Signal separation with almost periodic components: a wavelets based method

O.A. Rosso[a], A. Figliola[a], S. Blanco[a], and P.M. Jacovkis[a,b]

[a]*Instituto de Cálculo, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires,*
Pabellón II, Ciudad Universitaria, 1428 Buenos Aires, Argentina,
e-mail: rosso@ic.fcen.uba.ar, rosso@ba.net
[b]*Departamento de Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires,*
Pabellón I, Ciudad Universitaria, 1428 Buenos Aires, Argentina.

Natural time series usually show either a combination of periodic phenomena with stochastic components or chaotic behavior. In many cases, when nonlinear characteristics are computed, they will essentially indicate the most remarkable effects and the results will underestimate or overestimate the real complexity of the system. For that reason signal separation of the frequency bands representing well known phenomena, like periodic or almost periodic behaviors, allows comprehension of the hidden nonlinear or stochastic phenomena involved. In this work a signal separation method based on trigonometric wavelet packets is described. The method has been applied, as an example, to a time series of daily mean discharges of the Atuel river in Argentina, that presents strong annual and semiannual oscillations due to meteorological effects. The correlation dimension and the maximum Lyapunov exponent of the residual time series were obtained taking away its known almost periodic components.

Keywords: Time-frequency signal analysis; wavelet analysis; signal separation; meteorological time series.

Las series de tiempo representan una combinación de fenómenos periódicos y componentes estocásticas o comportamiento caótico. En muchos casos, cuando se computan cuantificadores no lineales para dichas series temporales, es de desear que éstos resalten las características más notables de las mismas y que sus resultados no subvaloren o sobrestimen la complejidad real del sistema. Por esa razón, la separación de bandas de frecuencia que representan fenómenos bien conocidos, tales como el caso de comportamientos periódicos o cuasi-periódicos, permite la comprensión de fenómenos no-lineales y/o fenómenos estocásticos ocultos involucrados en la ge/-neración de dichas series temporales. En este trabajo un método de separación de señales basado en paquetes wavelet trigonométricos es descrito. El método ha sido aplicado, como un ejemplo, a una serie temporal de descargas media diarias del río de Atuel en Argentina. Esta serie temporal presenta una fuerte oscilación anual y semestral debido a efectos meteorológicos. La dimensión de la correlación y el máximo exponente de Lyapunov correspondientes a la serie de tiempo residual fueron obtenidos luego de eliminar las componentes cuasi-periódicas conocidas.

Descriptores: Análisis de señales en tiempo-frecuencia; análisis de wavelets; separación de señales; series temporales meteorológicas.

PACS: 05.45.Tp; 02.70.Hm; 92.40.Fb

1. Introduction

Time series corresponding to observable data of natural systems are frequently analyzed by a combination of linear and nonlinear dynamics. As Kantz and Schreiber emphasize in the introduction of their book [1], linear methods interpret all regular structures in a data set as a set of pure frequencies. That means that the intrinsic dynamics of the system is governed by the linear paradigm: small causes lead to small effects. Since linear equations with constant coefficients can only lead to exponentially growing (or decreasing) or periodically oscillating functions, all the irregular behaviors of the system are usually attributed to some random external input. Now, chaos theory has taught us that random input is not the only possible source of irregularity in the system's output: nonlinear chaotic systems can produce very irregular data with purely deterministic equations of motion. Of course, a system with both, nonlinear and random inputs, will most likely produce irregular data as well.

Meteorological time series are typical examples of observational data which usually present strong almost periodic components. In consequence the nonlinear contributions may be masked by them. In the evaluation of nonlinear metric parameters this situation is reflected, for instance, in a decreasing value of the correlation dimension and Lyapunov exponents. These facts are related to the algorithms employed for the evaluation of the nonlinear metric parameters and are independent of the intrinsic characteristics of the signals. It must be remarked that, in practice, many nonlinear phenomena are incorrectly evaluated when the periodic effects are still present.

The use of traditional filtering methods, such as signal separation methods, is specially difficult when we are trying to clean processes associated with low and medium frequencies, because the traditional filtering methods usually drag in the processes unwanted portions of the signal. In particular, the use of band-pass filters based on the Fourier transform introduces modifications in the signal (ringing effects, spurious oscillations, etc.) [1-3]. While filtering and signal processing have a long history for linear systems, new techniques are required for the analysis of signals generated by nonlinear systems, because traditional filtering processes alter the nonlinear metric invariants [4-10].

Wavelet analysis is a method which relies on the introduction of an appropriate Hilbert space basis and spans the signal in it. If the wavelets are required to form a proper orthogonal basis, it has the advantage that an arbitrary function can be uniquely decomposed and the decomposition can be inverted [11-18]. At this point we can also refer to Mallat [16]. He proved that in a one dimensional signal denoising process (assuming additive noise), an orthogonal wavelet based method is better than a method based on Fourier transforms, because wavelets do not change the original signal [16]. For this reason we use a signal separation method based on orthogonal wavelets: we try to analyze the remaining signal with a minimum modification in the associated dynamics. Observe that we have not yet assumed anything about the characteristics of this signal. On the other hand, if the associated dynamics is a chaotic one, the use of non orthogonal wavelets as a signal separation method, like traditional filters, could also change the dynamics [1,2,4-10].

As an example we show how a periodic low frequency signal with strong intensity, added to a chaotic Lorenz time series, modifies the values of the correlation dimension (D_2) and of the maximum Lyapunov exponent (Λ_m) computed for the corresponding chaotic attractor associated with the pure signal. Both invariant parameters were computed with different metric algorithms usually employed for this purpose in the literature [19-23]. To avoid this kind of signal contamination, a signal separation method with almost periodic components based on trigonometric wavelet packets [24,25] is introduced. An example applies this procedure to a natural time series with strong periodic effects (annual and semiannual) corresponding to the daily mean discharges of the Atuel river in Argentina. These two almost periodic effects are removed using the proposed method based on trigonometric wavelet packets. The signal separation process lets us observe a considerable increment in the values obtained for the correlation dimension and for the maximum Lyapunov exponent for the residual dynamical system.

This paper is organized as follow: Sec. 2 is devoted to an exposition of the signal separation method. In Sec. 3 we describe theoretical examples. We analyze the consequences of overimposing a strong periodic signal to a chaotic one generated by the integration of the Lorenz equations. We discuss an example of the signal separation method and compute D_2 and Λ_m over the natural time series before and after signal separation. Sec. 4 presents our conclusions.

2. Time–frequency analysis

2.1. Wavelet transform

Wavelet analysis is a method which relies on the introduction of an appropriate basis and a characterization of the signal by the distribution of amplitude in this basis. If the basis is required to be a proper orthogonal basis, any arbitrary function can be uniquely decomposed and the decomposition can be inverted [11-18]. Wavelet analysis is a suitable tool

for detecting and characterizing specific phenomena in time and frequency planes. The wavelet is a smooth and quickly vanishing oscillating function with good localization in both frequency and time. A wavelet family $\psi_{a,b}$ is the set of elementary functions generated by dilations and translations of a unique admissible mother wavelet $\psi(t)$:

$$\psi_{a,b}(t) \;=\; |a|^{-1/2}\psi\left(\frac{t-b}{a}\right), \tag{1}$$

where $a, b \in \mathcal{R}$ (the set of real numbers), $a \neq 0$ are the scale and translation parameters respectively, and t is the time.

The correlated decimated discrete wavelet transform (DWT) provides a nonredundant representation of the signal $\mathcal{S}$, and the values $\langle \mathcal{S}, \psi_{a,b} \rangle$ constitute the coefficients in a wavelet series. These wavelet coefficients provide relevant information in a simple way and a direct estimation of local energies at the different scales. Moreover, the information can be organized in a hierarchical scheme of nested subspaces called multiresolution analysis in $L^2(\mathcal{R})$. In the present work, we employ orthogonal cubic spline functions as mother wavelets. Among several alternatives, cubic spline functions are symmetric and combine in a suitable proportion smoothness with numerical advantages. They have become a suitable tool for representing natural signals [26,27].

In the following we assume that the signal is given by the sampled values $\{s_0(n), \; n = 1, \cdots, M\}$, which correspond to an uniform time grid with sampling time Δt. If the decomposition is carried out over all resolution levels $N = \ln_2(M)$, the wavelet expansion will be

$$\mathcal{S}(t) = \sum_{j=-N}^{-1} \sum_{k} C_j(k)\psi_{j,k}(t) = \sum_{j=-N}^{-1} r_j(t). \tag{2}$$

where the wavelet coefficients $C_j(k)$ can be interpreted as the local residual error between successive signal approximation at scales j and $j + 1$, and $r_j(t)$ is the detail of the signal at scale j. It contains the information of the signal $\mathcal{S}(t)$ corresponding to frequencies $2^j\pi \leq |\omega| \leq 2^{j+1}\pi$.

In the wavelet multiresolution framework it is possible to evaluate the energy corresponding to each level, that can be used for the detection of the characteristic events [24,25]. Since the family $\{\psi_{j,k}(t)\}$ is an *orthonormal* basis for $L^2(\mathcal{R})$, the concept of energy is linked with the usual notions derived from Fourier theory. The wavelet coefficients are given by $C_j(k) = \langle \mathcal{S}, \psi_{j,k} \rangle$ and the energy, at each resolution level $j = -1, \cdots, -N_J$, will be the energy of the detail signal

$$E_j = \|r_j\|^2 = \sum_{k} |C_j(k)|^2. \tag{3}$$

The total energy can be obtained as

$$E_{tot} = \|\mathcal{S}\|^2 = \sum_{j<0} \sum_{k} |C_j(k)|^2 = \sum_{j<0} E_j. \tag{4}$$

2.2. Trigonometric Wavelet Packets

As is well known, wavelet analysis provides a *time-scale* description for any finite energy signal. Essentially, it is a successive decomposition of the signal at different scales. At each step, the corresponding details are separated providing useful information for detecting and characterizing short time phenomena or abrupt changes of energy. However, since wavelets are not well defined in frequency, they are not well suited for describing and characterizing stationary phenomena or for detecting time-frequency structures. This is an important limitation because significant events often involve joint variations of time and frequency. Wavelet packets analysis, a natural extension of wavelet analysis, overcomes this problem. Moreover, this technique allows a time-scale-frequency description of the signals.

A family of trigonometric wavelet packets is a collection of elementary signals obtained from appropriate linear combination of wavelets [24,25]. They look as locally oscillating wave forms resembling modulated sines or cosines. Moreover, they can be organized as an orthonormal basis of the space of finite energy signals. The main advantage of using wavelet packets is that standard wavelet analysis can be extended with a flexible strategy. So, the description of the given signal can be well adapted to the significant structures. Several families of wavelet packets have been proposed in the literature [11-18]. Here we apply trigonometric spline wavelet packets [24,25]. First, let us have a brief review of the proposed technique.

Given a finite energy signal $s_0(t)$, using spline wavelet analysis we can successively decompose it with the following recursive scheme:

$$s_{j+1}(t) = s_j(t) \oplus r_j(t), \qquad (5)$$

for each scale $j = 0, -1, \cdots, -N$. As we already mentioned, the components $s_{j+1}(t)$ and $s_j(t)$ summarize the information of the signal corresponding to the frequency bands $2^{j+1}\pi \leq \omega \leq 2^j\pi$ and $2^j\pi \leq \omega \leq 2^{j-1}\pi$, respectively. This means that the decomposition at level j consists of filtering the components $s_{j+1}(t)$, giving the details corresponding to the remaining frequencies $2^j\pi \leq |\omega| \leq 2^{j+1}\pi$. The component $r_j(t)$ summarizes this information and we can describe the signal $\mathcal{S}(t)$ in term of detail signals as

$$\mathcal{S}(t) = \sum_{j<0} r_j(t). \qquad (6)$$

The detail components can be described in terms of wavelets atoms

$$r_j(t) = \sum_k C_j(k)\psi_{j,k}(t). \qquad (7)$$

Since each wavelet $\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$ is well localized in the interval $2^{-j}k \leq t \leq 2^{-j}(k+1)$, the corresponding coefficient $C_j(k)$ summarizes the local information of the detail. However, as we mentioned above, these coefficients

average all the involved frequencies $2^j\pi \leq |\omega| \leq 2^{j+1}\pi$ and we shall not have explicit information about stationary structures.

Now, we are interested in improving the frequency precision. The main idea is to decompose the components $r_j(t)$ in portions, each one covering a longer interval. We define any portion or local signal as

$$r_j^{(m,l)}(t) = \sum_{k=l}^{l+2^m-1} C_j(k)\psi_{j,k}(t), \qquad (8)$$

where the parameters m and l are chosen so that $r_j^{(m,l)}(t)$ covers the full time interval $2^{-j}l \leq t \leq 2^{-j}(l+2^m)$, which is a relative long interval of length 2^{m-j}. Note that we defined the local wavelet packet with 2^m basic functions $\psi_{j,k}(t)$ for $k = l, \cdots, l+2^m - 1$.

Now, we define the set of fundamental frequencies

$$\omega_{mh} = \pi + 2h\pi/2^m, \qquad (9)$$

with $0 \leq h \leq 2^{m-1}$ and associated Fourier matrix $\mathbf{M}^{(m)}$ given by

$$M_{n,k}^{(m)} = 2^{-\frac{m}{2}} \cdot \begin{cases} \sin[\pi(k+\frac{1}{2})], & \text{if } n=1; \\ 2^{\frac{1}{2}}\cos[\omega_{mh}(k+\frac{1}{2})], & \text{if } n \text{ is even}; \\ 2^{\frac{1}{2}}\sin[\omega_{mh}(k+\frac{1}{2})], & \text{if } n \text{ is odd}; \\ \cos[2\pi(k+\frac{1}{2})], & \text{if } n=2^m; \end{cases} \qquad (10)$$

with $1 \leq n \leq 2^m$, $0 \leq k < 2^m$ and $h = [\![n/2]\!]$, where $[\![\ ]\!]$ denotes the integer part. It can be demonstrated that $\mathbf{M}^{(m)}$ is a $2^m \times 2^m$ dimensional orthogonal matrix [24,25].

Then, we can define the new set of elemental functions in order to expand $r_j^{(m,l)}(t)$ as a 2^m dimensional vector obtained from

$$\theta_{j,n}^{(m,l)}(t) = \sum_{k=l}^{l+2^m-1} M_{n,k}^{(m)} \psi_{j,k}(t), \qquad (11)$$

for $1 \leq n \leq 2^m$.

Clearly, these functions constitute a new local orthonormal basis covering the interval under analysis $2^{-j}l \leq t \leq 2^{-j}(l+2^m)$. Therefore we can give a second description of the local signal as

$$r_j^{(m,l)}(t) = \sum_{n=1}^{2^m} D_j^{(m,l)}(n)\,\theta_{j,n}^{(m,l)}(t). \qquad (12)$$

The corresponding coefficients are easily computed as

$$D_j^{(m,l)}(n) = \sum_{k=l}^{l+2^m-1} M_{n,k}^{(m)} C_j(k), \qquad (13)$$

where $1 \leq n \leq 2^m$.

The trigonometric wavelet packets $\theta_{j,n}^{(m,l)}(t)$ have zero mean, oscillate on the interval $2^{-j}l \leq t \leq 2^{-j}(l+2^m)$

and decay with exponential ratio. Moreover, their waveforms resemble modulate sines or cosines. In fact, it can be demonstrated that each Fourier transform $\hat{\theta}_{j,n}^{(m,l)}(\omega)$ is centered at the fundamental frequency ω_{mh}, when $n = 2h$ or $n = 2h + 1$. Moreover, $\hat{\theta}_{j,n}^{(m,l)}(\omega) = 0$ on the other fundamental frequencies and has a fast decay outside the range $2^j \pi \leq |\omega| \leq 2^{j+1}\pi$.

In other words, the coefficients $\{D_j^{(m,l)}(n)\}$ can be considered as the *discrete Fourier spectrum* for the local signal $r_j^{(m,l)}(t)$. Summing up, we can resume in the double set of coefficients $\{C_j(k), D_j^{(m,l)}(n)\}$ the time-scale-frequency information of the local signal $r_j^{(m,l)}(t)$.

Finally, to analyze the complete function $r_j(t)$, that is, the details at level j, we choose some partition in local components $r_j^{(m_i, l_i)}(t)$, according to the structure of the signal

$$r_j(t) = \sum_{m_i} r_j^{(m_i, l_i)}(t), \qquad (14)$$

where the sequence of index l_i verifies $l_{i+1} = l_i + 2^{m_i}$. Then, we implement the above referred time-scale-frequency technique for each local signal.

The wavelet packets $\theta_{j,n}^{(m,l)}$ have their time definition in the interval $2^{-j}l \leq t \leq 2^{-j}(l + 2^m)$, that we can make as small as the sample rate allows us to do. We use this characteristic to separate only the undesirable frequencies. So, combining the expression given by Eq. (7) with the packet $\theta_{j,n_0}^{(m_0, l)}$ of our interest, we build the quasimonofrequency time series with the wavelet packet coefficients and the corresponding inverse transformation. This is feasible because in all cases we used orthonormal bases for our calculations.

2.3. Signal separation scheme

To sum up, the signal separation method that we propose consists of the following steps:

- To apply the DWT to the data series under study, according to Eq. (2), obtaining the signal expansion in a wavelet basis.

- To calculate, with the computed wavelet coefficients, the corresponding energies for each frequency band and along the series [see Eq. (3)].

- To identify the bands with higher energies and the j wavelet resolution level (these bands will be in general associated to "almost periodic" phenomena).

- To divide a j wavelet level energy band just identified in wavelet packets according to Eq. (13) (note that each packet corresponds to a narrow frequency band, whose width can be reduced when increasing the number of packets).

- To eliminate the chosen packets and then to reconstruct the corresponding wavelet packet without them [see Eq. (14)].

- To reconstruct with this and the other wavelet levels the total signal, without the periodic portion.

3. Numerical evaluation

3.1. Synthesized example

To test the hypothesis of this work about the incidence of the periodic or almost periodic effects and the metric algorithms of the nonlinear dynamics over time series, we used the Lorenz dynamical system. We integrated the system using a fourth-order Runge-Kutta with adaptive step size control [28] for chaotic parameters ($R = 45.92$, $\sigma = 16.0$ and $b = 4.0$) adding to the result a strong periodic signal. In Fig. 1 we show the convergence of Λ_m with the iteration process: the solid line corresponds to the exponent for the Lorenz system (x-component), $\Lambda_m = 2.19$, performed using the modified Wolf algorithm for time series [21–23]. We can see that the value obtained corresponds to the theoretical predictions $\Lambda_m = 2.16$ obtained with the Wolf algorithm for ordinary differential Eqs. [21].

In the same figure but in dot-dashed and dashed line we display the convergence of Λ_m evaluated for the same signal plus a periodic component (sinusoidal signal with frequency $\omega_0 = 1.0$) with an amplitude twice ($\Lambda_m = 0.29$) and four times as large as the Lorenz signal ($\Lambda_m = 0.008$), respectively. In the last two cases the largest Lyapunov exponents go to zero as they would in a quasiperiodic situation, indicating that the algorithm detects only the strong perturbation included. The original Lorenz system and the perturbed time series were constructed with the length of the natural time series to be analyzed in the next section. The election of the
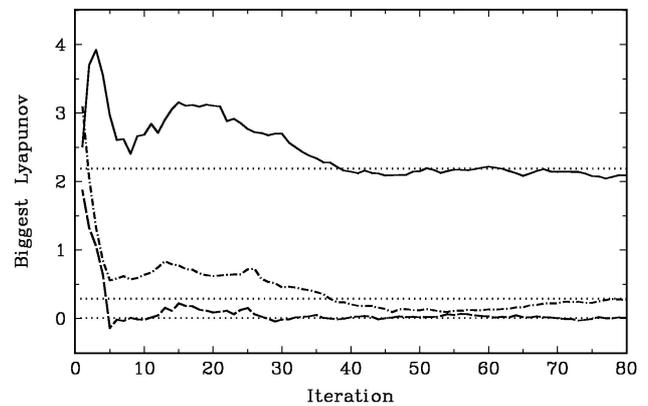

FIGURE 1. Convergence of maximum Lyapunov exponent Λ_m with the iteration process for its evaluation according with the modified Wolf algorithm for time series [21–23]. Solid line: Lorenz dynamical system with chaotic parameters (x-component). Dot-dash line: Lorenz dynamical system plus a sinusoidal signal of frequency $\omega_0 = 1.0$ and an amplitude twice the amplitude of the Lorenz system. Dash line: same as before but with a sinusoidal signal with an amplitude four times the amplitude of the Lorenz signal. The horizontal lines represent the convergent value for the three cases considered; they are $\Lambda_m = 2.19$, 0.29 and 0.008 respectively.

Wolf algorithm [21–23] is due to the fact that it is the most frequently employed algorithm in the analysis of natural time series. Similar results were obtained using the method proposed by Rosenstein *et al.* [29].

3.2. Natural time series

As an application of the signal separation method proposed and to analyze the remaining dynamics of a system with the characteristic mentioned above, we used a time series of daily discharges of the Argentinian Atuel river. Argentinian rivers whose sources are in the Andean Cordillera are extremely important for the economic development of the region (north of $38°$ S). The climate of that area is arid or semi-arid and river discharges are used for hydropower and for irrigation. The regime of these rivers is governed by strong annual waves: minimum discharges take place in winter and maximum discharges in summer. Water comes almost exclusively from melting of snow accumulated in the high peaks of the Cordillera during snowstorms that occur mainly between mid-autumn and the beginning of spring, and whose maxima are in winter [30]. The contribution of rains to discharges is almost nil, especially for points located at or upstream the foothills, because moisture from the Pacific Ocean cannot cross the high peaks of the Cordillera and moisture from the Atlantic Ocean yields only isolated connective storms in summer in the plains. The area of glaciers causes a regulating effect over the regime of these rivers, because blizzards act as regulating storage places that inhibit melting and store rains from one year to the following. It may happen that the ice and snow created during a year of high precipitation have not vanished when a poor year arrives, compensating the extremes of oscillation in the annual discharge of the river.

In spite of its regulating effect, the discharge series exhibits an important component of interannual variability, due to the large variability in intensity and frequency of winter snowstorms in this region of the Cordillera. Previous works [31, 32] showed the existence of a connection between the presence of the ENSO (El Niño Southern Oscillation) event in the Equatorial Pacific Ocean and rainfall data over the mean in central Chile. On the other hand, Compagnucci [31] found a relationship between the rainfall in Santiago de Chile, Chile, and the area of the Cordillera north of latitude $40°$ S, that corresponds to the zone of accumulation of the snow of the basin of the Atuel river.

Therefore, discharges, being determined by the amount of snow stored, will be sensitive to the ENSO event occurrences. In the Southern hemisphere (SH) summer, in the presence of ENSO, that is, of the warm sea surface temperature (SST) anomaly phase in the eastern Equatorial Pacific Ocean, discharges will tend to be above the average. On the other hand, in summers when the opposite event happens, *i.e.*, cold SST anomaly phase, named La Niña by Philander [33], streamflows below the average occur.

For this reason, it is likely that, besides the variability in discharges due to the annual wave, components of variability could be present due to the influence of ENSO events. Frequencies found in studies of ENSO signal include a broad band of the spectrum. Highest frequencies are given by biennial annual variability (around 18–35 months) in the Southern Oscillation (SO) and in the SSTs in the Equatorial Pacific Ocean, according to Rasmunsson *et al.* [34]. Other low frequencies belonging to the band around 32–88 months are also present in the spectral analysis of atmospheric or oceanic variables influenced by ENSO. This band corresponds to a current period of the event, because a characteristic of the ocean-atmosphere system is that the system oscillates successively from one extreme to the other of the signal, going from the warm SST anomaly and negative phase in SSO to the cold SST anomaly and positive phase in SSO, as was mentioned by Philander [34].

The time series includes the period from January 1959 to June 1995, that is, $M = 10228$ data points, and is shown in Fig. 2. This signal is long enough for the metric algorithms of nonlinear dynamics to be applied. We tested the stationarity of the series with a procedure based on the weak stationarity criteria introduced by our group [35]. Figure 3a shows the spectrum of this signal and the high peaks corresponding to the annual ($\omega_a = 0.0027$ 1/day) and semiannual periodicity ($\omega_b = 0.0054$ 1/day) are clearly spotted. Note that these peaks have a low but finite frequency dispersion.

After a multiresolution analysis using the wavelet transform and in agreement with a daily sampling rate ($\omega_{sample} = 1.0$ 1/day), the annual wave appears at level $j = -8$ and the semiannual at level $j = -7$. Both these levels have the maximum energy per band too, as we can see in Table I. The corresponding levels were split with the wavelet packets technique in 3 and 5 packets respectively. In this way we obtained the same frequency resolution for both wavelet levels, $\Delta\omega = 1.9531\ 10^{-3}$ 1/day. The annual wave was isolated in the second packet of the level $j = -8$, *i.e.*, $\theta_{8,n}^{(2,3)}$ and the semiannual wave in the second packet of the level $j = -7$, *i.e.*, $\theta_{7,n}^{(2,5)}$.
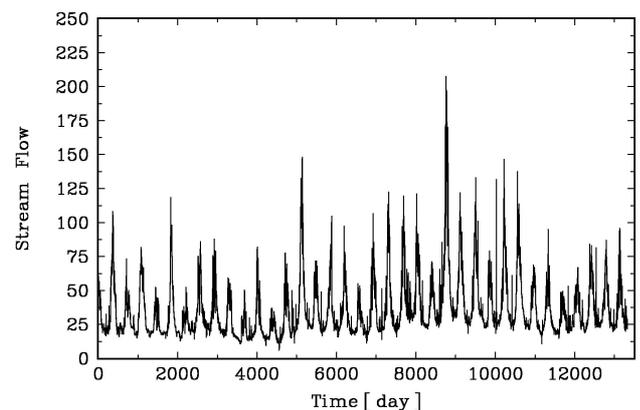

FIGURE 2. Daily mean discharges of the Atuel river in Argentina. The time series includes the period from January 1959 to June 1995. The number of data of the time series is $M = 10228$ and the sample rate is $\omega_{sample} = 1.0$ 1/day.
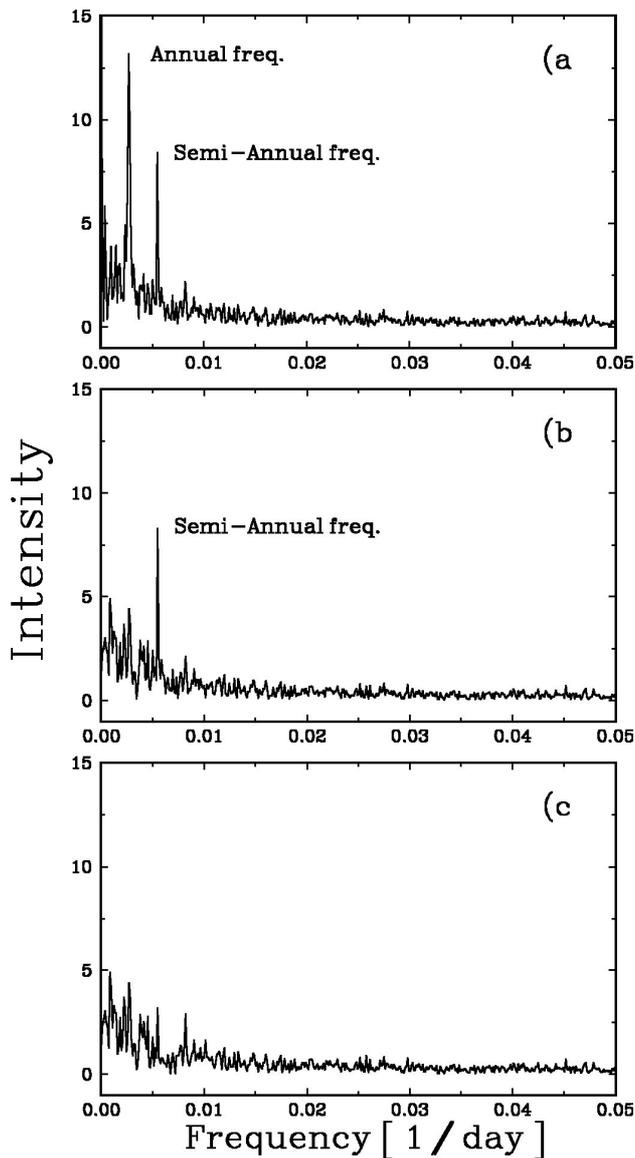
FIGURE 3. Power spectrum of the time series of daily discharges of Atuel River (Fig. 2) for: *a)* Original data. Note that the main peaks correspond to the annual and semi-annual frequencies. *b)* Residual time series without annual periodicity. *c)* Residual time series without annual and semiannual periodicity.

Starting with the coefficients of these trigonometric wavelet packets we reconstructed the two quasi-monofrequent time series and then subtracted them from the original time series. Figure 3b shows the spectrum of the Atuel river daily discharges after signal separation of the annual wave. It can be seen that the corresponding peak has disappeared without any visible alteration in the spectrum. In Fig. 3c we show the spectrum after signal separation of the semiannual frequency. The corresponding peak disappeared.

The minimum embedding dimension $(D_e^{(min)})$, time delay (τ), correlation dimension (D_2) and maximum Lyapunov exponent (Λ_m) were computed for the original and cleaned signal after the stationarity was checked [35]. The minimum embedding dimension $D_e^{(min)}$ for the attractors reconstruc-

TABLE I. Wavelet energies E_j and corresponding wavelets frequency bands associated ($\omega^{(min)}$, $\omega^{(max)}$) to a sample frequency $\omega_{sample} = 1.00$ 1/day for resolution wavelet level j.

j	E_j	$\omega^{(min)}$	$\omega^{(max)}$
-1	1585.1874×10^1	0.2500	0.5000
-2	4759.0858×10^1	0.1250	0.2500
-3	1011.6495×10^2	0.0625	0.1250
-4	1411.9919×10^2	0.0312	0.0625
-5	1862.6291×10^2	0.0156	0.0312
-6	2265.6863×10^2	0.0078	0.0156
-7	8891.0145×10^2	0.0039	0.0078
-8	3201.7304×10^3	0.0018	0.0039
-9	4893.5804×10^2	0.0009	0.0018

tion was calculated using the false neighbors method [2, 36] and the attractor reconstruction expansion was used as a geometry-based framework for choosing proper times delays [37]. The convergence of correlation dimension D_2 [19, 20] and maximum Lyapunov exponent Λ_m [21–23] were tested for five successive increasing values of the embedding dimension D_e taken as starting value $D_e^{(min)}$.

Figure 4 displays the convergence of the largest Lyapunov exponent of the two time series evaluated with the modified Wolf method for experimental data [21–23]. The dot-dashed line corresponds to the residual time series ($\Lambda_m = 1.32\ 10^{-2}$) and the solid line ($\Lambda_m = 6.18\ 10^{-4}$) to the original one. We see that the largest Lyapunov exponent falls about two orders of magnitude with respect to the original time series. The correlation dimension evaluated using Grassberger and Procaccia algorithm [19,20] for crude series is about $D_2=2.5$ and for the residual series is about $D_2=4.2$.
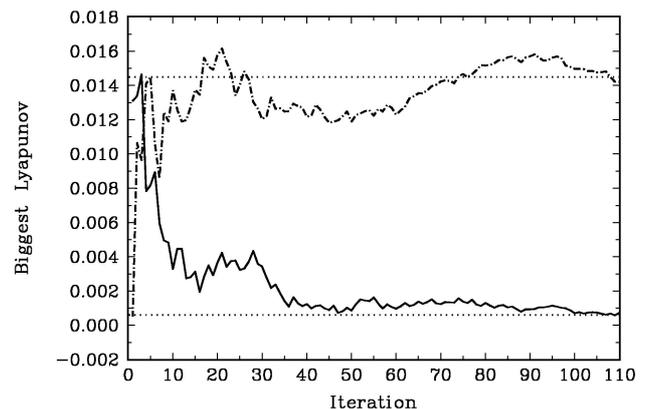

FIGURE 4. Convergence of maximum Lyapunov exponent Λ_m for time series of daily discharges of the Atuel River with the iteration process for its evaluation. Solid line: Original data, with $D_e^{(min)} = 6$ and $\tau = 16$ days. Dot-dash line: Residual time series without annual and biennial periodicity, with $D_e^{(min)} = 10$ and $\tau = 7$ days. The horizontal lines represent the convergent value for the both cases considered, $\Lambda_m = 6.18\ 10^{-4}$ and $1.32\ 10^{-2}$ for original and residual time series, respectively.

The minimum embedding dimension estimated through the false neighbors method [2, 36] again changes with the signal separation: it increases from $D_e^{(min)} = 6$ to $D_e^{(min)} = 10$. The time-lag employed [37] was $\tau = 16$ days and $\tau = 7$ days for the original and residual time series, respectively.

Physical quantities are almost never known exactly; therefore it is essential to provide an error estimate for all quantities, measured or calculated. In the case of nonlinear metric invariants (correlation dimension, largest Lyapunov exponent, etc.) this is surprisingly difficult and, according to some authors, almost impossible except in favorable situation [1, 38]. This problem is closely linked to the fundamental problems of time series analysis: given only a finite amount of data all conclusions are tentative. Numerical algorithms to estimate generalized dimension, Lyapunov exponents or other measures of the dynamical properties of the time series will usually produce some numbers without giving any information about the accuracy of results. It would be desirable not only to estimate the statistical error of the result, but also to assess whether the result was biased by numerical artifacts and systematical sources of errors. That means: could the same results be obtained by other time series which are similar in some respect to the original data, but for which we note that results would be biased or completely spurious? This question leads to the more general problem of statistical hypothesis testing: giving the data, which is the probability that the data are a realization from a certain class of processes? Surrogate data tests have been designed to provide some answers to this complex of problems and questions. For a very good discussion on this topics see Ref. 8.

4. Conclusions

Taking into account these results, the analysis of the dynamics associated to the Atuel river can be posed as the superposition of two phenomena, one "almost periodic", that explains the annual and semiannual variabilities, and the other corresponding to a phenomenon unknown but, most likely, nonlinear. As periodic phenomena tend to have higher energy, they mask the nonlinear dynamics, so that we consider that the proposed method is a good manner of separating them without altering the existent dynamics. On the other hand, the regional economies depend on the flood regime of the Atuel river, due to its direct incidence in the farming and cattle raising development of the region, so that a detailed study of it has economic importance.

The evaluation of nonlinear metric invariants is a difficult business, which can not be employed in the same way as other standard tools of time series analysis, for example, Fourier power spectrum estimation. This implies that metric algorithms must be used carefully in the analysis of natural time series that present strong periodicity and, therefore, the laborious method implemented for separated the known periodicity from the original signal are justified.

Regarding the results obtained for the natural time series studied (daily discharges of the Atuel river in Argentina) we can infer that the system's dynamics may be described as a superposition of phenomena ruled by a linear dynamics associated to annual and semiannual periodicities, plus other residual phenomena. In particular, the remaining phenomena, according with the previous results, present slightly chaotic characteristics, that could also be associated with colored noise. A more specific analysis using for example surrogate methods or similar (work in this direction are in progress) could be employed. Again we stress that the river streamflow data were used as an example of the proposed signal separation method described in this work.

In summary, in this paper we present a method, based on the wavelet transform, to separate signals according to specific frequencies. The method permits to remove exactly the desired frequency without perturbing other parts of the signal and leaving unaltered the dynamics of the remaining series. This methodology is appropriate to treat natural series in which there is a strong presence of a periodicity that has a significant part of the energy of the signal. Removing this portion of the signal usually affects the remaining components of frequency, above all when we are treating a slow periodicity.

Many natural phenomena exist that present signals with almost periodic components; for instance, fluctuations in pulsars, movements of planets and satellites, meteorological series affected by the solar activity, that forces a 24 hour period, etc. In ionospheric series usually the 11 years solar cycle appears. Something similar happens with physiologic series, for instance series that record presence of hormones or enzymes in the body (*i.e.*, circadian phenomena). In short, any series with a strong pulsation may be treated with this method. On the other hand, the method can easily be numerically implemented and it does not need much computer time.

Acknowledgments

1. H. Kantz and T. Schreiber, *Nonlinear time series analysis* (Cambridge University Press, New York, 1997).

2. H.D.I. Abarbanel, *Analysis of observed chaotic data* (Springer, New York, 1996).

3. R. Quian Quiroga, O.W. Sakowitz, E. Basar, and M. Schürmann, *Brain Research Protocols* **8** (2001) p 16.

4. R. Badii *et al.*, *Phys. Rev. Lett.* **60** (1988) 979.

5. F. Mitschake, M. Möller, and W. Large, *Phys. Rev. A* **37** (1988) 4518.

6. F. Mitschake, *Phys. Rev. A* **41** (1990) 1169.

7. A. Chennaoui, K. Pawelzik, W. Liebert, H.G. Schuster, and G. Pfister, *Phys. Rev. A* **41** (1990) 4151.

8. J. Theiler, B. Galdrikian, A. Longtin, S. Eubank, and J.D. Farmer, *Nonlinear Modeling and Forecasting.* Proceedings Volume XII, Santa Fe Institute Studies in the Sciences of Complexity. Edited by M. Casdagli and S. Eubank (Addison-Wesley, Redwood City, 1992) p. 163.

9. J. Theiler and S. Eubank, *Chaos* **3** (1993) 771.

10. M. Rosenstein and J.J. Collins, *Comput. & Graphics* **18** (1994) 587.

11. Y. Meyer, *Ondelettes and Operateurs,* Books I-II (Hermann, Paris, 1990).

12. C.K. Chui, *An Introduction to Wavelets* (Academic Press, San Diego, 1992).

13. I. Daubechies, *Ten Lectures on Wavelets* (SIAM, Philadelphia, 1992).

14. A. Aldroubi, M. Unser (Eds). *Wavelets in Medicine and Biology* (CRC Press, Boca Raton, 1996).

15. M. Akay, *Time frequency and wavelets in biomedical signal processing* (IEEE Press, Piscataway, 1998).

16. S. Mallat, *A Wavelet Tour of Signal Processing,* Second Edition (Academic Press, San Diego, 1999).

17. V. Samar, A. Bopardikar, R. Rao, and K. Swartz, *Brain and Language* **66** (1999) 7.

18. D.B. Percival and A.T. Walden, *Wavelet methods for time series analysis* (Cambridge University Press, New York, 2000).

19. P. Grassberger and I. Procaccia, *Phys. Rev. Lett.* **50** (1983) 346.

20. P. Grassberger and I. Procaccia, *Phyica D* **9** (1983) 189.

21. A. Wolf, J.B. Swift, H.L. Swinney, and J.A. Vastano, *Physica D* **6** (1985) 285.

22. L.D. Iasemedis, J.C. Sackellares, H.P. Zaveri, and W.J. Williams., *Brain Topography* **2** (1990) 187.

23. L.D. Iasemedis and J.C. Sackellares, "The evolution with the time of spatial distribution of largest Lyapunov exponent on the human epileptic cortex", in D. Duke and W. Pritchards (Eds), *Measuring Chaos in Human Brain* (World Scientific, Singapore, 1991) p. 49.

24. E. Serrano. Ph.D. Thesis, Department of Mathematics, FCEyN, Universidad de Buenos Aires, Buenos Aires, Argentina, 1996.

25. S. Blanco, A. Figliola, R. Quian Quiroga, O.A. Rosso, and E. Serrano, *Phys. Rev. E* **57** (1998) 932.

26. M. Unser, *IEEE Signal Processing Magazine* **16** (1999) 22.

27. P. Thévenaz, T. Blue, and M. Unser, *IEEE Trans. on Medical Imaging* **19** (2000) 739.

28. W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipies in FORTRAN,* Second Edition (Cambridge University Press, New York, 1995).

29. M.T. Rosenstein, J.J. Collins, and C.J. De Luca, *Physica D* **65** (1993) 117.

30. K.J. Prohaska, "Climates of Argentina, Uruguay and Paraguay" *Climates of Central and South America,* 12 (Elsevier, Amsterdam, 1976) p. 13.

31. R.H. Compagnucci and W.M. Vargas, *International Journal of Climatology* **18** (1998) 1593.

32. R.H. Compagnucci, S. Blanco, A. Figliola, and P.M. Jacovkis, *Environmetrics* **11** (1998) 251.

33. P.H. Philander, *Journal of Atmospheric Sciences* **42** (1985) 2652.

34. E.M. Rasmusson, X. Wang, and C.P. Ropelewesk, *Journal of Maritime Systems* **1** (1990) 71.

35. S. Blanco, H. García, R. Quian Quiroga, L. Romanelli, and O.A. Rosso, *IEEE Eng. in Med. and Biol.* **14** (1995) 395.

36. M.B. Kennel, R. Brawn, and H.D.I. Abarbanel, *Phys. Rev. A* **45** (1992) 3403.

37. M.T. Rosenstein, J.J. Collins, and C.J. De Luca, *Physica D* **73** (1994) 82.

38. A. Galka, *Topics in nonlinear time series analysis* (World Scientific, Singapore, 2000).