td>
<td>1514</td>
<td>1024</td>
<td>8940</td>
<td>206</td>
<td>01/01/1935–31/12/1985</td>
<td>18628</td>
</tr>
</tbody>
</table>

Fig. 3 Comparison of a common part of the three time series: (a) Dora Baltea, (b) Tanaro, and (c) Po. Notice that the analysed time periods are longer (see Table 1).

are quite close in space, they present fairly distinct hydroclimatic regimes. A common part of the three time series is shown for comparison in Fig. 3. For the Dora Baltea River, one observes a basic annual cycle due to the presence of snow and glaciers. This recurrent behaviour becomes less marked for the Tanaro time series and almost disappears for the Po time series. For the latter, in particular, the series is less correlated at longer time scales and flood events are more frequent and irregular, due to the many sub-basins with different dimensions and climate regimes.
Poincaré sections

Poincaré sections are employed here to analyse the issue of nonlinearity, especially during recessions. Figure 4 shows a superposition of the Poincaré sections at different levels on the main diagonal for the three time series. Their general shape confirms the one sketched in Fig. 1. One can notice that the points in the third quadrant, which correspond to discharge triplets during recessions, tend to collapse on a straight line. Only for the first parts of recessions, when discharge is still high (i.e. the farthest points from the origin on the third quadrant), will such curves start to deviate downward. This fact is more evident for the Dora and Tanaro time series. The relatively larger spreading of points for low discharge values (especially for the Po River) should be attributable to a higher dimensionality of the dynamics and to the consequent stronger stochastic character of the time series (see also the values of the KS entropy that will be discussed below).

By means of equation (5), the collapse of the Poincaré sections onto a straight line allows a new mathematical interpretation of recession curves and proves that, for the time series under examination, at medium-low discharge values the recession curves are the result of a mostly stationary linear system. Only for higher discharge values, when the points start deviating downward from the straight line, will the dynamics arguably become nonlinear.

The results on recession curves comply with the fact that linear behaviours of recessions have been often reported in the literature (e.g. Vogel & Kroll, 1992), even if exceptions cannot be ruled out (e.g. Brutsaert & Nieber, 1977; Wittenberg, 1999). The method based on the Poincaré sections proposed here is expected to be a good tool to discriminate between linearity and nonlinearity in recessions, an important unsolved problem of hydrology (Furey & Gupta, 2000).

Redundancies and KS entropy

The nonlinear redundancies were computed through the second-order correlation integral according to equations (8) and (9). The correlation integral was computed employing the optimized box-assisted algorithm of Grassberger (1991) with correc-
Figure 5 shows the normalized redundancies $R_2(m-1)$ vs delay time for different embedding dimensions ($m = 2–9$): (a) Dora Baltea, (b) Tanaro, and (c) Po.

Figure 6 shows the marginal redundancies for the Dora Baltea River. These are qualitatively similar to normalized redundancies shown in Fig. 5(a). Moreover, since the behaviour is similar for the three rivers, only the results for the Dora Baltea are considered for oversampled data (Theiler, 1986). Details of the computation are given in Porporato & Ridolfi (1997).

Figure 5 shows the normalized redundancies for the three time series considered. Strong seasonal components (at a 12, 6, and 3–4 months time scale) appear evident for the Dora Baltea series, whilst the presence of seasonal components becomes less important for the Tanaro time series and is almost absent in the Po time series. For all of the three time series, the correlation level is remarkably higher when nonlinear links are considered: nonlinear correlation is important also at smaller time scales when the time series are dominated by flood dynamics.
shown here for reasons of space. Again, one can immediately see the great degree of nonlinear correlation among data, with a linear, slowly decreasing trend due to a persistent nonlinear correlation between data at a daily/weekly time scale. Curves relative to different embedding dimensions tend to follow this typical behaviour, until they deviate downward. This last behaviour is probably due both to the increasing reduction of point density with embedding dimension and to the seasonal components that are first felt by delay vectors with higher embedding dimensions.

The general trend of nonlinear marginal redundancy allows for the estimation (at least qualitatively) of the KS entropy. Similar trends with a more rapid decay are found for the Tanaro and Po time series. As expected, the values of the KS entropy are higher for these series ($K_2 = 0.037$ bits per day for the Dora Baltea time series, $0.096$ bits per day for the Tanaro time series, and $0.19$ bits per day for the Po time series). From a physical point of view, such results are quite interesting. As previously seen, KS entropy expresses the average degree of complexity of the time series and is related to the average degree of predictability of the time series. Thus, the values of the KS entropy seem to indicate that the complexity increases with the basin dimension and with the reduction of the importance of seasonal components due to snow and glaciers. As far as the predictability of the series is concerned, the estimated values of KS entropy are in agreement with the values of prediction error provided by nonlinear prediction. This can be seen by comparing the minima of the mean forecasting errors in DVS plots in Fig. 7 below.

**DVS plots**

In order to avoid problems due to possible undersampling of the three time series, especially during flood events (Porporato & Ridolfi, 1997), before evaluating the DVS plots, the three time series were interpolated using natural cubic splines and resampled at a rate equal to 12 h. Moreover, rather than plotting $E_m = E_m(k)$ for different embedding, it was decided here to compute DVS diagrams relative to different
Fig. 7 DVS plots \( (m = 9, \tau = 0.5 \text{ days}) \) for different discharge thresholds. (a) Dora Baltea, (b) Tanaro, and (c) Po. From the lowest to the highest curve, the thresholds considered for the three series are respectively: 0, 100, 200, 300, 400, 500, 600 m\(^3\) s\(^{-1}\) for the Dora Baltea and Tanaro, and 0, 1000, 2000, 3000, 4000, 5000 m\(^3\) s\(^{-1}\) for the Po time series.

discharge thresholds, i.e. considering for a given curve only data above a given discharge, so as to focus on different parts of attractor corresponding to different runoff conditions of the basin. Figure 7 shows DVS plots considering testing sets of 600 days, chosen in correspondence of flood events. The plots refer to \( m = 9 \) and \( \tau = 0.5 \text{ days} \) because these are the values for which, on average, the predictions are better; however, the behaviour of the DVS plot is similar (although less clear) for different values of the parameters. When the threshold is low or absent, there is a very wide plateau without clear minimum. In contrast, when the threshold is increased, a minimum appears in correspondence of medium-low values of \( k \). Its magnitude varies
for the three rivers, but this feature appears to be typical in all cases and it is always present. The minima are more pronounced for higher threshold values and at the minimum level the forecast quality is always at least 20% lower than that corresponding to the equivalent stochastic model at high \( k \) values. According to the interpretation of DVS plots, this behaviour would indicate that the nonlinearities and the determinism of the system only emerge at medium and high discharges, while, when the whole signal is taken into account, its dynamics would appear to be linear. One possible explanation could be that the deterministic action of the basin is always present, but for low discharges either the dynamics become more linear or it is increasingly submerged by noise of various kinds. Instead, for high discharges, the fundamental effects of the basin on runoff are more and more relevant compared to the background noise, so that the nonlinear determinism becomes evident. These results are in agreement with those of the Poincaré sections, which showed both linearity and stronger presence of noise at low discharge values during recessions. Thus the conclusions reached by analysing the attractor in dimension 3, remain valid also for higher embedding dimensions at which the attractor becomes (slightly) better unfolded.

**Wavering product**

The degree of local complexity on the attractor, \( W_i(m, \tau) \), was computed for each point of the time series (equations (14) and (15)), and then plotted as a function of the corresponding discharge value. The delay time was chosen equal to 1, and the embedding dimension was varied from 4 to 20. Since the results are qualitatively similar for the three time series, only the results for the Dora Baltea are shown here. Figure 8 presents an example for the Dora Baltea time series using \( m = 9 \). It can be observed that the mean value of \( W_i \) is greater than 0, meaning that the attractor is not entirely unfolded.
The progressive flattening of the points around zero when the discharge grows indicates that the portion of attractor corresponding to low discharge needs greater embedding dimensions to be properly unfolded, while, for the portion corresponding to high discharges, low values of \( m \) seem to be sufficient. When higher embedding dimensions are considered, no significant changes in the behaviour are noticed. The values of \( W_i \) corresponding to higher discharges remain on the zero axis (i.e. good attractor reconstruction) and only a weak lowering of the mean value of \( W_i \) takes place, caused by a slight thinning of the cloud at low discharge values. The general behaviour is in agreement with the conclusions of the previous methods, that for high discharge values the dynamics of rainfall–runoff transformation appears to be governed by a lower number of variables compared to low discharge values.

CONCLUSIONS

A variety of methods from nonlinear time series analysis has been used here to detect low-dimensional nonlinear components and the presence of nonlinear dynamics. Results from all the employed techniques show evidence of these components in all the time series, especially at high discharge values. Additional deterministic and nonlinear components of longer time scales are found when glaciers and snowmelt dynamics are important.

Regardless of the type of hydroclimatic regime of the river, at low discharges the time series become dominated by high dimensional fluctuations and nonlinearity in the dynamics becomes less important. As a consequence, the dynamics appears indistinguishable from a purely random noise, probably because noise from various sources (measurement errors, incomplete reconstruction due to the univariate approach, etc.) attains higher relative importance. As shown by the analysis of the Poincaré sections, the dynamics of recession curves at low discharge values appears to be the outcome of a mostly linear system, which is increasingly masked by noise as the discharge diminishes.

The results obtained are important as a guide for modelling: at least in the case of the three rivers considered, nonlinear low-dimensional models are more suitable for high flows, as high-dimensional models would only risk to overfit the underlying disturbances. At low flows simple linear stochastic models can be employed on which a main deterministic component of the form suggested by equation (5) is grafted. The analyses show that the main deterministic components are already evident in dimension 3 (see Poincaré sections), while the distinction between linearity and nonlinearity as a function of discharge becomes more precise as the attractor is better unfolded at higher embedding dimensions (up to \( m = 9 \)). The existence of a separation in the state space of the system between linear and nonlinear behaviour and between low-dimensional and high-dimensional components is important for a proper modelling of the system and suggests the use of local models whose boundaries are chosen on suitable embedding spaces of relatively low dimension. To all the effects, the remaining unresolved components will necessarily be seen as separate random components.
REFERENCES


Received 10 September 2002; accepted 19 July 2003