UNIVERSIDADE DE SÃO PAULO
FFCLRP - DEPARTAMENTO DE FÍSICA
PÓS-GRADUAÇÃO EM FÍSICA APLICADA À MEDICINA E BIOLOGIA

ANDERSON IVAN RINCON SOLER

**Impact of artifact correction methods on R-R
interbeat signals to quantifying heart rate
variability (HRV) according to linear and nonlinear
methods.**

Dissertation presented to Faculdade de
Filosofia, Ciências e Letras de Ribeirão
Preto of Universidade de São Paulo, as part
of requirements for acquirement the grade of
Master of Sciences, Area: Physics applied to
Medicine and Biology.

Ribeirão Preto - SP
2016

ANDERSON IVAN RINCON SOLER

**Impact of artifact correction methods on R-R
interbeat signals to quantifying heart rate
variability (HRV) according to linear and nonlinear
methods.**

Dissertation presented to Faculdade de
Filosofia, Ciências e Letras de Ribeirão
Preto of Universidade de São Paulo, as part
of requirements for acquirement the grade of
Master of Sciences.

Concentration area:

Física aplicada à Medicina e Biologia.

Advisor:

Ph.D. Luiz Otavio Murta.

Rectified version

Original version available at FFCLRP - USP

Ribeirão Preto - SP

2016

I authorize partial and total reproduction of this work, by any conventional or electronic means, for the purpose of study and research, provided the source is cited.

FICHA CATALOGRÁFICA

Rincon Soler, Anderson Ivan

Impact of artifact correction methods on R-R interbeat signals to quantifying heart rate variability (HRV) according to linear and nonlinear methods. / Anderson Ivan Rincon Soler; orientador Ph.D. Luiz Otavio Murta. Ribeirão Preto - SP, 2016.

50 f.:il.

Dissertação (Mestrado - Programa de Pós-graduação em Física aplicada à Medicina e Biologia) - Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto da Universidade de São Paulo, 2016.

1. Variabilidade da frequência cardíaca. 2. Correções de artefato. 3. Métodos lineares e não lineares. 4. Processamento de sinais biomédicas.

Nome: RINCON SOLER, Anderson Ivan

Título: Impact of artifact correction methods on R-R interbeat signals to quantifying heart rate variability (HRV) according to linear and nonlinear methods.

Dissertation presented to Faculdade de Filosofia,
Ciências e Letras de Ribeirão Preto of
Universidade de São Paulo, as part of requirements
for acquirement the grade of Master of Sciences.

Approved in: ____/____/____.

Examination Board

Prof. Dr. : _____ Institution: _____

Judgement: _____ Signature: _____

Prof. Dr. : _____ Institution: _____

Judgement: _____ Signature: _____

Prof. Dr. : _____ Institution: _____

Judgement: _____ Signature: _____

To my family and my wife.

ACKNOWLEDGEMENTS

First, I want to thank my mother, my father, and my brother, *Carmen, Francisco and Michael*, for giving their unconditional love and support. Thanks for the laughs, the uncountable chats and for always being there. Also, I want to express my warmest thanks to my friend and wife *Yineth*. Thank for your love and support, and for always be with me and give your smile in the hardest moments. Thank for always be you.

I wish to express my huge gratitude to my tutor and friend Prof. Luiz Otavio Murta, PhD, who gave me the opportunity to start this academic stage in Brazil and who was always willing to discuss and clarify any doubts that arose during this process. Thank you also for sharing your knowledge and experience with humility and joy.

I am very grateful to Luiz Eduardo V. Silva, PhD, for to be always available to talk and discuss about the wonderful world of the Biomedical Signal Processing. Thanks to guide me and give me important advises along all this time in order to obtain always the best results.

A heartfelt thanks to my colleagues and friends of the research group **CSIM** for sharing with me everyday, for the excellent talks and to always make me feel comfortable even away from home. I wish to thank my friend Eduard A. Hincapié, M.Sc., who was always available to discuss the ideas about physics or simply available to talk.

Finally, I want to thank **CAPES** for the financial support that allowed the execution of this work.

A. I. Rincón Soler.

RESUMO

RINCON SOLER, A. I. **Impact of artifact correction methods on R-R interbeat signals to quantifying heart rate variability (HRV) according to linear and nonlinear methods..** 2016. 50 f. Dissertação (Mestrado - Programa de Pós-graduação em Física aplicada à Medicina e Biologia) - Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto - SP, 2016.

Na análise da variabilidade da frequência cardíaca (Heart Rate Variability - HRV) são usadas séries temporais que contém as distancias entre batimentos cardíacos sucessivos, com o fim de avaliar a regulação autonômica do sistema cardiovascular. Estas séries são obtidas a partir da análise de sinais de eletrocardiograma (ECG), as quais podem ser afetados por distintos tipos de artefatos, levando a interpretações incorretas nas análises feitas sob as séries da HRV. Abordagem clássica para lidar com esses artefatos implica a utilização de métodos de correção, alguns deles com base na interpolação, substituição ou técnicas estatísticas. No entanto, existem poucos estudos que mostram a precisão e desempenho destes métodos de correção em sinais reais da HRV. Assim, o presente estudo tem como objetivo determinar como os diferentes níveis de artefatos presentes no sinal afetam as características da mesma, utilizando-se diferentes métodos lineares e não lineares de correção e posteriormente quantificação dos parâmetros da HRV. Como parte da metodologia utilizada, sinais ECG de ratos obtidas mediante a técnica da telemetria foram usadas para gerar séries de HRV reais sem nenhum tipo de erro. Nestas séries foram simulados batimentos perdidos para diferentes taxas de pontos a fim de emular a situação real com a maior precisão possível. Adicionalmente, foram aplicados os métodos de eliminação de segmentos (DEL), interpolação linear (LI) e cúbica (CI), janela

de média móvel (MAW) e interpolação preditiva não lineal (NPI) como métodos de correção dos artefatos simulados sob as séries com erros. A precisão de cada método de correção foi conhecida através dos resultados obtidos com a quantificação do valor médio da série (AVNN), desvio padrão (SDNN), erro quadrático médio das diferenças entre batimentos sucessivos (RMSSD), periodograma de Lomb (LSP), análise de flutuações destendenciadas (DFA), entropia multiescala (MSE) e dinâmica simbólica (SD) sob cada sinal de HRV com e sem erros. Os resultados obtidos mostram que para baixos níveis de perdas de batimentos o desempenho das técnicas de correção é similar, com valores muito semelhantes para cada parâmetro quantificado da HRV. Não obstante, em níveis de perdas maiores só NPI permite obter valores muito próximos e sem muitas diferenças significativas para os mesmos parâmetros da HRV, em comparação com os valores calculados para as séries sem perdas.

Palavras-chave: 1. Variabilidade da frequência cardíaca. 2. Correções de artefato. 3. Métodos lineares e não lineares. 4. Processamento de sinais biomédicas.

ABSTRACT

RINCON SOLER, A. I. **Impact of artifact correction methods on R-R interbeat signals to quantifying heart rate variability (HRV) according to linear and nonlinear methods..** 2016. 50 f. Dissertation (M.Sc. - Postgraduate program in Physics applied to Medicine and Biology) - Faculty of Philosophy, Sciences and Literature, University of São Paulo, Ribeirão Preto - SP, 2016.

In the analysis of heart rate variability (HRV) are used temporal series that contains the distances between successive heartbeats in order to assess autonomic regulation of the cardiovascular system. These series are obtained from the electrocardiogram (ECG) signal analysis, which can be affected by different types of artifacts leading to incorrect interpretations in the analysis of the HRV signals. Classic approach to deal with these artifacts implies the use of correction methods, some of them based on interpolation, substitution or statistical techniques. However, there are few studies that shows the accuracy and performance of these correction methods on real HRV signals. This study aims to determine the performance of some linear and non-linear correction methods on HRV signals with induced artefacts by quantification of its linear and nonlinear HRV parameters. As part of the methodology, ECG signals of rats measured using the technique of telemetry were used to generate real heart rate variability signals without any error. In these series were simulated missing points (beats) in different quantities in order to emulate a real experimental situation as accurately as possible. In order to compare recovering efficiency, deletion (DEL), linear interpolation (LI), cubic spline interpolation (CI), moving average window (MAW) and nonlinear predictive interpolation (NPI) were used as correction methods for the series with induced artifacts. The accuracy of each correction

method was known through the results obtained after the measurement of the mean value of the series (AVNN), standard deviation (SDNN), root mean square error of the differences between successive heartbeats (RMSSD), Lomb's periodogram (LSP), Detrended Fluctuation Analysis (DFA), multiscale entropy (MSE) and symbolic dynamics (SD) on each HRV signal with and without artifacts. The results show that, at low levels of missing points the performance of all correction techniques are very similar with very close values for each HRV parameter. However, at higher levels of losses only the NPI method allows to obtain HRV parameters with low error values and low quantity of significant differences in comparison to the values calculated for the same signals without the presence of missing points.

Key-words: 1. Heart Rate Variability. 2. Artifact correction. 3. Linear and nonlinear methods. 4. Biomedical signal processing.

LIST OF FIGURES

2.1	Schematic representation of the Heart Rate modulation mediated by the sympathetic and parasympathetic systems.	4
2.2	Schematic representation of the HRV signal generation from an ECG recording.	5
3.1	Parameters configuration of the LabChart ECG module to detect the QRS complex in a real ECG signal. The values are standard measures for rat ECG signals.	10
3.2	Schematic representation of deletion method performance.	12
3.3	Example of linear interpolation.	13
3.4	Schematic representation of an cubic spline interpolation process. . .	15
3.5	Element test array definition.	18
3.6	Local trend for each box of size n	23
3.7	Example of a crossover phenomena for a DFA analysis of an RR time series for a normal subject. Red line correspond to short-range correlations and green line to long-range correlations.	25
3.8	Illustration of the coarse-graining procedure	26
3.9	Level's definition and symbolic conversion of a time series	27
4.1	Stage of correction for RR time series and their results.	33
4.2	DFA results obtained from subject <i>Rat-01</i>	38
4.3	MSE results obtained from subject <i>Rat-01</i>	39
4.4	SymDyn results obtained from subject <i>Rat-01</i>	40
4.5	RMSE bar plot for time domain parameters	44
4.6	RMSE bar plot for frequency domain parameters	47
4.7	RMSE bar plot for nonlinear parameters	49

B.1 Counting of patterns in the process to calculate the SampEn on an RR time series.	79
--	----

LIST OF TABLES

3.1	Known scaling exponents for different time series behavior.	24
4.1	Type and amount of losses for long RR time series found after visual inspection and calculated for signal with a length of <i>10.000</i> points. .	32
4.2	Time parameters for control group.	33
4.3	Time parameters calculated after application of each correction technique.	34
4.4	Frequency domain parameters for control group.	35
4.5	Frequency domain parameters after the application of correction methods.	36
4.6	Average value for Nonlinear parameters obtained from control group.	41
4.7	Average values for nonlinear parameters after application of correction methods.	42
A.1	Time parameters for control group.	62
A.2	Time parameters after correction with deletion method.	62
A.3	Time parameters after correction using LI.	63
A.4	Time parameters after correction using CI.	63
A.5	Time parameters after correction using MAW.	64
A.6	Time parameters after correction using modified version of MAW algorithm.	64
A.7	Time parameters after correction using NPI algorithm.	65
A.8	Frequency parameters for control group.	65
A.9	Frequency domain parameters after correction using deletion.	66
A.10	Frequency domain parameters after correction using LI.	66
A.11	Frequency domain parameters after correction using CI.	67

A.12 Frequency domain parameters after correction using MAW.	67
A.13 Frequency domain parameters after correction using mMAW.	68
A.14 Frequency domain parameters after correction using NPI.	68
A.15 Frequency parameters for control group.	69
A.16 Nonlinear parameters after correction using deletion.	70
A.17 Nonlinear parameters after correction using LI.	71
A.18 Nonlinear parameters after correction using CI.	72
A.19 Nonlinear parameters after correction using MAW.	73
A.20 Nonlinear parameters after correction using modified MAW algorithm.	74
A.21 Nonlinear parameters after correction using NPI algorithm.	75

LIST OF ABBREVIATIONS

α_1	Short-term scaling exponent of fractal-like correlation properties.
α_2	Long-term scaling exponent of fractal-like correlation properties.
ANS	autonomic nervous system.
AVRR	Average Value of RR interval time series.
CI	Cubic Spline Interpolation.
DFA	Detrended Fluctuation Analysis.
ECG	Electrocardiogra/m, -phic, -phy.
FFT	Fast Fourier transform.
HF	Power in the high frequency range.
HRV	Heart Rate Variability.
LI	Linear Interpolation.
LF	Power in the low frequency range.
LF/HF	The ratio of the power in the low frequency range to that in the high frequency range.
LP	Lomb's Periodogram.
LSP	Lomb-Scargle Periodogram.
MAW	Moving Average Window.

MSE	MultiScale Entropy.
NPI	Nonlinear Predictive Interpolation.
RMSSD	Root Mean Squared value of Successive Differences for an RR interval time series.
SDRR	Standard Deviation of RR interval time series.
SymDyn	Symbolic Dynamics.
VLF	Power in the very-low frequency range.

CONTENTS

List of Figures	xiii
List of Tables	xv
List of Abbreviations	xvii
1 Introduction	1
2 Theoretical Background	3
2.1 Heart Rate Variability	3
2.2 Generation of the HRV Signal	4
2.3 Assessment of the HRV signal	6
2.4 Artifacts in HRV signals	8
3 Materials and Methods	9
3.1 Database	9
3.2 Correction methods	12
3.2.1 Deletion (DEL)	12
3.2.2 Linear Interpolation (LI)	13
3.2.3 Cubic Spline Interpolation (CI)	14
3.2.4 Moving Average Window (MAW)	15
3.2.5 modified Moving Average Window (mMAW)	16
3.2.6 Nonlinear Predictive Interpolation (NPI)	16
3.3 HRV analysis methods	18
3.3.1 Linear methods for HRV analysis	18
3.3.1.1 AVRR	18

3.3.1.2	SDRR	19
3.3.1.3	RMSSD	19
3.3.1.4	Lomb-Scargle Periodogram (LSP)	19
3.3.2	Nonlinear methods for HRV analysis	22
3.3.2.1	Detrended Fluctuation Analysis (DFA)	22
3.3.2.2	MultiScale Entropy (MSE)	25
3.3.2.3	Symbolic Dynamics (SymDyn)	26
3.4	Statistical Tests	28
3.4.1	Paired t-test	28
3.4.2	The Wilcoxon Signed Rank Sum Test	29
4	Results	31
4.1	Missing points quantification	31
4.2	Correction stage	32
4.3	Quantification of HRV parameters	32
4.3.1	Time domain	33
4.3.2	Frequency domain	35
4.3.3	Nonlinear Domain Parameters	36
5	Conclusions	51
	Bibliography	53
	Appendix A - Long series results	61
A.1	Time domain results	61
A.2	Frequency domain results	61
A.3	Nonlinear domain results	69
	Appendix B - Sample Entropy (SampEn)	77

INTRODUCTION

The electrocardiogram (ECG) is a measure of the electrical activity of the heart that can be obtained from electrodes placed on the skin, which allows the description of the depolarization processes of the atria and ventricles [1]. From ECG records can be extracted measurements between successive heartbeats known as RR intervals, which are widely used to assess heart rate variability (HRV). HRV and their changes have been associated with the pathophysiology of some cardiac failures and potential risk of heart disease associated with obesity, epilepsy, diabetes, hypertension, and sudden death [2, 3, 4].

Most problems related to HRV signals are concerned to spurious interbeat intervals, which will lead to misinterpreted results. For human HRV signals, the major problem is related to ectopic beats, atrial fibrillation, sinus tachycardia, sinus bradycardia, ventricular tachycardia, and some others [5, 6]. However, in experimental field (rats and mice), it is very common to find poor quality ECG signals, related to animal movements, poorly fastened electrodes, source power noise and other influences; resulting in HRV signals with a great amount of missing beats. In order to solve the problems related to the presence of artifacts in HRV signals, different correcting methods have been proposed. Some of the most common correction artifact methods used in R-R time series involve process of deletion and interpolation of the problematic segments [7, 8, 5]. However, some other methods have been proposed for pre-processing these time series and eliminated the artifact interference. Some of this methods correspond to: Comparison and merging [9], predictive autocorrelation method [10], non-linear predictive interpolation [7], exclusion of R-R interval segments with divergent duration [5], impulse rejection [11], integral

pulse frequency model (IPFM) [11], sliding window average filter [12] and threshold filtering using wavelet [12, 13].

Despite this number of correction methods, there is not a well-defined methodology about which of them are the most suitable choice to deal with artifacts on HRV signals, and therefore, in the experimental field researchers often have to deal with HRV signals with an intermediate quality. In this way, it is very important to understand the effects of different correcting methods applied on HRV signals in situations closer to those found in the experimental field. In this order of ideas, the main objective of this study is:

- To apply a combination of linear and nonlinear correcting methods for different levels of missing data and evaluate its influence on the calculations of linear and nonlinear HRV indices.

and the specific objectives have been defined as:

- To establish the maximum missing beats that might be present in a signal in order to produce reliable analysis.
- To find the better correcting method that works for different combinations of length and rate missing points.
- To determine the importance of correcting methods on analysis methods (linear and nonlinear).

Finally, this master thesis is organized as follows: In chapter 2 a background of heart rate variability (HRV) and artifacts that affect it are presented. Then, in chapter 3 methods of correction and analysis, including both linear and nonlinear techniques, are described as well as the generation of the signal database for this study. In chapter 4 the results are described and discussed in two stages. First, a sequence of missing points were induced on the signals at two different levels. To follow, corrections were made by linear and nonlinear techniques and quantification of HRV indices take place using linear and nonlinear methods. Secondly, a comparative analysis involving the different level of losses, the correction methods and the analysis methods are presented in detail. In the last part of this document conclusions and some important appendices are stated.

THEORETICAL BACKGROUND

2.1 Heart Rate Variability

Heart rate variability (HRV) is defined as the inter-beat variability between successive heart beats in a determined time interval. This variability is mediated directly by the polarization and depolarization process of the *sinus node* (**SN**), which at the same time is regulated by the interaction of the *sympathetic* and *parasympathetic* branches of the *autonomic nervous system* (**ANS**). An increase in the *parasympathetic* activity implies a *heart rate* (**HR**) diminution mediated by liberation of acetylcholine; while, an increase of the **HR** is a direct consequence of an increase in the sympathetic activity that in this case is mediated through norepinephrine liberation on the heart beat regulatory mechanisms [2, 14, 4, 5]. Then, it can be established that the dynamical balance between sympathetic and parasympathetic activity has strong influence on **HR** causing oscillations around its average value, in other words the **HRV** phenomenon. In this order of ideas, **HRV** is used as noninvasive method to evaluate the sympathetic and parasympathetic functions of the **ANS** and the cardiovascular regulation [15, 16, 17]. Figure 2.1 explain in a graphical way, the influence of the sympathetic and parasympathetic systems on the HR and the result of their interaction.

In this field of study many pathologies, not necessarily from cardiovascular origin, and physiologic factors that directly disturbs the regulation of the **ANS**. These situations cause constant and abrupt changes in the **HR** and the **HRV**. In this way, HRV can be used to study many types of diseases as: Myocardial Infarction [5, 16, 7], sudden cardiac death [7, 17, 5], ventricular arrhythmias [5, 18, 19], congestive

Figure 2.1: Schematic representation of the Heart Rate modulation mediated by the sympathetic and parasympathetic systems.

heart failure [5, 20, 21], coronary artery disease [22], diabetes mellitus [5], epilepsy [23], obesity [14, 18], among others. Additionally, it is necessary to take into account other factors that affect the normal operation of the HR, among which are: Age, gender, body position, ingestion of alcohol or caffeine and stress [5, 16, 24, 14]. Thus, it is clear that a proper analysis of HRV in individuals with health problems mentioned above, can provide specific information for the detection and control of it. Furthermore, in healthy individuals HRV analysis provides information about the adaptability of the human body to physical and mental training [25].

2.2 Generation of the HRV Signal

The Heart Rate Variability (**HRV**) signals are the result of quantifying the distances between consecutive heart beats for a certain period of time. These periods of time are usually estimated through analysis of ECG signals with a duration of minutes or hours. In practice, the most accurate method is to identify all *QRS* complexes contained in the ECG signal and then mark as a reference point the R peak, because it is readily distinguishable from the other components of the complex. As a result, over the years it have been proposed a lot of methods for detecting R peaks, some of them based on: *Hilbert* transform [26], signal filtering

(*Pan-Tompkins algorithm*) [27, 28], *pattern recognition* [29] and *Wavelet transform* [12, 30]. Although the accuracy of these methods, there is no standard methodology for the R-peak detection phase, and this part of the processing is always left to the investigator's choice.

With all R peaks detected the next step is to calculate the time difference between two consecutive marks (R peaks) in order to generate a time series of RR intervals. After calculate these differences for the entire signal the obtained results is a discrete time series knowing as *RR tachogram* or *Heart Rate Variability* signal. It is important to note that this series of variability lacks uniformity in the distance between the points due to temporary differences between successive heartbeats, a feature that reflects the interaction of the sympathetic and parasympathetic system on heart activity. Figure 2.2 shows a graphical representation of the process involved to generate an RR time series from an ECG signal.

Figure 2.2: Schematic representation of the HRV signal generation from an ECG recording.

2.3 Assessment of the HRV signal

Heart Rate Variability (**HRV**) analysis is normally performed through a combination of linear and nonlinear methods. Linear methods have been categorized as methods in the time domain and methods in the frequency domain, while nonlinear methods correspond to a set of techniques to study the nonlinear dynamics of heart rate variability series.

The analysis in the time domain are the simplest way to extract features of the heart rate variability (HRV) signals by quantifying some indexes based on the statistics of the data contained in the signal. The most used time domain indices correspond to the *average value of all RR intervals* (**AVRR**), the *standard deviation of all RR intervals* (**SDNN**), the *square root of the mean squared differences of the successive RR intervals* (**rMSSD**), the *percentage of differences between adjacent N-N intervals that are by more than 50ms* (**pNN-50**). It is important to state that previous studies have shown that these time parameters are highly correlated with high frequency variations in heart rate (HR) [5, 31].

HRV signals exhibit an oscillatory behavior in which components of high and low frequency as a result of cardiovascular modulations performed by the sympathetic and parasympathetic nervous systems are mixed. Thus, methods of analysis in the frequency domain are used in order to quantify this type of information from the estimate of the power spectrum as a function of the frequencies contained in the signal [5, 32]. The calculation of the power spectrum or the power spectral density (**PSD**) of the HRV signal can be performed using parametric and nonparametric methods. Parametric methods usually estimate power spectrum through autoregressive models (AR) applied to the signal, while nonparametric methods using algorithms based on Fourier transform: *FFT* and *periodograms*. However, these methods require that the input signal will be evenly sampled, it means that all samples will be equally spaced in time. Then, in order to fulfill this requirement, it is necessary to perform a process of resampling on the RR series before make the spectral estimations. For HRV series it is recommended to make a cubic spline interpolation over the data using 4 *Hz* as a value for the re-sampling frequency (f_{r-s}) [5, 32, 33]. To avoid the resampling process over the HRV signals

and obtain an PSD estimate directly from the unevenly data, the algorithm proposed by Lomb in 1976 [34] and modified by Scargle in 1982 [35] can be used. This method estimates the PSD performing a normalization of sine or cosine functions independently in each sample of the input signal, allowing to retain all frequency characteristics and avoiding the induction of errors due to the interpolation process. Once the (PSD) is estimated, it is possible to quantify reliable information in the frequency domain integrating the spectrum in frequency bands previously defined. Four frequency bands, directly related to some physiological phenomena, have been considered as standard values in the frequency analysis of the HRV [5, 36, 37]. The values of these frequency bands correspond to:

- **VLF** $\rightarrow$ *Power in the very-low frequency range: 0.003-0.04 Hz for humans and 0.00-0.20 Hz for rats.*
- **LF** $\rightarrow$ *Power in the low frequency range: 0.04-0.15 Hz for humans and 0.20-0.75 Hz for rats.*
- **HF** $\rightarrow$ *Power in the high frequency range: 0.15-0.40 Hz for humans and 0.75-3.00 Hz for rats.*
- **LF/HF** $\rightarrow$ *The ratio of the power in the low frequency range to that in the high frequency range.*

In addition to the methods of analysis in the time and frequency domain, there are nonlinear methods which have been demonstrated to be extremely useful due to the non-stationary characteristics of the HRV signals. Most of these methods have their foundation in chaos theory and nonlinear dynamics, which allows to analyze HRV signals in a more complete way. In this regard, as was established by Peltola *et. al.* [17]: “*The basic concept of the nonlinear HRV methods is to try to capture the non-periodic behavior of the HRV and the complexity that exists inside the R-R interval dynamics*”. Among the most commonly used nonlinear methods for HRV analysis are: Poincaré plots [3], detrended fluctuations analysis (DFA) [20, 38], Lyapunov exponent [5, 39], sampling (SampEn) and multiscale entropy (MSE) [27, 40, 41, 42, 43, 44] and symbolic dynamics (SymDyn) [45, 46].

2.4 Artifacts in HRV signals

In ideal situations, adequate HRV analysis should be performed using RR interval series without any errors. However, there are several types of physiological and technical *artifacts* that constantly interfere in the generation and measurement processes of the electrocardiographic (ECG) signals. Thus, RR time series are obtained with low quality and erroneous information about the cardiovascular condition of the subject of analysis. Physiological artifacts are originated by erroneous behavior of the sinus node during the polarization and depolarization stages or by problems in the contractility of the heart tissue. These problems are normally presented by subjects with any type of cardiac disease; however, it can be appearing on extensive ECG records (greater than 2 hours) from healthy subjects [5]. However, there are some cardiac conditions under the presence of such physiological artifacts are normally found, some of these conditions are: arrhythmia, premature beats, atrial fibrillation, sinus tachycardia, sinus bradycardia, ventricular tachycardia and ventricular fibrillation [6, 17, 47]. It should also include the phenomena caused by electrical conduction problems as atrioventricular blocks (AV) or sinoatrial blocks (SA). Furthermore, technical artifacts often occur during the ECG measurement and processing stages in order to obtain the RR time series. In this way, poorly fastened electrodes, patient movement and sweating, poor contact between the patient-electrode interface, source power contamination, malfunction on R peaks detection algorithms and induction of electronic noise during conduction stage are some of the most common technical artifacts when working with ECG signals and they have strong influence during the quantification of HRV indices [5, 18, 28, 31]. The presence of these artifacts requires the visual examination of an expert in the area of ECG signals in order to make corrections on the data and obtain an adequate analysis of HRV. However, the most used approach is the use of *artifact correction methods* based on linear and nonlinear techniques of signal processing. Most of these methods inspect the time series removing, replacing or estimating new RR segments for those with defects. Despite, at this point there is not a well-defined methodology about which methods are the most suitable choice to deal with artifact corrections on HRV signals.

MATERIALS AND METHODS

3.1 Database

Real ECG signals with a length greater than ninety minutes were analyzed in order to generate an inter-beat interval (RR) time series data base. These signals were recorded with a sampling frequency of 1.000 Hz in continuous mode from three different groups of animals (Healthy rats, hypertensive rats and heart failure rats) using a telemetry measurement system (*PowerLab* system model ML870) associated directly to *LabChart Pro* Software *version 8.0* from **ADInstruments**. In this process a transmitter device is inserted surgically into the animal and the record signal is done remotely. These procedures of measurement, register and filtering were performed using *LabChart Pro* Software *version 8.0*.

Next step involves the QRS complexes detection, and more specifically all R peaks on each ECG signal in order to quantify the RR distances and generated the RR time series. This process was carried out using the processing modules incorporated in the *LabChart Pro* Software. Figure 3.1 shows an ECG signal from a healthy rat, the parameters used on the software to detect the QRS complexes and positions of the R peaks.

Figure 3.1: Parameters configuration of the LabChart ECG module to detect the QRS complex in a real ECG signal. The values are standard measures for rat ECG signals.

After processing all ECG signals and detect the R peaks the next step was extract the inter-beat interval distances, also known as RR intervals, and generate a set of time series that will be used further in our analysis of artifact correction and quantification of the Heart Rate Variability (HRV) parameters [6, 17, 33]. Using the **RR module** from *LabChat Pro* software this process could be done for each ECG signal analyzed in the previous stage. The same procedure was carried over thirty-seven (37) different signals.

Finally, all RR time series generated with this procedure previously described above were visually inspected with the aim to find segments without missing points with a length, of at least, ten-thousand (10.000) points. As a result, it was obtained sixteen series from the three different group of subjects, which were labelled from *Rat-01* until *Rat-16*.

3.2 Correction methods

In this study, we have selected the five most common artefact correction methods found in the literature about HRV analysis. These methods usually have been implemented to deal with problems like ectopic beats, noise and non-uniform sampling of the RR time series. However, it is well known that deletion, linear interpolation and cubic spline interpolation methods have been used in some investigations to correct missing beats in “artificial RR time series” [4, 16, 19]. In this work, these methods will evaluate the accuracy of this techniques working on the missing points case on real signals. A brief description of each correction method is presented to follow.

3.2.1 Deletion (DEL)

Deletion method simply removes the missing RR intervals in the time series and replaces each removed point(s) shifting the following RR intervals to the place of the deleted ones. After that, the corrected time series are shorter than the original. Figure corresponds to a schematic representation of the deletion method performance when the signal presents some missing points.

Figure 3.2: Schematic representation of deletion method performance.

3.2.2 Linear Interpolation (LI)

Interpolation means to compute points or values between ones that are known using the surrounding data ¹. In this order of ideas, **Linear interpolation** is the simplest method of interpolate values using straight line segments. Formally, given two known points located at (x_0, y_0) and (x_1, y_1) , and knowing that the linear interpolation is a straight line between these points, we can find for any x value in the range (x_0, x_1) their respective and unknown pair y using the equation 3.1. These process can be better understood looking the figure 3.3.

$$y = y_0 + (y_1 - y_0) * \frac{x - x_0}{x_1 - x_0} \quad (3.1)$$

Figure 3.3: Example of linear interpolation.

¹ *Formal definition of interpolation:* “Given a univariate function $f = f(x)$, interpolation is the process of using known values $f(x_0), f(x_1), f(x_2), \dots, f(x_n)$ to find values for $f(x)$ at points $x = x_i, i = 0, 1, 2, \dots, n$. In general, this technique involves the construction of a function $L(x)$ called the interpolant which agrees with f at the points $x = x_i$ and which is then used to compute the desired values.”

3.2.3 Cubic Spline Interpolation (CI)

The idea of a *spline* is basically joint two consecutive elements in a data series using a specific mathematical function. This function is used on each interval between all data points when the number of elements in a series is greater than two. The simplest spline is obtained connecting the data with a straight line as can be seen in section 3.2.2. The next simplest type of function is quadratic, and so on. In this order of ideas, the **cubic spline interpolation** is a method to joint/determine points in a data series using different cubic functions on each interval between data points. In general, this cubic functions correspond to third-order polynomials with the stipulation that the curve obtained be continuous and smooth.

The procedure for a cubic spline interpolation is to fit a piece-wise function of the form:

$$S(x) = \begin{cases} S_0(x) & \text{if } x_0 \leq x < x_1 \\ S_2(x) & \text{if } x_2 \leq x < x_3 \\ \vdots & \\ S_{n-1}(x) & \text{if } x_{n-1} \leq x < x_n \end{cases} \quad (3.2)$$

where S_i is a third degree polynomial defined by equation 3.3, for $i = 1, 2, \dots, n-1$.

$$S_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i \quad (3.3)$$

In this process the first and second derivatives of these $n-1$ equations are fundamental, and they are calculated as:

$$S'_i(x) = 3a_i(x - x_i)^2 + 2b_i(x - x_i) + c_i \quad (3.4)$$

$$S''_i(x) = 6a_i(x - x_i) + 2b_i \quad (3.5)$$

It is very important that all previous calculations must meet the following properties:

1. The piecewise function $S(x)$ will interpolate all data points.
2. $S(x)$ will be continuous on the interval $[x_0, x_n]$.
3. $S'(x)$ will be continuous on the interval $[x_0, x_n]$.
4. $S''(x)$ will be continuous on the interval $[x_0, x_n]$.

To follow, it is presented a schematic representation of an cubic spline interpolation process over a generic curve.

Figure 3.4: Schematic representation of an cubic spline interpolation process.

3.2.4 Moving Average Window (MAW)

The Moving Average Window is an algorithm that calculates the unweighted mean of the last n samples in order to predict the next point in a data series. The parameter n is often called the window size, because the algorithm can be thought of as a window that slides over the data points. Equation shows how to implement the procedure explained before, where $y[i]$ is the predicted point based on the $x[j]$ previous samples.

$$y[i] = \frac{1}{n} \sum_{j=i-n}^{i-1} x[j] \quad (3.6)$$

As an alternative, the group of points from the input signal can be chosen *symmetrically* around the predicted point as can be seen in the following example for a window size of $n = 6$

$$y[i] = \frac{x[i-3] + x[i-2] + x[i-1] + x[i+1] + x[i+2] + x[i+3]}{6}$$

To achieve this it is just necessary to change the limits of the summation in equation 3.6. In this case the the window length needs to be sufficiently wide in order to have sufficient points around predicted points and obtain a better estimate. It is important to know that symmetrical averaging requires that n be an odd number. In this research it has been used a symmetrical MAW with length width $n = 8$, it means that the average value for each point is calculated using four points at left and four at right. It is important to understand that this technique will average every point in the signal and not only the missing segments as occur with interpolation methods.

3.2.5 modified Moving Average Window (mMAW)

The modified Moving Average Window is a correction technique based on the moving average window method describes in 3.2.4, that uses a symmetrical moving average window only in the segments where the RR time series presents missing points. In this case the meaning objective is reduce at maximum the processing over the entire data series, and offer an alternative method to correct RR time series based on a well-known technique.

3.2.6 Nonlinear Predictive Interpolation (NPI)

The Nonlinear Predictive Interpolation (NPI) method is an algorithm designed by N. Lippman on 1994 [7], in order to solve the problem of ectopic beats present the analysis of an RR time series. It is able to perform corrections for single or sequences of ectopic beats with any length. In this research, the NPI method has been modified and used to correct the missing beats problem in inter-beat interval (RR) time series.

To perform the NPI method over RR time series is necessary to follow this steps:

1. Scan forward RR time series until the first missing point (RR interval) is found.
2. Define the segment length to be replaced, beginning with the first RR interval until the next RR intervals are found. The total amount of RR intervals to be replaced are called “beats to fill” (**B**), and becomes an input for the next steps.

3. A sequence of M RR intervals immediately before and N RR intervals immediately following the missing segment are used to define an $(M + N)_{ini}$ element test array. The M and N values could be different between them and they must to be specified as input parameters for the NPI algorithm. In figure 3.5 can be seen an example of how the element test array is conformed.
4. All available RR intervals are scanned, searching for segments of length $[M + B + N]$ without any missing RR interval. The M and N RR intervals in these sequences are used to construct $M + N$ element comparison arrays as:

$$\begin{bmatrix} (M + N)_1 \\ (M + N)_2 \\ \vdots \\ (M + N)_n \end{bmatrix} \quad (3.7)$$

5. All element comparison arrays of length $M + N$ found on the previous step are compared with the $(M + N)_{ini}$ element test array using a Cartesian distance metric, and the closest matching array is stored.

$$\begin{bmatrix} (M + N)_1^2 \\ (M + N)_2^2 \\ \vdots \\ (M + N)_n^2 \end{bmatrix} \leq (M + N)_{ini}^2 \quad (3.8)$$

6. From the closest matching comparison array determined above, the B RR intervals are extract and used to replace the missing segment found in the first step.
7. Repeat the previous procedure in order to find more missing segments or until reach the end of the time series.

Figure 3.5: *Element test array definition.*

3.3 HRV analysis methods

For analysis of the HRV series have been chosen some linear and nonlinear methods of great importance, based on a comprehensive literature review. Linear methods enable extraction and analysis of features in the time and frequency domain; However, they have problems to deal with factors such as non-stationary and different types of noise (noisy nature). In order to solve such problems, it is necessary to use nonlinear methods to achieve a comprehensive approach in the analysis of HRV.

3.3.1 Linear methods for HRV analysis

3.3.1.1 AVRR

This parameter corresponds to the mean value of all RR intervals in a data series. As we know, mean is a parameter for a distribution random variable, which is defined as a weighted average of its distribution. The **AVRR** is calculated as:

$$AVRR = \frac{1}{N} \sum_{i=1}^N RR_i \quad (3.9)$$

where N is the total number of all RR intervals in the time series.

3.3.1.2 SDRR

Standard Deviation of RR intervals is the measure of the variability or dispersion of a data set. This is a global index that correlates strongly with the total power of the time series, often $r > 0.9$ [48]. **SDRR** is calculated as

$$SDRR = \sqrt{\frac{1}{N} \sum_{i=1}^N (RR_i - \overline{RR})^2} \quad (3.10)$$

where $\overline{RR}$ is the arithmetic mean of the values RR_i , defined by equation 3.9.

3.3.1.3 RMSSD

RMSSD is the square root mean of successive RR intervals calculated using the equation 3.11, where N is the total number of all RR intervals in the series. This parameter plays an important role on heart rate variability analysis and it has been used in previous investigations as a significant indicator of both atrial fibrillation (*AF*) and sudden unexplained death in epilepsy (*SUDEP*) [31, 48].

$$RMSSD = \sqrt{\frac{1}{N-2} \sum_{i=3}^N (RR_{i+1} - RR_i)^2} \quad (3.11)$$

3.3.1.4 Lomb-Scargle Periodogram (LSP)

This method belongs to **Frequency-Domain** set of techniques used to analyze heart rate variability (HRV). Introduced in 1976 by Lomb [34] and modified in (1982) by Scargle [35], the method is used to estimate the *Power Spectral Density* (PSD) of an unevenly sampled signal. Lomb method has advantage over traditional methods based on the Fast Fourier Transform (*FFT*), because no re-sampling process is needed in order to create an evenly sampled representation and in the evaluation of the power spectrum the data are weighted on a point by point basis instead of weighted on a time interval basis [33, 49, 50]. Bellow, it is presented a short mathematical description about Lomb's method. Specific details can be consulted by the reader on references [32, 34, 35, 49, 50].

The procedure consist in fit a time series $x(n)$ *unevenly sampled* at times t_n by a weighted pair of cosine and sine waves, where each function is weighted by coefficients

a and b respectively. This fitting procedure must to be performed over the N samples of $x(n)$ obtained at times t_n and repeated for each frequency f . Equation 3.12 represents the fitting function to use in this approach, where coefficients a and b must to be determined in some point of the fitting procedure.

$$P(a, b, f, t_n) = a \cos(2\pi f t_n) + b \sin(2\pi f t_n) \quad (3.12)$$

Now, we fit P to signal x using minimization of the squared difference between them over all n samples (equation 3.13) and repeating this procedure for each frequency (f) data.

$$\epsilon^2 = \sum_{n=1}^N [P - x(n)]^2 \quad (3.13)$$

In order to find the minimum error in the minimization process, the next step sets to zero the partial derivatives of equation (3.13) respect to coefficients a and b , that is:

$$\frac{\partial \epsilon^2}{\partial a} = 0 \quad (3.14)$$

and

$$\frac{\partial \epsilon^2}{\partial b} = 0 \quad (3.15)$$

After evaluation of equations 3.14 and 3.15, and using some algebra over results; we obtain the following representations:

$$\sum_{n=1}^N x(n) \cos(2\pi f t_n) = a \sum_{n=1}^N \cos(2\pi f t_n)^2 + b \sum_{n=1}^N \cos(2\pi f t_n) \sin(2\pi f t_n) \quad (3.16)$$

and

$$\sum_{n=1}^N x(n) \sin(2\pi f t_n) = a \sum_{n=1}^N \cos(2\pi f t_n) \sin(2\pi f t_n) + b \sum_{n=1}^N \sin(2\pi f t_n)^2 \quad (3.17)$$

At this points, it is introduced the special feature of Lomb's algorithm: “*For each frequency f , every sample located at times t_n is shifted by an amount τ* ”. Then, in equations 3.16 and 3.17 t_n becomes $t_n - \tau$. In order to avoid errors introduced by the sine-cosine cross-terms, it can be chosen an optimal time shift value (τ) that makes them zero. This value is set as:

$$\tau = \frac{\tan^{-1} \left[\frac{\sum_{n=1}^N \sin(4\pi f t_n)}{\sum_{n=1}^N \cos(4\pi f t_n)} \right]}{4\pi f} \quad (3.18)$$

The value of variable τ is used in equations 3.16 and 3.17 in order to determine the form of a and b coefficients for each frequency. Making the respective algebra in those equations, it is found that:

$$a = \frac{\sum_{n=1}^N x(n) \cos(2\pi f(t_n - \tau))}{\sum_{n=1}^N \cos^2(2\pi f(t_n - \tau))} \quad (3.19)$$

$$b = \frac{\sum_{n=1}^N x(n) \sin(2\pi f(t_n - \tau))}{\sum_{n=1}^N \sin^2(2\pi f(t_n - \tau))} \quad (3.20)$$

Next step involves computation of the sum of squares of the sinusoidal signal form equation 3.12, this is $P^2(a, b, f, t - n)$, in order to obtain a representation that be proportional with the power spectrum S of $x(n)$ as a function of f :

$$S(f, a, b) = \sum_{n=1}^N P^2(a, b, f, t_n) = \sum_{n=1}^N [a^2 \cos^2(2\pi f(t_n - \tau)) + b^2 \sin^2(2\pi f(t_n - \tau)) + 0] \quad (3.21)$$

As the shift term τ was introduced in the mathematical formulation of S , all cross-terms are set to zero following the previous definitions. Now, the expressions for the coefficients a and b are replaced on equation 3.21 obtaining as result an expression for the power spectrum. This expression has the following form:

$$S(f) = \frac{\left[\sum_{n=1}^N x(n) \cos(2\pi f(t_n - \tau)) \right]^2}{\sum_{n=1}^N \cos^2(2\pi f(t_n - \tau))} + \frac{\left[\sum_{n=1}^N x(n) \sin(2\pi f(t_n - \tau)) \right]^2}{\sum_{n=1}^N \sin^2(2\pi f(t_n - \tau))} \quad (3.22)$$

In some cases the power spectrum expression from equation 3.22 is divided by 2 in order to get a power spectrum representation similar to the representaion obtained by the Fourier transform, or by $2\sigma^2$ ($\sigma^2 \rightarrow$ variance of $x(n)$) in order to obtain the standard normalized power spectrum and determine the statistical significance of the every peak in the spectral representation. Last representation corresponds to the modification of Lomb's algorithm made by Scargle in 1982 [35]

and used by other authors to estimate the power spectral components in the heart rate variability (HRV) [32, 33, 47, 49].

In the case of HRV it is well known that power spectrum covers a wide range of frequencies that contain relevant information about cardiovascular autonomic regulation. This information is located in some specific frequency bands and the power on each band is calculated by integrating the power spectral representation over the associated frequency range. Commonly used measures on HRV are [5, 13, 51]:

- **VLF** $\rightarrow$ *Power in the very-low frequency range: 0.003-0.04 Hz for humans and 0.00-0.20 Hz for rats.*
- **LF** $\rightarrow$ *Power in the low frequency range: 0.04-0.15 Hz for humans and 0.20-0.75 Hz for rats.*
- **HF** $\rightarrow$ *Power in the high frequency range: 0.15-0.40 Hz for humans and 0.75-3.00 Hz for rats.*
- **LF/HF** $\rightarrow$ *The ratio of the power in the low frequency range to that in the high frequency range.*

3.3.2 Nonlinear methods for HRV analysis

3.3.2.1 Detrended Fluctuation Analysis (DFA)

First introduced by Peng *et. al.* in 1995, the **DFA** method allows quantification of correlations on non-stationary time series [20]. These correlations are expressed and characterized by scaling properties and fractal structures. Along time it has been widely used in physiological time series in order to determine internal correlations that can be associated to some pathological conditions in the case of RR time series [2, 3, 21, 39, 52].

To calculate **DFA** for a given a time series $x(t)$ with $t = 1, 2, \dots, N$, the following steps should be performed:

1. The times series of length N is integrated using equation 3.23, where $\bar{x}$ is the

mean value of the original time series calculated by equation 3.24.

$$y(k) = \sum_{i=1}^k [x(i) - \bar{x}] \quad (3.23)$$

$$\bar{x} = \frac{1}{N} \sum_{j=1}^N x(j) \quad (3.24)$$

2. The integrated time series ($y(k)$) is divided into boxes of equal length n as shown in figure 3.6. On each box is calculated the local trend by fitting a regression line $y_n(k)$ in this data segment.

Figure 3.6: Local trend for each box of size n .

3. Next, the integrated time series ($y(k)$) is detrended by subtraction the local mean trend ($y_n(k)$) in each box. At this point the root mean square (RMS) fluctuation of this integrated and detrended time series is calculated using equation 3.25.

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - y_n(k)]^2} \quad (3.25)$$

4. This procedure is repeated for several different scales (all possible box sizes of length n) in order to provide a relationship between $F(n)$ and the box size n . The relationship is interpreted as the average fluctuation of the time series as a function of box size [20, 52].

Typically, **fluctuations** increases when the box size n increases. Now, if $\log(F(n))$ increases linearly as a function of $\log(n)$, the time series follows a scaling law. Under such conditions the fluctuation can be characterized by a *scaling exponent* α , the slope of the line relating $\log(F(n))$ vs $\log(n)$. Different values of α from specific types of time series as presented in table 3.1.

Scaling Exponent	Description of the Signal
$0 < \alpha < 0.5$	Small value followed more likely by a larger value and vice versa
$\alpha = 0.5$	Completely uncorrelated time series, that is, whitenoise
$0.5 < \alpha < 1.0$	Small value followed more likely by a small value and large value followed more likely by a large value (correlated).
$\alpha = 1.0$	$1/f$ type noise
$1.0 < \alpha < 1.5$	Noise of variable type
$\alpha = 1.5$	Brownian $1/f^2$ noise (integral of white noise)

Table 3.1: Known scaling exponents for different time series behavior.

Some times in short time series ($N \leq 10.000$) data depicted by $\log(F(n))$ vs $\log(n)$ cannot be fitted by only one *scaling coefficient* (α). This kind of situations conducts to a well known *crossover phenomena*, usually attributed to changes in the correlation properties of the time series at different time or space scales, though it can also be a result of nonstationarities in the time series [8, 53]. In this phenomenon the angular coefficient of the fitted line is altered from a given value n , from which it is necessary to make a new fit of the values $\log(F(n))$ vs $\log(n)$. After that, two scaling exponents are obtained in order to quantify the short-range (α_1) and long-range (α_2) correlations of the series.

Figure 3.7 shows an example of the *crossover phenomena* on DFA analysis from an RR inter-beat time series.

Figure 3.7: Example of a crossover phenomena for a DFA analysis of an RR time series for a normal subject. Red line correspond to short-range correlations and green line to long-range correlations.

3.3.2.2 MultiScale Entropy (MSE)

Multiscale entropy (**MSE**) analysis is a method to measure the complexity of a finite length time series [41, 42]. This method was developed for the analysis of any physiological time series, but principal applications are in the field of HRV analysis [54, 3, 18, 55, 56]. **MSE** analysis is based on the sample entropy (*SampEn*) measure but capable to perform calculations with any other type of entropy measures (*i.e.* *Approximate Entropy*). *SampEn* is a refinement of the approximate entropy family of statistics introduced by Pincus *et. al.* in 1991 [57], and widely used in the analysis of physiologic signals [40, 43, 44, 58]. However, in order to have good statistical reliability at higher scales, the number of data points must be at least 10.000 which becomes a limitation to use of **MSE** in many clinical studies.

Some investigations using surrogate time series obtained from inter-beat interval (RR) time series have demonstrated the advantages in the use of **MSE** over traditional methods of entropy analysis. It is established that only one measure of entropy is not enough to describe the complexity of a data set, and only using

a scaled approach the truly complexity can be described [41, 44, 58, 59].

In practice, the algorithm to calculate the **MSE** follow tow steps:

1. Given a time series x_i , $i = 1, \dots, N$, a *coarse-graining* process is applied over it. In this procedure, multiple coarse-grained time series ($y_j^{(\tau)}$) are constructed by averaging the data points with non-overlapping windows of increasing length τ . Each element of the coarse-grained series, $y_j^{(\tau)}$, is calculated using equation 3.26, where τ represents the scale factor and $1 \leq j \leq N/\tau$. The length of each coarse-grained time series is N/τ . For scale 1, the coarse-grained time series is simply the original time series.

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)+1}^{j\tau} x_i \quad (3.26)$$

To better comprehension of the coarse-graining process see figure 3.8.

2. Then, *SampEn* is calculated for each coarse-grained time series and plotted as a function of the scale factor τ .

Figure 3.8: Illustration of the coarse-graining procedure for scales 2 and 3 (Taken from [41])

3.3.2.3 Symbolic Dynamics (SymDyn)

Physiological time series usually shows complex structures which cannot be quantified or interpreted by linear methods due to the limited information about the underline dynamic system, whereas the nonlinear approach suffers from the curse

of dimensionality [14, 60]. In order to solve such problems the *Symbolic Dynamics* (**SymDyn**) method was proposed. The method attempts to characterize the original time series in a simple and coarser symbolic notation capable to retain the essential dynamic characteristics of the original time series. During this process there is loss of information contained in the time series, however, dynamical features are retained by the new coarse-grained representation[45, 46, 61].

The conversion of a time series into a set of symbols begins dividing the original signal into two or more levels, depending on how many symbols you want to use. There are several methods to make this process of conversion but the most used protocol involves the use of the signal average or the standard deviation. Figure 3.9 shows an example of symbolic quantification for an artificial signal.

Figure 3.9: Level's definition and symbolic conversion of a time series (Taken from [62])

As can be seen in the figure 3.9 the levels are selected as *A*, *B*, *C* and *D*; nonetheless, numbers or any other representations can be used in order to define a set of levels to quantify the signals. The next step involves the conversion of a time series into a symbol string and grouping these symbols in *words* of length *L*. A new *word* is always formed by stepping forward one step in the symbol string. An example of this process can be seen in figure 3.9, where the string formed by

A , B , C and D symbols is divided into *words* of length 3 for further analysis. This graphical representation is a very good representation of the original dynamics for any time series. In the final stage of the process, the formed patterns were grouped into four families identified according their variations from one symbol to the next. These families are identified as: 1) Patterns with no variations (**0V**: all symbols are equal, *i.e.* 111, 222, *etc.*), 2) patterns with one variation (**1V**: two consecutive symbols are equal and the remain is different, *i.e.* 112, 343, *etc.*), 3) patterns with two like variations (**2LV**: the three symbols form an ascending or descending ramp, *i.e.* 421, 542, *etc.*), and 4) patterns with two unlike variations (**2UV**: the three symbols forms a peak or a valley, *i.e.* 142, 315, *etc.*). The evaluation of the signal is made calculating the rates of occurrence of these families identified as **0V%**, **1V%**, **2LV%** and **2UV%**.

3.4 Statistical Tests

In order to compare the results obtained after performing some experimental procedure the field of statistics offers some useful techniques, that are based on the relationship between the data. In this order of ideas, sometimes it is better to use an independent test over the data set or use a paired test to know is our *null hypotheses* is rejected or not. As previously stated, there are many statistical methods to compare the properties of two or more groups; However, due to the nature of our data and the way in what the experiment was carried out, to follow the two comparison test used will be described.

3.4.1 Paired t-test

A paired t-test is normally used to compared tow groups or populations means, in where the observations of one population are paired with the observations of the other population. Examples of this type of situations are: students diagnostic test results before and after a particular module or course, or, a comparison of the efficiency of some treatment applied on the same subjects for a period of time [63, 64].

Let x and y datasets measured on the same test subjects before and after application

of some treatment. Then, in order to perform a paired t-test to know if the the null hypothesis that the true mean difference is zero, it means that there is not difference induced by the treatment, the procedure is as follows:

1. Calculate the difference ($d_i = y_i - x_i$) between the two observations on each pair, making sure you distinguish between positive and negative differences.
2. Calculate the mean difference, $\bar{d}$.
3. Calculate the standard deviation of the differences, s_d , and use this to calculate the standard error of the mean difference, $SE(\bar{d}) = \frac{s_d}{\sqrt{n}}$.
4. Calculate the t-statistic, which is given by $T = \frac{\bar{d}}{SE(\bar{d})}$. Under the null hypothesis, this statistic follows a t-distribution with $n - 1$ degrees of freedom.
5. Use tables of the t-distribution to compare your value for T to the t_{n-1} distribution. This will give the p-value for the paired t-test.

3.4.2 The Wilcoxon Signed Rank Sum Test

The Wilcoxon signed rank sum test is a nonparametric alternative to the two sample paired t-test, and it is known to be part of the family of distribution free test. This method is used to test the null hypotheses that the median of a distribution is equal to some value, normally zero. It can be used: (a) in place of a one sample t-test, (b) in place of a paired t-test or (c) for ordered categorical data where a numerical scale is inappropriate but where it is possible to rank the observations [63, 64].

To carried out the Wilcoxon signed rank sum test in the case of paired data, the correct procedure is:

1. State the null hypotheses, in this case that the median difference (M), is equal to zero.
2. Calculate each paired difference, $d_i = x_i - y_i$, where x_i and y_i are the pairs of observation.
3. Rank the d_i variable, ignoring the signs.

4. Label each rank with its sign, according to the sign of d_i .
5. Calculate W^+ , the sum of the ranks of the positives d_i , and W^- , the sum of the ranks of the negative d_i . As a check the sum of W^+ and W^- , should be equal to $\frac{n(n+1)}{2}$, where n is the number of pairs of observations in the sample.

Under the null hypothesis, we would expect the distribution of the differences to be approximately symmetric around zero and the the distribution of positives and negatives to be distributed at random among the ranks. Under this assumption, it is possible to work out the exact probability of every possible outcome for W . To carry out the test, we therefore proceed as follows:

6. Choose $W = \min(W^-, W^+)$.
7. Use tables of critical values for the Wilcoxon signed rank sum test to find the probability of observing a value of W or more extreme. Most tables give both one-sided and two-sided p-values.

On our data, these statistical tests were applied according to the result obtained after applied a *Shapiro Wilk* normality tests over each corrected and uncorrected signal. If the normality was preserved after each correction procedure, then a *paired t-test* was used; but, if the normality was losing after each correction, then the *Wilcoxon signed rank sum test* was choice. Then, any of the statistical procedures were used following the steps described before with the statistical hypotheses that the median of the difference of the paired data were zero, against that it was not zero. These procedure was carried out the fifteen HRV indexes calculated from corrected and uncorrected time series.

This statistical procedure shows what HRV indexes are more vulnerable to corrections and which correction method is more robust to the problem of the missing beats.

RESULTS

In the present section are presented the results obtained from two important stages. Firstly, the quantification of real amounts and types of missing points after analyze the RR time series described in chapter 3 was presented and used as parameters to validate the correction artifact techniques selected in this work. Secondly, indices of HRV are calculated for long RR time series before and after the presence of missing points. These data are compared via RMSE, and hypothesis tests are also used in order to determine the accuracy of each correction technique.

4.1 Missing points quantification

The identification and quantification of all missing points was made by visual inspection of the 37 RR time series generated from the ECG recordings. The total length of each signal was used in this inspections. The results shows that maximum quantity of missing points was around 5% of the total signal length and the most representative types of losses can be defined as: one non-consecutive random beat, three consecutive random beats and ten consecutive random beats. This analysis also allow us to determine that these three set of losses are distributed as:

- 1 non-consecutive beat = 79%
- 3 consecutive beats = 11%
- 10 consecutive beats = 10%

Notice that the sum of these loss percentages must be equal at maximum quantity of losses determined before.

Knowing that the maximum percentage of losses corresponds to 5% of the total signal length, it was selected another level of losses in order to know the correction methods performance in at least two different situations. fixed to 2.5% of points in total signal length. Table 4.1 describe the relation between signal length, maximum number of missing points and types of losses for the two chosen levels. These values are very important because they allow to work with parameters that describe real situations when in the field of HRV analysis.

Signal length (Points)	Max. percentage of losses	Quantity and type of points to be removed			
		Total points to remove	1 non-consecutive point (79%)	3 consecutive points (11%)	10 consecutive points (10%)
10.000	2.5%	250	198	27	25
	5.0%	500	395	55	50

Table 4.1: *Type and amount of losses for long RR time series found after visual inspection and calculated for signal with a length of 10.000 points.*

At this point we obtain thirty-two RR times series with random missing points divided in two levels, it means that each level contain sixteen time series and a length of 10.000 points.

4.2 Correction stage

Signals with missing points were processed, each one, with every correction method described in chapter 3, and this corrected version was used as inputs of the HRV analysis stage. Figure 4.1 shows the labels assigned to each signal after correction process. It is necessary remember that each correction technique was applied in both levels of losses.

4.3 Quantification of HRV parameters

Methods of analysis described in chapter 3 were tested on the group of signals the before the removal of points (*i.e. intact series or hereby control group*) and on each group of signals corrected by every method described in previous chapters. These results were divided, compared and analyzed in three parts: Time domain,

Figure 4.1: Stage of correction for RR time series and their results.

frequency domain and non-linearity. For each part results are presented in the form **mean** $\pm$ *standard deviation* ($m \pm \sigma$). However, the reader can find in the appendix 1 the results obtained for each subject before and after application of every correction techniques for the three parts mentioned above.

4.3.1 Time domain

AVRR, **SDRR** and **RMSSD** methods described in chapter (3) were used on the group of signals before the presence of missing points and over each corrected signal, for the two levels of losses. Table 4.2 shows mean values of the “control group” for the three parameters mentioned above. These data will be used for further comparisons against corrected values obtained after application of the respective correction techniques.

<i>Group \ Parameter</i>	<i>AVRR (ms)</i>	<i>SDRR (ms)</i>	<i>RMSSD (ms)</i>
Control	200.24 ± 17.41	8.06 ± 3.37	3.64 ± 1.09

Table 4.2: Time parameters for control group.

Following the quantification of time parameters for the “control” group (signals without the presence of missing points), the next steps involve quantification of same temporal parameters for the corrected groups of signals taking as reference

the two levels of losses. First correction was done using the method of *deletion* followed by *linear interpolation*, *cubic interpolation*, *moving average window*, *modified moving average window* and *nonlinear predictive interpolation* methods. The average value for each time parameter according the correction techniques are presented on table 4.3.

Percentage of losses:		2.5%		
<i>Correction method \ Parameter</i>		<i>AVRR</i> (ms)	<i>SDRR</i> (ms)	<i>RMSSD</i> (ms)
Deletion		200.24 $\pm$ 17.41	8.06 $\pm$ 3.38	3.63 $\pm$ 1.07
LI		200.24 $\pm$ 17.42	8.06 $\pm$ 3.37	3.57 $\pm$ 1.06
CI		200.24 $\pm$ 17.42	8.06 $\pm$ 3.37	3.57 $\pm$ 1.06
MAW		200.24 $\pm$ 14.41	7.56 $\pm$ 3.51	0.38 $\pm$ 0.09
mMAW		200.24 $\pm$ 17.41	8.05 $\pm$ 3.37	3.59 $\pm$ 1.06
NPI		200.24 $\pm$ 17.41	8.06 $\pm$ 3.37	3.62 $\pm$ 1.08

Percentage of losses:		5.0%		
<i>Correction method \ Parameter</i>		<i>AVRR</i> (ms)	<i>SDRR</i> (ms)	<i>RMSSD</i> (ms)
Deletion		200.23 $\pm$ 17.40	8.06 $\pm$ 3.37	3.62 $\pm$ 1.07
LI		200.24 $\pm$ 17.41	8.05 $\pm$ 3.37	3.49 $\pm$ 1.03
CI		200.24 $\pm$ 17.41	8.05 $\pm$ 3.37	3.49 $\pm$ 1.03
MAW		200.24 $\pm$ 17.41	7.56 $\pm$ 3.51	0.39 $\pm$ 0.09
mMAW		200.24 $\pm$ 17.41	8.04 $\pm$ 3.38	3.53 $\pm$ 1.04
NPI		200.24 $\pm$ 17.41	8.05 $\pm$ 3.38	3.60 $\pm$ 1.07

Table 4.3: *Time parameters calculated after application of each correction technique.*

From these results it can be observed a good corrections performed by, almost, all methods with nearly values in comparison with control group. At the same time we note that **MAW** method has appreciable differences for *SDRR* and *RMSSD* parameters but maintain a very good correlation on *AVRR* in comparison with the other techniques. These differences may be attributed to the process of smoothing carried out when the moving average is applied over the data, producing a drastically

variance reduction.

4.3.2 Frequency domain

At the same time, *Lomb's Periodogram* method was applied on each corrected and uncorrected signal in order to obtain their respective power spectral density (**PSD**) representations. Then, using the frequency band limits described in chapter 1 and calculating the area under the curve for these segments, it was possible to quantify the normalized power on very low frequency (**VLF**), low frequency (**LF**), high frequency (**HF**) and the low/high ratio (**LF/HF**) parameters, that allowed us the *HRV* characterization in the frequency domain.

Results are organized similarly as in time domain. First we show in the table 4.4 the average value of the parameters calculated for the signals without losses (“control group”) followed by the values obtained after each correction process. The corrected average values are presented on table 4.5 with their respective standard deviations.

Group \ Parameter	<i>VLF</i> (ms^2)	<i>LF</i> (ms^2)	<i>HF</i> (ms^2)	<i>LF/HF</i> (<i>n.u</i>)
Control	71.69 ± 65.02	1.95 ± 0.95	5.44 ± 2.48	0.36 ± 0.22

Table 4.4: *Frequency domain parameters for control group.*

From data on table 4.5 can be seen that the frequency parameter values for the corrected time series are very close in comparison with those obtained from control group, with exception that **LF**, **HF** and **LF/HF** values calculated by *MAW*, whose values are quite apart from the expected it. This difference is more appreciable when the amount of losses increase from 2.5% to 5.0%, given us to determine that this correction technique is not adequate to perform corrections over RR time series with the presence of missing points.

On the other hand, corrections using deletion gives parameter values greater than the control values when the level of losses increase; However, the interpolation methods and the modified moving average (mMAW) have a superior performance in both levels of losses. In the case of LI and CI, the results are in concordance with previous investigations on editing heart beats [4, 17, 24].

Percentage of losses:		2.5%			
<i>Correction method</i> \ Parameter		<i>VLF</i> (ms^2)	<i>LF</i> (ms^2)	<i>HF</i> (ms^2)	<i>LF/HF</i> (<i>n.u</i>)
Deletion		71.69 ± 65.50	1.95 ± 1.15	5.44 ± 2.49	0.38 ± 0.19
LI		71.27 ± 64.99	1.49 ± 0.96	4.41 ± 2.39	0.38 ± 0.23
CI		71.27 ± 64.99	1.49 ± 0.96	4.41 ± 2.39	0.38 ± 0.23
MAW		70.02 ± 64.46	0.17 ± 0.11	0.08 ± 0.06	2.65 ± 1.70
mMAW		71.25 ± 64.99	1.41 ± 0.92	4.43 ± 2.40	0.36 ± 0.22
NPI		71.25 ± 64.99	1.44 ± 0.94	4.52 ± 2.45	0.36 ± 0.22

Percentage of losses:		5.0%			
<i>Correction method</i> \ Parameter		<i>VLF</i> (ms^2)	<i>LF</i> (ms^2)	<i>HF</i> (ms^2)	<i>LF/HF</i> (<i>n.u</i>)
Deletion		71.48 ± 65.42	2.49 ± 1.55	6.42 ± 2.83	0.39 ± 0.17
LI		71.33 ± 64.99	1.54 ± 0.96	4.25 ± 2.30	0.41 ± 0.23
CI		71.34 ± 65.00	1.56 ± 0.96	4.26 ± 2.31	0.41 ± 0.23
MAW		70.02 ± 64.47	0.19 ± 0.12	0.09 ± 0.06	2.66 ± 1.53
mMAW		71.20 ± 64.96	1.39 ± 0.90	4.30 ± 2.32	0.36 ± 0.21
NPI		71.26 ± 65.12	1.45 ± 0.93	4.49 ± 2.49	0.36 ± 0.21

Table 4.5: *Frequency domain parameters after the application of correction methods.*

4.3.3 Nonlinear Domain Parameters

Nonlinear parameters for control group and corrected time series were calculated using *Detrended Fluctuation Analysis* (**DFA**), *MultiScale Entropy* (**MSE**) and *Symbolic Dynamics* (**SymDyn**) methods, all described previously in chapter (3). In DFA analysis the short (α_1) and long term (α_2) indices were calculated with a fixed crossover point equal to $n = 10$. This value is result of previous analysis in which more than 10 signals were analyzed by an minimization error algorithm specifically design to find the better crossover point. On the other hand, MSE measures were performed using as input parameters a tolerance factor (r) value of 0.15 the standard deviation of the time series, and a maximum number

of scales (τ) equal to 20. Finally, symbolic dynamic analysis was carried out using 6 levels ($\zeta = 6$) to quantify the time series, *words* of length $L = 3$ and non-overlapped windows with 300 points. On each window was performed the symbolic analysis and the median of the patterns were used to report the behavior of the analyzed data. These input parameters for DFA, MSE and SymDyn were kept for all time series analyzed in this work.

Figures 4.2, 4.3 and 4.4 presents individual plots for DFA, MSE and SymDyn analysis for one subject (*Rat-01*) before the presence of missing points and after application of each correction technique.

In the case of the MSE analysis were used the scale one (MSE_1) and the sum of the all scales (MSE_T) as representative parameters for our analysis; while in the symbolic dynamics the selected parameters were the indices for zero variations (OV) and two upper variations ($2UV$).

Figure 4.2: DFA results for Rat-01: (a) Before the presence of missing points and after perform corrections by (b) deletion, (c) LI, (d) CI, (e) MAW, (f) mMAW and (g) NPI.

Figure 4.3: MSE results for Rat-01: (a) Before the presence of missing points and after perform corrections by (b) deletion, (c) LI, (d) CI, (e) MAW, (f) mMAW and (g) NPI.

Figure 4.4: *SymDyn* results for Rat-01: (a) Before the presence of missing points and after perform corrections by (b) deletion, (c) LI, (d) CI, (e) MAW, (f) mMAW and (g) NPI.

In the three groups of figures, it can be seen a remarkable difference between results of the MAW and the other correction techniques, where an nonlinear behavior is clearly broken regarding the control results. Similar behavior was observed for all analyzed time series, because of this, the plots correspond only to one animal from our set of subjects.

DFA plots show a short and long term indices very similar to those obtained from control, and it is not easy to determine which technique has a better performance. However, quantification of the parameters for all subjects allows us to find significant differences on these measures. From MSE and SymDyn figures can be seen that deletion, mMAW and interpolation methods exhibits very similar patterns in comparison with control, but looking carefully, only the NPI method reproduce a close enough pattern in comparison with the data from the uncorrected time series (control).

Table 4.6 contents the average values of the nonlinear parameters for the control group and table 4.7 the average values for the same parameters quantified after every correction. These tables shows corrected values very close to control values with exception of those obtained by MAW technique. At same time, the small differences between the corrected data from first level of losses and the control data became greater when the amount of losses increased. Despite this, NPI exhibits better performance on both levels of losses with accurate values than those obtained by deletion, linear interpolation (LI), cubic interpolation (CI) and modified moving average (mMAW).

Group \ Parameter	$\alpha_1 (a.u)$	$\alpha_2 (a.u)$	$MSE_1 (a.u)$	$MSE_T (a.u)$	$\theta V (a.u)$	$2UV(a.u)$
Control	0.73 ± 0.24	1.13 ± 0.10	1.43 ± 0.50	20.13 ± 4.20	24.23 ± 11.90	34.94 ± 14.20

Table 4.6: Average value for Nonlinear parameters obtained from control group.

Percentage of losses:		2.5%						
<i>Correction method \ Parameter</i>	<i>Parameter</i>	$\alpha_1(a.u)$	$\alpha_2(a.u)$	$MSE_1(a.u)$	$MSE_T(a.u)$	$OV(a.u)$	$2UV(a.u)$	
Deletion		0.74 $\pm$ 0.22	1.12 $\pm$ 0.10	1.44 $\pm$ 0.50	20.37 $\pm$ 4.22	24.74 $\pm$ 11.62	34.51 $\pm$ 13.62	
	LI	0.75 $\pm$ 0.23	1.12 $\pm$ 0.10	1.42 $\pm$ 0.48	20.35 $\pm$ 4.21	25.53 $\pm$ 11.87	33.35 $\pm$ 13.47	
	CI	0.75 $\pm$ 0.23	1.12 $\pm$ 0.10	1.42 $\pm$ 0.48	20.37 $\pm$ 4.21	25.54 $\pm$ 11.95	33.25 $\pm$ 13.38	
	MAW	1.95 $\pm$ 0.12	1.20 $\pm$ 0.10	0.18 $\pm$ 0.07	15.79 $\pm$ 3.55	82.54 $\pm$ 4.25	2.15 $\pm$ 1.68	
	mMAW	0.73 $\pm$ 0.23	1.07 $\pm$ 0.22	1.41 $\pm$ 0.48	20.11 $\pm$ 4.12	25.30 $\pm$ 11.93	34.30 $\pm$ 13.78	
NPI		0.71 $\pm$ 0.27	1.13 $\pm$ 0.10	1.43 $\pm$ 0.50	20.24 $\pm$ 4.24	24.30 $\pm$ 11.73	34.73 $\pm$ 13.98	

Percentage of losses:		5.0%						
<i>Correction method \ Parameter</i>	<i>Parameter</i>	$\alpha_1(a.u)$	$\alpha_2(a.u)$	$MSE_1(a.u)$	$MSE_T(a.u)$	$OV(a.u)$	$2UV(a.u)$	
Deletion		0.74 $\pm$ 0.21	1.12 $\pm$ 0.12	1.44 $\pm$ 0.50	20.51 $\pm$ 4.16	25.67 $\pm$ 11.87	33.87 $\pm$ 13.63	
	LI	0.77 $\pm$ 0.22	1.12 $\pm$ 0.10	1.40 $\pm$ 0.47	20.53 $\pm$ 4.15	26.85 $\pm$ 11.56	32.00 $\pm$ 12.95	
	CI	0.77 $\pm$ 0.22	1.12 $\pm$ 0.10	1.41 $\pm$ 0.47	20.57 $\pm$ 4.14	26.97 $\pm$ 11.67	31.78 $\pm$ 12.75	
	MAW	1.94 $\pm$ 0.12	1.20 $\pm$ 0.10	0.18 $\pm$ 0.07	15.81 $\pm$ 3.46	82.27 $\pm$ 4.26	2.19 $\pm$ 1.69	
	mMAW	0.74 $\pm$ 0.23	1.53 $\pm$ 0.10	1.39 $\pm$ 0.47	20.14 $\pm$ 4.11	26.16 $\pm$ 11.60	33.74 $\pm$ 13.46	
NPI		0.74 $\pm$ 0.23	1.13 $\pm$ 0.10	1.43 $\pm$ 0.50	20.29 $\pm$ 4.25	25.00 $\pm$ 11.71	34.55 $\pm$ 13.82	

Table 4.7: Average values for nonlinear parameters after application of correction methods.

Corrected RR time series were compared against the control group data looking for statistical differences in order to determine what is the better correction technique to deal with the case of missing points on RR time series. **Two sample paired t-test** and **Wilcoxon signed-rank test** were performed for each nonlinear parameter using measures obtained before and after corrections [63, 64]. The use of a parametric or non-parametric comparison test was based on the decision of the **Shapiro-Wilk** normality test, used to check if the time series followed a normal distribution before corrections and after it. In addition, Root Mean Squared Error (**RMSE**) between control and corrected series was calculated using equation 4.1, where y_i are the control values and $\hat{y}_i$ the corrected values in reference to the same HRV parameter. All these results were used to compare the performance and accuracy of every correction technique on both levels of losses.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4.1)$$

Figure 4.5 displays RMSE bars for time domain parameters corrected for two levels of losses. It can be seen from 4.5-a and 4.5-b that corrections made using deletion has a negative impact on the **AVRR** parameter, causing a considerably error amount respect to other techniques. This behavior is attributed to the elimination of data made during the correction procedure. This results are similar with those reported previously by *Peltola et. al.* [17] and *Kim et. al.* [31] using artificial RR time series and the same method to edit artifacts that were induced manually. Other correction techniques presents error bars with similar trends and low variance at two levels of losses; but, it is interesting to see that correction using MAW for 5.0% shows an statistical difference in comparison with 2.5% due to the increased amount of missing points. On the other hand, **SDRR** and **RMSSD** parameters have a highly error when MAW is used to correct the time series. This error is statistically significant for the two levels of losses and extremely different in comparison with other techniques. However, we can see that the number of statistical differences are more in the second level of losses (5.0%) respect to the first level (2.5%).

In order to have a better analysis of the time domain parameters with low

Figure 4.5: RMSE bar plot for time domain parameters showing the performance for: (a) all techniques at 2.5% of losses, (b) all techniques at 5.0% of losses, (c) all techniques without MAW at 2.5% of losses and (d) all techniques without MAW at 5.0% of losses. (*: $p\text{-value} < 0.05$)

values of RMSE, it is presented on figures 4.5-c and 4.5-d these values without results calculated using the MAW method. It can be seen that RMSE values for 5.0% are bigger than those obtained for 2.5%; as well as the number of statistical differences found is greater as the number of losses increases. In the case of **AVRR** only deletion method has a big RMSE value compared with all other techniques. This behavior that was observed in both levels of losses. Despite this, no other method have significant differences in common and it can be established for this time domain parameter that only deletion and MAW methods are not acceptable for corrections. **SDRR** shows the same behavior for both levels of losses and, avoiding to use MAW, any other method give similar results. In contrast, **RMSSD** has better results when the time series are corrected by deletion or NPI methods with a considerable short RMSE value respect to the other techniques. Nevertheless, when the amount of missing points goes up all corrections exhibit more statistical differences in contrast to control measurements, showing that time parameters are very sensible to this type of artifact. In general, for time domain parameters all correction techniques, with exception of MAW, demonstrate a good performance for low and high level of losses; but, the NPI method shows more robustness to deal with this kind of artifacts, despite its statistical differences generated during the correction procedure.

Figures 4.6-a and 4.6-b show the behavior of the RMSE for the different frequency domain parameters as a function of the correction techniques in the two levels of losses selected in this work. Figures 4.6-c and 4.6-d have the same information, excepting the MAW technique due to the high error values that do not allow clearly analyze the behavior of other techniques when the RMSE value is very low. In general, it can be seen that for both levels of losses the corrections using MAW generate a much larger error than other techniques with statistical differences in all parameters when losses are increased from 2.5% to 5.0%. Additionally, a high RMSE value for deletion technique on both levels of losses is observed; with a greater contribution in the second level compared to the first. Furthermore, RMSE behavior for LI, CI, mMAW and NPI is very similar and only analyzing the figures 4.6-c and 4.6-d can be noticed a slight difference between these methods. Finally, it can be stated that, with the exception of deletion and MAW, any method considered in this

work can be used to correct RR time series with missing beats in order to obtain minimal changes in the frequency domain parameters.

(a) All techniques at 2.5%

(b) All techniques at 5.0%

(c) Without MAW at 2.5%

(d) Without MAW at 5.0%

Figure 4.6: RMSE bar plot for frequency domain parameters showing the performance for: (a) all techniques at 2.5% of losses, (b) all techniques at 5.0% of losses, (c) all techniques without MAW at 2.5% of losses and (d) all techniques without MAW at 5.0% of losses. (*: $p\text{-value} < 0.05$)

Finally, the behavior of RMSE for nonlinear parameters as a function of correction methods was analyzed following the same methodology used for the time and frequency domain parameters. Figure 4.7-*a* shows the results for the first level (2.5%) of losses and figure 4.7-*b* for the second level (5.0%). In both figures it can be seen again that the worst result is obtained for corrections with MAW, with exception of the parameter α_2 where the worst performance was obtained for mMAW. Also, it is interesting to note that these two techniques produce a considerable increase in the number of statistical differences when the level of losses are augmented from first to second level in all parameters considered.

Figures 4.7-*c* and 4.7-*d* present RMSE values without the presence of MAW method, in order to analyze the effects of other correction techniques that have low error values. In both cases, higher performance is exhibited by NPI method in all analyzed parameters, leading to low error values and fewer statistical differences. Furthermore, interpolation methods that showed good results in time and frequency are less promising for use with nonlinear parameters. In general, the great variance found in the nonlinear parameters makes it difficult to classify accurately the correction methods based on their RMSE values.

Figure 4.7: *RMSE bar plot for nonlinear parameters showing the performance for: (a) all techniques at 2.5% of losses, (b) all techniques at 5.0% of losses, (c) all techniques without MAW at 2.5% of losses and (d) all techniques without MAW at 5.0% of losses. (*: p -value < 0.05)*

Afterwards, analyzing these results is easy to notice that MAW is not a recommended correction method to apply in RR series with segment losses. In contrast, the modified version of MAW (mMAW) fits better the needs of this work giving an outstanding performance in the time and frequency domain, as well as in almost all non-linear parameters for the two levels of losses. Moreover, the corrections using linear and cubic interpolation work very well for the time and frequency parameters at both levels of losses, but the results do not follow this trend in the field of non-linearity. Therefore, its use is not recommended in this kind of corrections. Surprisingly, the deletion method showed better results than expected, showing in some cases RMSE values very close to those obtained by interpolation techniques. However, due to their behavior in the frequency domain it is not recommended to be used as a method of correction in such situations. Finally, the nonlinear predictive interpolation algorithm (NPI) presented better performance than all other methods and obtaining very low RMSE values in time and frequency domain as in non-linearity. Likewise, it can be seen that the number of significant differences in all parameters before and after corrections are less than those obtained with all other techniques.

CONCLUSIONS

Different correction techniques for cases of missing data (beats) in RR time series were studied. The effectiveness of each correction was determined by quantification of HRV indices in each time series with and without the presence of such type artifacts. The main findings found during the development of this work are presented:

- † Visual inspection of each RR time series allow us to determine that the maximum number of losses in signals recorded by telemetry, does not exceed 5% of the total data. This shows an improved performance of this method of signal registration in comparison with conventional methods reported in the literature.
- † The most representative types of losses for each analyzed RR time series corresponds to: one non-consecutive point, three consecutive points and ten consecutive points. Additionally it was found that these types of losses represent, respectively, 79%, 11% and 10% of the total amount of losses initially found (5%).
- † Using two different levels of loss was possible to determine that the moving average window (**MAW**) method is not suitable to make corrections in inter-beat interval (RR) time series. However, a slightly modified version of it gives very close results in comparison to those found using deletion and classical methods based on interpolation.
- † Correction procedures for the low level of losses using DEL, LI, CI and mMAW

exhibited values for each HRV index on time domain, frequency domain and non-linearity very similar to those obtained for the control series. Because of this, we recommended use these algorithms only when the level of losses not overpass a 2.5% of the total signal length.

† RMSE values between corrected and uncorrected time series calculated for HRV indexes in the time domain, frequency and non-linearity, allow us to determine that the method of predictive nonlinear interpolation (NPI) is best suited to recover lost segments of different sizes. Similarly, this method proved to be better correcting artifacts from both levels of losses (2.5% and 5.0%) used in this study.

† Finally, this work shows that a suitable correction of the RR interval time series reconstructs the dynamic characteristics of the records, without the need to reconsider data or perform new measurements in the subjects of work. Therefore, it is possible to obtain an adequate quantification of HRV indices in time, frequency and nonlinearity, which are comparable to those values obtained for RR time series without the presence of such artifacts.

BIBLIOGRAPHY*

- [1] SORNMO, L.; LAGUNA, P. *Bioelectrical signal processing in cardiac and neurological applications*. Amsterdam, Boston, Paris..., et al.: Elsevier Academic Press, 2005. ISBN 0-12-437552-9.
- [2] MÄKIKALLIO, T. H. et al. Clinical applicability of heart rate variability analysis by methods based on nonlinear dynamics. *Card. Electrophysiol. Rev.*, v. 6, p. 250–255, 2002. ISSN 13852264.
- [3] VOSS, A. et al. Methods derived from nonlinear dynamics for analysing heart rate variability. *Philos. Trans. A. Math. Phys. Eng. Sci.*, v. 367, n. 1887, p. 277–296, 2009. ISSN 1364-503X.
- [4] PELTOLA, M. A.; HUIKURI, H. V. Role of editing of R-R intervals in the analysis of heart rate variability. *Clin. Transl. Physiol.*, v. 3, p. 1–10, 2012.
- [5] MALIK, M. et al. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *European Heart Journal*, v. 17, p. 354–381, 1996. ISSN 0195-668X.
- [6] BERNTSON, G. G. et al. *Heart rate variability Origins, methods, and interpretive caveats*. [S.l.]: Psychophysiology, 1997. 623–648 p.
- [7] LIPPMAN, N.; STEIN, K. M.; LERMAN, B. B. Comparison of methods for removal of ectopy in measurement of heart rate variability. *Am. J. Physiol.*, v. 267, 1994.

*De acordo com a Associação Brasileira de Normas Técnicas. NBR 6023.

- [8] CHEN, Z. et al. Effect of nonstationarities on detrended fluctuation analysis. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, v. 65, n. 4, p. 1–15, 2002. ISSN 15393755.
- [9] CHEUNG, M. N. Detection of and Recovery from Errors in Cardiac Interbeat Intervals. *Psychophysiology*, v. 18(3), p. 341–346, 1981.
- [10] ALBRECHT, P.; COHEN, R. Estimation of heart rate power spectrum bands from real-world data: dealing with ectopic beats and noisy data. *Proceedings. Computers in Cardiology 1988*, 1988. ISSN 02766574.
- [11] MCNAMES, J.; THONG, T.; ABOY, M. Impulse rejection filter for artifact removal in spectral analysis of biomedical signals. *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, v. 1, p. 145–148, 2004. ISSN 1557-170X.
- [12] LEE, M.-Y.; YU, S.-N. Improving Discriminality in Heart Rate Variability Analysis Using Simple Artifact and Trend Removal Preprocessors. In: *32nd Annu. Int. Conf. IEEE EMBS*. [S.l.: s.n.], 2010.
- [13] THURAISINGHAM, R. A. Preprocessing RR interval time series for heart rate variability analysis and estimates of standard deviation of RR intervals. *Comput. Methods Programs Biomed.*, v. 83, n. 1, p. 78–82, 2006. ISSN 01692607.
- [14] MONTANO, N. et al. Heart rate variability explored in the frequency domain: A tool to investigate the link between heart and behavior. *Neurosci. Biobehav. Rev.*, v. 33, p. 71–80, 2009. ISSN 01497634.
- [15] VIKMAN, S. et al. Altered Complexity and Correlation Properties of R-R Interval Dynamics Before the Spontaneous Onset of Paroxysmal Atrial Fibrillation. *Circulation*, v. 100, p. 2079–2084, 1999.
- [16] SALO, M.; HUIKURI, H.; SEPPÄNEN, T. Ectopic Beats in Heart Rate Variability Analysis: Effects on Time and Frequency Domain Measures. *Ann. Noninvasive Electrocardiol.*, v. 6, n. 1, p. 5–17, 2001.

- [17] PELTOLA, M. A. et al. Effects and Significance of Premature Beats on Fractal Correlation Properties of R-R Interval Dynamics. *Ann. Noninvasive Electrocardiol.*, v. 9, n. 2, p. 127–135, 2004.
- [18] JARRIN, D. C. et al. Measurement fidelity of heart rate variability signal processing: The devil is in the details. *Int. J. Psychophysiol.*, v. 86, n. 1, p. 88–97, 2012. ISSN 01678760. Disponível em: <<http://dx.doi.org/10.1016/j.ijpsycho.2012.07.004>>.
- [19] WEJER, D. et al. IMPACT OF THE EDITING OF PATTERNS WITH ABNORMAL RR-INTERVALS ON THE ASSESSMENT OF HEART RATE VARIABILITY. v. 45, n. 8718, 2014.
- [20] PENG, C.-K. et al. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos*, v. 5, 1995.
- [21] KIM, K. K. et al. Effect of missing RR-interval data on nonlinear heart rate variability analysis. *Comput. Methods Programs Biomed.*, v. 106, p. 210–218, 2010.
- [22] SANTOS, L. D. et al. Application of an automatic adaptive filter for Heart Rate Variability analysis. *Med. Eng. Phys.*, v. 35, p. 1778–1785, 2013.
- [23] FERRI, R. et al. Heart rate variability during sleep in children with partial epilepsy. *Journal of Sleep Research*, v. 11, n. 2, p. 153–160, 2002. ISSN 09621105.
- [24] KIM, K. K. et al. The effect of missing RR-interval data on heart rate variability analysis in the frequency domain. *Physiol. Meas.*, v. 30, n. 10, p. 1039–1050, 2009. ISSN 0967-3334. Disponível em: <<http://iopscience.iop.org/0967-3334/30/10/005>>.
- [25] ACHTEN, J.; JEUKENDRUP, A. E. Heart rate monitoring: applications and limitations. *Sports medicine (Auckland, N.Z.)*, v. 33, n. 7, p. 517–538, 2003. ISSN 0112-1642.
- [26] BENITEZ, D. et al. The use of the Hilbert transform in ECG signal analysis. *Computers in Biology and Medicine*, v. 31, n. 5, p. 399–406, 2001. ISSN 00104825.

- [27] PAN, J.; TOMPKINS, W. J. A real-time QRS detection algorithm. *IEEE transactions on bio-medical engineering*, v. 32, n. 3, p. 230–236, 1985. ISSN 0018-9294.
- [28] RANGAYYAN, R. *Biomedical signal analysis : a case-study approach*. Piscataway, New Jersey: IEEE Press New York, 2002. (IEEE Press series on biomedical engineering). ISBN 0-471-20811-6.
- [29] MEHTA, S. S.; SAXENA, S. C.; VERMA, H. K. Computer-aided interpretation of ecg for diagnostics. *International Journal of Systems Science*, v. 27, n. 1, p. 43–58, 1996.
- [30] SASIKALA, P.; WAHIDABANU, R. Robust R Peak and QRS detection in Electrocardiogram using Wavelet Transform. *International Journal of Advanced Computer Science and Applications*, v. 1, p. 48–53, 2010. ISSN 21565570.
- [31] KIM, K. K. et al. Effect of missing RR-interval data on heart rate variability analysis in the time domain. *Physiol. Meas.*, v. 28, n. 28, p. 1485–1494, 2007. ISSN 0967-3334. Disponível em: <<http://iopscience.iop.org/0967-3334/28/12/003>>.
- [32] MOODY, G. B. Spectral Analysis of Heart Rate Without Resampling. In: *Comput. Cardiol. (2010)*. [S.l.: s.n.], 1993. p. 715–718.
- [33] CLIFFORD, G. D.; TARASSENKO, L. Quantifying errors in spectral estimates of HRV due to beat replacement and resampling. *IEEE Trans. Biomed. Eng.*, v. 52, n. 4, p. 630–638, 2005. ISSN 00189294.
- [34] LOMB, N. R. Least-squares frequency analysis of unequally spaced data. *Ap. Space Sci.*, v. 39, 1976.
- [35] SCARGLE, J. D. Studies in astronomical time series analysis. ii. statistical aspects of spectral analysis of unevenly spaced data. *Astrophys. J.*, v. 263, p. 835–853, 1982.
- [36] AKSELROD, S. et al. Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *Science*, v. 213, n. 4504, p. 220–222, 1981. ISSN 0036-8075.

- [37] HUIKURI, H. V. et al. Frequency domain measures of heart rate variability before the onset of nonsustained and sustained ventricular tachycardia in patients with coronary artery disease. *Circulation*, v. 87, n. 4, p. 1220–8, 1993. ISSN 0009-7322. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/8462148>>.
- [38] Rajendra Acharya, U. et al. Heart rate variability: a review. *Medical & biological engineering & computing*, v. 44, n. 12, p. 1031–51, 2006. ISSN 0140-0118. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/17111118>>.
- [39] ACHARYA, R. et al. Heart rate analysis in normal subjects of various age groups. *Biomedical engineering online*, v. 3, n. 1, p. 24, 2004. ISSN 1475-925X.
- [40] PAN, W.-Y. et al. Multiscale entropy analysis of heart rate variability for assessing the severity of sleep disordered breathing. *Entropy*, v. 17, n. 1, p. 231–243, 2015. ISSN 1099-4300. Disponível em: <<http://www.mdpi.com/1099-4300/17/1/231/>>.
- [41] COSTA, M.; GOLDBERGER, A. L.; PENG, C.-K. Multiscale entropy analysis of complex physiologic time series. *Physical Review Letters*, v. 89, n. 6, p. 068102, 2002. ISSN 0031-9007. Disponível em: <<http://link.aps.org/doi/10.1103/PhysRevLett.89.068102>>.
- [42] COSTA, M.; GOLDBERGER, a. L.; PENG, C.-K. Multiscale entropy to distinguish physiologic and synthetic rr time series. *Computers in cardiology*, v. 29, n. 1, p. 137–40, 2002. ISSN 0276-6574. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/14686448>>.
- [43] COSTA, M.; GOLDBERGER, A. L.; PENG, C.-K. Multiscale entropy analysis of biological signals. *Physical Review E*, v. 71, n. 2, p. 021906, 2005. ISSN 1539-3755. Disponível em: <<http://link.aps.org/doi/10.1103/PhysRevE.71.021906>>.
- [44] COSTA, M. D.; PENG, C.-K.; GOLDBERGER, A. L. Multiscale analysis of heart rate dynamics: Entropy and time irreversibility measures. *Cardiovascular Engineering*, v. 8, n. 2, p. 88–93, 2008. ISSN 1567-8822. Disponível em: <<http://link.springer.com/10.1007/s10558-007-9049-1>>.

- [45] PORTA, A. et al. Assessment of cardiac autonomic modulation during graded head-up tilt by symbolic analysis of heart rate variability. *AJP: Heart and Circulatory Physiology*, v. 293, n. 1, p. H702–H708, 2007. ISSN 0363-6135.
- [46] GUZZETTI, S. et al. Symbolic dynamics of heart rate variability: A probe to investigate cardiac autonomic modulation. *Circulation*, v. 112, n. 4, p. 465–470, 2005. ISSN 00097322.
- [47] SKOTTE, J. H.; KRISTIANSEN, J. Heart rate variability analysis using robust period detection. *Biomed. Eng. Online*, v. 13, 2014.
- [48] WANG, H.-M.; HUANG, S.-C. SDNN/RMSSD as a Surrogate for LF/HF: A Revised Investigation. *Model. Simul. Eng.*, Hindawi Publishing Corporation, v. 8, 2012.
- [49] CLIFFORD, G.; AZUAJE, F.; MCSHARRY, P. *Advanced Methods And Tools for ECG Data Analysis*. [S.l.]: Artech House, Inc., 2006. ISBN 1580539661.
- [50] FONSECA, D. et al. Lomb-Scargle Periodogram Applied to Heart Rate Variability Study. In: *Biosignals Biorobotics Conf.* [S.l.: s.n.], 2013. p. 1 – 4.
- [51] ERNST, G. *Heart Rate Variability*. [S.l.]: Springer-Verlag London, 2014.
- [52] CASTIGLIONI, P. et al. Local scale exponents of blood pressure and heart rate variability by detrended fluctuation analysis: effects of posture, exercise, and aging. *IEEE transactions on bio-medical engineering*, v. 56, n. 3, p. 675–84, 2009. ISSN 1558-2531. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/19389684>>.
- [53] MA, S. H. Crossover Phenomena in Detrended Fluctuation Analysis Used in Financial Markets. *Communications in Theoretical Physics*, v. 51, n. 2, p. 358–362, 2009. ISSN 02536102.
- [54] SILVA, L. E. V. et al. Heart rate complexity in sinoaortic-denervated mice. *Exp. Physiol.*, v. 100, p. 156–163, 2015. ISSN 09580670. Disponível em: <<http://doi.wiley.com/10.1113/expphysiol.2014.082222>>.

- [55] JOVIC, A.; BOGUNOVIC, N. Electrocardiogram analysis using a combination of statistical, geometric, and nonlinear heart rate variability features. *Artif. Intell. Med.*, v. 51, n. 3, p. 175–186, 2011. ISSN 09333657. Disponível em: <<http://dx.doi.org/10.1016/j.artmed.2010.09.005>>.
- [56] SILVA, L. E. V. et al. Multiscale entropy-based methods for heart rate variability complexity analysis. *Phys. A Stat. Mech. its Appl.*, v. 422, p. 143–152, 2015. ISSN 03784371. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0378437114010462>>.
- [57] PINCUS, S. M. Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences of the United States of America*, v. 88, n. 6, p. 2297–2301, 1991. ISSN 0027-8424. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/11607165>>.
- [58] COSTA, M.; HEALEY, J. Multiscale entropy analysis of complex heart rate dynamics: discrimination of age and heart failure effects. *Computers in Cardiology, 2003*, p. 705–708, 2003. ISSN 0276-6547.
- [59] COSTA, M.; GOLDBERGER, A. Generalized multiscale entropy analysis: Application to quantifying the complex volatility of human heartbeat time series. *Entropy*, v. 17, n. 3, p. 1197–1203, 2015. ISSN 1099-4300. Disponível em: <<http://www.mdpi.com/1099-4300/17/3/1197/>>.
- [60] VOSS, A. et al. *Symbolic Dynamics - a Powerful Tool in Non-Invasive Biomedical Signal Processing*. 2009.
- [61] TARKIAINEN, T. H. et al. Comparison of methods for editing of ectopic beats in measurements of short-term non-linear heart rate dynamics. *Clin. Physiol. Funct. Imaging*, v. 27, n. 2, p. 126–133, 2007. ISSN 14750961.
- [62] KAMATH, M.; MA, W.; UPTON, M. *Heart Rate Variability (HRV) Signal Analysis: Clinical Applications*. [S.l.]: CRC Press., 2012. ISBN 1439849803.
- [63] GIBBONS, J.; CHAKRABORTI, S. *Nonparametric Statistical Inference*. [S.l.]: Chapman and Hall/CRC, 2010. ISBN 9781420077612 - CAT# C7619.

- [64] DAVIS, C. *Statistical Methods for the Analysis of Repeated Measurements*. [S.l.]: Springer-Verlag New York, 2002. (Springer Texts in Statistics). ISBN 978-0-387-21573-0.
- [65] GOLDBERGER, A. et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*, v. 101, n. 23, p. E215–20, 2000. ISSN 1524-4539. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/10851218>>.

LONG SERIES RESULTS

In this part the reader can find the quantification of HRV parameters for every subject analysed in this work, classified in three stages: *Time domain*, *frequency domain* and *non-linear domain*. Also, we presented the results for the uncorrected animals before the corrected results.

A.1 Time domain results

First correction was done using the method of *deletion* and results are presented on table (A.2). In same way, *linear interpolation*, *cubic interpolation*, *moving average window*, *modified moving average window* and *nonlinear predictive interpolation* methods were used over the same signals containing missing points, and results are presented respectively on tables (A.3), (A.4), (A.5), (A.6) and (A.7).

A.2 Frequency domain results

Here we show in the table (A.8) the parameters calculated for the signals without losses followed by the values obtained after each correction process. In this order of ideas, table (A.9) contains the data corrected by deletion, table (A.10) by linear interpolation, table (A.11) by cubic interpolation, table (A.12) by moving average filter, table (A.13) by modified moving average filter and table (A.14) by nonlinear predictive interpolation.

Subject \ Parameter	<i>AVRR (ms)</i>	<i>SDRR (ms)</i>	<i>RMSSD (ms)</i>
Rat-01	218.145	9.832	5.142
Rat-02	215.897	12.352	2.875
Rat-03	215.688	6.770	4.074
Rat-04	215.241	14.394	4.684
Rat-05	194.812	10.627	2.589
Rat-06	189.636	14.556	2.526
Rat-07	177.319	5.756	3.893
Rat-08	188.456	8.266	5.550
Rat-09	203.592	5.446	5.049
Rat-10	194.616	4.351	2.312
Rat-11	194.060	8.407	2.875
Rat-12	205.800	6.002	4.321
Rat-13	168.634	5.605	2.917
Rat-14	229.322	4.998	3.727
Rat-15	215.376	4.683	3.499
Rat-16	177.280	6.956	2.175

Table A.1: *Time parameters for control group.*

Correction technique:				<i>Deletion</i>		
Percentage of losses:				<i>2.5%</i>		
Subject \ Parameter	<i>AVRR</i>	<i>SDRR</i>	<i>RMSSD</i>	<i>AVRR</i>	<i>SDRR</i>	<i>RMSSD</i>
Rat-01-DEL	218.168	9.828	5.110	218.124	9.827	5.085
Rat-02-DEL	215.887	12.358	2.876	215.790	12.377	2.862
Rat-03-DEL	215.694	6.792	4.053	215.692	6.770	4.028
Rat-04-DEL	215.224	14.401	4.670	215.255	14.372	4.668
Rat-05-DEL	194.843	10.655	2.589	194.783	10.598	2.610
Rat-06-DEL	189.588	14.567	2.536	189.643	14.574	2.524
Rat-07-DEL	177.324	5.755	3.863	177.332	5.750	3.857
Rat-08-DEL	188.470	8.247	5.524	188.498	8.255	5.484
Rat-09-DEL	203.613	5.452	5.044	203.593	5.460	5.030
Rat-10-DEL	194.621	4.355	2.322	194.621	4.344	2.313
Rat-11-DEL	194.069	8.404	2.883	194.042	8.433	2.882
Rat-12-DEL	205.803	6.011	4.327	205.771	5.981	4.314
Rat-13-DEL	168.638	5.599	2.936	168.636	5.609	2.939
Rat-14-DEL	229.307	4.986	3.727	229.322	4.998	3.727
Rat-15-DEL	215.357	4.678	3.473	215.358	4.674	3.445
Rat-16-DEL	177.261	6.943	2.183	177.266	6.977	2.193

Table A.2: *Time parameters after correction with deletion method.*

Correction technique:		<i>Linear Interpolation</i>					
Percentage of losses:		<i>2.5%</i>			<i>5.0%</i>		
Subject \Parameter		<i>AVRR</i>	<i>SDRR</i>	<i>RMSSD</i>	<i>AVRR</i>	<i>SDRR</i>	<i>RMSSD</i>
Rat-01-LI		218.160	9.818	5.029	218.147	9.821	4.910
Rat-02-LI		215.897	12.350	2.826	215.902	12.340	2.758
Rat-03-LI		215.680	6.771	3.989	215.666	6.748	3.900
Rat-04-LI		215.234	14.384	4.586	215.244	14.382	4.497
Rat-05-LI		194.810	10.628	2.536	194.820	10.622	2.506
Rat-06-LI		189.626	14.547	2.483	189.631	14.545	2.426
Rat-07-LI		177.319	5.754	3.806	177.316	5.743	3.739
Rat-08-LI		188.451	8.252	5.432	188.472	8.238	5.306
Rat-09-LI		203.595	5.432	4.957	203.610	5.436	4.862
Rat-10-LI		194.610	4.346	2.277	194.614	4.351	2.223
Rat-11-LI		194.055	8.412	2.825	194.068	8.406	2.772
Rat-12-LI		205.802	5.987	4.247	205.791	5.951	4.138
Rat-13-LI		168.633	5.605	2.877	168.635	5.594	2.821
Rat-14-LI		229.325	4.981	3.647	229.316	4.972	3.590
Rat-15-LI		215.375	4.670	3.420	215.372	4.668	3.338
Rat-16-LI		177.276	6.954	2.138	177.279	6.950	2.107

Table A.3: *Time parameters after correction using LI.*

Correction technique:		Cubic Interpolation					
Percentage of losses:		2.5%			5.0%		
Subject \Parameter		AVRR	SDRR	RMSSD	AVRR	SDRR	RMSSD
Rat-01-CI		218.160	9.819	5.029	218.150	9.824	4.911
Rat-02-CI		215.896	12.351	2.827	215.905	12.343	2.759
Rat-03-CI		215.679	6.773	3.989	215.665	6.752	3.901
Rat-04-CI		215.231	14.384	4.586	215.245	14.382	4.498
Rat-05-CI		194.810	10.628	2.536	194.821	10.624	2.507
Rat-06-CI		189.626	14.547	2.483	189.632	14.547	2.426
Rat-07-CI		177.319	5.756	3.806	177.317	5.745	3.739
Rat-08-CI		188.451	8.253	5.432	188.472	8.240	5.306
Rat-09-CI		203.596	5.432	4.958	203.614	5.441	4.863
Rat-10-CI		194.609	4.348	2.277	194.614	4.351	2.223
Rat-11-CI		194.054	8.412	2.825	194.069	8.407	2.773
Rat-12-CI		205.802	5.987	4.247	205.790	5.952	4.139
Rat-13-CI		168.632	5.606	2.878	168.636	5.595	2.821
Rat-14-CI		229.326	4.986	3.648	229.318	4.975	3.591
Rat-15-CI		215.374	4.669	3.420	215.371	4.670	3.339
Rat-16-CI		177.276	6.953	2.138	177.280	6.950	2.107

Table A.4: *Time parameters after correction using CI.*

Correction technique:		<i>Moving Average Window</i>					
Percentage of losses:		<i>2.5%</i>			<i>5.0%</i>		
Subject \ Parameter		<i>AVRR</i>	<i>SDRR</i>	<i>RMSSD</i>	<i>AVRR</i>	<i>SDRR</i>	<i>RMSSD</i>
Rat-01-MAW		218.156	9.150	0.544	218.135	9.180	0.577
Rat-02-MAW		215.892	12.113	0.404	215.887	12.107	0.409
Rat-03-MAW		215.684	6.193	0.409	215.675	6.192	0.420
Rat-04-MAW		215.235	13.971	0.555	215.241	13.974	0.566
Rat-05-MAW		194.810	10.386	0.359	194.812	10.376	0.364
Rat-06-MAW		189.629	14.388	0.369	189.637	14.389	0.377
Rat-07-MAW		177.319	5.282	0.299	177.319	5.279	0.306
Rat-08-MAW		188.454	7.588	0.432	188.467	7.590	0.446
Rat-09-MAW		203.589	4.554	0.363	203.605	4.560	0.373
Rat-10-MAW		194.615	4.041	0.242	194.614	4.056	0.242
Rat-11-MAW		194.056	8.134	0.343	194.066	8.122	0.344
Rat-12-MAW		205.803	5.119	0.389	205.793	5.118	0.400
Rat-13-MAW		168.636	5.079	0.320	168.635	5.082	0.324
Rat-14-MAW		229.318	4.043	0.445	229.312	4.045	0.435
Rat-15-MAW		215.375	4.143	0.328	215.374	4.150	0.335
Rat-16-MAW		177.276	6.752	0.268	177.277	6.751	0.268

Table A.5: Time parameters after correction using MAW.

Correction technique:		<i>Modified Moving Average Window</i>					
Percentage of losses:		<i>2.5%</i>			<i>5.0%</i>		
Subject \ Parameter		<i>AVRR</i>	<i>SDRR</i>	<i>RMSSD</i>	<i>AVRR</i>	<i>SDRR</i>	<i>RMSSD</i>
Rat-01-mMAW		218.133	9.804	5.064	218.150	9.815	4.985
Rat-02-mMAW		215.897	12.350	2.843	215.893	12.338	2.804
Rat-03-mMAW		215.691	6.755	3.997	215.677	6.729	3.936
Rat-04-mMAW		215.230	14.379	4.605	215.235	14.374	4.544
Rat-05-mMAW		194.809	10.622	2.554	194.814	10.606	2.548
Rat-06-mMAW		189.645	14.552	2.501	189.638	14.541	2.460
Rat-07-mMAW		177.319	5.746	3.823	177.315	5.742	3.743
Rat-08-mMAW		188.456	8.250	5.450	188.461	8.238	5.374
Rat-09-mMAW		203.588	5.421	4.963	203.589	5.400	4.871
Rat-10-mMAW		194.611	4.343	2.279	194.620	4.338	2.252
Rat-11-mMAW		194.056	8.397	2.838	194.059	8.396	2.787
Rat-12-mMAW		205.812	5.997	4.283	205.812	5.961	4.193
Rat-13-mMAW		168.630	5.593	2.888	168.629	5.590	2.863
Rat-14-mMAW		229.320	4.972	3.696	229.317	4.947	3.646
Rat-15-mMAW		215.375	4.672	3.436	215.373	4.659	3.380
Rat-16-mMAW		177.285	6.950	2.147	177.284	6.944	2.120

Table A.6: Time parameters after correction using modified version of MAW algorithm.

Correction technique:		<i>Nonlinear Predictive Interpolation</i>					
Percentage of losses:		<i>2.5%</i>			<i>5.0%</i>		
Subject \ Parameter		<i>AVRR</i>	<i>SDRR</i>	<i>RMSSD</i>	<i>AVRR</i>	<i>SDRR</i>	<i>RMSSD</i>
Rat-01-NPI		218.151	9.828	5.112	218.129	9.838	5.092
Rat-02-NPI		215.887	12.349	2.872	215.885	12.333	2.870
Rat-03-NPI		215.683	6.774	4.061	215.663	6.745	4.028
Rat-04-NPI		215.240	14.383	4.639	215.242	14.396	4.621
Rat-05-NPI		194.809	10.629	2.578	194.818	10.621	2.587
Rat-06-NPI		189.633	14.555	2.529	189.634	14.556	2.516
Rat-07-NPI		177.321	5.749	3.864	177.320	5.740	3.837
Rat-08-NPI		188.460	8.266	5.528	188.475	8.253	5.494
Rat-09-NPI		203.588	5.444	5.031	203.596	5.433	4.995
Rat-10-NPI		194.619	4.346	2.309	194.613	4.362	2.290
Rat-11-NPI		194.047	8.426	2.877	194.071	8.404	2.861
Rat-12-NPI		205.799	5.995	4.306	205.781	5.957	4.255
Rat-13-NPI		168.636	5.600	2.924	168.630	5.607	2.906
Rat-14-NPI		229.318	4.976	3.698	229.316	4.972	3.697
Rat-15-NPI		215.378	4.676	3.470	215.376	4.661	3.438
Rat-16-NPI		177.273	6.953	2.177	177.278	6.953	2.169

Table A.7: Time parameters after correction using NPI algorithm.

Subject \ Parameter	<i>VLF</i> (ms^2)	<i>LF</i> (ms^2)	<i>HF</i> (ms^2)	<i>LF/HF</i> (<i>n.u</i>)
Rat-01	92.221	3.091	8.183	0.378
Rat-02	155.139	1.582	2.422	0.653
Rat-03	41.752	1.270	4.741	0.268
Rat-04	201.401	2.311	7.176	0.322
Rat-05	107.053	1.626	2.344	0.694
Rat-06	209.845	1.327	2.312	0.574
Rat-07	28.672	0.546	4.537	0.120
Rat-08	59.556	0.763	9.180	0.083
Rat-09	21.266	0.580	8.233	0.070
Rat-10	17.272	0.316	1.891	0.167
Rat-11	70.170	1.117	2.931	0.381
Rat-12	26.199	2.955	6.187	0.478
Rat-13	26.224	1.923	3.399	0.566
Rat-14	17.985	2.853	4.160	0.686
Rat-15	18.263	0.408	3.472	0.118
Rat-16	47.456	0.391	1.797	0.218

Table A.8: Frequency parameters for control group.

Correction technique:					<i>Deletion</i>			
Percentage of losses:					2.5%			
Subject \ Parameter	<i>VLF</i> (ms^2)	<i>LF</i> (ms^2)	<i>HF</i> (ms^2)	<i>LF/HF</i> (<i>n.u</i>)	<i>VLF</i> (ms^2)	<i>LF</i> (ms^2)	<i>HF</i> (ms^2)	<i>LF/HF</i> (<i>n.u</i>)
Rat-01-DEL	92.403	3.907	8.607	0.454	91.011	4.952	9.580	0.517
Rat-02-DEL	156.270	2.544	4.396	0.579	155.779	3.863	6.326	0.611
Rat-03-DEL	41.477	1.706	5.186	0.329	42.269	1.777	5.433	0.327
Rat-04-DEL	202.405	3.777	9.976	0.379	202.184	5.595	12.458	0.449
Rat-05-DEL	108.428	2.307	3.723	0.620	107.103	3.100	5.770	0.537
Rat-06-DEL	211.541	2.603	5.237	0.497	211.934	3.702	8.576	0.432
Rat-07-DEL	28.784	0.742	4.521	0.164	8.916	0.929	4.423	0.210
Rat-08-DEL	60.134	1.123	9.927	0.113	60.345	1.579	10.804	0.146
Rat-09-DEL	21.232	0.826	8.209	0.101	21.271	0.943	8.444	0.112
Rat-10-DEL	17.122	0.480	2.184	0.220	17.336	0.618	2.388	0.259
Rat-11-DEL	70.094	1.722	4.052	0.425	69.544	2.389	4.981	0.480
Rat-12-DEL	26.357	3.149	6.594	0.478	26.342	3.448	6.985	0.494
Rat-13-DEL	26.286	2.150	3.799	0.566	5.835	2.231	4.545	0.491
Rat-14-DEL	18.266	2.977	4.366	0.682	17.980	3.119	4.648	0.671
Rat-15-DEL	18.567	0.524	3.646	0.144	18.350	0.684	3.976	0.172
Rat-16-DEL	47.723	0.723	2.545	0.284	47.479	0.957	3.443	0.278

Table A.9: Frequency domain parameters after correction using deletion.

Correction technique:					<i>Linear Interpolation</i>			
Percentage of losses:					2.5%			
Subject \ Parameter	<i>VLF</i> (ms^2)	<i>LF</i> (ms^2)	<i>HF</i> (ms^2)	<i>LF/HF</i> (<i>n.u</i>)	<i>VLF</i> (ms^2)	<i>LF</i> (ms^2)	<i>HF</i> (ms^2)	<i>LF/HF</i> (<i>n.u</i>)
Rat-01-LI	92.055	3.248	7.863	0.413	92.407	3.302	7.622	0.433
Rat-02-LI	155.297	1.605	2.359	0.680	155.444	1.627	2.290	0.710
Rat-03-LI	41.744	1.332	4.566	0.292	41.674	1.363	4.429	0.308
Rat-04-LI	201.297	2.359	6.909	0.341	201.364	2.476	6.633	0.373
Rat-05-LI	107.134	1.652	2.260	0.731	107.092	1.613	2.218	0.728
Rat-06-LI	209.632	1.323	2.243	0.590	209.630	1.354	2.151	0.629
Rat-07-LI	28.764	0.605	4.422	0.137	28.732	0.667	4.231	0.158
Rat-08-LI	59.588	0.875	8.894	0.098	59.752	1.068	8.540	0.125
Rat-09-LI	21.224	0.695	7.925	0.088	21.190	0.803	7.708	0.104
Rat-10-LI	17.238	0.324	1.826	0.177	17.272	0.343	1.752	0.196
Rat-11-LI	70.255	1.133	2.844	0.398	70.045	1.174	2.723	0.431
Rat-12-LI	26.154	2.945	6.037	0.488	26.370	3.032	5.757	0.527
Rat-13-LI	26.145	1.925	3.294	0.584	26.330	1.976	3.194	0.619
Rat-14-LI	18.017	2.875	4.028	0.714	18.131	2.875	3.860	0.745
Rat-15-LI	18.254	0.456	3.335	0.137	18.259	0.514	3.261	0.158
Rat-16-LI	47.509	0.408	1.739	0.235	47.510	0.411	1.685	0.244

Table A.10: Frequency domain parameters after correction using LI.

Correction technique:					<i>Cubic Interpolation</i>			
Percentage of losses:					2.5%			
Subject \ Parameter	<i>VLF</i> (ms^2)	<i>LF</i> (ms^2)	<i>HF</i> (ms^2)	<i>LF/HF</i> (<i>n.u</i>)	<i>VLF</i> (ms^2)	<i>LF</i> (ms^2)	<i>HF</i> (ms^2)	<i>LF/HF</i> (<i>n.u</i>)
Rat-01-CI	92.047	3.272	7.870	0.416	92.530	3.344	7.638	0.438
Rat-02-CI	155.312	1.612	2.362	0.683	155.528	1.642	2.300	0.714
Rat-03-CI	41.748	1.341	4.571	0.293	41.701	1.382	4.438	0.311
Rat-04-CI	201.281	2.372	6.915	0.343	201.306	2.496	6.644	0.376
Rat-05-CI	107.139	1.657	2.264	0.732	107.125	1.627	2.224	0.732
Rat-06-CI	209.631	1.329	2.247	0.592	209.651	1.360	2.157	0.630
Rat-07-CI	28.761	0.611	4.429	0.138	28.733	0.681	4.244	0.161
Rat-08-CI	59.589	0.886	8.900	0.100	59.775	1.098	8.557	0.128
Rat-09-CI	21.247	0.705	7.933	0.089	21.183	0.822	7.723	0.106
Rat-10-CI	17.227	0.328	1.828	0.179	17.273	0.346	1.755	0.197
Rat-11-CI	70.265	1.135	2.847	0.399	70.036	1.185	2.730	0.434
Rat-12-CI	26.165	2.957	6.042	0.489	26.394	3.061	5.768	0.531
Rat-13-CI	26.150	1.926	3.299	0.584	26.340	2.011	3.204	0.628
Rat-14-CI	18.024	2.890	4.031	0.717	18.137	2.897	3.871	0.748
Rat-15-CI	18.256	0.459	3.338	0.138	18.262	0.525	3.266	0.161
Rat-16-CI	47.502	0.410	1.740	0.235	47.518	0.413	1.689	0.245

Table A.11: Frequency domain parameters after correction using CI.

Correction technique:					<i>Moving Average Window</i>			
Percentage of losses:					2.5%			
Subject \ Parameter	<i>VLF</i> (ms^2)	<i>LF</i> (ms^2)	<i>HF</i> (ms^2)	<i>LF/HF</i> (<i>n.u</i>)	<i>VLF</i> (ms^2)	<i>LF</i> (ms^2)	<i>HF</i> (ms^2)	<i>LF/HF</i> (<i>n.u</i>)
Rat-01-MAW	89.206	0.484	0.166	2.906	89.323	0.508	0.166	3.052
Rat-02-MAW	153.278	0.160	0.118	1.362	153.296	0.172	0.117	1.472
Rat-03-MAW	39.978	0.127	0.045	2.812	40.044	0.144	0.047	3.052
Rat-04-MAW	198.161	0.270	0.214	1.259	198.269	0.304	0.211	1.442
Rat-05-MAW	105.890	0.185	0.110	1.680	105.672	0.187	0.110	1.708
Rat-06-MAW	208.417	0.161	0.181	0.888	208.412	0.172	0.186	0.924
Rat-07-MAW	28.417	0.085	0.052	1.635	28.372	0.088	0.052	1.701
Rat-08-MAW	58.568	0.170	0.118	1.438	58.655	0.191	0.117	1.627
Rat-09-MAW	20.862	0.048	0.061	0.779	20.856	0.054	0.062	0.867
Rat-10-MAW	16.775	0.052	0.025	2.058	16.789	0.049	0.025	1.945
Rat-11-MAW	69.121	0.234	0.068	3.443	68.985	0.243	0.068	3.572
Rat-12-MAW	25.485	0.192	0.057	3.352	25.605	0.199	0.057	3.479
Rat-13-MAW	25.671	0.264	0.038	7.005	25.731	0.281	0.045	6.247
Rat-14-MAW	16.255	0.200	0.038	5.307	16.248	0.208	0.040	5.139
Rat-15-MAW	17.284	0.073	0.032	2.323	17.226	0.079	0.033	2.423
Rat-16-MAW	46.893	0.074	0.018	4.135	46.827	0.074	0.019	3.950

Table A.12: Frequency domain parameters after correction using MAW.

Correction technique:		<i>modified Moving Average Window</i>							
Percentage of losses:		<i>2.5%</i>				<i>5.0%</i>			
Subject \ Parameter		<i>VLF (ms²)</i>	<i>LF (ms²)</i>	<i>HF (ms²)</i>	<i>LF/HF (n.u)</i>	<i>VLF (ms²)</i>	<i>LF (ms²)</i>	<i>HF (ms²)</i>	<i>LF/HF (n.u)</i>
Rat-01-mMAW		92.066	3.055	7.927	0.385	92.326	3.058	7.720	0.396
Rat-02-mMAW		155.189	1.541	2.383	0.647	154.948	1.497	2.339	0.640
Rat-03-mMAW		41.701	1.257	4.583	0.274	41.538	1.193	4.471	0.267
Rat-04-mMAW		201.175	2.287	6.953	0.329	201.306	2.240	6.763	0.331
Rat-05-mMAW		107.062	1.591	2.280	0.697	106.821	1.527	2.278	0.670
Rat-06-mMAW		209.785	1.258	2.277	0.552	209.597	1.233	2.201	0.560
Rat-07-mMAW		28.754	0.546	4.416	0.124	28.715	0.567	4.222	0.134
Rat-08-mMAW		59.544	0.784	8.899	0.088	59.680	0.879	8.549	0.103
Rat-09-mMAW		21.189	0.611	7.937	0.077	21.194	0.624	7.741	0.081
Rat-10-mMAW		17.272	0.304	1.832	0.166	17.262	0.313	1.762	0.178
Rat-11-mMAW		70.203	1.099	2.854	0.385	69.962	1.095	2.743	0.399
Rat-12-mMAW		26.119	2.836	6.094	0.465	26.203	2.785	5.853	0.476
Rat-13-mMAW		26.152	1.886	3.306	0.570	26.193	1.805	3.267	0.552
Rat-14-mMAW		17.982	2.737	4.090	0.669	17.937	2.683	3.905	0.687
Rat-15-mMAW		18.233	0.413	3.341	0.124	18.167	0.428	3.275	0.131
Rat-16-mMAW		47.532	0.392	1.742	0.225	47.415	0.376	1.696	0.222

Table A.13: *Frequency domain parameters after correction using mMAW.*

Correction technique:		<i>Nonlinear Predictive Interpolation</i>							
Percentage of losses:		<i>2.5%</i>				<i>5.0%</i>			
Subject \ Parameter		<i>VLF (ms²)</i>	<i>LF (ms²)</i>	<i>HF (ms²)</i>	<i>LF/HF (n.u)</i>	<i>VLF (ms²)</i>	<i>LF (ms²)</i>	<i>HF (ms²)</i>	<i>LF/HF (n.u)</i>
Rat-01-NPI		92.190	3.118	8.103	0.385	92.297	3.173	8.067	0.393
Rat-02-NPI		155.002	1.605	2.423	0.662	155.501	1.558	2.408	0.647
Rat-03-NPI		41.686	1.266	4.698	0.269	41.512	1.242	4.674	0.266
Rat-04-NPI		201.196	2.332	7.042	0.331	201.597	2.363	6.967	0.339
Rat-05-NPI		107.117	1.635	2.312	0.707	107.014	1.582	2.332	0.679
Rat-06-NPI		209.829	1.290	2.319	0.556	209.881	1.290	2.299	0.561
Rat-07-NPI		28.720	0.570	4.509	0.126	28.662	0.593	4.413	0.134
Rat-08-NPI		59.521	0.794	9.105	0.087	59.491	0.878	9.043	0.097
Rat-09-NPI		21.231	0.622	8.134	4.884	21.073	0.648	8.168	0.079
Rat-10-NPI		17.266	0.315	1.867	0.169	17.245	0.327	1.843	0.177
Rat-11-NPI		70.298	1.119	2.909	0.385	70.251	1.139	2.883	0.395
Rat-12-NPI		26.136	2.908	6.209	0.468	26.165	2.931	6.111	0.480
Rat-13-NPI		26.152	1.901	3.363	0.565	26.201	1.897	3.374	0.562
Rat-14-NPI		17.985	2.804	4.137	0.678	17.848	2.695	4.083	0.660
Rat-15-NPI		18.222	0.427	3.412	0.125	18.027	0.468	3.422	0.137
Rat-16-NPI		47.486	0.396	1.782	0.222	47.454	0.394	1.776	0.222

Table A.14: *Frequency domain parameters after correction using NPI.*

A.3 Nonlinear domain results

In this section it is presented data tables containing the values for the “control” group on table (A.15), and the data after every correction process organized as: *deletion* on table (A.16), *linear interpolation* on table (A.17), *cubic interpolation* on table (A.4), *moving average window* on table (A.19), *modified moving average window* on table (A.20) and *nonlinear predictive interpolation* on table (A.21).

Subject \ Parameter	$\alpha_1 (a.u)$	$\alpha_2 (a.u)$	$MSE_1 (a.u)$	$MSE_T (a.u)$	$\theta V (a.u)$	$2UV(a.u)$
Rat-01	0.761	1.082	1.375	19.978	21.107	44.291
Rat-02	1.061	1.168	0.669	16.162	39.007	26.596
Rat-03	0.727	1.091	1.339	22.270	14.189	53.716
Rat-04	0.798	1.106	1.060	19.794	37.755	24.490
Rat-05	1.044	1.258	0.873	17.168	39.310	19.310
Rat-06	1.013	1.328	0.608	14.164	40.604	20.134
Rat-07	0.445	1.188	1.490	22.043	10.959	57.192
Rat-08	0.351	1.202	1.553	15.873	8.997	48.789
Rat-09	0.321	1.147	1.906	19.890	10.204	45.578
Rat-10	0.544	1.026	1.661	19.886	17.958	30.282
Rat-11	0.726	1.161	1.269	17.444	24.555	26.690
Rat-12	0.785	1.048	2.155	24.306	29.861	27.083
Rat-13	0.904	1.041	1.928	26.625	34.812	19.454
Rat-14	0.943	0.928	2.197	29.831	27.815	32.119
Rat-15	0.555	1.038	1.794	20.555	6.485	59.727
Rat-16	0.642	1.192	1.008	16.006	24.014	23.656

Table A.15: Frequency parameters for control group.

Correction technique:		Deletion											
Percentage of losses:		2.5%						5.0%					
Subject	Parameter	$\alpha_1(a.u)$	$\alpha_2(a.u)$	$MSE_{E_1}(a.u)$	$MSE_T(a.u)$	$OV(a.u)$	$2UV(a.u)$	$\alpha_1(a.u)$	$\alpha_2(a.u)$	$MSE_{E_1}(a.u)$	$MSE_T(a.u)$	$OV(a.u)$	$2UV(a.u)$
Rat-01-DEL		0.817	1.074	1.397	20.241	24.915	39.932	0.829	1.075	1.410	20.210	25.888	39.425
Rat-02-DEL		1.040	1.165	0.678	16.318	40.339	25.424	1.055	1.168	0.685	16.627	38.503	25.668
Rat-03-DEL		0.780	1.089	1.362	22.423	15.541	53.041	0.755	1.094	1.381	22.347	14.757	51.389
Rat-04-DEL		0.809	1.100	1.059	20.128	37.884	24.232	0.834	1.094	1.058	19.904	39.693	22.487
Rat-05-DEL		1.023	1.253	0.874	17.348	40.678	18.983	0.984	1.253	0.879	17.513	39.726	19.349
Rat-06-DEL		0.993	1.329	0.610	14.339	37.884	21.160	1.027	1.323	0.609	14.467	42.007	19.218
Rat-07-DEL		0.483	1.193	1.509	22.566	10.848	55.932	0.485	1.191	1.530	22.452	11.765	54.152
Rat-08-DEL		0.375	1.197	1.564	16.087	9.622	48.454	0.381	1.197	1.562	16.574	10.866	47.538
Rat-09-DEL		0.337	1.141	1.902	20.214	10.170	45.424	0.360	1.130	1.910	20.707	9.310	47.069
Rat-10-DEL		0.590	1.021	1.660	20.011	20.206	30.822	0.569	1.015	1.657	20.429	21.993	29.553
Rat-11-DEL		0.737	1.157	1.266	17.589	24.555	26.335	0.747	1.150	1.267	17.604	26.148	27.562
Rat-12-DEL		0.753	1.051	2.159	24.719	29.932	26.871	0.753	1.053	2.160	24.647	31.100	26.289
Rat-13-DEL		0.902	1.030	1.934	26.737	34.915	20.678	0.874	1.024	1.934	26.805	34.828	20.345
Rat-14-DEL		0.912	0.931	2.201	29.997	26.370	32.192	0.894	0.924	2.200	30.194	28.328	30.375
Rat-15-DEL		0.573	1.033	1.823	20.884	7.167	58.703	0.615	1.029	1.841	21.138	7.107	58.714
Rat-16-DEL		0.673	1.199	1.009	16.257	24.739	24.042	0.675	1.193	1.011	16.485	28.671	22.727

Table A.16: Nonlinear parameters after correction using deletion.

Correction technique:		Linear Interpolation											
Percentage of losses:		2.5%						5.0%					
Subject	Parameter	$\alpha_1(a.u)$	$\alpha_2(a.u)$	$MSE_{E_1}(a.u)$	$MSE_T(a.u)$	$OV(a.u)$	$2UV(a.u)$	$\alpha_1(a.u)$	$\alpha_2(a.u)$	$MSE_{E_1}(a.u)$	$MSE_T(a.u)$	$OV(a.u)$	$2UV(a.u)$
	Rat-01-LI	0.781	1.081	1.390	20.255	21.799	41.177	0.798	1.080	1.386	20.403	24.222	39.792
	Rat-02-LI	1.076	1.168	0.666	16.163	40.894	25.086	1.080	1.168	0.657	16.199	41.667	22.917
	Rat-03-LI	0.747	1.091	1.373	22.639	15.101	52.013	0.768	1.090	1.393	22.854	16.667	48.640
	Rat-04-LI	0.807	1.107	1.045	19.992	39.041	23.288	0.834	1.105	1.025	19.734	40.339	21.695
	Rat-05-LI	1.053	1.258	0.855	17.202	40.273	18.089	1.070	1.259	0.843	17.446	41.497	18.027
	Rat-06-LI	1.022	1.328	0.600	14.213	42.034	18.983	1.033	1.328	0.584	14.153	41.837	18.708
	Rat-07-LI	0.486	1.184	1.499	22.566	12.069	53.793	0.509	1.182	1.500	22.690	14.384	50.685
	Rat-08-LI	0.389	1.199	1.575	16.327	10.345	46.552	0.417	1.197	1.572	17.188	11.905	45.238
	Rat-09-LI	0.350	1.143	1.888	20.445	10.922	44.369	0.384	1.138	1.879	21.369	11.225	44.558
	Rat-10-LI	0.560	1.026	1.654	20.121	21.035	28.966	0.588	1.025	1.651	20.371	21.951	28.571
	Rat-11-LI	0.749	1.160	1.251	17.523	25.899	25.899	0.766	1.160	1.233	17.463	27.500	25.357
	Rat-12-LI	0.794	1.046	2.055	24.464	30.877	26.316	0.813	1.048	1.982	24.416	31.724	24.828
	Rat-13-LI	0.918	1.041	1.834	26.578	35.274	19.521	0.916	1.041	1.763	26.326	35.836	18.771
	Rat-14-LI	0.960	0.928	2.151	29.926	29.195	30.537	0.965	0.928	2.087	30.018	30.612	28.231
	Rat-15-LI	0.582	1.038	1.843	21.163	8.247	56.701	0.618	1.035	1.863	21.537	10.000	54.828
	Rat-16-LI	0.665	1.191	1.011	16.094	25.540	22.302	0.685	1.191	1.009	16.344	28.269	21.201

Table A.17: Nonlinear parameters after correction using LI.

Correction technique:		Cubic Interpolation											
Percentage of losses:		2.5%					5.0%						
Subject \ Parameter		$\alpha_1(a.u)$	$\alpha_2(a.u)$	$MSE_{E_1}(a.u)$	$MSE_T(a.u)$	$OV(a.u)$	$2UV(a.u)$	$\alpha_1(a.u)$	$\alpha_2(a.u)$	$MSE_{E_1}(a.u)$	$MSE_T(a.u)$	$OV(a.u)$	$2UV(a.u)$
Rat-01-CI		0.783	1.081	1.392	20.303	21.528	41.319	0.801	1.080	1.391	20.490	24.567	39.100
Rat-02-CI		1.077	1.168	0.667	16.165	41.035	25.172	1.078	1.168	0.660	16.244	41.869	22.837
Rat-03-CI		0.747	1.091	1.375	22.677	15.101	51.678	0.770	1.090	1.397	22.897	16.724	48.464
Rat-04-CI		0.807	1.107	1.046	19.988	39.175	23.024	0.834	1.105	1.028	19.766	40.203	21.960
Rat-05-CI		1.054	1.258	0.856	17.209	40.411	18.151	1.070	1.258	0.845	17.496	41.497	17.687
Rat-06-CI		1.023	1.328	0.601	14.229	42.034	18.983	1.033	1.327	0.586	14.161	41.695	18.983
Rat-07-CI		0.489	1.183	1.501	22.599	11.945	53.925	0.513	1.181	1.506	22.745	14.384	51.027
Rat-08-CI		0.391	1.199	1.576	16.344	10.000	46.552	0.420	1.197	1.574	17.259	11.864	45.085
Rat-09-CI		0.352	1.142	1.888	20.463	10.884	43.878	0.388	1.137	1.886	21.477	11.263	44.369
Rat-10-CI		0.561	1.026	1.654	20.131	20.962	28.866	0.593	1.024	1.652	20.400	21.799	28.374
Rat-11-CI		0.750	1.159	1.252	17.544	25.632	25.993	0.769	1.160	1.235	17.491	27.500	25.357
Rat-12-CI		0.794	1.046	2.060	24.481	31.959	25.773	0.813	1.048	1.981	24.413	33.560	24.407
Rat-13-CI		0.918	1.041	1.840	26.594	35.593	19.322	0.917	1.041	1.778	26.321	36.610	18.305
Rat-14-CI		0.963	0.927	2.150	29.941	27.304	31.058	0.965	0.927	2.092	30.006	30.640	27.946
Rat-15-CI		0.583	1.037	1.845	21.209	8.532	55.973	0.621	1.034	1.862	21.633	9.797	53.041
Rat-16-CI		0.666	1.191	1.011	16.096	26.481	22.300	0.688	1.190	1.011	16.385	27.562	21.555

Table A.18: Nonlinear parameters after correction using CI.

Correction technique:		Moving Average Window											
Percentage of losses:		2.5%						5.0%					
Subject \ Parameter		$\alpha_1(a.u)$	$\alpha_2(a.u)$	$MSE_{E_1}(a.u)$	$MSE_T(a.u)$	$OV(a.u)$	$2UV(a.u)$	$\alpha_1(a.u)$	$\alpha_2(a.u)$	$MSE_{E_1}(a.u)$	$MSE_T(a.u)$	$OV(a.u)$	$2UV(a.u)$
Rat-01-MAW		2.003	1.150	0.174	14.958	83.729	2.034	1.987	1.149	0.176	14.882	82.432	2.703
Rat-02-MAW		2.052	1.229	0.110	12.945	85.859	0.673	2.035	1.229	0.110	12.966	84.849	1.010
Rat-03-MAW		2.016	1.164	0.195	17.775	80.000	2.667	1.997	1.163	0.198	17.780	80.405	2.365
Rat-04-MAW		2.036	1.163	0.147	16.125	87.333	0.667	2.015	1.160	0.148	15.949	87.879	0.673
Rat-05-MAW		2.013	1.325	0.111	13.403	85.284	0.669	2.010	1.325	0.113	13.463	84.564	0.671
Rat-06-MAW		2.049	1.385	0.087	11.529	87.542	0.337	2.055	1.384	0.087	11.531	86.622	0.669
Rat-07-MAW		1.814	1.280	0.176	17.612	76.768	4.714	1.820	1.280	0.178	17.392	76.610	5.085
Rat-08-MAW		1.890	1.278	0.170	11.498	82.886	2.349	1.874	1.278	0.173	11.840	82.550	2.685
Rat-09-MAW		1.627	1.259	0.251	14.871	74.150	6.122	1.628	1.257	0.257	15.423	73.490	6.376
Rat-10-MAW		1.992	1.085	0.167	15.137	85.571	1.678	1.952	1.084	0.169	15.273	85.522	1.347
Rat-11-MAW		2.020	1.225	0.128	13.105	85.185	1.347	2.012	1.225	0.131	13.138	85.667	1.000
Rat-12-MAW		1.769	1.172	0.240	18.511	76.014	2.365	1.771	1.173	0.237	18.522	75.421	2.357
Rat-13-MAW		1.905	1.138	0.240	20.548	82.770	1.351	1.896	1.137	0.236	20.467	83.108	1.351
Rat-14-MAW		1.979	1.019	0.387	24.999	78.716	1.689	1.964	1.019	0.382	24.855	78.788	1.684
Rat-15-MAW		1.994	1.096	0.190	16.451	81.879	4.698	1.982	1.093	0.197	16.282	82.095	4.054
Rat-16-MAW		1.990	1.240	0.123	13.155	86.869	1.010	2.009	1.242	0.124	13.260	86.242	1.007

Table A.19: Nonlinear parameters after correction using MAW.

Correction technique:		modified Moving Average Window											
Percentage of losses:		2.5%						5.0%					
Subject	Parameter	$\alpha_1(a.u)$	$\alpha_2(a.u)$	$MSE_{E_1}(a.u)$	$MSE_T(a.u)$	$OV(a.u)$	$2UV(a.u)$	$\alpha_1(a.u)$	$\alpha_2(a.u)$	$MSE_{E_1}(a.u)$	$MSE_T(a.u)$	$OV(a.u)$	$2UV(a.u)$
Rat-01-mMAW		0.765	1.077	1.368	20.010	22.145	42.907	0.781	1.077	1.360	20.014	22.569	42.014
Rat-02-mMAW		1.059	1.165	0.662	16.113	40.569	25.267	1.056	1.167	0.651	16.080	40.210	25.175
Rat-03-mMAW		0.735	1.088	1.361	22.303	15.254	51.864	0.739	1.090	1.368	22.381	17.057	51.171
Rat-04-mMAW		0.796	1.102	1.045	19.664	38.908	23.549	0.808	1.100	1.027	19.709	39.661	23.051
Rat-05-mMAW		1.043	1.257	0.854	17.078	40.339	18.983	1.047	1.259	0.844	17.052	40.678	19.661
Rat-06-mMAW		1.014	1.327	0.597	14.096	41.216	19.932	1.003	1.329	0.583	14.063	41.751	19.865
Rat-07-mMAW		0.460	1.190	1.481	22.171	11.724	55.862	0.465	1.192	1.472	22.251	12.371	54.983
Rat-08-mMAW		0.368	1.201	1.563	16.118	9.966	47.766	0.378	1.202	1.565	15.988	10.764	47.569
Rat-09-mMAW		0.332	0.332	1.889	20.057	10.690	45.862	0.351	1.147	1.874	20.640	11.684	44.674
Rat-10-mMAW		0.550	1.019	1.660	19.878	19.649	29.123	0.564	1.019	1.652	20.074	22.069	28.621
Rat-11-mMAW		0.735	1.157	1.256	17.476	25.623	26.690	0.736	1.158	1.237	17.341	27.931	26.207
Rat-12-mMAW		0.779	1.049	2.066	24.305	31.081	26.689	0.782	1.050	1.983	24.206	30.903	26.389
Rat-13-mMAW		0.902	1.043	1.849	26.477	35.374	20.748	0.894	1.045	1.743	26.220	37.201	19.795
Rat-14-mMAW		0.937	0.930	2.130	29.329	27.852	31.544	0.936	0.930	2.083	29.322	28.378	30.743
Rat-15-mMAW		0.571	1.038	1.822	20.751	7.509	58.703	0.575	1.039	1.830	20.951	9.343	57.093
Rat-16-mMAW		0.653	1.192	1.011	15.984	26.829	23.345	0.655	1.193	1.018	15.974	25.965	22.807

Table A.20: Nonlinear parameters after correction using modified MAW algorithm.

Correction technique:		Nonlinear Predictive Interpolation											
Percentage of losses:		2.5%						5.0%					
Subject \ Parameter		$\alpha_1(a.u)$	$\alpha_2(a.u)$	$MSE_{E_1}(a.u)$	$MSE_T(a.u)$	$OV(a.u)$	$2UV(a.u)$	$\alpha_1(a.u)$	$\alpha_2(a.u)$	$MSE_{E_1}(a.u)$	$MSE_T(a.u)$	$OV(a.u)$	$2UV(a.u)$
Rat-01-NPI		0.767	1.082	1.379	20.201	19.929	43.772	0.768	1.082	1.376	20.096	22.373	42.373
Rat-02-NPI		1.057	1.169	0.673	16.181	39.929	25.442	1.055	1.169	0.669	16.235	39.437	25.704
Rat-03-NPI		0.729	1.092	1.346	22.407	14.237	53.898	0.733	1.093	1.364	22.446	14.384	52.740
Rat-04-NPI		0.798	1.107	1.055	19.928	37.755	24.150	0.811	1.106	1.053	19.600	38.567	23.891
Rat-05-NPI		1.044	1.258	0.870	17.154	38.328	19.164	1.049	1.259	0.866	17.124	39.041	20.548
Rat-06-NPI		1.004	1.329	0.609	14.159	40.604	20.134	1.001	1.329	0.607	14.019	39.796	20.408
Rat-07-NPI		0.460	1.188	1.494	22.231	10.922	56.655	0.461	1.188	1.493	22.285	11.824	55.405
Rat-08-NPI		0.360	1.201	1.562	15.915	9.831	48.475	0.366	1.201	1.566	16.372	9.831	48.814
Rat-09-NPI		0.076	1.145	1.900	20.121	10.274	45.548	0.348	1.142	1.903	20.430	9.556	47.099
Rat-10-NPI		0.548	1.026	1.659	20.110	18.310	30.282	0.560	1.027	1.660	20.057	19.504	29.078
Rat-11-NPI		0.732	1.161	1.268	17.427	25.357	26.786	0.741	1.161	1.264	17.461	26.481	26.481
Rat-12-NPI		0.780	1.046	2.154	24.579	29.310	27.931	0.794	1.048	2.137	24.585	32.107	26.421
Rat-13-NPI		0.903	1.041	1.922	26.558	34.589	20.548	0.894	1.041	1.924	26.340	34.014	21.088
Rat-14-NPI		0.946	0.928	2.194	29.944	27.458	31.864	0.943	0.928	2.185	30.262	28.716	31.419
Rat-15-NPI		0.567	1.038	1.818	20.805	7.167	58.362	0.584	1.038	1.834	21.176	7.534	58.904
Rat-16-NPI		0.645	1.193	1.010	16.041	24.818	22.628	0.656	1.192	1.010	16.205	26.644	22.491

Table A.21: Nonlinear parameters after correction using NPI algorithm.

SAMPLE ENTROPY (SAMPEN)

Sample entropy (*SampEn*) is a modification of approximate entropy (*ApEn*), used extensively for assessing the complexity of a physiological time-series signal, thereby diagnosing diseased state. Unlike ApEn, SampEn shows good traits such as data length independence and trouble-free implementation. In ApEn the comparison between the template vector and the rest of the vectors also includes comparison with itself, causing a bias in the complexity estimation. SampEn avoids this self-match error eliminating the bias in the computation of the signal complexity.

The procedure to calculate the SampEn could be describe as follows: Assume a time series of length N , $X = x_1, x_2, \dots, x_N$. Now, define a template vector of length m of the form $U_m(i) = x_i, x_{i+1}, x_{i+1}, \dots, x_{i+m-1}$ and a distant function $d[U_m(i), U_m(j)]$, ($i \neq j$), that could be the “*Chebyshev distance function*”. The next step implies to count the number of vector pairs in template vectors of length m and $m + 1$ having a $d[U_m(i), U_m(j)] \leq r$, and denoted by B and A respectively. We define then the sample entropy as:

$$SampEn = -\log \frac{A}{B} \quad (\text{B.1})$$

where,

- A is the number of template vector pairs having $d[U_{m+1}(i), U_{m+1}(j)] < r$ of length $m + 1$.
- B is the number of template vector pairs having $d[U_m(i), U_m(j)] < r$ of length m .

From the definition is easy to see that A will always have a value smaller or equal to B . Therefore, $SampEn(m, r, \tau)$ will be always either be zero or positive value. In the case of HRV the value of m is set to be 2 and the value of r to be $0.15 \times std$. Where std corresponds to the standard deviation of all dataset.

An example taken from [65] using a simulated series and presented in the figure B.1, illustrates the procedure for calculating sample entropy (*SampEn*) for the case in which the pattern length, m , is 2, and the similarity criterion, r , is 20. Dotted horizontal lines around data points $u[1]$, $u[2]$ and $u[3]$ represent $u[1] \pm r$, $u[2] \pm r$, and $u[3] \pm r$, respectively. Two data values match each other, that is, they are indistinguishable, if the absolute difference between them is $\leq r$. All green points represent data points that match the data point $u[1]$. Similarly, all red and blue points match the data points $u[2]$ and $u[3]$, respectively. Consider the 2-component green-red template sequence $(u[1], u[2])$ and the 3-component green-red-blue $(u[1], u[2], u[3])$ template sequence. For the segment shown, there are two green-red sequences, $(u[13], u[14])$ and $(u[43], u[44])$, that match the template sequence $(u[1], u[2])$ but only one green-red-blue sequence that matches the template sequence $(u[1], u[2], u[3])$. Therefore, in this case, the number of sequences matching the 2-component template sequences is two and the number of sequences matching the 3-component template sequence is 1. These calculations are repeated for the next 2-component and 3-component template sequence, which are, $(u[2], u[3])$ and $(u[2], u[3], u[4])$, respectively. The numbers of sequences that match each of the 2- and 3-component template sequences are again counted and added to the previous values. This procedure is then repeated for all other possible template sequences, $(u[3], u[4], u[5]), \dots, (u[N-2], u[N-1], u[N])$, to determine the ratio between the total number of 2-component template matches and the total number of 3-component template matches. **SampEn** is the natural logarithm of this ratio and reflects the probability that sequences that match each other for the first two data points will also match for the next point.

Figure B.1: *Counting of patterns in the process to calculate the SampEn on an RR time series. Taken from [65].*