ECEA 2019
5th International Electronic Conference on Entropy and Its Applications
18–30 November 2019

sciforum

Conference Proceedings Paper

Detection of Arrhythmic Cardiac Signals from ECG Recordings Using the Entropy-Complexity Plane

Pablo Martinez Coq [1,*], Walter Legnani [2] and Ricardo Armentano [2,3]

[1] Signals and Images Processing Centre (CPSI), Facultad Regional Buenos Aires, Universidad Tecnológica Nacional, Medrano 951, C1179AAQ, City of Buenos Aires, Argentine
[2] Signals and Images Processing Centre (CPSI), Facultad Regional Buenos Aires, Universidad Tecnológica Nacional, Medrano 951, C1179AAQ, City of Buenos Aires, Argentine
[3] Bioengineering Development and Research Laboratory (GIBIO), Facultad Regional Buenos Aires, Universidad Tecnológica Nacional, Medrano 951, C1179AAQ, City of Buenos Aires, Argentine
* Correspondence: Correspondence: pablomartinezcoq@gmail.com; Tel.: +54 11 4867 7579

Abstract: The aim of this work was to analyze in the Entropy-Complexity plane (**H**x**C**) time series coming from ECG, with the objective to discriminate recordings from two different groups of patients: normal sinus rhythm and cardiac arrhythmias.

The **HxC** plane used in this study was constituted by the Shannon's Entropy as one of its axis, and the other was composed by the statistical complexity. To compute the entropy the probability distribution function (**PDF**) of the observed data was obtained using the methodology proposed by Bandt & Pompe (2002) [1].

The database used in the present study was the ECG recordings obtained from PhysioNet, 47 long-term signals of patients with diagnosed cardiac arrhythmias and 18 long-term signals from normal sinus rhythm patients were processed. Average values of statistical complexity and normalized Shannon entropy were calculated and analyzed in the **HxC** plane for each time series.

The average values of complexity of **ECG** of patients with diagnosed arrhythmias were bigger than normal sinus rhythm group. On the other hand, the Shannon entropy average values for arrhythmias patients were lower than the normal sinus rhythm group. This characteristic made possible discriminate the position of both signals' groups in the **HxC** plane. The results were analyzed through a multivariate statistical test hypothesis.

The methodology proposed has a remarkable conceptual simplicity, and shows a promissory efficiency in the detection of cardiovascular pathologies.

Keywords: entropy-complexity plane; ECG; arrhythmic cardiac signals; Shannon's entropy; permutation entropy; Bandt & Pompe

---

1. Introduction

As its published by World Health Organization (**WHO**) of the 56.9 million deaths worldwide in 2016, light more than a half (54%) were due to the top 10 causes. Ischemic heart disease and stroke are the world's biggest killers, accounting for a combined 15.2 million deaths in 2016. These diseases have remained the leading causes of death globally in the last 15 years [2].

The American Heart Association explain that some arrhythmias are so brief (for example, a temporary pause or premature beat) that the overall heart rate or rhythm is not affected at clinical level. But if arrhythmias last longer, they may cause the heart rate to be too slow or too fast or the heart rhythm to be erratic—so the heart pumps less effectively [3].

The most important clinical significance of arrhythmias is related to an association with sudden cardiac death (Goldstein et al.; Recommended General Bibliography p. xvii; 1994) [4]. It is also important to remember that frequently arrhythmias (especially atrial fibrillation) may lead to embolism, including cerebral emboli, often with severe consequences. Also, it must remember that sometimes fast arrhythmias may trigger or worsen a heart failure.

For the analysis of certain time series coming from biomedical signals, such as an electrocardiogram (**ECG**) [5–7], there is a need to characterize the degree of organization or complexity of it. For this reason, the theory of information and the analysis of dynamics of the systems under study provide a way to obtain relevant results.

In order to analyze these time series, the first step is the computation of the **PDF**. For this, Bandt and Pompe introduced a simple and robust symbolic method that takes into account the time causality connected with the dynamics of the system. Without any kind of assumption of the dynamics originating the signal, the permutation method proposed by Bandt & Pompe for the computation of the probability distribution function allows them, to compute the entropy. Adding to the last, the complexity, computed from the Jensen Shannon information divergence, constitutes the system's behavior localization in an entropy-complexity plane [8–14].

As is mentioned in Olivares et al. [15], the concept of entropy has many interpretations arising from a wide diversity of scientific and technological fields. Among them is associated with disorder, with the volume of state space, and with a lack of information too. There are various definitions according to ways of computing this important magnitude to study the dynamics of the systems, and one of the most frequent that could be considered of foundational definition is the denominated Shannon entropy [16], which can be interpreted as a measure of uncertainty. The *Shannon entropy* can be considered as one of the most representative examples of information quantifiers.

Let be a continuous Probability Distribution Function associated, (**PDF**), be noted by $\rho(x)$ with $x \in \Omega \subset R$, a continuous variable measured, and $\int_\Omega \rho(x)dx = 1$; its associated *Shannon Entropy* $S[\rho]$ is defined by [17] as:

$$\mathrm{S}[\rho] = -\int_\Omega \rho(x)\ln\big(\rho(x)\big)\,dx. \tag{1}$$

This concept means a global measure of the information contained in the time series; it has a low degree of sensitivity to strong changes in the distribution originating from a small-sized region of the set $\mathbf{\Omega}$.

For a time series $X(t) \equiv \{x_t; t = 1, \ldots, M\}$, a set of $\boldsymbol{M}$ measures of the observable $\boldsymbol{X}$ and the associated **PDF**, given by $P = \{p_i; i = 1, \ldots, N\}$, with $\sum_{i=1}^{N} p_i = 1$ and **N** as the number of possible states of the system under study, the *Shannon entropy* [17] is defined in (2).

$$\mathrm{S}[\rho] = -\textstyle\sum_{i=1}^{N} p_i \ln(p_i). \tag{2}$$

Equation (2) constitutes a function of the probability $P = \{p_i; i = 1, \ldots, N\}$, which is equal to zero when the outcomes of a certain experiment denoted by the index k associated with probabilities $p_k \approx 1$ will occur. Therefore, the known dynamics developed by the dynamical system under study is complete. If the knowledge of the system dynamics is minimal, all the states of the system can occur with equal probability; thus, this probability can be modeled by a uniform distribution $P_e = \{p_i = 1/N; \forall i = 1, \ldots, N\}$.

It is useful to define the so-called *normalized Shannon entropy*, denoted as $H[P]$ which its expression is (3).

$$\mathrm{H}[P] = \frac{S[P]}{S_{max}}. \tag{3}$$

Another information quantifier applied in this work was the *Statistical Complexity Measure* (**SCM**) which is a global informational quantifier. All the computations made in the present work were done with the definitions introduced in [18] and improved by [19]. For a discrete probability distribution function $P = \{p_i; i = 1, \ldots, N\}$, associated with a time series, this functional *C[P]* is given by (4).

$$\mathrm{C}[P] = Q_J[P, P_e].H[P], \tag{4}$$

where H denotes the amount of "disorder" given by the normalized Shannon entropy (3) and Q_J is called "disequilibrium", defined in terms of the Jensen-Shannon divergence, given by (5).

$$Q_J[P, P_e] = Q_0 J[P, P_e] = Q_0 \left\{ S\left[\frac{P+P_e}{2}\right] - \frac{S[P]}{2} - \frac{S[P_e]}{2} \right\}, \tag{5}$$

and, Q_0 denotes the normalization condition for the disequilibrium which corresponds to the inverse of the maximum possible value of Jensen-Shannon divergence, that is $Q_{0=}J[P_0, P_e]$.

The *C[P]* quantifies the existence of correlational structures giving a measure of the complexity of a time series. In the case of perfect order or total randomness of a signal coming of a dynamical system, the value of the *C[P]* is identically null that means the signal possesses no structure. Between these two extreme instances, a large range of possible stages of physical structure may be realized by a dynamical system. These stages should be reflected in the features of the obtained PDF and quantified by a no-null *C[P]*.

The global character of the **SCM** is due that its value does not change with different orderings computed by the **PDF**. So, the *C[P]* quantifies the disorder but also the degree of correlational structures. Given that the statistical complexity do not only quantifies randomness but also the degree of correlation between structures and consequently is not a trivial function of entropy, in the sense that, for a given value of **H**, there is a range of possible values of C between a minimum value C_{min} and a maximum value C_{max} [20]. The corresponding continuous curves are shown in the Figure 3 and Figure 4.

Once evaluated the Shannon's entropy and the statistical complexity, the results can be displayed in the **HxC** plane (global quantifier) with the objective to discriminate an ECG coming from an arrhythmic diagnosed patient from a normal sinus rhythm one.

In general terms, these characterization method shows to be efficient to distinguish between stochastic nature and the deterministic chaos from different groups of time series because it displays typical and specific features associated with its dynamics' nature [21].

2. Results and Discussion

The groups of time series analyzed in this work were obtained from the PhysioNet [22] platform (managed by members of the Computational Physiology Laboratory of the M.I.T., Massachusetts Institute of Technology), which are available at https://physionet.org/ .

As it is shown in the Table 1, the registry of patients with normal sinus rhythm is made up of eighteen **ECG** recordings registered at the Beth Israel Deaconess Medical Center. On the other hand, the registry of patients with cardiac arrhythmias is made up of forty-seven **ECG** recordings, where twenty-five of them were taking Digoxin, a medication that is used to treat heart failure and certain types of irregular heartbeat, such as chronic atrial fibrillation. Likewise, an example of a normal sinus rhythm and a cardiac arrhythmia recording are represented in the Figure 1, and the age distribution of each group of patients is shown in Figure 2.

Table 1. Distribution of each group of interest.

	Normal sinus rhythm	Cardiac Arrhythmias
Recordings	18	47 [1]
Males	5	26
Females	13	21
Taking medication	not available	36

[1] Signals denominated from origin *221* and *234* were discarded because there was no relevant information in it.

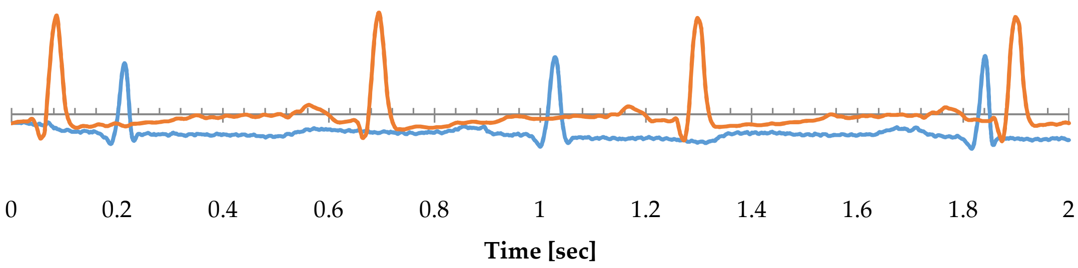

Figure 1. Orange trace represent record named from origin *16273* from MIT-BIH Normal sinus rhythm database, and light blue trace represents record *100* from MIT-BIH Arrhythmia database.

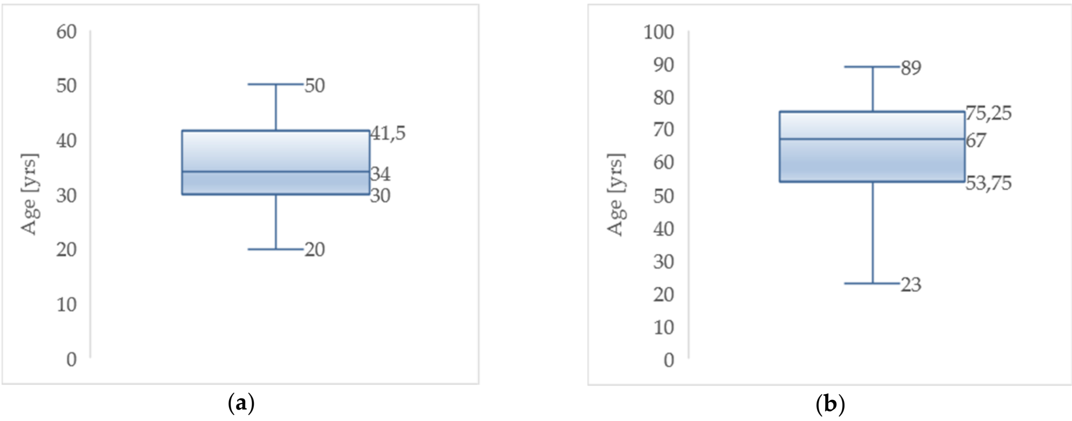

Figure 2. Age distribution of each group of patients: (**a**) Normal sinus rhythm recordings; (**b**) Cardiac arrhythmias recordings.

Several computational codes were developed to evaluate the values of statistical complexity *C* and the normalized Shannon entropy *H* for each one of the time series. Mean values of every pair of that calculus was represented in the **HxC** plane; where the statistical complexity was represented on **Y**-axis, and the normalized Shannon entropy was represented on **X**-axis.

Tables 2, 3, 4, 5, 6, and 7, are the summary of different comparisons of mean values, standard deviation, and the mean error of statistical complexity and the normalized Shannon entropy of the records under analysis.

Table 2. Mean **HxC** values for the whole database under analysis.

	Normal Sinus Rhythm		Cardiac Arrhythmias	
	Entropy - H	Complexity - C	Entropy - H	Complexity - C
Mean	0,89511	0,19060	0,81984	0,28982
S.D.	0,03002	0,04562	0,04924	0,05739
Mean Error	0,00728	0,01106	0,00726	0,00846

Table 3. Mean **HxC** values for records belonging to female patients of both groups of interest.

	Normal Sinus Rhythm		Cardiac Arrhythmias	
	Entropy - H	Complexity - C	Entropy - H	Complexity - C
Mean	0,90743	0,17234	0,82457	0,27700
S.D.	0,01765	0,02725	0,04656	0,04939
Mean Error	0,00509	0,00787	0,01016	0,01078

Table 4. Mean HxC Values for records belonging to male patients of both groups of interest.

	Normal Sinus Rhythm		Cardiac Arrhythmias	
	Entropy - H	Complexity - C	Entropy - H	Complexity - C
Mean	0,86553	0,23441	0,81588	0,30059
S.D.	0,03463	0,05370	0,05199	0,06228
Mean Error	0,01549	0,02402	0,01040	0,01246

Table 5. Mean Values of HxC for records belonging to patients up to 40 years of both groups of interest.

	Normal Sinus Rhythm		Cardiac Arrhythmias	
	Entropy - H	Complexity - C	Entropy - H	Complexity - C
Mean	0,90264	0,18140	0,82980	0,28171
S.D.	0,02859	0,04538	0,02211	0,02109
Mean Error	0,00825	0,01310	0,00903	0,00861

Table 6. Mean HxC Values for records belonging to patients from 40 up to 60 years of both groups of interest.

	Normal Sinus Rhythm		Cardiac Arrhythmias	
	Entropy - H	Complexity - C	Entropy - H	Complexity - C
Mean	0,87704	0,21269	0,82029	0,30795
S.D.	0,02789	0,04240	0,03341	0,05456
Mean Error	0,01247	0,01896	0,01181	0,01929

Table 7. Mean HxC Values for all records taking and not taking medication. Only for cardiac arrhythmias.

	Taking Medications		Not Taking Medications	
	Entropy - H	Complexity - C	Entropy - H	Complexity - C
Mean	0,81376	0,29037	0,82709	0,28917
S.D.	0,05606	0,06020	0,03976	0,05533
Mean Error	0,01121	0,01204	0,00868	0,01207

Results arising from Tables from 3 to 7 are represented in the **HxC** plane as it shown in Figure 3.

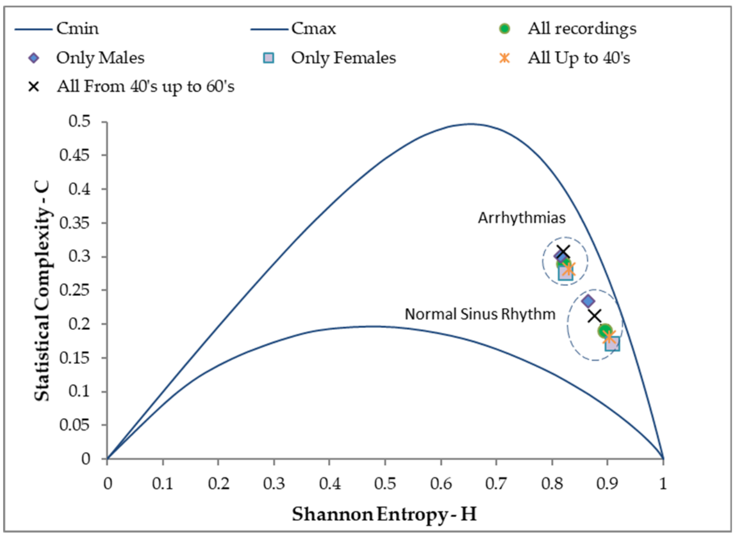

Figure 3. HxC plane. Mean values of statistical complexity and the normalized Shannon entropy of different comparisons of sub-groups of interest.

3. Conclusions

As is shown detailed in Figure 4, the average values of Statistical Complexity obtained for the group of patients with arrhythmias were higher than those of the normal heart rate group. On the other hand, the average values of normalized Shannon Entropy for the group of patients with arrhythmias were smaller than those of the normal heart rate group. The combination of the above characteristics allowed to discriminate the mean values of both groups of patients in the **HxC** plane. The difference between two groups of interest was analyzed through multivariate statistical tests (Test de Royston and Hotelling for samples of different sizes and with binormal distribution [23]). See Tables 8 and 9.

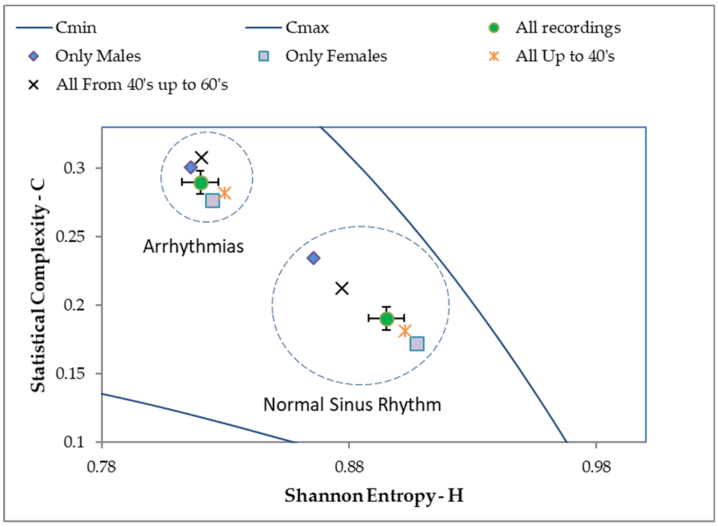

Figure 4. Detailed HxC for discriminating the mean values of both groups of interest.

Table 8. Royston Test for both groups of patients with a given significance of 0.050.

	Normal	**Cardiac Arrhythmias**
Royston's statistic	1,82783	0,15397
Equivalent deg. of freedom	0,527507	0,92392
P-value	0,082330	0,662806

According to values shown in Table 8, both groups of interest have a normal distribution.

Table 9. Hotelling Test for both groups of interest.

	T^2	$F(2,60)$	p
Normal vs Arrhythmias	40,141	19,741	0,0000002583

From values obtained from previous multivariate statistical tests, there was accepted the hypothesis than mean values of both groups of interest in the HxC plane are different.

This is showing that the signals from patients with normal **ECG** have less statistical complexity in their waveform, measured from divergence of information of Jensen-Shannon and higher entropy values. This result confirms, as studied in the literature [24], that a normal **ECG** would be closer to the regions of the entropy complexity plane that usually brings together the dynamic systems called dissipative (by decreasing their n-dimensional volume in the phase space) and that they comprise the systems that develop deterministic chaos (characteristic of the control system of a healthy heart), while an **ECG** from a patient with arrhythmia is located in a region of the **HxC** plane that characterizes the so-called k-noise [21] and fractional Brownian motion. This last result is indicating that cardiac arrhythmia could occur due to the loss of non-linear control (with the development of deterministic chaos behavior) of the heart giving a way to a more disorderly (random) dynamics typical of the pathology.

As seen in Table 7, about take Digoxin, mean values of Complexity and Entropy are quite similar, which could be interpreted as the information contained in the time series of the **ECG** do not allow to reject the hypothesis of different means between the interest groups in order to possible study the impact of the medication effects.

With the results obtained, an objective interpretation of a cardiac pathology can be made from the informational measures (complexity and entropy) when they are used together, forming what is agreed to be called the **HxC** plane, explaining in part of the disease process in itself.

The representation in the **HxC** plane is a very useful method for the analysis of signals from the ECG of patients with this kind of pathology, providing an efficient way to make a statistical and deterministic analysis. Finally, the proposed methodology has a remarkable conceptual simplicity and shows a promising efficacy in the detection of cardiovascular pathologies.

Acknowledgments: This work was partially supported by the (CPSI) Signals and Images Processing Centre of the Facultad Regional Buenos Aires of the Universidad Tecnológica Nacional, City of Buenos Aires, Argentine. and Universidad Tecnológica Nacional Grant PID UTN 4729.

Author Contributions: P. Martinez C. and W. Legnani conceived and designed the experiments; P. Martinez C. and W. Legnani performed the experiments; P. Martinez C. and W. Legnani analyzed the data; W. Legnani and R. Armentano contributed with his biomedical expertise; P. Martinez C, W. Legnani and R. Armentano wrote this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bandt, C.; Pompe, B. Permutation entropy: a natural complexity measure for time series. *Phys. Rev. Lett*. **2002**, 88, 174102.
2. WHO. The top 10 causes of death. **2018**. Last visited, June 2019. Available online at: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death
3. American Heart Association. About Arrhythmia. **2016**. Last visited, May 2019. Available online: https://www.heart.org/en
4. Bayés de Luna, A; Baranchuk, A. Clinical Arrhythmology. **2017**. *John Wiley & Sons Ltd*.
5. Khan, G. Rapid ECG interpretation. **2008.** *Humana Press Inc*.
6. Bayés de Luna, A. Basic Electrocardiography: Normal and Abnormal ECG patterns. **2007.** *Blackwell Publishing*.
7. Heart rate variability - Task Force of The European Society of Cardiology and The North American Society of Pacing and electrophysiology. **1996**. *European Heart Journal 17*, 354-381
8. Kolmogorov, A. A new metric invariant for transitive dynamical systems and automorphisms in Lebesgue spaces. **1959.** *Doklady Akademii Nauk SSSR.*, 119:861-864
9. Sinai Y. On the concept of entropy for a dynamical system. **1959.** *Doklady Akademii Nauk SSSR.*, 124:768-771
10. Olivares, F; Plastino, A; Rosso, O. Ambiguities in the Bandt and Pompe's methodology for local entropic quantifiers. **2012.** *Physica A: Statistical Mechanics and Its Application.*; 391:2518-2526.
11. Rosso, O.A; Larrondo, H.A; Martin, M.T; Plastino, A; Fuentes, M.A. Distinguishing noise from chaos. **2007.** *Physical review letters*, 99(15):154102.
12. Rosso, O; Olivares, F; Plastino, A. Noise versus chaos in a causal Fisher- Shannon plane. **2015.** *Papers in Physics*; 7:070006.
13. Rosso, O.A; Olivares, F; Zunino, L; De Micco, L; Aquino, A.L.L; Plastino, A; Larrondo, H.A. Characterization of chaotic maps using the permutation Bandt–Pompe probability-distribution. **2013.** *The European Physics Journal*, B, 86(4):116–128.
14. Amigó, J. Permutation Complexity in Dynamical Systems. **2010**. *Springer-Verlag*.
15. Olivares F; Souza L; Legnani W; and Rosso O. Informational Time Causal Planes: A Tool for Chaotic Map Dynamic Visualization. **2019**. *Nonlinear Systems - Volume 2*. DOI: 10.5772/intechopen.88107.

16. Shannon, C.E.; Weaver, W. The mathematical theory of communication. **1948.** *Bell Syst. Tech. J*, 27, 379–423.
17. Brissaud, JB. The meaning of entropy. **2005.** *Entropy*, 7:68-96
18. López-Ruiz R; Mancini H. L, Calbet X. A statistical measure of complexity. **1995. P***hys. Lett,* A, 209:321–326.
19. Lamberti, P; Martín, M; Plastino, A; Rosso, O. Intensive entropic non-triviality measure. **2004.** *Physica A: Statistical Mechanics and Its Applications*; 334:119-131
20. Martin, M.T.; Plastino, A.; Rosso, O.A. Generalized statistical complexity measures: Geometrical and analytical properties. **2006.** *Phys. A Stat. Mech. its Appl*, 369, 439–462.
21. Mateos, D. Medidas de Complejidad y de Información como herramientas para el análisis de series temporales. **2016.** *Aplicaciones al estudio de señales de origen electrofisiológicas*.
22. PhysioNet, the repository of freely-available medical research data, managed by the MIT Laboratory for Computational Physiology. Available online: https://physionet.org/about/database/
23. Mardia, K; Kent, J; Bibby, J. Multivariate Analysis. **1995**. *Academic Press*.
24. Clinton Sprott J. Chaos and Time-Series analysis. **2003**. *Oxford University Pres*.