ANALYSIS OF HEART RATE VARIABILITY COMPLEXITY THROUGH FRACTAL AND MULTIVARIATE APPROACHES

Roberto Sassi

Politecnico di Milano

DOTTORATO DI RICERCA IN BIOINGEGNERIA

Sede Amministrativa: Politecnico di Milano
Sedi Consorziate: Università di Genova, Padova, Pisa

ANALYSIS OF HEART RATE VARIABILITY COMPLEXITY THROUGH FRACTAL AND MULTIVARIATE APPROACHES

Tesi di Dottorato di:
Roberto Sassi
XIII Ciclo

Il Tutor: _____
(Prof. S. Cerutti)

Il Coordinatore: _____
(Prof. E. Biondi)

Anno 1997 - Anno 2000

Ph.D. Thesis in Biomedical Engineering

PUBLICATION DATA:
Analysis of Heart Rate Variability Complexity
through Fractal and Multivariate Approaches
Ph.D. Thesis - *Tesi di Dottorato*
Dottorato in Bioingegneria, XIII ciclo

Roberto Sassi
Dipartimento di Bioingegneria
Politecnico di Milano
Piazza Leonardo da Vinci, 32
20133 Milano, Italy
email: sassi@biomed.polimi.it

COLOPHON: This book has been typeset with LaTeX2ϵ and many useful tools present in the free MiKTeX distribution (www.miktex.org). The pictures on the cover illustrate the path to synchronization in a population of nonlinear coupled oscillators. It was obtained solving numerically the Kuramoto equations for the continuous case. It is a simplified model of the synchronous firing occurring in the heart pacemakers cells.

Dedico questa tesi a tutti coloro che l'hanno resa possibile.
Grazie per l'amicizia, il supporto e i consigli ricevuti.

SOMMARIO

L'analisi del segnale di variabilità cardiaca è diventata strumento sempre più impiegato nel processo di diagnosi di disturbi cardiaci. Un recente contributo, [Task Force, 1996], ha tentato di standardizzare le metodiche di utilizzo dei più comuni strumenti di indagine nel dominio del tempo e delle frequenze. Ancora molto da fare rimane, invece, nell'ambito dei parametri nonlineari.

Molte nonlinearità sono sicuramente in gioco nel sistema di controllo cardiovascolare, ma se esse producano o meno una modulazione apprezzabile (e rilevabile) nella serie RR è questione tuttora dibattuta.

Un contributo in questo senso è dato nel lavoro di tesi, dall'applicazione del metodo delle funzioni di struttura generalizzate per il calcolo dello spettro multifrattale delle serie RR. Uno spettro multifrattale decrescente è indice di correlazioni a lungo termine nel segnale e di una possibile natura nonlineare del processo generante. L'analisi è stata eseguita sui dati raccolti dalla multicentrica NOLTISALIS*; si è verificato che per tutte le* 50 *serie RR nel database, lo spettro degli esponenti è decrescente e che, in* 32 *casi su* 50*, è significativamente diverso da quello ottenuto surrogando le serie originali attraverso il metodo "amplitude adjusted Fourier transform". Si ritiene però che il basso valore degli esponenti stessi non permetta di concludere univocamente che il segnale RR sia di natura nonlineare; l'effetto di eventuali non stazionarietà, esterne al sistema, va ulteriormente investigato. Un passo in questa direzione è stata l'applicazione del metodo "wavelet transform modulus maxima" che, sfruttando le proprietà delle wavelet di eliminare andamenti polinomiali, tenta di ridurre la non stazionarietà.*

Gli esponenti multifrattali h_q si sono inoltre rivelati uno strumento molto efficace per distinguere la popolazione dei soggetti sani da quelle delle differenti patologie. Nella tesi questi parametri sono confrontati con molti altri indici, nel dominio del tempo, geometrici e monofrattali (esponente α dello spettro, parametri della "detrended fluctuation analysis"), in base alla significatività della differenza delle medie. Gli esponenti multifrattali sono risultati essere i migliori.

Al fine di sviluppare concrete applicazioni diagnostiche, diventa interessante un approccio multiparametrico al problema della classificazione di un soggetto all'interno di una serie di possibili popolazioni patologiche. Nell'analisi multiparametrica, invece di un solo indice, se ne considera un gruppo, sfruttando le caratteristiche di ciascuno e trasformando il problema da mono a multidimensionale.

Diversi tipi di classificatori sono stati sperimentati su una popolazione di 362 *registrazioni cardiotocografiche di frequenza cardiaca fetale (FHR); i* 50 *soggetti del*

database NOLTISALIS *non sono infatti numericamente sufficienti per una corretta valutazione statistica. Si è verificato che, utilizzando una rete neurale, avente in ingresso* 5 *parametri (accelerazioni grandi e piccole, Long Term Irregularity [LTI], rapporto spettrale [LF/(MF+HF)] e entropia approssimata) l'*80*% delle serie è assegnata correttamente alla popolazione di appartenenza (dati validati attraverso una "*7*-fold cross validation"; i feti sono stati distinti in "normali" e "potenzialmente patologici" in base alla diagnosi stilata alla nascita dal medico attraverso i tradizionali parametri clinici).*

L'approccio multiparametrico è complementare a quello multifrattale. Entrambi esplorano due delle direzioni significative in cui si sta muovendo la ricerca sul segnale HRV.

ABSTRACT

Heart rate variability signal analysis has became a widely employed tool in the diagnosis process of cardiovascular diseases. A recent study, [Task Force, 1996], attempted a standardization of the methods which are more widely employed, both in time and frequency domains. On the other side, it is recognized that further investigations are still necessary in the field of nonlinear parameters.

Certainly, several nonlinearities play a role in the cardiovascular control system, but it is not clear yet if they induce a significative (and detectable) modulation of the RR series.

A contribution in this direction is given in the thesis, by the computation of the multifractal spectrum of the RR signal through generalized structure functions. A decreasing multifractal spectrum indicates long term correlations in the signal and contributes an index of a possible nonlinear nature of the generating process. The analysis was performed on the data collected by the Noltisalis *multicentric research program; it was verified that the multifractal spectrum was decreasing for each of the* 50 *series in the database. In* 32 *cases it was significatively different from the spectrum of the surrogate data obtained through the amplitude adjusted Fourier transform method. The exponents were retained too small in absolute value to univocally conclude that the RR signal has a nonlinear nature. To further verify the non-stationarity influence, we also applied the wavelet-transform modulus-maxima method, which should eliminate polynomial trends from the data.*

Multifractal exponent proved to be a very effective tool in discriminating the population of healthy subjects from the four pathological ones. In the thesis, they have been compared with several other indexes: time domain, geometrical and monofractal parameters (α spectrum exponent, detrended fluctuation analysis). Multifractal exponent performed better than any other index.

To develop actual diagnostic applications, a multivariate approach to classification is also considered. In multivariate analysis a set of parameters is employed to allocate a subject in one out of several possible diagnostic groups. The classification problem becomes multidimensional.

Several supervised classifiers have been tested on a population of 362 *fetal heart rate (FHR) cardiotocographical recordings (the series in* Noltisalis *database were not numerically sufficient to ensure a correct statistical validation). The fetuses were divided in two groups: "normal" and "potentially pathological" according to the physician statement at delivery, which based on traditional clinical parameters. We verified that employing a feed-forward neural network and five input-indexes (large*

and small accelerations, Long Term Irregularity [LTI], spectral ratio [LF/(MF+HF)] and approximate entropy) 80% *of the cases were allocated to correct group (a* 7*-fold cross validation was used).*

Multivariate and multifractal approaches are complementary. Both explore two interesting directions towards which HRV research is moving.

CONTENTS

INTRODUCTION

Research on heart rate variability (HRV) started several decades ago. IN recent years, new methods, for the analysis of time series, allowed to understand new characteristics of the RR signal (being the most commonly used signal of HRV, detected as the successive values of RR interval - the instantaneous heart cycle) and, therefore, of the cardiovascular system. Simultaneously, the same numerical procedures started to be employed in the diagnostic process and their importance has always increased. Parametric spectrum analysis was, probably, one of the biggest improvements of the last two decades. As it permits an indirect but reliable measure of the autonomic activities in a very simple and elegant manner, it shares now a great popularity among researchers and physicians.

At the begin of the '80s, largely due to the introduction of methods like Grassberger and Procaccia's correlation dimension and entropy (K_2) [1983*b*; 1983*a*], the study of deterministic chaos underwent a transition from purely theoretical results to the quantitative determination of chaotic effects in experimental data. Then, by analyzing time series, coming from a large range of natural or biological systems, several researchers reported evidences of low-dimensional chaos. These studies supported the hypothesis that unpredictable patterns, shown by the signals, were produced by low-order dynamical systems and not by stochastic effects. But when, finally, a few tests (*e.g.*, surrogate data) meant to verify the consistence of non-linear indexes became available, most of the previous claims were disproved. Starting from the work of Osborne and Provenzale [1989], the scientific community developed the belief that non-linear time series methods must be trusted only when applied to controlled natural or laboratory systems for which reasonable models are known. In all the other cases, they are still important hints on the structure of the system but appropriate tests should be always performed to get a correct interpretation of the data.

The cardiovascular system is a complex mechanism, where several nonlinear mechanisms are accounted for; but the fact that a given system includes nonlinear components does not necessarily implies that the nonlinearities are

also contained in a specific signal measured from it. The quest for the nonlinear nature of the RR series started with the work of Goldberger *et al.* [1990], and a large literature developed subsequently. The "all chaotic" paradigm took over the previous "all stochastic" (or "pseudo-stochastic") belief. When the first tests became popular a bit of criticism developed around the diffused evidence of nonlinearity, previously registered. Nevertheless the work went on and more methods were proposed and applied. Many groups belonging to experimental fields like biology, chemistry and engineering contributed to these researches. Even inside the Biomedical Engineering Department in Milan, nonlinear time series analysis were performed along this direction, thus considering the hypothesis that cardiovascular regulation could be controlled both by linear and nonlinear mechanisms.

In 1996, a ESC[1] and NASPE[2] joint effort lead to a first standardization of HRV's analysis, which fixed several procedures to be used with classical time domain and frequency domain methods. The Task Force acknowledged nonlinear methods as "potential tools for HRV assessment", but stated also that advances in the interpretation of the results of these methods were necessary and highly recommended.

This thesis is meant to study the mechanisms underlying heart rate variability. The employed tools are time series methods and the analysis is conducted on RR signals. Quite a classical framework! It might be surprising, that a problem, stated so long ago and with so many published related papers, is now the topic of a Ph.D. thesis, which should be innovative by definition. The original contribution lies in its complete change of perspective. Usually, methods have been the key point; here, even if we employed several of them, the focus is on the mechanisms, which generate variability, and on their footprints, "impressed" on the RR signal.

We made just few initially physiological assumptions, not to be conditioned by prejudices. The main one was that there are many biological mechanisms, acting on HRV on different timescales, both in the cardiovascular control system (*e.g.*, vagally-mediated respiratory influences and baroceptor modulation of sinus-node activity) and in the whole body (*e.g.*, humoral factor, hormonal systems, posture, activity level, meals, sleep-wake cycle and circadian rhythm). Then we try to address two main general issues.

The first one is correlated to the investigation of the possible non-linear nature of the HRV signal, on long time scale ($\approx$ 24 hours). Is there any statistical significant evidence of a multifractal spectrum in the RR series?

[1] European Society of Cardiology
[2] North American Society of Pacing and Electrophysiology

It sounds ambitious to face such a long *querelle*, which many researchers struggled with. Certainly we did not claim we solved it; nevertheless, we think the work here presented may be a useful contribution to the topic. Besides, the new information, collected on the behaviour of the cardiovascular system, might be a useful basis for constructing a new model of the generating mechanism of the interbeat variability.

The second issue is more applicative. Omitting formal considerations, is it possible to use variability indexes with a diagnostic purpose? Among the several possible answers, we preferred to address a multivariate approach. Not a single index is used in the classification among diseased and healthy subjects, but a vector of different parameters, each one with its own peculiarity. Linear, spectral and nonlinear parameters are used together to build up a robust classifier. Then, the experimental work accomplishes two tasks: (i) the investigation on the mechanisms generating the HRV structure and (ii) starting from the evidence on the existence of nonlinear contributions, uses this information in the prediction of diseases.

The physiological bases of the work are presented in chapter 1. The cardiovascular system is briefly discussed starting from the muscular cells excitation. RR series are, also, introduced and discussed.

Chapter 2 is a basic introduction to the several methods used during the analysis. Techniques were organized, mainly, in three categories. The first one comprises classical time domain statistical methods; the distribution of the derivative signal $I^{\Delta}(i) = RR(i+1) - RR(i)$ is in depth analyzed; its long tails have connections with the probability density of some fractal set and stock market indexes. Then frequency domain techniques are discussed for the importance they hold in clinical diagnostic processes. Finally fractal-related (self-affine) approaches are introduced; in the literature a large number of these indexes is present. We tried to relate them, showing how similar their aims are. We also presented a new technique (multifractal analysis performed via generalized structure functions) previously employed in the study of fully developed turbulence. The chapter ends with a discussion of the hypothesis tests which were introduced to disprove the evidence of low-dimensional chaos in time series. Along the sections, for each methods, hypotheses and implications, related to the HRV series, were investigated; in some cases, relationships among indexes were stated, where relevant.

In chapter 3, multifractal analysis is applied to 50 different RR series, extracted from 24-hours Holter recordings, collected in the NOLTISALIS database. Nonlinear methods require, usually, very long time series; for this

reason, 24-hours Holter were selected for the analysis. Signals are separated in five different populations, corresponding to different cardiac pathological states. Initially, RR series were analyzed by means of the methods proposed in chapter 2. Then, the first question, among the two ones previously stated, was addressed and the hypothesis was validated through surrogate data tests. The discriminating ability of the multifractal exponents was also compared with the other considered techniques (on a singular index basis).

Multivariate analysis is described in chapter 4. Starting from the experience (made on HRV data and illustrated in chapter 3), we started with a second approach. If nonlinear mechanisms really contribute in the HRV behaviour, then they could improve the discrimination ability in diagnoses. Therefore, a second experimental protocol was introduced: 362 cardiotocographical exams have been collected and analyzed. The multifractal approach can not be applied (it is nearly impossible to record long enough series of this kind) therefore approximate entropy, a regularity index, became the nonlinear statistic of use. For each recording classical cardiotocographic index and spectrum components were also computed; with approximate entropy, they constitute the vector of parameters entering the classification process. The whole population of fetuses has been separated in two groups, according to the health state of the baby at delivery (as recorded in the physician diagnosis): "healthy" and "potentially pathological". Then several classification techniques (linear, quadratic and discriminant analysis, k-nearest neighbour and feed-forward neural networks) have been used to automatically allocate each exam to one of the two groups, using the information given by the parameters vector. The idea is that nonlinear and linear parameters, together, can characterize HRV much better than singularly. This would increase the number of possible diagnostic applications and the robustness of the allocations, as well.

Three appendixes of the thesis, contain more or less complex methodological material that would have distract the attention of the reader. In appendix A, Lévy stable distributions are discussed in depth; they play an important role in chapters 3 and 4. Appendix B, contains an introduction to the Kuramoto model, which is a possible mathematical model of synchronization in a group of nonlinear oscillators, globally coupled. it is a schematic explanation of what happens in the sinoatrial pacemaker as long as cells synchronously fire and give the pace to the whole muscle (chapter 1). Finally, appendix C reports a few notes on approximate entropy and sketches an original efficient algorithm for its numerical estimations (very useful on long time series in chapter 3).

The thesis work contributes to explore the complex control mechanisms of the cardiovascular system over long time scales; it also studies parameters which can be potentially effective in the actual clinical practise.

1. CARDIOVASCULAR CONTROL MECHANISMS

The heart and the blood vessels form a transportation system that delivers to all cells of the body the materials needed for their proper function and that carries away the waste products of their metabolism.

Specific function of the heart is to maintain the circulation so that an adequate supply of blood at a sufficient pressure reaches the body tissues [Milnor, 1968; Marshall, 1968]. Schematically, it is composed of four chambers made of muscular tissue with autonomous contractile capabilities. Each chamber is functionally equivalent to a mechanical pump with a two steps working cycle. At first, the cavity is stretched and filled ("diastolic phase"), then the contraction of the chamber pumps the blood out of the cavity ("systolic phase").

The two upper smaller cavities are named "atria" and their task is essentially to fill the two lower "ventricles" (the left atrium is connected through a valve to the left ventricles and, in the same way, the right atrium is connected to the right ventricles). The left ventricles supplies the pressure into the systemic circulation (nearly the whole body, ending into the right atrium) and for this reason it is the biggest chamber; the right ventricles pumps the blood though the pulmonary circulation back to the left atrium. In a healthy heart, the four chamber activities are coordinated so that the systolic contractions of the atria synchronize with the diastolic expansions of the ventricles. The *heartbeat*, easy to hear with a stethoscope, corresponds to the simultaneous contraction of the two ventricles.

Despite this naive mechanical description, such a complex organ like the heart can behave properly only if the cells activities are well directed and organized. First of all, this is necessary for the correct operation of each chamber (cells have to contract all together); then it is important for the functionality of the heart as a whole (atria and ventricles have to work coordinately to perform an efficient pumping function). To explore how synchronization develops, it is useful to explore heart cellular organization.

1.1 Cardiac Cells

Myocardial muscular cells are roughly cylindrical, typically 100 μm long and with a diameter of 15 μ; they are multinucleated and joined together at the *intercalated disk* [Keener and Sneyd, 1998]. Each disk comprises the plasma membranes of two separate cells and it is in this dense intracellular material that myofibrils terminate. Disks transect the fiber in a stepwise manner providing the mechanical adhesion among myofibrils of adjacent cells (*desmosomes* or *tight junctions*) [Marshall, 1968]. While the cellular membranes are typically separated by about 250Å, in this junction regions they are practically fused together. Electrical coupling of cells is provided by small nonselective channels, usually called *gap junction* or *electrical synapses*, that form direct intercellular connections through which ions and other small molecules can flow. Usually 20Å in diameter, they are formed by the joining of two *connexons*, which are hexagonal arrays of *connexin* protein molecules. The resistance gap junctions provide to conduction is low, compared to what would result from two cell membranes put together, but it is high if compared to the intracellular cytoplasm (the cross-sectional area for conduction is greatly reduced).

In general, the control of cell's ionic concentrations results in a potential difference across the cell membrane, causing ionic currents to flow through channels present into the membrane. Cardiac cells, as many kind of cells, use this potential as a signal, and their contraction depend on the generation and propagation of electrical signals. Cells of this sort are said to be *excitable* because by applying a sufficiently strong current, the membrane potential goes through a large excursion, called *action potential*, before eventually returning to rest.

> *The most obvious advantage of excitability is that an excitable cell either responds in full to a stimulus or not at all, and thus a stimulus of sufficient amplitude may be reliably distinguished from background noise. In this way, noise is filtered out, and a signal is reliably transmitted* [Keener and Sneyd, 1998].

How an action potential is generated? This issue was firstly addressed in 1952 by the works of A. L. Hodgkin and A. F. Huxley (with the collaboration of B. Katz) studying conduction in the squid giant axon . They developed one of the more important quantitative model in all physiology history, the so called Hodgkin-Huxley model. At rest, the potential measured inside a squid giant axon cell is negative with respect to the intracellular liquid, due mainly to the differences in the concentration of sodium and potassium

across the membrane (the equilibrium potential it is close to the K^+ Nernst potential, being other ions membrane conductances small). If small currents (from nearby cells, for example) are applied, for a short period of time, the potential return suddenly to its equilibrium after the current ceases. But if the stimulating current is large enough to rise the potential over a certain threshold, membrane permeability to Na^+ increase dramatically (voltage dependent Na^+ gate) and a fast sodium influx starts. The potential inside the cell displays a sudden increase becoming positive. In comparison with the rest condition, there's a nearly sign reversal in the potential, therefore, it is said that the cell *depolarize*. Sodium conductance then begins to decrease and drives the potential back to the equilibrium value; but at about the same time that the Na^+ current is inactivated, an outward potassium current starts and the potential decreases under the rest value and eventually return to the equilibrium situation; this second phase is called *repolarization*. After the Na^+ current is deactivated, the cell is in a *refractory* period when additional stimuli evoke no substantial response.

The qualitative behaviour of a cardiac cell is very similar to what described by Hodgkin-Huxley for the squid giant axon; at least it can give a clue on how an action potential is generated. The primary difficulties, that rise trying to develop a quantitative model of cardiac cells activities, are related to the presence of many different cell types and many different types of ionic channels (not only K^+ and Na^+ but also Cl^-, Ca^{2+} and H^+ to name but a few). As will become clear in the next paragraph, heart's cells can be collected in several families and these differences reflect on the different duration and shape of the action potentials. For these reasons quantitative models of cardiac cell structure are far from complete: some effects are retained, other discarded and some other collected into conglomerates (something like the *leakage* current into the Hodgkin-Huxley model).

Moreover, myocardial cells are contractile muscular cells: the action potential causes the cells to contract, thereby enabling the pumping functionality [Marshall, 1968; Keener and Sneyd, 1998]. This effect has to be taken into account, in a model, by considering the action of calcium currents.

1.2 Cardiac Propagation

The electrical activity of the heart starts in a collection of cells known as the *sinoatrial node* (SA node) located just below the superior vena cava on the right atrium. The cells in the SA node are *nodal* fiber: autonomous oscillators, with few myofibrils. They do not need a stimulating current to start

the depolarization and then generate an action potential. Being the region which have the greatest degree of auto-rhythmicity, it will determine the rate of beating of the entire organ. The signal then propagates by the atrial cells through the atria, which contract pumping blood into the ventricles.

Atria and ventricles are separated by a layer of non-excitable cells, that acts like an insulant, preventing electrical waves to be propagated to ventricular cells. Only at the base of the right atrium there is a pathway for an action potential to propagate across the septum to ventricles; it is a collection of nodal cells called the *atrioventricular node*[1] (AV node).

The propagation velocity in the AV node is slow compared to other cardiac cells, due to a decreased density of sodium channels; but when the action potential emerges from this gate, it is quickly delivered, to the inner surface of the ventricles, by a specialized collection of fibers called the *bundle of HIS*. The bundle of HIS is composed of *Purkinje* fibers, cells with reduced contractile capabilities and small auto-rhythmicity; the bundle divides into two branches which run down either side of the intraventricular septum and, eventually, it form a network spreading throughout the interior of the ventricles. As soon as action potentials emerge from the Purkinje fibers, they activate the contraction of the myocardial cells. The contraction wave propagates through the ventricular wall outward to the epicardial surface. Figure 1.1 shows a schematic presentation of the myocardial conduction system. In pathological conditions, several different regions of the myocardium can take over and perform, even temporarily, as the leading pacemaker.

In summary, in the heart there are three different groups of fiber: nodal fiber, like those composing the SA and AV node, Purkinje fiber and muscular tissue. Nodal cells have the capability of spontaneously excite and generate action potentials; on the other hand they have reduced contractile capabilities and small propagation velocity. Purkinje fiber can spontaneously excite too, even if this happens not frequently in normal conditions. They show small contractile capabilities but high propagation velocity of action potentials. Auto-rhythmicity is nearly absent in muscular fiber; they are characterized by good propagation velocity and riches of myofibrils[2].

[1] The AV node acts like as a reserve pacemaker and it is able to take over when the impulse generation of the SA node is suppressed (*nodal rhythms*).

[2] For quantitative models of electrical activities in the different kind of tissues see: Yanagihara *et al.* [1980] and Noble and Noble [1984] for the SA node; DiFrancesco and Noble [1985] for Purkinje fiber cells; Luo and Rudy [1991] for ventricular cells.

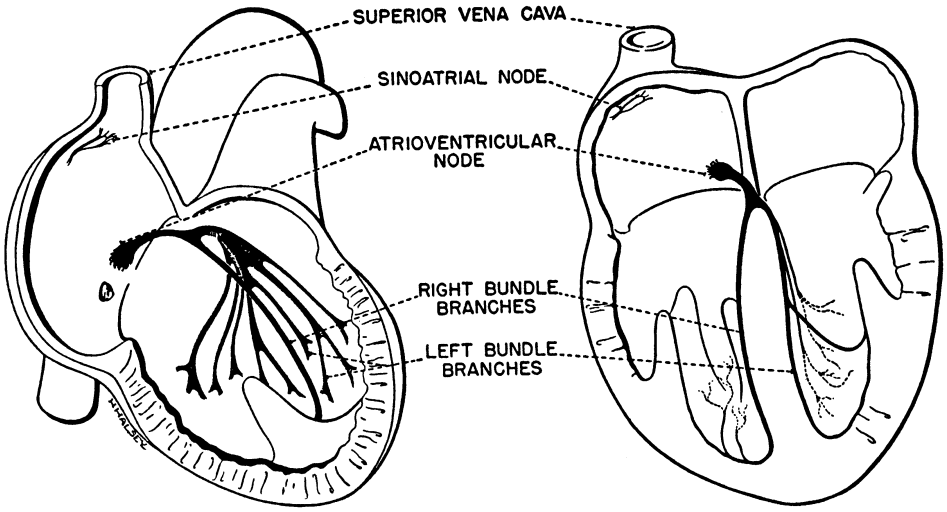


Fig. 1.1: *Schematic diagram of the cardiac conduction system [Keener and Sneyd, 1998].*

1.3 Arterial Pressure Regulation

The mechanic contraction described in the previous paragraph is spontaneous and, at least in non pathologic conditions, do not require the intervention of a control system located outside the heart. The depolarization in the sinoatrial node is "automatic" and, by itself, is sufficient to generate an heartbeat. This would lead to a nearly constant heart rate, as depolarization in the SA node happens with constant velocity. But the total amount of blood pumped by the heart (*cardiac output*), it is not constant and varies dramatically in response to the metabolic needs of the body; for example, during exercise the oxygen demand increases and the same have to do the cardiac output.

The primary nervous mechanism for the control of the cardiac output is the so called *baroceptor reflex*; it is a complex feedback scheme that include the whole circulatory system, not only the heart.

A rise in arterial pressure is detected by stretch receptors, called *baroreceptors*, located in the walls of large arteries of the systemic circulation (*carotid sinus* and *aortic arch*). This causes a signal to be sent to the central nervous system from which feedback signals are sent, through the autonomic nervous system, back to the circulatory system, eventually enabling the regulation of the arterial pressure.

The autonomic nervous system controls those functions normally ex-

cluded from the voluntary control, like the circulatory system, the gastrointestinal system and the respiratory system. It is composed of two main parts, the *sympathetic* (SS) and the *parasympathetic* (PS) nerves. The heart receives its nerve supply from both these components of the autonomic nervous system. The parasympathetic stimulations are inhibitory in their action, while the sympathetic are activating.

The parasympathetic nerve acting on the heart is the *vagus*; it innervates the nodal tissue of both SA and AV node, the bundle of His and the atrial muscle. Stimulation of the vagus nerves result in a slowing of the heart rate, a diminution in the strength of atrial contraction and a marked reduction of the conduction velocity through the AV node. The sympathetic nerves act directly on the ventricular muscle and slightly on SA, AV nodes and on the Bundle of His [Marshall, 1968]. Their stimulation results in a great increase in both rate and force of contraction of the heart and an increase in the conduction velocity through atria, AV node and ventricles.

The two autonomous nervous systems work synergically to regulate the heart activity, and the actual heart rate is much more a balance among their actions. An increase in heart rate can be the result of a stronger stimulation by the sympathetic system, a decrease activity of the parasympathetic or both. Moreover, their regulatory intervention is not relegated to special situation but it is always present; in fact both sets of nerves exhibit tonic activity under normal conditions.

Despite this brief description could lead to think so, the actions of the two collections of nerves are not exactly symmetrical; the sinoatrial node is prevalently innervated by vagal fibers, therefore, in normal conditions, the parasympathetic system has a leading action on the control of the heart activity. Moreover, vagal activity is much faster and can explicate its effects on the timescale of one beat; on the contrary, sympathetic activation is slower and requires from 2 to 4 beats to be noticed, but its effects are more persistent in time [Berne and Levy, 1996].

Experiments in which PS and SS inputs are blocked reveal that the interbeat intervals (τ), only directed by SA node firing, are very regular and average about 600 ms. Suppression of the SS stimuli can results in an increase of τ up to 1500 ms; when only the SS is active, τ can decrease to less then 300 ms [Berne and Levy, 1996].

A part from the baroceptor reflex, the nervous system activities is stimulated by a long list of other factors like elevation of carbon dioxide blood concentration, hormones release, emotional and phycological stimuli and many others; therefore, many causes influence the heart regulation.

1.4 Heart Rate Variability Signal

During the cardiac cycle, the action potentials propagation generates an electrical potential field that can be measured on the body surface. In fact along the depolarization wavefront, while the membrane potential experiences a sharp increase, the extracellular potential displays a sudden decrease. The body is a volume conductor, therefore the current (ions going in and out from cells membranes) spread throughout the body. For this reason, during heart depolarization it is possible to measure a voltage difference between any two points on the body surface (on average not bigger then 4 mV); from a certain distance, the action potential appears like a step in potential.

The *electrocardiogram* (ECG) is the continuous recording of the body-surface potential. Cardiologists have settled 12 standard leads along which recording are routinely made in hospitals every day. Then clinicians, looking at the tracing, glean information about the heart's conditions of patients.

Figure 1.2 presents the evolution of the surface potential during a heart cycle. When the action potential is spreading through the atria, the first signal, called the *P wave*, is detected. The propagation of the action potential through the wall of the ventricles, due to the big amount of muscular tissue involved, produces the largest deflection, called the *QRS complex*. Then, the repolarization of the ventricles yields to the *T wave*; atrial recovery is not detected, as it is hidden by the ventricles depolarization. Only those electrical phenomena producing enough extracellular current and involving adequate muscle mass are visible on the electrocardiogram; so SA nodal firing, AV node conduction and propagation through the Purkinje network, described in section 1.2, are not recognizable on the ECG.

In the previous paragraph, we referred to the interbeat distance τ; physiologically it correspond to the distance among two subsequent firing of the SA pacemaker, or, as it happen nearly at the same time, to the activation of the atria. But on the ECG recording the P wave, related to atrial depolarization, is difficult to detect (especially with automatic detection algorithm); for this reason, is usual to refer to the interbeat distance as to period of time between two following R peaks on the electrocardiogram (*RR interval*). The proposition holds under the hypothesis that the distance between the P and R peaks (*PR interval*) is constant and this's not a terrible approximation in normal physiological ranges.

Due to the presence of the control mechanisms, τ varies continuously and *Heart Rate Variability* (HRV) has become the conventionally accepted term to describe variations of RR intervals [Task Force, 1996]. In the last twenty years a strong relationship has been drown among cardiovascular

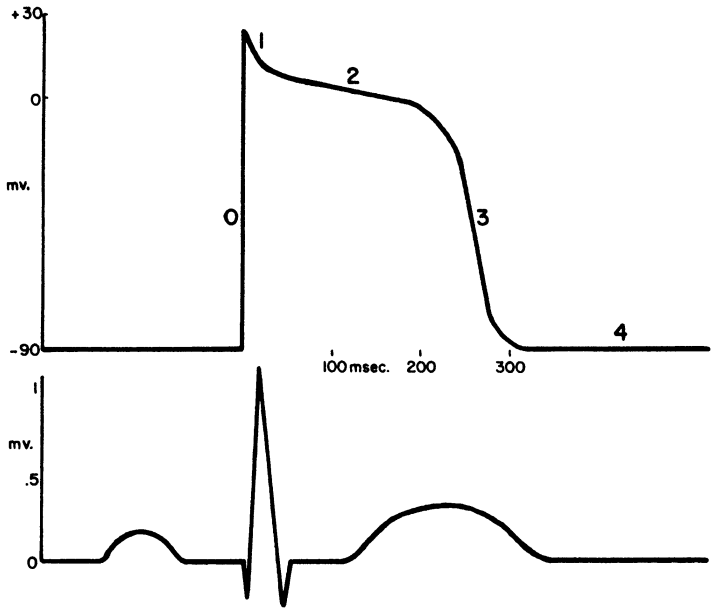


Fig. 1.2: *Cellular transmembrane potential and electrocardiogram. The upper tracing represents the transmembrane potential of a single ventricular muscular cell and the lower tracing shows the body surface potential during the same electrical event. The numbers on the upper tracing designate phases in the action potential cycle: 0: the upstroke, 1: the brief spike, 2: the plateau, 3: the rapid recovery, 4: resting potential. [Keener and Sneyd, 1998].*

mortality, including sudden cardiac death, and the autonomic control system. Clinicians and researchers are still on a quest for quantitative marker of autonomic activity: HRV represents one of the most promising.

In 1996, the European Society of Cardiology and the North American Society of Pacing and Electrophysiology collected their forces and expressed a *Task Force* with the responsibility of developing appropriate standards in HRV analysis. The document they produced, Task Force [1996], is now a fundamental reference for those working on Heart Rate Variability; this is how they described the history of HRV's clinical relevance:

> *The clinical relevance of HRV was first appreciated in 1965 when Hon and Lee noted that fetal distress was preceded by alterations in interbeat intervals before any appreciable change occurred in heart rate itself. Twenty year ago, Sayers and others focused attention on the existence of physiological rhythms imbedded in the beat-to-beat heart rate signal. During 1970s, Ewing et al devised*

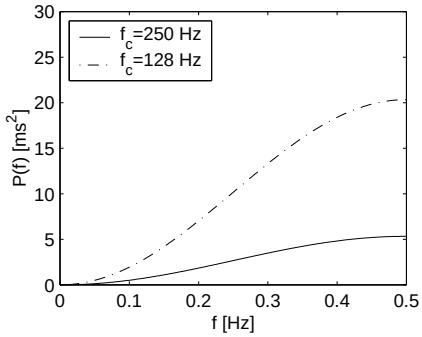


Fig. 1.3: *The finite sampling frequencies, in the ECG, implies that estimates of the interbeat intervals are affected by noise. In figure, the power spectrum of the noise t_i is plotted for two different values of the sampling frequency f_c (see the legend).*

> *a number of simple bedside tests of short-term RR differences to detect autonomic neuropathy in diabetic patients. The association of higher risk of postinfarction mortality with reduced HRV was first shown by Wolf et al in 1977. In 1981, Akselrod et al introduced power spectral analysis of heart rate fluctuations to quantitatively evaluate beat-to-beat cardiovascular control.*
>
> *These frequency domain analysis contributed to the understanding of autonomic background of RR interval fluctuations in the heart rate record. The clinical importance of HRV became appreciated in the late 1980s, when it was confirmed that HRV was a strong and independent predictor of mortality after an acute myocardial infarction.*
>
> *. . .*

1.5 Tachogram

The series of subsequent RR intervals is called *tachogram*. The sequence of interbeat distances is usually extracted from the ECG via automatic algorithms; a fiducial point is chosen on the QRS complex (normally the R-wave peak) and located with a threshold on the derivative or with template matching techniques. Only *normal* intervals are considered, that is, intervals between adjacent QRS complex resulting from sinoatrial node depolarization (presence of the P-wave).

The electrocardiogram is generally recorded with a sampling rate (f_c) in the range 128-1024 Hz and each point is then stored with a 10 or 12

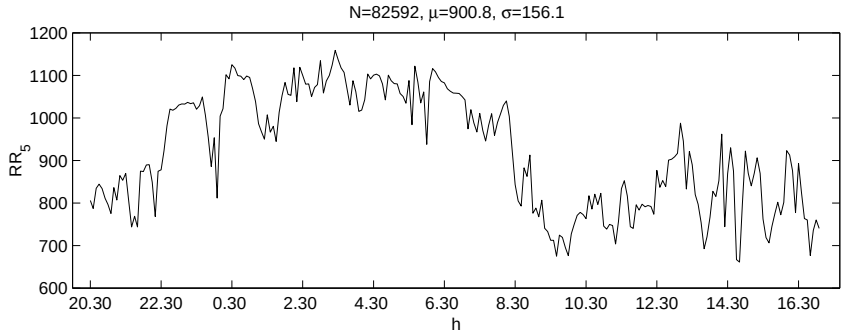


Fig. 1.4: *A tachogram, extracted from a 24-hours Holter ECG recording ($f_c = 250$ Hz) belonging to a non-pathological 38 years old subject. The signal has been averaged over 5-minutes windows. The recording started at 8.30pm and lasted nearly 21 hours.*

bits quantization; being the resolution in time limited by the sampling frequencies, on the ECG sequence, the R-wave peak is located with a precision proportional to f_c. Therefore, the finite sampling frequency of the recording device induces a quantization error on the RR series (*induced quantization error*).

Typically a sampling frequency $f_c = 250$ Hz is considered enough for the correct reconstruction of the ECG signal; on the other hand this means an uncertainty on the position of the R-wave peak of $1000/250 = 4$ milliseconds, in average. Putting this in formulae

$$RR_i = R_{i+1} + u_{i+1} - R_i - u_i = RR_i^* + t_i$$

where u_i is the δ-correlated quantization error, uniformly distributed with $E[u_i] = 0$ ms, $E[u_i^2] = (1000/f_c)^2/12$ ms^2.

The resulting error $t_i = u_{i+1} - u_i$, superimposed on the *real* RR_i interval, is *triangularly* distributed with $E[t_i] = E[u_i]$, $E[t_i^2] = 2E[u_i^2]$ and $E[t_{i+1}t_i] = -E[u_i^2]$. It is a moving average (MA) stochastic process with $H(z) = z - 1$ and power spectrum $S(f) = 4[\sin(\pi f)]^2 E[u_i^2]$ ms^2. Figure 1.3 plots the spectrum for two different value of f_c; as it can be noticed, the noise t_i is not white, but it affects mainly highest frequencies.

In most applications the jitter in the estimation of the R-wave peak can be accepted without drawbacks[3]; unfortunately with spectral analysis and many other techniques, derived from nonlinear approaches, the quantization noise can become a problem [Task Force, 1996]. In this cases, it is common to

[3] At the end of the day, the ratio, among the absolute value of the noise and the mean RR interval (≈ 800 ms), is small ($\approx 1\%$ for f_c=128 Hz).

use a sufficiently high sampling frequency (at least 250 Hz); another possibility is to refine the position of the R-wave peak with parabolic interpolation over nearby points (5 points are often a good compromise).

In figure 1.4 a tachogram extracted from a 24-hours Holter ECG recording; as expected, during the resting period the heart frequency decreases, as it is mainly influenced by vagal stimulations, and conversely the interbeat interval increases.

2. EXTRACTING INFORMATION FROM THE HRV SIGNAL

In this chapter we present several parameters proposed to analyze the heart rate variability signal (HRV). Some of them are quite widespread, like the classical time-domain statistics or the spectrum analysis, and now employed every day in the diagnostic process. Some others are still research matter, as the application of multifractal analysis, which seems promising.

The study of HRV is made to improve the physical and physiological knowledge on control mechanism of the beat generation (determined by the interactions of hemodynamic, electrophysiological and humoral factors as well as by the autonomic and central nervous regulations). Besides, certain parameters (*e.g.* approximate entropy) are considered only for their ability to summarize information and allow a comparison among different recordings. In diagnostic practice, in fact, the power of an index to discriminate a diseased patient from an healthy subject, is fundamental.

Heart rate variability analysis is usually performed on two different scales: short-term, of 2 to 5 minutes, and long-term, up to 24 hours, recordings, depending on the aim of the research. For example, central nervous regulations act on time scales of seconds, so short-term recording are preferred. On the contrary circadian rhythms modulate the firing rate over scales of hours, so a long-term window becomes necessary.

2.1 Statistical Characterization

Statistical methods have been, traditionally, used to extract information from RR series: in many cases they summarize characteristics of the signal in a single number, easy to compare and store. Nowadays they are still in a leading position, and perhaps among the most used by physicians.

Except a few cases, these methods do not take into account the temporal evolution of the sequence. They only display some aggregate information. Nevertheless them fundamental importance and usefulness concerns the preliminary screen of the data, before starting with any other more complex approach.

2.1.1 Classical Statistics

The minimal statistical characterization is achieved by computing the mean and the standard deviation ($\hat{\sigma}$ or SDNN[1]) of the entire RR series. SDNN increases with the length of the analyzed recording as the RR measures are not independent of one another[2] [Kaplan and Glass, 1995; Task Force, 1996]. Thus, it is inappropriate to compare variances obtained from sequences of different durations. Alternative measures of dispersion around the mean are the interquartile range $IQR = \xi_{3/4} - \xi_{1/4}$ (where ξ_p is the p quantile) and $MAD = \text{median}_i\{|RR_i - \mu|\}$. IQR and MAD indexes are not very efficient but much more robust than SDNN against outliers [Venables and Ripley, 1999] like spikes, very frequent in long series. At the normal $MAD \approx 0.6745\text{SDNN}$ so it is usual to refer to $MAD/0.6745$ as the MAD estimator of the standard deviation[3]. SDNN, IQR and MAD give information on the overall heart rate variability.

To neglect the influences arising from changes in heart rate due to cycles shorter than 5 minutes, it is useful to average the entire sequence over 5-minutes windows (see figure 1.4) and then to compute the standard deviation of the reduced series. The obtained index is only influenced by long-term

[1] We follow the terminology in [Task Force, 1996], where the interbeat difference is called *normal-to-normal* interval, referring to the fact that only normal intervals are considered. Thereby the suffix *NN*.

[2] It has been suggested that the heartbeat displays $1/f$ noise: for $1/f$ noise the variance increases as the length of the sequence N increases, and in the limit of an infinite series it goes to infinity.

[3] Although $\hat{\sigma}^2$ is the optimal estimator of the variance of the normal distribution, it can be substantially sub-optimal for distribution close to the normal; in fact, is any value $RR_i \to \infty$, then $\hat{\sigma} \to \infty$. Artifacts with large amplitude are not so uncommon in RR sequence before correction techniques are employed.

components of HRV and is usually called *SDANN*.

2.1.2 Density Estimation

Heart Rate Variability is a complex *signal* produced by a highly nonstationary system: a man or a woman involved in their daily activities, subjected to all the abrupt changes of rhythm that it requires. Therefore, it is very unlike the HRV probability density to be a regular and smooth function, like a Gaussian or a Gamma function. Unless it is possible to make reasonable hypothesis and to assume a particular functional form for the HRV density, all we can do is a non-parametric estimation of the distribution.

Of course, the normalized histogram is an estimator of the HRV density function, but it depends (i) on the bin width b and (ii) on the starting point of the grid of bins. Shifting the grid can result in large effects on the density [Venables and Ripley, 1999], in particular when a low sampling frequency has been used recording the ECG (128 Hz), as underlined in figure 2.1. The Task Force [1996] suggests a standard bin width of $b = \Delta_{128} = 1000/128$ ms, which is the lower precision in current commercial equipments. This recipe should permit comparisons among histograms computed on different RR series. To really fulfil the standardization it is also necessary to set a starting point for the bins grid either. We suggest, as a rule of thumb, to center the first bin on the lower RR value available in the sequence. Among all the possible starting points, this one accomplishes a good smoothness and does not overstate the maximum value of the density (see the caption of figure 2.1 for more numerical details).

If recordings were made using odd sampling frequencies (250 instead of 256 Hz for example) or bizarre rounding algorithms, even the prescription just described can not be enough to ensure smoothness. The variance of the histogram depends inversely on b, so in situations like these may be useful to enlarge the standard bin width 3 or 4 times, as illustrated in figure 2.2.

Otherwise, if smoothness is the key point, or if more robustness and resolution are necessary, a better approach is to consider the *kernel* density statistic [Venables and Ripley, 1999]:

$$\hat{f}(x) = \frac{1}{Nb} \sum_{j=1}^{N} K\left(\frac{x - RR_j}{b}\right).$$

N is the number of sample in the RR series; $K()$ is the kernel, normally chosen to be a probability density function. Common alternatives are the normal or the "cosine" function $(1 + cos\pi x)/2$ for $x \in [-1, 1]$. In this case,

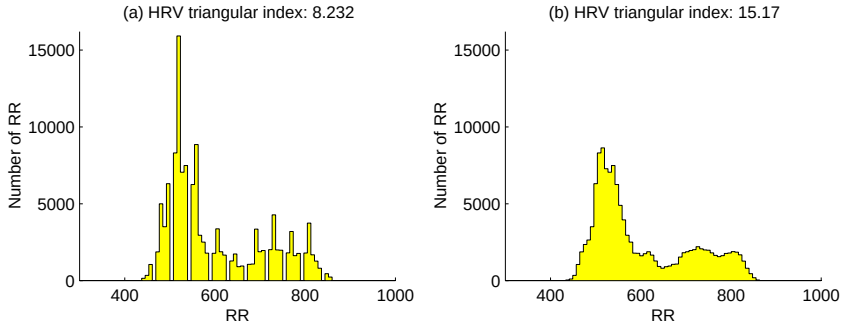


Fig. 2.1: *Histogram of the RR intervals for a 24-hours Holter series belonging to a 60 years old heart transplanted patient (f_c=128Hz). Both figure where constructed using the recommended [Task Force, 1996] bin width $\Delta_{128} = 1000/128 = 7.8125$ ms but different starting points were used for the grid of bins . In figure* (a) *the first bin starts at 297 (a total of 160 bins were used, up to 1547 ms), while in* (b) *at 297-$\Delta_{128}/2$. A naive half-bin shift produces in* (a) *unrealistic regular spike-like bumps and predicts a maximum for the density quite uncorrect, leading to a smaller HTI value. The RR intervals are usually measured on a discrete scale, to reduce storage requirement, rounding to the nearest millisecond. When, like in this case, the sampling frequency is given as a power of 2, the base-10 rounding "conflicts" with the base-2 bin width producing an irregularly spaced quantization scale. For $f_c = 128$ the sequence of distances between adjacent possible value of RR is something like: ...-8-8-8-8-7-8-8-8-8-8-7-8-8-8-8-7-.... With other f_c the scale is different but still problematic, unless the frequency used is very high (≥ 1024). The solution to produce comparable histograms is to place the first bin so that the minimum RR value, in the Holter series, lies exactly in the middle of it.*

the ideal density, which is made by a series of delta functions (one for each RR value), is filtered out by the kernel (through a proper convolution) so obtaining the smoothing.

Like with the histogram, the choice of a bandwidth value b is a compromise between different degrees of smoothing: enough to remove insignificant bumps and not too much to cancel out real peaks. A good choice for the normal kernel may be

$$\hat{b} = 1.06 \min(\hat{\sigma}, R/1.34) N^{-1/5},$$

where $\hat{\sigma}$ is the sample standard deviation and IQR the interquartile range. This is a variant of the (so called) Silverman "rule of thumb" [Venables and Ripley, 1999].

Once the density is estimated, it is possible to evaluate the HRV by measuring geometric properties of the density itself. These indexes are called *geometric*. The *Heart triangular index* (HTI) is defined as the total number of RR intervals in the series divided by the number of intervals in the modal

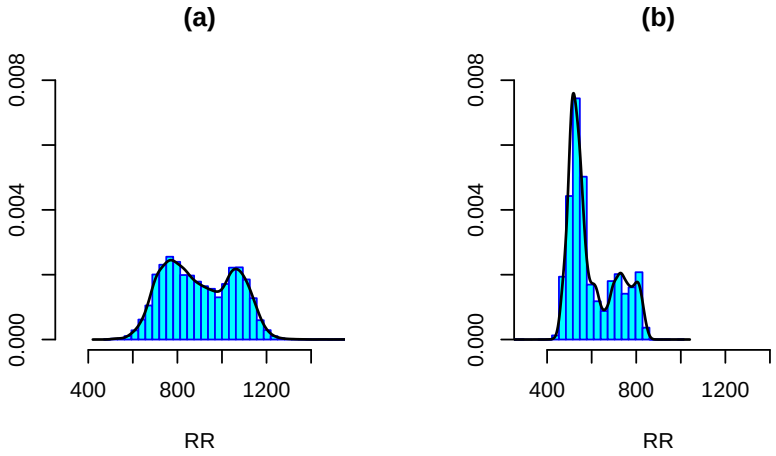


Fig. 2.2: *Probability densities estimated for two RR 24-hours Holter series. In panel (a) normal subject (see figure 1.4); panel (b) a heart transplanted patient (see figure 2.1). Histograms have been computed using a bin width of $4\Delta_{128}$ ms; the superimposed continuous lines are kernel density estimations with Gaussian kernel [Venables and Ripley, 1999]. The transplanted patient's HRV range is greatly reduced; nevertheless a day-night variability is still recognizable.*

bin of the standard histogram (using $b = \Delta_{128}$). In the same way, if a density $\hat{f}(x)$ has been constructed instead of an histogram (that is, its integral is 1), HTI=$[\max(\hat{f})\Delta_{128}]^{-1}$.

In figure 2.2 the probability density of two RR sequence, extracted from 24-hours Holter recordings, are estimated via normalized histograms and using a Gaussian kernel. The density in *panel (a)* corresponds to the series in figure 1.4; day and night activities are marked by the two distinct bumps, respectively; density in *panel (b)* corresponds to the series in figure 2.1.

During transplantation, all heart innervations from the autonomous nervous system are removed by the surgical procedure (reinnervation may occur as early as 1 to 2 years after the surgery, but it is usually of sympathetical origin). Depleted of nervous control contributions, HRV activity is very reduced, affected only by hormonal and mechanical influences. Therefore the RR range is decreased ($\approx$ 450 vs. $\approx$ 900 ms in the normal case) and most of the probability gathers around a single RR value, likely close to the sinoatrial pacemaker firing frequency (the peak at $\approx$ 500 ms). Nevertheless a weak night-day variability is still noticeable.

2.1.3 Intervals Differences

A common estimate of short-term components of HRV is the square root of the mean squared differences of successive RR intervals (*RMSSD*) [Task Force, 1996]

$$RMSSD = \sqrt{\frac{\sum_i (RR_{i+1} - RR_i)^2}{N-1}} \approx \hat{\sigma}_{I^1}.$$

It is a standard deviation estimator and it is not robust against artifacts in the series. Then, in this case, the same alternatives described for SDNN in section 2.1.1 can be employed. The signal

$$I_i^{\Delta} = RR_{i+\Delta} - RR_i,$$

where Δ is a time lag, is usually took into account because, unlike the RR series, it can be considered nearly *stationary*[4].

Peng *et al.* [1993] suggest that the probability density of I^1, considered as a random variable, can be well described by the Lévy symmetrical stable distribution

$$f_X(x, \alpha, \gamma) = \frac{1}{\pi} \int_0^{+\infty} e^{-\gamma q^{\alpha}} \cos(qx) dq. \tag{2.1}$$

Lévy stable distribution is a long-tail generalization of the normal distribution, making more likely the appearance of unusual events. It is the most general probability distribution of a sum of identically distributed random variables that looks like the distribution of each variable [Feller, 1971]. A comparison with the standard normal distribution can be drawn, by considering the Lévy distribution with $\gamma = 1/2$ and one free parameter, α (see appendix A for a brief introduction to stable distributions).

Lévy distributions are closely related to the theory of fractals [Mandelbrot, 1977, 1983]. Mantegna [1991] showed that price indices of the Milan stock exchange have statistical properties compatible with a Lévy density. Penna *et al.* [1995] found that intervals between successive drops from a leaky faucet display scale-invariant characteristics, typical of a Lévy distribution.

In their work, Peng *et al.* [1993] fit the distribution (2.1) to several I^1 series extracted form ECG recordings belonging to healthy subjects and to patients with a severe cardiac disease (dilated cardiomyopathy). Both

[4] A random process is stationary if its statistical characteristics are invariant under time shifts, that is, if they remain the same when t is replaced by $t + \Delta$, where Δ is arbitrary. The probability density, together with the mean and the variance, are independent on the absolute position of the points on the time axis.

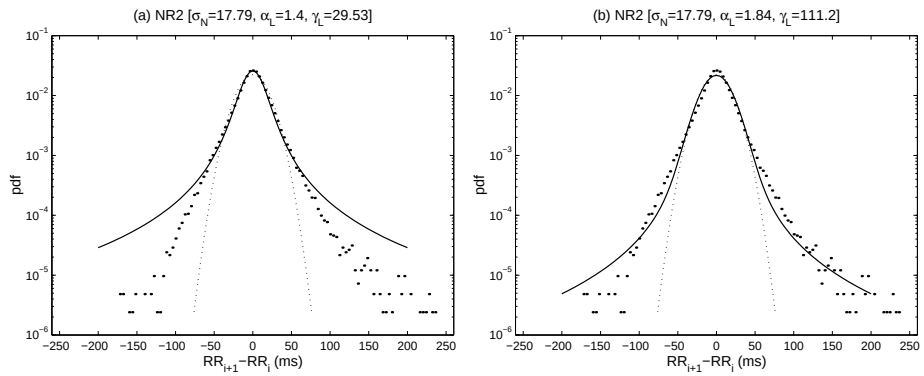


Fig. 2.3: *Probability density of $I_i^1 = RR_{i+1} - RR_i$ for a 46 years old normal subject (dots); density has been estimated via normalized histogram using a bin width of 4 ms. The continuous lines are fitted Lévy stable densities with parameters α_L and γ_L displayed in the title; the sketched lines are Gaussian distributions with zero mean and standard deviation σ_N. Fittings were performed in (a) minimizing (2.2) and in (b) minimizing (2.3); the match in (a) lacks of precision at the extremes of the tails. The standard deviation of the data increments I^1, σ_N, have been computed with the MAD estimator, to avoid outliers influence.*

populations led to a fit with $\alpha = 1.7$ (γ is a scale factor and it was considered not relevant).

In chapter 3 we perform a similar fitting to sequences belonging to a broad spectrum of different pathology, but we reach similar conclusions: no distinctions arise among different populations. The slow decay of the density for large increments values may be of physiological importance as it could be related to dynamical properties of the system that do not vary under pathological conditions.

The probability density $h_X(x)$ of I^1 for a 46 years old normal subject is presented in figure 2.3. It was estimated with a normalized histograms as described in the previous section for the RR signal. The hypothesized density (2.1), following Mantegna's work on Milan stock exchange [1991], is fitted to $h_X(x)$ in panel (a). The error sum of squares

$$\epsilon^2(\alpha, \gamma) = \sum_k [h_X(I_k^1) - f_X(I_k^1, \alpha, \gamma)]^2, \tag{2.2}$$

with $I_k^1 = bk$ (b is the histogram's bin width), was minimized in the interval $I_k^1 \in (-200, 200)$. The integration (2.1) was performed by solving the re-

lated ODE via a Runge-Kutta algorithm with automatic adjustment of the integration step.

By using this technique, the fitting is generally good on the entire interval, but sometimes it lacks in precision at the range extremes. The reason is that the values of the density function, being are so small, have the same relevance of the errors in the sum (2.2). This is exactly the case in panel (a). To overcame this problem, we suggest to minimize

$$\xi^2(\alpha,\gamma) = \sum_k \left[\frac{\log_{10}\left(\frac{bh_X(I_k^1)}{F_X(I_k^1+b/2,\alpha,\gamma)-F_X(I_k^1-b/2,\alpha,\gamma)} \right)}{1+\sqrt{|I_k^1|}} \right]^2, \qquad (2.3)$$

where F_X is the Lévy stable cumulative distribution function. The fitting is performed on a logarithmic scale, with a weighting function which slightly decrease relevance of the tails. Results are displayed for a comparison in panel (b); the fitting is more intuitive.

Evaluation of the fitting, through standard statistical test, is an interesting point; we will do a deeper investigation on this point in chapter 3.

2.1.4 Cumulative Variation Amplitude Analysis

In section 2.1.2 the most common techniques employed to estimate the probability density from the beat-to-beat intervals series have been described. It has been pointed out that the RR sequence is highly non-stationary[5], due to abrupt changes in patient's activity, and parametric estimations are far to be feasible. Recently, a method [Ivanov *et al.*, 1996, 1998; Havlin *et al.*, 1999] was proposed to overcame such difficulties and try to address a *universal* probability density for normal patients. We report it here for completeness.

The method is called *cumulative variation amplitude analysis*; it involves the sequential application of a set of algorithms based on wavelet and Hilbert transform analysis. The wavelet transform is attractive because it can eliminate local polynomial behavior in the non-stationary signal by an appropriate choice of the analyzing wavelet. The Hilbert transform also does not require stationarity. The idea is to "filter" away non-stationarity effects from the RR series it is possible to uncover intrinsic characteristics of the dynamics. Eventually the transformed series is suitable for a parametric density estimation.

[5] A typical feature of such non-stationary signals is the presence of "patchy" patterns which change over time; heterogeneous properties may be even more strongly expressed in certain cases of abnormal heart activity.

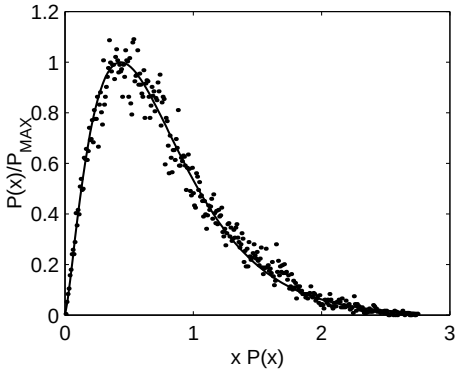


Fig. 2.4: *Probability density of heart rate variations $x \equiv A(t)$. The amplitude series was obtained with the* cumulative variation amplitude analysis *from the same RR data used in figure 1.4; here only 6 daily hours have been considered (from 12pm to 6pm). As exploring wavelet we used $\psi(\tau) = -2e^{[-\tau^2/2]}(\tau^2 - 1)/(\sqrt{3}\pi^{1/4})$ with scale $a = 8$ (second derivatives of the Gaussian function scaled to have unitary 2-norm). The continuous line is the Gamma probability density $P(x)$ that better fit the data; the parameters $\nu = 1.34$ and $b = 0.012$ are maximum likely estimates. Density have been rescaled by P_{MAX}, preserving normalization to unit area, where $P_{MAX} = b\nu^{\nu}e^{-\nu}/\Gamma(\nu + 1)$ is the Gamma distribution peak at $x_0 = b/\nu$.*

First, the wavelet transform is applied. The wavelet transform of a time series $s(t)$ is defined as

$$T_{\psi}(t, a) \equiv \frac{1}{a} \int_{-\infty}^{\infty} s(\tau)\psi\left(\frac{\tau - t}{a}\right) d\tau$$

where the analyzing wavelet ψ has a width of the order of the scale a and is centered at t. The scale a has to be chosen as a compromise among good localization (small a) and filtering needs: $a = 8$ beats is often a good choice, as it smooths locally very high-frequency variations. As analyzing wavelet, derivatives of the Gaussian function: $\psi^{(n)} \equiv d^n/d\tau^n e^{-1/2\tau^2}$ have been proposed [Ivanov *et al.*, 1996].

The second step is to extract the instantaneous variation amplitudes of the wavelet transformed signal by means of an *analytic signal* approach [Oppenheim and Schafer, 1989; Rocca, 1998]. Let $s(t)$ represent a real signal; the analytic signal is defined by $S(t) \equiv s(t) + i\tilde{s}(t)$, where $\hat{s}(t)$ is the Hilbert transform of $s(t)$ (the original real sequence with a $\pi/2$ phase shift). The instantaneous amplitude A(t) is then defined as $A(t) \equiv \sqrt{s^2(t) + \tilde{s}^2(t)}$; it can be considered, approximatively, as the envelope of the function $s(t)$.

Figure 2.4 analyzes the distributions of the beat-to-beat variation amplitudes $A(t)$ of a healthy subject during day hours (6-hours record of RR intervals, from 12pm to 6pm). The amplitudes are well fitted by the Gamma distribution function

$$P(x,b) = \frac{b^{\nu+1}}{\Gamma(\nu+1)} x^{\nu} e^{-bx}$$

(see the caption of the figure for more details).

Ivanov *et al.* [1996] claimed that the distributions of the variations, obtained with the cumulative variation amplitude analysis from RR series of normal subjects, when suitably rescaled, are described by a single Gamma distribution, which is stable over a wide range of timescales (*data collapse*). The authors attributed the functional form of the scaling observed in the healthy subjects to underlying nonlinear dynamics, essential to the normal heart function and possibly destroyed under abnormal conditions. As a first prove of this hypotheses, they found that data collapse was not present in subjects with obstructive sleep apnoea.

Ultra Low Frequencies	ULF	$f \leq 0.003$ c/b
Very Low Frequencies	VLF	$0.003 < f \leq 0.04$ c/b
Low Frequencies	LF	$0.04 < f \leq 0.15$ c/b
High Frequencies	HF	$0.15 < f \leq 0.4$ c/b

Tab. 2.1: *Standard frequency range employed in the spectrum analysis of the interbeat RR series [Task Force, 1996].*

2.2 Spectrum Analysis

Spectrum analysis (SPAN) is used to study how the power (variance) of a given signal is distributed among the frequencies spectrum. It is a classical tool in signal processing [Oppenheim and Schafer, 1989; Kay and Marple, 1981]. At first approximation the Spectral Density of a signal is proportional to the square module of its Fourier transform. Why is SPAN useful in the analysis of beat-to-beat interval series? If a mechanism (e.g., a control loop) is responsible for heart period modulation of a certain frequency, the corresponding frequency component of HRV may be used as a measure of these modulations. At least as long as the modulation is stable.

Interbeat sequences are discrete signals expressed as duration of RR segments (on the ordinates, measured in a proper time unit) versus number of progressive beat (on the abscissas). The Fourier transform of a RR series in the conjugate space is therefore measured as a function of non-dimensional units, usually called *number of cycles over beat* (c/b). But spectral analysis has been so largely employed for the analysis of signals evolving with time, that for historical reason, it is very common to express frequencies in Hertz ([Hz]=1/[s]) also with other signals; Hz unit is frequently misused with PSD of RR series too (and we'll sometimes follow this habit).

A bit of care is necessary in that situations in which, for research reasons it is necessary to compare external modulations at a given frequency (properly expressed in Hz) and the HRV spectrum. The make the comparison feasible, the PSD has to be properly rescaled by the mean RR value (in seconds). The rescaled unit is often called *Hz equivalent* (Hz eq.).

The frequencies for an RR series ranges between 0 and 1/2 c/b (and also from $-1/2$ and 0 for symmetry reasons). A standard subdivision of this range, motivated by theoretical and physiological arguments, has been defined by the Task Force [1996] and it is reported in table 2.1.

Computing the PSD of an RR series means we made some assumption on the system generating it. The interbeat series is considered the realization

of a stochastic process which is (i) stationary and (ii) ergodic in the first and second order. The energy of such a process may be infinite but its power (variance), which is the quantity of interest, is not.

The basis of spectrum analysis are provided by the autocorrelation function, defined for a stochastic process as

$$\mathcal{R}(\tau) = E[RR_{t+\tau}RR_t].$$

The Wiener-Khinchin theorem state that the Power Spectral Density $\mathcal{P}(f)$ is the Fourier transform of $\mathcal{R}(\tau)$; that is

$$\mathcal{P}(f) \equiv \mathscr{F}\{\mathcal{R}(\tau)\} \qquad \mathcal{R}(\tau) \equiv \mathscr{F}^{-1}\{\mathcal{P}(f)\}.$$

As long as we do not have an infinite amount of data, PSD has to be properly estimated. Several technique have been proposed, but they can be generally classified into two categories: non-parametric and parametric ones. The former do not require the formulation of a specific generating model for the signal but they need a longer series to be accurate. The second ones are feasible only assuming the underling model as stationary and therefore cannot deal with long-time recordings. A good review of spectrum analysis techniques can be found in [Kay and Marple, 1981].

2.2.1 Non-Parametric Estimation

Non-parametric estimations are based on Fourier transform operations. Historically two main techniques have been developed: the direct approach via discrete Fourier transform of the data, usually called *periodogram*, and the indirect approach via autocorrelation estimate. Only the former will be described in the remain of the section. Indirect methods are numerically less efficient.

The periodogram is given by

$$\hat{\mathcal{P}}(k\Delta_f) = \frac{T_\mu}{N}\left|\sum_{n=0}^{N-1} RR^0_{n+1}e^{-j2\pi kn/N}\right|^2 \equiv \frac{T_\mu}{N}\left|\mathrm{DFT}(RR^0)\right|^2 \tag{2.4}$$

where $\Delta_f = 1/[NT_\mu]$, $k = 0, \cdots, N-1$, $T_\mu = 1$ ($T_\mu = \mathrm{mean}(RR)$ seconds, if the frequency axis is expressed in Hz eq.) and $RR^0 = RR - \mathrm{mean}(RR)$. Δ_f is the frequency resolution so that f ranges from 0 to $1/T_\mu$. The RR series

has been subtracted by its mean value, therefore the integral of the PSD is equal to the variance of the signal, that is

$$\mathrm{var}(RR) = \sum_{k=0}^{N-1} \hat{\mathcal{P}}(k\Delta_f)\Delta_f \equiv \frac{N-1}{N}\hat{\sigma}^2, \tag{2.5}$$

where $\hat{\sigma}$ is the unbiased estimator of the standard deviation.

When the length of the RR sequence, N, is a power of two, a high-speed radix-2 fast Fourier transform algorithm can be employed (this is the main reason why the periodogram is so widely used). Otherwise one can compute the mixed-radix discrete Fourier transforms: sequences whose lengths have many small prime factors are processed quickly. Those that have few are not[6] [Press *et al.*, 1992]. Mixed-radix algorithms are especially useful with very long sequence when reducing the lengths to the nearest power of two would result in a big loss of data.

A discrete frequency spectrum implies periodicity in time. By using the periodogram as defined by (2.4), the RR series is implicitly considered periodic, that is, after R_N we suppose to have another R_1 and so on. On the contrary, if zero values are more realistic boundary conditions than periodicity, PSD have to be computed on the series QQ^0, which is obtained by padding RR^0 with at least $N-1$ zeros, to avoid circular convolution effects. Actually, more than $N-1$ zeros are usually added to make the new length of the series, M, an exact power of 2. *Caveat*: (i) when assuming periodicity, *zero padding* must be absolutely avoided; (ii) when zero padding is necessary, even if the new series length is M, PSD in (2.4) has still to be normalized by N to comply with (2.5)[7].

The Periodogram's inverse discrete Fourier transform (IDFT) is equiva-

[6] When the sequence length is not an exact power of two, the prime factors of the sequence length are found and the mixed-radix discrete Fourier transforms of the shorter sequences is computed. The computational load is roughly proportional to $N\sum_i p_i$ where p_i are the prime factors composing N. Series with power of length are processed quickly, but even power of 3 and 5 ones perform very well.

[7] On the contrary Δ_f is affected by zero padding, shrinking from $1/[N\Delta_t]$ to $1/[M\Delta_t]$. However actual frequency resolution is definitely not improved by zero padding. The extra $M-N$ points in the Fourier space are obtained by interpolation between the N anyway existing.

lent to the biased autocorrelation estimator[8]

$$\hat{\mathcal{R}}(m) = \frac{1}{N} \sum_{n=1}^{N-m} RR_{n+m} RR_n \equiv \mathrm{DFT}[\hat{\mathcal{P}}(k)], \tag{2.6}$$

(the formula is correct in case of zero values boundary conditions. With periodic boundary conditions it needs to be properly adjusted, but the relation with the IDFT of the periodogram still holds).

The periodogram estimator offers a high spectral resolution but, as a drawback, the PSD is definitely not smooth. Decrease in the variance of the PSD estimator can be obtained by multiplying the autocorrelation function by a proper window, that is convolving the periodogram with a smoothing filter. Filters based on the Bartlett triangular window $w_B(\tau)$ (zero for $\tau \geq \tau_{MAX}$) and the Hanning window $w_H(\tau) = (1/2)[1 + \cos(\pi\tau/\tau_{MAX})]$ are commonly used. See [Rocca, 1998; Oppenheim and Schafer, 1989] for more details.

The periodogram approach to PSD estimation is (i) computationally efficient and (ii) it makes use of simply algorithms. Main drawbacks are: (iii) suppression of weak periodic signal and (iv) frequency resolution limited by the available data length. For more details see [Kay and Marple, 1981; Task Force, 1996]

2.2.2 Parametric Estimation

Spectral analysis, based on a parametric approach, becomes a three step procedure. At first a proper time series model has to be selected; then the parameters of the chosen model need to be estimated from data. Finally the spectral estimate is obtained by substituting the computed parameters into the theoretical PSD implied by the model.

Various families of models can be suitably chosen to perform a parametric spectral estimation on the RR tachogram. Among these, autoregressive models (AR) are generally preferred [Kay and Marple, 1981; Zetterberg, 1969; Baselli *et al.*, 1997] at least for two reason: (i) each MA and ARMA stationary model, with finite variance, is equivalent to an AR model of proper order (at most infinite) [Bittanti, 1981]; (ii) parameters of AR models can be estimated by solving convenient linear equations.

[8] The biased autocorrelation estimator is usually preferred to the unbiased one $\hat{\mathcal{R}}'(m) = (N/N-m)\hat{\mathcal{R}}(m)$ since it tends to have less mean square error.

Auto-regressive are purely non-deterministic processes. They are obtained by filtering a white gaussian noise through discrete dynamical systems with asymptotically stable rational transfer functions [Bittanti, 1990]. Thus, an AR process, of order M, is the only stationary solution of the stable equation

$$RR_n^0 = -\sum_{i=1}^{M} a_i RR_{n-i}^0 + w_n, \tag{2.7}$$

where $w_n \stackrel{d}{=} WGN(0, \sigma_G^2)$ is the white Gaussian noise. $RR^0 = RR - \text{mean}(RR)$ is the RR sequence subtracted by its mean value.

The first step in AR model identification is the estimation of the autocorrelation function, usually through the biased estimator reported in equation (2.6). Model parameters estimation is achieved solving recursively the Jule-Walker equation

$$\hat{R}(\tau) = \begin{cases} -\sum_{i=1}^{M} a_i \hat{R}(\tau - i) & \text{for } \tau > 0 \\ -\sum_{i=1}^{M} a_i \hat{R}(\tau - i) + \sigma_G^2 & \text{for } \tau = 0; \end{cases} \tag{2.8}$$

via the Levinson-Durbin algorithm[9]

$$\begin{aligned} a_{1,1} &= -\hat{R}(1)/\hat{R}(0) \\ \sigma_1^2 &= (1 - |a_{1,1}|^2)\hat{R}(0) \\ &\cdots \\ a_{k,k} &= -\frac{1}{\sigma_{k-1}^2}\left(\hat{R}(k) + \sum_{i=1}^{k-1} a_{k-1,i}\hat{R}(k-i)\right) \\ a_{k,j} &= a_{k-1,j} + a_{k,k}a_{k-1,k-j}^* \\ \sigma_k^2 &= (1 - |a_{k,k}|^2)\sigma_{k-1}^2 \end{aligned}$$

where $a_{k,j}$ is the j^{th} coefficient at level k (model of order k). σ_k^2 is the variance of the one-step forward prediction error

$$\xi_n^k = RR_n^0 - \hat{RR}_n^0 = RR_n^0 + \sum_{i=1}^{k} a_i RR_{n-i}^0.$$

The choice of the model order, for RR series, usually have to fulfill the three conditions: (i) $M \in [8, 20]$ to avoid both under and over-fitting of the

[9] Alternatively one can solve directly the first equations in (2.8), for $\tau = 1, 2, \cdots, M$, employing the symmetry of the autocorrelation function $R(\tau) = R^*(-\tau)$; then by plugging the model coefficients a_i in the second equation of (2.8) the standard deviation σ_G of the white noise entering the model can be computed.

model to the data; (ii) $M \leq M_{AIC}$ where M_{AIC} is the order which minimizes the Akaike information criterium [Bittanti, 1981; Akaike, 1974]

$$AIC(k) = \frac{2k}{N} + \log \sigma_k^2;$$

(iii) the prediction error ξ^M satisfies a whiteness test, like the Anderson[10] test, the Box and Pierce [1970] *portmanteau* test[11] or the Ljung and Box [1978] Q^* statistics[12].

Power spectral density for an AR process is given by

$$\text{PSD}(f) = \frac{\sigma_G^2 T_\mu}{|1 + \sum_{k=1}^{M} a_k e^{-j2\pi k f T_\mu}|^2} = \frac{\sigma_G^2 T_\mu}{A(e^{j2\pi f T_\mu}) A^*(e^{j2\pi f T_\mu})}, \qquad (2.9)$$

[10] Anderson test is performed on the biased estimator of the prediction error's autocorrelation function

$$\hat{R}_\xi(\tau) = \frac{1}{N-M} \sum_{i=1}^{N-M-\tau} \xi i + \tau \xi_i.$$

Under the hypothesis that ξ is a white Gaussian noise $\stackrel{d}{=} WGN(0, \sigma_\xi^2)$, the statistics $\hat{R}_\xi(\tau)$, for $\tau \neq 0$, is normally distributed with zero mean and variance $E[\hat{R}_\xi^2(\tau)] = [(N-M-\tau)/(N-M)^2]\sigma_\xi^4 = \sigma_{\hat{R}}^2(\tau)$. Defined a significance level α (usually 5%), let's be $\mathcal{O}$ the number of $\hat{R}(\tau)$ values, for $1 \leq \tau \leq \bar{\tau} \leq N-M-1$, outside the α% confidence limit of the autocorrelation estimator, that is

$$|\hat{R}_\xi(\tau)| > \frac{\sqrt{N-M-\tau}}{N-M} \Phi\left(1 - \frac{\alpha}{2}\right) \sigma^2,$$

where $\Phi(\cdot)$ is the normal cumulative distribution function. If $\mathcal{O} < \alpha\bar{\tau}$, then can not be rejected the null hypothesis that the prediction error ξ is a white Gaussian noise at the significance level α. $\bar{\tau}$ is usually set to $\approx N/10$.

[11] The Box & Pierce statistics is defined as

$$Q_{\bar{\tau}} = (N-M) \sum_{\tau=1}^{\bar{\tau}} \hat{R}_\xi(\tau)$$

(see note (10) for details on $\bar{\tau}$ and $\hat{R}_\xi(\tau)$). Under the null hypothesis that $\hat{R}_\xi(\tau) = 0$ $\forall \tau \neq 0$, $Q_{\bar{\tau}}$ has a χ^2 distribution with $\bar{\tau}$ degrees of freedom [Venables and Ripley, 1999]. See also [Box *et al.*, 1994].

[12] The Ljung & Box statistics is defined as

$$Q_{\bar{\tau}}^* = (N-M)(N-M+2) \sum_{\tau=1}^{\bar{\tau}} \frac{\hat{R}_\xi(\tau)}{N-M-\tau}$$

(see note (10) for details on $\bar{\tau}$ and $\hat{R}_\xi(\tau)$). Under the null hypothesis that $\hat{R}_\xi(\tau) = 0$ $\forall \tau \neq 0$, $Q_{\bar{\tau}}^*$ has a χ^2 distribution with $\bar{\tau}$ degrees of freedom. See also [Box *et al.*, 1994].

where $T_\mu = 1$ ($T_\mu = \text{mean}(RR)$ if frequencies are expressed in Hz eq.) and $A(z)$ is the zeta transform of the transfer function of model (2.7).

The total power σ^2_{RR} can be expressed [Baselli *et al.*, 1997] by

$$\sigma^2_{RR} = \int_{-1/[2T_\mu]}^{1/[2T_\mu]} \frac{\sigma^2_G T_\mu}{A(e^{j2\pi f T_\mu})A(e^{-j2\pi f T_\mu})} df = \frac{\sigma^2_G}{j2\pi} \oint_{|z|=1} \frac{z^{-1}}{A(z)A(z^{-1})} dz,$$

with the change of variable $z = \exp(j2\pi f T_\mu)$. The classical residual theorem states that the last integral is equivalent to the sum of the residuals γ_k, one for each pole p_k zeros of $H(z) = A(z)A(z^{-1})$ with $|p_k| < 1$. That is

$$\sigma^2_{RR} = \sigma^2_G \sum_{k=1}^{M} \gamma_k = \sigma^2_G \sum_{k=1}^{M} \frac{1}{p_k H'(p_k)} = \sigma^2_G \sum_{k=1}^{M} \frac{1}{\prod_{i=1\neq k}^{M}(1 - p_i p_k^{-1}) \prod_{i=1}^{M}(1 - p_i p_k)}.$$

Residuals γ_k relative to a real poles are real, and residuals of complex conjugated poles are complex conjugated as well. In the same way the PSD (2.9) can be divided in contribution relative to each pole:

$$\mathrm{S}_k(f) = T_\mu \sigma^2_G \left(\frac{\gamma_k p_k}{e^{-j2\pi f T_\mu} - p_k} + \gamma_k + \frac{\gamma_k p_k}{e^{j2\pi f T_\mu} - p_k} \right), \tag{2.10}$$

and

$$\mathrm{PSD}(f) = \sum_{k=1}^{M} \mathrm{S}_k(f).$$

The frequency $f_{p_k} = \measuredangle(p_k)/(2\pi T_\mu)$ of p_k allows to easily allocate each pole, and its associate power $\sigma^2_G \gamma_k$, to ones of the frequency band of table 2.1. Moreover f_{p_k} can be interpreted as the central frequency of the HRV modulating mechanism. *Caveat*: the power $\sigma^2_G \gamma_k$ is real (for two complex conjugated poles it is real the sum $\sigma^2_G[\gamma_k + \gamma_k^*]$) but it is not necessary a positive value[13].

2.2.3 Remarks

Spectrum analysis requires stationary of the time series under study; on the contrary the interbeat series is a highly non-stationary signal, continuously

[13] Usually negative power components are considered an index of model over-fitting to the data (too a high order N). With an automated procedure, when negative power components can lead to unexpected results, the power in each spectral band can be more safely computed through direct integration of the PSD,

$$\int_{f_a}^{f_b} \mathrm{S}_k(f) df = -\Delta \gamma_k \sigma^2 (f_b - f_a) - \frac{j\gamma_k \sigma^2}{2\pi} \log \left[\frac{p_k - e^{j2\pi f_b \Delta}}{p_k - e^{-j2\pi f_b \Delta}} \cdot \frac{p_k - e^{-2j\pi f_a \Delta}}{p_k - e^{j2\pi f_a \Delta}} \right].$$

modified by the action of neural control system and hormonal releases. The physiological mechanisms responsible for LF and HF power components can not be considered stationary on time periods longer then 2-5 minutes [Task Force, 1996]. Therefor, usually spectral analysis is performed on two different timescales, according to the investigatory motivations of the study.

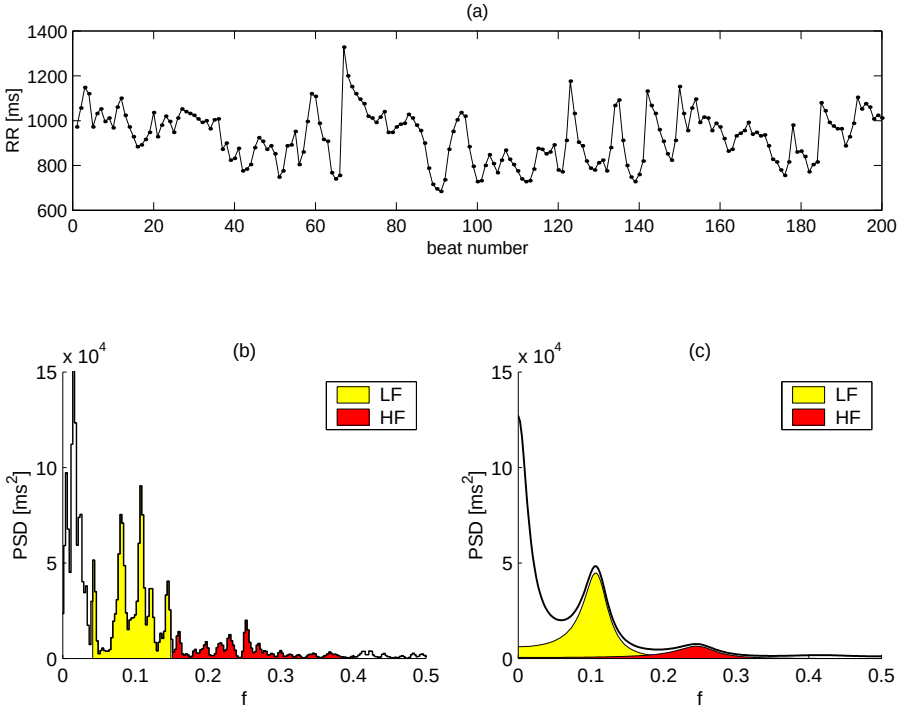
Short term recordings of 2-5 minutes analyzes power components in the LF and HF bands[14]. Commonly used indexes are the LF/HF ratio and the normalized powers: LF norm [LF/(σ^2_{RR}-VLF)] and HF norm [HF/(σ^2_{RR}-VLF)]. The efferent vagal activity is a major contributor to the HF component, while LF components are considered a marker of sympathetic modulation [Pagani *et al.*, 1986; Lombardi *et al.*, 1987; Pagani *et al.*, 1991; Task Force, 1996]. The LF/HF ratio characterizes the autonomic regulation of the heart period, and often termed "symphato-vagal balance" [Malliani *et al.*, 1991]. The physiological explanation of the VLF component is much less defined. It is commonly accepted to consider its major constituent nonharmonic components, without coherent properties and affected by algorithms of baseline and trend removal [Task Force, 1996]. On 5-minutes sequence long term regulatory mechanism are detected as noise or small trends.

Parametric techniques are basically employed with short term SPAN. The interbeat series can be considered stationary on such short timescales, moreover AR spectrum components can be easily allocated to each frequency band. In figure 2.5, the spectrum is computed with both parametric and non-parametric techniques on a short 3 minutes RR series, extracted from a 24-hours Holter recording.

Longer time series (usually up to 24-hours) can be analyzed to explore the power contents in VLF and ULF bands[15]. Only non-parametric technique can be applied. VLF and ULF physiological correlates are still unknown [Task Force, 1996]; they can be searched in the long term regulatory mechanisms of the organism like thermoregulation and humoral factors. An interesting approach is the examination of the PSD's slope at the lower frequencies (see section 2.3.1).

[14] Hereafter we'll use LF, HF, VLF and ULF to name both the frequency band and the spectral powers [ms^2] contained into that band

[15] This analysis provides only averages measure of the modulations attributable to LF and HF components



(d)

	PERIODOGRAM			AR MODEL		
	power [ms^2]	**f$_C$**	**power** [nu]	**power** [ms^2]	**f$_C$**	**power** [nu]
VLF	5174.7	0	-	5972.9	0	-
LF	6047.4	0.096	73.5%	5641.1	0.109	75.9%
HF	1866.7	0.243	22.7%	1320.4	0.249	17.8%

Fig. 2.5: *Spectral analysis was performed on a short 3 minutes RR series extracted from a 24-hours Holter recording of a 43 years old healthy subject (diurnal activities); the* 200 *RR values are displayed in panel (a). In panel (b), PSD estimate obtained with the periodogram and slightly smoothed with a Bartlett window ($\tau_{MAX} = 400$); the* 200 *original points have been padded with* 200 *extra zeros. Panel (c) reports the PSD computed with a parametric AR model (order $M = 8$, complying with Anderson whiteness test). In both the panels, yellow area is the power associated with LF; red with HF. In panel (d) some numeric values: "nu" are normalized unit (LF norm and HF norm); f_C is the central (barycenter) frequency in each band (this frequency can be related to the underling physiological mechanism). The total power is $\sigma^2_{RR} = 13401.3$ ms^2; $LF/HF = 3.24$ with the periodogram and $LF/HF = 4.27$ with the parametric estimation.*

2.3 Fractal and Multifractal Analysis

In 1975, Mandelbrot coined the term "fractal", which up to now has been used more informally to describe a basic concept, rather than being defined in a mathematical rigorous way. Roughly speaking, a fractal entity is characterized by the inherent, ubiquitous occurrence of irregularities which governs its shape and complexity. Quoting Mandelbrot,

> *I coined the term fractal from the Latin fractus, which describes the appearance of a broken stone: irregular and fragmented. Etymology cannot force an actual stone's surface to be fractal, but it is surely not a standard surface, and it should be fractal if it is scaling. The science of wear and of friction* [...] *supports the belief that fractional Brown surfaces provide first approximation representations* [...] *for many natural surfaces* [Mandelbrot, 1983].

It has become generally accepted that the theory of fractals is certainly more suitable for a comprehensive description of the physical world than many other theories which mainly handle completely regular phenomena. The simplest fractal sets are characterized by some form of self-similarity, in which parts, when magnified by a constant r, appear similar to the original whole. The more general class of fractals are really multi-scale fractals, or multifractals, which are characterized by multiple subdivisions of the original into N objects, each magnified by a different factor r_i, with i=1,2,...,N.

The best known fractal process is the Brownian motion which can be constructed through a simple iteration, a property shared by many fractal objects which can be studied analytically.

2.3.1 $1/f$ Noise and Scaling Exponent

In 1982, Kobayashy and Musha [1982], analyzing a 10 hours RR series recorded from an awake and at rest healthy subject, found that the power spectrum was displaying a "power law" behaviour for time scales from few minutes to hours; The PSD at frequency f was proportional to $1/f^{\alpha}$ with $\alpha > 0$. This preliminary observation was then confirmed by Saul *et al.* [1987] that reported the $1/f$ relationship in 24-hours recording during ordinary daily life in a larger population of subjects. It is found empirically that in heart rate $\alpha \approx 1$ as firstly reported by Kobayashy & Musha.

The value of the α exponent is usually computed by linear regression technique in a log-log plane. In the frequencies range $(0, 0.04]$ Hz [Task

Force, 1996], the model $\log \mathrm{PSD}(f) = C - \alpha \log f$ is fitted on the power spectrum $\mathrm{PSD}(f)$ obtained via FFT. The goodness-of-fit can be assessed by R^2 statistics. This technique, though it is very elementary, is one of the most reliable [Pilgram and Kaplan, 1998], especially when the number of samples is big enough ($N > 1000$).

The study of the power spectrum slope has gained popularity after Bigger *et al.* [1996] proved it has clinical relevance. α seems to be an excellent predictor of patient mortality, superior to other traditional power spectral parameters quantifying the RR variability.

Generically a power law PSD form is called *one-over-f noise*; it is the hallmark of long term correlations in the underling signal and it has been observed in many types of time series from physical, biological, physiological, economic, technological and sociological systems. With $1/f$ noise, there exists no well-defined temporal scale for the correlation time; this implies that the current value of the heart rate co-varies not only with its most recent values but also with its long-term history in a scale-invariant manner.

Stationary stochastic processes displaying power law PSD of the form $\mathrm{PSD}(f) = Cf^{-\alpha}$ are also called *scaling* [Mandelbrot, 1983] or *self-affine* [Mandelbrot, 1983; Osborne and Provenzale, 1989]. For a self-affine signals X_n, by definition,

$$\Delta^{-H}[X_{n+\Delta} - X_n] \stackrel{d}{=} [X_{n+1} - X_n] \tag{2.11}$$

independently on time ($\stackrel{d}{=}$ means equality in the sense of distributions). $H \in [0, 1]$ is the *scaling exponent*. From a practical point of view, self-affinity means that if the time scale is rescaled by a factor Δ and the signal itself is rescaled by a factor Δ^{-H}, then the transformed time series has the same statistical properties as the original one. Self-affine signal are truly fractal curves (in the continuous case they are everywhere continuous and non-differentiable). Examples are the classical Brownian motion (Bm) [Einstein, 1985; Mandelbrot, 1983], for which $H = 1/2$ and the *fractional Brownian motion* (fBm) introduced by Mandelbrot and Van Ness [Mandelbrot, 1983; Fabani and Sassi, 1996] motivated by the analysis of Nile river annual discharges ($0 \leq H \leq 1$). In the classical case $H = 1/2$ the correlation $R(\tau)$ vanishes for $\tau \neq 0$ as expected (in Bm, successive increments are independent and identically distributed Gaussian random variables). With fBm, for $H > 1/2$ the correlation is positive, expressing persistence, and it becomes 1 when $H = 1$. For $H < 1/2$, the correlation is negative, expressing anti-persistence, and it becomes $-1/2$ when $H = 0$.

There is a close theoretical relationship between self-affine signals and their power spectra. Remember that a *structure function* of the random process X_n is defined as

$$\mathrm{SF}(\Delta) \equiv \langle |X_{n+\Delta} - X_n|^2 \rangle; \tag{2.12}$$

the structure function of a self-affine process scales as $\mathrm{SF}(\Delta) \sim \Delta^{2H}$. A classical argument relates the structure function of a stationary process with its spectrum and indicates that a self-affine process has a power law spectrum $\mathrm{PSD}(f) = Cf^{-\alpha}$ with $\alpha = 2H+1$ [Osborne and Provenzale, 1989; Provenzale *et al.*, 1991].

The fractal dimension of the fractal curve can be computed by the equation

$$D = \frac{2}{\alpha - 1};$$

this relationship is valid only for $0 < H < 1$, therefore for $1 < \alpha < 3$. For $\alpha > 3$ the curve ceases to be fractal (its Hausdorff dimension equals its topological dimension) and for $0 \leq \alpha \leq 1$ the scaling exponent is zero and the fractal dimension is infinite (as expected for white noise when $\alpha = 0$).

Besides the spectrum regression technique, it is possible to analyze the self-affine properties of the RR series by computing directly the scaling exponent H and then inferring the parameter α. From (2.11), by supposing the signal self-affine, then

$$\langle |RR_{n+\Delta} - RR_n| \rangle = \Delta^H \langle |RR_{n+1} - RR_n| \rangle;$$

thus a graph of $\langle |RR_{n+\Delta} - RR_n| \rangle$ versus Δ on a log-log plot is a straight line whose slope is the value of H.

The scaling exponent H and the Hurst exponent, though related, assume different values. The Hurst exponent is usually computed with the *rescaled range* R/S statistics. Only skipping the integration step, at the beginning of the computation of R/S, the Hurst exponent and H are identical (in the assumption that the RR series can be modelled as fBM) [Mandelbrot, 1983, chap. 39].

In the dynamical system theory, a strange attractor with fractal dimension $D > 1$ is an object which is differentiable along the direction of motion and which can be fractal in some direction perpendicular to the direction of motion. Are the fractal evidence of the RR series related to an underling low dimensional dynamical system? Starting from a similar argument, a lot of work has been made in the recent years to assess if the interbeat variability can be modelled as produced by a low dimensional dynamical system. This

is certainly still an open question (see chapter 5 for a proper discussion); but the $1/f$ power spectrum is not a sufficient condition to state that a low dimension dynamical system is a good model for the system. In fact, as we realized in this section, proper self-affine colored noises display power law spectral behaviour and a finite fractal dimension.

Generating Mechanisms

It is widely accepted that the HRV signal displays a $1/f$ power spectrum, at the lower frequencies. But what is the mechanism which is able to generate such a widespread behaviour? Two main groups of hypothesis have been formulated [Rienzo *et al.*, 1997]. Some researchers think that the $1/f$ trend is a result of an underlying $1/f$ modulation of a number of factors affecting blood pressure and heart rate variability. Some of this factors are:

- fluctuations in the electric potentials of the pacemaker cell membrane at the level of the sinus node [Kobayashy and Musha, 1982];
- modulation induced by the renin-angiotensin system (that regulates extracellular fluid volume) [Calcagnini *et al.*, 1995];
- fluctuation at the organ level, including oscillations in metabolism, changes in respiration and body temperature, changes in the state of the autonomic neural centers;
- physical activities;
- modulation in the baroreflex sensitivity.

"The second hypothesis on the nature of the $1/f$ trend may be regarded as a sort of "general law" which can apply also to fields other then cardiology. According to this hypothesis the observed $1/f$ power law is the result of the complex interaction between different precesses characterized by a variety of time scales and simultaneously acting on the same system. This hypothesis actually fits the reality of the cardiovascular system which is under the concomitant action of many control mechanism with time scales that ranges from seconds to several hours" [Rienzo *et al.*, 1997]. Two recent works deserve to be cited.

Hausdorff and Peng [1996] prepared a stochastic process made by simply summing up 8 independent inputs x_i. At each time step the state of x_i changes with probability $1/\tau_1$, where τ_1 is the time constant of the input. The time constants were chosen to be consistent with the physiological

mechanisms, which are relevant to HRV variability modulation (e.g. respiration, blood pressure). The model can mimic $1/f$ behaviour for about 4 decades if the amplitude of each noise input is identical; otherwise if each input noise is allowed to vary, $1/f$ scaling is no longer obtained consistently. The study suggests that the balance between different noise inputs is also crucial to produce self-similar behaviour.

Pilgram and Kaplan [1999] tried a similar approach. They concatenated short segments of Gaussian white noise ($\alpha = 0$) and Brownian motion ($\alpha = 2$), selecting randomly among the two typology. In addition, they superimposed to each segment a linear trend with a random direction (matched at the endpoint), accounting for background environmental modifications. None of these three processes individually produce $1/f$ noise. They found that, for model segment 32 points long, this simple model is capable to produce a realistic time series with realistic $1/f$ power spectra. The result suggests that the power-law structure of the HR signal may be non-stationary over fairly intervals, similar to the 5-min segments commonly used in short-term HRV analysis. Finally, in term of control system, Brownian motion corresponds to control being turned off and white noise indicates that control is on. The model indicates that HR series can be composed of periods during which the system drifts from an equilibrium point (Brownian motion) interrupted by intervals during which HR is locked to a required state by control mechanisms (white noise).

2.3.2 Detrended Fluctuation Analysis

Detrended fluctuation analysis (DFA) is a fractal-related method that provides for estimation of the scaling exponents α (the slope of the power spectrum). As previously underlined, the heartbeat time series is highly non-stationary. One of the main causes of non-stationarity is the continuous change in the environmental conditions. If the interest of the researcher is into the cardiovascular control system, which is mostly the case, modification in the activity due to external stimuli should be regarded as noise and eliminated. These are the aims of the "detrended fluctuation analysis" (DFA), a method introduced by Peng *et al.* [1995].

The RR series is integrated, after subtraction of the mean value

$$Y_k = \sum_{i=1}^{k} RR_i - \mu NN;$$

next the integrated time series is divided into boxes of equal length, n. In each box a least-square line, representing the trend in that box, is fit to the

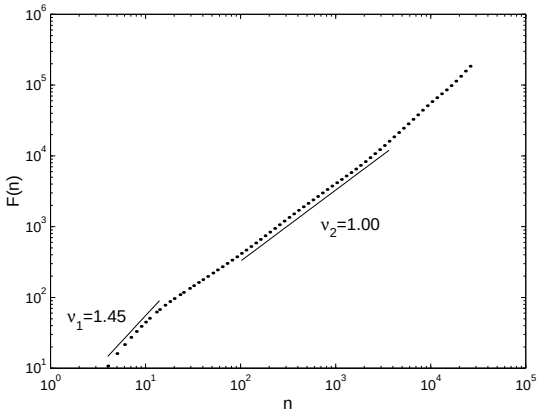


Fig. 2.6: *Detrended Fluctuation Analysis: the standard deviation of the integrated and detrended series $F(n)$ is plotted vs. n for a $\approx$ 24-hours RR sequence (healthy 41-years old subject). The slope of $F(n)$ over the range $n \in [100, 4000)$ is called "long-term fractal scaling exponent" and termed ν_2. Over short scale the graph of $F(n)$ displays a different evolution: this second slope ν_1 ("short-term fractal exponent"), usually computed for $n < 16$, is normally bigger than ν_2 (at least for normal subjects).*

data and then subtracted. The standard deviation of the integrated and detrended series (after being divided into blocks of length n) is named $F(n)$.

The computation is repeated over all time scale (box sizes). The graph of $F(n)$ versus n on a log-log plot is a straight line whose slope is the value of ν_2, the so called "long-term fractal exponent". A slope greater than 0.5 (white noise) and less than 1.0 ($1/f$ noise) indicates long-range correlations[16]

The fractal exponent can be related to the slope of the power spectrum α by the equation $\alpha = 2\nu_2 - 1$ [Pilgram and Kaplan, 1998].

Figure 2.6 presents the results for a DFA computed on a healthy subject. The scaling exponent is $\nu_2 = 1.00$, corresponding to $1/f$ noise ($\alpha = 1$).

The different scaling at the lower scales ($n < 16$) suggested the introduction of a second exponent ν_1 (see [Stanley *et al.*, 1999] and references therein). For normal subject it is usually larger than ν_2.

[16] Detrended fluctuation analysis is generally performed through an alternative and equivalent approach, numerically more efficient. The procedure is described in footnote 14 of [Peng *et al.*, 1994]. We computed the DFA adapting for MATLAB® the code present in the PHYSIOTOOLKIT database [Goldberger *et al.*, 2000].

2.3.3 Generalized Structure Functions

The word "fractal" is used to express a basic concept rather than being rigorously defined. Several definitions can be set out, each of them tailored to a specific problem[17]. Some general aspects of a fractal curves are: (i) self-similarity (the entity contains scaled copy of itself) and (ii) the quality of being everywhere continuous but not differentiable.

RR sequence are discrete series of events. A fractal sequence is essentially a self-affine curve (see equation (2.11)). Therefor in the following fractal and self-affine will used as synonymous.

It has been recognized that most fractals in nature are actually composed of an infinite set of interwoven subfractals. This structure becomes apparent when a particular measure μ supported by the set is considered. If the measure has different fractal dimension on different parts of the support, the measure is named *multifractal* [Feder, 1988]. For a formal definition see Halsey *et al.* [1986] and Muzy *et al.* [1993].

Multifractality is a property of a measure which is supported by a fractal set. In the following, for simplicity, we will refer to fractal sets supporting mono-fractal measures as *monofractal* and to fractal sets supporting multi-fractal measures as *multi-fractal*.

The most important request for a measure is that it should be positive definite. For a sequence X_n, like the RR series, proper measures may be $\mu_n = X_n^2$ or $\mu_n = |X_{n+1} - X_n|^2$. After the definition of a measure, the extraction of the generalized multifractal dimension is, commonly, performed by mean of a box-counting method [von Hardenberg *et al.*, 2000]. As a first step, the measure is integrated over intervals of length Δ, obtaining a new series $r_i(\Delta)$, given by

$$r_i(\Delta) = \sum_{j=1}^{\Delta} \mu_{(i-1)\Delta+j}$$

for $i \in [1, N_\Delta = N/\Delta]$; then the partition function

$$B(\Delta, q) = \left\{ \frac{\sum_{i=1}^{N_\Delta} [r_i(\Delta)]^q}{\left[\sum_{n=1}^{N} \mu_n\right]^q} \right\}^{\frac{1}{q-1}}$$

is constructed. In the limit for $\Delta \to 0$, a fractal measure is characterized by a scaling behaviour of the partition function,

$$B(\Delta, q) \sim \Delta^{D_q},$$

[17] Mandelbrot [1983] underlines the difficulty of a universal definition.

where D_q is the generalized fractal dimension. If all the D_q coincide with D_0, the measure is monofractal. On the contrary, if the spectrum of the generalized dimension decreases, $D_q < D_p$ for $q > p$, the measure is multifractal. The exponents D_q are usually computed as slopes of least square lines in a log-log graph of $B(\Delta, q)$ vs Δ.

The spectrum of generalized fractal dimensions can be associated with the presence of nonlinear correlations in the signal. Therefore a discussion on the presence of a multifractal spectrum in the HRV signal implies debating about its nonlinear nature. Being the topic really controversy, we would like to employ a robust method but the box-counting technique is not good enough. In fact, as von Hardenberg *et al.* [2000] recently pointed out, spurious multifractality can be detected even if the input sequence is: (i) a white noise with exponential or hyper-exponential distribution; (ii) a nonlinearly-filtered, linear autoregressive process.

A more robust approach is through of the *generalized structure functions* that not require the introduction of a measure. The method is a generalization of equation (2.12); for a random process X_n, GSF are defined[18] as

$$\mathrm{GSF}(\Delta, q) \equiv \langle |X_{n+\Delta} - X_n|^q \rangle.$$

If the analyzed signal has fractal nature, it does exist a scaling region where $\mathrm{GSF}(\Delta, q) \sim \Delta^{qh_q}$. When h_q is constant with q ($h_q = h_2 = H$, $\forall q$), the signal is monofractal (or self-affine), like in the case of standard Brownian motion and white noise[19]. A multifractal signal is instead characterized by

[18] The works on this subject were pioneered by Frisch and Parisi [1976], in the context of fully developed turbulence. Given a velocity field $v(x)$, they defined *generalized structure function* in the limit $l \to 0$

$$S(l, q) = \langle |v(x + l) - v(x)|^q \rangle_x \sim l^{qh_q}, \tag{2.13}$$

where $\langle\ \rangle_x$ is the mean value computed over x.

[19] A couple of examples. Consider a Gaussian white noise $WGN \stackrel{d}{=} \mathrm{N}(\mu, \sigma)$; the generalized structure functions can be computed analytically and are constant. In fact

$$\mathrm{GSF}(\Delta, q) = \frac{\sigma^q 2^q}{\sqrt{\pi}} \Gamma\left(\frac{q+1}{2}\right) \sim \Delta^0,$$

where $\Gamma()$ is the Gamma function, therefor $h_q = 0$ for each q. Equivalently, for the self-affine random walk $x_{i+1} = x_i + w_{i+1}$, where $w_i \stackrel{d}{=} \mathrm{N}(0, \sigma)$,

$$\mathrm{GSF}(\Delta, q) = \frac{\sigma^q 2^{q/2}}{\sqrt{\pi}} \Gamma\left(\frac{q+1}{2}\right) \Delta^{q/2} \sim \Delta^{q/2},$$

that is, $h_q = 1/2$ for every q.

the scaling exponent decrease for increasing q ($h_q < h_p$ for $q > p$). Once more, it is useful to underline that monofractality or multifractality do not necessary imply the presence of deterministic chaos.

The generalized fractal dimension spectrum D_k, computed through the box-counting method, can be obtained also from the scaling exponents h_q [Muzy *et al.*, 1993] via the Legendre transform

$$D_k = \min_q [qk - qh_q + 1]. \tag{2.14}$$

2.3.4 Wavelet-Transform Modulus-Maxima

The generalized structure functions approach requires the process X_n to be stationary, at least at the scales of interest. To overcome this possible difficulty, recent works [Amaral *et al.*, 1998; Ivanov *et al.*, 1999; Stanley *et al.*, 1999] use a different method to compute the generalized dimensions spectrum from a time series. The technique is called *wavelet-transform modulus-maxima* (WTMM) and it was introduced by Muzy *et al.* [1993] (see also [Muzy *et al.*, 1994]). Wavelets can remove polynomial trends that could cause box-counting techniques to fail in quantifying the local scaling of the signal.

The Wavelet transform of the sequence X_i is defined as:

$$T_\psi[X_i](b,a) = \frac{1}{a} \sum_i \psi\left(\frac{i-b}{a}\right) X_i,$$

where $a > 0$ is a scale parameter and b a space position. The *analyzing wavelet* ψ is chosen to be orthogonal to some low-order polynomials:

$$\int_{-\infty}^{\infty} x^m \psi(x) dx = 0, \qquad \forall m : 0 \le m \le N.$$

A class of commonly used analyzing wavelet satisfying this condition is given by the successive derivatives of the Gaussian function

$$\psi^{(N)}(x) = \frac{d^N(e^{-x^2/2})}{dx^N}.$$

The higher the order, N, of the derivative, the higher the order of the polynomial trends removed.

The WTMM method defines the *partition function*

$$Z(a,q) = \left[\sum \max_b |T_\psi[X_i](b,a)|\right]^q,$$

where $\max_b f(b)$ are *all* the local maxima of the function $f(b)$. In the limit $a \to 0$, the partition function

$$Z(a, q) \sim a^{\tau(q)}.$$

Multifractal processes are characterized by a non linear function $\tau(q)$, monofractal by a linear one. The generalized dimension can be computed by mean of the Legendre transform

$$D_k = \min_q [qk - \tau(q)]. \tag{2.15}$$

Comparing equation (2.15) with (2.14), the relation among the two scaling exponent $\tau(q)$ and h_q (the structure function scaling parameter defined in the previous section) results

$$\tau(q) = qh_q - 1.$$

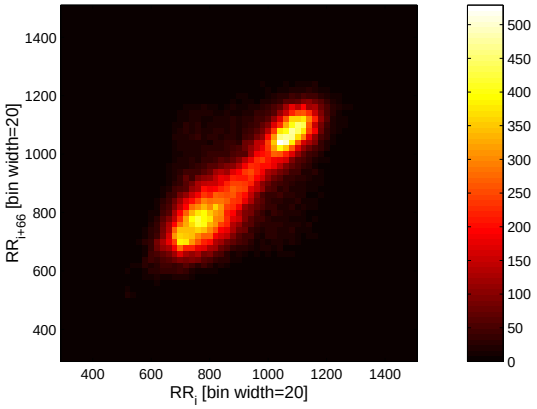


Fig. 2.7: *Two-dimensional histogram computed on a ≈ 24-hours Holter recording. Night and day variability are well recognizable. Δ is the first numerical minimum of the RR autocorrelation function. The series is the same employed in panel* (a) *of figure 2.2.*

2.4 Other Techniques: Recurrence Maps

Recurrence maps (or "scattergram") represents a very common tool, which can be useful in understanding the attractor morphology with low order dynamical system, in the hypothesis that the system is bounded onto the stable nonlinear manifold.

A recurrence plot of a series X_n is a graph of X_n against $X_{n+\Delta}$, where Δ is a time lag. Δ is usually chosen to be 1; otherwise, if the correlation between adjacent values in the sequence is too strong (it would hide any other useful information), it is better to use larger value (for example the first minimum of the autocorrelation function or the first minimum of the mutual information function).

The slope α_R of the line $RR_{i+1} = \alpha_R RR_i + \beta$ is a classical heart rate variability index. It is fitted with a standard linear regression technique onto the variables (RR_{i+1}, RR_i).

In figure 2.7, a two-dimensional histogram is computed for the variable $(RR_{i+\Delta}, RR_i)$. The graph is basically equivalent to a recurrence map.

2.5 Other Techniques: Approximate Entropy

Approximate Entropy (ApEn) is a "regularity statistic" that quantifies the unpredictability of fluctuations in a time series. The presence of repetitive patterns makes the time series more predictable.

The statistic was introduced by Pincus [1991]. Starting from a sequence u_i, with $i = 1, 2, \cdots, N$, the first step is the construction of the vectors

$$x_i = \begin{bmatrix} u_i & u_{i+1} & u_{i+2} & \dots u_{i+m-1} \end{bmatrix},$$

where $m \in \mathbb{N}$. The distance between vectors x_i e x_j is defined as the maximum difference in their respective scalar components, $d[x_j, x_j] = \|x_i - x_j\|_\infty$. For each vector x_i, $\mathcal{O}_i(m, r)$ is the number of vectors x_j, $j \leq N - m + 1$, such that $d[x_j, x_j] \leq r$, where $r \in \mathbb{R}^+$. Then construct

$$\mathcal{C}_i(m, r) = \frac{\mathcal{O}_i(m, r)}{N - m + 1} \qquad \Phi(m, r) = \frac{\sum_{i=1}^{N-m+1} \log \mathcal{C}_i(m, r)}{N - m + 1}$$

$\mathcal{C}_i(m, r)$ defines within a tolerance r the regularity, of frequency, of patter similar to a given patter of length m.

Pincus named *approximate entropy*

$$\text{ApEn}(m, r) = \lim_{N \to \infty} [\Phi(m, r) - \Phi(m + 1, r)].$$

It measures the likelihood that runs of patterns that are close for m observations remain close on next incremental comparison [Pincus, 1995]. A time series containing many repetitive patterns has a relatively small ApEn, a less predictable process has a higher ApEn .

The parameter m specifies the length of patters that are compared. The bigger m, the higher the level of detail required. r defines the "noise" level accepted in the comparison: differences, smaller than r, between two vectors, in absolute value, are considered not relevant; r is usually expressed as a figure of the standard deviation of u_i.

The definition is clear after reorganizing the terms

$$\begin{aligned} \text{ApEn}(m, r) &= \lim_{N \to \infty} \left[\frac{\sum_{i=1}^{N-m+1} \log \mathcal{C}_i(m, r)}{N - m + 1} - \frac{\sum_{i=1}^{N-m} \log \mathcal{C}_i(m + 1, r)}{N - m} \right] \\ &= \lim_{N \to \infty} \sum_{i=1}^{N-m} \frac{\log[\mathcal{C}_i(m, r)/\mathcal{C}_i(m + 1, r)]}{N - m} \\ &= -\text{mean}_i \left[\log \frac{\mathcal{C}_i(m + 1, r)}{\mathcal{C}_i(m, r)} \right] \end{aligned}$$

The quantity $\mathcal{C}_i(m+1, r)/\mathcal{C}_i(m, r)$ is the conditional probability that $|u_{j+m} - u_{i+m}| \leq r$ given that $|u_{j+k} - u_{i+k}| \leq r$ for $k = 0, 1, \ldots, m - 1$.

The definition is given in the limit $N \to \infty$, and in any practical situation it is necessary to employ an estimator. The most common one is

$$\mathrm{ApEn}^*(m,r) = \frac{\sum_{i=1}^{N-m+1} \log \mathcal{C}_i(m,r)}{N-m+1} - \frac{\sum_{i=1}^{N-m} \log \mathcal{C}_i(m+1,r)}{N-m}$$

[Pincus and Huang, 1992], where the limit is simply skipped.

Approximate entropy has been fruitfully employed to quantify the regularity of RR series and has proved to be a reliable index[20]. Nevertheless, it is important to underline that ApEn is only a regularity statistic and *not* a consistent estimate of the K-S entropy, though similarity in the definition. For this reason any speculation on the deterministic nature of the process under analysis using ApEn is absolutely not well posed. Its only application is the comparison among time series based on their *regularity*. The intuition motivating ApEn is that if joint probability measures that describe each of two process are different, then their marginal distribution on a fixed partition, given by conditional probability, are likely different.

[20] A review on approximate entropy, containing a fast numerical code, is [Fabani and Sassi, 1996]; see also appendix C.

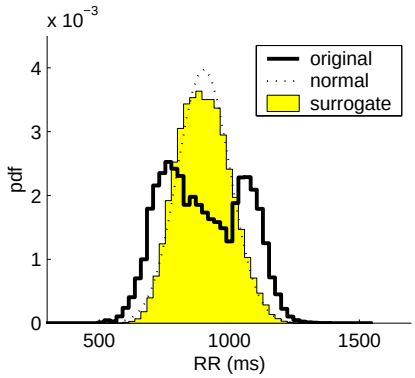


Fig. 2.8: *The probability density function (pdf) for a RR series belonging to a 38-years old healthy subject (thick line) is compared with the pdf of a set of surrogate data obtained with phase-randomization (yellow) . The surrogation process, though maintaining the autocorrelation function, renders the pdf normal. The dotted line is the normal distribution obtained with a MLE-estimator from the surrogate data.*

2.6 Testing for Nonlinearity

Two distinct reasons can motivate a nonlinear approach to the analysis of a time series, like the interbeat sequence. The first may be that the arsenal of linear methods has been exploited thoroughly but all the efforts left certain structures in the time series unaccounted for. Otherwise it is possible that certain *a priori* knowledge on the structure of the system leads to the inclusion of nonlinear components. The latter may be the case of the cardiovascular system, where several nonlinear mechanisms are accounted for; but the fact that a given system includes nonlinear components does not necessarily imply that nonlinearities are also contained in the specific signal we measured from it (e.g. reasonable nonlinearity in the heart control do not hint that the RR series contains itself this nonlinearity). Consequently, for a data driven analysis, the application of nonlinear time series methods has to be justified by establishing nonlinearity in the time series [Schreiber and Schmitz, 2000].

The nonlinear nature of the RR signal was firstly suggested by Goldberger *et al.* [1990]. After him, given the importance of the problem, several studies were performed to inspect the speculated nonlinearity in the RR series. Nevertheless the issue is still an open question.

In time series analysis literature, the most popular tool developed to verify the nonlinear nature of a stationary process is the *surrogate data*

method. Stated the classical Occam's razor that given two explanations for data we should favor the "simpler" one, surrogate data techniques allow to find the more uncomplicated explanation that cannot be ruled out based on the data we have in hands. The task compares a nonlinear discriminating statistic, computed on the measured data, to its empirical distribution on a collection of Monte Carlo realizations of the *null hypothesis*.

To apply the method, it is, firstly, necessary to postulate the null hypothesis (NH), on the data under analysis, which the test should, if necessary, falsify. The null may be, *e.g.*, "the data are colored noise". Then suitable synthetic series consistent with the NH, the surrogate data, are prepared. These sequences share with the original ones certain properties, invariant under the NH. Finally the null is attempted to be rejected by comparing the value of a nonlinear parameter taken on by the data with the same parameter computed on the surrogate series. If no differences arise among the two it means that the index is measuring characteristics participated by *all* the sequences, thus the null cannot be rejected safely.

The described process of using surrogate data to find the range of values for data consistent with NH is called *bootstrapping*.

Several surrogation methods have been developed, each suitable for particular null hypotheses. In the following section we present a few of them[21].

2.6.1 Phase-Randomization

Any stationary linear process is fully specified by its autocorrelation function, therefor by the its power spectrum.

Thus (i) computing the Fourier transform of the original sequence, (ii) replacing the phases with random number uniformly distributed between 0 and 2π and (iii) transforming back the result, shall produce a surrogate sequence having the same autocorrelation function of the original one. In fact the procedure does not alter anyhow the power spectrum of the process [Kaplan and Glass, 1995].

The method of creating surrogate data by adding random phases to the Fourier transform of a process is called *phase-randomization*. The null hypothesis it tries to nullify is: "the signal is generated by a linear precess (colored noise)".

[21] The technique presented are meant to construct *constrained realizations* [Theiler and Prichard, 1996]: measurable properties of a time series are conserved rather than a speculated underlying model. They are certainly appropriate for a broader class of problems.

2.6.2 Amplitude Adjusted Fourier Transform

Unfortunately, surrogate data obtained using phase-randomization of the Fourier transform have a Gaussian distribution of values. Many time series, among which RR sequence, do not have a normal distribution, and therefore several nonlinear techniques, influenced by non-normal distribution of data, are cheated by the surrogated sequences. An example is presented in figure 2.8 where we compare the original distribution with the one of the surrogate data. The Gaussian distribution, fitted via MLE estimator, approximates very well the surrogate data one.

Non-Gaussian distribution can be created by a linear stochastic process passed through a nonlinear static (not depending on time) filter: non linear measurement distortion are a typical example of static filter use.

Theiler *et al.* [1992] proposed a method, called *Amplitude Adjusted Fourier Transform* (AAFT), that avoids being mislead by nonlinear static filters. The null hypothesis is in this case: "the signal is generated by a linear stochastic process distorted by a nonlinear, monotonically increasing, filter".

The route to obtain the surrogate series $\tilde{s}_n$ from s_n, based on AAFT, is: (i) s_n is rank-ordered according to a set of Gaussian random number; this makes the resulting series $g_n = r(s_n)$ Gaussian, though following the time evolution of s_n; (ii) g_n is phase-randomized to obtain $\tilde{g}_n$; (iii) $\tilde{g}_n$ is rank-ordered according to the distribution of the original data s_n leading to the final surrogate series $\tilde{s}_n = \hat{r}(\tilde{g}_n)$.

$\tilde{s}_n$ has, by construction, the same distribution of amplitudes of s_n but only approximately the same power spectrum[22]. Such discrepancy can lead to false rejection of the null hypothesis. To overcome these difficulties, [Schreiber and Schmitz, 1996] developed a further refinement (IAAFT; "I" stands for "iterated") of the algorithm. In this way they correct iteratively the mismatch among the spectra under a desired threshold. The improvement lies only in the preparation of the surrogate data, thus the null hypothesis remains unchanged.

Being (a) l_n a sorted version of s_n and (b) S_n the amplitude of the Fourier transform $|\mathscr{F}\{s_n\}|$, IAAFT surrogate data are obtained: (i) s_n's amplitudes

[22] The idea underlying AAFT is that the data may have been produced by filtering an original gaussian sequence w_n through a nonlinear static filter expressed by the monotonically increasing function $h(x)$. In this hypothesis and in the limit of $N \to \infty$, where N is the number of points in the sequence, $r(x)$, the rank-ordering procedure, approaches the inverse of $h(x)$ with mean square fluctuation proportional to $1/N$. Discrepancies in the power spectra are due to the differences between the two nonlinear rank-ordering functions $r(x)$ and $\hat{r}(x)$; the AAFT algorithm introduces a bias towards a flat spectrum [Schreiber and Schmitz, 1996].

are randomly shuffled to obtain $s_n^{(0)}$. (ii) As the Fourier transform $\tilde{s}_n^{(i)}$ of $s_n^{(i)}$ is prepared, the Fourier amplitude of $\tilde{s}_n$ are replaced with S_n. The spectrum is transformed back: $\bar{s}_n^{(i)} = \mathscr{F}^{-1}\{S_n \exp[j\omega^{(i)}]\}$, where $\omega^{(i)}$ are the phases of $\tilde{s}^{(i)}$. At this stage, the power spectra of the intermediate surrogate series $\bar{s}_n^{(i)}$ and of the original s_n are identical; (iii) $\bar{s}_n^{(i)}$ is rank-ordered according to the original amplitudes l_n to obtain $s_n^{(i+1)}$. Steps (ii) and (iii) have to be repeated until the differences between $|\tilde{s}^{(i+1)}|$ and S_n can be considered negligible. The final surrogate series, $s_n^{(i+1)}$, has, by construction, the same amplitude distribution of the s_n and the same power spectrum, with good approximation.

2.6.3 Remarks

The null hypothesis of a Gaussian linear process filtered through a monotonic function (AAFT and IAAFT) is quite general, nevertheless its rejection does not imply nonlinear dynamics. A statistical test can only be falsified; it is improper to make positive statement.

When producing surrogate data, *no* zero-padding can be employed to perform quickly any Fourier transform. Padding introduces structures into a time series that will not exists in the surrogates any more, leading necessarily to misleading results. Mixed-radix-fft algorithms may be useful to reduce the number of points truncated at the end of each series, without losing too much computational performances.

Any mismatch between the beginning and the end of a time series poses problems with Fourier transform. Usually, edge effects are treated by windowing and zero padding, both to avoid with surrogate data. The problem is specially effective with phase-randomization, when the addition of high frequency components to the series is evident. To prevent edge mismatch a sub-interval of the recording can be chosen so that the discrepancies among the begin and the end of the series are minimized [Schreiber and Schmitz, 2000][23].

[23] Schreiber and Schmitz [2000] suggest to measure the end point mismatch by

$$\gamma_{jump} = \frac{(s_1 - s_N)^2}{\sum_{i=1}^{N}(s_n - \langle s_n \rangle)^2},$$

where s_1 and s_N are the first and the last point in the sequence s_n.

3. WORKING ON THE NOLTISALIS DATABASE

Abstract

In this chapter, we present the results obtained analyzing the $\approx$ 24-hours RR series, in the NOLTISALIS database. NOLTISALIS is a multicentric research program which aims at the nonlinear analysis of heart rate variability series. it is composed by several italian university departments and rehabilitation clinics. We focused on the multifractal characteristics of the HRV signal through two approaches: generalized structure functions, a tool proposed in the context of fully developed turbulence, and wavelet-transform modulus-maxima. This last method, recently introduced, should be less sensitive to non-stationarity. A large number of classical parameters, for the analysis of the HRV signal over long time scales, have been considered to set up a proper comparison. We considered classical time-domain indexes, "monofractal" characteristics ($1/f^{\alpha}$ spectrum; detrended fluctuation analysis) and a regularity statistic (approximate entropy).

The spectrum of generalized fractal dimensions can be associated with the presence of nonlinear correlations in the signal, and possibly, with the nonlinear nature of the signal HRV. To verify this last hypothesis, we computed the generalized structure function also on a set of surrogate data (amplitude adjusted surrogate data). In 32 cases, the multifractal spectrum of the original RR series differs from those obtained from the surrogate signals significatively (t-test), this supporting the idea of a RR nonlinear nature. However, absolute values of the multifractal exponents h_q are very small and non-stationarity effects should be further investigated.

3.1 Introduction: the NOLTISALIS Database

The NOLTISALIS[1] database was collected by the cooperation of several university departments and rehabilitation clinics[2] in Italy. The acronym NOLTISALIS stands for "NOnLinear TIme Series AnaLysIS" and highlights the objectives of the multicentric project: to study the nonlinear nature of the heart rate variability signal from a time series perspective.

The database is composed by several signals:

1. 50 RR series extracted from 24-hours Holter recordings (they have been collected in five different subsets; see table 3.1);

2. 5 RR series extracted from regular ECG recordings of normal subjects in stationary and controlled conditions (seated on armchair, no external noises, low-intensity lights, room at constant temperature);

3. synthetic RR series (and their surrogate data [Chang *et al.*, 1994]), obtained with the Ornstein-Uhlenbeck process [Theiler *et al.*, 1992] (2 series) and with the baroreflex-regulation model [Cavalcanti and Belardinelli, 1996] (2 series);

4. test series obtained from classical low-dimensional nonlinear deterministic systems (Henon, Lorenz and Rossler; 5 series each).

Five subsets compose the 24-hours Holter series data set: each subset refers to a particular health state of the subject on whom the recording was made. Apart from healthy adults, other four pathologic cardiac conditions are present as reported in table 3.1.

3.2 Methods

ECG data were recorded using different Holter devices. they are reported in table 3.1; The last column of table 3.2 displays, for each recording, the sampling frequency. Beats were labelled using automated procedures through

[1] The works presented in this chapter were performed on the first version of the database, but have been made coherent with the third released (March 2000). The database can be requested writing to `signorini@biomed.polimi.it`; it is distributed only on CD-rom support.

[2] Dipartimento di Elettronica, Informatica e Sistemistica (Università di Bologna), Istituto per i Circuiti Elettronici (C.N.R., Genova), Dipartimento di Bioingegneria (Politecnico, Milano), Fondazione Salvatore Maugeri (Pavia), Istituto di Fisiologia Clinica (C.N.R., Pisa), Istituto di Biofisica (C.N.R., Pisa), Dipartimento di Informatica e Sistemistica (Università "La Sapienza" di Roma).

Population	Code	Series	Location	Device
normal	NR	10	Veruno	custom
hypertension	IP	10	Roma	Del Mar Avionics
(after) myocardial infarction	MI	10	Veruno	custom
heart failure (A)	SC	5	Veruno	custom
heart failure (B)	SC	5	Montescano	Oxford Medilog Excel
heart transplanted	TR	10	Montescano	Oxford Medilog Excel

Tab. 3.1: *The 50 RR 24-hours series in the* NOLTISALIS *database are classified into 5 populations, according to the health state of the subject to whom the recording belongs. In table the five populations are indicated along with the number of subjects in each of them; "Location" refers to the locality where the recording were performed. Some recording were obtained with custom Holter recorder ("custom").*

a proper analysis software. Labels for detected beats were: N (normal), V (ventricular ectopic), S (supraventricular ectopic) or X (artifacts). Subsequently, experienced Holter scanning technicians manually verified the annotations.

Using this procedure, R-wave peaks were determined with a resolution value in the range 1-8 ms. The finite resolution implies that estimates of the interbeat intervals are affected by a noise due to estimation error (see section 1.5 for more details). The signal-to-noise ratio[3] over the whole database, ranges from 25 to 51 dB.

The populations are not exactly age-matched[4]: this is mostly due to the fact that most cardiac pathological states are age dependent.

3.2.1 Artifacts Detection

The quality of the Holter recordings is high. Nevertheless the acquisition process lasts many hours, during which the subject moves or even exercises. The sweat or the movement of the electrodes can, therefore, corrupt the ECG signal. The peaks detection procedure marks low quality intervals,

[3] The signal to noise ratio was computed on the spectrum of the signals: $\mathrm{SNR} = 10\log_{10}\mathrm{SDNN}^2/\mathrm{E}[u_i^2]$. See sections 2.1.1 and 1.5 for the definitions of SDNN and $\mathrm{E}[u_i^2]$, respectively.

[4] The mean ages for the five populations are (standard deviation within brackets):

NR	42.2(±6.4)
IP	40.7(±1.1)
MI	50(±10.2)
SC	53.6(±11.2)
TR	44.9(±14.8).

	f_c (Hz)	start (h)	length (h)	RR	S	V	X	Age
NR01	250	20:30	21:40	82592	1	1	77	38
NR02	250	12:40	23:50	103821	8	0	15	46
NR03	250	8:40	24:00	102493	224	0	27	56
NR04	250	10:40	23:10	105142	12	3	256	40
NR05	250	11:40	21:10	117284	3	2	127	34
NR06	250	11:20	23:10	104495	14	0	288	35
NR07	250	8:50	24:30	107056	18	1	205	41
NR08	250	8:30	24:50	104589	7	0	36	46
NR09	250	9:30	24:40	111285	9	432	17	43
NR10	250	11:40	19:10	81774	75	0	5	43
IP01	512	9:41	24:46	117022	9	0	127	40
IP02	512	10:50	22:06	106386	12	0	49	43
IP03	512	10:01	23:56	116720	41	0	224	42
IP04	512	12:00	21:48	103079	39	332	595	41
IP05	512	9:33	20:23	97820	7	1	213	41
IP06	512	9:30	23:48	110618	5	4	95	40
IP07	512	11:47	22:56	113641	28	2	318	40
IP08	512	12:12	22:47	98770	5	0	83	40
IP09	512	11:42	22:47	100255	4	1	24	40
IP10	512	15:17	21:14	113242	2	0	21	40
MI01	250	9:20	24:30	87636	11	10	188	60
MI02	250	9:30	22:40	108721	11	197	189	33
MI03	250	9:10	24:30	108243	33	5	50	56
MI04	250	10:30	23:00	125628	119	242	47	34
MI05	250	8:40	22:60	85599	6	31	249	61
MI06	250	14:50	24:30	100499	44	9	97	52
MI07	250	15:10	23:00	115571	428	197	23	48
MI08	250	11:10	23:30	102814	18	25	68	43
MI09	250	15:20	18:00	74967	35	38	87	57
MI10	250	15:40	24:40	76963	46	5	6	56
SC01	250	11:10	23:0	122775	71	88	67	47
SC02	250	8:50	23:20	103783	28	37	0	64
SC03	250	9:30	24:40	97335	21	301	13	68
SC04	250	9:20	23:10	102990	33	192	55	64
SC05	250	11:30	21:40	95138	10	4	37	53
SC06	128	10:25	22:07	95869	0	926	0	56
SC07	128	8:45	22:56	139625	0	122	0	36
SC08	128	9:27	22:55	127445	0	40	0	36
SC09	128	11:36	22:57	90730	0	123	0	60
SC10	128	11:47	22:57	95755	0	106	12	52
TR01	128	9:44	22:55	122042	0	26	0	18
TR02	128	11:10	22:56	130619	0	6	1	46
TR03	128	12:13	22:55	131055	0	1	0	60
TR04	1024	9:33	24:50	104097	0	0	20	35
TR05	1024	11:18	23:58	117020	0	5	0	54
TR06	1024	10:30	23:17	97639	0	5	30	46
TR07	128	9:44	22:56	103678	0	3	0	62
TR08	128	9:00	22:57	112149	0	21	2	24
TR09	128	10:27	23:35	128230	0	0	54	53
TR10	128	11:10	22:55	113901	0	43	2	51

Tab. 3.2: *Available information for each subject belonging to the* NOLTISALIS *database. f_c is the sampling frequency (Hz); RR is the number of points in the series; S, V e X are the number of intervals labelled, respectively, as supraventricular, ventricular and artifacts.*

for which it is impossible to locate any QRS complex. In the RR series they appear in distances, in milliseconds, among the last R peak, *before* the spoiled segment, and the first one, *after* it. This choice should ensure that the original temporal structure of the signal is preserved, as it is expressly required to employ several of the techniques presented in the following.

Dramatically, several of the series contain artifacts which are marked as normal beats (N), and not with the proper label (X). This is due to a defect in the labelling software employed in one of the rehabilitation clinic. Therefore, it becomes necessary to set on a standard procedure able to detect artifacts masked as normal beats and to apply the same criterium to the whole database.

Ventricular ectopic and supraventricular ectopic beats need also a correction. In fact they generate in the myocardial muscular tissue or, anyway, outside the proper pacemaker, the sinoatrial node. An ectopic beat usually happens just after the end of the refractory period and before the next impulse from the SA node (SAim1). When SAim1 arrives, the muscular tissue is in refractory period as well and there's no consequent contraction. In the ECG recording this ends in a shorter first RR interval (the ectopic beat arrives before SAim1) and in a longer second RR (SAim2 is fired autonomously from the ectopic beat). If the topic of the study is the regulation of the heart rate, ectopic beats must be regarded as artifacts, being independent from any regulation input. The labelling software recognizes ectopic beats by the absence of the P wave, marking them correctly.

In this section we present the algorithm we developed to detect artifacts. Obviously, there is a compromise between specificity and sensitivity: we can afford the presence of a few not-detected artifact, if they are comparable to the surrounding RR intervals, but we do not want the algorithm to correct too many true-normal intervals. The procedure is described step by step:

0^{th} PREPROCESSING: A preprocessing step avoids the presence of RR points considered meaningless. Intervals which undergo one of the two conditions $RR_i > 10000$ or $RR_i \leq 0$ are considered anomalous and replaced with the mean value of the whole sequence.

1^{st} LOCALIZATION: The points that need to be fixed are marked if they undergo one of following criteria:

1. they are already marked as artifacts ("X") in the labels file linked to the data;

2. are marked as artifacts the RR_i points that undergo the condition $|RR_i - B_i^1| > 10$ std$(RR - B^1)$ where std() is the standard deviation. Ectopic beats ("V" ventricular or "P" post-ventricular) are excluded *a priori*. The baseline B_i^1 is computed by filtering the RR series with a 17 points rectangular window, at first, applied with a forward step, and then with a backward step to avoid the introduction of delay. This procedure is necessary to eliminate spurious trains of spike.
3. A RR_0 point, equal to the mean value of the signal, is added at the begin of the sequence, then the series $I_i = RR_{i+1} - RR_i$ is computed. The distribution of I_k is estimated with a normalized histogram with bins width of 20 ms (see section 2.1.2). A Lévy stable distribution is fitted to the normalized histograms.

   Using the Lévy distribution so estimated, is computed the value $I_{1/1000}$ such that $P(I_{1/1000}, \alpha, \gamma) = 1/1000$. If $I_{1/1000}$ is greater than 1000, than it is set $I_{1/1000} = 1000$.

   Are marked as artifacts the RR_i points that undergo both the two conditions $|I_i| > I_{1/1000}$ and $|RR_i - B_i^1| > 4$ std$(RR - B^1)$. As before, ectopic beats are excluded *a priori*.

2^{nd} CORRECTION: A first correction is performed on ectopic beats and artifacts with a linear interpolation on the two surrounding points, leading to a new RR^1 series.

A low pass filter mask with cutting frequency 0.01 Hzeq is constructed with standard windowing technique using an Hanning five points window. Then the filter is applied to the RR^1 sequence in two steps, the first forward and the second backward. This is made to avoid, like before, the introduction of any delay. We get in this way the RR^2 signal.

Points needing correction in RR are now replaced with the correspondent in RR^2. In this way we can make corrections with a moving average mean, avoiding the bias coming from the presence of points very far from a reasonable value.

The algorithm has been employed on every RR series in the database. Table 3.3 reports the corrections made on each signal.

	E	X	A	X^*	$Total$
NR1	2	77	0	86	88
NR2	8	15	0	67	75
NR3	224	27	0	34	258
NR4	15	256	0	212	227
NR5	5	127	0	187	192
NR6	14	288	0	64	78
NR7	19	205	0	147	166
NR8	7	36	0	110	117
NR9	441	17	0	14	455
NR10	75	5	0	97	172
IP1	9	127	0	216	225
IP2	12	49	0	88	100
IP3	41	224	2448	327	2816
IP4	371	595	0	638	1009
IP5	8	213	0	212	220
IP6	9	95	0	156	165
IP7	30	318	1	363	394
IP8	5	83	0	86	91
IP9	5	24	0	24	29
IP10	2	21	0	118	120
MI1	21	188	0	107	128
MI2	208	189	0	143	351
MI3	38	50	0	77	115
MI4	361	47	0	392	753
MI5	37	249	0	278	315
MI6	53	97	0	113	166
MI7	625	23	0	622	1247
MI8	43	68	0	153	196
MI9	73	87	0	105	178
MI10	51	6	0	42	93
SC1	159	67	0	602	761
SC2	65	0	0	273	338
SC3	322	13	0	329	651
SC4	225	55	0	219	444
SC5	14	37	0	55	69
SC6	926	0	0	875	1801
SC7	122	0	0	45	167
SC8	40	0	0	81	121
SC9	123	0	0	209	332
SC10	106	12	0	112	218
TR1	26	0	0	80	106
TR2	6	1	0	97	103
TR3	1	0	0	163	164
TR4	0	20	0	112	112
TR5	5	0	0	277	282
TR6	5	30	0	43	48
TR7	3	0	0	260	263
TR8	21	2	0	119	140
TR9	0	54	0	144	144
TR10	43	2	0	123	166

Tab. 3.3: *Using the technique described, the RR data have been corrected. In this table, for each series, the amount of changes made are reported. The labels mean: E ectopic beats, X declared artifacts, A anomalies and X^* effectively found artifacts. $Total$ is the total number of changes made on the series.*

	NR	IP	MI	SC	TR
μNN	799.3(±65.8)	736.7(±35.7)	836.1(±135.0)	766.0(±108.2)	703.7(±88.4)
SDNN	133.5(±21.3)	135.4(±38.4)	102.0(±17.8)	86.9(±19.3)	70.4(±27.6)
RMSSD	32.3(±13.4)	31.4(±13.4)	23.0(±6.5)	17.8(±6.5)	10.8(±4.5)
SDANN	119.4(±23.8)	124.9(±42.7)	89.9(±16.7)	79.1(±17.0)	70.0(±28.8)
HTI	37.5(±7.5)	37.0(±7.8)	29.1(±5.0)	25.5(±7.8)	17.8(±5.7)
α Lévy	1.64(±0.28)	1.61(±0.30)	1.82(±0.13)	1.85(±0.14)	1.84(±0.12)
α spectrum	0.99(±0.06)	1.07(±0.10)	1.07(±0.13)	1.14(±0.28)	1.54(±0.26)
ν_1 DFA	1.15(±0.17)	1.06(±0.13)	1.15(±0.27)	1.10(±0.50)	0.44(±0.26)
ν_2 DFA	1.01(±0.05)	1.09(±0.04)	1.07(±0.07)	1.09(±0.11)	1.39(±0.16)
$h_1 \equiv$ H	0.09(±0.02)	0.13(±0.04)	0.13(±0.06)	0.17(±0.05)	0.32(±0.09)
h_2	0.07(±0.02)	0.12(±0.04)	0.12(±0.06)	0.15(±0.04)	0.29(±0.09)
h_3	0.06(±0.02)	0.11(±0.04)	0.10(±0.06)	0.13(±0.04)	0.25(±0.10)
$h_{\mathcal{R}}$	0.06(±0.04)	0.09(±0.04)	0.09(±0.02)	0.10(±0.05)	0.18(±0.10)
τ_3	0.86(±0.07)	0.83(±0.05)	0.78(±0.11)	0.78(±0.16)	0.39(±0.24)
τ_3	0.26(±0.12)	0.24(±0.16)	0.16(±0.12)	0.16(±0.21)	0.02(±0.03)
$\tau_{\mathcal{R}}$	0.87(±0.22)	0.99(±0.36)	0.81(±0.14)	0.82(±0.23)	0.38(±0.23)
α_S recurrence	0.83(±0.06)	0.85(±0.05)	0.80(±0.08)	0.86(±0.13)	0.97(±0.03)
ApEn night	1.09(±0.12)	1.22(±0.24)	1.20(±0.26)	0.89(±0.33)	0.82(±0.29)
ApEn day	1.05(±0.29)	0.95(±0.23)	1.03(±0.14)	1.05(±0.33)	1.04(±0.18)

Tab. 3.4: *Mean values, of parameters discussed in the chapter, obtained averaging on each population, are presented in table; standard deviations are reported within brackets. An empty line divides different groups of parameters; (from the top) (i) statistical & geometrical, (ii) fractal, (iii) multifractal and (iv) others. See caption of figure 3.6 for further details.*

3.3 Results

The work on the NOLTISALIS database has two main objectives: (i) to gain additional information on the long period behaviour of the cardiovascular system and (ii) to test statistical indexes, able to significantly discriminate pathological subjects, with a possible diagnostic application.

The new information gained within objective (i) could be used, in the future, to construct a new model of the heart rate generating mechanism or to refine an existing one.

The two aims would be pursued at the same time, following the order we used along chapter 2. Multifractal analysis is the more advanced technique presented; it will be introduced at the end of this section. Nevertheless the sections coming first are fundament to focus the problem and make possible a real comparison. In fact, a more complex study would be useless if simpler approaches were satisfactory. We will prove that multifractal analysis is a real improvement with respect to the other techniques considered.

3.3.1 Statistical Characterization

Classical Statistics

Classical statistics are time-domain measures of heart rate variability, commonly employed in medical researches. As described in section 2.1.1, the Task Force [1996] suggested, among the huge number of possible time-domain indexes, a set of four measures, that highly summarize all the others.

We computed the values of SDNN, RMSSD, SDANN and HTI for each recording in the database. In table 3.4 are reported the mean values and the standard deviations. The value of the mean RR interval, μNN, is included as a reference, but it can not be considered relevant as it depends largely on parameters which are external to the current analysis (age, fitness, ...).

The estimation of the probability density of the RR intervals is necessary to compute the "Heart Triangular Index". The task was performed by means of normalized histograms (with a bin width of $3 \cdot 7.8125$ms). In figure 3.1 the mean densities for the five populations are presented. The HTI index, inversely proportional to the peak of the density, in average decreases from NR to TR subjects.

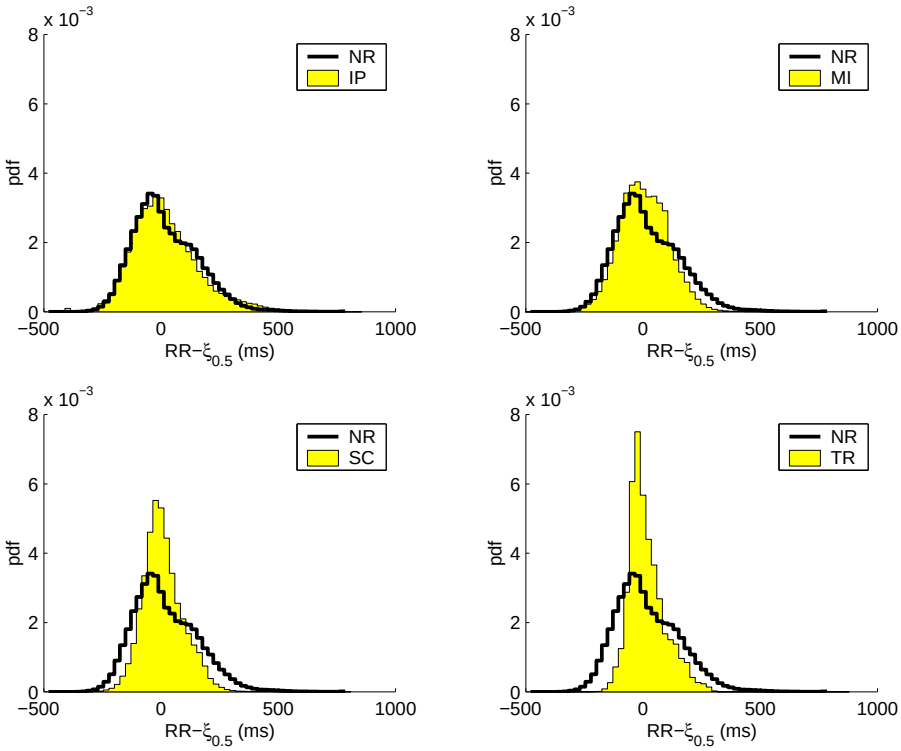


Fig. 3.1: *For each RR recording in the* NOLTISALIS *database, a normalized histogram has been computed using a bin width of* $3 \cdot 7.8125$*ms (see caption of figure 2.1 for further details). Each density was, firstly, re-scaled by the median value* ($RR^* = RR - \xi_{0.5}$) *and then a "mean" histogram was computed averaging over the densities within each populations. In the four panels the NR mean distribution is compared with the four pathological ones. The NR histogram (thick line) displays a slight bimodal behaviour, produced by the day-night variability. IP subjects are practically undistinguishable from healthy ones. MI, SC and TR densities's width decrease with a consequent peaks increase.*

The Distribution of $RR_{i+1} - RR_i$

The RR series can not be considered stationary; for this reason it is a common choice to study the difference signal[5]

$$I_i^{\Delta} = RR_{i+\Delta} - RR_i,$$

which is, by construction, *nearly* stationary.

[5] The difference signal is commonly considered; ad example, RMSSD, a classical index introduced in the previous section, is just the standard deviation of I^1.

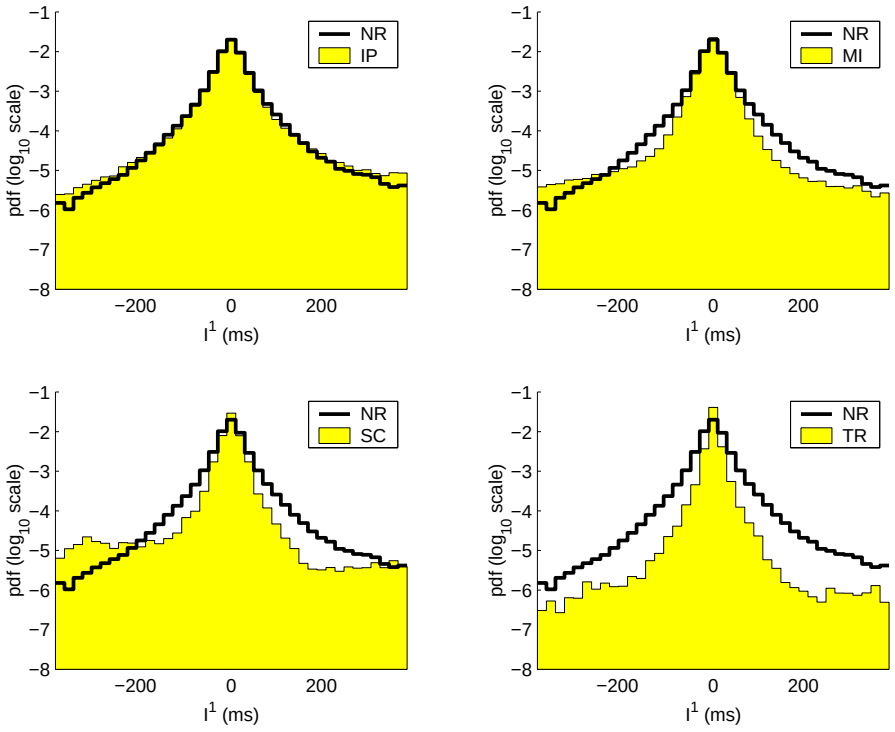


Fig. 3.2: *For each RR recording in the* NOLTISALIS *database, a normalized histogram of $I_i^1 = RR_{i+1} - RR_i$ has been computed using 20 ms bins from the raw data (no artifacts nor ectopic beats corrected). Within the same population, all the histograms have been averaged, producing a mean distribution. In the four panels, the mean normal distribution is compared with the four pathological ones. The asymmetry in SC is due to the high number of ectopic beats present in the series (see table 3.3).*

As described in section 2.1.3, the probability distribution of I_i^1 is well approximated by a Lévy stable function.

To verify if this approximation holds for the series in the NOLTISALIS database, we fitted the distribution

$$f_{I^1}(x, \alpha, \gamma) = \frac{1}{\pi} \int_0^{+\infty} e^{-\gamma q^\alpha} \cos(qx) dq,$$

(see also equation (2.1)) to an estimated probability density. We adopted, as statistic of the distribution, the normalized histogram $h_{I^1}(x)$ computed using a bin width dependent on the ECG sampling frequency[6].

[6] For populations MR, MI and SC_A we used a bin width of 4 ms; for IP series, 6 ms;

The fitting of the distribution was obtained minimizing the error $\xi^2(\alpha, \gamma)$, defined in equation (2.3), over the interval $I^1 \in (-200, 200)$. ξ^2 was preferred, over ϵ^2, to ensure a satisfactory match along the tails of the function. A non-linear least squares algorithm has been employed; more details on the numerical technique can be found in section 2.1.3. The fitting, verified by visual inspection, is generally good (see, for example, figure 2.3).

To assess the statistical significance of the match, two tests are available: the *chi-square test*, for binned distribution, and the *Kolmogorov-Smirnov test* for continuous data [Mood *et al.*, 1988].

Unfortunately, it is difficult to employ the Kolmogorov-Smirnov test when the supposed distribution is characterized by parameters depending on the data set. In fact, no analytical estimates of the distribution function of the KS statistics are available and Monte Carlo techniques become necessary [Press *et al.*, 1992, chap. 14.3].

Thus, we applied a chi-square test. The test is *not* significant for any series in the database. This is not surprising: the Lévy stable is only a *model*, and deviation from the model can be expected with such large data sets (each sequence contains nearly 100000 RR intervals). Nevertheless, it is positive that the values of the χ^2 statistic are, in average, 20 times smaller than those obtained when fitting a normal distribution: a Lèvy stable function represents the statistical properties of the interval differences I^1, better than a normal distribution.

In the Lévy stable distribution, the parameter γ is fundamentally a scale factor; the *dispersion* $\sigma_L = [2\gamma]^{(1/\alpha)}$ (see appendix A) depends on γ, at constant α levels. On the other hand, α controls the shape of the distribution. For $\alpha = 2$ we have a normal distribution, but for $0 < \alpha < 2$ the tails decay slower (*heavy tails*). In figure 3.3 a comparison is established among the α parameters in the five populations. Out-layers are present in at least three populations (NR, IP and SC), but the remaining patients have $1.55 < \alpha < 1.98$ with a mean value of approximately 1.83. With a t-test, it has been verified[7] that it is not possible to reject the null hypothesis that the α values were extracted from distributions with the same mean.

We can conclude that the Lévy stable is a kind of "universal" distribution of the difference signal I^1, and it doesn't change under pathological cardiac conditions. Similar results were previously observed also by Peng

for SC_B and TR_A ($f_c = 128$), 1000/128 ms and, for TR_B ($f_c = 1024$), 2 ms.

[7] Using a t-test, each pathological population has been compared with the healthy one. The null hypothesis was that the means (of the α parameters) were undistinguishable. (‡) a t-test, making the assumption that the standard deviations were different, was used.

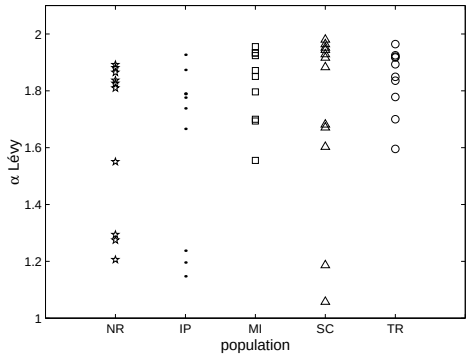


Fig. 3.3: *For each sequence in the database a Lévy stable distribution has been fitted on the normalized histograms of the variable I^1. The parameters α, obtained for each series, are compared in figure. No differences arise between recording coming from normal and pathological subjects.*

et al. [1993], but only by comparing normal and suffering from dilated cardiomyopathy patients.

By now, we just investigated the statistical characteristics of the difference variable I^1. Moreover, the approximation is very robust, in fact we also verify that long tails are present in $I^\Delta = RR_{i+\Delta} - RR_i$ with $\Delta \neq 1$. The Δ evolutions, of the variables α, γ and ξ^2, are displayed in figure 3.4 for a normal subject. For $\approx 16 < \Delta <\approx 1000$, parameters values can be considered constant. Similar patterns were found for nearly all the series in the database (in 40 over 50 sequences).

3.3.2 Monofractal Approach

We considered as "monofractal", methods that provide for estimation of the scaling exponents H, directly or indirectly. The three techniques described in this section are all meant to quantify the same characteristic, the *self-affinity* in the signal, despite through different approaches (a self-affine signal is, by definition, fractal; see section 2.3.1).

In first place, the scaling exponent H can be computed directly on the

NR vs/	F-test	t-test
IP	0.4376	0.8239
MI	0.0171	0.0978
SC	0.0262	‡0.0576
TR	0.0069	‡0.0699

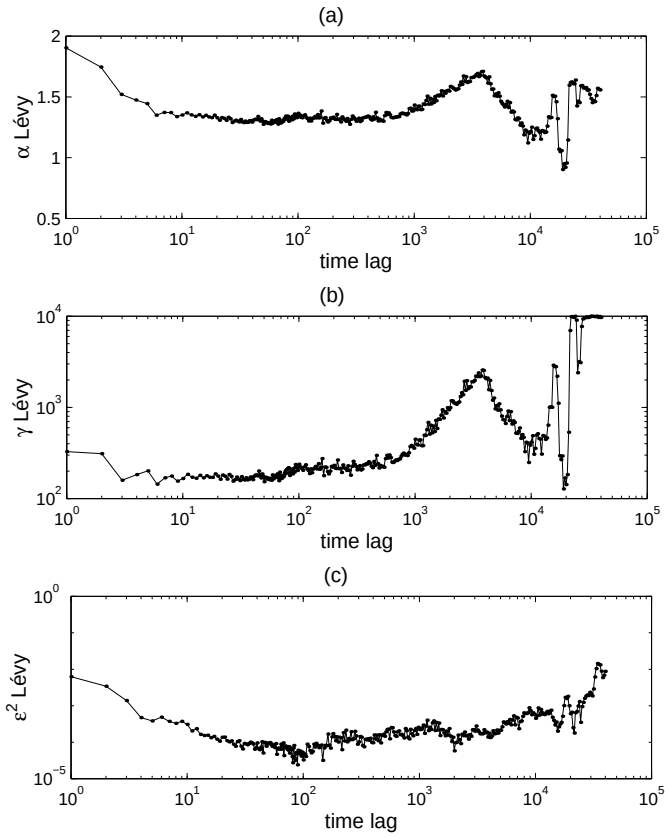


Fig. 3.4: NOLTISALIS *database, subject* NR 1. *A Lévy stable distribution has been fitted onto the data* $I_i^{\Delta} = RR_{i+\Delta} - RR_i$ *where* Δ *is a fixed time lag. This has been repeated for several values of the time lag. In figure are plotted the evolution of* α *(*panel a*)* γ *(*panel b*), the two parameters characterizing the distribution. In panel* c *the mean square fitting error* $\xi^2(\alpha,\gamma) = \sum[(\log_{10} h_X(I_k))/(1+\sqrt{|I_k|}) - (\log_{10} f_X(I_k,\alpha,\gamma))(1+\sqrt{|I_k|})]^2$ *is reported.* h_I *is the normalized histogram computed onto the* I *series and* f_I *a Levy stable distribution.*

RR series. From (2.11),

$$\langle |RR_{n+\Delta} - RR_n| \rangle = \Delta^H \langle |RR_{n+1} - RR_n| \rangle;$$

thus a graph of $\langle |RR_{n+\Delta} - RR_n| \rangle$ versus Δ on a log-log plot is a straight line whose slope is the value of H. We'll use an extension of this methods to calculate the full multifractal spectrum in the next section; we postpone there any further discussion.

The slope α, of the power spectrum on a log-log plot in the lower frequen-

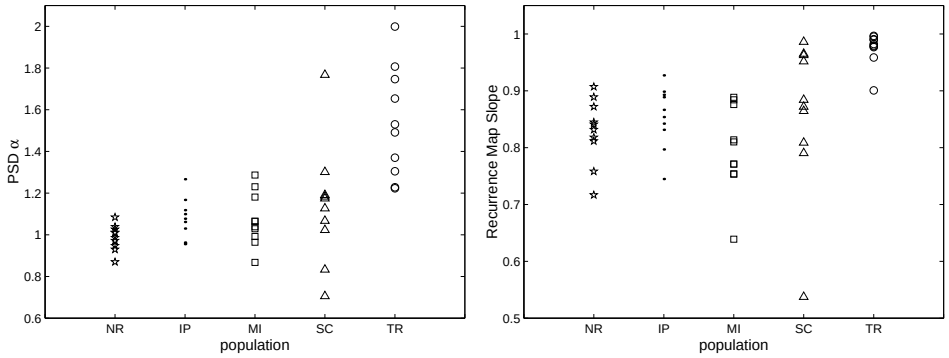


Fig. 3.5: *Panel (a): in the frequencies range* $(0, 0.04]$ *Hzeq, the model* $\log S(f) = \beta - \alpha \log f$ *was fitted on the power spectrum* $S(f)$*. The values of* α *for the five* NOLTISALIS *populations are compared. Panel (b): the model* $RR_{i+1} = \alpha_S RR_i + \beta$ *is fitted with a standard linear regression technique onto the variables* (RR_{i+1}, RR_i)*. The values of* α_S *for the five* NOLTISALIS *populations are compared.*

cies, is also an hallmark of self-affinity. We found that the RR series in the database display a very good scaling in the frequencies range $(0, 0.04]$ Hzeq. The model $\log S(f) = \beta - \alpha \log f$ has been fitted on the power spectrum $S(f)$. In figure 3.5 panel (a), the α values obtained for the five NOLTISALIS populations are displayed; IR and TR are significantly different (t-test)[8] from healthy subjects.

A third approach is based on classical random-walk analysis. Introduced by Peng *et al.* [1995], the method is named "detrended fluctuation analysis" (DFA). The computational technique is complex and it is described in section 2.3.2; it is enough to say that it provides a couple of scaling exponent ν_1 and ν_2 which should be robust against the to non-stationarity in the RR data.

The first, ν_1 (called "short-term fractal exponent"), is computed on short scales ($\Delta < 16$); the second, ν_2 ("long-term fractal exponent"), expresses the

[8] Using a t-test, each pathological population has been compared with the healthy one. The null hypothesis was that the means (of the α slopes) were undistinguishable. (‡) a t-test, making the assumption that the standard deviations were different, was used.

NR vs/	F-test	t-test
IP	0.0785	0.0398
MI	0.0197	‡0.0825
SC	4.63e-5	‡0.1374
TR	8.87e-5	‡1.22e-4

scaling on longer scales ($100 \leq \Delta < 4000$); it is related to the slope of the power spectrum by the relation $\alpha = 2\nu_2 - 1$.

We computed both exponents on the entire database. Results are reported in tables 3.4 and 3.6.

3.3.3 Multifractal Approach

The computation of the multifractal spectrum of a time series is generally performed through a box-counting method. In section 2.3.3 we described how the technique can lead to misleading results. Thus, we preferred to follow, for the work on the NOLTISALIS series, the generalized structure function approach. WTMM, a wavelet based technique, that should decrease non-stationarity effects, was applied too.

Generalized Structure Functions

Generalized structure functions, described in section 2.3.3,

$$\mathrm{GSF}(\Delta, q) \equiv \langle |X_{n+\Delta} - X_n|^q \rangle$$

were computed for each RR series in the database, with indexes $q = 1, \cdots, 10$ and[9] $\Delta < N/2$. GSF for a 38-years old healthy subject are displayed in figure 3.6 panel (a).

Typically three different regions can be recognized: (i) a steep growth for the smallest time lags ($\approx\leq 16$); (ii) a good linear scaling region at intermediate scales ($\Delta \in [100; 5000]$) and (iii) a third region at large scales where the stationarity hypothesis breaks down, the behaviour gets unpredictable and stops being interesting ($\Delta > 5000$).

Two hypotheses may be formulated to explain the existence of the first region. In first place the quantization error induced on the RR series by the finite sampling frequency of the recording Holter device can be responsible for modulation of the GSF for the smallest Δ values[10]. Additionally, it is possible that at scales $\Delta < 16$, the HRV modulating systems is not strong enough to prevent the RR series from randomly walking. In fact, the random walk $RR_{i+1} = RR_i + w_{i+1}$ would produce GSF proportional to $\Delta^{1/2}$.

[9] The Δ-axis was constructed with: (i) $\Delta = \{1, 2, \cdots, 100\}$ (ii) 90 points in each subsequent decade, at approximately the same distance in logarithmic scale, up to $\Delta = N/2$, where N is the number of points in the series.

[10] A similar situation was described by Ivanov *et al.* [1999]. Unfortunately the mechanism, by itself, is not sufficient to explain the ramp for $\Delta <\approx 16$; in fact, for the noise $t_i = u_{i+1} - u_i$, $\mathrm{GSF}(\Delta, 1) \neq \mathrm{GSF}(\Delta, q)$, $\forall q \neq 1$ but for $q \geq 2$ the GSF are expected to be flat. See also section 1.5 for further details on the superimposed quantization error t_i.

Also the characteristic exponent α of the approximating Lévy distribution (figure 3.4) decreases for the smaller scales thus suggesting some changes in the modulating mechanisms.

The linear scaling region is the more important one. We expect a fractal process to present GSF linearly scaling. In this region, for Δ in the interval $[100, 5000]$, the model $\mathrm{GSF}(\Delta, q) = A\Delta^{qh_q}$ was fitted onto each structure function. Moreover a true multifractal process is characterized by decreasing scaling exponents for increasing q ($h_q < h_p$ for $q > p$). Panel (b) of figure 3.6 illustrates the h_q exponents, computed on the GSF of panel (a). They show a clear decreasing behaviour supporting the hypothesis of multi-fractality in the RR series.

Every series in the NOLTISALIS database shows generalized structure functions similar to the ones presented in figure 3.6; they all have a distinctive three regions behaviour, with good linear scaling and decreasing h_q.

To further verify the nature of the observed multi-fractality, we prepared, for each sequence, a set of 10 amplitude-adjusted fourier-transform (AAFT) surrogate data [Theiler *et al.*, 1992]. As von Hardenberg *et al.* [2000] points out, the technique is able to correctly identify the spurious origin of multifractality in most cases. The hypothesis we would like to nullify is: "the signal is generated by a linear stochastic process distorted by a nonlinear filter, expressed by monotonically increasing function". We employed the iterative technique introduced by Schreiber and Schmitz [1996] and described in section 2.6.2, using $j = 5$ iterations. The differences among the spectra of the original and surrogate series were measured by

$$\delta = \frac{\sum_{i=1}^{N}[|S_n| - |\tilde{s}_n^{(j)}|]^2}{\sum_{i=1}^{N}[|S_n|]^2},$$

where S_n and $\tilde{s}_n^{(j)}$ are, respectively, the Fourier transforms of the original signal and of the surrogate data, at the j^{th} iteration. We verified that $\delta < 1.4 \cdot 10^{-3}$ and $\mathrm{mean}(\delta) = 2.6 \cdot 10^{-5}$. Besides, the end mismatch was verified through

$$\gamma_{jump} = \frac{(RR_1 - RR_N)^2}{\sum_{i=1}^{N}(RR_i - \langle RR_i \rangle)^2},$$

to be less than $1.4 \cdot 10^{-4}$. We truncated up to 1000 points at the end of each original series (actually we did the truncation before computing *any* structure functions to avoid bias) to obtain a length suitable for fast compu-

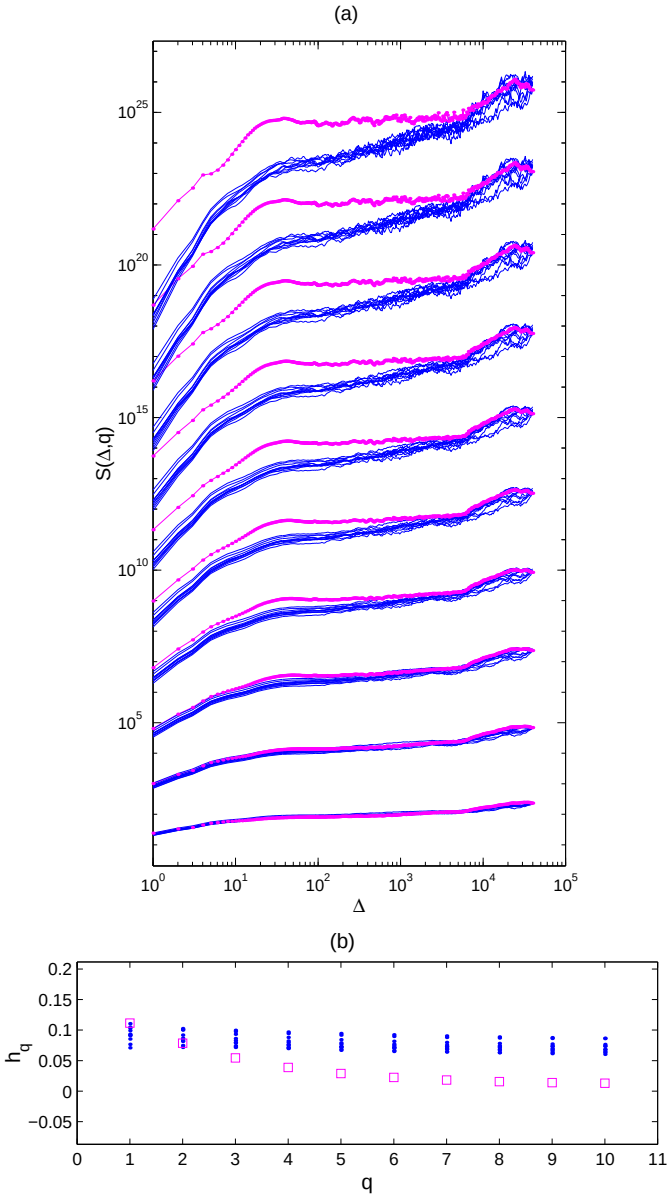


Fig. 3.6: NOLTISALIS *database, subject* NR1. *Structure functions* $S(\Delta, q)$ *for the RR series (*upper panel*) are computed for* $q = 1..10$ *(*thick lines*); structure functions for 10 surrogates data are displayed as well. The* $q = 1$ *case is on the bottom of the figure,* $q = 10$ *on the top. The model* $S_q(\Delta) = A\Delta^{qh_q}$ *is fitted onto each structure function in the range* $\Delta \in [100, 5000]$*. The values of* h_q *are displayed in the lower panel; the squares are the results obtained from the RR series; the dots come from the surrogates data.*

NORM1	✓	IPER1	✓	POST1	-	SCOM1	-	TRAP1	✓
NORM2	✓	IPER2	-	POST2	✓	SCOM2	✓	TRAP2	-
NORM3	-	IPER3	✓	POST3	✓	SCOM3	-	TRAP3	✓
NORM4	✓	IPER4	✓	POST4	-	SCOM4	✓	TRAP4	✓
NORM5	✓	IPER5	-	POST5	✓	SCOM5	✓	TRAP5	-
NORM6	✓	IPER6	✓	POST6	✓	SCOM6	✓	TRAP6	✓
NORM7	-	IPER7	✓	POST7	✓	SCOM7	-	TRAP7	✓
NORM8	-	IPER8	✓	POST8	-	SCOM8	-	TRAP8	-
NORM9	-	IPER9	✓	POST9	✓	SCOM9	✓	TRAP9	✓
NORM10	✓	IPER10	✓	POST10	✓	SCOM10	-	TRAP10	-

Tab. 3.5: *Using a statistical test, it is been evaluated the null hypothesis:* "the value of h_{10}, computed on the *RR* series, and the parameters h_{10}, extracted from the surrogates data, come from the same normal distribution"*. The significance level is* $p < 0.01$*; The check-mark means that the null hypothesis can be safely rejected.*

tation through mixed-radix-fft algorithm[11]. Generalized structure function were computed on each surrogate sequence, and $h_q = h_q^*$ values as well. In figure 3.6, those relative to the subject NR1 are reported. Surrogates GSF show a common behaviour, with constant $h_q^* \neq 0$ as expected: though having the same power spectrum of the original RR series, they lack of phase correlations and are only colored noise.

By means of a statistical test, we tried to nullify the hypothesis: "h_q and h_q^* were extracted from the same distribution, supposed normal" (see table 3.5). We select $q = 10$ (if the original series were only colored noise the choice of q wouldn't make any difference). In 32 out of the 50 cases, surrogate data are significantly different ($p < 0.01$) from the original series. The results should support the conclusion that RR series are true multifractal processes.

Unfortunately the values of the exponents h_q, we found, assume a very small absolute value (close to zero) and thus non-stationarity effects must be further investigated.

The mean values of the h_q exponents in the five populations are compared in figures 3.7 and 3.8; multifractal spectra of normal subjects display a much lower value than any other population. To assess the statistical significance of the differences in the means, we perform a t-test[12] among

[11] The actual number $N^* < N$ of points employed was chosen to fulfil the three conditions: (i) $0 \leq N - N^* \leq 1000$; (ii) $p_i \leq 17$, $\forall i$, where p_i are the prime factors composing N^*; (iii) N^* minimizes the approximate computational load (of the mixed-radix-fft) $\mathcal{L} = N^* \sum_i p_i$.

[12] Using a t-test, each pathological population has been compared with the healthy one. The null hypothesis was that the means (of each h_q exponent) were undistinguishable. (†)

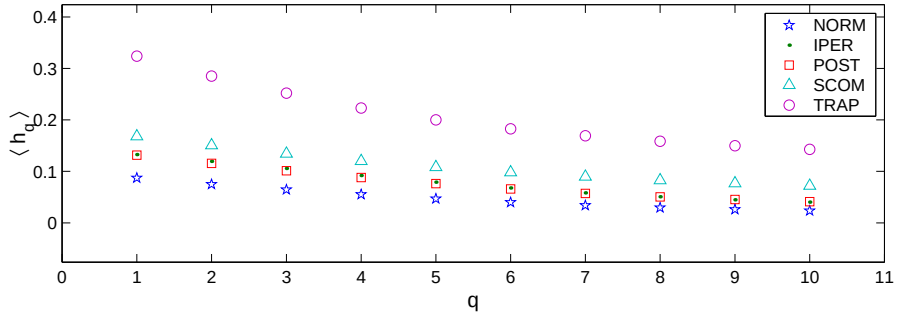


Fig. 3.7: *The mean values of h_q, slopes of the structure functions in a log-log plot, have been computed among the five different populations; they are compared in the plot.*

the h_q values of healthy and pathological subjects. For small q values, h_q resulted very effective in distinguish among NR and the other populations. h_1, which matches the H scaling exponent (often refer to as the "Hurst" exponent, see section 2.3.1) proved to be the only index (among all we tested in our analysis) able to statistically discriminate among NR and *any* other population.

Wavelet-Transform Modulus-Maxima Method

Non-stationarity effects may highly influence the values of the scaling exponents. Recently, a new approach, wavelet-transform modulus-maxima, has been introduced (see section 2.3.4). In fact, wavelet transform can remove polynomial trends.

To further limit non-stationarity, we considered only six night hours (from 24:00 to 6:00) from each RR series: during the night environmen-

a t-test, making the assumption that the standard deviations were equal, was used.

NR vs/	IP	MI	SC	TR
h_1	†0.005	0.049	3.7e-4	1.3e-5
h_2	0.008	0.074	3.0e-4	6.1e-5
h_3	0.011	0.105	4.8e-4	1.8e-4
h_4	0.016	0.138	7.0e-4	5.0e-4
h_5	0.023	0.176	0.001	0.001
h_6	0.040	0.222	0.001	0.002
h_7	†0.065	0.267	0.002	0.004
h_8	†0.107	0.308	0.004	0.006
h_9	†0.153	0.342	0.007	0.008
h_{10}	†0.196	0.374	0.010	0.010

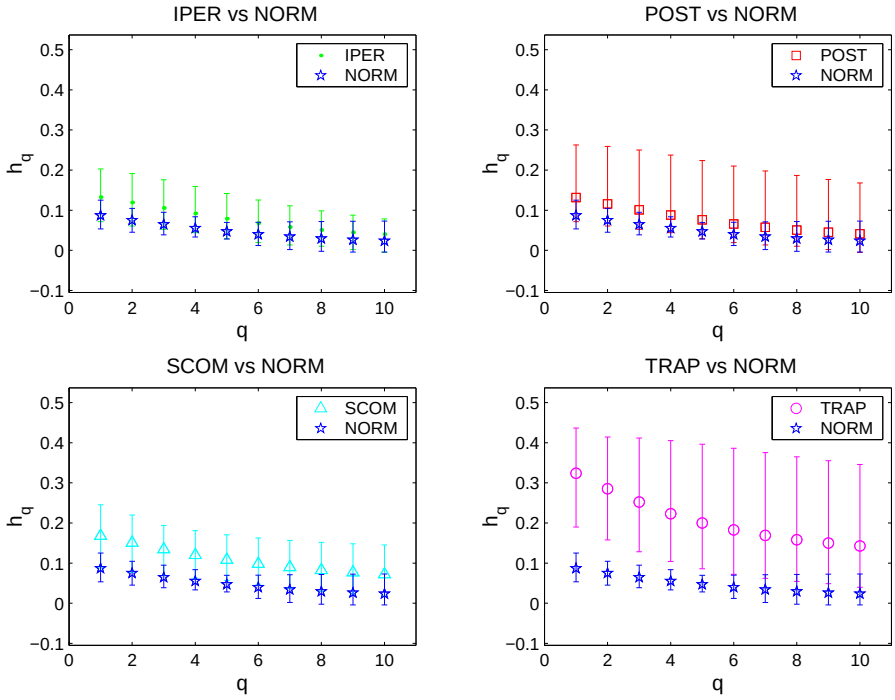


Fig. 3.8: *Mean values of h_q as in figure 3.7. Each panel compares the values obtained from a pathological population with those coming from the normal group. The error bar displayed are the ranges (minimum–maximum value) of each h_q group.*

tal stimuli should be highly reduced. The wavelet transform, for a discrete series RR_i, is defined by

$$T_\psi[RR_i](b,a) = \frac{1}{a}\sum_i \psi\left(\frac{i-b}{a}\right) RR_i,$$

where $a > 0$ is a scale parameter and b a space position. At fixed scale a, $T_\psi[X_i](b,a)$ is essentially obtained through a convolution, among the analyzing wavelet and the sequence RR_i, and can be quickly computed by means of Fourier transforms.

The analyzing wavelet, ψ, used here was the third derivatives of the Gaussian function, $g(x) = \exp(-0.5x^2)$,

$$\psi(x) = -\frac{2\sqrt{30}}{15\sqrt[4]{\pi}}\,(t^3 - 3t)e^{-1/2t^2}$$

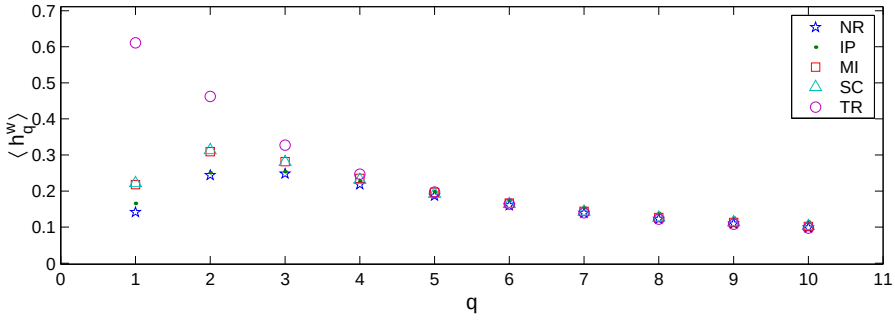


Fig. 3.9: *The mean values of $h_q^w = [\tau(q) + 1]/q$, GSF scaling exponents equivalent, obtained through the WTMM method, have been computed among the five* NOLTISALIS *populations. See also figure 3.7.*

which was normalized to have unit 2-norm[13], on the finite support $x \in [-5000, 5000]$. We considered scales $a = 2 \cdot 1.15^i$, with $i = 0, 1, \cdots, 41$.

Even at the largest scale considered, the value of ψ, evaluated in $5000/a$, is smaller than $1 \cdot 10^{-11}$. The wavelet can be considered zero out of its support thus enabling the use of a fft algorithm with zero-padding for the convolution.

The partition function $Z(a, q)$ was computed for $q = 1, 2, \cdots, 10$. At the scales of interest, the wavelet transform $T_\psi[RR_i](b, a)$ is a smooth function in b; local maxima have been localized simply by thresholds on numerical derivatives of T_ψ. Each $\tau(q)$ exponent was computed as the slope of the least square line in a log-log graph of $Z(a, q)$ toward a, with $a \in (16, 700)$.

Equivalent values of the GSF scaling exponents, $h_q = h_q^w$, were computed through the relation $h_q^w = [\tau(q) + 1]/q$. Mean h_q^w values are presented in figure 3.9. If they are compared to h_q exponent, they show a different behaviour, with a starting growing trend (with the exception of TR subjects).

Mean $\tau(q)$ values (and equivalently mean h_q^w values) of each pathological population have been compared with those of the healthy subjects. Unfortunately, except from TR patients, no pathological population can be significantly distinguished from the NR ones by means of $\tau(q)$ indexes. See tables 3.6 and 3.6.

[13] The 2-norm of the function $f(t)$ is defined as

$$\|f(t)\|_2 = \left[\int_{-\infty}^{\infty} f^2(t)dt\right]^{\frac{1}{2}}.$$

3.3.4 Other Techniques

Recurrence Maps

A two-dimensional recurrence map has been constructed from the RR data of each subject. The model $RR_{i+\Delta} = \alpha RR_i + \beta$ was then fitted with a standard linear regression technique to the variables $(RR_{i+\Delta}, RR_i)$. Δ was chosen to be the first numerical minimum of the autocorrelation function and it is different from 1. In fact the high correlation of RR_{i+1} with RR_i would have masked any other information.

The α_S parameters, which are the slope of the regression lines computed for each populations in the NOLTISALIS database, are displayed in figure 3.5, panel (b). Significant differences (t-test)[14] arise only among normal and transplanted patients.

Approximate Entropy

Approximate Entropy (ApEn) has been successfully employed in a series of discrimination tasks on the RR signal (see for example [Fabani and Sassi, 1996] and references therein).

Two values of ApEn have been computed for each series in the NOLTISALIS database: (i) a first index was computed on six night hours (from 0:00 to 6:00, approximatively 22000 consecutive points); (ii) a second was calculated on each available day hour (approximatively 4400 points each), in the interval 12:00-18:00; then the values have been averaged. The choice of the parameters m and r is critical; we followed the suggestions in [Fabani and Sassi, 1996], setting $m = 2$ and $r = 0.2 \cdot \text{SDNN}$.

Unfortunately the ApEn values computed on the five populations are nearly identical and no significance distinctions can be made. The only exception is between the night-ApEn of normal and transplanted patients. The last ones have smaller approximate entropy and therefore lower heart rate variability. Numerical results are reported in tables 3.4 and 3.6.

[14] Using a t-test, each pathological population has been compared with the healthy one. The null hypothesis was that the mean values of the α slopes were undistinguishable. The double dagger (‡) indicates that it was used a t-test with the assumption that the standard deviations were different.

NR vs/	F-test	t-test
IP	0.0226	‡0.1540
MI	0.0154	‡0.2481
SC	0.0026	‡0.1118
TR	0.0074	‡0.0188

3.4 Discussion

At the begin of this chapter we stated two objectives we would have liked to pursue during the analysis of the NOLTISALIS database.

Firstly, the attention was on gaining new knowledge on the structure of the long RR series and thus on the behaviour of the cardiovascular system over large time scales. The results we obtained on this point concern, mainly, the statistical description of the variable $I^{\Delta} = RR_{i+\Delta} - RR_i$ and the multifractal characteristics of the RR series.

We verified (section 3.3.1) that the normalized histogram of the intervals difference variable I_i^1, is well approximated by a Lévy stable distribution. The approximation holds as well for differences performed on larger scales (approximately $\Delta < 1000$), for the variables I_i^{Δ}. The result may appear surprising because non Gaussian Lévy stable distribution ($\alpha \neq 2$) have infinite second moments. However this is not in contradiction with any physiological or physical knowledge. Statistical distributions with power-law tails are often present in open complex system in which the action of any part influences the whole system, as may be the case in the heart rate regulation system. In the RR series, long tails may be generated by the competition among the sympathetic and parasympathetic systems, that try to increase and decrease, respectively, the sinoatrial pace maker frequency [Penna *et al.*, 1995].

No significative differences arise among the histograms of I^{Δ} of healthy and diseased subjects. All of them are well fitted by a Lévy stable distribution, with approximatively a value of $\alpha \approx 1.83$. The different scaling pattern we found, for example with the multifractal analysis, must relate to the ordering and correlation of the I^{Δ} variables.

We also verified (section 3.3.3) that the multifractal spectrum of the exponents h_q (computed through the structure functions method) show a decreasing course, for increasing q. This behaviour is often the hallmark of multifractal processes and might be associated with the presence of nonlinear correlations in the signal, and possibly, with the nonlinear nature of the HRV signal. The chance of spurious multi-fractality, due only to nonlinearly filtered colored noise, was investigated through surrogate data technique (amplitude-adjusted fourier transform). 32 over 50 h_q spectra were statistically different ($p < 0.01$) from their surrogate. Both this findings would support the thesis of multifractal RR series.

Unfortunately, the values we obtained for the h_q exponents were small (approximately < 0.12 for normal subjects). Non-stationarity effects may

have influenced the RR series much more then expected, thus producing decreasing trends in the multifractal spectrum.

To further investigate the non-stationarity influence, we also applied the wavelet-transform modulus-maxima method. The method should eliminate polynomial trends, up to the third order, from the data. The h_q^w equivalent multifractal spectrum shows a distinctive behaviour, with a first growing tendency before decreasing, eventually. The exponents values are slightly increased, compared to the structure function method.

The different techniques employed, though supporting the hypothesis of dynamical nonlinearities in the RR series, do not provide a definitive answer to the debate on the nonlinear nature of the interbeat series. Besides, the question is really complex for sure, but we believe that the results presented in the chapter might be useful to proceed in the analysis of the problem.

The second aim of the work was that of testing statistical indexes, able to significantly discriminate pathological patients from healthy subjects.

Each multifractal exponent has been compared along the whole database. For every population, a mean value was computed and then we verified the null hypothesis that means of normal and diseased patients were undistinguishable by means of a t-test. h_q proved to be very effective in this discrimination task, in particular with $q = 1, 2$ and 3.

A large number of classical parameters, for the analysis of the HRV signal over long time scales, have been considered to set up a proper comparison. We considered classical time-domain indexes (section 3.3.1), "monofractal" characteristics ($1/f^\alpha$ spectrum; detrended fluctuation analysis; see section 3.3.2) and a regularity statistic (approximate entropy; see section 3.3.4). They are all reported in tables 3.4 and 3.6. The multifractal exponents h_q were the best performing: $h_1 \equiv H$ was the only index, among those considered, able to distinguish all the four pathologic populations from the normal one. On the contrary, $\tau(q)$ WTMM exponents were quite disappointing.

A possible diagnostic application of the described multifractal indexes, might be obtained in a multivariate contest, putting them together with classical time-domain indexes. The use of a pool of statistics is taken into account in the next chapter. In fact, the computation of discriminating statistics requires a very large number of RR series.

NORM vs/	IP	MI	SC	TR
μNN	✓	-	-	✓
SDNN	-	✓✓	✓✓✓	✓✓✓
RMSSD	-	-	✓✓	✓✓✓
SDANN	-	✓✓	✓✓✓	✓✓✓
HTI	-	✓✓	✓✓	✓✓✓
α Lévy	-	-	-	-
α spectrum	✓	-	-	✓✓✓
ν_1 DFA	-	-	-	✓✓✓
ν_2 DFA	✓✓✓	-	-	✓✓✓
$h_1 \equiv$ H	✓✓	✓	✓✓✓	✓✓✓
h_2	✓✓	-	✓✓✓	✓✓✓
h_3	✓	-	✓✓✓	✓✓✓
$h_{\mathcal{R}}$	-	-	-	✓✓
τ_1	-	-	-	✓✓✓
τ_3	-	-	-	✓✓✓
$\tau_{\mathcal{R}}$	-	-	-	✓✓✓
α_S recurrence	-	-	-	✓✓✓
ApEn night	-	-	-	✓
ApEn day	-	-	-	-

Tab. 3.6: *Several parameters computed all along the chapter are compared. For each of them the ability to discriminate a pathological population from the healthy one is evaluated by means of a t-test. First of all, a F-test was performed to assess differences among the variances of the groups ($p < 0.05$) and then the proper t-test was employed. The null hypothesis was "the two populations (supposed sampled from a normal distribution) have the same mean". In the table the significance levels are reported: ✓✓✓ ($p < 0.001$), ✓✓ ($p < 0.01$) and ✓ ($p < 0.05$). An empty line divides different groups of parameters: (from the top) (i) statistical & geometrical, (ii) fractal, (iii) multifractal and (iv) various. Only the scaling parameter H is able to discriminate all the four pathologic populations. "μNN" is the mean RR interval; "SDNN", "RMSSD", "SDANN" and "HTI" are the four classical time-domain measures, selected by Task Force [1996]; "α Lévy" is the characteristic exponent for the Lévy distribution that best fits the I^1 series; "α" is the Power spectrum slope (log-log graph) in the lower frequencies; "ν_1" & "ν_2" are the DFA fractal exponents; "h_q" are the q^{th} scaling exponents obtained through GSF approach; "$h_{\mathcal{R}}$" $= max_q[h_q] - min_q[h_q]$; "$\tau_q$" are the exponents $\tau(q)$ obtained via WTMM method; "$\tau_{\mathcal{R}}$" $= max_q[\tau(q)] - min_q[\tau(q)]$; "$\alpha_S$" is the slope of the least-square line fitted to the spectrogram of the variable $(RR_{i+\Delta}, RR_i)$, where Δ is the first numerical minimum of the autocorrelation function of the interbeat sequence; "ApEn night" and "ApEn day" are the approximate entropies ($m = 2$, $r = 0.2 \cdot$SDNN) computed in six night hours (from 24.00 to 6.00) and six day hours (from 12.00 to 18.00) of RR series, respectively.*

4. MULTIVARIATE FETAL HRV ANALYSIS

Abstract

We considered 362 separate cardiotocographic exams (200 obtained from fetuses regarded as "normal", at delivery, and 162 from "pathological" ones); the recordings were performed at a gestational age in the interval 28–42 weeks. Data belong to a database of 252 subjects and the quality, of each recording, was preventively verified.

After artifacts and zero data correction, from each exam, a set of 16 parameters (x) was extracted. Mantel's approach was used to calculate time domain variables, such as the baseline, accelerations, and decelerations; Power spectrum components, in LF (activity of the sympathetic nervous system), MF (fetal movements) and HF (fetal breathing) bands, were obtained through parametric analysis with autoregressive models from the fetal heart rate signals (FHR). Also approximate entropy, a regularity statistics, was considered; in fact, it is suggested that non-linear mechanism may influence the FHR generation process.

We investigated the possibility to automatically allocate, by means of a classification technique, a fetus to a fetal suffering condition (one out of a fixed number of types) according to the value of x; the approach might be a first step towards the development of a diagnosis support device. Several supervised classification systems have been applied to the parameter set: linear, quadratic and logistic discriminant analysis, k-nearest neighbour classifiers and multi-layers feed-forward neural networks.

Only feed forward neural networks were able to distinguish normal and pathological fetuses: the method correctly classified 80% of the exams. The applied technique, though needing further work and tuning, is promising. The multivariate approach proved to be very effective.

4.1 Introduction

The knowledge of fetus health conditions during pregnancy is of great importance to prevent dangerous events. In fact, serious malformations, irreversible neurological damages and the death of the fetus can be avoided by timely diagnosing some growth and nutrition alterations that induce fetal suffering conditions. Nevertheless, the early identification of many among these disease states is very difficult to obtain. In this case the knowledge of data on the fetal health conditions can strongly help the diagnostic process. Moreover, the fragile structure of the fetus requires the use of non-invasive methods to extract the necessary data. Cardiotocography (CTG) is the clinical, traditional, noninvasive approach to monitor antepartum fetal conditions but rarely it detects emergencies of fetal pathologies. There are still many problems connected with interindividual and intraindividual variability concerning the interpretation of the curves. A number of systems to extract parameters starting from the CTG signal were developed, highly reducing intraindividual variability. The Fetal Heart Rate signal (FHR), upon which CTG is based, provides reliable indications on fetal status. In fact, several conditions such as hypoxia, acidemia and drug induction produce noticeable variations of FHR. Recent results on FHR signal demonstrated that both linear and non-linear mechanism are involved in the signal generation suggesting the usefulness of a multivariate approach [Task Force, 1996].

In this chapter we propose to classify FHR signals through a set of indexes including (i) time domain, (ii) frequency domain and (iii) regularity (approximate entropy) parameters. This set is used as the input of an automatic system, whose goal is to detect the risk for the fetus to enter a pathological state.

In the work we focused on four potential pathological states: (i) nutrition alterations caused by maternal hypertension (H), (ii) intra-uterine growth retardation (IUGR), (iii) nutrition alterations caused by maternal diabetes (DG), and (iv) fetal macrosomia (MACRO). States (ii) and (iv) are defined at delivery, comparing the actual weight of the newborn baby (W) with standard weights tables (obtained by averaging over huge populations of infants); state (ii) is characterized by $p(W) < 10\%$ and (iv) by $p(W) > 90\%$, where $p(W)$ is the probability of a given weight W. On the contrary states (i) and (iii) are *potentially* dangerous situations, in which a special attention is necessary during pregnancy, as they can lead at delivery to states (ii) and (iv), respectively. Both maternal hypertension and diabetes are pathologies induced in the mother by pregnancy; the generating mechanisms are widely discussed and, yet, no general explanation has been proposed.

Usually specific tests, other than cardiotocography, are available to detect the two maternal pathological states; nevertheless, being able to localize these problems with a widely employed technique such as CTG would be of immediate impact in reducing the time of the diagnosis process.

4.2 A Short Fetal Monitoring History

Until[1] the second half of the 20th century, assessment of the fetal condition depended on very limited means: the growth of the uterus and its contents, the movements of the fetus perceived by the mother and the listening of the fetal heart beat with a mono or binaural stethoscope.

Obstetric auscultation was mentioned as early as the eighteenth century; in 1821, a Parisian doctor, Jean-Alexandra La Jumeau, was the first person to demonstrate fetal heart tone auscultation in public. Eventually, fetal heart tone auscultation became a standard for fetal surveillance until late 1960's. In 1961, the 12th edition of Williams' Obstetrics stated, "The most characteristic sign (of fetal distress) is afforded by change in the fetal pulse rate. [...] For practical purpose it is well to assume that a pulse rate of 100 or less for any great length of time is incompatible with life of fetus, and under such circumstances delivery is indicated, provided it can be accomplished without risk to the mother". The experience of many obstetricians indicated that the statement was not always true. In 1968, Benson published a collaborative study on the value of fetal heart tone auscultation, and concluded that there was no reliable simple auscultatory indicator of fetal disease in term of fetal heart rate except in the extreme degree.

Sudden absence of fetal movements in the second half of pregnancy was a serious diagnostic problem. Usually one had to wait for some weeks in order to observe if the uterus would grow before the decision could be reached to induce labour. The sign of Spalding at X-ray examination, showing overlapping of the fetal skull bones like roof tiles as a result of advanced maceration of the fetal tissues, was one of the few helpful objective diagnostic signs of fetal demise.

This recurrent dilemma, whether or not the fetus had died in utero, formed the major impulse for the development of cardiotocography. Since the recording of fetal electrocardiograms through abdominal wall electrodes was reported by Cremer in 1906, there have been attempts to apply this technique in fetal surveillance, but the results have been limited. Initially

[1] The section has been adapted from [van Geijn, 1998] and [Yeh, 1998]

fetal abdominal electrocardiography and phonocardiotocography were pursued, but failed, primarily due to technical problems.

It was only at the end of the 60's, when the fetal heartbeat could be rather easily detected by means of ultrasound (the Doppler-shift) or through the application of direct electrocardiography, that cardiotocography became popular as the method to monitor the condition of the fetus. This new modality provides not only continuous heart rate information, but also fetal heart rate changes in response to uterine contractions. Currently the majority of obstetric decisions to assist delivery of the baby by artificial means (caesarean section, forceps or vacuum extraction) for reasons of suspected fetal distress, relies on information gathered through the application of cardiotocography.

It is the obstetrician's reassurance that if the fetal heart rate (FHR) pattern is normal then there is the nearly 100% certainty that the fetus is in a good condition, which has made cardiotocography so attractive and has induced its widespread use.

4.3 Fetal Heart Rate Signal

4.3.1 Background

There[2] are three signal sources available when monitoring the FHR: (i) electrical impulses which are obtained by internal monitoring via Direct ECG (DECG) or externally via Abdominal ECG (AECG); (ii) movement of the heart which is externally detected via the ultrasound (US)-Doppler method; (iii) sound, obtained by phonographic methods (Phonocardiography).

FHR monitoring via DECG is the most accurate method for obtaining beat-to-beat FHR information, as the beat-to-beat measurements are calculated from the electrical signals of the fetal heart. Therefore, the closer FHR recorded by ultrasound gets to the DECG measurements, the more accurate the measurement is.

The Ultrasound-Doppler Method

In the earlier 1970's, the ultrasound continuous wave (CW) Doppler introduced a new method in FHR monitoring. A few years later, due to an improved signal-to-noise ratio and the possibility to detect signals from a defined depth, the ultrasound pulsed-Doppler method was introduced.

[2] A part of this section was adapted from: Hewlett Packard, B oblingen, Application Note 5964-1562E, 1995

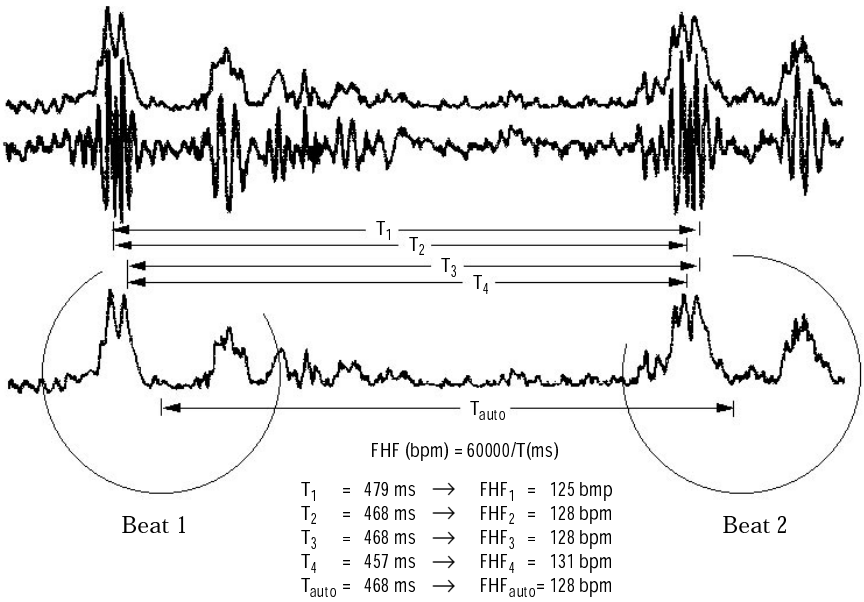


Fig. 4.1: *Envelope Doppler-signal (upper trace), derived from the demodulated Doppler-signal (second trace and lower trace) four possible heart periods T1–T4 , derived by peak detection. Symbolized by the two circles the heart rate interval is calculated by autocorrelation of the waveform complexes. (Hewlett Packard, Böblingen, Application Note 5964–1562E, 1995).*

Using this method, a pulse typically consisting of about 100 cycles of 1 MHz is transmitted towards the fetal heart. Then the ultrasound crystal is electrically switched to the receiving mode. The reflected pulse, slightly shifted in frequency (Doppler shift) by contractions of the fetal heart, will be compared with the transmitted pulse (demodulation). Figure 4.1 shows the demodulated signal (second trace from top) and the corresponding envelope signal (first and third trace). In earlier fetal monitors, beat-to-beat fetal heart rate was calculated using peak detection (time T from highest peak in the first heart beat to highest peak in second heart beat). Since these two heart beats in figure have similar double peaks, up to four different time intervals T1 to T4 and, consequently, four different FHR values are possible. However peak-to-peak detection lacked accuracy due to "jitter" (artificial variability) affecting the beat-to-beat FHR.

With the introduction of the autocorrelation method, successive heart signals are compared and tested for their similarity.

Thus, not one point in time within a heart action, but the complete waveform complex, symbolized by a circle in figure 4.1, is compared to the following one.

Problems of Autocorrelation

Even slight movements of the fetus, of the mother, and other sources will change the detected Doppler signal. To overcome these problems, an average buffer with several weighted heart beats is built up and the most probable heart rate will be generated. The actual FHR signal produced by the CTG device (which is printed or recorded) is obtained by sampling the buffer at a constant rate (usually 2Hz), independently from the value of the buffer itself. This means that a FHR value can be sampled twice and the resulting series displays high persistence. Thus the advantage of easily picking up fetal heart activity must be balanced with the disadvantage of a limited beat-to-beat accuracy.

4.3.2 Signal Recording & Preprocessing

Due to historical reasons, fetal CTG monitors, commercially available, display *only* the fetal heart rate expressed in number of beats per minute (bpm) and do not offer the series of interbeat intervals, usually employed in HRV analysis (see section 1.5).

If compared to standard Holter recordings, the buffering procedure highly reduces the precision of the RR sequence as generated by inverting the FHR signal (60000/FHR ms). Besides, it is possible that the CTG device erroneously locks on the slower maternal fetal beat, even the autocorrelation method, described previously, is employed. This leads to an abrupt decrease into the FHR signal and it influences the evaluation of variability indexes. Therefor, a proper artifact detection technique has to be employed. The one we developed relays on the work of van Geijn *et al.* [1980]; the main concept is that an acceleration of heart rate develops more slowly than a deceleration does, thus the limit for the acceptance of the point $S(i+1)$ differs according to whether $S(i+1)$ is smaller or greater then $S(i)$.

In details, three requirements are set up: (i) acceptance of FHR values which satisfy the criterium:

$$\frac{200S(i)}{400 - S(i)} < S(i+1) < \frac{200S(i)}{114 + 0.43S(i)}; \tag{4.1}$$

the whole series is processed many times and, at each run, the number of points rejected is counted; the process stops when in a entire run no

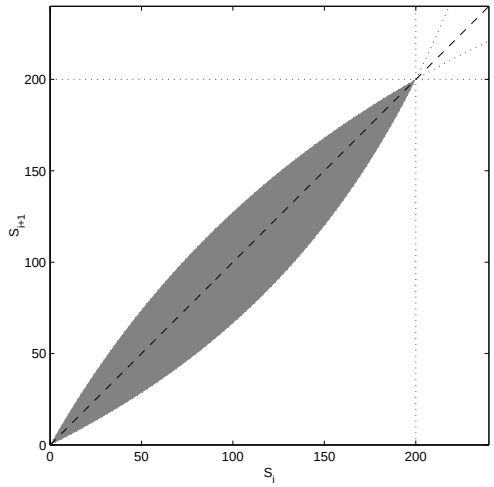


Fig. 4.2: *Grey region identifies the FHR points, corresponding to equation (4.1). They are considered invalid and marked as artifacts for successive correction. Accelerations of heart rate develop more slowly than decelerations do, thus the region below the* $S(i) = S(i+1)$ *line is bigger than the above one.*

further points were discarded[3]; (ii) a minimum of three intervals that qualify to (4.1) must be present in succession ($S(i-1)$, $S(i)$ and $S(i+1)$ for final acceptance of $S(i+1)$); (iii) short intervals of valid points, contained between invalid sequences, are rejected if their length is equal or smaller than 20 points. Figure 4.2 presents the region where FHR values are not considered acceptable according to rule (ii).

Ranges for acceptance of $S(i+1)$ are comparable with those applied in commercial monitors ($= S(i) \pm 20$ bpm). Nevertheless the applied criterium is definitely more selective and precise.

4.3.3 Baseline Detection

Interpretation of the heart rate pattern is usually performed by the physician who analyzes the deviations of the signal from an imaginary line, the *baseline*. He hypothetically constructs it as a running average of the heart rate. Accelerations and decelerations are defined as deviations from the baseline, and more than one quantitative definition is available.

[3] Already rejected points are substituted by means of linear interpolation. If a point fails to comply to (4.1) and it is already marked as invalid, a still valid point is sought forward and marked as invalid (this is particularly useful with artifact series at unusually low values of FHR).

	New	Old
Probability before peak P	0.167	1/8
Minimum length of accelerations	6 points (15s)	5 points (12.5s)
Minimum length of decelerations	12 points (30s)	5 points (12.5s)

Tab. 4.1: *Changes we made at the Mantel* et al. *[1990*a,b*] algorithms. A reference P value is computed at the beginning of the baseline computation procedure, scanning a histogram of the RR values from high to low values. The peak is located after at least 0.167 (it was 1/8) of the whole area of the histogram is scanned. This ensures, still, a baseline fit near the lower level of the basal heart rate but avoid problems with subject suffering from fetal tachycardia (or showing episodes of tachycardia, with basal heart rate highly over the average). Also, we changed the minimum length a deviation from the baseline has to display to be considered either an acceleration or a deceleration. This avoid short episodes to be regarded as meaningful deviations.*

Therefore, the analysis of the same tracing to different physicians can leads to different baseline subjective determinations, and consequently to different locations for the deviations. To make things even more complex, during accelerations and decelerations the baseline represents a theoretical frequency which might have occurred if the accelerations didn't take place; this statement is by itself difficult to translate in practical rules.

In the construction of an automated system for the evaluation of the CTG recordings, a reproducible determination of the baseline is a fundamental starting point. Several attempts in this direction have been made starting from the work of Dawes *et al.* [1982]; the approach we followed was that suggested by Mantel *et al.* [1990*a*], but we tuned the parameters of the algorithm, to make the outcome fully compliant with the opinions of a team of physicians expert on CTG analysis.

The algorithm is very complex, and a full description can be found in the cited reference. We followed completely the indications therein; the changes in the parameters values we introduced are reported in table 4.1. As a preparation step to baseline computation, after acquisition and pre-processing, FHR sequences were averaged over 2.5s periods (5 points) and decimated; in the following we'll call this signal $S_{24}(i)$ as in every minute, there are 24 points.

The changes we made to the original Mantel algorithm are meant to fix situations where, during periods of high basal fetal heart rate, the baseline was underestimated.

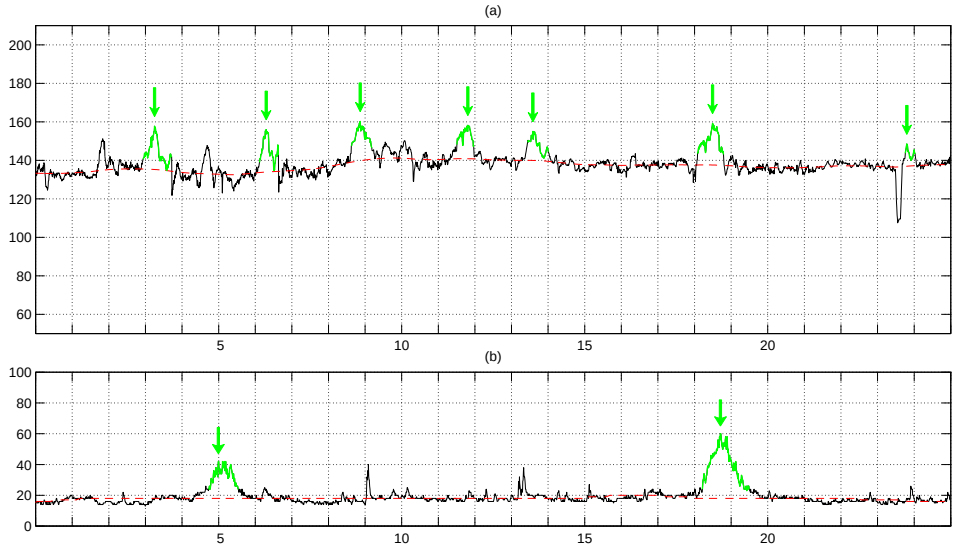


Fig. 4.3: *A 25-minutes cardiotocographic tracing is illustrated; the signals have been recorded at the* 33^{rd} *week and the baby, at delivery, has been classified as "normal". Panel* a *shows fetal heart rate in bpm; the sketched line is the baseline. Arrows mark accelerations, while decelerations are not present. Panel* b *presents the toco signal recorded on the mother: the baseline is presented as a sketched line. Arrows marks contractions.*

Accelerations & Decelerations

Accelerations and decelerations are deviations of the fetal heart rate from the baseline lasting a sufficient amount of time (accelerations are positive deviations, decelerations negative). They are correlated with the normal activities of the fetus, who "trains", moves and exercises to breath. Decelerations are unfrequent and usually correlated with uterine contraction. Since the arteries supplying the placenta follow a winding course through the complex uterine musculature, each contraction restricts the supply of blood, reducing the oxygen support , thus leading to a decrease of the heart frequency. Moreover, it can not be excluded that the autonomic control of the fetus itself, through the action of the baroceptor system, contributes to the observed decrease of the FHR.

Unfortunately, different quantifications of the words "deviations" and "sufficient" led each medical school to develop their own method to evaluate, by means of a ruler, these quantities on the monitoring strip. We tried to develop a procedure fully consistent with the definition of Mantel *et al.* [1990*b*], but also holding the suggestions present in Arduini *et al.* [1993]. At

this purpose, the minimal length necessary to qualify a deviation to be an acceleration or a deceleration has been increase as reported in table 4.1.

Also, we divided the accelerations in two populations according to their maximum deviation (L_{MAX}) over the baseline. Arduini *et al.* [1993] suggests that the two groups might have different diagnostic importance. Large accelerations have $L_{MAX} > 15$ bpm, while the small ones, $10 < L_{MAX} \leq 15$ bpm.

In figure 4.3(a), a fetal heart rate signal, recorded by means of a cardiotocographical device is illustrated. The same figure displays the baseline and a few accelerations, computed with the discussed techniques.

4.3.4 Classical Statistical Characterization

With the recognition of the significance of fetal heart rate variability, starting from the work of De Haan *et al.* [1971], several attempts were made to develop statistical indexes that quantify variabilities. Although there have been many methods proposed by various researchers, we selected a few of them, which we retained the most widely accepted. They became, over the years, integral part in fetal evaluation. They could have been introduced in chapter 2, but as long as they concern only fetal heart rate variability analysis, we preferred to postpone here their description. Classical FHR statistics are truly time-domain measures.

In the following, interbeat sequences (loosely speaking called RR series) will be used in place of heart rates; for the reasons already stated, in cardiotocography, they are computed as: $T(i) = 60000/S(i)$ ms and $T_{24}(i) = 60000/S_{24}(i)$ ms.

Long Term Irregularity

LTI, Long Term Irregularity, was the first index ever introduced; it was proposed by De Haan *et al.* [1971]. It is usually computed on a short segment of RR signal (initially 512 points for computational reasons).

Given three minutes of RR signal, $T_{24}(i)$, in ms ($i \in [1; 72]$), we defined LTI as the interquartile range $[1/4; 3/4]$ of the distribution of the modula $m_{24}(j)$, where

$$m_{24}(j) = \sqrt{T_{24}^2(j+1) + T_{24}^2(j)}. \tag{4.2}$$

We excluded from the computation large accelerations and decelerations, as suggested by Arduini *et al.* [1993], to avoid deviations from providing spurious measures of variability. The three minutes, after the removal of the

undesired parts, must contain, at least, a continuous segment of 30 seconds (12 points). Even for the computation of Delta, STV and II we used the same approach. The only exception is Delta, in this case we did not require the 30 seconds to be consecutive.

Delta

Delta is the rawest measure of variability, being define as the range of the signal in a given time interval. "Normality" of a FHR tracing, in clinical routine, is often assessed by means of the values of Delta.

Given one minute of RR signal, $T_{24}(i)$, in ms ($i \in [1;24]$), we defined *Delta* as

$$\text{Delta} = \max_i[T_{24}(i)] - \min_i[T_{24}(i)]. \tag{4.3}$$

Rarely, Delta is expressed in bpm; being $y = 60000/x$ a monotonically decreasing function, DeltaBPM can be obtained as

$$\text{DeltaBPM} = \frac{60000\ \text{Delta}}{\max_i[T_{24}(i)]\,\min_i[T_{24}(i)]}.$$

Even in this case, exclusion of large accelerations and decelerations applies, as already discussed in the LTI case.

Short Term Variability

Short Term Variability, STV, quantifies FHR variability over a very short time scale, usually on a beat to beat basis. We refer to the definitions provided by Dalton *et al.* [1977] (even if we used a scale factor of 12) and by Arduini *et al.* [1993].

By considering one minute of RR signal, $T_{24}(i)$, in ms ($i \in [1;24]$), we defined STV as

$$\text{STV} = \text{mean}[|T_{24}(i+1) - T_{24}(i)|] = \frac{\sum_{i=1}^{23} |T_{24}(i+1) - T_{24}(i)|}{23}. \tag{4.4}$$

We must underline that the definition (4.4) is also similar to that introduced by Organ *et al.* [1978] (see also [Parer *et al.*, 1985]), except, in that case STV was computed on 30 seconds of heart rate signal in bpm.

Even in this case, exclusion of large accelerations and decelerations applies, as already discussed in the LTI case.

Interval Index

Historically, Interval Index, II, was introduced just after LTI and it is certainly one of the most used variability index. It was proposed by Yeh *et al.* [1973] as a long term variability statistic (see also [Parer *et al.*, 1985]); the original definition was,

$$\mathrm{II_Y} = \frac{\mathrm{std}[T_{24}(i)]}{\mathrm{mean}[T_{24}(i)]},$$

where $T_{24}(i)$ corresponds to 30 seconds of RR signal.

Arduini *et al.* [1993] obtained a better short term parameter by changing the definition to

$$\mathrm{II} = \frac{\mathrm{std}[|T_{24}(i+1) - T_{24}(i)|]}{\mathrm{mean}[|T_{24}(i+1) - T_{24}(i)|]} = \frac{\mathrm{std}[|T_{24}(i+1) - T_{24}(i)|]}{\mathrm{STV}}, \tag{4.5}$$

where $T_{24}(i)$ ($i \in [1;24]$) is a minute of RR signal. We followed this second approach.

Even in this case, exclusion of large accelerations and decelerations applies, as already discussed in the LTI case.

Remarks

Classical statistical FHR indexes can be divided into two main families: (i) long term and (ii) short term statistics. The first ones measure the variability over the whole segment of the considered signal. Both LTI and Delta (and $\mathrm{II_Y}$ as well) are among these. The second one detect local properties: longer time series are used only to gain better statistical estimations by means of averages; STV and II pertain to the latter group.

4.4 Toco Signal

During a CTG recording session, while the fetal heart rate is detected by means of an ultrasound-Doppler probe, the intrauterine pressure is also measured through a pressure transducer, located on the mother's abdomen.

The signal, acquired at a sapling frequency of usually 2Hz (the same of the cardiographic device), estimates the variations of intrauterine pressure (conventionally in mmHg) with respect to a certain initial value, manually configured at the beginning of the recording session (conventionally indicated with the value of 20mmHg). Therefore, it is only a relative measure.

The pressure time-series is also called "toco" signal; in the following we will indicate it with $Q(i)$.

4.4.1 Reference Line Detection

Usually, no baseline is computed for the toco signal; at least nothing equivalent to what described for the FHR series. In fact, the toco signal is more predictable and, often a straight line, drawn at the reference value (20mmHg), is used to detect deviations during the visual inspection of the monitoring strip.

On the other hand, if the mother moves during the recording, or the belt (to which the transducer is connected) is not properly fasten or is loosen by movements, the reference value may not be properly calibrated anymore. In these cases, sudden steps appear in the pressure signal, and the reference value moves, usually to a different level. While developing a fully automatic procedure, these variations in the reference value have to be taken into account, in such a way to not be confused with uterine contractions. Thus, an original algorithm was developed to construct a toco-baseline.

The technique is composed of two different phases.

- Given a certain point, i, a distribution, of the values of the toco signal $Q(j)$ in a 8-minutes window centered on i ($j \in [i-480, i+480]$) was constructed by means of an histogram. The value of the 0.5 quantile, $\xi_{0.5}(i)$, was then computed; it constitutes a first rough value for the baseline in i.

  The procedure was repeated for each point in the toco series (dealing with the distributions computed at the extreme of the series, we considered null the signal for $j < 1$ and $j > N$, where N is the toco-length).

- The quantile series $\xi_{0.5}(i)$ was then filtered through the low-pass FIR filter

$$h(j) \equiv h(N_f - j + 1) = \frac{1}{(\alpha+1)^2},$$

  where $\alpha = (N_f - 1)/2$ and N_f is the order of the filter. The filter has exact linear phase; we used $N_f = 15$, which correspond to a time window of 7.5s (band -3dB, approximatively $f = 0.08$Hz). The filter was applied twice, the first time forward and the second backward.

A toco signal and the baseline computed with this procedure is reported in figure 4.3(b).

Contraction

As long as the gestational process evolves, the uterine musculature undergoes occasional contractions; the number of contractions gradually increases as the delivery date approaches. Uterine contractions are physiological and prepare the musculature for delivery labour.

A contraction is revealed, in the pressure signal, by a triangular-like pattern. To automatically detect contractions, we sought for positive deviations from the baseline, that:

- have a minimum distance from the baseline of 7.5 mmHg, for at least one minute (120 points);
- have a minimum distance from the baseline of 10 mmHg, for at least 30 seconds (60 points).

The algorithm, we developed, was obtained by adapting the approach to localization of FHR accelerations described in Mantel *et al.* [1990*b*]. In particular only phase one was considered relevant and two subsequent contractions were collected together if the gap among them was shorter then 25 points (12.5s; the signal must be over the baseline for the duration of the gap). The toco signal was preliminary filtered through a low-pass FIR filter[4] to eliminate the high frequency jitter. Each contraction, conventionally, starts when the distance of the toco signal from the baseline exceeds 5mmHg. Similarly, it ends when the distance goes below 5mmHg. In figure 4.3(*b*), two contractions, located with this technique, are illustrated.

4.5 Experimental Protocol

4.5.1 Data collection

The data were collected by means of a joint research effort[5], including university departments and a private company, in Italy. The project aim is to develop new tools, diagnostic or clinic, in the field of cardiotocography.

In a temporal interval of two years, 815 CTG recordings were collected, under the supervision of prof. Domenico Arduini. The data were acquired

[4] The filter is defined by: $h_1 \equiv h_5 = -3/35$, $h_2 \equiv h_4 = 12/35$ and $h_3 = 17/35$; it has exact linear phase, with a -3dB band ending at $f \approx 0.48$Hz.

[5] Dipartimento di Informatica (Università di Pavia), Dipartimento di Bioingegneria (Politecnico di Milano), Istituto di Ginecologia ed Ostetricia (Università di Roma "Tor Vergata"), Agilent Technologies Italia.

with four identical devices (HP M135XA), located in various university clinics in Roma. For 549 of them, it has been possible[6] to obtain the physician's diagnosis about the health conditions of the baby at delivery (weight, type of delivery, Apgar score, ...).

Each recording lasts, at least, 30 minutes and it is composed of both the cardiographic series and the toco trace. For the FHR sequence, it is also available (and it has been collected) a quality index signal, corresponding to the output of the autocorrelation procedure, previously described. It is quantified in three different levels ("green" or 32, optimal quality – high correlation; "yellow" or 64, acceptable quality; "red" or 128, insufficient quality – signal unavailable). Signals were recorded at the highest available sampling frequency (2Hz).

After removal of possible red-quality points at the beginning of the sequence, each FHR series has been subdivided into 3-minutes segments (360 points). We considered a segment to be of "sufficient" quality (SQ) if it contains at most 17 red-quality points ($< 5\%$ of the length of the segment). Out of the 549 recordings, only those with, at least, 5 segments of sufficient quality were further considered, discarding the others (22 recording were discarded, 4% of the available ones).

We chose this quality-assessment procedure to ensure that:

i. sequences with long red-quality segments, but with still a reasonable quantity of good data wouldn't be discarded (it would be the case setting a maximum rate of red-quality points over the whole signal);

ii. the quality of the data further employed in parameters estimation was sufficient;

iii. the technique could be easily employed in a CTG device, actually used in monitoring practice (more strict requirements would have led to a high number of rejection, with a subsequent increase of the monitoring time).

We further restricted the number of recording employed, by activating the three following criteria:

1. Only 4 of the possible pathological states, of the mother-fetus system, were taken into account. The other ones were considered numerically

[6] Often, the delivery doesn't occur in the same hospital where the fetal monitoring was performed; in all these cases, it has been very difficult to have all the information required by the experimental protocol.

too small and statistically intractable. We focused on: (i) alterations caused by maternal hypertension (H), (ii) intra-uterine growth retardation (IUGR), (iii) alterations caused by maternal diabetes (DG), and (iv) fetal macrosomia (MACRO). Normal fetus (N) as well were included. Each fetus was allocated to one of these categories according to the statements of the obstetrician.

2. Only patients with a gestational age ranging between 28 and 42 weeks were considered. Older and younger ages were considered not statistically relevant (with a very small numerical incidence on clinical practice).

3. Only one recording per week, per patient, was taken into account. In fact in few cases the monitor session was performed several times in the same week leading to a possible polarization of the whole work on that subjects. When more then one signal was available, the one with the larger number of sufficient-quality segments was selected.

Table 4.2 reports a subdivision, of the considered recordings, with pathological state and gestational week. The total number of signals further considered into the analysis is 362 belonging to 252 different patients[7].

[7] The distribution of the 252 patients, with respect to the pathological state, is:

ID	Pathol.State	Patients#	Recordings#	A priori prob.(%)
1	N	154	200	49.1
2	H	32	53	12.7
3	IUGR	23	40	10.2
4	DG	19	38	7.4
5	MACRO	24	31	9.7
	Total	252	362	89.0

"A priori prob." is the *a priori* probability for each pathological state computed on the whole set of diagnoses, without quality considerations. It, loosely speaking, represents the population of patients that access to the screening test in the considered hospitals.

ID	Population \ Week	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	Total
1	**N**	-	-	1	5	7	9	12	13	25	23	37	25	17	18	8	200
2	**H**	1	1	1	-	2	2	6	4	4	8	11	8	4	1	-	53
3	**IUGR**	-	-	-	2	4	4	6	4	5	2	5	7	1	-	-	40
4	**DG**	-	-	-	1	-	1	2	7	8	7	3	6	1	2	-	38
5	**MACRO**	-	1	-	-	-	1	1	1	2	1	10	8	2	3	1	31
	Total	1	2	2	8	13	17	27	29	44	41	66	54	25	24	9	362
	(OTHER)	-	1	5	4	11	6	7	7	11	9	9	8	7	1	2	88

Tab. 4.2: *Out of the 527 sufficient-quality recordings collected, only a few have been selected, according to the criteria specified in the text. In the table, signals are allocated to the gestational week and to the pathological state they belong to. The pathological states taken into account are: (i) normal healthy fetus (N); (ii) alterations caused by maternal hypertension (H), (iii) intra-uterine growth retardation (IUGR), (iv) alterations caused by maternal diabetes (DG), and (v) fetal macrosomia (MACRO). All the subjects belonging to other pathological states have been collected into the variable (OTHER) and reported for comparison purposes only.*

4.5.2 Parameters Computation

On each of the selected recordings, we computed some parameters. The aim of the research is to investigate if a group of indexes, x, is able to characterize the signal, that is, if it is possible to automatically allocate, by means of a classification technique, a fetus to a pathological state (one out of a fixed number of types) according to the value of x.

The classification approach is described in the next paragraphs, here we focus on the parameters x_i composing x.

For each recording: (i) some indexes were computed on the whole signal, (ii) others on each segment of sufficient-quality and, finally, (iii) a few others on each minute composing each sufficient-quality segment. This is related to the typical definition of each index (*e.g.*, accelerations are counted over the whole signal while LTI is computed on one-minute basis).

a. We started by computing the baseline for both FHR and toco signals. Then, we counted the number of: (1) large accelerations per hour, (2) small accelerations per hour, (3) decelerations per hour and (4) uterine contractions per hour. The numerical details were illustrated in sections 4.3.3 and 4.4.1.

b. On each sufficient-quality (SQ) RR segment (the decimated series T_{24}, sampled at 0.4Hz, was employed) we considered: (5) the mean $T_{24}(i)$ value (ms), (6) the standard deviation of $T_{24}(i)$ (ms^2) and (7) the Long Term Variability (LTI, measured in ms; described in section 4.3.4) value.

   Also spectrum analysis (see section 2.2) was performed on each SQ-segment (over a 3-minutes interval, the FHR signal should be sufficiently stationary). After artifacts detection, the original FHR series, $S(i)$ was processed for correction; missing or discarded points were substituted by means of linear interpolation over nearby points. Spectrum analysis was carried out on the RR sequences obtained by $T(i) = 60000/S(i)$. A parametric approach was followed [Kay and Marple, 1981; Zetterberg, 1969; Baselli *et al.*, 1997]: as a first step an autoregressive model (AR) fitted the data[8]; then the frequency axis was divided into the 4 adjacent bands[9]:

[8] The order was selected in the fixed interval [8; 20]; the order had to minimize the Akaike [1974] information criterium and also fulfill the Anderson whiteness test (see note (10) in chapter 2).

[9] VLF=*Very Low Frequency*, LF=*Low Frequency*, MF=*Middle Frequency* and HF=*High Frequency*.

	Start	End
VLF	0	0.03
LF	0.03	0.15
MF	0.15	0.5
HF	0.5	$f_{Nyquist} = 0.5/\Delta$,

where the frequencies are expressed in Hz-equivalent units and Δ is the mean $T_{24}(i)$ value, in seconds. Ranges are different from those suggested by the Task Force [1996] and reported in table 2.1. In fact, the basal value of FHR is usually twice as big as the mean frequency of an healthy adult. Therefore, the ranges must be enlarged to reach higher frequencies where, reasonably, physiological mechanisms play a role. Besides, a fetus typically moves and the movements contribute in the dynamics of the heart system by adding a frequency component in the range, christened "middle frequency". Summarizing, we considered four bands: VLF, LF (collecting neural sympathetic activity), MF (depending on fetal movements and maternal breathing) and HF (marking the presence of fetal breathing). Finally, we computed the power in each spectral band, by direct integration of the power spectral density of the fitted AR-model (see note (13) in chapter 2). The obtained parameters were: (8) LF-power (ms^2), (9) MF-power (ms^2), (10) HF-power (ms^2) and (11) the ratio LF/(MF+HF) (the autonomic "sympatho-vagal balance" described in section 2.2.3)

Also, on each $T(i)$ SQ-segment, approximate entropy (ApEn), a regularity statistics [Pincus, 1991, 1995], was computed. Several previous works, on HRV of fetus or new-born children [Fabani and Sassi, 1996; Pincus and Viscarello, 1992], verified that, among the nonlinear statistics, ApEn could reliably discriminate healthy and pathological subjects employing shorter time series. The multifractal approach, described in chapter 3, though very powerful, requires large numbers of points to be effective. Unfortunately, through cardiotocography, it is practically impossible to obtain a sufficiently long time series[10] and ApEn is preferable. A few notes on approximate entropy can be found in section 2.5; we followed the indications in Fabani and Sassi [1996], setting $m = 1$ e $r = 0.2$, thus the computed parameter was: (12) ApEn$(1, 0.2)$. The algorithm, actually employed, is described in appendix C.

[10] Cardiotocography technology produces poor quality signals with many signal losses; this makes difficult to collect long time series. Secondary, the recording time must be limited not to maintain the mother in a still position too long; this could induce in the signal anomalous patterns.

c. On each minute of each sufficient-quality segment we computed: (13) Delta (ms), (14) short term variability (STV, measure in ms) and (15) interval index (II). Numerical details are described in section 4.3.4).

For each parameter, described at point (b), several numerical values are available, one for each SQ-segment. Equivalently, parameters at point (c) have an estimate for each minute. To ensure a better statistical reliability, we suggest to average multiple values, thus decreasing intraindividual variability.

Summary

Among several cardiotocographical recordings, 362 exams have been selected to be of sufficient quality, with a gestational age from 28 to 42 weeks. They were associated, by expert physicians, to 5 pathological states.

On every recording, several parameters have been computed; when possible, they have been calculated more then once (on each sufficient quality 3-minutes-segment and on each minute) and subsequently averaged. The indexes were:

computed on the whole signal
- (1) large accelerations per hour
- (2) small accelerations per hour
- (3) decelerations per hour
- (4) uterine contractions per hour

computed on each 3-minutes SQ-segments
- (5) $T_{24}(i)$ mean (ms)
- (6) $T_{24}(i)$ standard deviation (ms^2)
- (7) LTI (ms)
- (8) LF-power (ms^2)
- (9) MF-power (ms^2)
- (10) HF-power (ms^2)
- (11) LF/(MF+HF)
- (12) ApEn$(1, 0.2)$

computed on each minute in each 3-minutes SQ-segments
- (13) Delta (ms)
- (14) STV (ms)
- (15) II

The 15 parameters[11] (when applicable their average) plus the (16) ges-

[11] The 15 parameters might be grouped also as: morphological (large and small acceler-

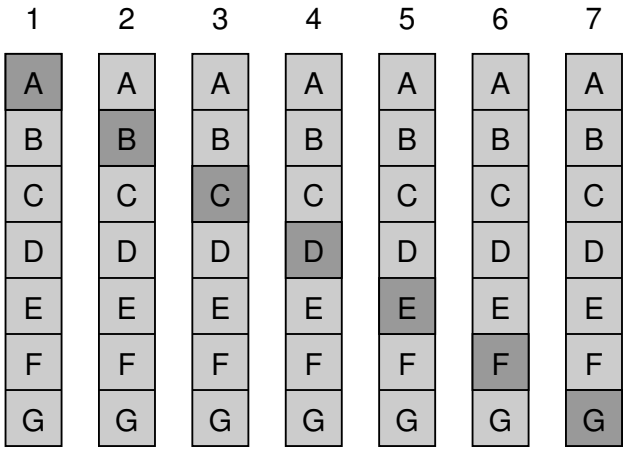


Fig. 4.4: *Structure of the 7-fold cross-validation. The original data set $\mathcal{X}$ is divided into 7 subsets, here labelled ranging from A to G. At first (column 1 in figure), A is regarded as the test set (darker grey) while the training set is composed by all the other subsets (lighter grey). Thus, according to the supervised technique in use, a classifier is determined and cases in A are classified. Then, in turn, each other group is taken as a test set (columns ranging from 2 to 7) and the process is repeated, overall, 7 times. Eventually, each case in $\mathcal{X}$ is classified and the predicted group can be compared with the actual one. The technique ensure that the data employed in the training and test processes are separated; nevertheless the training set is not as small as it would be by splitting the data set in two parts.*

tational age of the fetus, constituted the multivariate variable x (in the following also called "16-Set"), which was used in the classification process.

4.6 Classification With Multivariate Methods

"Multivariate statistical analysis is concerned with data that consist of sets of measurements on a number of individuals or objects" [Anderson, 1984]. The subjects under analysis might be flowers, described by the length and color of their petals or, as in our case, human fetuses characterized by variables obtained by means of cardiotocographic equipments. The data set can be represented by the matrices $\mathcal{X}$ $(n \times p)$, where n is the number of observations (362 recordings) and p the number of variables (the 16 parameters computed

ations per hour, decelerations per hour and contractions per hour), time domain ($T_{24}(i)$ mean, $T_{24}(i)$ standard deviation, LTI, Delta, STV and II), frequency domain (LF-power, MF-power, HF-power and LF/(MF+HF)) and regularity parameters (approximate entropy).

on each recording). A single row of $\mathcal{X}$ may be though as an observation extracted from a multivariate distribution.

Multivariate methods can be separated in two main groups: (i) those methods that assume a given structure. They usually divide the observations into g groups and specify to which of them each case belongs; (ii) those methods that seek for discovering a possible structure in the data matrix, for example, eventually obtaining a separation into groups [Venables and Ripley, 1999, chap. 11]. Following the typical terminology of pattern-recognition, the first ones are called *supervised* methods and the second ones *unsupervised*. A classical unsupervised method is cluster analysis, widely employed in social sciences, where investigators are interested in finding the smallest number of homogeneous groups (*cluster*) in which a certain population of individuals or behavioral patterns can be rearranged.

On the other hand, supervised methods try to allocate future cases (for example, future CTG recordings) to one of the g pre-specified classes in which the current observations are collected. Discriminant analysis and neural networks are classical supervised techniques: current data are used to determine the discriminant function or to train the network. Modern statistics refers to the process of case allocating into predefined classes (medical diagnosis, for example) as "classification" [Ripley, 1997].

Almost all classification methods can be seen as ways to approximate an optimal classifier, the Bayes rule. Given a future case x, the classifier finds the class k with the largest posterior probability $p(k|x)$ and allocates the case to this class (this is correct if all mis-classifications are considered equally bad and we do not take into account the possibility of "doubt" classifications). The posterior probability are learned from a training set, a collection of examples, already classified (by experts or physicians, for example) mimicking the way humans use to perform classifications by means of comparison with some summary of past experiences. This approach, where the estimated probabilities $p(k|x)$ are used as true probabilities, is called the *plug-in* approach and can result in *over-fitting*, by performing very well only on the training set but not on any future cases. To avoid this problem, the available data are usually split into two sub-set, a *training* and a *test* set. The first one is used to estimate the classification model; the second acts as a group of future cases and is classified with the model previously obtained. In this way over-fitting is excluded (the second set was not employed when the classifier was constructed) and a reliable estimation of the performances of the classification process is achieved.

Several supervised techniques were applied on the parameters set com-

Normal			Pathological		
Excluded Var.	Cor.Max	RCond	Excluded Var.	Cor.Max	RCond
NONE	-	1.69e-3	NONE	-	1.77e-3
Delta	0.94	1.96e-3	Delta	0.94	4.08e-3
LF	0.92	4.35e-3	LF	0.91	6.65e-3
STV	0.90	6.50e-3	STV	0.89	1.31e-2
std	0.77	2.51e-2	MF	0.71	1.54e-2
MF	0.67	3.22e-2	std	0.65	3.26e-2
LTI	0.43	4.43e-2	LTI	0.51	4.64e-2
decelerations	0.39	4.76e-2	Gest. weeks	0.39	5.37e-2
Gest. weeks	0.34	5.17e-2	HF	0.35	8.02e-2
large acc.	0.32	8.79e-2	large acc.	0.20	1.23e-1
HF	0.28	1.11e-1	mean	0.16	1.40e-1
mean	0.15	1.21e-1	contractions	0.09	1.44e-1
II	0.05	1.47e-1	II	0.04	1.67e-1
contractions	0.00	1.65e-1	small acc.	0.03	1.85e-1
small acc.	0.00	2.02e-1	decelerations	0.00	2.24e-1

Tab. 4.3: *Starting from the full set of parameters (*16-Set*), the correlation matrices have been computed for both normal and pathological populations. The* LINPACK *reciprocal condition estimator for the full matrices is reported in the first row. Badly conditioned matrices have "RCond" values near zero. In turn the more correlated variable has been excluded from the set (having the largest correlation coefficient, "Cor.Max", and the largest sum of correlation coefficients) until only two variables remained [LF/(MF+HF) and ApEn(1,0.2)]. The "RCond" value is always acceptable.*

puted in section 4.5.2 and the results are reported below. As already stated, the aim of this research was to develop an automated tool that could help physicians in the diagnostic process. A classification technique, which is able to classify with a sufficient precision a CTG recording, would certainly fulfill our goal.

At first, 7 non overlapping subsets, of 50 recordings each, were randomly chosen from the full set of 362 exams. Then, with each supervised method, a 7-fold cross-validation technique was employed, using the same subsets partition (12 exams never entered any test set, though they were always contained in the training partition). This procedure ensures a fair comparison among different methods. The validation technique is described in figure 4.4. Besides, the whole population was divided in two groups: *normal* (labelled "1"), if the baby at delivery was regarded as N, and *pathological* (labelled "2") when the fetus was included in states H, IUGR, DG and MACRO (pathological states were described in section 4.5.1).

A few standard statistical analysis were performed on the parameters

Normal		Pathological	
Excluded Var.	R^2	Excluded Var.	R^2
STV	0.968	Delta	0.961
LF	0.959	LF	0.928
Delta	0.914	STV	0.896
std	0.907	std	0.801
MF	0.704	MF	0.728
LTI	0.670	ApEn(1, 0.2)	0.657
ApEn(1, 0.2)	0.619	LTI	0.604
large acc.	0.453	HF	0.483
Gest. weeks	0.272	Gest. weeks	0.289
small acc.	0.266	large acc.	0.159
HF	0.200	mean	0.086
contractions	0.075	LF/(MF+HF)	0.032
mean	0.028	contractions	0.011
LF/(MF+HF)	0.007	decelerations	0.002
II	0.001	II	$5e^{-5}$
decelerations	-	small acc.	-

Tab. 4.4: *Starting from the full set of parameters (*16-Set*), in turn, a regression model was constructed with response variable x_i and explanatory variables the remaining indexes x_j $(j \neq i)$. The parameter x_i, better explained by the remaining variables (leading to the model with the largest value of the R^2 "goodness-of-fit" statistic), was then excluded; it is reported in the first row of the table. The process was iterated until all the variables, except one, were eliminated (as done in usual backward stepwise regression). Setting a tolerance value ($tv = 1 - R^2$) of* 0.3*, approximatively* 5 *variables resulted redundant for a linear model in both populations. Caveat: the table reflects the linear correlation in the parameters; it becomes only a guideline for variables selection with a possible nonlinear model.*

set to verify the degree of linear dependence. As part of the computations involved in several methods, the covariance matrix of the variables in the model is inverted. Variables linearly dependent on the other ones would lead to ill-conditioned matrices (characterized by very small condition numbers), which can not be inverted. Moreover, completely redundant variables would only make computations more complex.

From the analysis of the covariance matrix, reported in table 4.3, the condition number resulted always acceptable. Also a backward stepwise multi-regression was performed (see table 4.4), but at this stage we considered incorrect to exclude any variable (it would be acceptable if we were considering only linear models. For an introduction to multivariate regression see [Everitt and Dunn, 1991, chap. 8]).

	LDA		QDA		LOGDA		KNN1	
(A)\(P)	1	2	1	2	1	2	1	2
1	132	65	124	73	132	65	96	101
2	86	67	93	60	86	67	81	72
Sens.	43.8		39.2		43.8		47.0	
Spec.	67.0		63.0		67.0		48.7	

Tab. 4.5: *Statistical classifiers are compared on the* 16-Set *data. For each supervised method a contingency matrix is reported: rows contain actual classes (A), while predicted groups are in the columns (P); "1"≡ "normal", "2" ≡ "pathologic". Sensitivity ("Sens.") and specificity ("Spec.") are also reported in %. For a definition see footnotes (16) and (17).*

4.6.1 Discriminant Analysis

Discriminant analysis is a classic parametric approach to discrimination. Excluding neural network, it is, perhaps, the best known method of classification. The decision regions, produced by the classifiers in the parameters multidimensional space, have boundaries with very smooth surfaces.

Linear & Quadratic Discriminant Analysis

Linear discriminant analysis (LDA), as first derived by Fisher, does not imply any normality of the multivariate probability distribution of the two populations. Fisher's idea was to find a linear combination of the variables x having maximal squared difference between the two sample mean values (divided by the pooled estimate of the variance of that difference) [Everitt and Dunn, 1991, chap. 12]. The recording x is then associated to population 1 if

$$\left(\bar{x}_1^T - \bar{x}_2^T\right) S^{-1} \left(x - \frac{\bar{x}_1 - \bar{x}_2}{2}\right) > \log \frac{\pi_2}{\pi_1}, \tag{4.6}$$

and otherwise to population 2, where $\bar{x}_1$ and $\bar{x}_2$ are the sample mean values in each population. S is the pooled within-group sample covariance matrix given by $S = (n_1 S_1 + n_2 S_2)/(n_1 + n_2)$ where $n_1 = 200$ and $n_2 = 162$ are the populations sizes and S_1 and S_2 the group covariance matrices. If the two classes are *a priori* equal likely ($\pi_1 = \pi_2$), then x is classified as coming from the nearest class, in the sense of having the smallest Mahalanobis distance to its mean (see also Ripley [1996*a*], for a complete description of the method and numerical details).

By making the assumption that the distribution of the vector x, belonging to the k class, is multivariate normal with mean μ_k and covariance matrix Σ (all classes share the same covariance matrix $\Sigma \equiv \Sigma_K$) the rule in equation (4.6) can be re-derived from the classical *Bayes rule* [Ripley, 1996*a*]. On the other hand, by supposing a different covariance matrix Σ_K for each class, quadratic discriminant analysis (QDA) is obtained. With this second classifier, the variable x is allocated to the class k that minimizes the quantity

$$Q_k \equiv \frac{1}{2} \log |S_k| + \frac{1}{2} (x - \bar{x}_k)^T S_k^{-1} (x - \bar{x}_k) - \log \pi_k, \tag{4.7}$$

Means and covariances have been substituted by their sample correspondent values. The analysis is termed "quadratic" because Q_k is a quadratic form in x, while equation (4.6) is linear. The boundaries of the decision regions are, respectively, quadratic and linear surfaces in the x space.

The results for the computations[12], on the *16-Set* data, are reported in table 4.5. The *prior* probabilities for the two classes, $\pi_1 = 200/362$ and $\pi_2 = 162/362$, were unknown and were extracted directly from each class proportion. Both the classifiers perform very poorly[13].

Logistic Discriminant Analysis

Making the assumption of two normally distributed classes with mean μ_k and common covariance matrix Σ, posterior probabilities obey the log-linear model

$$\log \left[\frac{p(k=1|x)}{1 - p(k=1|x)} \right] = \alpha + \beta^T x \tag{4.8}$$

usually known as *logistic regression* [Ripley, 1996*a*]. The recording x is allocated to class 1 if $p(k = 1|x) > 0.5$, otherwise to class 2. The whole classification process is called *logistic discriminant analysis* (LOGDA) and it is usually used as a benchmark for more complex classifiers.

With the *16-Set* data, the parameters α and β were computed by maximum likelihood using a generalized linear model approach [Everitt and Dunn,

[12] Computations described in sections 4.6.1 and 4.6.2 were performed largely using routines present in the MASS library, developed by Venables and Ripley [1999], in the R environment. R is a free system for statistical computation and graphics [Ihaka and Gentleman, 1996], that can be found at `www.r-project.org`. Caveat: all the results were obtained using a 7-fold cross-validation, as described in the previous section.

[13] In both linear and quadratic discriminant analysis we used just a *plug in* approach. Results obtained using a *predictive* rule [Venables and Ripley, 1999, sec. 11.4] do not show statistically significant differences. They are not reported here.

1991]. The estimation process is independent of the form assumed for the class density function, so deviation from normality are less effective. The results are reported in table 4.5: they provide a very poor classification.

Linear discriminant analysis can, also, be derived from the same hypothesis supporting LOGDA; nevertheless the two estimates may give quite different classifiers.

4.6.2 k-nearest neighbour classifiers

A typical non-parametric approach to classification is *via* k-nearest neighbour classifiers (KNN). KNN is based on finding the k nearest vectors to x (in correspondence to a definition of distance), in the parameters space, and by taking a majority vote among the classes of these k points. The rule is equivalent to estimate the posterior probability $p(k|x)$, by the proportion of the classes among the k examples [Venables and Ripley, 1999]. In both KNN and LOGDA, the probability of the actual set $p(x)$ is never computed, as not needed for classification. In statistical science this is called *diagnostic* approach [Ripley, 1996*b*], apposite to the *sampling* paradigm where $p(x)$ is obtained from sample values (as with LDA and QDA, for example).

In the computation we performed, several values of k have been tested. We found the best choice to be $k = 1$ with a Euclidean measure of the distances. Classification scores are reported in table 4.5.

4.6.3 Neural Networks

The classification approaches, attempted in the previous sections, weren't successful at all. Indeed, more powerful techniques, able to construct complex decision regions in the parameters space, are needed. Thus, we decided to employ a neural network [Ripley, 1996*a*; Hudson and Cohen, 2000].

As a first step, the original set, composed by 16 variables was reduced. It is possible that a few variables were not relevant to the classification process and were acting as noise; moreover, reducing the input variables would also decrease the network training time, thus permitting a more accurate estimation of the model parameters.

Unfortunately, a neural network is essentially a nonlinear system and it does not allow to uniquely distinguish which parameters are less important then the other ones inside the classification process. Therefore, several different approaches were attempted. Most of them are relevant to the construction of a linear model. Nevertheless they can give interesting insight

Variable	t-test p
Gest. weeks	0.016
Contractions	0.026
MF	0.062
Delta	0.070
LTI	0.077
LF	0.119
II	0.121
small accelerations	0.131
STV	0.139
ApEn(1,0.2)	0.636
std	0.653
large accelerations	0.654
decelerations	0.816
HF	0.827
mean	0.936
LF/(MF+HF)	0.980

Tab. 4.6: *Normal and pathological fetal populations are compared by means of a t-test. The null hypothesis: "the two populations (sampled from a normal distribution) have the same mean" is tested for each variable. The significance levels are reported in ascending order.*

Added Variable	F-test p	F	T^2
Gest. weeks	0.016	5.83	5.83
II	0.009	4.79	9.62
LTI	0.007	4.06	12.23
std	0.002	4.33	17.48
LF	3.1e-4	4.76	24.09
ApEn(1,0.2)	3.0e-4	4.35	26.45
contractions	4.5e-4	3.87	27.53
large accelerations	6.8e-4	3.49	28.47
Delta	9.5e-4	3.21	29.53
decelerations	1.3e-3	2.97	30.50
MF	0.002	2.78	31.51
small accelerations	0.003	2.56	31.72
LF/(MF+HF)	0.005	2.38	31.98
HF	0.007	2.20	32.02
STV	0.012	2.06	32.09
mean	0.018	1.92	32.10

Tab. 4.7: *Normal and pathological fetal populations are compared by means of a stepwise multivariate F-test. The null hypothesis is "the two normal multivariate sets have the same mean value". At each step, the variable, which minimize the significance level, is added to the set employed in the comparison. The added variable, the significance level p, the F statistic and the Hotelling's T^2 statistic are reported in each line.*

	Normal			Pathological		
Comp.	% σ^2	$\sum$ % σ^2	Var.	% σ^2	$\sum$ % σ^2	Var.
1	39.22	39.22	LF	38.85	38.85	LF
2	14.31	53.53	LF/(MF+HF)	13.73	52.58	LF/(MF+HF)
3	10.35	63.89	small acc.	8.40	60.98	mean
4	9.56	73.45	Gest. weeks	8.22	69.20	Gest. weeks
5	6.50	79.95	II	7.52	76.72	decelerations
6	5.11	85.06	mean	6.30	83.02	II
7	3.82	88.88	contractions	4.91	87.93	large acc.
8	3.43	92.31	decelerations	3.42	91.35	contractions
9	2.29	94.60	HF	2.45	93.80	std
10	1.80	96.40	ApEn(1,0.2)	2.01	95.81	HF
11	1.27	97.68	LTI	1.49	97.30	LTI
12	1.19	98.87	large acc.	1.14	98.45	MF
13	0.52	99.39	std	0.65	99.09	ApEn(1,0.2)
14	0.35	99.74	MF	0.48	99.57	STV
15	0.15	99.89	STV	0.28	99.85	Delta
16	0.11	100.00	Delta	0.15	100.00	small acc.

Tab. 4.8: *Principal Component Analysis is performed on the 16 variables set characterizing both normal and pathological population; each population is analyzed separately. Every component is associated with the variable, not already chosen, which has the highest coefficient in absolute value on that component [Everitt and Dunn, 1991]. "% σ^2" is the proportion of variance, "$\sum$ % σ^2" is the cumulative proportion of variance and "Var." is the associated variable.*

and a possible starting point in the variable selection process which must be performed by successive experiments, anyhow.

- Mono-variate t-test (see table 4.6); for each single variable, two sets were considered: x^N and x^P, which correspond, respectively, to normal and pathological fetuses. A standard t-test was performed to determine whether x^N and x^P (we supposed them sampled from normal distributions, with unknown variances) could have the same mean value. A standard F-test was preliminarily applied to compare the two variances (significance level $< 5\%$), then the suitable t-test was selected. The technique aimed at verify with respect to which variable pathological fetuses were separated from the normal ones.

- Multi-variate F-test (see table 4.7). The full *16-Set* was separated into two parts, $\mathcal{X}^N$ and $\mathcal{X}^P$, corresponding to normal and pathological fetuses, respectively. Initially only one parameter was considered. In this case the multivariate F-test coincides with a classical t-test for

two normal monovariate samples with common variances (the selection process started from the variable "Gest. weeks" (gestational weeks), showing the smaller p value in table 4.6). Then, at each step, the variable, minimizing the significance level for a multivariate F-test [Everitt and Dunn, 1991], was added to the set employed in the comparison, until the full 16-dimensional set was obtained again. Multivariate mean values were always significantly different and the minimum p value was obtained by taking into account 5 or 6 variables.

- Principal Component Analysis (PCA) (see table 4.8); PCA [Everitt and Dunn, 1991] was performed on the two matrices $\mathcal{X}^N$ and $\mathcal{X}^P$, separately. Correlation matrices were used. Every component was then associated with the variable, not chosen already, which had the highest coefficient in term of absolute value for that component. Principal components analysis is one of the most classical multivariate statistical method. It seeks new variables, which are rotations of the old parameters, to better explain variation in the data set (in the rotated space, the new correlation matrix is diagonal). The components are orthogonal and each one is responsible for a proportion of the global variance. With the *16-Set* data, only 5-6 components were sufficient to explain $\approx 80\%$ of the variance.

- In a recent study [Paglione, 2000]), performed on the same data set (*16-Set*) at the "Università di Pavia[14]", an ANFIS (Adaptive Neuro-Fuzzy Inference System) classifier (Sugeno-type)[15] was considered. In the work the sensitivity of the classification process to each considered parameters was analyzed and six variables were retained (decelerations, large accelerations, small accelerations, (LF/MF+HF), ApEn(1,0.2) and gestational age); unfortunately, as two different validation processes were employed, is quite impossible a comparison with the results presented in this chapter).

By combining all the collected information together, the original parameter set was reduced to a new one composed by 5 variables only (*5-Set*):

[14] The institution is a member of the joint research project.

[15] The parameters of the membership function (only linear or constant for Sugeno-type inference system) are tuned using a backpropagation algorithm. ANFIS systems, like neural network, "learn" from the data.

computed on the whole signal	(1)	large accelerations per hour
	(2)	small accelerations per hour
computed on each 3-minutes SQ-segments	(3)	LTI (ms)
	(4)	LF/(MF+HF)
	(5)	ApEn$(1, 0.2)$

Subsequently, a static neural network (NN), which had considered the *5-Set* as input, was constructed. Three layers had been employed, composed by 12, 8 and 1 neurons, respectively. The transfer function used for each neuron was the hyperbolic tangent sigmoid $y = 2/[1 + \exp(-2x)] - 1$; the output of the network was quantized in two values, with a static threshold set at zero ($-1 \equiv$ "pathological" and $1 \equiv$ "normal").

Input CTG parameters in each training sets and the corresponding actual output groups were used to train the network (up to 30000 training epochs), until an acceptable error goal was achieved (a back-propagation algorithm was employed).

The classification performances of the NN are reported in table 4.9. The classifiers described in sections 4.6.1 and 4.6.2 were also computed on the reduced set. The NN performed better then any other technique which has been evaluated in this work, showing a 20% misclassification rate and an appreciable sensitivity[16] and specificity[17], both reaching approximatively 80%.

4.7 Discussion

At the present, automated methods have limited clinical application in cardiotocography. A relevant amount of this unsatisfactory performance resides on the weakness of methods used for classifying fetal condition generating risk alarms during pregnancy. Moreover, even if variability became an integral part in fetal evaluation, from the clinical point of view the lack of standardization makes any comparison very difficult. Often, researchers use different numerical expressions for indexes sharing the same name. In the present chapter we tried to move a step forward towards an automated CTG

[16] Sensitivity or "true positive rate" is the probability that the test is positive given that the person has the disease.

[17] Specificity or "true negative rate" is the probability that the test is negative given that the person does not have the disease.

	5-Set			*16-Set*
	Misclas. Rate	Sensitivity	Specificity	Misclas. Rate
LDA	48.9	18.3	76.6	43.1
QDA	48.3	52.3	51.3	47.4
LOGDA	49.1	19.0	75.6	43.1
KNN1	46.0	46.4	59.9	52.0
NNET	20.0	76.1	83.3	-

Tab. 4.9: *The five classification techniques employed on the FHR dataset are compared. Two sets of variables are considered:* 16-Set*, the full one, and* 5-Set*, the reduced set suggested in section 4.6.3. A 7-fold cross-validation technique has been employed; for the neural network (12+8+1 neurons with hyperbolic tangent transfer function, see section 4.6.3), at each classification step, the previous network has been used as a starting point. Only the feed-forward neural network is able to separate normal fetuses from pathological ones.*

risk alerts generator, that might help the physician in drawing the final diagnosis. The work was performed on two different levels.

In first place, parameters selection was conducted with large attention, by comparing the different definitions in literature and by clearly stating any modification introduced in the numerical procedures. New algorithms were developed for toco signal analysis, to improve robustness during clinical usage. Also, FHR signal quality assessment was considered essential. Numerical indexes were computed on short 3 minutes windows and averaged to reduce intraindividual variability.

Classical supervised classifiers fail to distinguish pathological and normal fetuses. It may be possible that the normality hypothesis, required by quadratic discriminant analysis (DA) and logistic DA, is not appropriate for a few variables composing the parameter set.

The poor value of the true classification rates obtained also with linear DA, probably suggest that the two populations lie in very convoluted and intermingled regions in the parameters space. Direct inspection of the data set confirmed such assumptions. Therefore, only methods able to shape very complex decision regions are eligible to succeed.

The proposed feed forward network achieved a 80% true classifications rate with sufficient high sensitivity and specificity. We acknowledge that the method is still not suitable for a direct usage in the clinical environment. Nevertheless the approach is promising, and, in addition, it was achieved by

an automatic procedure.

The use of an even larger data set and a classification performed on potential pathological states basis (with a different group for each pathological state), are only a few among the possible future improvements.

5. DISCUSSION & CONCLUSIONS

In this concluding chapter, the results, obtained during the thesis work, are summarized. Sections dedicated to discussion are also contained in chapters 2 and 3, respectively; we refer to them for specific results. Here, we want simply to summarize the main topics of the work and the view is more general.

The thesis was mainly focused on the study of the mechanisms which generate heart rate variability. The RR signal reflects, partially, the characteristics of the cardiovascular system by which is modulated; at the same time, it is indirectly influenced by several others physiological subsystems, acting on different time scales. The focus of the work was set on the complexity of HRV signal, and it was not devoted to any particular method.

Two main complementary approaches have been chosen: multifractal and multivariate analysis.

In the introduction, we sketched two main questions, which the thesis was meant to face. The first issue dealt with the possible nonlinear nature of the HRV signal and the general debate inside the scientific community about the topic. We underlined the widespread convincement that deterministic chaos may be proved, by means of time series methods, only when the natural or laboratory system under analysis is in controlled conditions and a reasonable model is available (the latter is a sufficient, but not necessary, proposition). All the other situations must be considered on the basis of each singular case after applying the most robust methods available in literature (to be sure that the results are reasonably verified). Certainly it is worth reminding that often it is possible neither to demonstrate the nonlinear nature of the system, nor to disprove it. The analysis fells in a Gödel-like logic contradiction, where *tertium datum est*.

In chapter 3, we tested the nonlinear hypothesis by means of generalized structure functions. The method was developed in the study of fully developed turbulence and leads to the computation of a multifractal spectrum, h_q; when the h_q values decrease with the order q, the process is multifractal;

thus a decreasing spectrum might be associated with the presence of nonlinear correlations in the signal, and possibly, with the nonlinear nature of the HRV signal. We found that most of the series in the NOLTISALIS data base showed a multifractal-like behaviour.

When working with RR series, the sufficient proposition recalled above is useless. The two requirements can not both be fulfilled: too many mechanisms are modulating the interbeat signal over long time scale to be able to construct a complete model. Moreover, extended time series would be necessary but it is difficult to maintain controlled conditions for long periods of time.

Therefor, it is necessary to proceed by a subtractive logical approach only, reducing the number of possible hypotheses. In order to do so, we applied several classical time series methods. We verified that the distribution of the incremental signal $I^{\Delta}(i) = RR(i+1) - RR(i)$ is *not* normal, but displays very long tails (in fact, we found the best fitting distribution to be a Lévy stable). Heavy tails are typical of anomalous mixing in turbulent fluids and other nonlinear processes [Mandelbrot, 1983]. We investigated the chance of spurious multi-fractality, due only to nonlinearly filtered colored noise, through a surrogate data technique (amplitude-adjusted fourier transform with iterative refinement). 32 over 50 h_q spectra resulted statistically different ($p < 0.01$) from their surrogate. This findings would support the thesis of multifractal RR series. Unfortunately, the values we obtained for the h_q exponents were small, thus non-stationarity effects could have played a relevant role. To further verify the non-stationarity influence, we also applied the wavelet-transform modulus-maxima method, which should eliminate polynomial trends from data. Also the latter methods did *not* allow a rejection of the hypothesis that the process underlying HRV is nonlinear.

In summary, the various methods employed did *not* contradict the statement that the RR series is generated by a nonlinear process. On the other hand, it is not possible to state with absolute certainty that the process *is* nonlinear, either. The question is still open, and we are confident that the followed path is correct and fruitful. Surely, some of the presented results are noteworthy and significant from a methodological point of view. For example, the multifractal exponents, when used simply as statistical indexes to discriminate between healthy and cardiac-diseased patients, performed better than any other considered methods. We retain this an evidence that new properties of the signal are taken into account and exploited.

Certainly, a formalistic approach and the study of basic issues, as the nature of the cardiovascular control system, are fundamental and intriguing.

But then, from a bioengineering point of view, it is necessary to move further. Even if the theoretical framework is not, still, completely clear (and it remains unclear, even after many distinguished efforts) useful statistical results should not be discarded but deeply analyzed and interpreted. An example will make clear this "pragmatic" statement. Approximate entropy (ApEn) is certainly connected with the Kolmogorov-Sinai entropy (KS), but while the latter has a clear theoretical derivation, the first one is not more than a statistical regularity index. But, KS is not well defined on time series; on the contrary ApEn, can be easily calculated and proved to be very effective in discrimination tasks. Should ApEn not be taken into account? For sure, it i's useless to make statements around the nonlinear nature of the process (it is only a vaguely approximation of KS). Conversely, we retain, it should be employed as a statistic index able to distinguish signals among them, on a regularity basis, as ApEn demonstrated a great efficiency on this.

The second question, which was addressed in the thesis, is about the opportunity to use variability indexes with a diagnostic purpose. Among the several possibilities, in chapter 4, we opted for multivariate analysis. We do not consider a single parameter at a time, but a set of variables was used to distinguish healthy from potentially pathological fetuses.

We analyzed 362 cardiotocographical recordings; the Holter signals used in chapter 3 were, in fact, numerically not sufficient to achieve statistical confidence. The nonlinear hypothesis about the nature of the RR series was not disproved, it is possible that nonlinearity are present, also, in the fetus cardiovascular system. ApEn was preferred to the multifractal spectrum: the latter can not be computed, because it is impossible to obtain long enough signals. Moreover, the regularity of a process is highly influenced by nonlinear correlations. Besides ApEn, other 15 variables were taken into account: classical time domain indexes, morphological parameters and spectral components.

Classical statistical classifiers (linear, quadratic and logistic discriminant analysis) and a nonparametric approach (k-nearest neighbour classifiers) were compared with a feed-forward neural network. Only the latter was able to draw complex enough decision regions, in the parameters space. The developed neural network correctly allocated, approximatively, 80% of the considered exams, with discrete sensitivity and specificity (also $\approx 80\%$).

We are aware that the method still needs further work and tuning, especially if the aim is to develop a device which must be actually used in the clinical environment; nevertheless, the approach is really promising and certainly more performing than any other technique, today available in car-

diotocography.

Even if 362 recording are a lot, to further increase the statistical robustness of the classification process, an even larger number of cases would be positive. For this reason, the research project is still making an effort to collect new recordings to reach a more clinically significant sample. Another future refinement will be the use of adaptive neuro-fuzzy classifiers and the construction of a new neural network to decrease the number of misclassifications and increase both sensitivity and specificity. The set of 16 parameters described in chapter 4, has been recognized worthy to be implemented in a commercial cardiotocographic monitor, develop by Agilent Technologies and now in prototypal stage. The monitor will be provided also with the classification routines when the precision achieved will be considered sufficient for a clinical use. The numerical libraries, which we wrote and employed during the thesis work, constitute the kernel of the monitoring software.

Regarding the work on the NOLTISALIS database, a further development could concern the analysis of very long RR series recorded from patients in controlled conditions (for example, normal subjects lying in bed at controlled temperature for several hours). Diminishing the external stimuli, the non-stationarity present in the data should decrease. Interesting would be to verify the differences arising from the results obtained with multifractal analysis in this thesis and with the new set of data; would the multifractal exponents increase in absolute value?

APPENDIX

A. LÉVY STABLE DISTRIBUTION

Consider N mutually independent random variables X_k with common distribution $F_X(x)$. F_X is called *stable* if, whatever N, does exist two values $c_N^* > 0$ and ζ_N such that

$$c_1 X_1 + c_2 X_2 + ... + c_N X_N = c_N^* X + \zeta_N \tag{A.1}$$

where c_k are real numbers; besides, F_X is said *stable in strict sense* if $\zeta_N = 0$.

Feller [1971] proved that, if $c_k = 1$ for each k, then $c_N^* = N^{1/\alpha}$ with $0 < \alpha \leq 2$; α is said the *characteristic exponent* for F_X. A strictly stable distribution is *fractal* in nature, as the sum of N independent variables extracted from it looks exactly the same as a single variables[1], once adjusted by a scale factor c_N^*.

The theory of random variables was formalized by Lévy [1937] (see also [Mantegna, 1991]); among other results, he showed that the family of symmetrical distributions

$$f_X(x, \alpha, \gamma) = \frac{1}{\pi} \int_0^{+\infty} e^{-\gamma q^\alpha} \cos(qx) dq \tag{A.2}$$

defines all the stable symmetrical distributions which solve the functional equation (A.1) combined with the auxiliary relation

$$c^\alpha = c_1^\alpha + c_2^\alpha + ... + c_N^\alpha.$$

[1] it is possible to derive a generalization of the box-counting dimension for a random variable X extracted from a strictly stable distribution D. X is equivalent to the sum of N random variables extracted from the same distribution and scaled by $c_k = 1/c_N^*$, that is X_k/c_N^*. The number of variables necessary to *cover* X, N, is proportional to the *diameter*, $\epsilon_N = 1/c_N^*$, so that $N = \epsilon_N^{-D}$, where D is the dimension. It follows that

$$D = \lim_{N \to \infty} -\frac{\log(N)}{\log(\epsilon_N)} = \lim_{N \to \infty} \frac{\log(N)}{\log(N^{1/\alpha})} = \alpha.$$

The value $\alpha \in (0, 2]$ is the characteristic exponent and $\gamma > 0$ a scaling constant; such distributions are called, after Lévy, *Lévy stable distributions*. The density f_X can be obtained as inverse Fourier transform of the function $e^{-\gamma|q|^\alpha}$ often called *characteristic function*; in fact

$$\begin{aligned} f_X(x, \alpha, \gamma) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\gamma|q|^\alpha} \cos(qx) dq \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\gamma|q|^\alpha} \cos(qx) dq + i \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\gamma|q|^\alpha} \sin(qx) dq \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\gamma|q|^\alpha} e^{iqx} dq = \mathscr{F}^{-1} \left\{ e^{-\gamma|q|^\alpha} \right\} \end{aligned}$$

The density has analytical form only for $\alpha = 1$, $3/2$ and 2; in $x = 0$

$$f_X(0, \alpha, \gamma) = \frac{\gamma^{-\frac{1}{\alpha}}}{\pi\alpha} \Gamma\left(\frac{1}{\alpha}\right), \tag{A.3}$$

where $\Gamma(t) = \int_0^{+\infty} x^{t-1} e^{-x} dx$ is the Gamma function. The cumulative distribution function can be obtained switching the order of the integrations in (A.2)

$$F_X = P[X \geq x] = \frac{1}{2} + \frac{1}{\pi} \int_0^{+\infty} e^{-\gamma q^\alpha} \frac{\sin(qx)}{q} dq, \tag{A.4}$$

The moment generating function (see [Mandelbrot, 1977]) is for the Lévy distribution[2]

$$m_{f_X}(t) = e^{-\gamma|t|^\alpha}. \tag{A.5}$$

The first moment is finite (E[X]=0) for $1 < \alpha \leq 2$ while the second only for $\alpha = 2$ ($E[X^2] = 2\gamma$). All superior moments are infinite.

The Cauchy (Lorentzian) and the Gaussian distribution are particular case of Lévy stable distribution. In fact with $\alpha = 1$

$$f_X(x, 1, \gamma) = \frac{\gamma}{\pi} \frac{1}{\gamma^2 + x^2}$$

[2] At this purpose it is useful to change the usual definition of the *moment generating function* in

$$m_{f_X}(t) = E[e^{ixt}],$$

the r^{th} moment can be obtained with the formula

$$E[X^r] = (i)^{-r} \lim_{t \to 0} \frac{d^r m_{f_X}(t)}{dt^r}.$$

α	$\Phi^{-1}(1)$	$\Phi^{-1}(2)$	$\Phi^{-1}(3)$
1.00	0.92	6.98	117.91
1.20	0.92	4.63	47.63
1.40	0.94	3.40	23.92
1.60	0.96	2.65	13.33
1.80	0.98	2.21	7.31
2.00	1.00	2.00	3.00

Tab. A.1: *Confidence intervals of zero for the standardized variable $\bar{x}$. $\Phi^{-1}(\bar{x})$ is the inverse of the standard normal cumulative distribution function. Values are expressed in standardized dispersion $\bar{\sigma}_L$.*

which is a Cauchy distribution; instead, setting $\alpha = 2$, it is possible to obtain the Gaussian distribution

$$f_X(x, 2, \gamma) = \frac{1}{2\sqrt{\pi\gamma}} \exp\left(\frac{-x^2}{4\gamma}\right) = \frac{1}{\sqrt{2\pi}\sqrt{2\gamma}} \exp\left(\frac{-x^2}{2(\sqrt{2\gamma})^2}\right)$$

with standard deviation $\sigma = \sqrt{2\gamma}$; the normal is the only stable distribution with finite second moment.

A scaling parameter, that can be considered the extension of the standard deviation to stable distributions, is the dispersion $\sigma_L = [2\gamma]^{(1/\alpha)}$. With the change of variable $\bar{x} = x/\sigma_L$ it can be obtained the *standardized* Lévy density

$$f_{\bar{X}}(\bar{x}, \alpha, 1/2) = \frac{1}{\pi} \int_0^{+\infty} e^{-(1/2)q^\alpha} \cos(q\bar{x}) dq,$$

with $\bar{\gamma} = 1/2$ and $\bar{\sigma}_L = 1$, very useful in calculation. A few confidence intervals of zero for the standardized variable $\bar{x}$ are reported in table A.1; they correspond to probability levels frequently used: $\Phi^{-1}(1)$, $\Phi^{-1}(2)$ and $\Phi^{-1}(3)$, where $\Phi^{-1}(\bar{x})$ is the inverse of the standard normal cumulative distribution function ($\approx 68.3\%$, $\approx 95.5\%$ and $\approx 99.7\%$ respectively). With small values of α the confidence intervals becomes incredibly large, accounting the long tails in the density.

B. THE KURAMOTO MODEL

B.1 Introduction

Mutual synchronization is a common phenomenon in biology. It occurs at different levels, ranging from the small scale of the cardiac pace-maker cells of the SA (sinoatrial) and AV (atrioventricular) nodes in the human hearth that synchronously fire and give the pace to the whole muscle, to the coordinated behaviours of crickets that chirp in unison and of fireflies that flash together in some parts of southeast Asia.

Winfree [1967] was the first to underline the generality of the problem, fixing the first assumptions for a mathematical model. In his work each *oscillating species* (cell, or cricket, or firefly) is modeled as a nonlinear oscillator with a globally attracting limit cycle; The oscillators were assumed to be weakly coupled and their natural frequencies to be randomly distributed across the population.

Kuramoto [1975] proposed the first model (called for this reason the *Kuramoto model*). His assumptions were that each oscillator is equal to the others, upto the frequency and phase, that the system has a *mean field* coupling and that the amplitudes of the oscillations are all the same (phase-only model). The equation of the model for the n oscillator (regarding its phase ϕ) is:

$$\frac{d\theta_n}{dt} = \omega_n + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_n) + \xi_n, \tag{B.1}$$

where K is the *coupling strength*, ω_n is a random variable with probability density function $g(\omega)$ and ξ_n is white noise.

Defining as *order parameter* the complex number,

$$re^{i\psi} = \frac{\sum_{j=1}^{N} e^{i\theta_j}}{N}, \tag{B.2}$$

it is possible to *measure* the synchronization among the oscillators phases: $r = 0$ corresponds to the completely incoherent state, finite r to synchronization.

Kuramoto determined that $r = 0$ is always a steady solution; but there exists, in the case of no added random noise, a critical value of the coupling parameter $K_c = 2/[\pi g(0)]$ below which only incoherent populations exist ($r = 0$). For $K > K_c$ a population of synchronized oscillators can exist ($r > 0$).

The results of the numerical simulations, performed solving equation (B.1) with $N = 256$ and two different values of the coupling parameter K, are shown in figure B.1; in the upper panels the time evolution of the discrete probability density function[1] is plotted; in the lower panels, the time evolution of the order parameter is displayed. The initial condition is, in both cases, a population of oscillators with phases uniformly distributed in $[0, 2\pi]$ and slightly perturbed.

When $K = 0.8 > K_c$ ($K_c = 0.739$), in a very short time, the phases of the oscillators *gather* together in a small range of angles and then begin drifting coherently. The order parameter grows quickly and exhibits small oscillations due to the random noise added to the system.

A different situation arises with $K = 0.65 < K_c$; a coherent behaviour never starts, even if small structures can be noticed: small population of oscillators synchronize and drift for short periods of time. This is reflected in the order parameter that oscillates between 0 and 0.3 and decreases only slowly.

B.1.1 A continuous model

Using the approach sketched in the previous paragraph, it is difficult to go much farther; it is not easy, for example, to answer questions such as *"Is the coherent state ($K > K_c$) stable?"*

Strogatz and Mirollo [1991] introduced a partial differential equation that describes the behaviour of the Kuramoto model in the limit $N \to \infty$.

The idea is that, in the continuous limit, the state is described by a probability density function: $\rho(\theta, \omega, t)$. The Kuramoto equation (B.1) becomes:

$$v = \omega + K \int_{-\infty}^{\infty} \int_{0}^{2\pi} \sin(\phi - \theta) \rho(\phi, \omega, t) g(\omega) d\phi d\omega, \tag{B.3}$$

[1] The θ axis is divided into 64 intervals and, at each instant of time, the normalized histogram of the phases of the oscillators is computed.

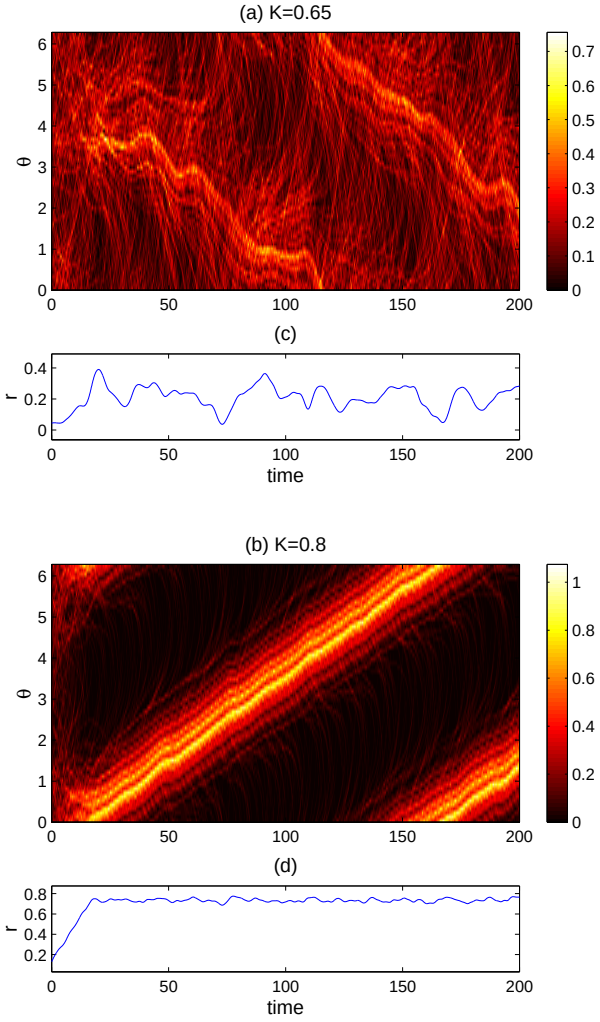


Fig. B.1: *Numerical simulations on the discrete Kuramoto model with* $N = 256$, $D = 0.01$, $K = 0.65 \approx 0.9K_c$ *(panel a and c) and* $K = 0.8 \approx 1.1K_c$ *(panel b and d). (*panel a and b*): time evolution of the probability density function computed on the trajectories of the system splitting up the* θ *axe in sub-intervals; (*panel c and d*): time evolution of the absolute value of the order parameter. The initial condition is a population of oscillators with phases chosen in* $[0, 2\pi]$ *according to the distribution* $\rho(\theta, \omega, 0) = 1/(2\pi) + 4\xi/[\pi(\omega^2 + 4)]$ *with* $\xi = 0.1$. *The natural frequencies* ω_n *are selected from the probability distribution* $g(\omega) = (1 - \omega^2)/[(\pi - 2)(1 + \omega^2)]$ *with* $|\omega| < 1$.

where v is the *velocity* at the point (θ, ω, t). Moreover, the density function ρ has to satisfy, for each given ω, a normalization law

$$\int_0^{2\pi} \rho d\theta = 1; \tag{B.4}$$

and a Fokker-Plank-type conservation law [2]

$$\frac{\partial}{\partial t}\rho(\theta, \omega, t) + \frac{\partial}{\partial \theta}\left[\rho(\theta, \omega, t)v(\theta, \omega\, t)\right] = D\frac{\partial^2}{\partial \theta^2}\rho(\theta, \omega, t), \tag{B.5}$$

where $\theta \in [0; 2\pi]$ and $\omega \in [-\infty; \infty]$.

The *order parameter* (B.2), in the continuous limit, becomes:

$$re^{i\psi} = \int_{-\infty}^{\infty}\int_0^{2\pi} e^{i\theta}\rho(\theta, \omega, t)g(\omega)d\theta d\omega. \tag{B.6}$$

The continuous model permits further analysis on the system; first of all the computation of the stability threshold for a wider number of cases. An in depth description of the topic, is provided by [Balmforth and Sassi, 2000; Sassi, 2000].

In chapter 1, the physiology of the myocardial cells was discussed. The electrical activity of the heart starts in the sinoatrial node. Pacemaker cells are autonomous oscillators and they do not need any external current to depolarize. On the other hand, they must fire synchronously (with the same

[2] The derivation of the two equations has the flavor of the BBGKY hierarchy in plasma physics and can be found in [Crawford and Davies, 1999]. Some rationalization of equation (B.5) can be given on recollecting that because the probability is conserved,

$$\begin{aligned}
\frac{\partial}{\partial t}\int_{\theta_1}^{\theta_2}\rho(\theta)d\theta &= \rho(\theta_1)v(\theta_1) - \rho(\theta_2)v(\theta_2) \\
&= -\int_{\theta_1}^{\theta_2}\frac{\partial}{\partial\theta}(\rho v)d\theta \\
\rho_t &= -\frac{\partial}{\partial\theta}(\rho v),
\end{aligned}$$

and on remembering Einstein's derivation of the diffusion equation in his work on the explanation of the Brownian motion [Einstein, 1985], which indicates that

$$\begin{aligned}
\rho_t &= 2\langle\xi^2\rangle\rho_{\theta\theta} \\
&= D\rho_{\theta\theta}
\end{aligned}$$

phase in the periodical firing cycle) to produce a coherent pacing signal for the rest of the muscular tissue.

Making the assumptions that (i) the number of nodal cell is small, thus the coupling among them can be described by a mean field interaction, (ii) each cell activity can be modelled with a nonlinear oscillator (see note (2) in chapter 1), the Kuramoto model is a possible description of the dynamic of the group of cells. It gives an account of the transition to synchrony and allows to compute how strong must be the interaction among oscillators to produce a coherent behaviour.

In pacemaker cells, like in figure B.1*(b)*, the cells phase-synchronize around the mean natural frequency. Fortunately, the coupling is strong enough. But if it decreases under a critical threshold (in pathological condition for example), synchronization is lost and, as in figure B.1*(a)*, the capability of the node to excite the heart is highly reduced.

C. NOTES ON APPROXIMATE ENTROPY COMPUTATION

Approximate entropy (ApEn) is a statistics quantifying regularity in a time series [Pincus, 1991]. The definition was reported in section 2.5 and the same terminology will be employed in this brief appendix.

Writing a computer code to compute ApEn from the definition is straightforward; unfortunately, for a series of length N, the computational complexity scales as N^2. Therefore, the time necessary to perform the numerics increases quadratically with the number of points in the signal. Approximate entropy is often employed to measure regularity with very long RR series ($N > 20000$); in all these situations, a fast algorithm *is* critical.

The crucial point in any ApEn computer code is the computation of the kernels $\mathcal{C}_i(m, r)$ and $\mathcal{C}_i(m+1, r)$, which requires a few multiplications but a large amount of comparisons between couples of floating point numbers. Proceeding from the definition,

$$\mathcal{M}_1 = m\frac{(N-m+1)(N-m)}{2} + (m+1)\frac{(N-m)(N-m-1)}{2}$$

comparisons seem necessary, but Fabani and Sassi [1996] showed that this value can be reduced to

$$\mathcal{M}_2 = \frac{N(N-1)}{2} - \frac{m(m-1)}{2}$$

which corresponds to compare each point u_i with, at most, all the others u_j. In this appendix we will show how it is possible to reduce the computational complexity below $\mathcal{M}_2$.

Common practise is to use a value of the parameter r which is proportional to the standard deviation σ of the signal. This is very useful when comparing signals with different powers, and it is equivalent to normalize the series with respect to σ. Suppose that the process u is a white Gaussian

i \ j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1	0	1	0	1	1	0	0	1	0	1	0	1	0	0	0	1	1	1	1
2		1	0	1	1	0	0	0	0	0	0	1	0	1	0	0	1	0	0	1
3			1	1	1	1	1	1	0	0	0	1	1	1	0	0	1	1	1	1
4				1	1	0	0	0	0	1	1	1	1	1	0	1	0	1	0	1
5					1	1	1	0	1	1	0	0	1	0	0	0	0	1	0	1
6						1	1	0	0	0	0	1	1	1	1	1	1	0	0	0
7							1	1	0	1	0	0	0	1	1	1	1	0	0	1
8								1	1	0	0	0	0	0	0	1	1	0	0	1
9									1	1	0	0	1	1	0	1	1	1	1	0
10										1	1	0	0	0	1	0	0	0	0	1
11											1	0	0	1	0	1	1	1	1	0

Fig. C.1: *The comparisons involved in ApEn computation can be illustrated as composing a symmetrical matrix; only a few rows are plotted while the lower half of the matrix is symmetrical and not reported. Each u_i is compared with each u_j (i varies with rows, while j with columns): 1 indicates that the two number are closes enough ($|u_i - u_j| \leq r$); 0 that they are different, at the current r level. Further details are presented in the text.*

noise $\stackrel{d}{=} N(0,1)$; it is straightforward to compute the probability that two points u_i and u_j are closer than r

$$P(|u_i - u_j| \leq r) = \int_r^r \frac{1}{2\sqrt{\pi}}\, e^{-x^2/4}\, dx.$$

For example, with $r = 0.2$ (a very recurrent value), $P = 0.112$. Thus, in this case, only about 11% of the comparisons affect the values of the kernels.

With a more correlated series the value of P would have been bigger, but generally $P \ll 1$ and skipping all the useless comparisons, we would end with an algorithm making only $\mathcal{M}_3 \approx P\mathcal{M}_2$ operations. This is the main idea underlying the algorithm we developed, which is composed, schematically, by the following steps (we describe only the computation of the kernels $\mathcal{O}_i(m,r)$):

a. Preliminarily, an ordered copy w of the signal u is built with a quick-sort algorithm, which is computationally proportional to $N \log_2 N$.

   Then for each point w_i, two indexes are calculated: $j_{min} < i$ and $j_{max} > i$. On the ordered sequence, j_{min} is the smallest value of j for which $w_j > w_i - r$; j_{max} is the largest value of j for which $w_j < w_i + r$. Outside the interval $j_{min} < i < j_{max}$, comparisons are useless.

   The two limiting indexes are computed using a fast bisection method, so that the number of comparisons required is proportional to $2N \log_2 N$.

b. The series w is ordered back to the original sequence u, just to construct two indexes maps: $i_s = h(i_o)$ and $i_o = H(i_s)$. Calling i_o the original position of the point u_i in the series u and i_s its position in the sorted w, the two maps are used to relate one series to the other *without* sorting anymore.

c. Finally the core routine is entered:

   i. Let's start from u_1. Only the w_j points with $j_{min} < j < j_{max}$ must be considered; thus, for each of these j a corresponding $k = H(j)$ is computed. Then the comparisons are performed along the diagonals: u_2 is compared with u_{k+1}, u_3 with u_{k+2} and so on, until a comparison fails. The process is schematically illustrated in figure C.1. Supposing $k = 11$, the couples of points compared are (u_2, u_{12}), (u_3, u_{13}) and (u_4, u_{14}); they are reported in figure with bold red numbers.

   ii. According to the values of m we are interested in, the opportune kernels $\mathcal{O}_i(m, r)$ must be increased. In the example, supposing $m = 2$, $\mathcal{O}_1(2, r)$, $\mathcal{O}_2(2, r)$, $\mathcal{O}_3(2, r)$ and, correspondingly, $\mathcal{O}_{11}(2, r)$, $\mathcal{O}_{12}(2, r)$, $\mathcal{O}_{13}(2, r)$ should be increased by one unit.

   iii. The process is repeated for each point u_i; duplicated comparisons must be avoided (in figure, for u_2, valid values of k are only 5 and 17, marked with bold blue numbers. The others were already considered with u_1).

   The main routine performs only "true" comparisons. They corresponds to the probability P, computed in the WGN case.

Overall, kernels computation requires $\mathcal{M}_3 \approx P\mathcal{M}_2 + 4N \log_2 N$ comparisons. The described routine was employed to compute approximate entropy both in chapters 3 and 4 and proved to be very effective. In average, it was approximatively $5 - 10$ times faster than the code reported in [Fabani and Sassi, 1996, app. D].

BIBLIOGRAPHY

[**Akaike, 1974**] H. Akaike, "A new look at the statistical model identification", IEEE Trans. Automat. Control. **19**, 716–723 (1974).

[**Amaral *et al.*, 1998**] L. A. N. Amaral, A. L. Goldberger, P. C. Ivanov and H. E. Stanley, "Scale-Independent Measures and Pathologic Cardiac Dynamics", Phys. Rev. Lett. **81**, 2388–2391 (1998).

[**Anderson, 1984**] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, second ed. (John Wiley & Sons, New York, 1984).

[**Arduini *et al.*, 1993**] D. Arduini, G. Rizzo, G. Piana, A. Bonalumi, P. Brambilla and C. Romanini, "Computerized Analysis of Fetal Heart Rate: I. Description of the System (2CTG)", Matern. Fetal. Invest. **3**, 159–163 (1993).

[**Balmforth and Sassi, 2000**] N. J. Balmforth and R. Sassi, "A shocking display of synchrony", Phys. D **143**, 21–55 (2000).

[**Baselli *et al.*, 1997**] G. Baselli, A. Porta, O. Rimoldi, M. Pagani and S. Cerutti, "Spectral Decomposition in Multichannel Recordings Based on Multivariate Parametric Identification", IEEE Trans. Biomed. Eng. **44**, 1092–1101 (1997).

[**Berne and Levy, 1996**] R. M. Berne and M. N. Levy, *Cardiovascular Physiology* (The C. V. Mosby Company, Saint Louis, 1996).

[**Bigger *et al.*, 1996**] J. T. Bigger, R. C. Steinman, L. M. Rolnitzky, J. L. Fleiss, P. Albrecht and R. J. Cohen, "Power law behavior of RR-interval variability in healthy middle-aged persons, patients with recent acute myocardial infarction, and patients with heart transplants", Circulation **93**, 2142–2151 (1996).

[**Bittanti, 1981**] S. Bittanti, *Identificazione parametrica* (CLUP, Milano, Italy, 1981).

[**Bittanti, 1990**] S. Bittanti, *Teoria della predizione e del filtraggio* (Pitagora, Bologna, Italy, 1990).

[**Box and Pierce, 1970**] G. E. P. Box and D. Pierce, "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Times Series Models", Journal of the American Statistical Association **65**, 1509–1526 (1970).

[**Box *et al.*, 1994**] G. E. P. Box, G. M. Jenkins and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, III ed. (Prentice Hall, Englewood Cliffs, NJ, 1994).

[**Calcagnini *et al.*, 1995**] G. Calcagnini, Jr, S. Strano, A. Marcheselli, S. Cerutti and G. Calcagnini, "Power spectrum density estimation ($1/f$ spectrum) of long term heart rate variability in acute myocardial infarction", in *Computers in Cardiology 1995*, edited by IEEE Engineering in Medicine and Biology Society, pp. 115–117 (IEEE Computer Society Press, New York, 1995).

[**Cavalcanti and Belardinelli, 1996**] S. Cavalcanti and E. Belardinelli, "Modeling of cardiovascular variability using a differential delay equation", IEEE Trans. Biomed. Eng. **43**, 982–989 (1996).

[**Chang *et al.*, 1994**] T. Chang, S. J. Schiff, T. Sauer, J. P. Gossard and R. E. Burke, "Stochastic versus deterministic variability in simple neuronal circuits: I. Monosynaptic spinal cord reflexes", Biophys. J. **67**, 671–683 (1994).

[**Crawford and Davies, 1999**] J. D. Crawford and K. T. R. Davies, "Synchronization of globally coupled phase oscillators: singularities and scaling for general couplings", Phys. D **125**, 1–46 (1999).

[**Dalton *et al.*, 1977**] K. J. Dalton, G. S. Dawes and J. E. Patrick, "Diurnal, respiratory and other rhythms of fetal heart rate in lambs", Am. J. Obstet. Gynecol. **127**, 414–424 (1977).

[**Dawes *et al.*, 1982**] G. S. Dawes, C. R. S. Houghton and C. W. G. Redman, "Baseline in human fetal heart-rate records", Br. J. Obstet. Gynaecol. **89**, 270–275 (1982).

[**De Haan *et al.*, 1971**] J. De Haan, J. H. Vam Bemmel, B. Versteeg, A. F. L. Veth, L. A. M. Stolte, J. Janssens and T. K. A. B. Eskes, "Quantitative evaluation of fetal heart rate. I. Processing methods", Eur. J. Obstet. Gynaecol. **3**, 95–102 (1971).

[**DiFrancesco and Noble, 1985**] D. DiFrancesco and D. Noble, "A model of cardiac electrical activity incorporating ionic pumps and concentration changes", Phil. Trans. R. Soc. B. **307**, 353–398 (1985).

[**Einstein, 1985**] A. Einstein, *Investigations on the Theory of the Brownian Movement* (Dover Publications, New York, 1985).

[**Everitt and Dunn, 1991**] B. S. Everitt and G. Dunn, *Applied Multivariate Data Analysis* (Arnold, London, 1991).

[**Fabani and Sassi, 1996**] E. Fabani and R. Sassi, *Studio dell'entropia approssimata per la classificazione di serie temporali: applicazioni al segnale di variabilita cardiaca*, Tesi di Laurea, Politecnico di Milano (1996).

[**Feder, 1988**] J. Feder, *Fractals* (Plenum, New York, 1988).

[**Feller, 1971**] W. Feller, *An Introduction to Probability Theory and Its Applications, 2nd Ed.* (Wiley, New York, 1971).

[**Frisch and Parisi, 1976**] U. Frisch and G. Parisi, "On the singularity structure of fully developed turbulence", in *Turbulence and predictability in geophysical fluid dynamics and climate dynamics*, edited by M. Ghil, R. Benzi and G. Parisi, vol. LXXXVIII of *Proceedings of the Enrico Fermi Physics School*, pp. 84–88 (North Holland, Amsterdam, Netherlands, 1976).

[**Goldberger *et al.*, 1990**] A. L. Goldberger, D. R. Rigney and B. J. West, "Chaos and Fractals in Human Physiology", Scientific American **262**, 42–49 (1990).

[**Goldberger *et al.*, 2000**] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet. Components of a New Research Resource for Complex Physiologic Signals", Circulation **101**, e215–e220 (2000). `www.physionet.org`.

[**Grassberger and Procaccia, 1983*a***] P. Grassberger and I. Procaccia, "Estimation of the Kolmogorov entropy from a chaotic signal.", Phys. Rev. A **28**, 2591–2593 (1983).

[**Grassberger and Procaccia, 1983*b***] P. Grassberger and I. Procaccia, "Measuring the strangeness of strange attractors.", Physica D **9**, 189–208 (1983).

[**Halsey *et al.*, 1986**] T. C. Halsey, M. H. Jensen, L. P. K. I. Procaccia and B. I. Shraiman, "Fractal measures and their singularities: The characterization of strange sets", Phys. Rev. A **33**, 1141–1151 (1986).

[**Hausdorff and Peng, 1996**] J. M. Hausdorff and C.-K. Peng, "Multiscale randomness: A possible source of $1/f$ noise in biology", Phys. Rev. E **54**, 2154–2157 (1996).

[**Havlin *et al.*, 1999**] S. Havlin, S. Buldyrev, A. Bunde, A. Goldberger, P. Ivanov, C.-K. Peng and H. Stanley, "Scaling in nature: from DNA through heartbeats to weather", Physica A **273**, 46–69 (1999).

[**Hodgkin and Huxley, 1952**] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve", Journal of Physiology **117**, 500–544 (1952).

[**Hodgkin *et al.*, 1952**] A. L. Hodgkin, A. F. Huxley and B. Katz, "Measurement of current-voltage relations in the membrane of the giant axon of *Loligo*", Journal of Physiology **116**, 424–448 (1952).

[**Hudson and Cohen, 2000**] D. L. Hudson and M. E. Cohen, *Neural Networks and Artificial Intelligence for Biomedical Engineering* (IEEE Press, New York, 2000).

[**Ihaka and Gentleman, 1996**] R. Ihaka and R. Gentleman, "R: A Language for Data Analysis and Graphics", Journal of Computational and Graphical Statistics **5**, 299–314 (1996).

[**Ivanov *et al.*, 1996**] P. Ivanov, M. Rosenblum, C.-K. Peng, J. Mietus, S. Havlin, H. Stanley and A. Goldberger, "Scaling behaviour of heartbeat intervals obtained by wavelet-based time-series analysis", Nature **383**, 323–327 (1996).

[**Ivanov *et al.*, 1998**] P. Ivanov, M. Rosenblum, C.-K. Peng, J. Meitus, S. Havlin, H. Stanley and A. Golberger, "Scaling and universality in heart rate variability distributions", Physica A **249**, 587–593 (1998).

[**Ivanov *et al.*, 1999**] P. C. Ivanov, L. A. N. Amaral, A. L. Goldberger, S. Havlin, M. G. Rosenblum, Z. R. Struzik and H. E. Stanley, "Multifractality in human heartbeat dynamics", Nature **399**, 461–465 (1999).

[**Kaplan and Glass, 1995**] D. Kaplan and L. Glass, *Understanding nonlinear dynamics* (Springer, New York, 1995).

[**Kay and Marple, 1981**] S. M. Kay and S. L. Marple, "Spectrum Analysis - A Modern Perspective", Proc. IEEE **69**, 1380–1419 (1981).

[**Keener and Sneyd, 1998**] J. Keener and J. Sneyd, *Mathematical Physiology* (Springer-Verlag, New York, 1998).

[**Kobayashy and Musha, 1982**] M. Kobayashy and T. Musha, "$1/f$ Fluctuation of Heartbeat Period", IEEE Trans. Biomed. Eng. **BME-29**, 456–457 (1982).

[**Kuramoto, 1975**] Y. Kuramoto, "Self-Entrainment of a population of coupled nonlinear oscillators", in *International symposium on mathematical problems in theoretical physics*, edited by H. Araki, no. 39 in Lecture Notes in Physics, pp. 420–422 (Springer, New York, 1975).

[**Lévy, 1937**] P. Lévy, *Théorie de l'Addition des Variables Aléatories* (Gauthier-Villars, Paris, 1937).

[**Ljung and Box, 1978**] G. M. Ljung and G. E. P. Box, "On a Measure of Lack of Fit in Time Series Models", Biometrika **65**, 297–303 (1978).

[**Lombardi *et al.*, 1987**] F. Lombardi, G. Sandrone, S. Pernpruner, R. Sala, M. Garimoldi, S. Cerutti, G. Baselli, M. Pagani and A. Malliani, "Heart rate variability as an index of sympathovagal interaction after acute myocardial infarction", Am. J. Cardiol. **60**, 1239–1245 (1987).

[**Luo and Rudy, 1991**] C. H. Luo and Y. Rudy, "A model of the ventricular cardiac action potential; depolarization, repolarization and their interaction", Circ. Res. **68**, 1501–1526 (1991).

[**Malliani *et al.*, 1991**] A. Malliani, M. Pagani, F. Lombardi and S. Cerutti, "Cardiovascular neural regulation explored in the frequency domain", Circulation **84**, 1482–1492 (1991).

[**Mandelbrot, 1977**] B. B. Mandelbrot, *Fractals: Form, Chance and Dimension* (Freeman, San Francisco, 1977).

[**Mandelbrot, 1983**] B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1983).

[**Mantegna, 1991**] R. N. Mantegna, "Lévy walks and Milan stock exchage", Phys. A **179**, 232–242 (1991).

[**Mantel *et al.*, 1990*a***] R. Mantel, H. P. van Geijn, F. J. M. Caron, J. M. Swartjes, E. E. van Woerden and H. W. Jongsma, "Computer analysis of antepartum fetal heart rate: 1. baseline determination.", Int. J. Biomed. Comput. **25**, 261–272 (1990).

[**Mantel *et al.*, 1990*b***] R. Mantel, H. P. van Geijn, F. J. M. Caron, J. M. Swartjes, E. E. van Woerden and H. W. Jongsma, "Computer analysis of antepartum fetal heart rate: 2. detection of accelerations and decelerations.", Int. J. Biomed. Comput. **25**, 273–286 (1990).

[**Marshall, 1968**] J. M. Marshall, "The heart", in *Medical physiology*, edited by V. B. Mountcastle, vol. I, XII ed., chap. 4, pp. 35–68 (The C. V. Mosby Company, Saint Louis, 1968).

[**Milnor, 1968**] W. R. Milnor, "Cardiovascular system", in *Medical physiology*, edited by V. B. Mountcastle, vol. I, XII ed., chap. 3, pp. 24–34 (The C. V. Mosby Company, Saint Louis, 1968).

[**Mood *et al.*, 1988**] A. M. Mood, F. A. Graybill and D. C. Boes, *Introduzione alla statistica* (McGraw-Hill, Milano, 1988).

[**Muzy *et al.*, 1993**] J. F. Muzy, E. Bacry and A. Arneodo, "Multifractal formalism for fractal signals: The structure-function approach versus the wavelet-transform modulus-maxima method", Phys. Rev. E **47**, 875–884 (1993).

[**Muzy *et al.*, 1994**] J. F. Muzy, E. Bacry and A. Arneodo, "The multifractal formalism revisited with wavelets", Int. J. of Bifurcation and Chaos **4**, 245–302 (1994).

[**Noble and Noble, 1984**] D. Noble and S. J. Noble, "A model of S.A. node electrical activity using a modification of the DiFrancesco-Noble (1984) equations", Proc. Roy. Soc. Lond. **222**, 295–304 (1984).

[**Oppenheim and Schafer, 1989**] A. Oppenheim and R. Schafer, *Discrete-Time Signal Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1989).

[**Organ *et al.*, 1978**] L. W. Organ, P. A. Hawrylyshyn, J. W. Goodwin, J. E. Milligan and A. Bernstein, "Quantitative indices of short- and long-term heart rate variability", Am. J. Obstet. Gynecol. **130**, 20–27 (1978).

[**Osborne and Provenzale, 1989**] A. R. Osborne and A. Provenzale, "Finite correlation dimension for stochastic systems with power-law spectra", Physica D **35**, 357–381 (1989).

[**Pagani *et al.*, 1986**] M. Pagani, F. Lombardi, S. Guzzetti, O. Rimoldi, R. Furlan, P. Pizzinelli, G. Sandrone, G. Malfatto, S. Dell'Orto, E. Piccaluga, M. Turiel, G. Baselli, S. Cerutti and A. Malliani, "Power spectral analysis of heart rate and arterial blood pressure variabilities as a marker of symphato-vagal interaction in man and conscious dog", Circ. Res. **59**, 178–193 (1986).

[**Pagani *et al.*, 1991**] M. Pagani, G. Mazzuero, A. Ferrari, D. Liberati, S. Cerutti, D. Vaidtl, L. Tavazzi and A. Malliani, "Sympathovagal interaction during mental stress", Circulation **83**, 43–51 (1991).

[**Paglione, 2000**] G. Paglione, *Algoritmi neuro-fuzzy per la cardiotocografia automatica*, Tesi di Laurea, Università degli Studi di Pavia (2000).

[**Parer *et al.*, 1985**] W. J. Parer, J. T. Parer, R. H. Holbrook and B. S. B. Block, "Validity of mathematical methods of quantitating fetal heart rate variability", Am. J. Obstet. Gynecol. **153**, 402–409 (1985).

[**Peng *et al.*, 1993**] C.-K. Peng, J. Mietus, J. M. Hausdorff, S. Havlin, H. E. Stanley and A. L. Goldberger, "Long-Range Anticorrelations and Non-Gaussian Behaviour of the Heartbeat", Phys. Rev. Lett. **70**, 1343–1346 (1993).

[**Peng *et al.*, 1994**] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley and A. L. Goldberger, "Mosaic organization of DNA nucleotides", Phys. Rev. E **49**, 1685–1689 (1994).

[**Peng *et al.*, 1995**] C.-K. Peng, S. Havlin, H. E. Stanley and A. L. Goldberger, "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series", Chaos **5**, 82–87 (1995).

[**Penna *et al.*, 1995**] T. J. P. Penna, P. M. C. Oliveira, J. C. Sartorelli, W. M. Gonçalves and R. D. Pinto, "Long-Range Anticorrelations and Non-Gaussian Behaviour of a Leaky Faucet", Phys. Rev. E **52**, 2168–2171 (1995).

[**Pilgram and Kaplan, 1998**] B. Pilgram and D. T. Kaplan, "A comparison of estimators for 1/f noise", Phys. D **114**, 108–122 (1998).

[**Pilgram and Kaplan, 1999**] B. Pilgram and D. T. Kaplan, "Nonstationarity and $1/f$ noise characteristics in heart rate", American Journal of Physiology **276**, R1–R9 (1999).

[**Pincus, 1991**] S. M. Pincus, "Approximate entropy as a measure of system complexity", in *Proc. Natl. Acad. Sci. USA*, vol. 88, pp. 2297–2301 (1991).

[**Pincus, 1995**] S. M. Pincus, "Approximate entropy (ApEn) as a complexity measure", Chaos **5**, 110–117 (1995).

[**Pincus and Huang, 1992**] S. M. Pincus and W. M. Huang, "Approximate entropy: statistical properties and applications", Community Statistic Theory and Methods **21**, 3061–3077 (1992).

[**Pincus and Viscarello, 1992**] S. M. Pincus and R. R. Viscarello, "Approximate entropy: a regularity measure for fetal heart rate analysis", Obstet. Gynecol. **79**, 249–255 (1992).

[**Press *et al.*, 1992**] W. H. Press, S. T. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes in C: the art of scientific computing* (Cambridge University Press, Cambridge, England, 1992).

[**Provenzale *et al.*, 1991**] A. Provenzale, A. R. Osborne and R. Soi, "Convergence of the K_2 entropy for random noises with power law spectra", Physica D **47**, 361–372 (1991).

[**Rienzo *et al.*, 1997**] M. d. Rienzo, G. Parati, A. Pedotti and P. Castiglioni, "$1/f$ Modelling of blood pressure and heart rate spectra", in *Frontiers of blood pressure and heart rate analysis*, edited by M. d. Rienzo, G. Mancia, G. Parati, A. Pedotti and A. Zanchetti, pp. 45–53 (IOS Press, Amsterdam, 1997).

[**Ripley, 1996*a***] B. D. Ripley, *Patter recognition and neural networks* (Cambridge University Press, Cambridge, 1996).

[**Ripley, 1996*b***] B. D. Ripley, "Pattern recognition via neural networks", (1996). To appear in a volume of Oxford Graduate Lectures on Neural Networks, Oxford University Press.

[**Ripley, 1997**] B. D. Ripley, "Classification (update)", in *Encyclopedia of Statistical Sciences*, edited by S. Kotz, C. B. Read and D. L. Banks, vol. Update 1, pp. 110–116 (John Wiley & Sons, New York, 1997).

[**Rocca, 1998**] F. Rocca, *Elaborazione Numerica dei Segnali* (CUSL, Milano, Italy, 1998).

[**Sassi, 2000**] R. Sassi, "Nonlinear coupled oscillators", GFD report WHOI-2000-07, Woods Hole Oceanog. Inst., Woods Hole, MA (2000).

[**Saul *et al.*, 1987**] J. P. Saul, P. Albrecht, R. D. Berger and R. J. Cohen, "Analysis of long term hearth rate variability: methods, $1/f$ scaling and implications", in *Computers in Cardiology 1987*, edited by IEEE Engineering in Medicine and Biology Society, pp. 419–422 (IEEE Computer Society Press, New York, 1987).

[**Schreiber and Schmitz, 1996**] T. Schreiber and A. Schmitz, "Improved surrogate data for nonlinearity tests", Phys. Rev. Lett. **77**, 635–638 (1996).

[**Schreiber and Schmitz, 2000**] T. Schreiber and A. Schmitz, "Surrogate time series", Phys. D **142**, 346–382 (2000).

[**Stanley *et al.*, 1999**] H. E. Stanley, L. A. N. Amaral, A. L. Goldberger, S. Havlin, P. Ivanov and C.-K. Peng, "Statistical physics and physiology: Monofractal and multifractal approaches", Phys. A **270**, 309–324 (1999).

[**Strogatz and Mirollo, 1991**] S. H. Strogatz and R. E. Mirollo, "Stability of incoherence of a population of coupled oscillators", J. Stat. Phys. **63**, 613–635 (1991).

[**Task Force, 1996**] Task Force, "Heart Rate Variability, Standards of Measurement, Physiological Interpretation and Clinical Use", Circulation **93**, 1043–1065 (1996). Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology.

[**Theiler and Prichard, 1996**] J. Theiler and D. Prichard, "Constrained-realization Monte-Carlo method for hypothesis testing", Phys. D **94**, 221–235 (1996).

[**Theiler *et al.*, 1992**] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian and J. D. Farmer, "Testing for nonlinearity in time series: the method of surrogate data", Physica D **58**, 77–94 (1992).

[**van Geijn *et al.*, 1980**] H. van Geijn, H. W. Jongsma, J. de Haan and T. K. A. B. Eskes, "Analysis of heart rate and beat-to-beat variability: Interval difference index", Am. J. Obstet. Gynecol. **138**, 246–252 (1980).

[**van Geijn, 1998**] H. P. van Geijn, "Cardiotocography", in *A Textbook of Perinatal Medicine*, edited by A. Kurjak, vol. 2, pp. 1424–1428 (Parthenon Publishing, New York, 1998).

[**Venables and Ripley, 1999**] W. N. Venables and B. D. Ripley, *Modern Applied Statistics with S-PLUS*, III ed. (Springer, New York, 1999).

[**von Hardenberg *et al.*, 2000**] J. G. von Hardenberg, R. Thieberger and A. Provenzale, "A box-counting red herring", Phys. Lett. A **269**, 303–308 (2000).

[**Winfree, 1967**] A. T. Winfree, "Biological rhythms and the behavior of populations of coupled oscillators", J. Theoret. Biol. **16**, 15–42 (1967).

[**Yanagihara *et al.*, 1980**] K. Yanagihara, A. Noma and H. Irisawa, "Recostruction of sino-atrial node pacemaker potential based on voltage clamp experiments", Jap. J. Physiology **30**, 841–857 (1980).

[**Yeh, 1998**] S. Yeh, "30 Years Of Electronic Fetal Monitoring. What Have We Learned From This?", Bulletin De Perinatologie **22** (1998). `http://matweb.hcuge.ch/matweb`.

[**Yeh *et al.*, 1973**] S. Yeh, A. Forsythe and E. H. Hon, "Quantification of fetal heart rate beat-to-beat interval differences", Obstet. Gynecol. **41**, 355–363 (1973).

[**Zetterberg, 1969**] L. H. Zetterberg, "Estimation of Parameters for a Linear Difference Equation with Application to EEG Analysis", Math. Biosci. **5**, 227–275 (1969).