Asociación Argentina de Mecánica Computacional

Mecánica Computacional Vol XXXVI, págs. 421-429 (artículo completo)
José G. Etse, Bibiana M. Luccioni, Martín A. Pucheta, Mario A. Storti (Eds.)
San Miguel de Tucumán, 6-9 Noviembre 2018

BROWNIAN CHARACTERIZATION OF NONLINEAR BEHAVIOR OF RED BLOOD CELLS FROM LEUKEMIA AND ANEMIA PATIENTS THROUGH THE DETERMINATION OF THE MAY-SUGIHARA COEFFICIENTS AND THE HURST EXPONENTS

Santiago A. Bortolato[a,c], Manuel A. Mancilla Canales[a,b], Bibiana D. Riquelme[b], Susana Pérez[a], Alcides J. Leguto[a,d], Juan P. Rebechi[a,d] and Ana M. Korol[a,b]

[a] *Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Suipacha 570 Rosario, Argentina.*

[b] *Instituto de Física Rosario (IFIR), CONICET, Ocampo y Esmeralda, Rosario, Argentina.*

[c] *Instituto de Química Rosario (IQUIR), CONICET, Suipacha 570, Rosario, Argentina.*

[d] *Instituto de Biología Rosario (IBR), CONICET, Ocampo y Esmeralda, Rosario, Argentina.*

Keywords: Ektacytometry, Hurst Exponent, May-Sugihara, anemia, leukemia.

Abstract. Blood diseases affect one or several parts of it and prevent blood to accomplish its function. These diseases could be chronic or acute. For example, leukemia is a type of blood cancer that starts in the bone marrow where blood cells are formed. Also, leukemia can be classified as lymphocytic or myelogenous. In some types of leukemia, it could be affected by any of the different precursors of the cell lines of the bone marrow as for example myeloid, erythroid or megakaryocytic precursors. Besides, iron deficiency anemia is one of the most common anemia and is caused by a deficiency in iron present in the hemoglobin of red blood cells (RBCs). Ektacytometry techniques quantify RBCs deformability by measuring the elongation of suspended RBCs subjected to shear stress. A patented optical system denominated Erythrocyte Rheometer was used to evaluate the viscoelastic properties of RBCs from patients with leukemia and iron deficiency anemia by ektacytometry. RBCs from healthy donors were used as a control. From the diffraction patterns of several of millions of RBCs (subjected to shear stress), photometric series were obtained. On this temporal series, it was evaluated non-lineal quantifiers in order to study fluctuations in the elongations of the RBCs. The dynamics of the cells were analyzed with the coefficient of non-linear correlation proposed by May and Sugihara and with the Hurst Exponent. Patterns of different behavior were observed in the utilized quantifiers. The patterns observed for the leukemia patients showed a chaotic behavior while patterns from iron deficiency anemia showed a random behavior like the control RBCs. The proposed quantifiers proved to be useful in the discernment of different pathologies associated with RBCs. Besides, the results shed light in the utilization of non-linear quantifiers on photometrically recorded series from erythrocytes subjected to shear stress for a quantitative characterization of the behavior of the system under study.

1 INTRODUCTION

Red blood cells (RBCs) are the most abundant cells in the blood and its primary function is the transport of oxygen and carbon dioxide. In order to carry out this task, they have to travel through capillaries of diameters smaller than themselves. This requires that the RBCs must be deformed in a greater extent. Due to these mechanical properties and their abundance, RBCs are also major determinants for the rheological behavior of blood.

Leukemia is a cancer of the bone marrow and blood. The four main types of leukemia are acute myeloid leukemia, chronic myeloid leukemia, acute lymphoblastic leukemia, and chronic lymphocytic leukemia (CLL). Acute leukemia is a rapidly progressive disease that affects cells that have not fully developed. These cells can not perform their normal functions. Chronic leukemia tends to progress more slowly, and patients have a greater number of mature cells. In general, these more mature cells can perform some of their normal functions. In the case of acute myeloid leukemia, the cancerous change begins in a cell in the bone marrow that normally forms certain blood cells, that is, RBCs, some types of white blood cells and platelets. Thus, this type of leukemia can arise from degenerated RBCs, which also may present morphological alterations (Miller et. al. 2005). With respect to CLL, it has not yet been possible to establish a link between this disease and the morphological alteration of RBCs. However, it has been shown that many serious chronic diseases can damage the erythrocyte membrane by hyperthermia, and generate some degree of mechanical injury of the RBCs when crossing injured tissues (Zerga, 2004).

Iron-deficiency anemia (IDA) is an anemia caused by a lack of iron. Anemia is defined as a decrease in the number of RBCs or the amount of hemoglobin in the blood. When onset is slow, symptoms are often vague, including feeling tired, weakness, shortness of breath, or poor ability to exercise. Anemia that comes on quickly often has greater symptoms, including confusion, feeling like one is going to pass out, and increased thirst. Recently, it was reported that the decrease of RBCs during anemia can be attributed to an increase in membrane stiffness and a decrease in deformability, which decreases the ability of the RBCs to pass through the spleen without being removed (Nagababu et al. 2008).

Whatever the case, both diseases could modify the rheological behavior of blood since they could produce possible changes in the morphology of RBCs. In this work, a feasible diagnostic tool, based on the Brownian characterization of the nonlinear behavior of RBCs from CLL or IDA patients was developed. The new method consists of the use of the May-Sugihara Coefficient and the Hurst Exponent for the characterization of the RBCs samples. This technique enables the discrimination of different groups according to their nature.

In order to develop our proposal, the manifestation of the complex behavior related to RBCs is described through a range of new concepts: the techniques of Time Delay Coordinates suggested by Takens (1981), False Nearest Neighbors proposed by Abarbanel et al. (1993) and the method proposed by May and Sugihara (1990). The latter provides an estimation of the number of degrees of freedom and also makes inferences about the dynamical nature of the system under study. This inference can be carried out by examining the correlation coefficient between predicted and observed time series through different prediction intervals, depending on the embedding dimension of the attractor as long as the photometric time series is identified as chaotic. Lastly, Hurst Exponent would allow characterizing ordinary Brownian motion and fractional Brownian motion of the systems from photometrically recorded series.

The method requires obtaining time series that account for the alteration of the erythrocyte membrane of the RBCs. For this, photometrically time series of cells under shear stress from RBCs samples of patients with CLL or IDA were recorded. The time series were obtained by Ektacytometry technique using a device developed and patented by our Group of Applied Optics at Biology from IFIR (CONICET-UNR) (Riquelme et al. 2005). The same procedure was carried out for healthy patients (RBCs control) for comparative purposes.

2 MATERIALS AND METHODS

2.1 Red blood cells (RBCs)

Human venous blood samples of red cells Group 0 were anticoagulated with Na_2EDTA and maintained at 4 ºC until they were processed. Whole blood was centrifuged at 800 g during 10 minutes. Then, plasma and buffer coat were removed. The remaining RBCs were washed three times with phosphate buffer saline (PBS, pH 7.4) at 25 ºC.

Erythrocyte suspensions were obtained according to the experimental procedure described above and following the International Committee for Standardization on Hematology.

2.2 Data acquisition

The Erythrocyte Rheometer based on Ektacytometry, was used to dada acquisition. In this instrument, light diffraction under Fraunhofer theory conditions may be applied to obtain quantitative information of diffracting particles such as suspended RBCs. Cells in dilute suspension under shear stress take a three axial ellipsoidal shape having the major axis in the same direction towards the shear field direction. A laser beam that transverse perpendicularly a thin layer of cells suspension is diffracted producing a Fraunhofer diffraction pattern that is either circular when the cells are at rest or elliptical when they become deformed by a shear stress field.

Normal RBCs being at rest can be considered as a monodisperse population having discoid shapes with almost the same size. RBCs suspension is placed between two flint glass disks, and a driving motor allows the rotation of the lower disk. Data is stored for being numerically processing. Based on previous reports, the time series obtained were splitted in two parts: the first one describes the stage when the erythrocytes are subjected to shear stress and the second one describes the stage when the erythrocytes recover their circular shape for being at rest.

An example of the time series for the creep process and for a healthy control sample obtained by the equipment, as well as the Fraunhofer diffraction patterns for rest and under shear stress, is shown in Fig.1.

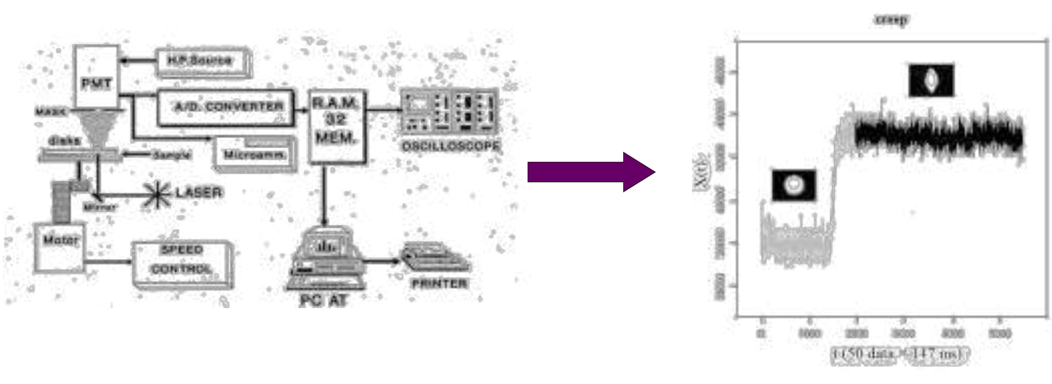

Figure 1. Flowchart of Erythrocyte Rheometer and a typical time series from the creep process of

erythrocytes under shear stress and Fraunhofer diffraction patterns for rest and under shear stress.

2.3 Software

The algorithms for estimate the quantifiers were developed in two different programming languages, R (2016) and MATLAB (2013); and are at disposal if required.

3 DATA ANALYSIS

3.1 The percentage of False Nearest Neighbors.

A particular time series, such as the dataset depicted in Fig. 1, can be connected in time, leading to an orbit or trajectory that represents the evolution of the system. The set of orbits starting from all possible initial conditions generates a flow in the state space and can be used to visualize the attractor of the system. However, limitations of such representation of the system include the conditions that every trajectory must be non-intersecting and that different trajectories originating from different initial conditions must not overlap or occupy the same space. This arises from the fact that a point in phase space representing the state of the system is considered to encode all the information about the system, including both its past and future history, which is a deterministic system must be unique.

In this phase space, points of an orbit acquire neighbors. These neighbors provide information on how phase space neighborhoods evolve over time. In an embedding dimension E, that is too small to unfold the attractor, not all the points that are close to each other will be actual neighbors due to the dynamics. Some of them will be far from each other and simply appear as neighbors because the geometric structure of the attractor has been projected onto a smaller space. Abarbanel and co-workers evaluated neighbors with increasing dimensions until no false neighbors remained (Abarbanel et al. 1993). They developed a specific method from geometrical considerations, known as percentage of false nearest neighbors (%FNN). The proposed method looks for a value for the minimum embedding dimension that correctly analyzes the dynamics of the process.

In D dimension each vector Y(k) will have a nearest neighbor $Y_{NN}(k)$, in the sense of the minimum Euclidean distance between them:

$$R_D^2(k) = \left|Y(k) - Y_{NN}(k)\right|^2 = \sum_{j=0}^{D-1} (y(k-j\tau) - y_{NN}(k-j\tau))^2 \tag{1}$$

$$\begin{aligned} R_{D+1}^2(k) &= \left|Y(k) - Y_{NN}(k)\right|^2 = \sum_{j=0}^{D} (y(k-j\tau) - y_{NN}(k-j\tau))^2 \\ &= \sum_{j=0}^{D-1} (y(k-j\tau) - y_{NN}(k-j\tau))^2 + (y(-D\tau) + y_{NN}(k-D\tau))^2 \\ &= R_D^2(k) + (y(-D\tau) + y_{NN}(k-D\tau)^2 \end{aligned} \tag{2}$$

From Eqs. (1-2), if R_D is the minimum distance between Y(k) and $Y_{NN}(k)$, then R_{D+1} must be smaller than R_D, but if it is bigger, then the attractor in R_D is not unfolded and those points appeared near because the real attractor has been projected on the smallest space on delay coordinates. This could be checked for increasing embedding dimensions until the %FNN is less than 1%. The calculation of the %FNN could be used as a test on measurements from the dynamical system in order to find the minimum embedding dimension E in which the attractor of the system is completely unfolded.

3.2 May and Sugihara Correlation Coefficient

First, the smallest embedding dimension E in which every trajectory must be non-intersecting was chosen, as obtained with the %FNN procedure. Time delay coordinates technique, suggested by Takens, was employed in order to generate the phase space portraiture for the system dynamic (Takens, 1981).

A convenient way to reconstruct the dynamics of the process is to unfold the time series by successively higher shifts. The shifts were defined as integer multiples of a fixed lag τ, where τ is an integer number defined as $\tau = m\ \Delta t$ (in our work, $m = 1$). Taking N equidistant points for creep and recovery process, we are able to define the phase space of all the possible states of the system variables under study.

For our time series, each sequence for which we wish to make a prediction is now to be regarded as an E-dimensional point, that is a vector, comprising the present value and the E-1 previous values each separated by one lag time.

May and Sugihara propose then that all nearby E-dimensional points are placed in the state space. For each of these points, a minimal neighborhood is defined to be such that the predicted one (Y*) is contained within the smallest simplex. A simplex containing E+1 vertices is the smallest simplex that can contain an E-dimensional point as an inside one. The lower dimensional simplex of nearest neighbors was used for points on the boundary.

Prediction is now obtained by projecting the domain of the simplex into its range, which is done by keeping track of where the points in the simplex end up after s time steps. To obtain the predicted value, we compute where the original predicted one has moved within the range of this simplex. This is a nonparametric method, which uses no prior information about the model used to generate the series; the only information is the output itself. It should apply to any stationary or quasi-ergodic dynamic process, including chaos.

Plotting the conventional statistical coefficient correlation $<Cs(Y,Y^*)>$ between predicted Y* and observed values Y as a function of *s*. If one obtains a decrease in the correlation coefficients with increasing prediction time then this is a characteristic feature of chaos. This property is noteworthy because it indicates a simple way to differentiate additive noise from deterministic chaos. The former is uncorrelated, regardless of how far or close into the future one tries to project the simplex, whereas predictions with the latter will tend to deteriorate as one tries to forecast further into the future.

Our predictions were generated by using the first half of the data series to construct an ensemble of points in an E-dimensional state space. This first half will be the library of past patterns; they were done avoiding the first 500 data points corresponding to the stationary process. Then the resulting information was used to depict the remaining second half values in the series.

3.3 Hurst Exponent

In a classical non-differentiable trajectory or, more generally, ordinary Brownian motion, past increments in displacement are uncorrelated with future increments, that is, the system has no memory. In a correlated random walk, or more generally, fractional Brownian motion, past increments in displacement are correlated with future increments, at least for the first steps of the process, hence the system has memory.

Hurst Exponent for a time series provides a measure of whether it is a pure white noise random process or has underlying trends. Dynamic process, that might naively characterized with purely white noise, sometimes turn out to exhibit Hurst Exponent statistics for long memory process, i.e., colored noise. A long memory process is a process where past events have a decaying effect on future ones. But those are forgotten as time moves forward.

A time dependent function X(t), as ours photometrically recorded time series, is said to be self-affine if fluctuations in different scales can be rescaled in order to obtained the original signal, to be statistical equivalent to the rescaled version (Simonsen, 2003):

$$X(t) = \lambda^{-H} x(\lambda t) \tag{3}$$

In Eq. (3), λ is a positive number and H is the Hurst exponent, which quantifies the degree of correlation (positive or negative) when the increments are $\Delta Y(t_i)=y(t_{i+1})-y(t_i)$.

It can be shown that for a process satisfying the self-affine property, the correlation function <C_s(Y,Y*)> between ΔY(t) of the real series and ΔY*(t) for the predicted one, can be best illustrated by considering the correlation coefficient for Brownian motion, proposed by Feder (1988), which is given by the expression:

$$C(t) = \frac{< \Delta Y(t) \Delta Y^*(t) >}{< [\Delta Y(t)]^2 >} = 2^{2H-1} - 1 \tag{4}$$

From Eq. (4), if H = 0.5, the increments in displacement are statistically independent, and then <C_s(Y,Y*)> = 0. This is the result expected for ordinary Brownian motion. For $H > 0.5$ past and future increments are positively correlated, this type of behavior is known as persistent. For $H < 0.5$ past and future increments are negatively correlated, this type of behavior is known as anti-persistent. In both cases <C_s(Y,Y*)> ≠ 0.

In order to obtain the correlation coefficient between the photometric time series of the second half of our series Y(t), and the theoretical Y*(t), we applied the Sugihara and May methodology. We correlated Y*(t), obtained from the series corresponding to the creep process, X(t), with Y(t) which is the second half, for the different steps increments s. For further details, please read the references (May and Sugihara, 1990; Feder, 1988).

4 RESULTS AND DISCUSSION

The random behavior is, as one expects, unpredictable, thus the question of randomness in a data series is more than a question of mixtures of determinism and randomness. Since noise is present in all physical measurements, determining if randomness is inherent in the system dynamics or in the measurement process is not always straightforward. In the study of the deformability of RBCs through the analysis of photometric series, much of what has just been said is present, which makes the use of nonlinear quantifiers a valuable contribution to unraveling what type of sample one is analyzing. Certainly, the use of a nonlinear quantifier is not intended to replace conventional analysis, but to provide further insights into the underlying RBCs deformation mechanisms.

At the start, over to all the series, the first differences ($x_{t+1} - x_t$) were applied in order to whiten the series, that is, to reduce the autocorrelation, and also to diminish any signals associated with simple cycles. In order to reconstruct the process dynamics, the technique of delay coordinates suggested by Takens and co-workers was used, and the Abarbanel method of false nearest neighbors was applied to finding the phase space dimension, as such explained in Section 3.1. Thus, the %FNN for seven embedding dimensions, one higher than the other, were obtained. For all the samples it was found that when E = 6 the %FNN was less than 1%, and therefore the system attractor should completely unfold. Thus, E = 6 was the chosen embedding dimension for our analysis. The results obtained are shown in Fig. 2.

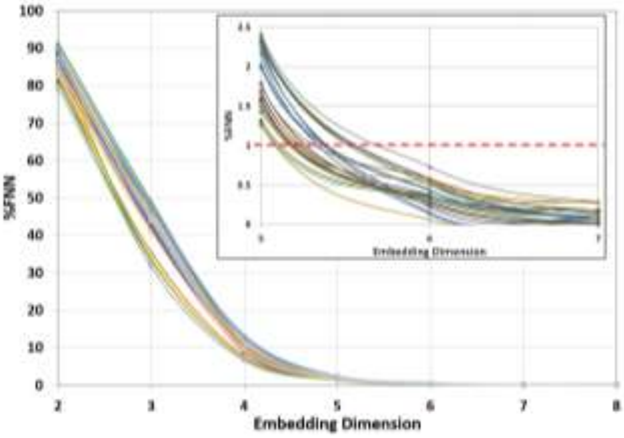

Fig. 2. Embedding dimension E vs. %FNN. Inset: %FNN for E = 5, 6 and 7.

Fig. 3 shows the outputs from the May and Sugihara analysis. Results showed differences comparing RBCs samples from CLL and IDA patients. During the first three steps, IDA samples generated lower correlations than the CLL ones (Fig. 3-B). However, some independence between the correlation and the step process could appear because we are dealing with uncorrelated additive noise.

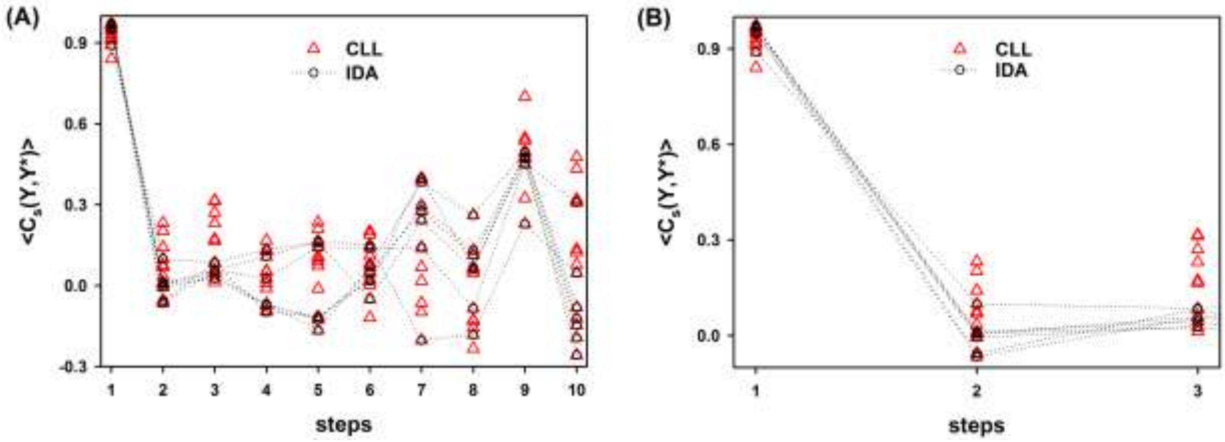

Fig. 3. A) Correlation coefficient versus steps (*s*) for RBCs from CLL (red triangle) and IDA patients (black circle and dashed line). B) Same figure, highlighting the first three steps.

The accuracy of the prediction for IDA samples, as measured by <Cs(Y,Y*)>, shows no systematic dependence on ***s***, between experimental and theoretical ones. By contrast, for CLL samples, the <Cs(Y,Y*)> does not decrease with increasing ***s***, which is characteristic of a chaotic dynamic.

Samples	Hurst (*s*=1)*	Hurst (*s*=2)*
Healthy controls	0.98 ± 0.01	0.46 ± 0.04
Leukemia (CLL)	0.98 ± 0.03	0.63 ± 0.07
Anemia (IDA)	0.98 ± 0.02	0.49 ± 0.02

Table 1: Mean Hurst Exponent for all the RBCs samples (*: mean H ± SD).

The mean Hurst exponents for all the RBCs samples (including healthy controls) at $s = 1$ and $s = 2$ are shown in Table 1. If the value of Hurst Exponent for healthy controls is observed (for $s = 2$), the dynamics of this system could be classified as an ordinary Brownian motion (i.e., H ca. 0.5), where statistical properties such as invariance or range are not related at all. Interestingly, the same goes for IDA RBCs samples ($H = 0.49 \pm 0.02$), which would indicate that deformation that this disease would produce on the RBCs (Nagababu et al. 2008) would not have a direct effect on the rheological properties of the blood. On the other hand, on CLL RBCs samples, the system dynamic could be characterized as a persistent fractional Brownian motion ($H > 0.5$), which is fractal in a statistical sense, that is, statistical properties are related over different time scales by way of a power law. In other words, the stress process gives us some special information of the relaxation one in a short time and the series exhibit a great sensitivity to initial conditions.

5 CONCLUSIONS

The proposed quantifiers proved to be useful in the discernment of different pathologies associated with RBCs. Besides, for a quantitative characterization of the behavior of the system under study the results shed light in the utilization of non-linear quantifiers on photometrical recorded series from erythrocytes subjected to shear stress.

Understanding the behavior of biological systems and how it is altered under pathological conditions is a promising way of diagnosis. Specifically, we have linked photometrical recorded time series of erythrocytes under shear stress with increasing embedding dimension E for the dynamical process and Hurts Exponent, to be able to compare and understand the changes on three different erythrocytes populations: healthy donor samples, patients with leukemia, and patients with iron-deficiency anemia. Different patterns were observed for the different populations studied. While in the case of leukemia the patterns showed a chaotic behavior (with deterministic component), in the case of the samples corresponding to iron-deficiency anemia patients, the behavior was random as well as the healthy control samples.

REFERENCES

Abarbanel, H.D.I., R. Brown, Sidorowich, J.J., Tsimring, L.S., The Analysis of Observed Chaotic Data in Physical Systems, *Reviews of Modern Physics* 65:1331–1392, 1993.

Feder, J., *Fractals*, Plenum Press, New York, 1988.

MATLAB. The Mathworks Inc., Natick, Massachusetts, USA, 2013.

Miller, K.B., Daoust, P.R. Clinical manifestations of acute myeloid leukemia. In: Hoffman, R, Benz, E.J., Shattil, S.J., Furie, B., Cohen, H.J., Silberstein, L.E., McGlave, P., eds. Hematology: Basic Principles and Practice. 4th ed. Philadelphia, Pa: Elsevier, 2005.

Nagababu, E., Gulyani, S., Earley, C.J., Rifkind, J.M., Mattson M.P., Cutler, R.J., Iron-Deficiency Anemia Enhances Red Blood Cell Oxidative Stress, *Free Radical Research* 42(9): 824–829, 2008.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (http://www.R-project.org/).

Riquelme, B., Foresto, F., D'Arrigo, M., Valverde, J. Rasia, R., A dynamic and stationary rheological study of erythrocytes incubated in a glucose medium, *Journal of Biochemical Biophysical Methods* 62:131–141, 2005.

Simonsen, I., Measuring anti-correlations in the nordic electricity spot market by wavelets *Physica A* 322:597–606, 2003.

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

Takens, F., Detecting strange attractors in turbulence. Dynamical systems and turbulence, Lecture Notes in Mathematics. 1:366-381, 1981.

Zerga, M.E., Anemia de los trastornos crónicos, Hematologia, 8:45–55, 2004.