Mohamed Boudiaf University - M'sila

FACULTY OF TECHNOLOGY

ELECTRONICS DEPARTMENT

Serial number …………………..…
Registration number ……….……

Thesis

Presented for the award of the degree of

DOCTORATE 3rd CYCLE LMD

Field: Electronics
Specialization: Instrumentation

THEME

Contribution to the development of tools for the treatment of time series and dynamical systems

Presented By

NABI Zahia

Defended on: 09/11/2024

In front of a jury composed of:

First & last name	**Grade**	**Institution**	**Position**
BOURAS Mounir	Professor	Univ. of M'sila	President
OUALI Mohammed Assam	MCA	Univ. of M'sila	Supervisor
LADJAL Mohamed	Professor	Univ. of M'sila	Co. Supervisor
DJERIOUI Mohamed	MCA	Univ. of M'sila	Examiner
AOUICHE Abdelaziz	MCA	Univ. of Tebessa	Examiner
MESSOUDI Noureddine	MCA	Univ. of Boumerdes	Examiner
BENNACER Hamza	Professor	Univ. of M'sila	Invited

Academic year: 2023/2024

Acknowledgements

"A sincere thank you is the most beautiful reward."

First and foremost, I express my deep gratitude and sincere thanks to my thesis advisors, Dr. ***OUALI Mohammed Assam*** *and* ***Pr. LADJAL Mohamed****, for proposing this research topic and for supervising this thesis. Working with you has been enriching, and I am grateful for your encouragement throughout my thesis.* ***Dr. OUALI Mohammed Assam****, thank you for your invaluable help, guidance, patience, understanding, kindness, and incredible availability.* ***Pr. LADJAL Mohamed****, thank you for all the assistance you have provided, your attentive listening, kindness, and availability.*

I would like to sincerely thank the jury members who have honored me by evaluating this work. Thank you to ***Pr. BOURAS Mounir*** *for accepting to be the president of the jury for this thesis, and to* ***Dr. DJERIOUI Mohamed****,* ***Dr. AOUICHE Abdelaziz****,* ***Dr. MESSOUDI Noureddine****, and* ***Pr. BENNACER Hamza*** *for honoring me by agreeing to be part of the jury.*

I extend my heartfelt thanks to ***Dr. OUADFEL Ghania*** *who continuously provided me with advice, encouragement, and both scientific and human support throughout my research.*

I would also like to thank my family members for their unwavering support throughout these years, especially my dear sister ***Nacima****, who is not only a constant source of support but also a true scientific role model. Her dedication as professor and her inspiration has guided and motivated me throughout my career.*

Finally, I wish to express my deep gratitude to all the people, near and far, who have provided their help, advice, and support throughout this work. Thank you all.

Dedication

With profound gratitude and sincere appreciation, I am honored to dedicate this present work to the dearest people in my heart:

To my father and my mother, whose unwavering support throughout my academic journey has allowed me to become the person I am today. I will be eternally grateful to them for their unconditional love, sacrifices, invaluable advice, trust and tenderness. Thanks to them, I have received an education of which I am immensely proud.

To a special person whom I do not name, but who recognizes himself

To my brothers and sisters, as well as my nieces and nephews, who have all contributed to my personal and academic growth.

To my dear friends and all those who are very dear to me, for their presence, encouragement, and sincere friendship.

Your support and love have been the pillars of my success. I dedicate this work to you with all my gratitude and affection.

Zahia

Abstract

This thesis focuses on two main areas of research. The first axis deals with the identification and modeling of dynamical systems and time series. Given this, an intelligent approach based on neural networks combined with metaheuristic algorithms has been proposed to obtain a valid and efficient model that represents and captures the characteristics of these systems and their chronological variation. One of the unique features of the proposed intelligent model is its ability to be used not only for identifying and modeling dynamical systems and time series, but also for data compression. The second axe concerns the analysis and processing of time series data, in particular data pre-processing, the main objective being the development of techniques for cleaning and preparing time series data, including the management of missing values, anomaly detection, data normalization and filtering in order to eliminate various noises and artefacts (such as biomedical or other signals). To this end, a number of methods for filtering ECG time series have been proposed. These include methods based on empirical mode decomposition and its graphical variant as well as wavelets.

Keywords: dynamical systems, time series, meta-heuristic algorithms, Artificial Neural Network, ARMA model, denoising ECG signal.

Résumé

Cette thèse présente principalement des travaux de recherche axés sur deux axes principaux. Le premier axe traite de l'identification et de la modélisation des systèmes dynamiques et des séries temporelles. Dans cette optique, une approche intelligente basée sur des réseaux neuronaux combinés à des algorithmes métaheuristiques a été proposée pour obtenir un modèle valide et efficace qui représente et capture les caractéristiques de ces systèmes et leur variation chronologique. L'une des caractéristiques uniques du modèle intelligent proposé est sa capacité à être utilisé non seulement pour identifier et modéliser des systèmes dynamiques et des séries chronologiques, mais aussi pour la compression des données. Le deuxième axe concerne l'analyse et le traitement des données de séries temporelles, notamment le prétraitement des données, l'objectif principal étant le développement des techniques pour nettoyer et préparer les données temporelles, y compris la gestion des valeurs manquantes, la détection des anomalies, la normalisation des données et le filtrage afin d'éliminer divers bruits et artefacts (tels que les signaux biomédicaux ou d'autres signaux). À cet effet, un certain nombre de méthodes de filtrage des séries temporelles d'ECG ont été proposées. Il s'agit notamment de méthodes basées sur la décomposition en modes empiriques et sa variante graphique ainsi que les ondelettes.

Les mots clés : systèmes dynamiques, séries temporelles, algorithmes méta-heuristiques, Réseau de Neurones Artificiels, le modèle ARMA, filtrage du signal ECG.

ملخص

تركز هذه الأطروحة على محورين رئيسيين للبحث. يتناول المحور الأول تحديد ونمذجة الأنظمة الديناميكية والسلاسل الزمنية. وفي ضوء ذلك، تم اقتراح نهج ذكي يعتمد على الشبكات العصبية مقترنة بخوارزميات ميتاهورستية للحصول على نموذج صالح وفعال يمثل ويلتقط خصائص هذه الأنظمة وتغيرها الزمني. تتمثل إحدى السمات الفريدة للنموذج الذكي المقترح في إمكانية استخدامه ليس فقط لتحديد ونمذجة الأنظمة الديناميكية والسلاسل الزمنية الديناميكية فحسب، بل أيضًا لضغط البيانات. ويتعلق المحور الثاني بتحليل بيانات السلاسل الزمنية ومعالجتها، ولاسيما المعالجة المسبقة للبيانات، والهدف الرئيسي هو تطوير تقنيات لتنظيف بيانات السلاسل الزمنية وإعدادها، بما في ذلك إدارة القيم المفقودة، والكشف عن الشذوذ، وتطبيع البيانات وتصفيتها من أجل القضاء على مختلف الضوضاء والتحف مثل الإشارات الطبية الحيوية أو غيرها من الإشارات. وتحقيقًا لهذه الغاية، تم اقتراح عدد من الطرق لتصفية السلاسل الزمنية لتخطيط كهربية القلب. وتشمل هذه الأساليب طرقًا تعتمد على تحلل النمط التجريبي ومتغيره البياني بالإضافة إلى الموجات.

الكلمات المفتاحية: الأنظمة الديناميكية، السلاسل الزمنية، الخوارزميات الميتاهورستية، الشبكة العصبية الاصطناعية، نموذج ARMA، إزالة الضوضاء من إشارة تخطيط القلب.

List of tables

Table of figures

List of abbreviations

AR: Autoregressive model.

MA: Moving average model.

ARMA: Autoregressive moving average model.

ARIMA: Auto regressive integrated moving average.

SARIMA: Seasonal auto regressive integrated moving average.

SARIMAX: Seasonal auto regressive integrated moving average with eXogenous variables.

ARCH: Auto regressive conditional heteroskedasticity.

GARCH: Generalized auto regressive conditional heteroskedasticity.

ETS: Exponential smoothing model.

VAR: Vector auto regressive.

ECG: Electrocardiograms.

EEG: Electroencephalograms.

SA: Sinus node.

AV: Atrioventricular node.

EMG: Electromyographic.

AI: Artificial intelligence.

ANN: Artificial neural network.

RNN: Recurrent neural networks.

ANNARMA: Artificial neural network autoregressive moving average.

EMD: Empirical mode decomposition.

GEMD: Graph empirical mode decomposition.

IWO: Invasive weed optimization.

PSO: Particle swarm optimization.

ICA: Imperialist competitive.

CMA-E: Covariance matrix adaptation evolution strategy.

ES: Evolution strategy.

EA: Evolutionary algorithm.

DWT: Discrete waveform transform.

FT: Fourier transform.

WT: Wavelets.

IMF: Intrinsic mode functions.

MSE: Main Square Error.

MAE: Mean absolute error.

SNR: Signal-to-noise ratio SNR.

CONTENTS

Chapter II: Applied methods

Chapter III: Modeling and identification of dynamical system and time series

Chapter IV: Preprocessing of ECG signals

INTRODUCTION

The modeling and identification of dynamical systems, along with time series analysis and processing, are of vital importance in a variety of scientific and practical areas. On one side, modeling and identification allow for the capturing of the intricate behavior of real systems, whether biological, mechanical, electrical, or otherwise. This capacity to represent and predict how systems will behave is crucial for designing, controlling, commanding, and optimizing various fields, spanning from engineering to medicine. On the other side, time series analysis and processing (modeling, prediction, filtering and classification) is useful because many examples of real data have a time series dimension, such as biomedical signals (ECG, EEG, EMG, etc.), drinking water analysis data, meteorological records and computer network monitoring data, etc., enabling the discovery of hidden patterns, noise reduction and the extraction of valuable information from complex time series data.

One notable challenge in modeling dynamical systems is the formulation of mathematical models for these systems, which is often hindered by the lack of complete knowledge regarding system parameters, in addition to other constraints namely the complexity of dynamic behaviors such as oscillations and chaos, and the non-linearity of relationships between input and output variables [1]. Similarly, the processing of time series data is often hampered by irregular data acquisition and the constant need for expert intervention. Thus, the interpretation of results and decision-making can be complicated by the presence of noisy, missing, or mislabeled data [2]. This complexity hampers the understanding and optimization of systems, limiting their efficiency and control, and reducing the ability to detect significant patterns and make informed decisions, especially in the case of time series. To overcome these obstacles, it is crucial to develop intelligent statistical learning models for dynamical systems, enabling accurate modeling despite their complexity. At the same time, innovative methodologies need to be developed to automate the processing of time series, incorporating machine learning techniques for more efficient data analysis, thereby reducing the need for recurrent human intervention.

The work outlined in this doctoral thesis is divided into two main parts.

I) Modeling and Identification of dynamical systems and time series.

II) Analysis and processing of time series data.

Modeling and Identification of dynamical systems and time series

Modeling entails constructing an abstract representation of a real system. It focuses on the mechanism of how the components in this system evolve across time. This step is fundamental to understanding and predicting the behavior of various physical and industrial processes [3]. In the literature, two main types of modeling can be encountered: the first is mathematical modeling [4] and the second is computational and intelligent modeling [5]. Mathematical modeling consists of representing a real system using precise mathematical equations, such as algebraic or differential/difference equations. These equations describe the relationships between the different variables in the system and enable its behavior under different conditions to be predicted. Mathematical modeling is based on sound mathematical principles and physical laws. Whereas intelligent modeling aims to represent a real system using techniques that are not strictly mathematical but rather than based on artificial intelligence, which are the most widely adopted by researchers in recent decades [6]. These methods include the principle of Neural Networks [7], Radial Basis Function networks [8, 9], Fuzzy Logic [10, 11], Neuro-Fuzzy Systems [12], Machine learning [13] and Deep learning methods [14]. These techniques allow complex and often non-linear relationships between system variables to be captured, using linguistic rules or models inspired by the workings of the human brain. Intelligent modeling is often used when the relationships between variables are ill-defined, imprecise or difficult to represent mathematically. These models aim to be as close as possible to the real system they represent. This involves adjusting the values of the model's parameters based on available data through the parameter identification process.

Analysis and processing of time series data

Time series represent data gathered successively throughout time, and their processing and analysis holds significant importance across various fields, including medicine, finance, meteorology, and others. The electrocardiogram (ECG) is a non-stationary, non-linear, quasi-periodic time series. It is a widely utilized tool for the examination of the muscular and electrical functions of the heart. It is a time-varying biosignal that reflects the flow of ionic current, which causes the cardiac fibers to contract and then relax, giving an indirect indication of blood flow to the heart muscle. It offers information on electrical activity, rhythm, and heart rate. The ECG provides both physiological and pathological information, which is essential to the diagnosis of cardiac disease [15]. Various kinds of noise such as electrosurgical noise, power line interference, motion artifacts, baseline drift, muscle contraction, baseline drift, ECG signal amplitude variation with respiration, and equipment noise may have an impact on the ECG recording phase [16]. These noises can alter the ECG's morphological characteristics and interval elements, resulting in incorrect patient diagnosis and treatment. ECG signal filtering is a fundamental discipline in time series analysis and processing due to the importance of the latter in the medical field. Its intent is to minimize and eliminate undesired noise and interference from ECG signal so as to retrieve and maintain the diagnostic data that may be utilized for the enhancement of the patient's quality of life.

In this thesis, we set out to accomplish dual objectives: firstly, our initial aim is to develop a robust and efficient modeling framework to model dynamical systems and time series by exploring neural networks and metaheuristic algorithms as a complementary intelligent modeling approach. Neural networks are capable of learning from complex, non-linear data, making them suitable for modeling and predicting dynamical systems and time series. Metaheuristic algorithms can enhance neural network models to accurately solve large and complex problems. By merging these two concepts, we can leverage the benefits of each in a synergistic way. Secondly, we investigated temporal preprocessing techniques for ECG signals in order to enhance their quality and precisely reduce noise while maintaining important data. The requirement to include the dynamic fluctuations over time observed in ECG signals, which are often not successfully captured by standard or frequency-based pretreatment techniques, influenced our choice of methods. A comparative study will be carried out to evaluate the effectiveness of each proposed filtering strategies.

To do this, this thesis is grouped into four chapters:

- ❖ The first chapter introduces the concepts of dynamical systems and time series, as well as heart's concept and its functioning, the origin and nature of ECG signals, and different types of noise that can affect this signal. In addition, it presents a state-of-the-art review of techniques for modeling and identifying dynamical systems and time series, data compression, as well as a review of ECG signal filtering techniques.
- ❖ The second chapter covers the theoretical foundations of neural networks and metaheuristic algorithms such as IWO, PSO, ICA and CEMA-ES. It also explores

different ECG signal decomposition and filtering techniques, including wavelets, EMD and GEMD.

- The third chapter presents the suggested methodology for modeling and identifying dynamical systems and time series using neural networks and metaheuristic algorithms.
- The fourth chapter focuses mainly on methods for denoising ECG signals, highlighting EMD, wavelets and GEMD. At the end of this chapter a comparative study will be carried out to evaluate the performance of each of the proposed filtering methods.

CHAPTER I

BASIC CONCEPTS AND STATE OF THE ART

This chapter is intended to be a starting point in comprehending the heart of this thesis. It covers in the first part the fundamental ideas of dynamical systems and time series, together with the key concepts of modeling and identification. It also discusses the analysis and processing of time series data, in particular filtering and choosing electrocardiograms (ECGs) which are absolutely essential biological signals classified as time series data as an illustrative example. A detailed explanation of the anatomy of the heart is provided to better understand the generation of these signals. The second part presents the state of the art in modeling and identification techniques for dynamical systems and time series, as well as filtering methods specifically applied to ECGs.

Contents

I.1. Introduction

Modeling and identification of dynamical systems and time series, as well as the pre-processing and analysis of temporal data, are crucial areas in many scientific and industrial fields. These disciplines enable us to understand, interpret, predict, make informed decisions and control complex evolving systems, such as biological systems, industrial processes, financial markets and many others. This first chapter will cover three important components:

- ✓ Dynamical systems concepts, modeling and identification.
- ✓ Notions of time series and the objectives of their analysis and processing.
- ✓ Basic understanding of the ECG signal, including various forms of noise and variability encountered in ECG recordings.

I.2. Part I: Basic concepts

I.2.1. Dynamical systems

I.2.1.1. Description

In the context of several fields, a system may be defined as a collection or grouping of components or elements that come into contact with each other according to particular norms and rules in order to generate a wished outcome and accomplish a specific intention. A dynamical system, on the other hand, can be characterized as the description of the dynamic behavior of various systems [17]. Dynamic behavior is explained by the idea of state variables and is associated with the temporal development of a system, in other words, how a system changes and evolves through time. State variables indicate the system's current state, attributes, or condition at any particular moment. By observing changes in these state variables, we can comprehend know how the system responds to external disturbances, how its components interact, and how complex patterns and behaviors arise. We can understand the underlying mechanisms, forecast future system states, and optimize system performance in a variety of sectors of study by analyzing dynamic behavior [17, 18].

I.2.1.2. Mathematical representation of dynamical systems

Dynamical systems are often described mathematically using differential equations or difference equations [19], and this might vary depending on the nature and complexity of the dynamical system as well as the aims of the research. These equations define the link between the state variables of the system and their evolution through time.

Differential equations are used in the case of continuous systems [20]. They describe how the rates of change of state variables are affected by their existing values and other variables. Differential equations can be linear or nonlinear, and they can explain continuous changes in the system:

$$\dot{x} = \frac{dx}{dt} = f(x, t, u, \beta) + d \tag{I.1}$$

Where f represents a set of functions $\{f_1, f_2, f_3 \ldots\ldots f_n\}$ to describe the dynamic of $\{\text{state}_1, \text{sate}_2 \ldots\ldots \text{state}_n\}$ of the represented system. More precisely we can say that it is a vector value function that for a given current state x and tells as how that state changes the next time and step. While x is the vector that indicates the state of the system or more precisely the vector that has the minimum number of values that describe the system and t represents time. Time is important because in lot of dynamical systems change with time. u on the other hand, indicates the variables that we can manipulate to try to change the behaviors of the systems (Command vector). β denotes the parameters of the systems and finally, d represents the system's disturbance or noise.

In dealing with discrete systems, difference equations are used [20]. These explain the manner in which the state variable values move from one time step to the next, and they are supplied by:

$$x_{t+1} = f\left(x_t, x_{t-1}, \ldots., x_{t-n}\right) \tag{I.2}$$

Where f represents a mathematical relation that expresses how the system evolves from its current and/or previous states and x_t is the state variable at time t. While xt_{+1} is the value of the state variable at time $t+1$ (i.e., the next state of the system) and lastly x_{t-1} to x_{t+n} are the previous values of the state variable, up to a certain order n (this order may vary depending on the system).

I.2.1.3. Notions of modeling and identification

Modeling and identification are mainly utilized to gain a deeper comprehension of complex systems and how they behave. By merging these couple approaches, we generate representational and accurate models that increase our understanding of the current world, boost structure efficiency, aid decision-making processes and control in extensive domains.

➢ Modeling

Modeling is an important process in several areas. It involves the simplification and meaningful representation of complex real-world occurrences. These representations, also referred to as models, allow to better understand, evaluate and predict the behavior of these phenomena [21].

➢ Model

A model is generally described as a design or planning that is generated for a specific intent, based on physical laws, mathematical equations, conceptual representations, or computer simulations. It is justified scientifically, mathematically, or rationally to ensure its validity and accuracy. A model is defined and differentiated by its various characteristics and attributes. These characteristics may relate to the model's structure, components, behavior, functionality, performance, or other specific elements, depending on the model's field of application [22].

➢ Types of models

Models can be classified based on their nature and purpose. Establishing a model typically requires identifying variables, parameters, causal or functional connections, as well as evaluating and validating the model with empirical information or outcomes from experiments. These classifications aid in comprehending the various models. Here are some of the common models [23]:

- ✓ **Mathematical models:** these models involve mathematical equations to illustrate naturally occurring phenomena and processes.
- ✓ **Statistical models:** employ statistical techniques to analyze data and make predictions.
- ✓ **Computer models:** computers are used to carry out these models in order to resolve complicated issues.
- ✓ **Intelligent models:** these are computational models, combine artificial intelligence and machine learning techniques.
- ✓ **Conceptual models:** the intention of employing them is to illustrate abstract ideas and conceptual links.
- ✓ **Physical models:** consists of physical representations of objects or systems, such as prototypes.
- ✓ **Experimental models:** their aim is to test theories in a controlled environment.

- **Identification**

Once the model has been established, the model's parameters must be estimated. Parameters are the unknown values that determine the model's behavior or properties; this process is termed identification. Identification is an important aspect of the procedure of modeling and entails estimating the unknown parameters of a model through observable data from the real system. This process attempts to determine the model parameter values that best match the actual system behavior. Statistical methodologies, optimization algorithms, and system identification methods are employed to perform this estimation. By estimating the parameters, we may develop and refine the model, making it more precise and representative of the real system [**24**].

I.2.1.4. Utilities and Advantage of models of dynamical systems

- ✓ Modeling permits us to evaluate the system without the requirement for physical construction. Analysis of the physical system may not always be possible due to a range of factors, including cost considerations.
- ✓ Dynamical system models incorporate essential elements to understand complex world's dynamics, allowing us to predict future outcomes and understand the interactions of various components within these systems, leading to a more optimistic and effective future.
- ✓ Models assist us in making informed decisions, design control strategies and developing innovative solutions by simulating real-world scenarios.

I.2.2. Time series

I.2.2.1. Definition

A time series (or chronic series) represents a set of observations of a variable recorded at successive intervals or point of time that provides important information about the observed phenomenon over time. In another way it can be said that it reflects the output obtained from monitoring and tracking specific occurrences or processes. These observations are analyzed to understand the past, in order to predict the future and assist in decision-making in different areas [**25**, **26**].

Mathematically it is defined by the functional relationship:

$$y_t = f\left(t\right) \qquad \text{(I.3)}$$

where y_t is the time series value at time t and t can be depending on the case (minute, hour, day, week, year, etc.).

The time series is frequently depicted graphically in various ways constructed as follows: the horizontal axis represents time and the vertical axis reflects the values of the variable investigated [27]. These typical representations are thought to be a basic step in comprehending a time series since they graphically show the behavior, trends, and fluctuations of the data across time and can create the groundwork for creating a trustworthy model thus helping to analyze the behavior of the phenomenon under study, understand its previous history, and predict its future behavior.

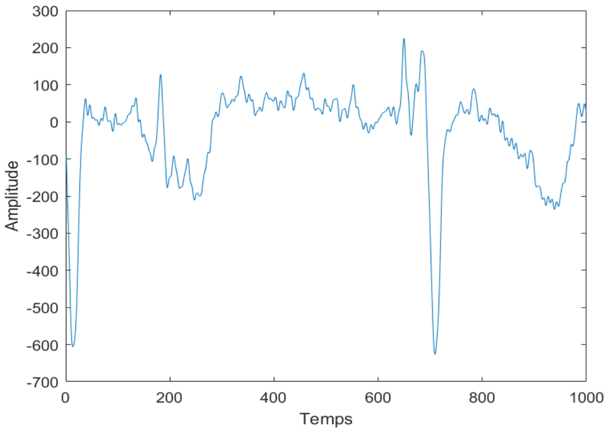

Figure I.1 – Example of graphical representation of time series.

I.2.2.2. Characteristics of time series

Mean, variance, autocorrelation, stationarity and non-stationarity are specific statistical characteristics associated with time series analysis. These statistical characteristics are crucial in exploratory time series analysis and are often used to make decisions about model selection and the transformations needed to make the data more suitable for modeling [28].

- **Mean**

The mean of a time series, being a measure at the first moment with respect to the origin, reflects the central position of values, thus providing an indication of the overall trend of the series over a specified period. It is also defined as the mathematical expectation of the variable being considered [29]. The general formula to calculate the mean of a time series is given as follows:

$$\overline{x} = \frac{\sum_{t=1}^{n} x_t}{n} \tag{I.4}$$

where x_t represents the value of the time series at time t.

- **Variance**

Variance measures the dispersion or variability of values in a time series. It indicates the extent to which individual observations deviate from the mean. A low variance suggests that the values are relatively close to the mean, while a high variance indicates wider dispersion.

When we compute variance, we estimating the dispersion of values in a data set. Two distinct scenarios are considered based on what we know about the population from which this data comes. If we have all the data of the entire population, we use what is called "theoretical variance"[29]. The theoretical variance formula is given as follows:

$$\mathrm{var}(x) = \frac{\sum_{t=1}^{n} (x_t - \bar{x})^2}{N} \tag{I.5}$$

where N is the size of the population, x_t et $\bar{x}$ are the individual values and the mean of population, respectively.

Otherwise, if we only have a representative sample of the population, we use the "empirical variance" to estimate the dispersion of the entire population from the sample [29]. The empirical variance formula is slightly different for the fact that we consider only a sample and is given as follows:

$$\mathrm{var}(x) = \frac{\sum_{t=1}^{n} (x_t - \bar{x})^2}{n-1} \tag{I.6}$$

Here, the division by $n-1$ instead of N, this corresponds to the use of a sample rather than the entire population.

In the case of calculating the variance of a time series, the "empirical variance" is generally used because it is rare to have all past and future values of the time series. In the statistical or economic literature, it is preferable to present the standard deviation rather than the variance because it offers a measure of dispersion expressed in the same units as the original observations, thus facilitating its interpretation and comparison. The standard deviation is given by the following formula [29]:

$$\sigma_x = \sqrt{\mathrm{var}(x)} \tag{I.7}$$

➢ Auto-covariance

Covariance is a statistical measure that assesses how two variables evolve together. It indicates the tendency of these variables to vary simultaneously from their respective averages. The covariance formula between two variables ***x*** and ***y*** is:

$$Cov(x,y)=\frac{1}{n}\sum_{t=1}^{n}(x_t-\bar{x})(y_t-\bar{y}) \tag{I.8}$$

Autocovariance measures the covariance between two values separated by a delay k (delay), providing insight into the variability of the series and the temporal relationships between the different components of the series. The function of autocovariance generally denoted γ, associates each pair of random variables Z to a real number [29].

$$\forall k,\ t\ \in\ Z,$$

$$\gamma_k=Cov(x_t,x_{t+k}) \tag{I.9}$$

➢ Auto-correlation

The correlation is defined as measures of the normalized covariance between two variables. A correlation coefficient close to 1 suggests a strong positive correlation, while a coefficient close to -1 indicates a strong negative correlation. A coefficient close to 0 indicates a weak linear correlation. When it comes to analyzing time series data, autocorrelation becomes a crucial extension of correlation. Autocorrelation describes the average dependence between values of the same series but shifted by a lag ‘ ’*k*’’. Autocorrelation is used to characterize linear dependencies in residual time series, which are series corrected for trend and seasonality [29]. The autocorrelation function is defined by the following formulas:

$$\rho_k=Corr(x_t,x_{t-k})$$

$$=\frac{Cov(x_t,x_{t-k})}{\sqrt{Var(x_t)\times Var(x_{t-k})}}$$

$$=\frac{\gamma_k}{\gamma_0},\quad \mathrm{k}\in Z \tag{I.10}$$

In addition to autocorrelation, there is also partial autocorrelation, which captures the direct correlation between observations at a specific lag while neutralizing the influence of intermediate lags up to *k*. This provides a more nuanced understanding of temporal dependencies in a time series.

- **Stationarity and Non-stationarity**

The concepts of stationarity and non-stationarity are primary concepts in time series analysis. A time series is deemed stationary if its statistical properties, such as mean, variance, autocorrelation, and others, remain invariant over time and do not undergo significant variations, as opposed to non-stationary series, whose statistical properties change with time and indicate a variability in the distribution of data [30].

I.2.2.3. Components of time series

A time series *y(t)* is defined by four fundamental components or variations which are Trend, Seasonality, Cyclicity and Irregularity or Residues. These elements, explain the overall structure of the time series and contribute to understanding the underlying behavior and models of a time series and are considered while analyzing, modeling, and predicting temporal data. The **trend** represents the long-term evolution of the investigated series, that is, the overall direction in which the time series evolves over a prolonged period of time. It might be increasing, decreasing, or steady and linear or non-linear, for instance, the long-term increase in the stock market in the past year, the temperature rising or falling in a single day, the longevity of people’s lives…etc. The **seasonality** refers to regular cyclical variations that occur in the time series at fixed time intervals. In general, it is a seasonal phenomenon that repeatedly occurs at the same period or point every year, such as climatic variations, sales of products during the cold and rainy or warm season (umbrellas, coats, ice cream…etc), energy consumption in summer and winter…etc. The **cyclicity** on the other hand, represents a periodic change in data over time. It is a sequence of ups and downs that reproduce regularly, usually over a longer period than seasonal variations. e.g. the year has cycles of spring, summer, autumn and winter that come back every year similarly. And finally, **irregularity or residues** that describe fluctuations or noise that are random, uncontrollable, and non-systematic in the time series. These fluctuations are unexplainable and unpredictable. It might be because it includes random influences such as earthquakes, floods, or strikes, as well as measurement errors. Noise can make time series analysis and modeling more complex [31].

I.2.2.4. Time series decomposition scheme

Decomposition schemes or models have been proposed and used; these models define the mathematical formula of the link between the components of the time series. This helps to better understand how each of the time series' components (trend, seasonality, cyclicity, and irregularity) contributes to the data set and to in-depth analysis. There are three types of models: **additive**, **multiplicative**, and **mixed** models, which will be introduced below [32].

- ✓ The additive scheme presumes that the four components are added collectively to create the complete series. They operate independently and are expressed as follows:

$$y_t = T_t + S_t + C_t + R_t \tag{I.11}$$

- ✓ In the multiplicative scheme the numerous components act in proportion to one another. According to this model:

$$y_t = T_t \times S_t \times C_t \times R_t \tag{I.12}$$

- ✓ For the mixed scheme, T_t and C_t are related to each other, but independent of S_t and R_t, and the model will be as follows:

$$y_t = T_t \times C_t + S_t + R_t \tag{I.13}$$

In these schemes, y_t represents the time series, while T_t, S_t, C_t and R_t are the trend, seasonality, cyclicity and irregularity at time t respectively. The determination of the series' components along with the kind of model that establishes the mathematical formula for the link between these components intends to determine the extent to which they affect the values of the studied phenomenon in order to isolate these external effects and determine the true or significant values.

I.2.2.5. Basic models of time series

Time series, being chronological data sequences, necessitate the use of particular analytical methods to understand underlying trends and patterns. At the heart of this investigation are critical statistical models that have been extensively acknowledged for their efficiency in a variety of study disciplines. They are among the oldest uses of time series analysis and are essential for many empirical studies. These models provide a profound understanding of the temporal structures of data, allowing for better prediction and interpretation of temporal occurrences. In this section, we will look at three key models that are important in time series analysis: the Autoregressive Model (AR), the Moving Average Model (MA), and the Autoregressive Moving Average Model (ARMA), detailed as follows [33, 34]:

➢ Auto-regressive Model (AR)

The AR model is among the first models employed in time series analysis and is based on the concept that the current value of a time series is a linear combination of past values. Mathematically, an AR model of order p (AR(p)) can be represented by the following equation:

$$x_{(t)} = \sum_{i=1}^{p} w_i x_{(t-i)} + \varepsilon_{(t)} \tag{I.14}$$

where $x(t)$ represents the value of the time series at time t, p is the order of the model, w_i are the autoregressive parameters, and $\varepsilon(t)$ denotes the error term.

➢ Moving Average Model (MA)

In contrast to the AR model, the MA model relies on past errors of the time series, assuming that the future value depends linearly on these past errors. An MA model (MA(q)) of order q is defined by the equation :

$$x(t) = \sum_{i=1}^{q} \theta_i \varepsilon(t-i) + \varepsilon(t) \tag{I.15}$$

where θ_i are the moving average parameters, q is the model's order, $\varepsilon(t-i)$ denotes the error term.

➢ Auto Regressive Moving Average Model (ARMA)

The ARMA model combines aspects of AR and MA models to provide a more comprehensive modeling of time series. For ARMA (p,q) model, output is modeled as a linear difference equation between the current and past inputs and past outputs as described in the following equation:

$$x(t) = \sum_{i=1}^{p} w_i x(t-i) + \sum_{i=1}^{q} \theta_i \varepsilon(t-i) + \varepsilon(t) \tag{I.16}$$

The previously mentioned models are among the most basic and widely recognized. Variations of these models are documented in the literature, we mention: ARIMA model (Auto Regressive Integrated Moving Average), SARIMA model (Seasonal Auto Regressive Integrated Moving Average), SARIMAX model (Seasonal Auto Regressive Integrated Moving Average with eXogenous variables), ARCH model (Auto Regressive Conditional Heteroskedasticity), GARCH model (Generalized Auto Regressive Conditional Heteroskedasticity), Exponential Smoothing model (ETS), VAR model (Vector Auto Regressive). As they will not be the focus of our study, we have briefly mentioned them and refer the reader to references [34, 35, 36] for a detailed explanation of these models.

The choice of model often depends on the specific nature of the time series being analyzed. It requires a thorough understanding of the time series' unique characteristics, as well as particular modeling objectives. This approach guarantees a precise depiction of temporal patterns and promotes accurate modeling and analysis of the time series.

I.2.2.6. Objective of Analyzing and processing time series

Time series analysis and processing are important areas of study in the research and investigation of chronological data that is frequently employed in various sciences and techniques. It entails modeling, prediction, filtering, and classification, with the objective of extracting important information and other data features, determines patterns and tendencies, and reveals underlying structures inherent in temporal data, while recognizing seasonal and cyclical variables. The principal objectives are summarized in the following points [25, 37]:

- ✓ Visualize and decompose the components of the series.
- ✓ Visualize and dissect the series' components.
- ✓ Identify the time series' normal and irregular oscillations.
- ✓ Recognize the temporal behavior of data.
- ✓ Estimate future aspects of a series using current and historical data (explain how the past influences the future).
- ✓ Create a model that can anticipate future observations based on available data.
- ✓ Filtering: identifying and extracting relevant temporal structures.
- ✓ Prediction of future achievements of the series based on its past values.
- ✓ Different series classification.
- ✓ Control and Anomaly detection: detect abnormal or unexpected data items.
- ✓ Visualize what factors impact certain variables from period to period.

I.2.2.7. Examples of time series

Time series is found in diverse fields such as engineering, science, sociology, and economics, public health, and beyond. This diversity of applications suggests that time series play a key role when examining data that varies over time, within dynamic phenomena studied in a multitude of fields. Among them, we can mention [38]:

- ✓ The meteorological time series that record variables like temperature, humidity, atmospheric pressure, etc. over a certain period of time.
- ✓ Financial data, including stock prices, change rates, and transaction volumes, are used to analyze market trends and models.
- ✓ Traffic data captured by sensors on highways and city routes can also be shown as time series to study traffic patterns, flow models, and other traffic-related phenomena.
- ✓ Production and analysis data for drinking water.
- ✓ Biomedical time series, such as electrocardiograms (ECGs), electroencephalograms (EEGs), blood pressure monitoring, etc., are data that are gathered at regular intervals in the biomedical area and show changes in specific biological or medical variables over time.

These examples are just a glimpse into the vast array of time series. As part of our study, we are particularly interested in ECG time series. In the following we discuss the concept of ECGs in detail, covering the anatomy of the heart, signal generation mechanisms and the various noises that can affect them.

A. ECG time series

ECG represents a worthwhile time series that record the electrical activity of the heart throughout time. ECG signals provide a dynamic representation of the heart's electrical behavior, revealing noteworthy insights in cardiac rhythm, anomalies, and overall cardiovascular function. Analyzing ECG time series data is essential for detecting irregularities, understanding temporal patterns, and constructing predictive models for cardiovascular incidents. Furthermore, the frequent availability of ECG data in medical research is an important and well-studied time series domain. In this study, our specific focus is on enhancing ECG analysis through filtering techniques. The denoising process is crucial to remove unnecessary noise or artifacts, allowing for a clearer representation of the underlying cardiac activity. By applying advanced filtering methods to ECG signals, we intend to increase signal quality and extract meaningful information, contributing to the broader field of cardiology research.

A.1. Cardiac Anatomy and electrical Activity

➢ Heart Structure

Heart is a vital organ in the human body, in the form of a hollow muscle of immense importance, located in the center of the thorax, positioned slightly to the left, concealed behind the sternum, between the lungs, and above the diaphragm, in front of the spinal column. Its acts as a pump, propelling blood through blood vessels to ensure the distribution of oxygen and nutrients to organs, tissues, and cells. Comprising four chambers (Figure I.2), the left and right atria, as well as the left and right ventricles, the heart operates in a coordinated manner. The atria receive blood from the veins, while the ventricles eject it into the arteries. Valves, such as the mitral valve and the tricuspid valve, regulate blood flow between the different chambers, ensuring efficient functioning. This constant movement of blood, propelled by the rhythmic contraction of the cardiac muscle, enables the body to maintain an adequate supply of oxygen and nutrients while eliminating metabolic waste. Thus, the heart embodies a complex and essential mechanism that ensures the necessary blood circulation for the proper functioning of the organism [39, 40].

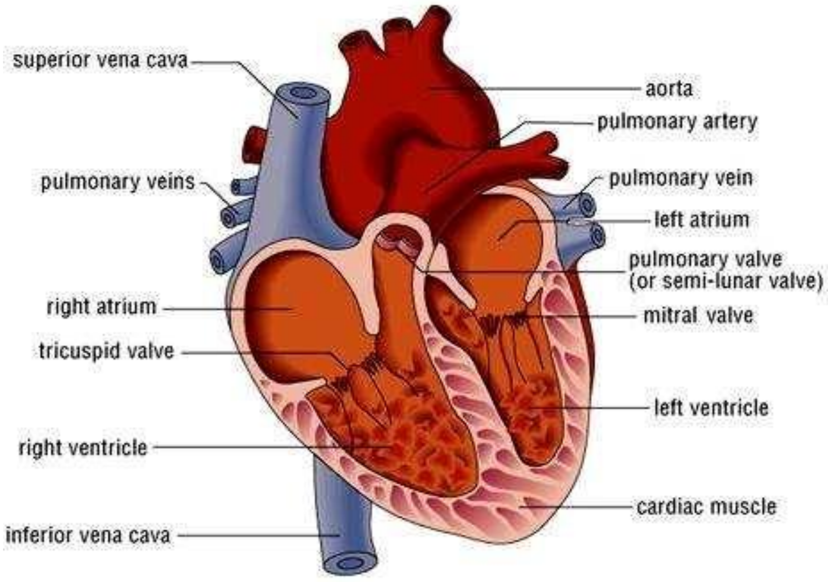

Figure I.2 – Structure of the human heart [41].

➢ Electrical functioning of the Heart

The heart has an integrated network of cells responsible for generating and transmitting electrical pulses, as well as cells capable of responding to these pulses by contracting. These include nodal cells, such as the sinus node (SA) in the right atrium taken from the entrance of the superior vena cava, the atrioventricular node (AV) located on the other side of the right atrium, and the His Bundle and Purkinje's network. On the other hand, cardiac muscle cells, also known as cardiomyocytes (Figure I.3). The sinus node is the pacemaker of the heart and it initiates all the heartbeats of the heart and determines the heart rate. The electrical signal generated in the SA node extends to the atria and has the effect that they contract. The function of the AV node is to transmit the electrical pulse from the atria to the ventricles with a small time delay. This delay is necessary to ensure that the atria have ejected all blood into the ventricles before the ventricles contract. The electrical signal then passes into the barter of the Bundle of His, which after a short journey divide into two branches, one for the right ventricle and the other for the left ventricle. The signal then passes into Purkinje network and extends to the ventricular myocardium. Finally, the pulse triggers the contraction of the ventricles, which expels blood to the lungs and the rest of the body. These intricate electrical processes, which represent the heart rhythm, can be recorded in the form of an electrocardiogram (ECG), providing a detailed insight into the electrical function of the heart [42, 43].

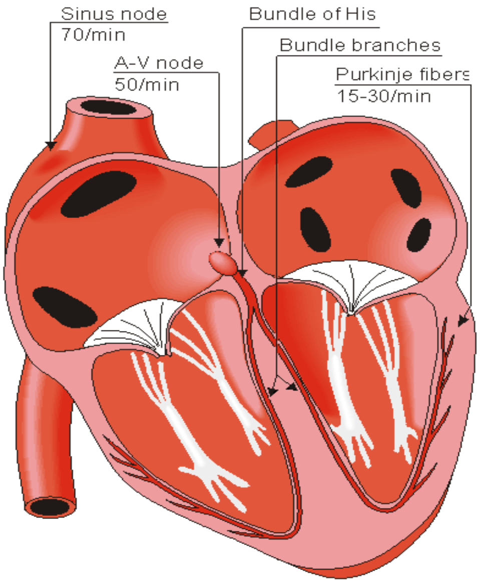

Figure I.3 – The electrical conduction system of the heart [44].

A.2. Electrocardiography

➢ Definition

Electrocardiogram, often abbreviated as ECG, is one of the key tools for interpreting and determining cardiovascular illness. It involves the graphic recording of the variations in electrical impulses originated by the heart's electrical activity over a period of time [45].

➢ ECG Recording Procedure

The human body conducts electricity, allowing the collection of potential actions generated during the electrical activity of the heart. This is done by placing electrodes on the skin, which record variations in electrical potential throughout the cardiac cycle. The precise location of electrodes corresponding to the different leads, as well as their number, depend on the type of electrocardiographic curve and the specific information the doctor wishes to obtain. There are two main types of leads: bipolar leads when both electrodes face sites with similar potential variations and unipolar leads when recording potential variations is between a virtual point (reference point) and an electrode [46].

The standard ECG includes both bipolar and unipolar leads, thus forming a system of 12 complementary leads standardized by an international convention (I, II and III which are indirect and bipolar; AVR, AVL, and AVF are indirect and unipolar; and V1 to V6 are semi-direct) [47, 48, 49]. The electrode placements associated with each type of lead will be outlined in the following paragraphs.

- **Frontal leads:** also called peripheral leads, these are leads that explore the heart in a frontal plane. They are placed on the limbs (ideally on the wrist for the arms and on the ankles for the legs) and there are six (06) of them:

  a. **Bipolar limb leads:** designated as DI, DII, and DIII, as determined by Einthoven in 1906. These leads reflect the potential difference between two limbs, thus forming Einthoven's triangle (Figure I.4). The electrode positions are as follow :

    - ✓ D1: connects the right arm to the left arm, with electrodes placed at the wrists.
    - ✓ D2: connects the right arm to the left leg (at the ankle).
    - ✓ D3: connects the left arm to the left leg.

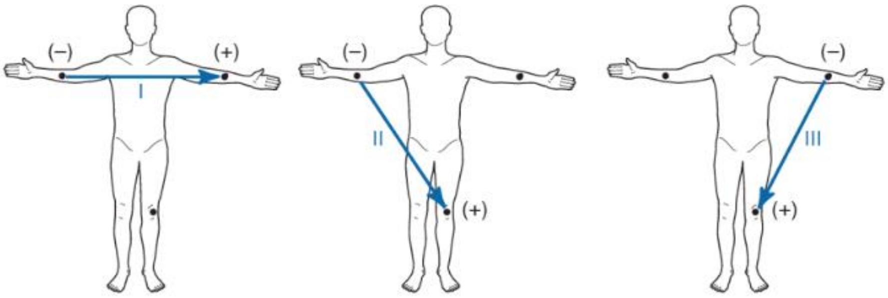

Figure I.4 – Bipolar limb leads [49].

The potential differences between these vertices are measured as follows:

- ✓ Lead I : DI= VL − VR
- ✓ Lead II : DII=VF − VR
- ✓ Lead III : DIII=VF – VL

with VL representing the potential on the left arm, VR the potential on the right arm, and VF the potential on the left leg:

b. **Unipolar limb leads:** designated as aVR, aVL, and aVF, these leads were initially introduced by Wilson in 1934. They measure the potentials of each point of Einthoven's triangle in relation to a reference point. Subsequently, Goldberg introduced the concept of augmented unipolar leads, known as aVL, aVR, and aVF (Figure 1.5). The unipolar limb leads are obtained by placing the electrodes as follows:

- ✓ aVR : one electrode on the right arm (wrist);
- ✓ aVL : one electrode on the left arm;
- ✓ aVF : one electrode on the left leg (ankle).

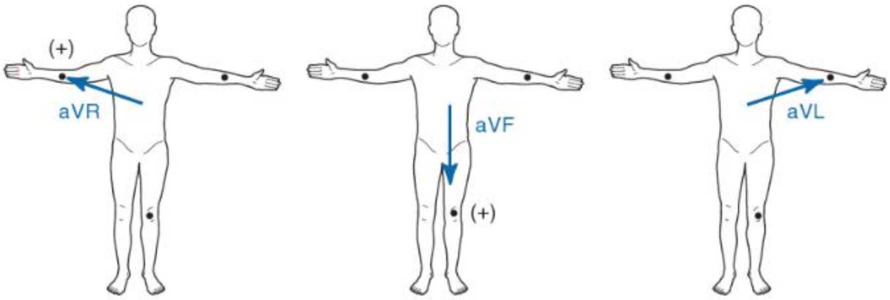

Figure I.5 – Unipolar limb leads [49].

- **Precordial Unipolar Leads:** Precordial leads are leads that explore the heart in a horizontal plane. They are placed in well-defined areas on the chest wall and there are six (v1 to v6) as well. The position of the precordial electrodes is therefore as follows (Figure I.6):

  - ✓ V1: Placed at the 4th intercostal space along the right border of the sternum,
  - ✓ V2: Placed at the 4th intercostal space along the left border of the sternum,
  - ✓ V3: Midway between electrode V2 and V4,
  - ✓ V4: Placed at the 5th intercostal space on the left, along the midsternal line,
  - ✓ V5: Placed at the 5th intercostal space on the anterior axillary line,
  - ✓ V6: Placed at the 5th intercostal space on the midaxillary line.

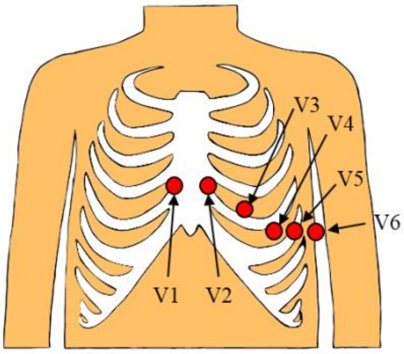

Figure I.6 – Location of precordial electrodes: V1 to V6 [50].

- ➢ **ECG Waves and intervals**

An ECG is a composite record of all action potentials. The action potentials represent electrical currents produced by the nodes and cells of the myocardium during the depolarization and repolarization process. The ECG signal is characterized by five waves named in alphabetical order the P, Q, R, S and T. Their shape, amplitude, and time intervals give crucial details regarding health and heart's condition (Figure I.7). Each wave or segment of the ECG corresponds to a certain event in the cardiac electrical cycle. When the atria are full of blood, the SA node triggers a pulse that is quickly transmitted to the atria and causes their depolarization. This is represented by the P wave on the ECG. The contraction of the atria, or atrial systole, begins about 100 milliseconds after the start of the P wave. The PQ segment represents the time required for the transmission of the electrical pulse from the SA node to the

AV node. The QRS complex marks the triggering of the AV node and represents ventricular depolarization (the Q wave corresponds to the depolarization of the interventricular septum, while the R wave is produced by the depolarization of the main mass of the ventricles, and finally the S wave represents the last phase of the ventricular depolarization at the base of the heart). Atrial repolarization also occurs during this period, but the signal is obscured by the large QRS complex. The ST segment reflects the action potential plateau of the myocardium. This is when the ventricles contract and pump blood. The T wave represents ventricular repolarization immediately before ventricular or diastole relaxation [51].

The cycle repeats with each heartbeat.

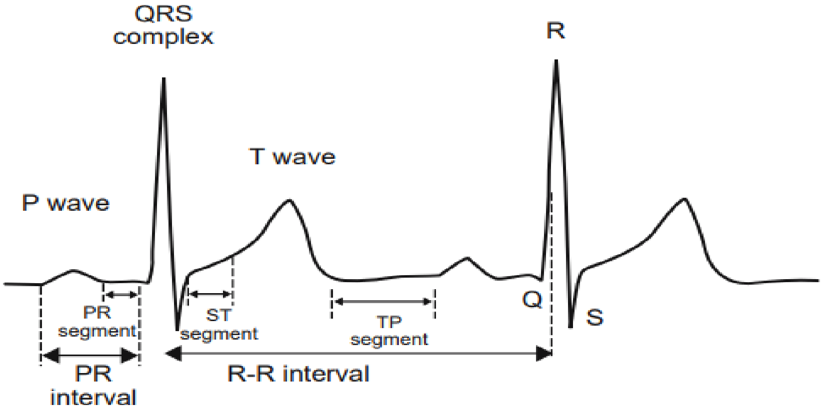

Figure I.7 – Normal ECG signal [52].

A.3. Noise and Variability in ECG recordings

ECG measurements can be altered by various types of noise During its recording originating both within and outside patients' bodies, making their analysis considerably more complicated. Among the disturbances identified as being of primary interest, we distinguish: Power Line Interference, Electrode Contact Noise, Motion Artifact, Muscle Contractions or electromyographic noise (EMG) and Noise Generated by Electronic Devices [53, 16, 54], detailed as follows:

- **Power Line Interference**

Power line interference refers to undesirable disturbances in electrocardiogram (ECG) coming from the power supply through the electrical distribution network, often exacerbated by inductive and capacitive coupling phenomena. Inductive coupling is generated by the existence of a mutual inductance between two conductors. When current flows through these wires, it generates a magnetic field capable of inducing current in adjacent circuits, while Capacitive coupling occurs when there is an electric field between power lines and nearby conductors. These disturbances are typically characterized by a signal frequency of 60 or 50 Hz depending on the region, and exhibit harmonics. These harmonics can be effectively modeled as sine waves or combinations thereof, with an amplitude that can reach as much as 50% of the ECG's amplitude from peak to peak. The origins of these disruptions often stem from the parasitic impacts of alternating current fields associated with loops in the patient's cables, as well as issues like loose contacts and unclean electrodes. Particularly noteworthy is the impact of a disconnected electrode, capable of generating a potent disturbance signal. Failure to promptly address these interferences can lead to the complete obscuring of the ECG waveform, thereby compromising the accuracy of measurements [**53**, **16**].

➢ Electrode Contact Noise

Electrode contact noise occurs when the electrode momentarily or permanently loses contact with the skin. This loss generates significant disturbances in the ECG signal due to capacitive coupling. This noise is characterized by a sudden, random disturbance of the basic signal, rapidly followed by a return to the initial position, with the addition of an extra component at a frequency of 60 Hz. This disconnection of the skin electrode leads to desynchronization of the measurement system [**53**, **16**].

➢ Motion Artifact

Motion artifacts are generated by electrode movements away from the contact zone on the skin. These disturbances, challenging to predict due to their unpredictable morphology and frequency, result in abrupt changes in the amplitude of the ECG signal and alterations in the low-frequency baseline. The typical duration of these artifacts is 100 to 500ms, with a magnitude reaching 500% of the ECG’s amplitude from peak-to-peak [**53**, **16**].

➢ Muscle Contractions or electromyographic noise (EMG)

Muscle contraction induces electromyographic noise resulting from the activation of muscles other than the heart. When neighboring muscles contract, they generate depolarization and repolarization waves captured by the ECG. These EMG signals, accounting for approximately 10% of the ECG amplitude, manifest as bursts of Gaussian noise lasting around 50ms, covering a frequency range from direct current to 10 kHz. The interference of these signals with the ECG produces artifacts, sharing a similar frequency spectrum and possessing significant energy [**53**, **16**].

- **Noise Generated by Electronic Devices**

Refers to instrumentation noise produced by the equipment composing the recording section (probes, cables, analog to digital converter, etc.). Obviously, these types of interferences could be significantly reduced by a careful choice of high-quality devices [**16**, **54**].

These interferences (or disturbances) can alter ECG measurements, cause changes in its characteristics and making its analysis particularly difficult and complicate, resulting in degradation of signal quality and frequency resolution. In addition, they generate large amplitude signals in the ECG, which can lead to confusion. As a result, these interferences mask subtle but crucial features for clinical monitoring and diagnosis. The key task is therefore to eliminate these artifacts from ECG signals, an approach considered crucial to improve diagnostic accuracy.

I.3. Part II: State of the art

I.3.1. State of the Art in modeling and identification of dynamical systems and time series

Identification and modeling issues arise from the need to comprehend and predict the complex and evolving behaviors of dynamical systems and chronological series. Over the years, researchers have suggested approaches to solve these issues. However, identification and modeling remain challenging issues that call for a cautious and adaptable approach to meet the particular challenges associated with each context. Among these methods are the following:

In 2000, Lee, C. H & Teng, C. C. [**55**] proposed a Recurrent Fuzzy Neural Network (RFNN) structure intended for the identification and control of nonlinear dynamic systems. The Recurrent Fuzzy Neural Network (RFNN) is a type of multilayered connectionist network that is specifically intended for fuzzy inference through the use of dynamic rules. The second layer of the network incorporates feedback connections to accommodate temporal relations. The network can better handle temporal problems as a result of this extension.

Also in 2000, a method for developing an evolutionary basis of optimal fuzzy rules for the modeling and control of dynamical systems was introduced by Kang, S. & *al.* [**56**]. Using evolutionary programming, the structure and parameter of a fuzzy rule base are simultaneously evolved for a specific task.

During the year 2002, Juang, C. F. [**57**] presented a TSK-type recurrent fuzzy network (TRFN) structure that can be created using genetic algorithms or neural networks, depending on the learning scenario. Recurrent fuzzy if-then rules with TSK-type consequent parts are incorporated into TRFN, which uses internal variables to improve learning and help with temporal history memorization. For various learning environments, two variants are suggested: TRFN-S (supervised learning) and TRFN-G (genetic learning). When compared to other recurrent networks and configurations, TRFN shows high learning efficiency and accuracy in dynamic system identification and control.

Subsequently in 2006, a novel fuzzy inference system for modeling nonlinear dynamic systems with input and output data affected by measurement noise is presented in [**58**] by Kim, J & *al.* The extended relevance vector machine (RVM) is used by the system to automatically produce membership function parameters and fuzzy rules. With its robust generalization ability and probabilistic Bayesian learning framework, the RVM optimizes the structure of the fuzzy system in a manner akin to the Takagi-Sugeno model. Notably, the suggested technique makes use of the gradient ascent method to optimize a marginal likelihood, thereby minimizing the quantity of fuzzy rules. The least squares method is used to determine the subsequent part coefficients.

Progressing to 2007, An effective technique for automatically creating Takagi-Sugeno (TS) fuzzy models which are essential in fuzzy sets theory has been presented by Chafaa, K & *al.* [**59**]. Structure and parameter identification are the two processes involved in the generation of the TS fuzzy model. For structure identification, a clustering technique based on the Gustafson-Kessel algorithm is suggested, and for parameter identification, the Kalman filter algorithm is applied twice. Based only on input-output data, this approach effectively creates TS fuzzy models automatically when used to identify a dynamic system.

In the same timeframe, A new method for identifying and controlling a nonlinear system was proposed in [**60**] by Singh, M & *al.* This approach relies on the use of a neural network model and a learning algorithm to update the parameters, thus simplifying the task of calculating the system parameters compared to traditional methods of extracting the dynamics of nonlinear systems.

In the year 2009, Zhao, H & Zhang, J [**61**], suggested an artificial neural recurrent network with functional connectivity and data flow (PFLARNN) as an efficient computational solution for the identification of non-linear dynamic systems. The PFLARNN is made up of multiple basic neural network modules with recurrent neurons at functional liaison (FLARNN). These modules, when run concurrently in pipeline mode, significantly increase the PFLARNN's overall computational efficiency. Additionally, a nonlinear functional expansion is added to the entry motif in order to introduce nonlinearity into each module.

Moving to 2010, a recurrent modeling method based on a neuro- fuzzy network named LRFNN-SVR (locally recurrent fuzzy neural network with support vector regression) was suggested by Juang & *al.* in [62]. This approach uses the properties of time to treat the problems and has a five-layer structure. The resulting couche is of the Takagi-Sugeno-Kang (TSK) type, which is a linear function of the current states. A linear and iterative approach known as Support Vector Regression (SVR) was used to modify the parameters of the fuzzy rules.

As we advance to 2012, Lin, Y. Y. & *al.* [63] have introduced the IRSFNN, a novel neural flow network for the prediction and identification of dynamic systems. Equipped with a recurrent structure based on external bounces and internal retroactions, the IRSFNN makes use of an FLNN in its significant portion for improved mapping capability. With the help of a Kalman filter, an online clustering algorithm, and simultaneous learning of structure and parameters, the IRSFNN is able to achieve better performance when compared to other neural networks with fuzzy recurrent connections.

Also in 2012, a novel neural network auto-organized with radial basis functions (SORBF) has been proposed by Qiao J-F and Han H-G [8] for the identification and modeling of non-linear dynamic systems. The SORBF allows for simultaneous network construction and parameter optimization, enabling dynamic adjustment of the number of cached neurons. With a convergence analysis, the simplified optimization algorithm significantly improves the model's performance. The simulations show that this approach is effective in identifying non-linear dynamical systems.

Remaining in the same year, Cavuslu & *al.* [64] introduced hardware implementation of artificial neural networks with learning ability on field programmable gate array (FPGA) for dynamical system identification using the improved particle swarm optimization. The enhanced PSO is obtained through modification of the velocity update function. By including an additional term in the velocity updating function decreased the possibility of getting trapped in a local minimum.

In 2013, Han, H. & *al.* have [65] introduced a novel A self-organizing fuzzy neural network with adaptive computation algorithm (SOFNN-ACA) method for effectively modeling a class of uncertain nonlinear systems. The proposed approach, combining structure design and parameter learning, was successfully applied to model a classical nonlinear dynamic system. The results demonstrated superior learning efficiency and performance compared to various other algorithms commonly used in the literature.

In 2013 as well, Rios, R. & *al.* [66], have put forth a strategy to improve time series modeling by taking into consideration both deterministic and stochastic influences. There are two steps to the approach. First, observations are split into two parts, one of which represents deterministic influences and the other, stochastic influences. For this decomposition, the Empirical Mode Decomposition (EMD) technique is used. Second, models are improved for each component independently before being combined to form a hybrid model that enhances time series analysis as a whole.

During the year 2014, a framework for modeling nonlinear dynamical systems involving Gaussian membership functions and linear differential equations known as the dynamic adaptive neuro-fuzzy inference system (DANFIS) is introduced by Yilmaz, S and Oysal, Y. [67]. The Broyden-Fletcher-Goldfarb-Shanno (BFGS) optimization algorithm is employed, and gradients for parameter optimization are computed using the adjoint sensitivity method. Through the use of simulations on a circuit featuring a tunnel diode and the Van der Pol oscillator, the viability of the suggested method is tested.

In the same year, Kayacan, E. & *al.* [68], presented a novel fully sliding mode parameter update rules for training interval T2FNNs to identify nonlinear and time-varying dynamic systems. Two nonlinear systems are identified in order to evaluate the learning algorithm's performance and show its potential for real-time applications. The suggested algorithm's high identification accuracy and much shorter computation time when compared to alternative approaches are highlighted by the simulation results. When computation time is critical, the parameter update rules can also be used for control.

Also in 2014, a neuro-fuzzy auto-organization system with an adaptive computation algorithm dubbed SOFNN-ACA (self-organizing fuzzy neural network with adaptive computation algorithm) was proposed by Honggui & *al.* in [69]. During the learning phase, a set of fuzzy rules can be reduced or added by using the following methodology: The rules that are fuzzy will be treated as a new set of rules if they have high point-to-value intensities (SI: Spiking Intensities). If, on the other hand, the rules that are fuzzy have a low relative mutual information value (RMI: Relative Mutual Information), these rules will be tailed in order to simplify the adopted neuro- fuzzy network structure, where a strategy based on learning rate has been employed to accelerate convergence.

Moving forward to 2016, Hichem Loussifi & *al.* [70] have proposed a novel hybrid intelligent approach to construct fuzzy neural networks (WK-FNN) by combining kernel approaches with wavelet techniques. This method makes use of multi-resolution wavelet analysis (MRA) as machine learning kernel functions as well as activation functions for neural network architecture. Compared to conventional techniques, this method enhances accuracy in identifying dynamic systems and predicting chaotic time series by using exact scale parameters based on wavelet analysis.

As we advance to 2017, Han, H. & *al.* [71], have presented the SOFNN-AGA, an auto-organized neural network with gradient adaptation algorithm for modeling non-linear systems. It employs a potentiality of rules floues (PFR) method for the training process and a PFR-based approach to structural learning to ascertain the network's size. A new gradient adaptive algorithm (AGA) is designed to adjust the parameters of the SOFNN-AGA, along with a theoretical analysis of its convergence in cases of fixed and auto-organized structure.

Remaining in 2017, Wang, Y. [72] in his paper explores the application of Long Short-Term Memory (LSTM) Recurrent Neural Networks (RNN) for identifying and controlling dynamic systems. The study delves into both the identification of dynamic systems and the design of controllers based on the identification.

During the same time frame, Hosein & *al.* [73], have developed another method for modeling temporal series based on fuzzy neural network inference systems (ANFIS: Adaptive Neuro-Fuzzy Inference Systems). To model non-linear functions, the ANFIS integrates the artificial neural network learning capability with the Takagi-Sugeno type fuzzy inference system. The proposed system's parameters have been adjusted through the use of a method that combines the least squares estimator (LSE) with the Bee Colony Optimization Algorithm.

In 2017 as well, a neuro- fuzzy auto-organization system, called SOFNN-AGA (Self-Organization Fuzzy Neural Network with Adaptive Gradient Algorithm), was proposed by Hanggui & *al.* in [74]. The authors have developed an approach to learning fuzzy neural structure that is based on the baptized method (PFR: Potentiality of Fuzzy Rules) and error reduction ratio (ERR: Error Reduction Ratio). A parameter adjustment strategy based on the gradient based adaptive learning rate method has been applied.

In 2019, Li, Q. & *al.* [75], has introduced a novel kind of floating-point neural network (FNN) called the self-constructing recurrent fuzzy neural network (SCRFNN) for the identification of non-linear dynamic systems. This method enhances the traditional FNN structure by adding a recurring connection to each node of the cached layer, thereby enhancing the network's capabilities and functionalities. The learning process consists of two main stages: learning the structure, which is based on splitting the entry space, and learning the parameters, which is done with the use of a gradient descent method supervised by a delta adaptation law. Three real-world studies using SCRFNN validate the superiority of the proposed method over some existing neural networks.

In the same year, Dass, A. and Srivastava, S. [76] have suggested an innovative approach for modeling and controlling dynamic system employing two distinct recurrent fuzzy systems (RFS) architectures and highlighting their superiority over type-1 fuzzy systems. The two RFS architectures successfully accomplish the goals of system identification and control. An additional benefit of RFS, particularly in the processing of higher-order systems where a significant reduction in calculation complexity is possible. The proposed scheme illustrates the relative advantages of the two proposed fuzzy recurrent systems in comparison to the type 1 conventional fuzzy systems.

Also in 2019, a Neuro-Fuzzy Inference System with Dynamic Neurons (NFIS-DN) for time series forecasting and dynamic system identification is presented in [77] by Samanta, S. & *al.* NFIS-DN makes use of Dynamic Neurons (DNs), which are based on a finite-memory discrete-time nonlinear state-space model. The five-layer model uses DNs in the layers that handle sharp values. An algorithm that uses time learning and self-regulated backpropagation is used to update the antecedent and consequent parameters.

During the year 2020, Rajendra, P. and Brahmajirao, V. [14] presented a review of modern perspective on dynamic systems, focusing on current challenges such as identifying dynamics from data and creating data-based representations for linear analysis of nonlinear systems. It highlights the crucial role of machine learning techniques, including dimensionality reduction and deep learning methods, in solving these challenges to efficiently model complex dynamic systems.

Additionally, in 2020, a new method for dynamic Systems Identification using a Broad Neural Network and Particle Swarm Optimization is presented by Han, R. & *al.* [78]. The proposed algorithm combines a Broad Learning System (BLS) with Particle Swarm Optimization (PSO) to identify nonlinear dynamical systems. The method involves dimension expansion of the dataset as BLS input, followed by model weight optimization using the PSO algorithm.

Remaining in the same year, a novel Strategy to addresses the challenge of system identification for a dynamical system, described by a nonlinear equation system, using discrete time series data is described by Niven, R. & *al.* in [**79**].

Progressing to 2021, Mao, W. & *al.* [**80**], suggested an approach that combines the Biogeography-Based Optimization (BBO) learning algorithm with a Multilayer Perceptron (MLP) structure. The MLP is used to model dynamic and nonlinear systems because of its universal approximation capability. In this method, the learning algorithm for the MLP is BBO instead of the conventional back-propagation algorithm, which is a popular gradient descent technique. When it comes to system identification, the suggested method with BBO learning outperforms the conventional gradient descent method.

I.3.2. State of the art in compressing ECG time series data

In recent years, advances in storage and transmission technologies have made it possible to process large volumes of data, particularly medical data such as ECGs, making compression essential. In simple terms, data compression is the action of reducing the physical size of a block of information. The primary objective of any compression technique is to achieve maximum data volume reduction while preserving the main morphological characteristics of the signal during reconstruction to ensure accurate medical diagnosis. A great deal of work has been carried out in this area, including :

In 2000, Ahmed, S.M. & *al.* [**81**], introduced a technique for compressing ECG signals using a novel category of discrete wavelet transforms (DWT) utilizing non-orthogonal bases. This method aims to minimize the distortion of the compressed signal, measured by the percentage root mean square difference (PRD), while retaining clinically significant features. DWT coefficients are calculated to minimize the difference between the original and reconstructed signals using the least squares criterion.

In 2001, a compression method Utilizing optimized quantification of discrete cosine transform coefficients (DCT) was put forward by Batista, L.V. & *al.* [**82**]. The method involves dividing the ECG signal into fixed blocks and then quantizing each DCT block using quantization vectors and thresholds especially determined for every signal. The estimated entropy is minimized by optimizing these vectors with Lagrange multipliers while respecting the reconstruction signal's distortion.

In the same year, Wei, J.J. & *al.* [**83**] also presented a method for compressing ECG data based on truncated singular value decomposition (SVD). SVD is used to extract the significant components of the ECG by decomposing the signal into basic patterns with associated scaling factors. As the signal information is mainly concentrated in a number of singular values because of the high correlation between ECG heartbeats, solely the relevant portions of the singular triplets are kept as compressed data. Insignificant components are truncated to eliminate redundancy.

The following year, in 2002, a wavelet-based ECG data compression algorithm is proposed by Rajoub, B.A. [**84**]. For this method, the ECG signal is pre-processed and then the discrete wavelet transform (DWT) is applied. DWT coefficients are divided into three groups and thresholded based on energy clustering efficiency. A binary significance map is generated to indicate significant coefficients. Compression is performed using variable length coding for the significance map and a direct binary representation for the significant coefficients.

In 2003, Al-Shrouf, A. & *al.* [**85**] suggested a new ECG compression algorithm. It compresses the linearly predicted residuals of the wavelet coefficients of the signal, using coding techniques. The error between the wavelet coefficients and the predicted coefficients is minimized to obtain the best predictor.

Subsequently in 2004, a novel filter bank-based algorithm for compressing ECG signals is proposed by Blanco-Velasco, M. & *al.* [**86**]. It consists of three steps, starting with the decomposition in sub-bands and contrasting the wavelet packets (WP) method's performance with that of a quasi-perfect reconstruction (N-PR) filter bank. The aim is to assess how well the reconstructed signal matches predefined precision criteria. According to the outcomes, the N-PR filter bank technique performs better in terms of efficiency and quality than the WP approach.

Also in 2004, Borsali, R. & *al.* [**87**] introduced an ECG compression technique. This technique combines both strategies: polynomial modeling and ECG beat alignment. To lessen high-frequency variations from one beat to the next, QRS complexes are first identified and then aligned. These alterations are modeled using polynomial projection.

Moving into 2006, Ku, C.T. & *al.* [**88**], explored the use of complete wavelets in their study for ECG data compression. This involves transforming the data into non-recursive unidimensional discrete periodized non-recursive wavelets (1-D NRDPWT) with minimal registration length. Additionally, a method with a high compression rate is developed for word length reduction. This algorithm uses small, large quantification scales along with high octave coefficients without additional division.

Advancing to 2007, Ouamri, A., & Naït-Ali, A. [89], proposed an ECG compression algorithm using a combination of Lorentzian functions. To estimate the parameters of the Lorentzian functions, the discrete Fourier transform (DFT) is first applied to an averaged ECG signal, retaining only the most significant DFT coefficients. These coefficients are then modeled as the sum of a number of exponentially damped sinusoids (EDS), identified by their amplitudes, real damping factors, frequencies and initial phases. The EDS parameters are then estimated using the SVD method and coded.

In 2008, an innovative denoising and compression schemes based on a modified Extended Kalman Filter (EKF) structure for ECG signals was introduced by Sayadi, O. & Shamsollahi, M. B. [90]. The new EKF structure is used to estimate the model parameters and reconstruct the signal with the dynamic equations of the model, improving both denoising and compression of ECG data.

By 2009, Ahmed, S. M. & *al.* [91], proposed a new two-stage hybrid approach for ECG signal compression based on the Modified Discrete Cosine Transform (MDCT) and the Discrete Wavelet Transform (DWT). The ECG signal is divided into blocks and MDCT is applied to each block to decorrelate the spectral information. DWT is then applied to the resulting MDCT coefficients. Spectral redundancy is eliminated by compressing subordinate components more than dominant components. The resulting wavelet coefficients are then thresholded and compressed using a meaningful bit map coding and energy clustering technique to save storage space.

In 2011, Abo-Zahhad, M. & *al.* [92] developed an effective technique for ECG signal compression based on discrete wavelet transform (DWT) and QRS complex estimation. The ECG signal is first preprocessed by normalization and mean removal. Next, an error signal is formed as the difference between the pre-processed ECG signal and the estimated QRS complex waveform, then wavelet transformed, and the resulting wavelet coefficients are thresholded by setting all coefficients below certain thresholds to zero. Threshold levels for all sub-bands are calculated as a function of energy pooling efficiency (EPE) to obtain the maximum compression ratio (CR) and the minimum percentage root mean square difference (PRD). A particular coding method is used to code the threshold DWT coefficients.

Also in 2011, an improved method of ECG signal compression based on the discrete cosine transform (DCT) was proposed by Bendifallah, A. & *al.* [93]. The method consists of segmenting the ECG signal into blocks and applying the DCT to each block. Then the DCT coefficients are quantized using a uniform scalar quantizer with a dead zone, which reduces the less significant coefficients to zero, thus effectively compressing the data. Finally, the quantized coefficients are encoded using arithmetic coding, a technique that efficiently compresses the data by reducing redundancy.

Moving into 2012, Ranjeet, K. & *al.* [94] presented an optimized ECG compression method using a wavelet filter bank derived by linear optimization and Kaiser windows, combined with run-length coding (RLE). RLE boosts compression without sacrificing any essential signal information, while Wavelets reduce compression distortion. This formulated technique uses modified thresholding, resulting in enhanced signal compression compared to other thresholding methods already in use.

In 2013, Abo-Zahhad, M. & *al.* [95] suggested an optimized ECG compression algorithm using wavelet filters and adjusted threshold levels to maximize data minimizing while ensuring quality in reconstructing. This algorithm segments ECG signal into frames, decomposes every frame into sub-bands, applies thresholding and then encodes the remaining coefficients with modified run-length coding. Threshold levels are adjusted to achieve a predefined compression ratio and signal quality.

Additionally, in 2014, Kanhe, R. K., & Hamde, S. T. [96] introduced a wavelet-based compression method for reducing ECG data while preserving the signal. Wavelets allow ECG signal decomposition and faithful reconstruction using analysis and synthesis filters. Analysis filter divides the signal utilizing high and low pass filters, while the decomposed parts are reconstructed through the synthesis filter. The reconstruction of the ECG signal using the synthesis filter is considered acceptable by cardiologists. Performance parameters such as the cross-correlation coefficient (CCC) support the effectiveness of this compression technique, even at high compression ratios.

Moving forward to 2015, a method of compressing ECG signals combines singular value decomposition (SVD) and wavelet difference reduction (WDR) was presented by Kumar, R. & *al.* [77]. It starts by using SVD for initial compression by transforming the ECG signal into a two-dimensional (2D) image, which reduces the data while maintaining good reconstruction quality. Next, WDR or Adaptive Scanning Wavelet Difference Reduction

(ASWDR) is applied for final compression. The method demonstrates high compression with low distortion of the reconstructed signal.

In 2016, Wang, X. & *al.* [98], presented a novel ECG compression technique that combines the Empirical Mode Decomposition (EMD) and Wavelet Transform. EMD decomposes the ECG signal into a set of intrinsic mode functions, abbreviated as IMFs. The approach proposes recomposing these functions to two distinct groups and then compressing each of them independently to optimize the characteristics of the data. The initial group is fully described through its extrema with low reconstruction error. Meanwhile, the second group undergoes further decomposition using the wavelet transform. This method performs competitively with existing ECG compression techniques by appropriately selecting the threshold and performing run-length coding and Huffman coding.

In the same year, an innovative ECG data compression technique using a combination of vector quantization and singular value decomposition (SVD) was proposed by Soussi, I., & Ouslim, M. [99]. The process begins by applying SVD to the 2D depiction of ECG signal. Next, using vector quantization, a finite set of weighted right singular vectors are quantized. Additionally, residual coding on the basis of the residual error's SVD is employed, which minimizes the reconstruction error. A new codebook generation scheme for vector quantization is also introduced. This scheme, based on the Lindth-Buzo-Gray algorithm, generates a new codebook at each vector quantization stage based solely on the singular vectors selected.

Additionally, during this time, Fathi, A., & Faraji-kheirabadi, F. [100], suggested an alternative compressing ECG signals method that combines the Wavelet Packet Transform (WPT) with a subband-specific quantization algorithm. Initially, WPT is applied to the ECG signal, followed by selection of the most significant sub-bands based on their Shannon entropy. Next, a content-based quantization and denoising technique is applied to the coefficients of the selected sub-bands. Finally, arithmetic coding is used to compress the data.

Also in 2016, A method using principal component analysis (PCA) and discrete cosine transform (DCT) for ECG signal compression is described by Benseguenі, S., & Bennia, A. [101]. First, PCA is applied to ECG data to obtain a low-dimensional representation by identifying orthogonal linear combinations with maximum variances. Then, variable DCT thresholds are applied to the different principal directions of the data to increase the compression ratio while preserving the signal morphology.

By 2017, Zhang, B. & *al.* [**102**] proposed a new strategy for ECG data compression combining short-time Fourier transform, neural network model optimized with multiple objectives and wavelet transform. First, the short-time Fourier transform is applied to reduce the real-time computation time. Next, the wavelet transform is used for lossless compression, allowing noise to be removed without losing data. The multi-objective optimized neural network model is then applied to the signal to extract the essential details.

The year 2021 saw new contribution. Jha, C. K., & Kolekar, M. H. [**103**] proposed an ECG data compression scheme which uses sifting function based empirical mode decomposition (EMD) and discrete wavelet transform (DWT). First, sifting function -based EMD is used to obtain the first intrinsic mode function (IMF). Then, the first IMF and four significant sifting function are combined together to remove many irrelevant components from the signal. The DWT with the wavelet mother 'bior4.4' is applied to this combination. The transform coefficients obtained after DWT are subjected to dead-zone quantization, which eliminates small transform coefficients around zero. Next, integer conversion of the coefficients and run-length coding are used to obtain a compressed form of the ECG data.

Moreover, in 2022, Hassan, A. M. A. & *al.* [**104**], suggested an effective compression method based on the compressive sensing technique for obtaining as well as recovering the ECG signals characterized by sparsity. This approach allows signal reconstruction even at sampling rates below the Nyquist frequency. An innovative analysis is proposed, where the initial step involves generating a sparse signal by applying a modified discrete cosine transform to the specified ECG signal. This transform is considered as a major contribution of this work and paves the way for further research in the field. A modest amount of wavelet elements is used to depict ECG signal, which is a novel approach to obtaining sparse ECG signals. A detection method for compressing ECG signals is also proposed, where ECG signals are randomly introduced between each successive heartbeat.

In 2022 as well, Verma, A. R. & *al.* [**105**] developed a new method for ECG signal compression based on Wavelet Packet Transform (WPT), exploiting Multi-Resolution Analysis (MRA). As a key tool for assessing cardiac health, the ECG signal may require prolonged storage for accurate analysis, which poses challenges in terms of storage space. The proposed compression technique overcomes this constraint. Using WPT, which generates a large amount of sparse data from the ECG signal, this method achieves significant compression by computing 2n sub-bands compared to the (n+1) sub-bands of the discrete wavelet transform.

Finally, in 2023, an algorithm for compressing ECG signals based on the discrete wavelet transform (DWT) and various optimization techniques inspired by nature is proposed by Singhai, P. & *al.* [106]. The algorithm presented uses optimization methods to determine optimal wavelet parameters and threshold levels. DWT is used to decompose the signal into sub-bands, the coefficients of which are then obtained. The threshold values for each sub-band are selected using optimization algorithms, and the coefficients are then further compressed using modified RLE encoding (MRLE). The results of this method are promising, effectively preserving the characteristics of the original signal after reconstruction.

Within the same year, Mohebbian, M. R., & Wahid, K. A. [107] presented an innovative semi-lossless compression technique designed for the monitoring and visualization of ECG signals. The method combines B-spline interpolation with optimization using ant colonies. By using the coefficients of the signal's B-spline as compressed data, it becomes possible to visualize the signal without the need for decompression, which can lead to time savings compared with conventional techniques.

I.3.3. State of the Art in ECG signal filtering approaches

ECG signal filtering encompasses a set of techniques designed to enhance the quality of electrocardiographic recordings by eliminating or mitigating unwanted interference, thereby achieving a more accurate representation of cardiac electrical activity. The process of denoising is exceptionally intricate due to the non-stationary nature of ECG signal. Additionally, there is a risk of information loss during the filtering process. Therefore, when applying any biomedical signal filtering, it is crucial to exercise caution to avoid altering the desired details. Numerous techniques have been proposed in the literature to contribute to the overall enhancement of cardiac data quality. Every denoising method aims to minimize information loss during effectively removing noise. In this context, several techniques have been proposed to deal with this problem:

In 2000, Novak, Daniel, & *al.* [108], have proposed an ECG signal denoising method using an adaptive wavelet technique. This involves detecting noise levels, and then applying a soft thresholding denoising method to individual signal sections based on varying noise levels, and utilizing different decomposition levels and wavelet families accordingly. The results are compared with classic filters, emphasizing improvements in signal-to-noise ratio (SNR) through the discussed wavelet denoising approach.

In 2001, a method based on utilization of successive decompositions, matrix manipulation, and specific transforms like Discrete cosine transform (DCT) and Wiener filtering in a translation-invariant wavelet domain, was introduced by Nikolaev, N. & *al.* [109] to effectively suppress EMG interference in ECG recordings, resulting in improved signal quality for diagnostic purposes.

During the year 2002, Ziarani, A. K., & Konrad, A. [110], presented a novel approach to eliminate power line noise from ECG signals, utilizing a recently developed signal processing algorithm designed for extracting and tracking specific signal components over time. The proposed method demonstrates superior performance in effectively eliminating noise, especially in scenarios with varying power line interference frequencies. Its simplicity and robust structure make it well-suited for practical constraints, including limited computational resources and low sampling frequencies, as evidenced by computer simulations.

The following year, in 2003, Esposito, A., & D'Andria, P. [111] suggested a new method that addresses the challenge of power line interference and electrical noise altering the morphology of ECG, a common issue in traditional filtering techniques that can distort the original waveform and complicate pathology interpretation. The proposed adaptive neural method for filtering ECGs effectively filters ECGs without causing information loss and outperforms a recent morphological filtering technique, as evidenced by tests on 110 segments from the European ST-T database.

Subsequently in 2004, Ercelebi, E. [112], proposed an ECG denoising technique based on the utilization of second-generation wavelets and a level-dependent threshold estimator, constructed through a lifting scheme. Unlike classical first-generation wavelets, which rely on translation and dilation of a fixed function and require a Fourier transform, the lifting scheme operates entirely in spatial domain. This method allows for a faster implementation of the wavelet transform (WT) with in-place calculation, eliminating the need for extra memory. The proposed approach demonstrates enhanced denoising capabilities compared to classical methods, making it a promising technique for signal processing applications.

Progressing to 2005, where a denoising algorithm for real ECG signals that addresses distortion issues in classical wavelet denoising, specifically in the R waves at the 4th level, by estimating and removing the corrupted white Gaussian noise (WGN) interfering with R waves was given by Chouakri, S. A. & *al.* [113]. The proposed denoising algorithm outperforms traditional methods, achieving a higher signal-to-noise ratio (SNR) of about 6 dB and a lower mean square error (MSE) of approximately 0.0011, as demonstrated in a comparison using the MIT-BIH arrhythmia database.

Also in 2005, a novel nonlinear filter designed to effectively reduce baseline wander and power interference noise in ECG signals, emphasizing the limitations of linear filters, such as signal distortion and long impulse response has been introduced by Łęski, J. M., & Henzel, N. [**114**].

Moving forward to 2006, H.E, Taigang & *al.* [**115**], have explored in thier stydy the use of independent component analysis (ICA) to remove noise and artefacts from routinely recorded electrocardiograms (ECGs). Testing the ICA algorithm using human participants' three-channel ECG recordings particularly in the cardiac care unit, the findings demonstrate its efficacy in detecting and removing various sources of noise and artefacts. The paper introduces a novel method based on basic statistical criteria to determine the order of independent components, successfully addressing a challenge in ICA application to ECG data as well as presenting the possibility of processing ECGs online using this technique.

As we advance to 2007, O. Sayadi and M.B. Shamsollahi [**116**], developed a multi-adaptive bionic wavelet transformation (MBWT) strategy to suppress AWGN and BW noise. This strategy is based on optimizing bionic wavelet parameters to find a new threshold value. Therefore, this technique provides good results compared to the traditional method based on the wavelet transform.

Also, en 2007, Sameni R. & *al.* [**117**], presented a nonlinear Bayesian filtering framework for noisy ECG recordings, utilizing a modified nonlinear dynamic model and introducing an automatic parameter selection method for adaptability to diverse ECGs. This proposed method demonstrates superior outcomes than traditional filtering methods, like wavelet denoising, adaptive filtering, and bandpass filtering, across various ECG signal-to-noise ratios. It is also effective in handling real nonstationary muscle artifact, showcasing its potential as an efficient model-based filtering framework for noisy ECG recordings.

In 2008, a novel ECG enhancement method based on empirical mode decomposition (EMD), specifically designed for both high-frequency noise and baseline wander (BW) removal was introduced by Blanco-Velasco, M. & *al.* [**118**]. Unlike previous approaches, the method doesn't rely on simple partial summation of intrinsic mode functions (IMFs), but rather selects and processes different IMFs to successfully address denoising and BW removal.

Shifting to 2009, Kaur, M., & Singh, B. [119], proposed a combination technique utilizing moving averages and IIR notch methods to mitigate power line interference. Results demonstrate a clear reduction in power line noise in the ECG signal. The proposed filter, with fewer coefficients, ensures faster computation, making it suitable for real-time processing.

Entering 2010, Al-Qawasmi, A. R., & Daqrouq, K. [120], introduced a novel approach using Wavelet Transform to filter ECG signal. The goal is to adapt discrete wavelet transform for ECG signal enhancement, showcasing superior results compared to conventional methods. This method outperforms the FIR filter and discrete wavelet thresholding coefficients established by Donoho in terms of performance.

In the year 2011, El-Dahshan [121] developed an effective technique for denoising the ECG data using the related genetic algorithm (GA) to remove non-stationary noise associated with wavelet transformation. This method consists of selecting the optimal parameters for denoising wavelets in order to maximize filtering performance.

Transitioning to 2012, Smital. L. & *al.* [122] proposed another procedure using adaptive wavelet-Wiener filtering. This technique is characterized by the application of the stationary wavelet transformation (SWT) in the Wiener filtering domain to eliminate the noise of EMG corroding ECG signals.

In 2012 as well, Dr. A. Kabir and C. Shahnaz [123] presented a new approach to ECG denoising based on a noise reduction algorithm in the domains of empirical decomposition (EMD) and discrete wavelet transformation (DWT). The authors performed windowing within the EMD domain to reduce the first IMFs' noise instead of eliminating them entirely, keeping QRS complex intact and producing an ECG signal that is comparatively clean. The resulting signal is then altered within the DWT domain, where a noise reduction algorithm based on adaptive soft thresholding is used taking into account the advantageous properties of DWT over EMD to preserve energy in a noisy environment and to reconstruct the initial ECG signal with enhanced time resolution.

Staying within the year 2012, another method has also been proposed by Chandrakar, C and Kowar, M. K. [124], employing the Recursive Least Squares (RLS) algorithm as an adaptive filtering technique to reduce noise in ECG and artifacts removal. This approach preserves the ECG's low-frequency components and subtle features. The algorithm employs least-squares techniques to minimize squared differences between the desired signal and the model filter output. Through recursive computation with new incoming signal samples at each iteration, it establishes the recursive least-squares (RLS) algorithms.

In 2013, Y. Yang and Y. Wei [125] suggested another filtering technique, based on random interpolation mean (RIA) using several wavelet bases, which was applied to reduce three types of noise like AWGN, MA and EM from ECG signals. It initially involves the Lagrange interpolation of the third order. Then, the conventional WT denoising is applied to each interpolated signal. Then, each denoted signal is reconstructed by inverse interpolation. As a result, the final denoted signal is obtained by averaging the reconstructed signals.

In the same timeframe, S. Ari & *al.* [126], proposed a filtering technique that applied to reduce unwanted noise components in the time-frequency domain using the Stockwell transform (S-Transform). The proposed method exhibits better and positive results in contrast to earlier published techniques such as wavelet transform with subband-dependent threshold (WT-Subband) and wavelet transform with soft thresholding (WT-Soft), confirming its superiority in preserving biological information contained in the upgraded ECG signal.

In addition, in 2014, M. Talbi [127] introduced a novel ECG signal denoising technique utilizing thresholding of the acquired coefficients through the implementation of forward wavelet transform translation invariant to each Bionic Wavelet coefficient. The free noise ECG signal is then reconstructed by employing the inverse of the Bionic Wavelet Transform (IBWT) to the de-noised Bionic Wavelet coefficients.

Within the same year, various filter types to enhance the noise reduction of ECG signals, namely, Median filter, FIR filter, Gaussian filter, Butterworth filter, were proposed by Subbiah, S. & *al.* [128].

Moving forward to 2015, where a denoising technique for ECG signals employing the wavelet packet transform in conjunction employing two adaptive filters namely recursive-least-square (RLS) and normalized least-mean-square (NLMS) was presented by Biswas, U and Maniruzzaman, M. [129]. The outcomes contrasted against a conventional notch filter, examining both the frequency and temporal domains.

In 2015 as well, Salih, S. K. & *al.* [130], introduced a novel approach for denoising ECG utilizing a multi-iteration moving average filter. The algorithm involves two key steps: first, estimating the ECG signal noise level, and second, removing the identified noise.

Advancing to 2016, Wang, Z. & *al.* [131], have introduced an ECG signal denoising algorithm based on Adaptive Fourier Decomposition (AFD), utilizing a stop criterion derived from the estimated Signal-to-Noise Ratio (SNR) of the noisy signal. By removing noise components from the energy domain, the method effectively separates signal and noise with overlapping frequency ranges but distinct energy distributions. The stop criterion aims to maximize the SNR of the reconstructed signal.

In the same year, Nguyen, P., and Kim, J. M. [132], proposed an adaptive ECG denoising methodology utilizing a genetic algorithm (GA)-based thresholding method with ensemble empirical mode decomposition (EEMD). The denoising methodology involves two critical steps. Firstly, an adaptive IMF selection method is applied, using Kullback-Leibler divergence and probability density functions to isolate signal-dominant IMFs from those heavily affected by noise. Secondly, a GA-based IMF thresholding technique is introduced to effectively reduce noise in the contaminated IMFs while preserving crucial information. The optimization of threshold parameters for each noisy IMF aims at achieving optimal denoising results in terms of signal-to-noise ratio improvement (SNRimp). The final denoised ECG signal is obtained through a reconstruction process that combines signal-dominant IMFs with the denoised IMFs.

Additionally, in the same timeframe, Panigrahy, D. and Sahu, P. K. [133], suggested an advanced noise cancellation approach ECG signals based on extended Kalman smoother (EKS) in conjunction with differential evolution (DE) method. DE is used to automatically select and optimize ten ECG signal's components. These optimized components are utilized in the reconstruction process to closely match an actual ECG signal. Then, these optimized parameters are incorporated into an extended Kalman smoother (EKS) framework, utilizing them for developing the mathematical representation of the ECG signal's evolution (state equation) and initializing various EKS parameters. The EKS framework is then applied to denoise the ECG signal, focusing on a single channel of ECG data.

Also in 2016, an effective method for denoising ECG signals, combining the adaptive dual threshold filter (ADTF) and the discrete wavelet transform (DWT) was introduces by Jenkal, W. & *al.* [134]. The proposed algorithm involves three denoising steps: DWT decomposition, the ADTF step, and a correction step for the highest peaks.

In 2017, Jain, S. & *al.* [135], proposed a novel method for ECG denoising. This method integrates the empirical mode decomposition algorithm with Riegmann Liouvelle (RL) fractional integral filtering and Savitzky-Golay (SG) filtering. The approach involves decomposing the noisy

ECG signal into intrinsic mode functions (IMFs), identifying noisy IMFs corrupted with high-frequency (HF) and low-frequency (LF) noises using proposed identification methodologies. Denoising is achieved by applying RL fractional integral filtering to noisy IMFs with HF noises and SG filtering to those with LF noises. The noise-free ECG signal is then reconstructed.

Remaining in the year of 2017, a novel method to eliminate baseline wander noise and powerline interference from ECG signals utilizing the Empirical Wavelet Transform (EWT) was given by Singh, O. and Sunkaria, R. K. [**136**]. EWT, a relatively recent method under active research, offers an innovative solution to traditional filtering challenges. The proposed approach involves decomposing the ECG signal contaminated with powerline interference into two modes using EWT with optimized boundaries. The last mode is then processed and subtracted from the noisy ECG to produce a noise-free ECG signal. For baseline wander correction, a similar EWT decomposition with different boundaries is applied to the noisy ECG. The first mode obtained helps estimate baseline drift, subsequently eliminated, resulting in an ECG signal free of baseline wander.

Advancing to 2018, Manas and Susmita [**137**] introduced an effective ECG denoising technique using Empirical Mode Decomposition (EMD) and Adaptive Switching Mean Filter (ASMF). The benefits of EMD and ASMF techniques are leveraged to reduce noise in ECG signals with minimal distortion. This technique is based on denoising using soft wavelet thresholding applied to the three initial Intrinsic Mode Functions (IMFs). This approach effectively reduces noise components in the QRS regions and enhances QRS complexes. Subsequently, an ASMF technique is employed to decrease the impact of noise in the low-frequency region of ECG signals, further improving signal quality. Due to the ASMF operation, the peaks of ECG signals are slightly attenuated. Therefore, a peak correction process assisted by positional information is used to restore them.

Similarly, in 2018, Ahmad & *al.* [**138**] conducted a study focused on the use of Genetic Algorithm (GA) with Wavelet Transform (WT) for ECG signal denoising. In this study, WT parameters are employed as inputs for GA to denoise the input signal corrupted by Gaussian white noise. The noisy signal undergoes a decomposition process to extract approximation and detail coefficients, and then the detail coefficients are thresholded using a threshold value to eliminate noise. Finally, the signal is reconstructed using the denoised approximation and detail coefficients.

In the same year, Bahaz and Benzid [**139**] conducted a study on the use of the well-known Discrete Fourier Series (DFS) to reduce the contribution of baseline wander (BW) and powerline interference (PLI) noise in ECG recordings.

Additionally, in 2018, an innovative approach to denoise ECG signals using variational mode decomposition (VMD) while incorporating the estimation of non-local means (NLM) in addition to discrete wavelet transform (DWT) filtering techniques introduced by Singh, P. and Pradhan, G. [**140**]. In this method, the noisy ECG signal undergoes decomposition to narrow-band variational mode functions (VMFs) utilizing the VMD approach and categorizing them into lower- and higher-frequency signal. Then a thresholding with DWT technique is applied to remove interference from VMFs with higher frequencies. The VMFs with lower frequencies undergo denoising utilizing NLM estimation technique. Finally, the denoised high-frequency and low-frequency VMFs are used for signal reconstruction.

During the same period, a new ECG denoising technique combining the efficiency of empirical mode decomposition (EMD) with non-local mean (NLM) was proposed by Kumar, S. [**141**]. The proposed methodology employs EMD to minimize noise in ECG signal, and EMD output undergoes NLM processing to preserve edges and remove noise introduced during the EMD process. In addition, for this method, the standard deviation is calculated to gather information about the incoming noise, allowing adjustments to be made within NLM and EMD, and guiding the noise removal process while optimizing the performance of the combined EMD-NLM method.

Progressing to 2019, Dengyong & *al.* [**142**] presented a method for ECG noise removal using a subband smoothing filter and the wavelet energy. In contrast to the traditional wavelet thresholding denoising method, wavelet coefficients requiring threshold noise reduction are selected based on wavelet energy, while the other wavelet coefficients remain unchanged. Furthermore, the adopted subband smoothing filter further denoises the ECG signal, enhancing the overall signal quality.

Staying within the same year, Wang, Z. & *al.* [**143**], introduced a modified wavelet design approach. Through the optimization of filter coefficients to closely approximate the ideal filter, a nearly symmetric orthogonal wavelet is derived and implemented for denoising ECG signals.

Also in 2019, Asma, T. & *al.* [144], suggested a straightforward and effective technique for eliminating noise from ECG signals which relies on common digital signal processing tools such as the moving average filter, median filter, baseline drift removal, and peak detection. The proposed algorithm exhibits superior performance compared to conventional median and moving average filters, positioning it as a compelling alternative to standard wavelet-based methods.

During the year 2020, numerous research studies were conducted to optimize and apply a new approach, namely, Ouali, M. A. & *al.* [145], proposed a new method for ECG signal processing based on the use of a type-2 fuzzy system and TLBO optimization. The proposed filter is described as a two-layer feedback system, where the first layer utilizes a type-2 fuzzy autoregressive filter model, and the second layer adjusts the membership function parameters using a Teaching Learning-Based Optimization algorithm (TLBO) to achieve the desired signal reconstruction by minimizing the criterion function. This proposed filter yields superior outcomes compared to other model-based and non-model-based methods detailed in the literature, while preserving the morphology of the ECG signal and maintaining diagnostic performance.

Additionally, in 2020, Zhang, D. & *al.* [146], developed an ECG denoising method by combining empirical mode decomposition (EMD), sample entropy, and an improved threshold function. The primary aim of this study is to integrate empirical mode decomposition (EMD) and sample entropy to identify and denoise the intrinsic mode functions (IMFs), avoiding the common practice of directly discarding the first IMF in traditional EMD denoising, which can lead to loss of crucial information and incomplete signal denoising. This method can better reduce the noise of ECG signals while also retaining more details of the original signal.

Additionally in 2020, Houamed, I. & *al.* [147] have introduced an efficient method for noise removal in ECG signals based on fractional wavelet decomposition combined with thresholding techniques. Instead of employing conventional low-pass and high-pass filters in the wavelet transform, this approach incorporates fractional-order filters. Fractional wavelets are introduced, simulated, and compared to other wavelets to assess their efficacy in ECG denoising. The denoising process is operationalized through a careful selection of wavelet transform coefficient thresholding and the signal's wavelet decomposition level.

Within the same year, Zhang, M. & *al.* [148] proposed an Integrated Empirical Mode Decomposition Adaptive Threshold Denoising method (IEMD-ATD) for ECG signal processing. The IEMD-ATD method comprises three key components. Firstly, an integrated EMD approach based on the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise

(CEEMDAN) framework is introduced to enhance the decomposition quality. Secondly, a novel grouping method for Intrinsic Mode Functions (IMFs) is developed to determine boundaries among IMFs dominated by high-frequency noise, those dominated by useful information, and IMFs with low-frequency noises. Finally, an adaptive threshold denoising method is formulated and employed to denoise IMFs predominantly affected by high-frequency noise.

The year 2021 was marked by new contributions in the field of ECG filtering. Among these contributions, we mention: Talbi, M. [**149**], which proposed a novel technique utilizing total variation minimization (TVM) and lifting wavelet transform (LWT) it involves utilizing LWT to extract three wavelet sub-bands (cD1, cD2, and cA2) from the noisy ECG data, then denoise the detail coefficients (cD1 and cD2) with soft thresholding and the approximation coefficient (cA2) using TVM, and finally, applying the inverse of LWT to reconstruct the denoised ECG signal.

We mention also, A. K. Dwivedi & *al.* [**150**], that they proposed a novel method for eliminating baseline interference from ECG signal through the combination of empirical mode decomposition (EMD) and stationary wavelet transform (SWT), as well as ensemble empirical mode decomposition (EEMD) and SWT. After decomposing the ECG signals to diverse intrinsic mode functions, SWT is employed to further minimize noise.

In 2021 as well, a novel method for the elimination of baseline wander (BW) and power line interference (PLI) from ECG signal, utilizing a hybrid method combining empirical wavelet transform (EWT) and empirical mode decomposition (EMD) was introduced by Boda, S. & *al.* [**151**]. The method comprises four distinct stages: PLI and BW estimates utilizing EWT, sub-signals of high and low frequencies constructed using the zero-crossing number ratio (RZCN), interference subtraction, and EMD-based ECG signal decomposition. Both qualitative and quantitative analyses indicate the superior performance of the proposed method over existing approaches.

Furthermore, in the same timeframe, Mohguen, W. and Bouguezel, S. [**152**] presented a novel method for denoising ECG signals based on Ensemble Empirical Mode Decomposition (EEMD) with the incorporation of a modified customized thresholding function. The method involves decomposing the noisy ECG signal into a series of Intrinsic Mode Functions (IMFs) using the EEMD algorithm. Additionally, a modified customized thresholding function is applied to reduce noise from the ECG signal while preserving the QRS complexes. The denoised signal is then reconstructed using all thresholded IMFs.

This trend of progress has continued in 2022, where a swarm-based intelligence approach for ECG denoising was proposed by Subbiah, B. [153]. It involves decomposition using Empirical Wavelet Transform (EWT), with an optimized window function determined by the Honey Badger Optimization (HBO) algorithm. Subsequently, the Adaptive Hybrid Filter is employed, and its weight parameters are fine-tuned using the HBO technique. The final step yields denoised ECG signals through this comprehensive approach. The proposed methodology contributes significantly to ECG signal denoising.

Moreover, in 2022, Hussein, A. & *al.* [154], introduced an empirical mode decomposition-based adaptive method for removing noise from ECG signals. Unlike conventional EMD-based approaches that may discard fundamental functions or utilize window-based methods for reducing high-frequency noise, the proposed method minimizes noise in ECG signals with minimal distortion.

In 2022 as well, Talbi, M. [155] developed a novel ECG denoising technique, this method applies the Wavelet Total-Variation (WATV) denoising method in the domain of the Stationary Bionic Wavelet Transform (SBWT). The technique significantly reduces noise while practically preserving the diverse waves PQRS-T of the original signal.

Similarly, in 2022, Das, M. and Sahana, B. C. [156] developed a method for ECG signals denoising enhancement performances by introducing an orthogonal wavelet-based filtering approach. The key principle involves utilizing an optimized filter bank to address limitations observed in standard wavelet-based filtering. The suggested wavelet filter's coefficients are adjusted through the use of a population-based optimization technique inspired by nature, specifically the cuckoo search optimization algorithm. Comparative analysis with other filtering techniques, such as empirical mode decomposition, empirical wavelet transforms, as well as Butterworth high-pass filter, demonstrates satisfactory performance in noise removal from ECG signals, even under low input SNR conditions.

Finally, in 2023, Malik, S. A. & *al.* [157], introduced a hybrid-based strategy, integrating discrete fast lifting-based wavelet transform (LWT) with an adaptive iterative filtering (IF) technique to effectively remove noise from ECG signals. Because of its proper mathematics foundation and convergence assured beforehand, the iterative adaptive filtering strategy was selected in preference to empirical mode decomposition (EMD), the generated IF's noisier modes are processed by a LWT framework. This disintegrates them into approximation and detail coefficients, which are then scaled to produce a noise-free signal using a threshold approach.

In addition, Yue, Y. & *al.* [158] presented a denoising algorithm for ECG signals, employing a combination of Ensemble Empirical Mode Decomposition (EEMD), Empirical Mode Decomposition (EMD), and wavelet packet (WP). The ECG signal undergoes a triple decomposition process, utilizing EEMD, EMD, and WP successively. The noise free ECG signal is attained by fusing the processed signal components, effectively eliminating baseline wander, powerline interference, and muscle artifacts while preserving signal details.

Also in 2023, a digital filtering technique was employed to enhance the electrocardiogram (ECG) signal by Prasad, V. [159]. Specifically, denoising the ECG signal was achieved utilizing discrete wavelet transforms (DWT) as well as low pass filter (LPF) methods. These filtering algorithms were tested for de-noising of the ECG signal in BLW noise, MA noise and EMG noise. Also, different smoothing techniques such as moving mean method, linear regression method and savitzky-golay smoothing techniques are applied for ECG signal processing.

I.4. Contribution

Part 1

Modeling of time series and dynamical systems is of paramount importance in many areas such as finance, meteorology, medicine, engineering and many others. These systems often exhibit complex behaviors and temporal dependencies, making their modeling and prediction particularly difficult. As part of our research work in this PhD thesis, we’ve put forth a novel structure for modeling dynamical systems and time series based on the artificial neural network autoregressive moving average (ANNARMA) and Metaheuristics algorithms to fine-tuning the parameters of this proposed neural model. This framework encompasses a parallel arrangement of two interconnected sub-models. The foremost sub-model is the primary model, characterized by standard specifications and a lower resolution, designed for the dynamical system or time series being examined. In order to address the resolution limitation and achieve heightened precision, a second sub-model, named the error model, is introduced. This error model captures the disparities between the primary model and the nonlinear dynamical system or time series data.

Part 2

The significance of the ECG lies in its ability to identify heart irregularities. One crucial job for biomedical engineering and science is to reduce noise in the ECG signal with the objective of providing accurate diagnosis, enhancing the quality of research data and contribute in the creation of innovative therapies within the domains of cardiology and cardiovascular health. The denoising process presents considerable complexity due to the ECG's non-stationarity and, in certain situations, the possible information loss during filtering. Hence, numerous methods have been suggested in the literature to tackle this problem. Every noise reduction method strives to minimize the loss of information while ensuring satisfactory noise elimination. The choice of each method is motivated by its distinct ability to address the specific characteristics of ECG signals, such as their non-linearity, non-stationarity and susceptibility to noise. For the second part, we explored various temporal approaches for preprocessing ECG signals with the aim of improving their quality and accurately reducing noise while preserving essential information. Our choice of methods was guided by the need to take into account the complex variations over time observed in ECG signals which are frequently not effectively captured by traditional or frequency-based pretreatment techniques.

Our initial contribution is based on the application of the Empirical Mode Decomposition (EMD) method which is an adaptive and local time-frequency analysis method. It is based on a data-driven approach and is particularly suited to the analysis of non-linear and non-stationary signals. Using this method, our aim is to decompose the ECG signal into its intrinsic components, which facilitates an in-depth analysis of the various aspects of the signal, including its low and high frequency components as well as its non-stationary characteristics. Furthermore, the idea of utilizing a method based on average wavelet coefficients alongside the Hurst exponent concept allows one to distinguish between high-frequency and low-frequency IMF, making it easier to eliminate unnecessary components from the signal.

As a second contribution, we also examined the utilization of wavelet transform method with the Hurst exponent. The wavelet provides a multi-scale representation of the signal, which is beneficial for the analysis of ECG signals and the Hurst exponent can be used for determining the ECG signal components that may be impacted by noise. By combining the advantages of the two approaches our aim is facilitating the detection and elimination of unwanted artefacts.

The third contribution lies in the application of a significant variant of the EMD family, referred to as graph EMD (GEMD). Graph representation offers a novel approach to the decomposition of complex signals like ECGs by exploiting the structural relationships and connections between the various signal components. This provides a robust structure that facilitates the separation of signal components even in the presence of perturbations, resulting in more reliable and resilient analysis and decomposition results.

I.5. Conclusion

In this chapter, we briefly reviewed the general concept of dynamical systems and time series. As an application of time series, we specifically focus on ECG data, which are non-stationary time series. We explored the anatomy of the heart, its electrical conduction system, the principle of ECG generation, along with potential sources of contamination noise. Subsequently, a state of the art of modeling and identifications techniques of dynamical systems and time series, as well as filtering methods for ECG signals were presented at the end of this chapter.

CHAPTER II

APPLIED METHODS

Neural networks alongside their ability to model intricate structures, and metaheuristic algorithms, offer the best possible solutions for challenging issues. Synergistically combining the potent learning capabilities of artificial neural networks with the adaptability and exploration potential of metaheuristic algorithms provides a symbiotic solution that surpasses the specific limitations of each approach. Together, these approaches create a strategic alliance that enables effective and efficient task resolution. Strategies based on these cutting-edge methods are applied in our research work for identifying and modeling dynamical systems and time series. In addition, in a different setting, we explore specific decomposition and filtering techniques designed for Processing and analyzing temporal series, specifically filtering the ECG signal. This second chapter offers a thorough overview of the key ideas related to neural networks, metaheuristic algorithms, and signal decomposition and filtering techniques, illuminating the path to a thorough understanding of how these methodologies are applied in the context of our research.

Contents

II.1. Introduction

Artificial neural networks, or simply abbreviated ANN, are computational system that intent to imitate the functionality of biological neural networks, introduced as a concept initially in 1940 [160]. ANN has evolved over the years to stand out as the most effective and versatile techniques in the field of automatic learning. The primary goal of ANN is to provide a model that is able to comprehend and represent complex relationships between input and output data sets. In contrast to traditional mathematical models based on equations, neural network models are particularly effective in handling extremely complex data sets that would typically be difficult to predict using traditional mathematical modeling.

Often, to adjust the parameters of ANNs and enhance their performance, methods such as metaheuristic algorithms are employed, and this relates to the ability of metaheuristics to effectively explore complex and high-dimensional solution spaces. Metaheuristics can be defined as an optimization algorithm emerged first in the 1950s [161], which later diversified into numerous variants to solve tricky optimization issues throughout a range of fields by borrowing principles from biological and biogeographic mechanisms. Its objective is to become closer to a global optimum by functioning as a search algorithm that discovers the characteristics of the issue and uses that knowledge to approximate the ideal solution.

In parallel, in the context of time series analysis, the utilization of dedicated decomposition and filtering approaches becomes vital for extracting and isolating intricate details or distinctive features from medical data, specifically ECG signals. These methods contribute to the overall interpretability of ECG signals, thereby bolstering the accuracy of diagnostic insights and making informed clinical decisions possible.

This chapter describes the methods that will be used in our study, whether for modeling and identifying time series and dynamical systems, or for filtering ECG signals.

II.2. Artificial Nouerons Networks concepts

II.2.1. Introduction to ANNs

Artificial Neural Network or ANN, by abbreviating it, as a branch of artificial intelligence (AI) modeled after the structure and operations of the human brain, are proving to be effective tools for handling non-structured and complex challenges. Dating back to 1940, the intention was to generate a set of elementary neurons, referred to as "artificial neurons," arranged into clusters and connected to transfer signals from input to output. These neurons are typically simple mathematical functions, but their network organization allows for complex tasks like image classification, chronology prediction and many others. The first ANN models were proposed by Warren McCulloch and Walter Pitts in 1943 [162], and the Hebbian learning rule was later introduced by Donald Hebb in 1949 [163]. Since 1980, ANNs have developed at an exponential rate, making significant strides in supervised and unsupervised learning, deep learning, and a variety of other fields, including robotics, computer vision, and natural language processing.

II.2.1.1. Biological neuron

A biological neuron is a sophisticated cell that processes chemical and electrical signals and is essential to the nervous system's functioning. Neural networks in the human brain are composed of countless interconnected neurons. Each neuron functions as an independent unit. It is composed of a soma (or cell body), which integrates the incoming data and serves as a control center. After processing this data, the neuron uses its axon to send the result, which takes the form of electrical signals, from the cell body to the entries of other neurons. The axons, which are in charge of connecting the neurons, have a crucial role for the nervous system's overall function. Furthermore, neurons have multiple projections known as dendrites, which serve as the main receptors of the neuron and transmit information from the external environment to the cell body. On the other hand, information from other neurons is received by the synapses of a neuron through the axon, enabling communication between them [164, 165]. (Figure II.1) illustrates the basic structural layout of a biological neuron.

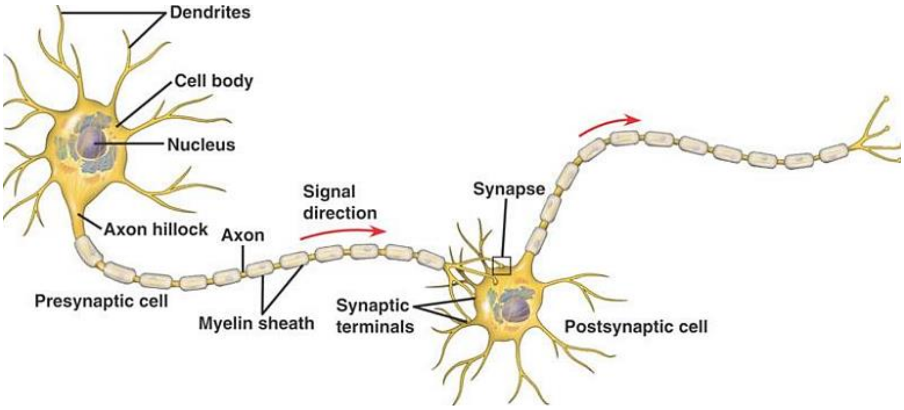

Figure II.1 – Biological neuron [166].

II.2.1.2. Artificial neuron

Artificial neuron incorporates two key properties of biological neurons: the capacity to transfer nerve impulses and the ability to react to outside factors and transform them into impulses. The process of modeling a formal neuron from a biological neuron established by corresponding dendrites in the biological neuron to the input signal of artificial neuron, and the axons matches the output signal. Furthermore, synapses stand in for the weights of the connection, while the bias stems from certain characteristics of biological neurons' operation, like the existence of activation thresholds that seeks to determining necessary excitation level and generates biological potential or electrical response. In this case, an artificial neuron can be described as a processing unit receiving data as an input vector and producing a real output. This output is influenced by connections inputs and weights [165]. The diagram depicting the general structure of an artificial neuron is illustrated as follow:

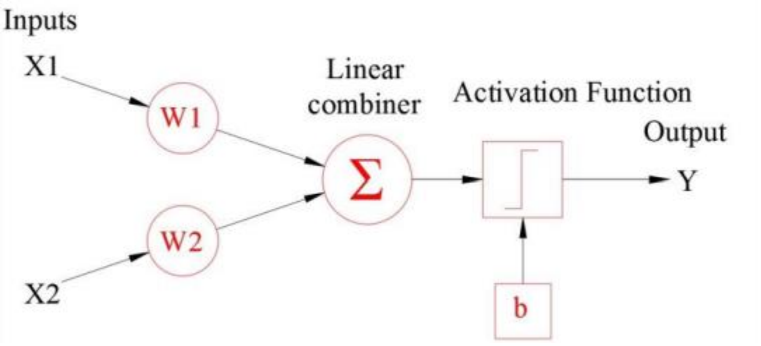

Figure II.2 – Artificial neuron [166].

Each neuron in a network can be implemented as shown in (Figure II.2) and its operation is divided into multiple stages. First of all, the data from the external environment, representing the input signals (Xi), are weighed through matching synaptic weights (wij). Due to these weights, the relevance of each external signal that a neuron receives can be quantified for its normal functioning. Then by combining linearly these signals an activation voltage is created and compared to a predefined activation threshold or bias (b) in order to figure out whether or not the neuron is activated. In the event that the voltage rises above this threshold, an activation potential (u) is produced by deducting the activation voltage from the linear assembler (∑), and then subjected it to an activation function (*f*) to restricting the neuron's output to a range of acceptable values. Ultimately, based on this activation potential the neuron generates an output signal (Y) that can be utilized as an input by other neurons within the network. Two expressions can be used to describe this process [165]:

- ✓ A linear combination of inputs

$$u = \sum_{i=1}^{n} Wij.Xi + b \tag{II.1}$$

- ✓ The output of the neuron

$$Y = f(u) \tag{II.2}$$

II.2.1.3. Activation functions

The activation function is a mathematical function whose primary role is to convert the result of the linear combination of entries of a neuron into an output value, it may also be referred to as a transfer function or threshold function. This function allows the network of neurons to capture complex relationships in the data thanks to the introduction of non-linearity in the behavior of the neuron. Neural networks use many activation functions, and here are some of the most common ones. [167, 168, 169]

- **Sigmoid**

The sigmoid is a continuous function whose output value is typically stated as a probability which means that if the input value is extremely large and positive, the function will convert it to a probability of 1. On the other hand, the function wil translate an input value to a probability of 0 if it is negative and very small. The following defines the sigmoid function :

$$f(x) = \frac{1}{1+e^{-x}} \tag{II.3}$$

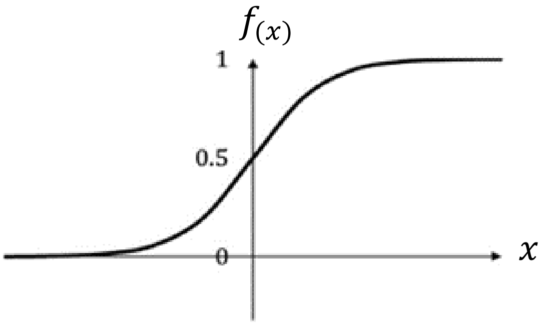

Figure II.3 – Sigmoid function.

- **Tanh (Hyperbolic Tangent)**

The function Tangent Hyperbolic (Tanh) is similar to the function Sigmoid but it produces a result between -1 and 1. Its ability to produce negative values as well as positive values offers better results than the Sigmoid function because of its symmetry. It is well suited to be used in multilayer perceptrons, especially for hidden layers.

$$f(x) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}} \tag{II.4}$$

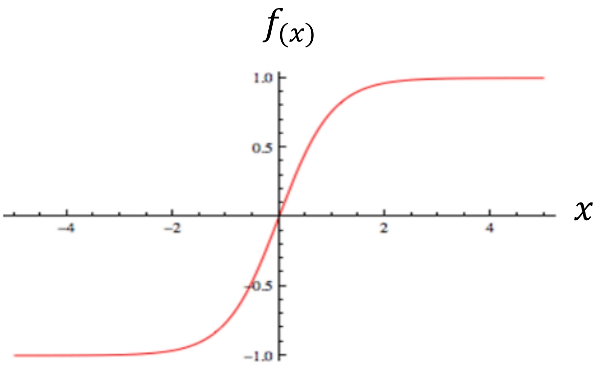

Figure II.4 – Tanh function.

- **ReLU (Rectified Linear Unit)**

Relu is a nonlinear real function that was created to solve the Sigmoid and Tanh functions' saturation problem, which is the circumstance in which the gradient of both functions stays almost zero after each learning period. This activation function does not saturate and greatly accelerates network convergence.

$$f(x) = \max(0, x) \tag{II.5}$$

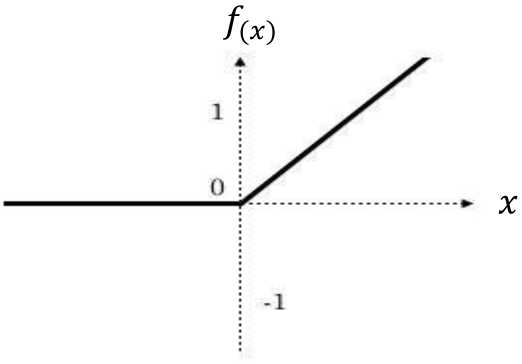

Figure II.5 – Relu function.

When the entry is negative, the output is zero, while if the entry is positive, the output is x.

- **Fonction Leaky Relu**

The Leaky Relu function tries to solve the problem of the Relu function when the input is negative by introducing a slight negative slope, usually fixed at 0.01 or higher, it solves the problem of inactivity of the Relu function for negative inputs. The Leaky Relu function is given by the formula:

$$f(x) = \max(0.01x, x) \tag{II.6}$$

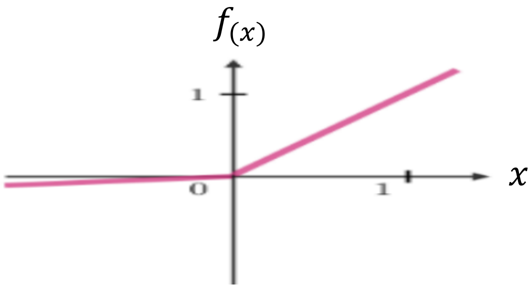

Figure II.6 – Leaky Relu function.

- **Softmax**

SoftMax is a function that handles multiple entries and assigns probability to each entry and act like the sigmoid by lowering the number of units from 0 to 1, simultaneously dividing each output so that the total sum of outputs is equal to 1. This feature allows faster learning and is often used in multi-classification networks classes. The softmax function can be interpreted by the following formula:

$$\sigma(Z_j) = \frac{e^{Z_j}}{\sum_{k=1}^{n} e^{Z_k}} \tag{II.7}$$

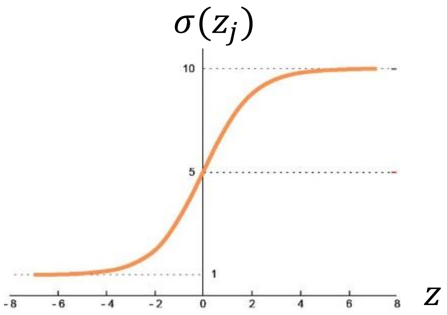

Figure II.7 – Softmax function.

where, Z is a vector representing the inputs for each unit in the output layer.

The choice of the activation function is often dictated by the specific requirements of the problem to be solved, the desired characteristics of the model and the empirical knowledge acquired.

II.2.2. Structure of ANNs

A neural network is a simultaneous mechanism of processing data, made up of three primary parts: the input layer, the hidden layer, and the output layer. The input receives external information and is standardized to boost numerical accuracy. While, most internal network processing is done by the intermediate layers, which extract patterns associated with the system or process under analysis. The output part is responsible for the final output.

II.2.2.1. Feedforward structure

A feedforward network, is a class of neuron network in which information flows from inputs to outputs without feedback or recurring connection loops (Figure II.8). It is classified as static networks and has two topologies: single-layer and multi-layered A single-layer neural network is the simplest form of a neuronal network, comprising a single set of neurons, where input data is transmitted directly to the output neurons without going through intermediate neurons. On the other hand, a multi-layer neural network, includes several layers of neurons, among which are hidden layer, which interpret input signals and transmit output results. Generally, feedforward neural networks are utilized for tasks like modeling non-linear static processes, classification, or approximating nonlinear functions [165, 170].

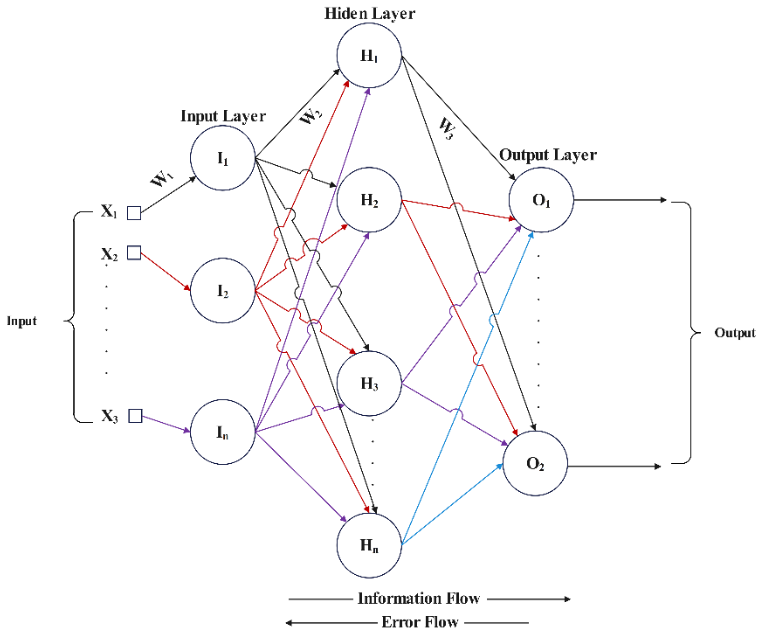

Figure II.8 – Feedforward structure.

II.2.2.2. Recurrent or feedback structure

Feedback networks, also known as a dynamic network, is the second type of neuron network in which information circulates cyclically, this means that neuron outputs can be used as feedback inputs for other neurons (Figure II.9). This makes them appropriate for dynamic information processing in systems that vary over time such as predicting chronological series, identifying and modeling systems, and controlling processes or filtering [165, 170].

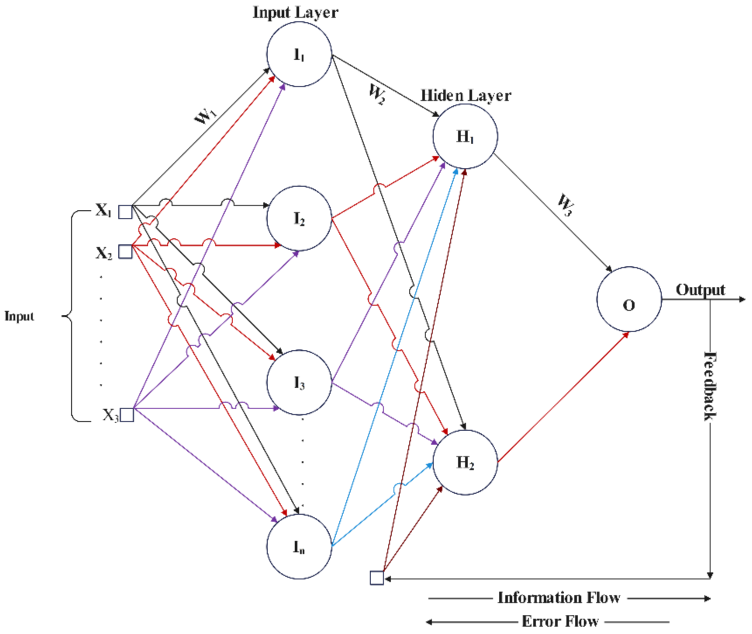

Figure II.9 – Recurrent structure.

II.2.3. Learning ANNs

Training ANN designates the mechanism of adjusting the weights of neuronal connections in order to optimize the network's efficiency in accomplishing a specific assignment. This entails iteratively updating the network's parameters in response to the data input and the intended result. Supervised learning, unsupervised learning and Reinforcement Learning are three commonly used algorithms for neural network learning described below.

II.2.3.1. Supervised learning

Supervised learning is a computational learning strategy where inputs data are presented together with expected responses. This works just like a coach guiding the system by showing it the appropriate results for each example. By analyzing the discrepancy between its predictions and the expected responses, the learning algorithm adjusts the internal system parameters. The goal is to minimize this difference so that the system can learn to generalize to new contexts and solve related problems [171, 172].

II.2.3.2. Unsupervised learning

In contract to supervised learning, unsupervised learning involves training the system on unlabeled data, meaning that there are no desired outputs for the input samples.

In this case, the neural network must independently, without explicit supervision, find interesting features or patterns in the data. To achieve this, they group similar data according to shared characteristics and organizing it. To reflect these groupings, the algorithm modifies Connection weights and thresholds; alternatively, the conceptor may specify the maximum number of groups based on their understanding of the issue [173, 172].

II.2.3.3. Reinforcement learning

Reinforcement learning is a behavioral psychology-based learning approach that learns through interactions with the outside world. It entails agents choosing actions based on both new possibilities and previous information, thus learning the consequences of their activities rather than taking explicit training. The key idea is to find a balance between exploring new opportunities and using existing knowledge to make decisions [172].

II.3. General Concepts of metaheuristic Algorithms

II.3.1. Optimization

Optimization falls within the realm of mathematics and computer science which seeks to model real-world problems as optimization problems, and then solve them to identify the best solution or the optimal set of solutions the most closely match a given objective while taking possible constraints into account. This solving procedure makes use of analytical or numerical approaches, or by applying heuristics or metaheuristics, specifically for complex problems where standard numerical methods may not be unworkable or for situation in which it is not practical to conduct a thorough search for the best solution [174, 175].

II.3.2. Optimization problem

Optimization problem is identified as the search among a set of solutions « S » belonging to the set of real numbers « R » for one or more of the best optimal solutions « x », according to a criterion that evaluates the quality of the solutions, this criterion is defined by a mathematical function « f », known as the objective function or cost function f(x). By minimizing or maximizing f(x), optimal solutions are obtained, and this varies depending on the type of problem. In a minimization problem, the goal is to find x that makes f(x) as small as possible, or x represents the decision variables. While in a maximization problem, the objective is to make f(x) as large as possible [176].

However, when tackling an optimization problem, there may be intermediate solutions in the search space, representing optima, also known as local optima, but their validity is limited to specific regions of the search space. Nevertheless, the primary objective remains the search for the best possible solution, often referred to as the global optimum. This distinction between global and local optimal optima adds significant complexity to solving optimization problems, as finding the global optimum requires careful navigation through these local solutions. The latter, although prima facie optimal in their restricted spaces, can sometimes be misleading in their apparent limited optimality. Taking into account the assumption that the search space « S » is a topological space, meaning it has a structure allowing the definition of neighborhoods. This assumption becomes necessary to describe local solutions to the optimization problem, as it facilitates determining points close to each other in this space, essential for understanding local optima. Thus, the assumption of a topological space plays a fundamental role in the analysis of local solutions when searching for the global optimum [176]. To illustrate this, an example of local and global minima and maxima displayed in (Figure II.10).

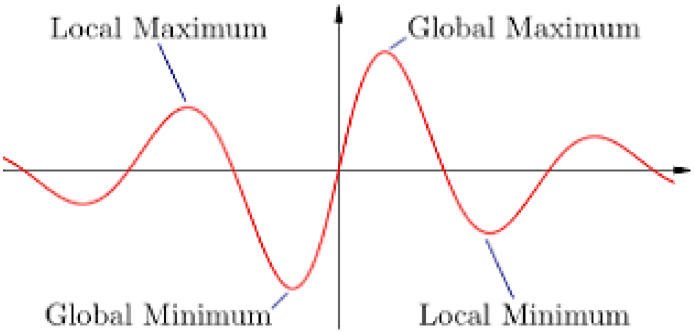

Figure II.10 – Local and global minima and maxima.

II.3.3. Classification of optimization problems

There are specific types of optimization problems [177, 178, 179]:

- ✓ **Constrained and unconstrained optimization problems:** Unconstrained problems are those in which the objective function is optimized over a set of possible values without any restrictions on the decision variables. Constrained problems, on the other hand, entail optimizing an objective function within a limited set of possible values subject to specific conditions or constraints.
- ✓ **Multi-objective and mono-objective optimization problems:** multi-objective optimization refers to the simultaneous optimization of multiple contradictory objectives; this means striking an equilibrium or a compromise between the different objectives in order to arrive at a solution that is generally satisfactory. As opposed to mono-objective optimization, which focuses on finding the best possible solution in relation to a single objective.
- ✓ **Multimodal and unimodal optimization problems:** multimodal optimization refers to problems where a function admits multiple local optima, that is, multiple points in the search space where the objective function reaches its maximum or minimum. In the other case, it is unimodal.
- ✓ **Combinatorial and continus optimization problems:** combinatorial optimization entails determining the optimal combination solution among a discrete and limited set of possible options. While continus optimization involves finding the values of the variables that optimize the objective function within the continuous space of choices.
- ✓ **Static and dynamic optimization problems:** in static optimization problems, the range of possible values that the problem's variables might have contained and the objective function to be optimized are fixed throughout the resolution process. However, in dynamic optimization problems, they may change over time depending on a number of factors.

II.3.4. Metaheuristics

Classical or exact optimization approaches thoroughly examine every possibility in a search space in order to identify the best possible (optimal) solution via deterministic algorithms. Amidst these methods, Newton's method, the simplex method, or the gradient method can be cited. However, their applicability to large-scale problems may be limited due to their frequently lengthy computation times and high mathematical complexity. To get around these restrictions, techniques that are heuristic or approximate have been introduced. The latter provide a more adaptable and effective method for resolving complicated issues in less time. These strategies do not guarantee optimality, despite this consideration, they persist to be helpful tools by yielding excellent results, which makes them indispensable for plenty of applications [180].

The term "heuristic" originates from the ancient Greek "eurisko," meaning "to find," and refers to anything that is helpful for research, invention, and discovery. Contrary to metaheuristics, these methods are often specific to a type of problem and are hence dependent on it. The term "metaheuristic" comes from the ancient Greek word "meta," which means "beyond" or "at a higher level" and the previously mentioned heuristic concept. In 1986 [181], Fred Glover made the first appearance.

Metaheuristic techniques share a number of vital characteristics in their strategy to deal with complex problems. They use exploration and exploitation strategies to take advantage of the solution space, frequently following random components to stay away from local optima. These techniques usually involve an iterative process until a satisfactory solution is found, and they are meant to be flexible enough to accommodate different kinds of problems. Despite their specific differences in how they explore the solution space and look for high-quality solutions using various principles and mechanisms, these approaches all aim to find high-quality solutions for challenging problems in a reasonable amount of time. The following flowchart represents a general process that numerous metaheuristics follow to solve a wide range of optimization problems. Each step is crucial for effectively guiding the search towards excellent solutions [182]:

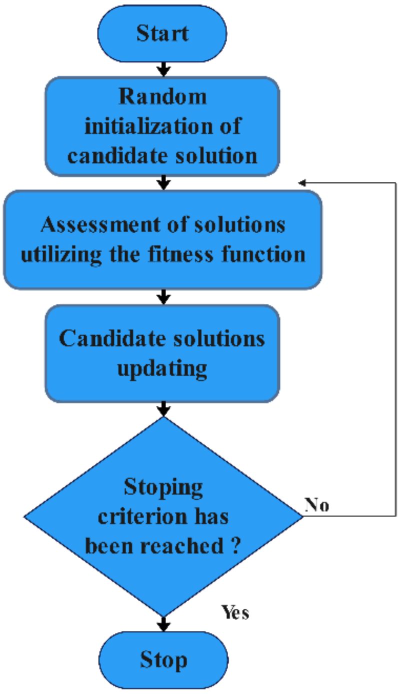

Figure II.11 – General operating flowchart of metaheuristic techniques.

II.3.5. Classification of metaheuristic techniques

There is a wide range of metaheuristic techniques, and they can be categorized along several criteria according to:

- ✓ **The total number of solutions:** manipulating a set of solutions, referred to as population-based, or one solution at a time, known as local search or trajectory methods.
- ✓ **The research history**: These methods utilize search history and previous findings in order to steer optimization, either long-term or short-term, using memory.
- ✓ **The areas of inspiration they are derived from:** Their origin, whether or not they drew inspiration from nature.

The selection of an optimization algorithm for a given problem is crafted by the kind and nature of the problem that has to be solved, specific constraints, the availability of data, and the anticipated performance in terms of computational time and solution quality. In what follows, I will proceed to a thorough explanation of the four specific metaheuristic methods I have selected for my study.

➢ Invasive weeds optimization

The Invasive Weed Optimization (IWO) algorithm was introduced by Alireza Mehrabian and Caro Lucas in 2006 [**183**], it is a population-based optimization technique which takes its cues from the natural behavior of invasive weeds to solve continuous optimization issues.

Weeds are distinguished by their vigor, rapid environmental adaption, and ability to spread, which makes them a serious threat to agricultural crops. Weeds infiltrate agricultural areas through airborne seed dispersal. With the help of the resources at hand, these seeds take up the available space and develop into flowering weeds. The process is repeated when new weeds that go through the same stages are haphazardly scattered over the field and grow into flowering weeds.

The following points outline the steps of the IWO algorithm:

- Randomly generating an initial population of *n* weeds.
- Generating the seed population.
- Evaluating each seed's fitness, ranking them according to their fitness; these seeds are now referred to as ‘’flowering weeds”.

- The previously flowering weeds produce new seeds based on their rank. From the lowest-ranked weed to the highest-ranked weed, the quantity of seeds generated by each weed varies between $S_{\min}$ and $S_{\max}$ and increases linearly.
- Utilizing the normally dispersed arbitrary values that have a mean equal to the location of the producing weeds, seeds are generated. the subsequent formula is utilized to vary the standard deviations:

$$\theta_{iter} = \frac{\left(iter_{\max} - iter\right)^{n}}{iter_{\max}^{n}} \left(\theta_i - \theta_f\right) + \theta_f \qquad \text{(II.8)}$$

Here, $iter_{\max}$ represents most iterations possible, while θ_i and θ_f denotes both the initial and final standard deviations, with n serving as the non-linear modulation index.

- Assessing the fitness of newly produced seeds which become flowering weeds, and they are subsequently ranked alongside their parent seeds based on their fitness levels.
- Removing weeds with lower fitness to attain the maximum allowable number of weeds in the colony ($P_{\max}$).
- The weeds that survive can generate new seeds based on their ranking, and this process continues until the stopping criterion is met. Typically, the stopping criterion is either the maximum number of iterations or a certain limit value of fitness.

➢ Particle swarm optimization

The Particle Swarm Optimization (PSO) algorithm was established by Kennedy and Eberhart in 1995 [184], it is a forceful meta-heuristic optimization technique that rely upon the comportment of flocking birds and schooling fish observed in nature.

PSO algorithm follows this concept: in the search area, a flock of birds is initialized at random, where each bird is referred to as a ‘’particle’’. These birds, or particles, find the ideal global position after a specific number of iterations. Each particle may alter its velocity vector for each iteration according to its momentum and the influence of its optimal position and the optimal position of the most competent individual. The particle then moves to a newly determined location, where the optimization problem's objective function can be utilized to assess its fitness. The best position previously reached by the particle (i) is called its personal best position (P_{best}). The best global position (g_{best}) is that of the particle with the best position in the whole swarm.

At each step, the particle's velocity and new position are assessed utilizing these two equations:

$$V_{(t+1)} = V_{(t)} + C_1 r_1 \left(pb_{(t)} - X_{(t)} \right) + C_2 r_2 \left(pg_{(t)} - X_{(t)} \right) \qquad \text{(II.9)}$$

$$X_{(t+1)} = X_{(t)} + V_{(t+1)} \qquad \text{(II.10)}$$

with C_1 and C_2 representing the acceleration coefficients, and r_1, r_2 are random variables within the a range of [0, 1].

The PSO algorithm is given below:

- Randomly generating a population swarm of N particles.
- Initializing the PSO's initial parameters (C_1 and C_2).
- Initializing particle with random position (x) and velocity (v).
- The following steps are repeated for each iteration until the stopping criterion is satisfied:
    - ✓ Fixing the target issue.
    - ✓ Computing the objective function.
    - ✓ Updating the values of g_{best} and P_{best}.
    - ✓ Updating the particle position (x) and velocity (v) in accordance with the velocity and position updating equations (4) and (5).
- Ending conditions are reached: the best global position and the optimal solution has been identified.

➢ Imperialist Competitive Algorithm

The Imperialist Competitive (ICA) Algorithm was proposed by Esmaeil Atashpaz Gargari and Caro Lucas in 2007 [185], it is a socio-political metaheuristic technique inspired by the historical process of colonization and empires competing to get more colonies.

The algorithm initiates with a random starting population (country). The strongest nation will serve as the imperial heart of the empire, while the others will form colonies. To assess an empire's general intent, we linearly combine the result expected by the imperialist with the average of the objective values of the empire's colonies. The empire with the highest level of vulnerability is identified after an assessment of all the empires. The other empires then compete to take possession of this empire's weakest colony.

The ICA algorithm is summarized below:

- Initializing the algorithm.
- Generating a collection of arbitrary solutions within the optimization problem's search space and create initial empires. Countries are randomly generated and their power is assessed utilizing the cost function.
- The countries possessing the lowest function values become Imperialists, seize control over other countries (or colonies), and founding the initial empires.
- Assimilation induces in seeking optimization space each empire's colonies to move closer to the imperialist state.
- Revolution guides to random and sudden changes in some countries' relative positions in the space of research.
- A colony may rise to a more advantageous position through assimilation and revolution, with the potential to overthrow the current imperialist state of the empire and seize control of the entire empire.
- All empires participate in the imperialist competition with the goal of winning and seizing control of other empires' territories. At every level of the algorithm, all empires have the chance to acquire one or more colonies of the weaker empire, depending on their strength.
- The algorithm keeps running through the previously listed stages (Assimilation, Revolution, Competition) until a stop condition is reached.

➢ Covariance Matrix Adaptation-Evolution Strategy

The covariance Matrix Adaptation Evolution Strategy (CMA-ES) algorithm was proposed by Nikolaus Hansen and Andreas Ostermeier in 2001 [187], it is a powerful versions of evolution strategy algorithm for solving continuous optimization problems. CMA-ES is built upon the evolution strategy (ES) which falls under the category of evolutionary algorithm (EA). The statistical and mathematical model underlying the construction of CMA-ES is rather intriguing, and sets it apart among other metaheuristics and evolutionary algorithms, it seeks to maximize an objective, or fitness function $f(x)$, in order to optimize a vector of parameters x. The exploration distribution's mean and covariance are regular variables that are not conditioned by the agent's present status.

CMA-ES stands out for its unique covariance matrix adaptation mechanism, which dynamically modifies the search distribution. This makes it particularly effective for optimizing complex, high-dimensional objective functions. It has been effectively employed across several fields, including robotics, engineering, as well as in contexts where traditional optimization methods might falter, such as parameter tuning, neural network architecture optimization and parameter tuning.

The CMA-ES algorithm is clarified as follows:

- Initializing CMA-ES parameters.
- Generating a population of candidate solution.
- Evaluating the fitness of each individual in the population.
- Selecting parents on the basis of their fitness values from the fittest individuals to set up the next generation's parent population.
- The following steps are repeated for each iteration until the stopping criterion is satisfied:
  - ✓ Updating the covariance matrix:

$$C^{t+1}=\left(1-c_1-c_\mu\right)C^t+c_1+P_c^{t+1}\left(P_c^{t+1}\right)^T+c_\mu\sum\nolimits_{i=1}^{\gamma}W_{Ri}\,\frac{X_i-m^t}{\sigma^t}\left(\frac{X_i-m^t}{\sigma^t}\right)^T \qquad \text{(II.11)}$$

Here c_1 and c_u represent learning rate parameters, W_{Ri} denotes the weight for the R^{ith} highest point, and P_c represents the evolution path.

$$P_c^{t+1}=\left(1-C_c\right)P_c^t+\sqrt{\frac{C_c\left(2-C_c\right)}{\sum\nolimits_{i=1}^{\gamma}w_i^2}}\,\frac{m^{t+1}-m^t}{\sigma^t} \qquad \text{(II.12)}$$

  - ✓ Updating the Step Size:

$$\sigma^{t+1}=\sigma^t\exp\left(\frac{c_\sigma}{d_\sigma}\,\frac{P_\sigma^{t+1}-X_d}{X_d}\right) \qquad \text{(II.13)}$$

Here c_σ represents the learning rate, while d_σ denotes the damping rate, with P_σ is the evolution path.

- ✓ Generating a sample population for the generation t+*1*.
- ✓ Updating mean for the generation t+*1*:

$$m^{t+1} = \sum_{i=1}^{\gamma} W_{Ri} X_i \tag{II.14}$$

- ✓ Updating the best ever solution.
- ✓ Termination criteria are satisfied: Outcomes.

II.4. Signal decomposition and filtering approach

II.4.1. Overview

Decomposition is the process of separating a given signal into its constituent parts in order to uncover previously unknown aspects of the data and to extract relevant information's from those components that is often unobtainable by taking into account the entire set of data. In many real-world signal processing processes, decomposition is an essential stage of pre-processing which aids with tasks like damage detection, noise or artifact removal, extraction of features and parameters identification through offering insightful details about the sub-relevant underlying data. There are several methods to carry out such a form of decomposition, which can be categorized into two classes: methods that decompose signals based on predefined mathematical functions or foundations, and methods that are data-driven, which directly analyze data to obtain intrinsic conformed signal components without making a lot of assumptions about the input data. Each of these techniques has its own specific advantages and applications depending on the characteristics of the signal and its nature as well as the analysis's aims, among them we briefly mention the most commonly used ones, such as Fourier transform (FT) [**187**], wavelets (WT) [**188**], Dynamical mod decomposition (DMD) [**189**] and Empirical Mode Decomposition (EMD) [**191**]. However, since most of the signals found in real-world applications have a non-stationary and non-linear nature, and this kind of signals requires special methods for their processing due to the variation over time and the relationships' complexity between the different components of the signal, making the use of conventional processing techniques difficult [**190**]. As is the case for Fourier transform, it inability to analyze non-stationary or non-linear signals may present challenges [**191**]. In such situations, considering the inherent limitations of certain techniques, it is appropriate to opt for a method that has fewer constraints and is specifically designed for the analysis of nonlinear and non-stationary signals, such as EMD and its variant [192], EMD is characterized by his adaptability to such type of signal and the fact that is empirical, meaning it is not based on predetermined models or theoretical presumptions, but rather on adaptive basis systems derived straight from the data, Moreover, this method stands out by its adaptive modes oriented towards the locally dominant frequencies or scales to efficiently extract the intrinsic components of the signal which provide more robust results without altering the crucial information contained in these signals, especially when dealing with complex and sensitive signals such as biomedical signals like ECG. The discrete waveform transform (DWT) [193] [194] is also well-suited for handling non-stationary and non-linear signals; it is widely employed for ECG signal filtering. It captures the complex temporal and frequency characteristics of ECG signals, making it a valuable method for denoising and detecting events of interest in the signal. Combining DWT with other filtering algorithms can sometimes yield better results in terms of the quality of the denoised signal.

The following will explain in detail the principal functioning of each of EMD, wavelets and GEMD which considered as a performant variant of EMD. These techniques will be adopted in our study for ECG signal denoising.

II.4.2. Empirical mode decomposition

II.4.2.1. Definition

Empirical mode decomposition (EMD) is a data processing innovative technique created by Huang *et al.* [6] for decomposing nonlinear and nonstationary signals into a collection of oscillatory components known as intrinsic mode functions (*IMFs*). An IMF is defined as a function that satisfies the following two conditions [191]:

(i) The number of zero crossings and extrema must be equal.

(ii) The envelopes' average values, as determined via local maxima and minima, are invariably zero.

II.4.2.2. Functioning of the EMD Method

EMD utilizes an approach that is iterative called sifting, which consists of the following phases [191]:

1. Local minima and maxima extraction of $x(t)$.
2. Location of both upper and lower envelopes utilizing cubic spline interpolation.

3. Average the upper and lower envelopes to get the mean envelope $m(t)$.
4. Calculation of the temporary local oscillation:

$$h(t) = x(t) - m(t) \tag{II.14}$$

5. Check the subsequent condition:
   - *The average h(t) is regarded as the first IMF (ci(t)) if it is near zero; if not, reprise steps (1-3).*
6. Computing residue :

$$r(t) = x(t) - c_i(t) \tag{II.15}$$

7. Repeat steps (1 to 5) using $r(t)$ instead of $x(t)$ to obtain the next IMF and residue.
8. The process of decomposition stops when the residual $r(t)$ becomes a constant that no longer satisfies the conditions of an IMF.

$$x(t) = \sum_{i=1}^{N} c_i(t) + r_N(t) \tag{II.16}$$

where $c_i(t)$ denotes the m^{th} IMF and $r_N(t)$ denotes the last residual obtained from EMD algorithm.

II.4.3. Graph empirical mode decomposition

Graph empirical mode decomposition, or GEMD is a graph-based variant of the well-known EMD [172]. It enhances several features of the halting criterion, interpolation, and extrema of EMD. In the GEMD decomposition process, an initial phase involves the construction of a graph and the determination of its adjacency matrix. The latter is a square matrix and represents the relationships between the graph nodes. Then, the fundamental steps of the classic EMD are implemented to obtain the IMFs.

II.4.3.1. Graph construction

A graph is made up of A collection of nodes (also known as vertices) V and a set of links (also known as edges) that join these nodes. Usually shown as G = (V, E). Graph data is often represented as a vector or scalar function connected to the graph's vertices. Let V be the graph's node set. When exploring options for defining the edges in the graph supporting the signal, we consider two approaches [196]:

1. **A weighted graph parameterised by δ:** only pairs of nodes $V_{(i,j)}$ at a distance $d_{i,j}$ shorter than δ are connected by an edge, with weight $w_{i,j} = exp(-d_{i,j}^2 / 2\delta^2)$ (II.17).

2. **A binary graph parameterised by *k*:** Every node is linked to its *k* nearest neighbours (k-NN).

Whichever approach we choose, we obtain a graph $G = (V, E)$. Let A be its adjacency matrix.

II.4.3.2. Definition of local extrema

A node i is a local maximum (or minimum) for a signal x specified on V if, for all of its neighbors k in G, $x(i) > x(k)$ (*resp.* $x(i) < x(k)$). Determining these local minimum points in a graph signal is crucial because they help identify the IMFs throughout the decomposition process.

II.4.3.3. Interpolation procedure

Once the local extrema have been defined, an interpolation procedure is used to obtain smooth upper and lower envelopes. These envelopes are used in the decomposition process to define the boundaries of the MFIs.

II.4.3.4. Choice of stopping criteria

When applying GEMD, it is necessary to define criteria to determine when to stop the decomposition process. These criteria can be based on predefined thresholds, such as the number of iterations, the convergence of the algorithm, or the quality of the approximation obtained.

II.4.3.5. Functioning of the GEMD Method

Similar to the classical EMD, the GEMD can be built in the following manner [195, 196]:

1. Creation of the adjacency matrix A for the graph G;
2. The process starts by Initializing $m = x_i$;
3. Conditional loop, wherein steps 4 through 8 are repeated as long as the condition (the stop criterion) is not achieved;
4. Detection of the local extreme of m;
5. Interpolation of the upper and lower extremes of m to obtain the envelop e_{max} and e_{min};
6. Computation of the average envelope $z = \frac{e_{min} + e_{max}}{2}$;

7. Subtraction of the average envelope from the signal: $m = m - z$;
8. Updates for the next iteration, include setting $d_{i+1} = m$ and $x_{i+1} = x_i - m$;
9. Stop the decomposition and terminate if m meets the stopping criteria, return stored IMFs, and obtain:

$$x(t) = m_K(t) + \sum_{k=1}^{K} d_k(t) \tag{II.18}$$

with K denotes the number of modes to be extracted.

II.4.4. Wavelets

In mathematics and science, Fourier analysis has long been one of the most popular and effective tools. This representation works well for processing stationary signals, or signals with certain features that are constant across time and is based on the physical concept of frequency. Nonetheless, the majority of signals in the actual world are not stationary, and the content of the information they carry is exactly determined by the evolution of their statistical, frequency, temporal, and spatial characteristics. Any concept of temporal location (or spatial for images) vanishes during Fourier's analysis; consequently, in order to obtain a time/frequency or space/scale representation of the data, a compromise or transformation that reveals information about the frequency content while maintaining the temporal location must be found. This is where the wavelet theory was introduced in 1980 [**200**, **197**, **198**, **199**]. The term "wavelet", symbolized by the Greek letter ‘’ψ ‘’ refers to oscillating function with zero mean, that exhibits a particular level of regularity and finite support. It’s used to localize a given function in both space and scaling. A family of wavelets $\psi_{a,b}(t)$ can be constructed from mother wavelet $\psi(t)$ by dilatation and translation depending on the nature of the signal, this mother wavelet is in the form [**198**]:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) \qquad a,b \in R \text{ and } a \neq 0 \tag{II.19}$$

The two parameters a and b, which are the scale factor and translation factor, respectively, determine the coefficients of oscillations. The analysis of the signal's lower frequency components is made possible by the large values of the factor ‘’a’’ enlarging the mother wavelet, while the analysis of the signal's higher frequency components is made possible by the factor ‘’a’’ low values contracting the mother wave.

II.4.4.1. Wavelet transform

Wavelet transforms (WT) are a popular and useful method for representing signals in the frequency and temporal domains. WT are often classified into two categories: discrete wavelet transforms (DWT) and continuous wavelet transforms (CWT), which are further explained below [200].

A. Continuous Wavelet Transform

The continuous wavelet transform (CWT) [201, 202, 203] involves dividing a continuous function in the time domain into a set of wavelets. This transformation allows for a unified representation of a signal in both the time and frequency domains, providing insights into the data's behavior across time and frequency ranges. The mathematical concept of this transformation can be expressed as follows:

$$X\left(a,b\right)=WT(x(t))=\left\langle x,\psi_{a,b}\right\rangle=\frac{1}{\sqrt{|a|}}\int_{-\infty}^{+\infty}x(t).\psi\left(\frac{t-b}{a}\right).dt \qquad \text{(II.20)}$$

$$X(a,b)=\int_{-\infty}^{+\infty}x(t).\psi_{a,b}(t).dt \qquad \text{(II.21)}$$

with a and b representing the scale and translation factors, respectively, and ψ a,b denoting the previously defined dilated and translated mother wavelet (eq).

B- Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT) [204, 205, 204, 206] stands out as one of the most renowned and widely used tools for wavelet transformations across a multitude of disciplines for numerous applications, such as signal processing, compression methods, noise reduction, feature extraction, and biological signal analysis. DWT decomposes signals into various frequency components across multiple resolutions in a discrete and non-continuous manner. It uses high-pass and low-pass filters to separate signals into detailed and approximate information. The signal's lower frequency components are represented by the approximation, while the higher frequency components are represented by the details. The number of decompositions, also known as the number of levels, is determined by the type of signal being considered and the problem that needs to be solved.

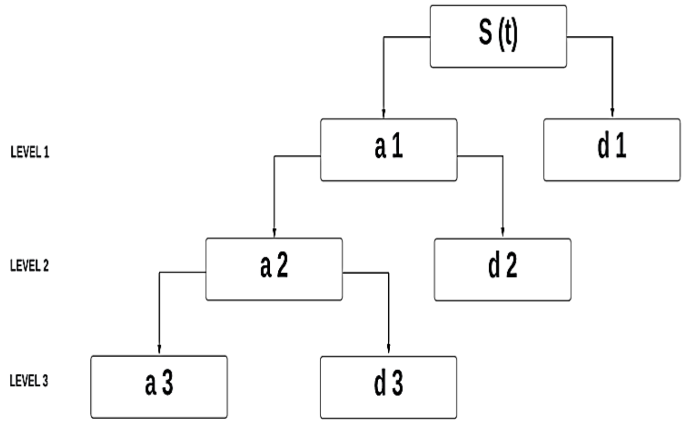

Figure II.12 – An example of decomposition of a signal into three levels of approximations and details.

After performing some processing on the wavelet components obtained by the decomposition, the resulting signal *S'(t)* will be obtained by the inverse wavelet transform. This operation is called "Reconstruction".

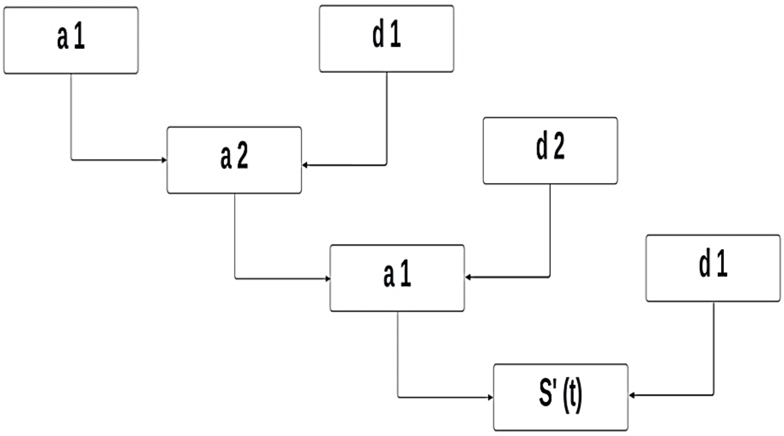

Figure II.13 – Reconstruction.

II.5. Conclusion

This chapter is devoted to presenting the fundamental methodologies and tools that will be used later on. These methods include the use of artificial neural networks (ANN) and metaheuristic algorithms such as IWO, PSO, ICA, and CEMA-ES. On the other hand, techniques for decomposition and filtering ECG signals include wavelets and EMD and its graph-based variants. The theoretical foundation of these approaches has been clearly laid out in order to set the stage for the strategies that will be proposed in the upcoming chapters.

CHAPTER III

MODELING AND IDENTIFICATION OF DYNAMICAL SYSTEM AND TIME SERIES

The objective of modeling and identification is to find simple and effective representations to describe the behavior of dynamical systems or time series, thus allowing a better understanding and practical applications such as classification, prediction and control. In this chapter a new method for modeling and identification of dynamical systems and time series based on artificial neural network autoregressive moving average (ANNARMA) and optimization algorithms will be presented.

Contents

III.1. Introduction

The learning capabilities, flexibility and symbolic reasoning make ANNs the predominant choice in many fields such as medicine, engineering, economics, navy, military and many others for optimization, prediction, modeling, identification and control of complex systems. One benefit of using ANNs in modeling is their ability to improve the accuracy and usability of complicated natural systems with a large number of inputs. This led a number of researchers to use it as a modeling tool in their studies rather than statistical modeling techniques [**207**, **208**, **209**], due to the fact that in combination with other strategies, they can be an extremely successful way to enhance modeling performance and yield better accuracy than when employed independently.

Any ANN-based technique that aims to achieve superior performance must have an algorithm for training and update its weights iteratively in order to minimize the error function, which is defined as the desired and target output, avoid getting stuck in local minimums, and accelerate the rate of convergence. Metaheuristic algorithms are effective due to their capability to enhance neural network models, enabling them to accurately address complex and large-scale problems [**183**]. They are presently considered state-of-the-art for numerous optimization challenges, particularly those that are exceptionally intricate and feature high dimensionality.

III.2. ANN-ARMA model

As described in the previous chapiters, ANN constitutes an artificial intelligence technique that mimic the human neurological system to solve complex problems. These networks consist of artificial neurons linked via communication lines, each with a weight and bias. The activation function determines the weights and biases of each neuron. ANN uses optimization algorithms like gradient descent to minimize the difference between expected and actual output. Alternatively, Autoregressive (AR) models and Moving Average (MA) models are commonly utilized in time series analysis. AR models forecast the next value in a series by relying on prior values, whereas MA models predict the next value based on the average of the previous values. ANN can also be used for time series analysis, including both AR and MA models. In an autoregressive artificial neural network (ARANN), the input to the network is a time series sequence, and the network uses prior values to predict the next value in the sequence. In a moving average artificial neural network (ANNMA), network's input is a moving window of the time series sequence, and the network forecasts the subsequent value based on the average of the previous values in the window. A commonly used combination of these two approaches is the

Autoregressive Moving Average (ARMA) model, leveraging the strengths of both techniques. In an ARMA model, the network utilizes both the previous values and the moving average of previous values to predict the next value in the time series sequence. AR, MA, and ARMA models can all be implemented employing ANNs for time series analysis. The selection of which model to utilize relies on the distinct characteristics of the time series data and the objectives of the analysis. For ARMA model, output is modeled as a linear difference equation between the current and past inputs and past outputs as described in the following equation [34]:

$$y(t) = \sum_{i=1}^{n} a_i x(t-i) + \sum_{j=1}^{n} b_j u(t-j) \tag{III.1}$$

where $x(t)$ and $u(t)$ denote the inputs and outputs, respectively, with a_i and b_j being the ARMA parameters.

These outputs being added to a neural network as inputs is equivalent to changing its architecture into a recurrent neural network. The aim of this hybridization of ANN and ARMA models is to merge their respective strengths to obtain a more reliable modeling outcome.

As a way to minimize the error between the model's output and actual output, an algorithm is utilized to update the model parameters. Creating an appropriate fitness function, alternately referred to as an objective function, is crucial for effectiveness of the system identification, it is formulated to find the control parameter values that most closely match the desired goal. Typically, the control parameters must be selected within certain restrictive limits. In this work, the Main Square Error (MSE) criterion function was used which is described in the following equation:

$$MSE = \frac{1}{N} \sum_{k=1}^{N} (y_k - \hat{y}_k)^2 \tag{III.2}$$

where y_k and $\hat{y}_k$ denote the actual measurement and its corresponding estimate, respectively, with N is is the data length.

III.3. The proposed modeling and identification method

This section will explore the suggested ANN approach to model and identify dynamical systems and time series. The suggested technique involves three key phases: firstly, identifying the primary model; secondly, identifying the error process; and lastly, creating the final model, which involves a parallel connection between the initial two steps.

III.3.1. Updating ANN parameters

Feedforward neural network is the type of neural network employed in this study. The used ANN structure throughout this present work is given by the (Figure III.1):

The ANN model's parameters to be trained by the optimization mechanisms are the weights (w_i), which are the parameters in a neural network's hidden layers, change the input data, and the Biases (b_n), which are the constants added to the product of features and weights. They are utilized to offset the outcome.

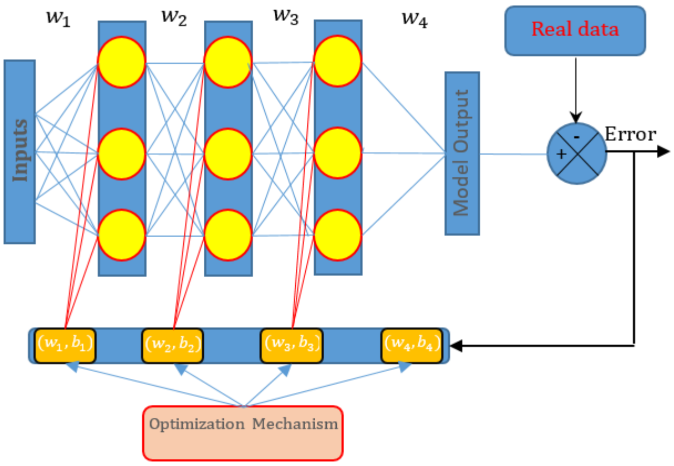

Figure III.1 – Structure of an ANN.

III.3.2. Primary model identification

At this stage, the primary artificial neural network model (Y_{PM}) for the given dynamical system or time series is established utilizing the input-output dataset (U_k, Y_k) as shown in (Figure III.2). The primary ANN model is constructed by incorporating an autoregressive moving average artificial neural network model (ANN-ARMA) that attempts to predict the new output based on the sum of past outputs and past inputs explicitly. The structure of the primary ANN-model primarily focuses on the online adaptation of the neural network's feed forward parameters. The block for parameter optimization illustrated in (Figure III.2) employing algorithms such as IWO, PSO, ICA, or CMA-ES, is responsible for adjusting the primary model's parameters in such a way that the error $E_k$between the system output Y_k and the estimated output $\widehat{Y}_k$ from the primary model reaches its minimum value.

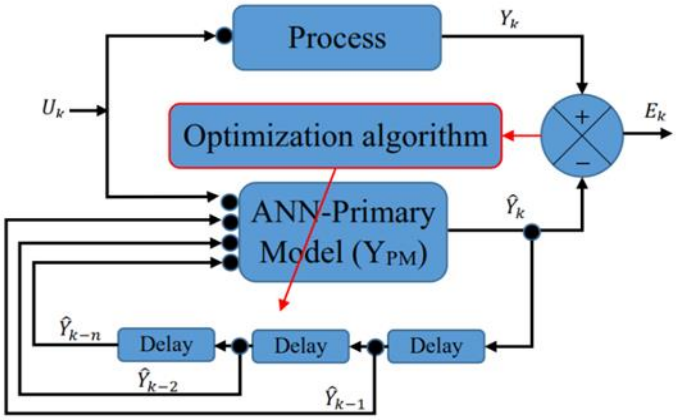

Figure III.2 – ANN-primary model.

III.3.3. Error process identification

During this subsequent stage, the process will mirror the initial step, with the emphasis now placed on determining the error of the first stage (E_k). This error arises from a parallel link between the pertinent dynamical system process or time series (y_k) and the output of the primary model ($\hat{y}_k$). The error E_k is explicitly characterized by:

$$E_k = y_k - \hat{y}_k \tag{III.3}$$

After obtaining the error process E_k, the subsequent step is its modeling through a second ANN model known specifically as the ANN error model (Y_{EM}). The error E_k may be viewed as a time series, so it is appropriate to design its model using an autoregressive model (AR), which attempts to predict the new output based on the previous results. The structure of this stage is depicted in (Figure III.3).

The ANN error model's structure mainly focuses on the online adaptation of the neural network's feed forward parameters. The block for parameter optimization shown in (Figure III.3) employs algorithms like IWO, PSO, ICA, or CMA-ES to fine-tune the parameters of the error model in a way that the error E_{1k}between the output of the error process E_k and the output's error model $\hat{E}_k$reaches its minimum value.

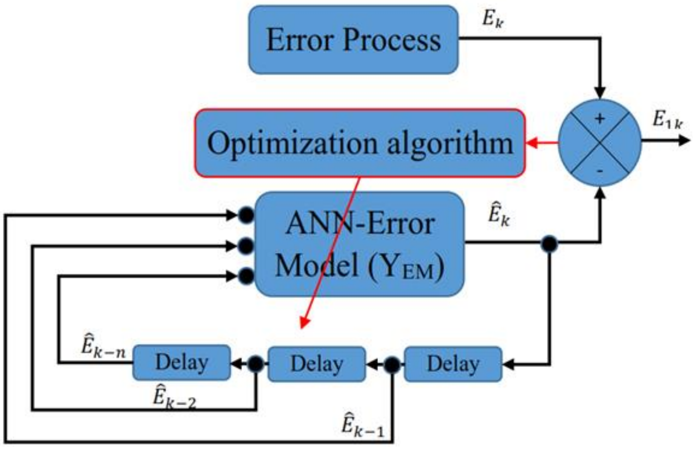

Figure III.3 – ANN-error model.

III.3.4. Final model design

Ultimately, the main model and the error model will be connected in parallel, forming the final ANN model as illustrated in (Figure III.4). This parallel interconnection is designed to diminish the modeling error and acquire a net final model.

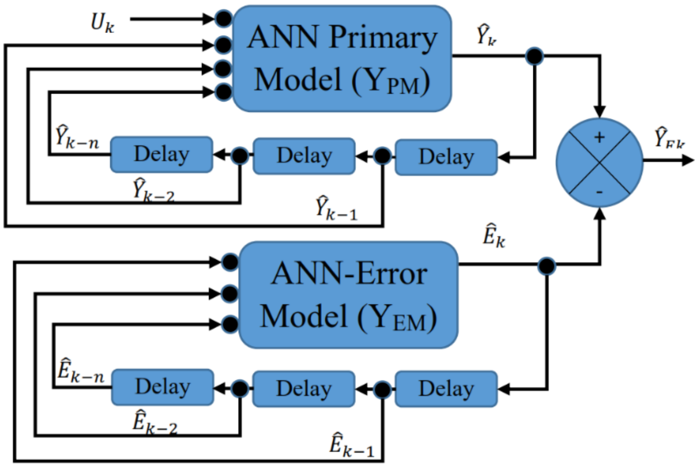

Figure III.4 – Design of the ANN-final model.

III.4. Results and discussion

The simulation outcomes of the suggested approach for dynamical system modeling and identification will be given and discussed in this section. To test the effectiveness of the suggested strategy, the three nonlinear dynamical systems listed below will be utilized [210]:

- System 1:

$$y_p(k+1) = f[y_p(k), y_p(k-1), ..., y_p(k-n+1)] + \sum_{(i=0)}^{(m-1)} \beta_i u(k-1) \tag{III.4}$$

- System 2:

$$y_p(k+1) = \sum_{i=0}^{n-1} \alpha_i y_p(k-1) + g[u(k), u(k-1), ..., u(k-m+1)] \tag{III.5}$$

- System 3:

$$y_p(k+1) = f[y_p(k), y_p(k-1), ..., y_p(k-n+1] + g[u(k), u(k-1), ..., u(k-m+1)] \tag{III.6}$$

After conducting a thorough comparative analysis, it has been determined that the IWO algorithm stands out as the most effective optimization algorithm among the four algorithms utilized. This superiority will be highlighted in the comparative study section. Subsequently, the simulation results of our approach based on the IWO algorithm will be presented. The weights and bias components of the proposed ANN model are adjustable through this algorithm. Here are the key parameters of the IWO algorithm utilized in our study:

- ✓ *Initial and final population size: 10 and 25 respectively.*
- ✓ *Maximum and minimum number of seeds: 0 and 5 respectively.*
- ✓ *Initial and final value of the standard deviation: 1.5 and -1.5 respectively.*

III.4.1. Modeling and identification of system I

We consider for this system the particular case governed by the differential equation given below:

$$y_P(k+1) = f\left[y_p(k), y_p(k-1)\right] + u(k) \tag{III.7}$$

with:

$$f\left[y_p, y_p(k-1)\right] = \frac{y_p(k) y_p(k-1)\left[y_p(k) + 2.5\right]}{1 + y_p^2(k) + y_p^2(k-1)} \tag{III.8}$$

$$u(k) = \sin\left(\frac{2\pi k}{25}\right) \tag{III.9}$$

where f is the segment of equation (III.7) that will be identified utilizing the primary model, with u being the input signal.

The outcomes of simulation for system I are depicted in (Figure III.5) as follows:

- ✓ Figure III.5.a: Illustrates the superposition between the output of system I and the output of the primary model.
- ✓ Figure III.5.b: Display figure 3.5.a. with zoomed segments.
- ✓ Figure III.5.c: Demonstrates the superposition between the modeling error and the model of modeling error.
- ✓ Figure III.5.d: Displays the superposition between the output of system I and the output of the final model.

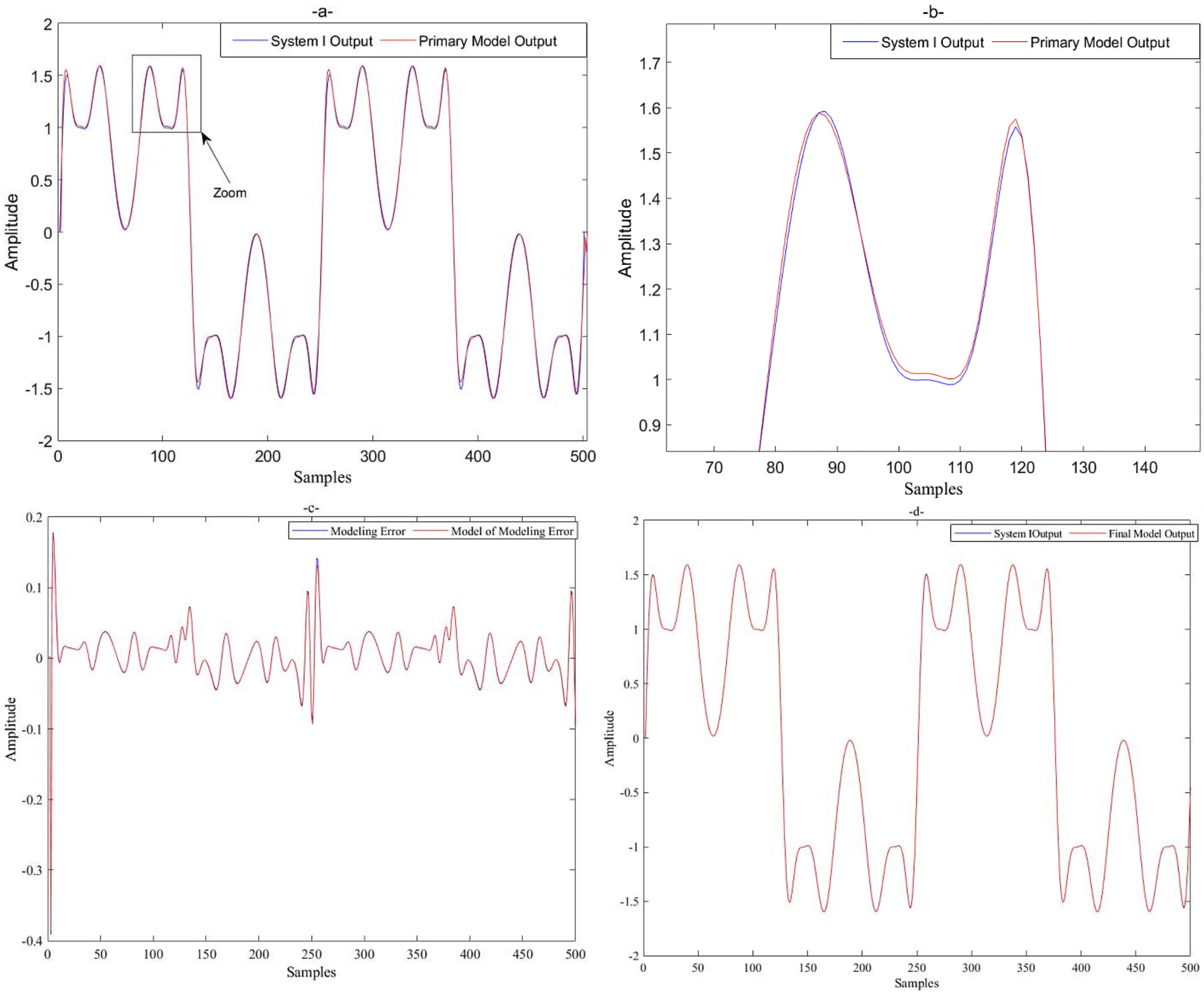

Figure III.5 – System I ANN-based IWO model.

III.4.2. Modeling and identification of system II

Regarding the second nonlinear dynamical system, the following difference equation describes the process to be identified:

$$y_p(k+1) = 0.3y_p(k) + 0.6y_p(k-1) + f\left[u(k)\right] \tag{III.10}$$

where the following is the form of the unknown function that needs to be identified:

$$f(u) = 0.6\sin(\pi u) + 0.3\sin(3\pi u) + 0.1\sin(5\pi u) \tag{III.11}$$

The signal input u is selected to be sinusoidal in the following manner

$$u(k) = \sin\left(\frac{2\pi k}{250}\right) \tag{III.12}$$

The second system modeling and identification is simulated using the same settings of the IWO optimization method as those utilized for the modeling and identification of system I; simulation outcomes are displayed in (Figure III.6).

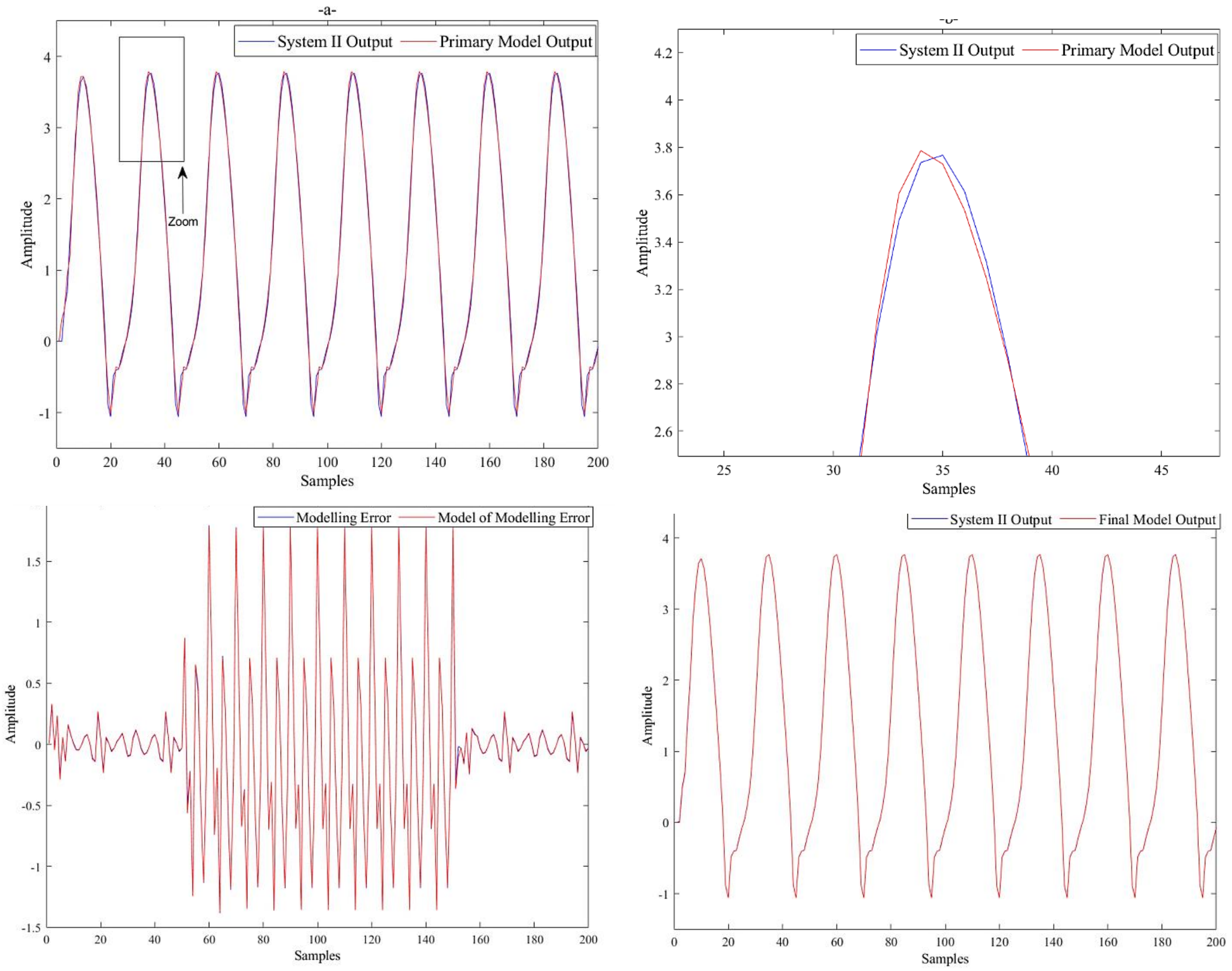

Figure III.6 – System II ANN-based IWO model.

III.4.3. Modeling and identification of system III

Concerning this system, we take into consideration the specific case that is expressed by the difference equation that follows:

$$y_p(k+1) = f(y_p(k), u(k)) = \frac{y_p(k)}{1+y_p(k)^2} + u(k)^3 \qquad \text{(III.13)}$$

$$u(k) = \sin\left(\frac{2\pi k}{25}\right) + \sin\left(\frac{2\pi k}{10}\right) \qquad \text{(III.14)}$$

Here $u(k)$ denotes the input signal. Utilizing the identical IWO optimization technique parameters as for systems I and II, we simulated this case. (Figure III.7) displays the simulation's findings:

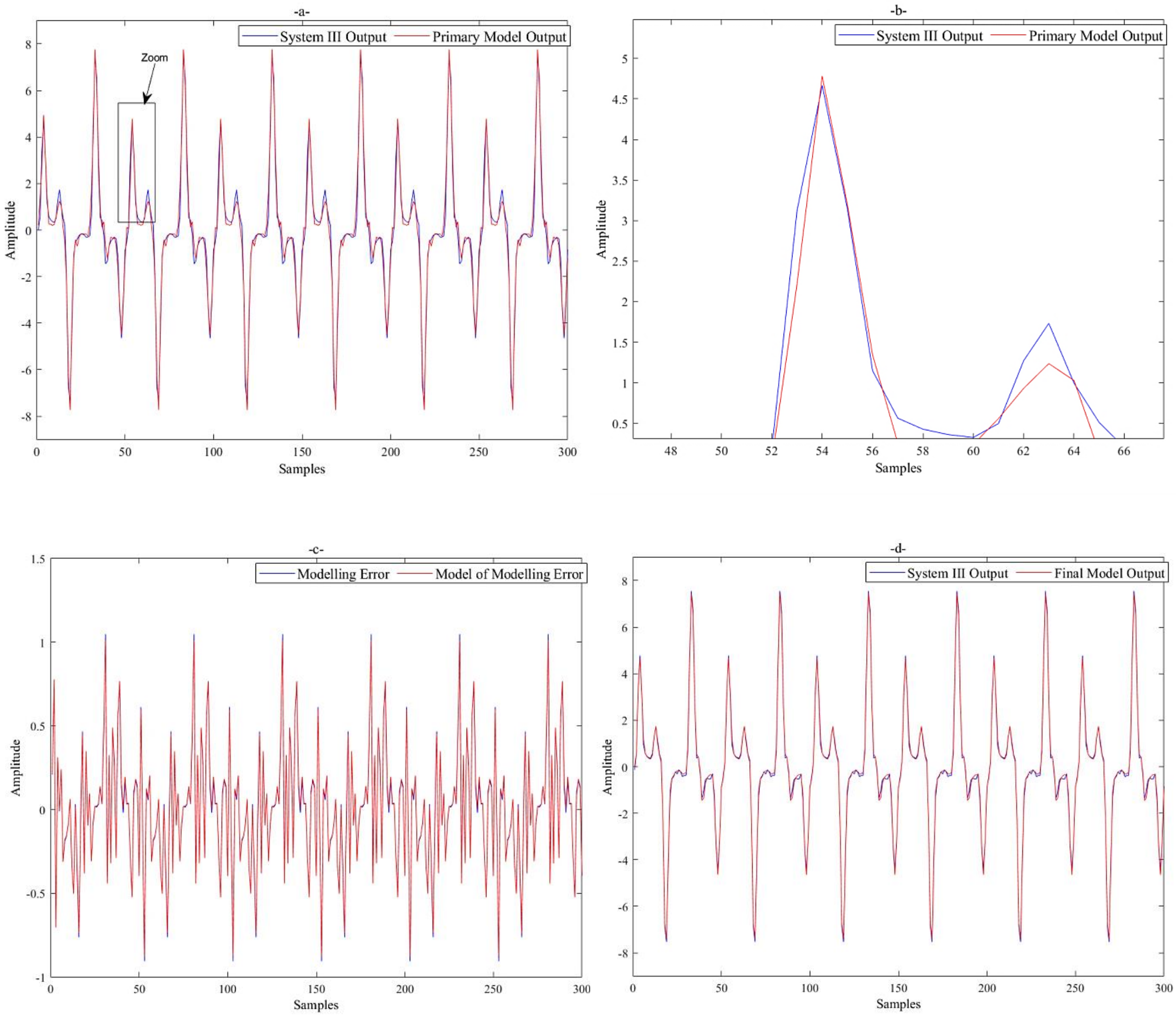

Figure III.7 – System III ANN-based IWO model.

Upon visually inspecting (Figures III.5, III.6 and III.7), it becomes apparent that the final model surpasses the primary model significantly. This observation confirms the effectiveness of the identification idea based on error models.

III.4.4. Validation and generalization tests

Validation tests have been conducted to ensure our approach's robustness and efficiency. This section provides a succinct explanation of these validation tests.

III.4.4.1. Generalization test

In the generalization process, the following steps are followed:

1. Validate the primary model using new input data u_2, generating a new error.
2. Utilize the resulting error in the error identification step.
3. Generate the final model, which represents a parallel interconnection of the two models (primary model and error model).

The results of this validation process are depicted in (Figure III.8) where:

- ✓ Figure III.8.a: Illustrates the primary model output with an input data u_1,which is defined as follows:

$$u_1(k) = \sin\left(\frac{2\pi k}{25}\right) \tag{III.15}$$

- ✓ Figure 3.8.b: shows the primary model output, with a novel input data u_2 as indicated by the subsequent equation:

$$u_2(k) = \sin\left(\frac{2\pi k}{25}\right) \quad for\ 1 \le k \le 50\ and\ 150 \le k \le 200 \tag{III.16}$$

$$u_2(k) = \sin\left(\frac{2\pi k}{10}\right) + \sin\left(\frac{2\pi k}{5}\right) \quad for\ 50 \le k \le 150 \tag{III.17}$$

- ✓ Figure III.8.c: depicts the error process model.
- ✓ Figure III.8.d: depicts the final model output.

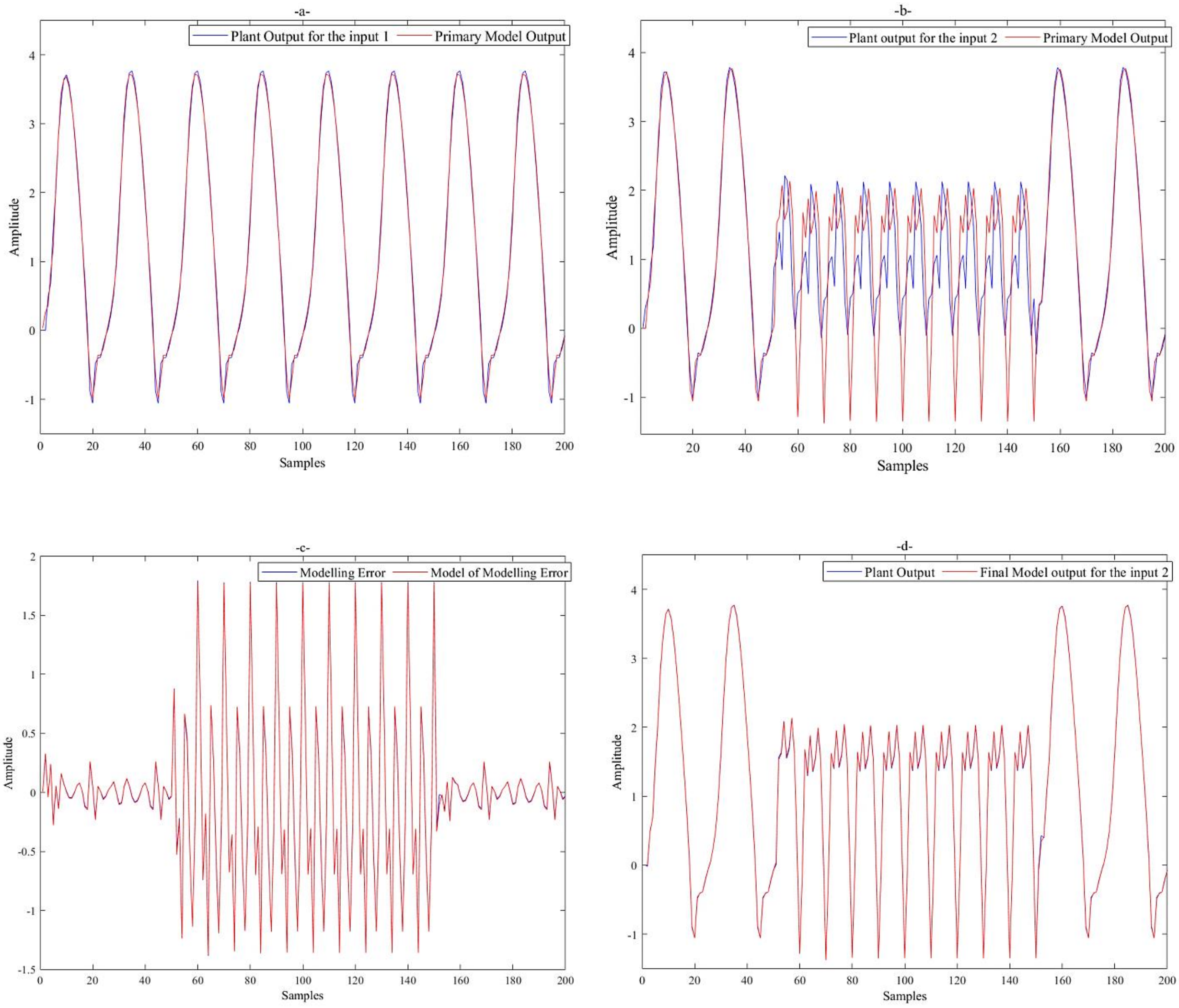

Figure III.8 – Generalization test.

The Models' generalization test is a pivotal step in evaluating machine learning and statistical. It evaluates a trained model's performance on newly acquired or unseen data. Upon visually examining (Figure III.8.d), it is evident that our model exhibits satisfactory performance when applied to novel input data, validating the efficacy of our suggested approach.

III.4.4.2. Validation test

Through conducting validation tests, we are able to evaluate the model's applicability for real-world applications, generalisation capability, and reliability. These tests play a vital role in guaranteeing that the model is not only accurate on the data it was trained on but also proves effective in making predictions on new, unseen data.

A. Modeling and identification of ECG signal

An ECG signal is one kind of time series data that depicts the heart's electrical activity over a given length of time. In the context of an ECG, the time series is made up of a sequence of voltage readings obtained at various periods throughout the cardiac cycle. This part applies our approach to the identification of two different kinds of ECG signals: the actual ECG signals acquired from the ECG PhysioNet database [211] and the synthetic ECG signal [212].

A.1. Real ECG signal

In the subsequent section, we delve into the implementation of the suggested strategy to actual ECG data. For this study, we acquired the ECG signal 100.dat dataset from the MIT-BIH normal sinus rhythm database [211], which was recorded with a sampling rate of 360 Hz and a resolution of 11 bits per sample. The results of applying the proposed technique to the real ECG signal are depicted in (Figure III.9).

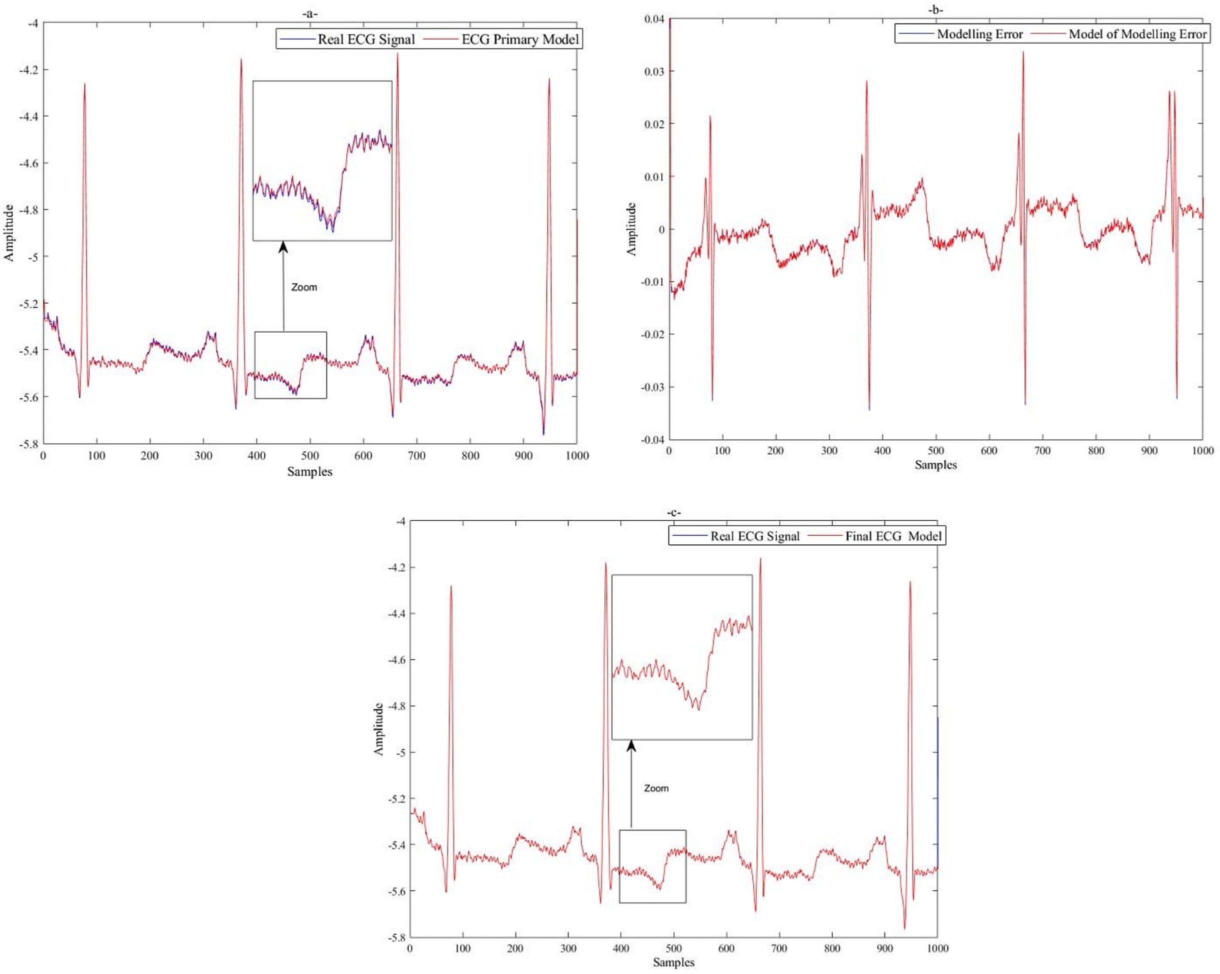

Figure III.9 – ANN based IWO model for real ECG signal: a) Real ECG signal vs primary ECG model, b) Modeling error vs Model of modeling error, c) Real ECG signal vs Final ECG model.

A.2. Synthetic ECG signal

Following the same precepts and processes as in the preceding sections to model the synthetic ECG signal data [**196**]. (Figure III.10) displays the outcome obtained.

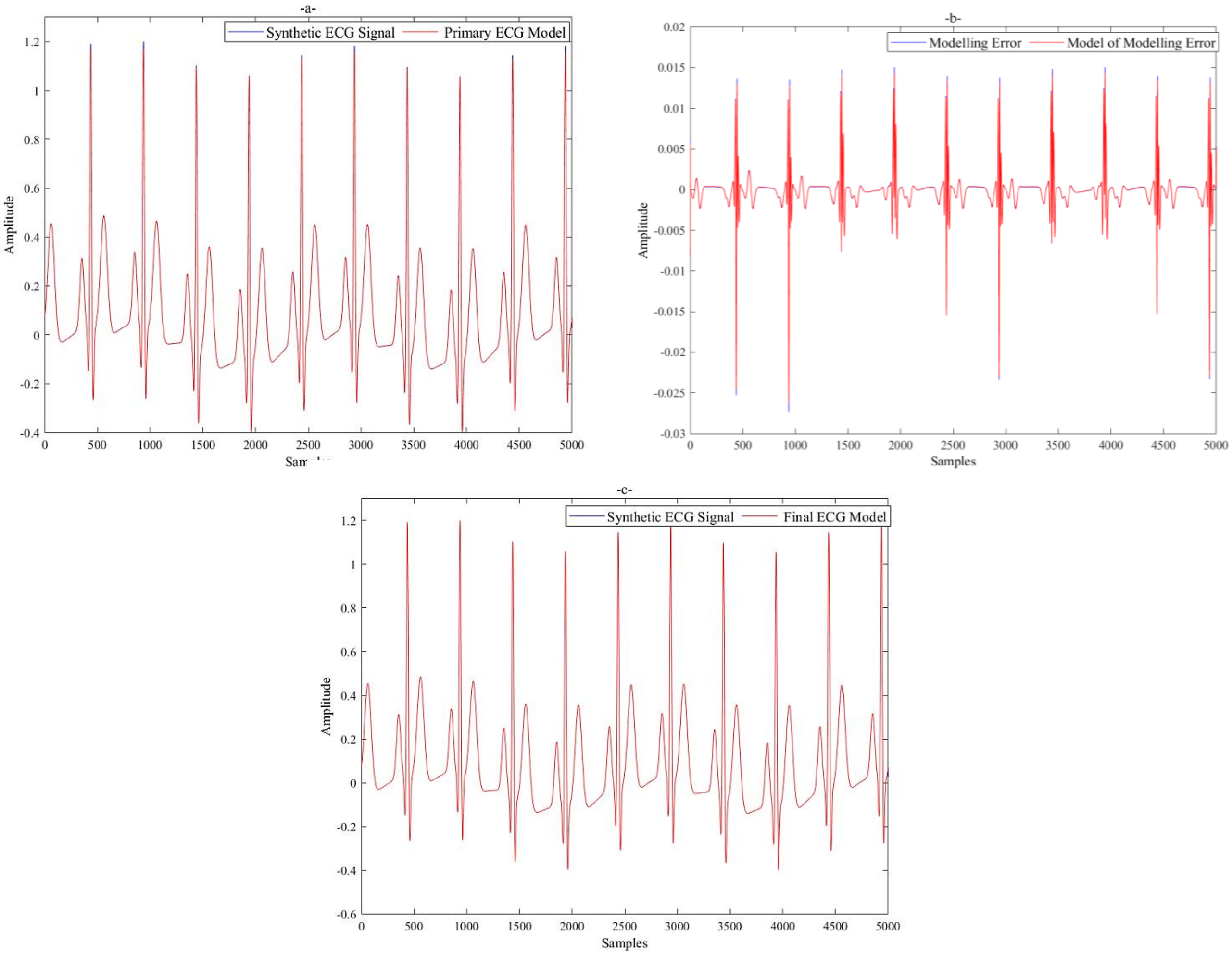

Figure III.10 – ANN based IWO model for synthetic ECG signal: a) Synthetic ECG signal vs Primary ECG model, b) Modeling error vs Model of modeling error, c) Synthetic ECG signal vs Final ECG model.

B. Mackey–Glass time series modeling and identification

Our approach is also applied to another data from a time series; remember that a time series is simply series of data points ordered in time. In a time series, time is often the independent variable and the objective is usually to make a forecast for the future. For this purpose, we consider a time series that is generated by the Mackey–Glass equation. The outcomes of simulation are shown in (Figure III.11).

The outcomes of the generalization and validation tests carried out earlier demonstrate that the modeling and identification of distinct data kinds have been successfully completed, indicating that the strategy we have suggested works well for modeling a variety of datasets.

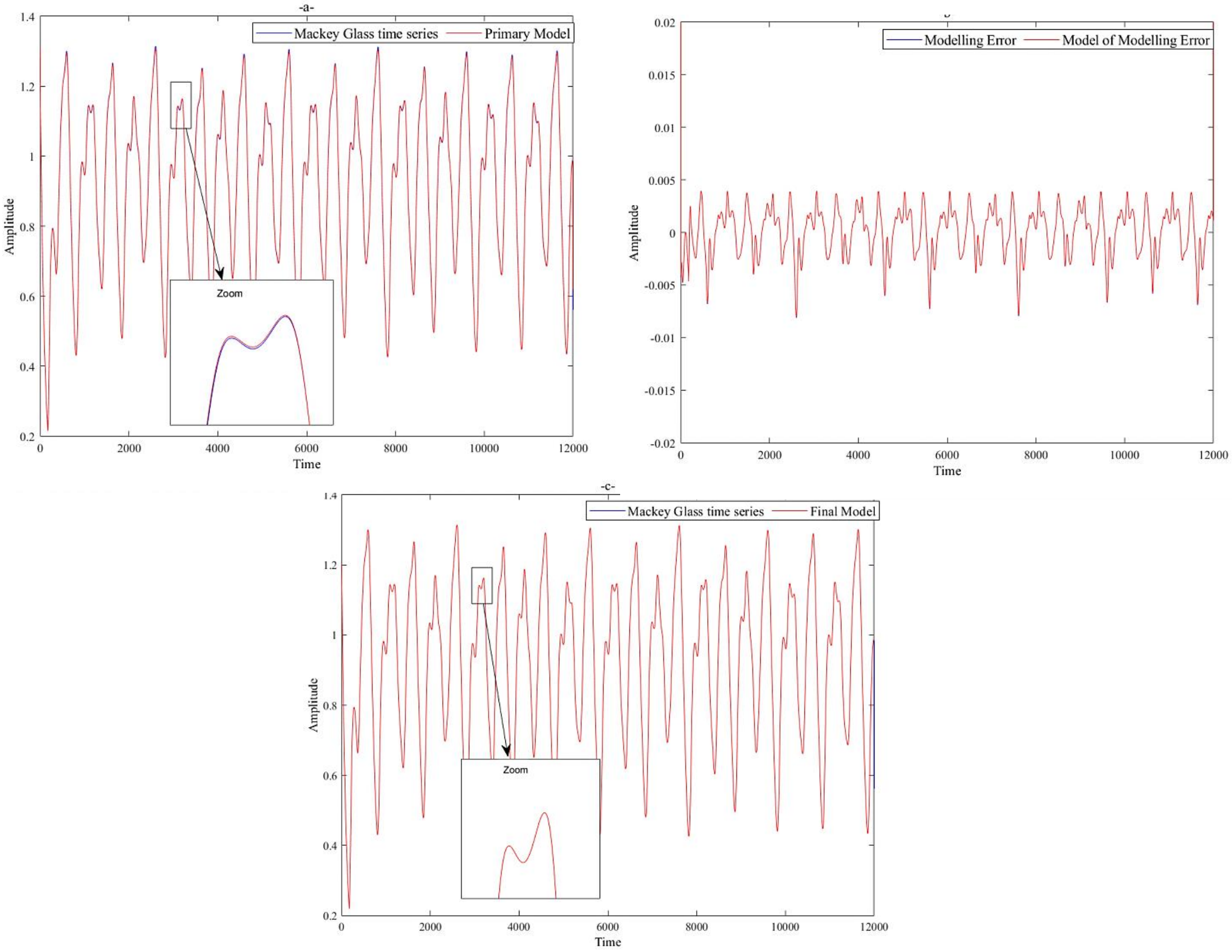

Figure III.11 – ANN based IWO model for Mackey Glass time series: a) Mackey Glass time series vs Primary model, b) Modeling error vs Model of modeling error, c) Mackey Glass time series vs Final model.

III.4.5. Comparative study

In this part, we conduct a comparative analysis to showcase the effectiveness of the IWO optimization algorithm in contrast to other optimization strategies. To achieve this, we have specifically chosen three algorithms: PSO, ICA, and ES-CMA, as detailed in section 1.2. The parameters for each optimization algorithm are defined as follows:

- **PSO algorithm parameters:** The acceleration constants are set as C_1=1.5 and C_2=2.5, with a coefficient of inertia ω = 0.48.

- **ICA algorithm parameters:** Parameters include α=1, β=1.5, and μ (revolution rate) = 0.1.
- **ES-CMA algorithm parameters:** The population size λ is set to 140, and the number of parents μ is set to 40.

Figures (III.12, III.13 and III.14) show the modeling and identification outcomes of the three dynamical systems previously discussed, utilizing the suggested modeling approach based on each of the four optimization algorithms (IWO, PSO, ICA, and ES-CMA).

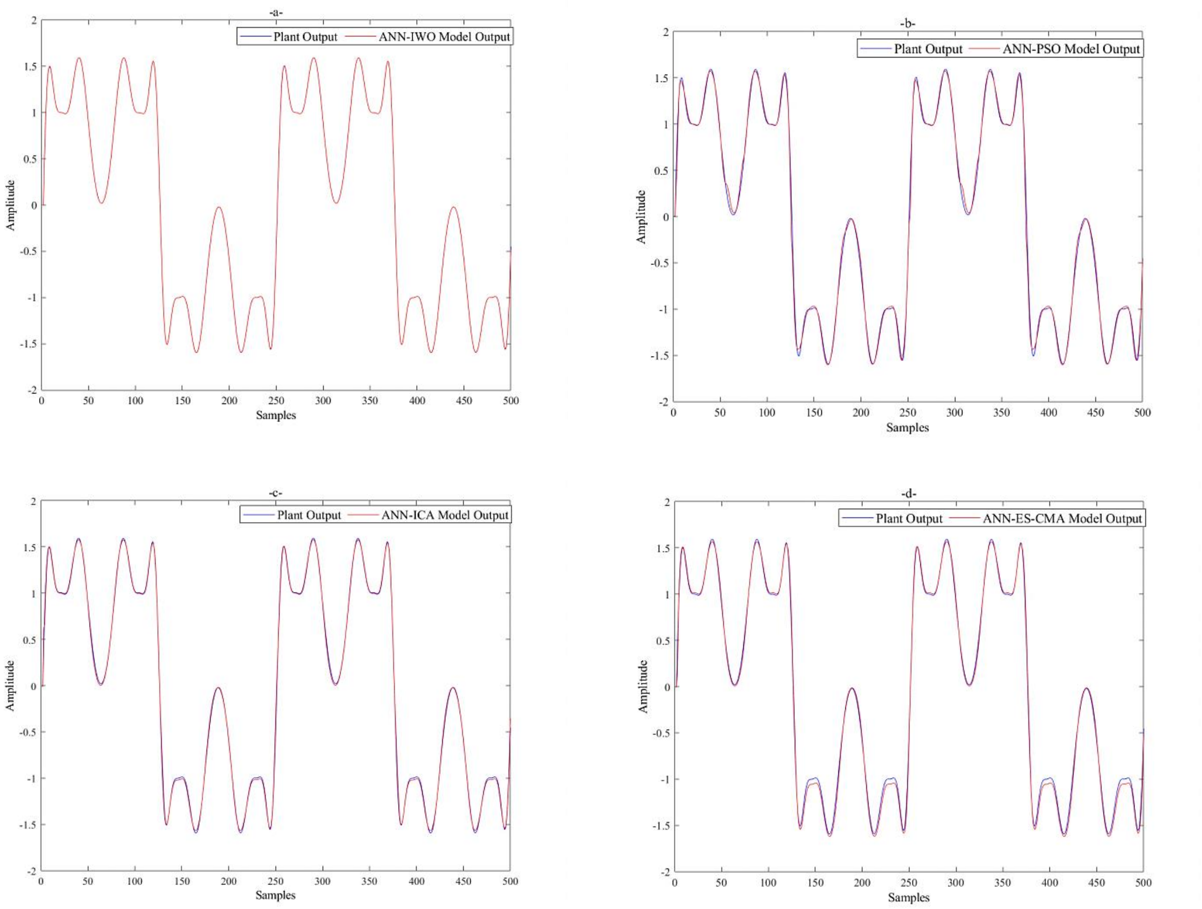

Figure III.12 – ANN-based model for system I: (a) IWO, (b) PSO, (c) ICA, (d) ES-CMA.

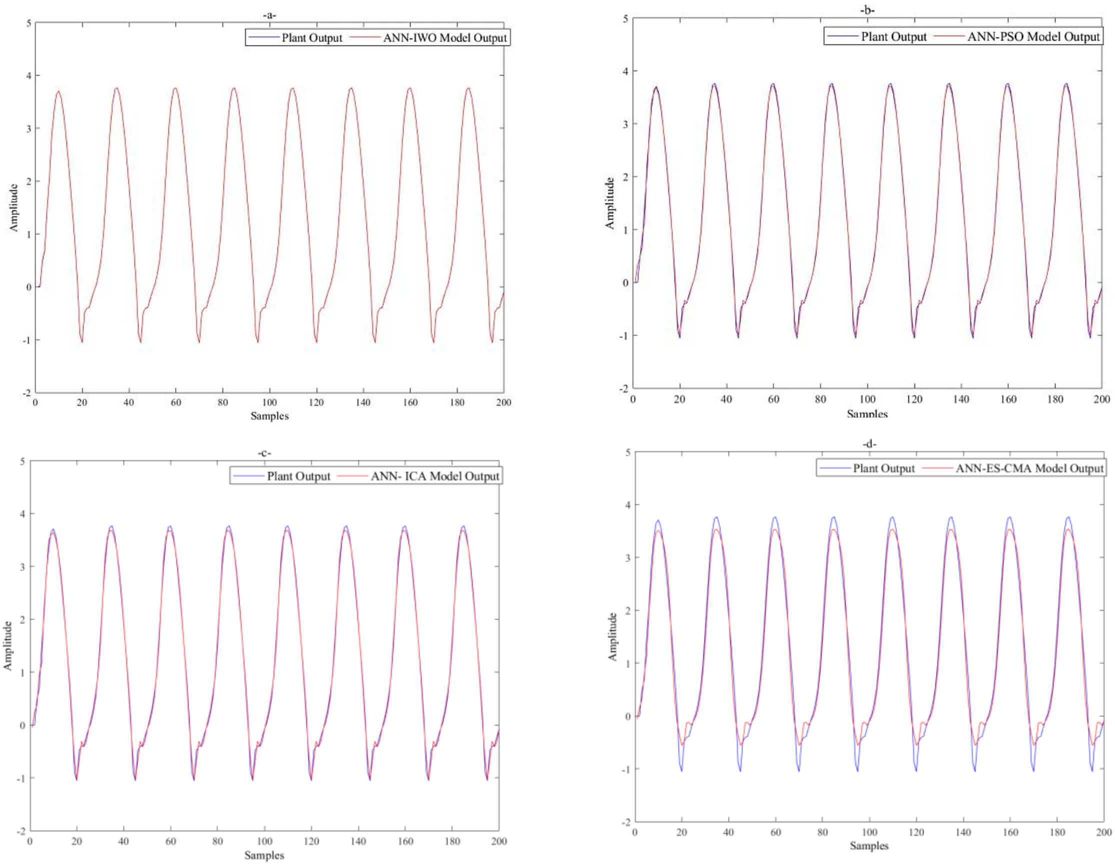

Figure III.13 – ANN-based model for system II: (a) IWO, (b) PSO, (c) ICA, (d) ES-CMA.

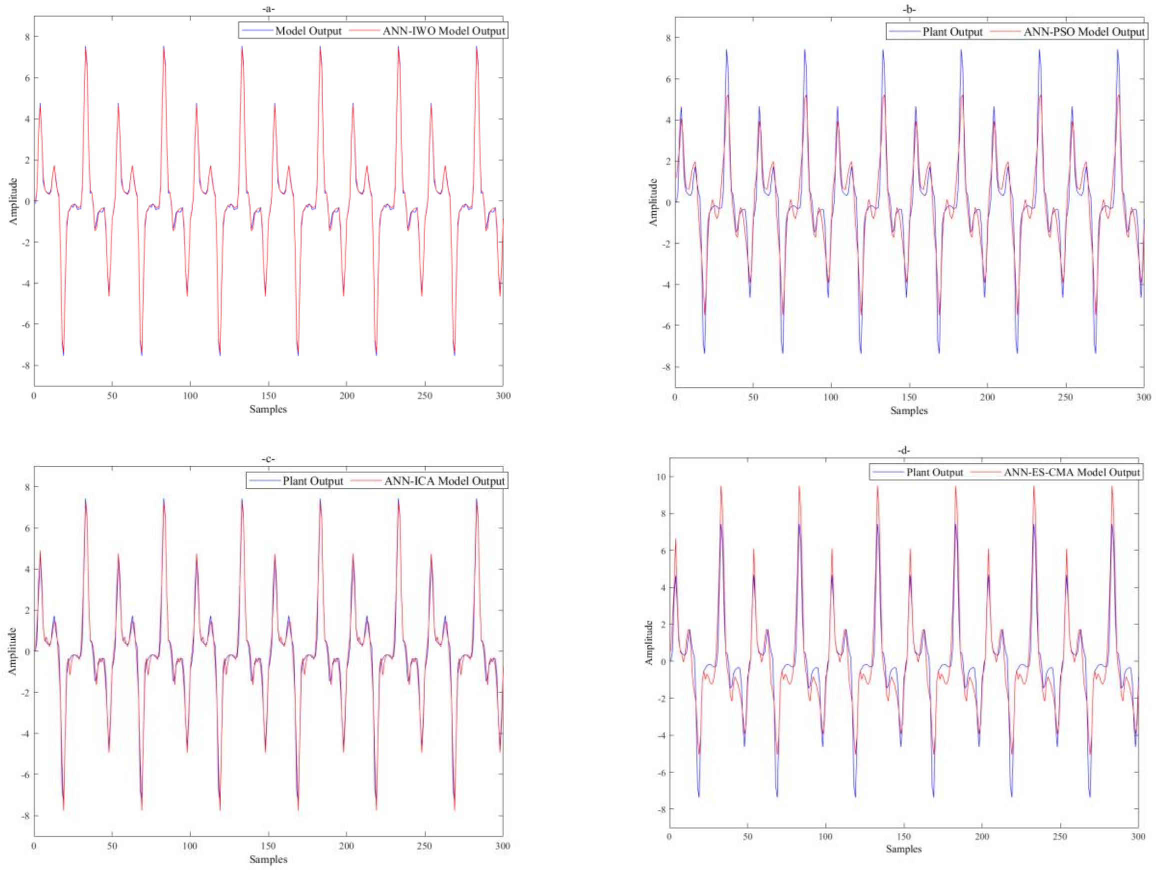

Figure III.14 – ANN-based model for system III: (a) IWO, (b) PSO, (c) ICA, (d) ES-CMA.

Upon examination of every figure, it is apparent that the IWO algorithm performs better than the PSO, ICA, and ES-CMA algorithms. According to the results, the IWO algorithm performs better at modeling and identifying the dynamical systems under consideration than the other optimization methods.

In our current evaluation, we are conducting a numerical assessment of the method's performance utilizing a fitness function known as the mean square error (MSE). To ensure the reliability of our findings, we have carried out 20 independent trials of our method and for each optimization algorithm. (Table III.1) showcases statistical performances metrics such as the worst and best values of the fitness function.

Optimisation algorithm	**Best value**	**Worst value**
IWO	**9.1127e-9**	**8.1002e-7**
PSO	8.2201e-6	5.9317e-4
ICA	4.2138e-6	1.2139e-5
ES-CMA	4.2138e-6	1.7293e-5

Table III.1 – Results of the fitness function (MSE) for twenty independent trials.

According to (Table III.1) and the strategies discussed, the IWO algorithm showed superior performance by attaining the best value during the course of twenty independent trials.

Furthermore, to enhance the depth of our statistical analysis, we have incorporated error bars for parameter optimization. This graphical method visually conveys the variability of the estimated parameters on graphs, providing insights into the uncertainty associated with the estimates and offering a comprehensive view of the accuracy of the parameter values. The error bars for both the primary model parameters and the error process modeling parameters are depicted in (Figures III.15 and III.16) respectively.

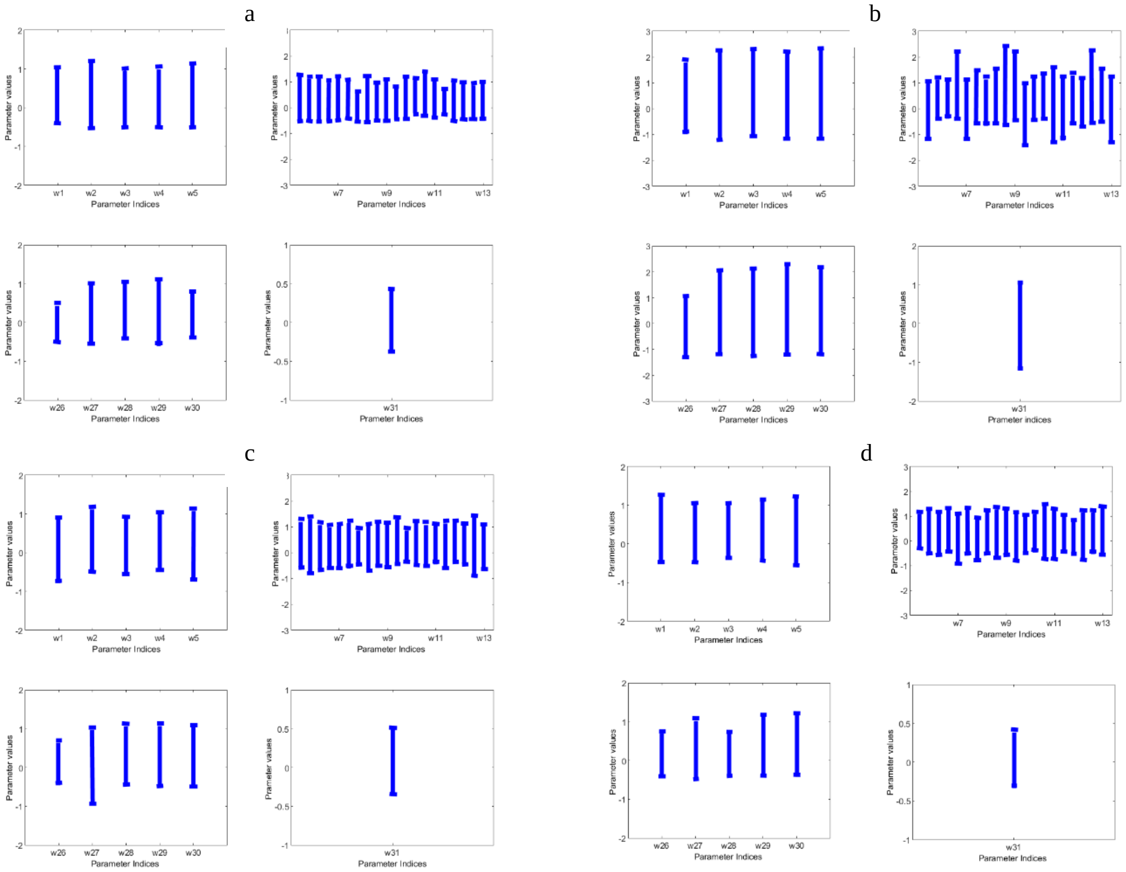

Figure III.15 – Error bars for primary model parameters: (a) IWO, (b) PSO, (c) ICA, (d) ES-CMA.

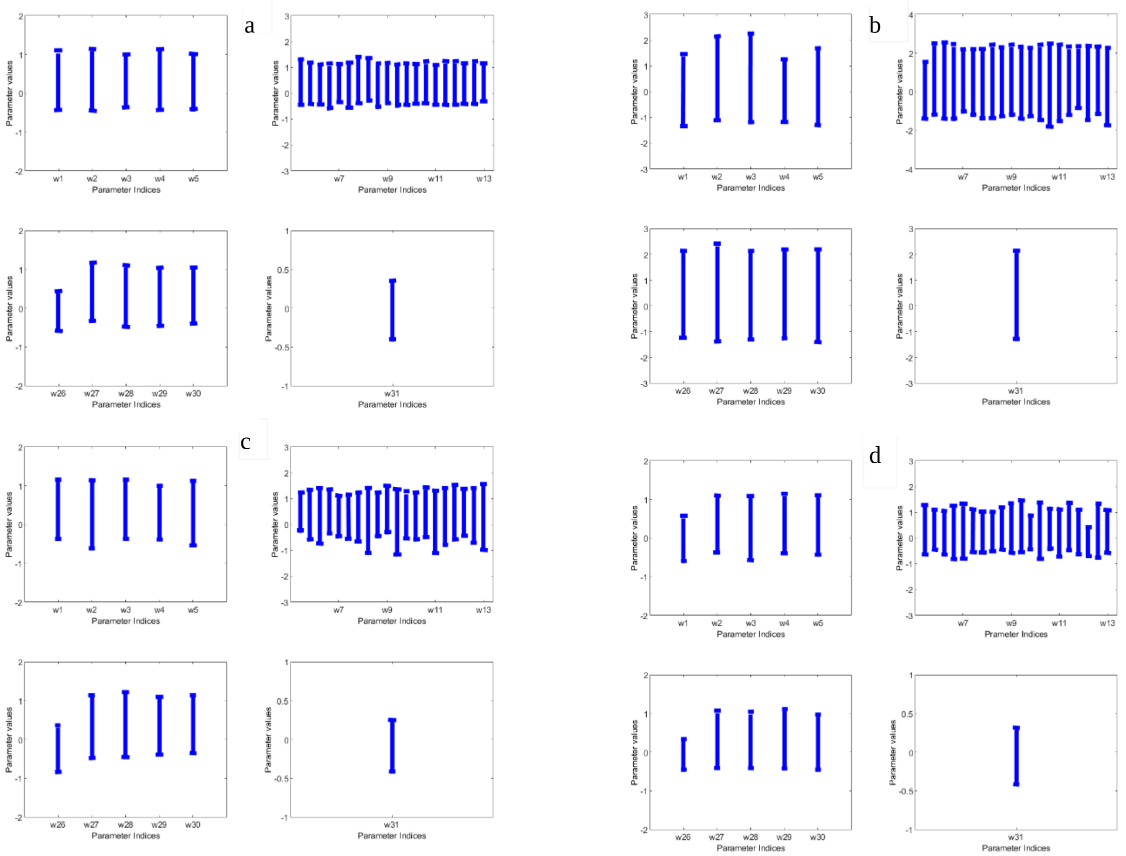

Figure III.16 – Error bars for error process parameters: (a) IWO, (b) PSO, (c) ICA, (d) ES-CMA.

A preliminary visual inspection of these figures reveals that the error bars widths corresponding to the IWO method are the smallest in comparison to the error bars of the PSO, ICA, and CMA-ES methods. This observation indicates that the IWO approach demonstrates greater precision and consistency in parameter optimization.

III.5. Conclusion

In this chapter, we have presented a novel strategy to address common challenges in modeling and identification of dynamical systems and times series. Our approach involves a hybrid Artificial Neural Network Autoregressive Moving Average (ANNARMA) model incorporating metaheuristic algorithms. Our goal with this integration is to successfully address the typical issues that crop up in this field. The provided approach introduces an innovative identification module referred to as the ”error model”. This module is an important addition to the primary model, improving its overall quality and leading to a more precise fit. As a result, the suggested approach yields a higher resolution model with improved accuracy. To achieve optimization in ANN identification, a variety of metaheuristic algorithms such as, ICA, PSO, CMA-ES, and IWO, have been applied. These algorithms play a crucial role in refining the ANN identification process and enhancing its efficiency. The suggested approach's effectiveness is validated via simulation results and comparative studies. The outcomes of these comparisons demonstrate that IWO outperformed the other metaheuristic algorithms utilized in this study, providing the best optimization results. The superiority of IWO further reinforces the credibility and efficiency of the proposed approach in modeling and identification of dynamical systems and time series.

CHAPTER IV

PREPROCESSING OF ECG SIGNALS

Divers types of noise stemming from varied sources can alter ECG signals and make their interpretation difficult, especially for diagnostic purposes. Filtering as a process, seeks to isolate and extract the essential components of a signal drowned in noise. This method is implemented in the specific context of the ECG signal to eliminate the noise surrounding the signal recorded by the electrocardiograph. In this chapter, we focus on the practical application of the methods described in chapter 2 for ECG denoising. Empirical mode decomposition, wavelets and graph empirical mode decomposition are implemented on noisy ECG data. Each method is applied according to a specific defined procedure, taking into account the parameters necessary for their optimal implementation. We also illustrate the results obtained using simulations carried out in the MATLAB environment and, in order to objectively assess the effectiveness of each denoising method applied, we have carried out a comparative study.

Contents

IV.1. Introduction

The presence of artefacts in the ECG signal can complicate diagnosis for specialists. These interferences, which are often relatively strong, may alter the shape of the signal or present with morphologies similar to those of the signal itself. This poses a significant challenge when analyzing ECG waveforms, as accurately identifying abnormalities is crucial for an accurate diagnosis of cardiovascular diseases. To overcome these difficulties, the first essential step in ECG preprocessing is denoising. This procedure aims to eliminate unwanted interference while preserving the essential characteristics of the cardiac signal. By reducing the impact of interference, denoising produces an ECG signal that is smoother and more consistent with actual cardiac activity, making it easier for specialist clinicians to interpret and diagnose.

IV.2. ECG Database

To validate our research work, we chose to use the MIT-BIH PhysioBank database, which provides actual electrocardiogram (ECG) data and is known for its comprehensive collection of ECG signals [213]. This database is made up of forty-eight ambulatory ECG excerpts, each lasting thirty minutes, collected from forty-seven patients. Varied in terms of age and gender, examined at the BIH Arrhythmia Laboratory from 1975 to 1979. At random, twenty-three recordings were chosen from 4000 ambulatory 24-hour ECG recordings collected at Beth Israel Hospital in Boston. From a similar aggregate, an additional 25 recordings were selected to represent less common but medically significant arrhythmias. The recordings were digitized with an 11-bit resolution for each channel at 360 samples per second. Every individual recording underwent annotation by two heart specialists, resolving any discrepancies and generating approximately 110,000 computer-readable reference annotations. This database has a major advantage due to its diversity of cardiac pathologies, thus offering in-depth validation of algorithms on a wide range of ECG signals. In addition, each record is associated with three distinct file types (.dat, .hea, .atr), containing: the digitized ECG signal data, the interpretation parameters, and the locations of the manually marked R-peaks by cardiologists, respectively.

IV.3. Performances evaluation criteria

Assessing the effectiveness of any filtering technique it is essential to analyse its performance and behavior. Various criteria are used to quantify the impact of filtering strategies on ECG signals. These criteria must be reliable and relevant in order to provide an accurate assessment. In our application, we use the mean square error (MSE), Mean absolute error (MAE) and the signal-to-noise ratio (SNR) to evaluate the proposed filtering strategies.

IV.3.1. Mean Squared Error

The most commonly used criterion in the field of signal processing is the mean square error (MSE). This measure evaluates the average difference between the filtered signal and the original signal. In the ECG context, where data precision is critical for medical diagnosis, minimizing MSE ensures that the filtered signal accurately reflects cardiac activity. MSE is calculated as follows [214]:

$$MSE = \frac{1}{N}\sum_{n=1}^{N}(x_n - \hat{x}_n)^2 \qquad \text{(IV.1)}$$

where N is the length of the ECG signal, x_n presents the original signal and $\hat{x}_n$ is the filtered signal.

IV.3.2. Mean absolute error

Mean absolute error (MAE), similar to MSE, is a metric used to evaluate denoising accuracy. The average of the absolute differences between the denoised and the original ECG signal is quantified by this measure. The quality of the denoising increases with decreasing MAE value [215]. The definition of MAE is as follows:

$$MAE = \frac{1}{N}\sum_{n=1}^{N}\left|\left(x_n - \hat{x}_n\right)\right| \qquad \text{(IV.2)}$$

IV.3.3. Signal-to-Noise Ratio

The quality of a signal, particularly in the context of ECG, is often assessed using the ratio between the signal of interest and the noise, known as the ‘’**signal-to-noise ratio**” (SNR). This measure quantifies the presence of noise in the ECG signal. The SNR is generally expressed on a logarithmic scale in decibels (dB) and it is expressed by [215]:

$$SNR = 10\log_{10}\left(\frac{\sum_{n=1}^{N}\left(x_n\right)^2}{\sum_{n=1}^{N}\left(x_n - \hat{x}_n\right)^2}\right) \qquad \text{(IV.3)}$$

where N is the length of the ECG signal, x_n is the original signal and $\hat{x}_n$ is the estimated signal.

IV.4. Denoising of ECG signal using diverse approaches: Description and results

IV.4.1. EMD-based method

IV.4.1.1. Description of the proposed method

The main flowchart of the suggested approach is illustrated in (Figure IV.1). It can be organized into the following three main parts:

- Decomposition of ECG data by EMD.
- Computation of the hurst exponent via Average Wavelet Coefficient method.
- Thresholding and reconstruction.

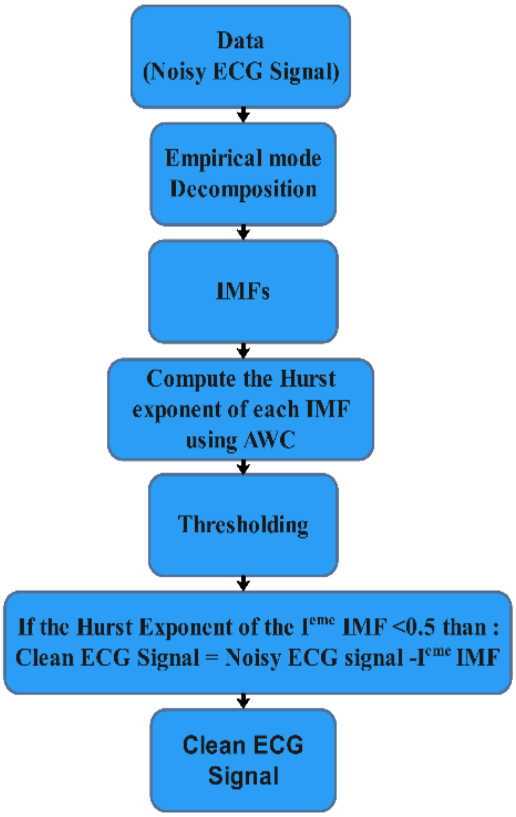

Figure IV.1 – The proposed denoising method’s flowchart.

A. Decomposition of ECG data by EMD

The ECG is decomposed according to the EMD algorithm described in (chapter 2). The initial step in the decomposition process involves extracting the ECG signal's local maxima and minima. Next, using cubic spline interpolation, upper and lower envelopes are obtained from these extrema, and an average envelope is then computed from these envelopes. By subtracting this average envelope from the original signal, a temporary local oscillation will be obtained. If this average is close to zero, the local oscillation is considered to be the first IMF. Otherwise, the process is repeated until this condition is met. After that, the residuals are calculated by subtracting the IMF from the original signal. This process is repeated iteratively on the residuals until the residual becomes a constant, meaning that it no longer satisfies the conditions for an IMF. Finally, the original signal is decomposed into a set of the IMFs and the ultimate residual. The intrinsic components of the ECG signal are extracted during this decomposition process, which facilitates analysis and comprehension of the signal's temporal and frequency characteristics.

B. Computation of the Hurst exponent via Average Wavelet Coefficient method

The AWC approach was first presented in [216] and is based on computing a signal’s coefficients of the wavelet. Following the signal alteration, an average coefficient "fv" for each scale is obtained by averaging the coefficients over the spatial dimension. This approach offers a reliable tool for Hurst exponent estimation. Hurst exponent, symbolized by (H), represents a statistical metric utilized to express how closely a time series' points correlate. Originally, it was developed around 1915 through the hydrologist Harold Edwin Hurst [217], the one suggested a strategy for optimizing the amount of storage capacity at the Nile Dam by taking into account the drought periods and the precipitation amount seen over an extended period. His approach is founded on the principle that plenty of natural occurrences exhibit behavior that may be explained by a biased random process in which observations are influenced by one another, or memory between observations. The values for the Hurst exponent are specified across the 0–1 range. Based on the value of H, any time series can be classified into one of the following three categories:

- $\mathrm{H} = 0.5$ indicates the complete lack of correlation in the series.
- $\mathrm{H} < 0.5$ indicates a time series data where long-term low and high values follow each other in adjacent pairs, suggesting a trend towards anti- persistency. This suggests that a high value is typically succeeded by lower ones, after that there is a tendency to

return to a long-term high value, thus showing an alternation between high and low values over the long term.

- H $>$ 0.5 suggests a time series with consistently positive autocorrelation over the long term. This indicates a higher value is succeeded by a higher one.

Using the AWC approach, the Hurst exponent value can be estimated as follows [217]:

- ✓ Transforming data to the wavelet domain using wavelet transform.
- ✓ Computing the average wavelet coefficient according to the following equation:

$$W[h]_{(a)} = \left\langle \left| W[h](a,b) \right| \right\rangle_b \tag{IV.4}$$

- ✓ The plotting of average wavelet coefficient on a log-log plot, against scale a.
- ✓ Estimating H's value based on the straight line's slope that is appropriate for the $\log\left(W[h](a)\right)$ plot. The slope of this line is: $1/2 + H$

C. Thresholding and reconstruction.

After we decomposing the noisy ECG signal and obtaining the set of IMFs, and we calculated the Hurst exponent of every IMF utilizing AWC approach in order to differentiate between low-frequency and high-frequency IMFs, and due to the fact that the relevant information of the signal is predominantly found within low-frequency IMFs, while noise is mainly located in high-frequency IMFs, whatever IMF of which the Hurst exponent is less than or equal to **0.5** will be considered as noise and eliminated. Ultimately, through reconstructing each of the chosen IMFs, we obtained a clean and noise-free ECG signal.

IV.4.1.2. Simulation results

This section presents the results of simulations and experimental tests carried out to assess the effectiveness of the proposed strategy by applying both the real ECG signal collected from the standard MIT-BIH database mentioned above and the synthetic ECG signal produced with the dynamic model in Cartesian space, as described by the authors in [212]. This dynamic model consists of the three coupled differential equations in the state space, as listed below:

$$\begin{cases} x = a\dot{x} + wy \\ y = a\dot{y} + wx \\ \dot{z} = \sum_{i\in(P,Q,R,S,T)} \frac{A_i w}{b_i^2} \Delta\theta_i \exp\left(-\frac{\Delta\theta_i^2}{2b_i^2}\right) - (z - z_0) \end{cases} \tag{IV.5}$$

where $\alpha = 1 - \sqrt{x^2 + y^2}$, $\Delta\theta_i = (\theta - \theta_i)\bmod(2\pi)$, and $\theta = a\tan 2(y,x)$. Here θ represents the four quadrat arctangent of the real parts of the elements of x and y, with $-\pi \le a\tan 2(y,x) \ge \pi$. Additionally, w is the angular velocity of the trajectory as it moves around the limit cycle. A_i, b_i and θ_i correspond to the amplitude, width and center parameters of the Gaussian (P, Q, R, S and T ECG signal waves). The baseline wander was introduced by coupling the baseline z_0 in (5) to the respiratory frequency f_2 as:

$$z_0 = A * \sin(2\pi f_2 t) \tag{IV.6}$$

where $A = 0.15mv$ and $f_2 = 0.25Hz$

Table (IV.1) provides the parameters' values for (Eq.IV.5).

Index(i)	**P**	**Q**	**R**	**S**	**T**
Time(sec)	-0.2	-0.05	0	0.05	0.3
$\boldsymbol{\theta_i}$ **(rads)**	$-\pi/3$	$-\pi/12$	0	$\pi/12$	$\pi/2$
$\boldsymbol{A_i}$	20	-50	300	-75	75
$\boldsymbol{b_i}$	0.25	0.1	0.1	0.1	0.4

Table IV.1 – Model parameters.

Figure (IV.2) depicts the ECG signal generated by the model (IV.5):

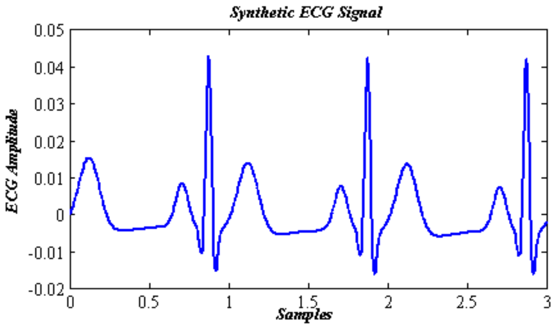

Figure IV.2 – Synthetic ECG Signal.

Gaussian white noise is added to the resulting ECG signal with $SNR_{input} = 5dB$ so as to produce a noisy ECG signal (Figure IV.3).

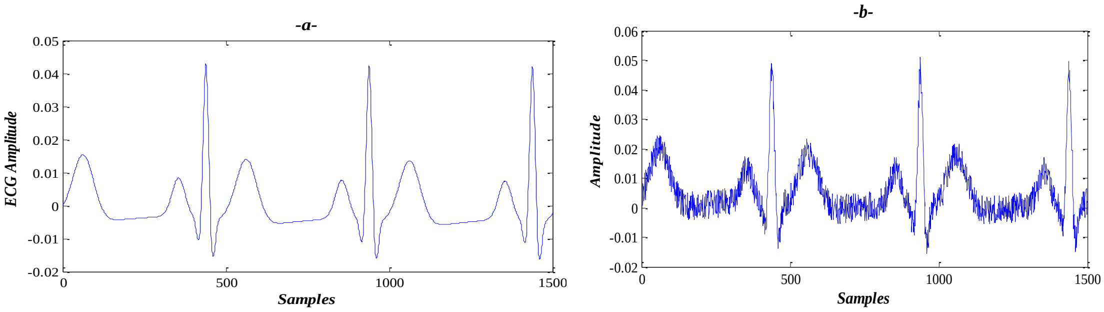

Figure IV.3 – (*a*) An Artificial ECG Signal, (*b*) an ECG signal containing noise with an SNR input of 5 dB.

In what follows, the synthetic noisy ECG signal is denoised utilizing the proposed approach. (Figure IV.4) and (Figure IV.5) depicts the outcomes of denoising.

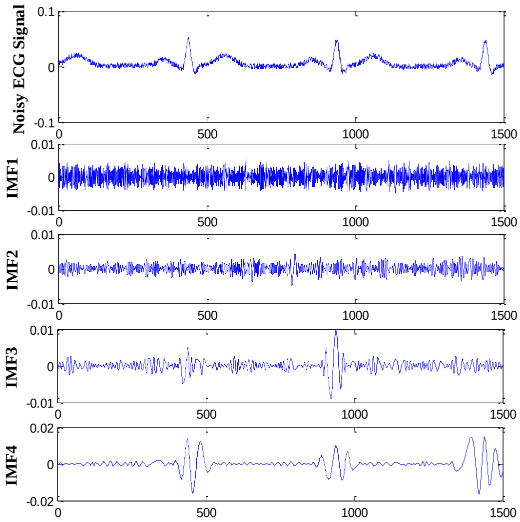

Figure IV.4 – Noisy ECG signal's EMD decomposition.

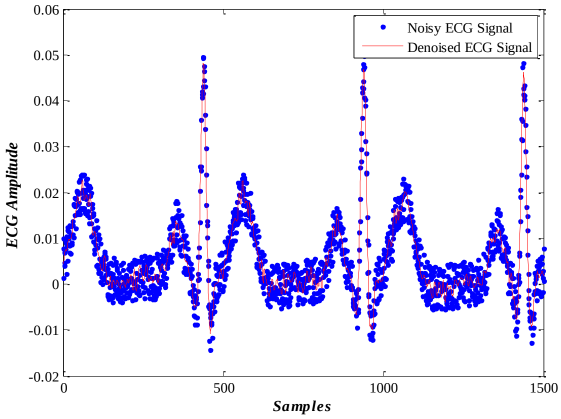

Figure IV.5 – Result of denoising.

By visual inspection, it can be seen that the proposed approach has the potential to give a very good denoising result. To assess the suggested approach's denoising performance quantitatively, $\boldsymbol{MSE}$ (Eq.IV.1) $\boldsymbol{MAE}$ (Eq.IV.2) and $\boldsymbol{SNR}$ (Eq.IV.3) are used as metrics for denoising performance. Performance measures obtained under various levels of noise ($\mathbf{5}\boldsymbol{dB}$, $\mathbf{10}\boldsymbol{dB}$ and $\mathbf{15}\boldsymbol{dB}$) are arranged according to the following table.

Noise intensity	$\boldsymbol{MSE}$	$\boldsymbol{MAE}$	$\boldsymbol{SNR_{output}}$
$\mathbf{5}\boldsymbol{dB}$	0.000007	0.002245	13 dB
$\mathbf{10}\boldsymbol{dB}$	0.000004	0.001655	18 dB
$\mathbf{15}\boldsymbol{dB}$	0.000002	0.001338	22 dB

Table IV.2 – Metrics of performance at different noise levels.

From (Table IV.2), we may observe that, the suggested approach has a lowest MSE value and a highest SNR_{output} across all levels of noise intensity. The findings demonstrate that the suggested method can lessen the impact of noise.

For validation purposes, we utilized an actual ECG signal data (100.dat) (Figure IV.6) to which Gaussian white noise was added. The noisy actual ECG signal's denoising outcomes applying the suggested filter are displayed in (Figure IV.7) as well as (Figure IV.8).

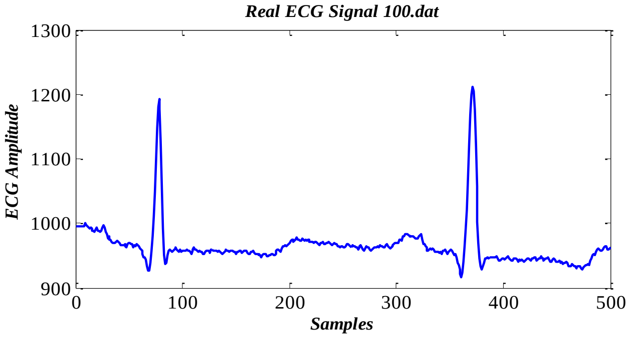

Figure IV.6 – 100.dat real ECG signal.

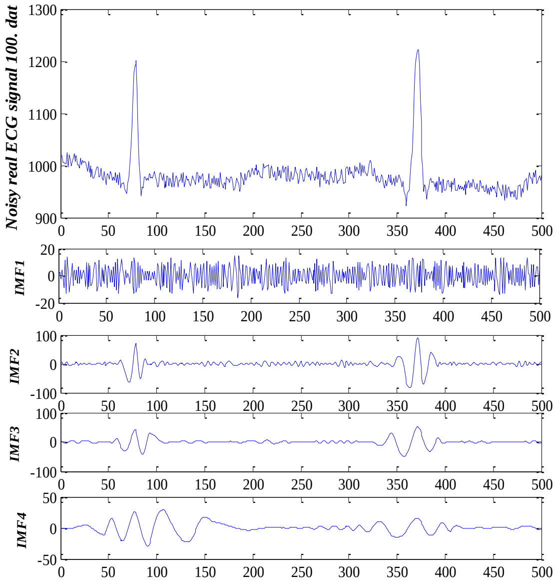

Figure IV.7 – Noisy real ECG signal's EMD decomposition.

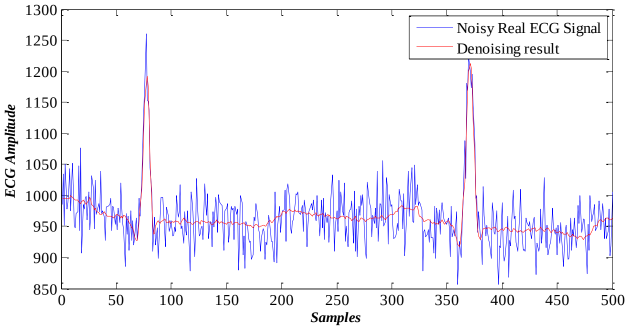

Figure IV.8 – Result of denoising.

The latest findings attest to the suggested approach's capacity to remove noises that could taint an ECG signal.

IV.4.2. Wavelet-based method

IV.4.2.1. Description of the proposed method

In this second experiment, we followed the same procedures as those presented for the first contribution, with the exception that we replaced the EMD decomposition by the wavelet transform (DWT) with four decomposition levels. The choice of the mother wavelet for ECG signals is a crucial decision that can greatly influence the quality of the results obtained. It is often necessary to carry out empirical studies and comparative evaluations to determine the most suitable wavelet for a specific application. Commonly used wavelet types include Daubechies, Symlet, Coiflet and biorthogonal wavelets. In our study, we opted for the second-order Daubechies mother wavelet.

The filtering process begins by decomposing the noisy ECG signal into wavelets and approximations at different resolution scales. Next, the Hurst exponent is calculated for each approximation and detail of the wavelet decomposition. The Hurst exponent gives an idea of the memory of the time series: when the Hurst exponent is close to 1, this indicates that the time series retains long-term information (which is desirable), while a Hurst exponent close to 0 suggests that the time series does not retain long-term information (which is undesirable).. Thus, if the Hurst exponent of each approximation or detail is below a predetermined threshold, that approximation or detail will be removed.

Finally, the clean ECG signal is obtained by summing the approximations and details whose Hurst exponent is greater than the predetermined threshold. This process is summarized in the following block diagram:

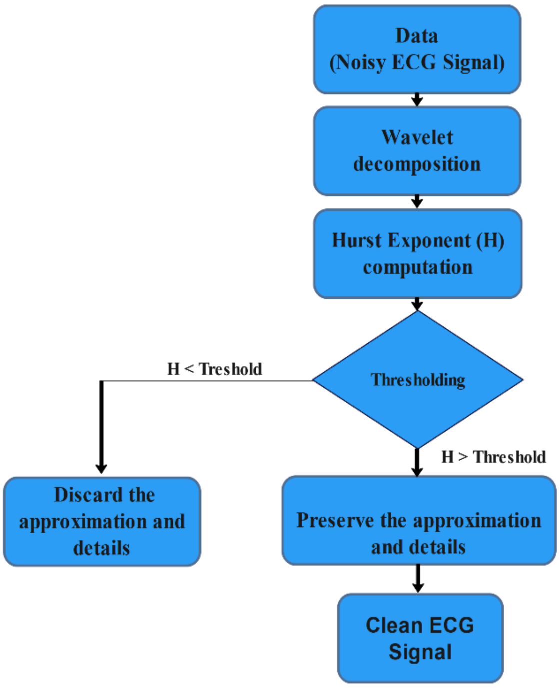

Figure IV.9 – DWT flowchart

IV.4.2.2. Simulation results

The results obtained in the second experiment, are summarized in this section. (Figure IV.10) shows the DWT decomposition of synthetic noisy ECG signal and (Figure IV.11) illustrates the outcomes of denoising.

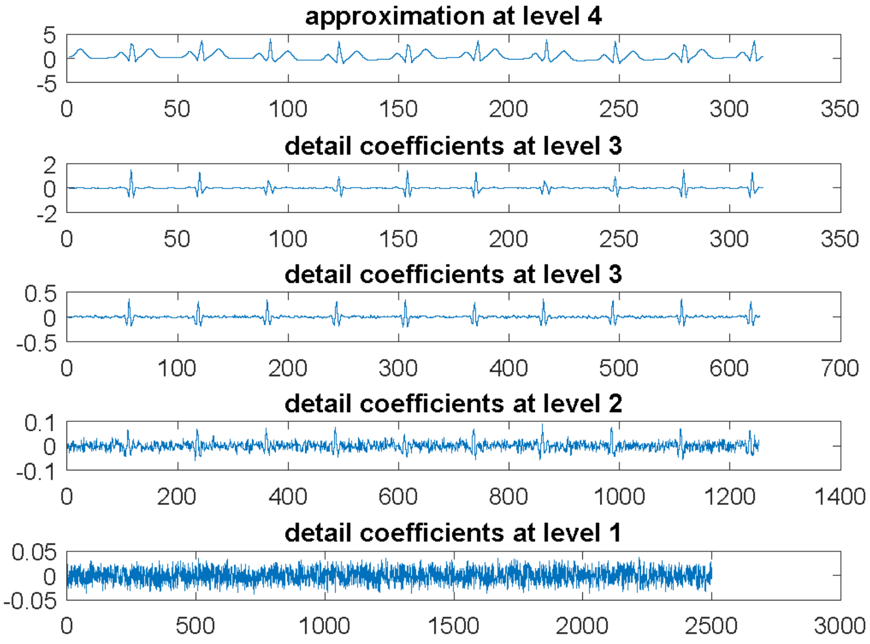

Figure IV.10 – Noisy synthetic ECG signal's DWT decomposition.

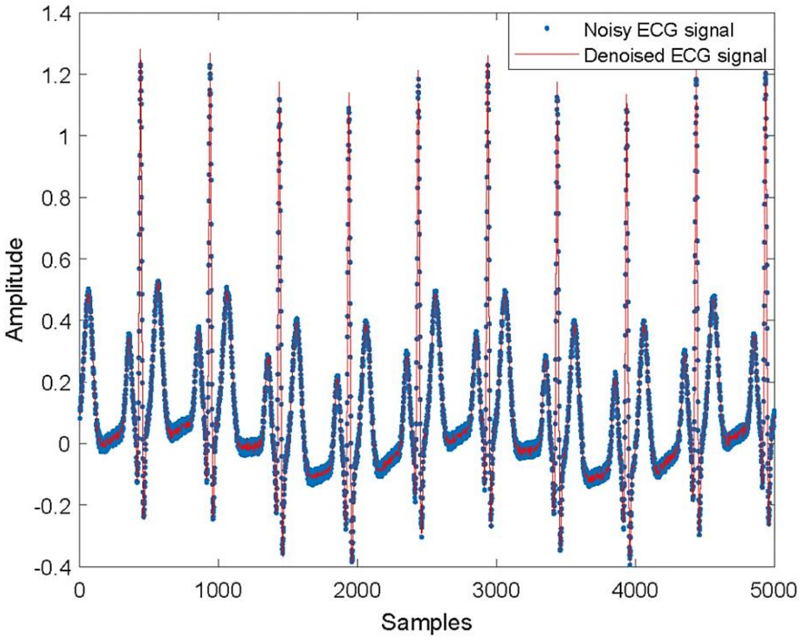

Figure IV.11 – Denoising outcome.

The performance metrics resulting from the second experiment are organized in the subsequent table.

Noise intensity	MSE	MAE	SNR_{output}
$\mathbf{5dB}$	0.016769	0.103217	7 dB
$\mathbf{10dB}$	0.004391	0.052856	12 dB
$\mathbf{15dB}$	0.001467	0.0330502	16 dB

Table IV.3 – Metrics of performance across different levels of noise.

The (Figures IV.12 and IV.13) display the denoising outcomes obtained from the application of the wavelet-based filter method to the noisy real ECG signal.

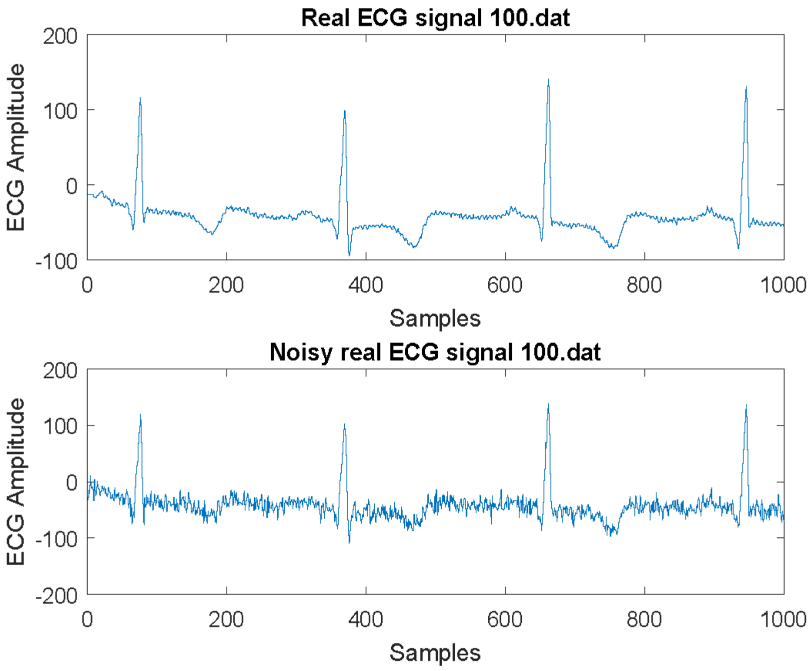

Figure IV.12 – Real ECG signal.

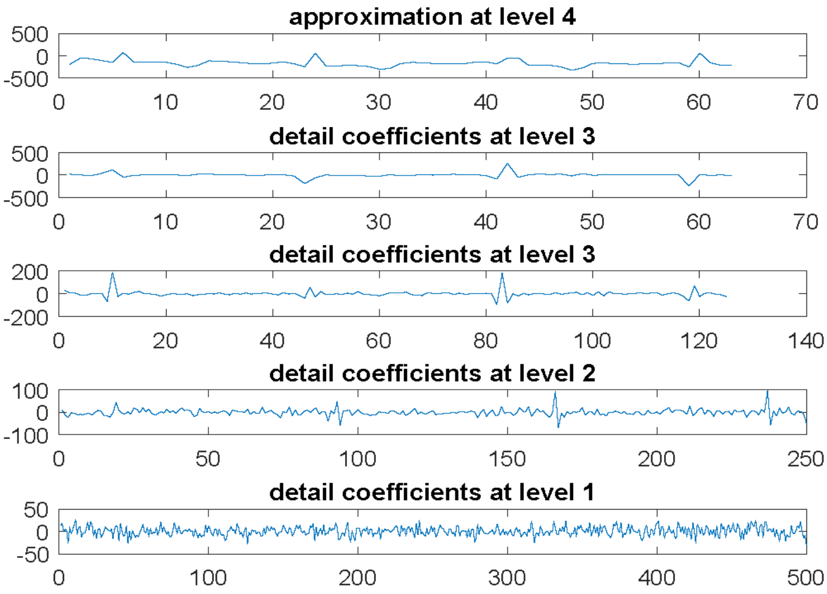

Figure IV.13 – DWT decomposition of noisy reel ECG signal.

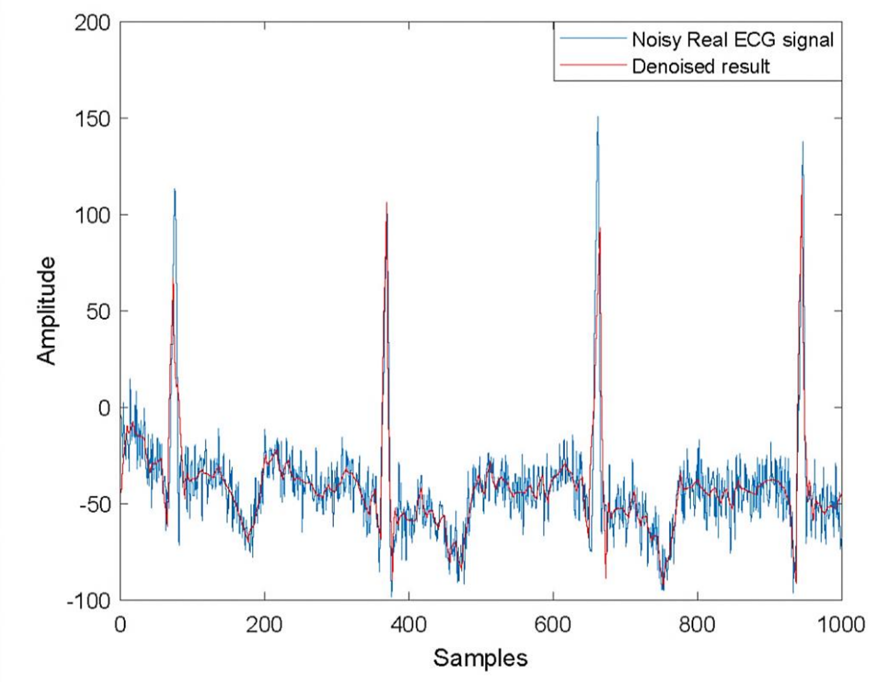

Figure IV.14 – Denoising outcome.

IV.4.3. GEMD- based method

Although EMD is an effective method for analyzing non-linear and non-stationary signals, it does have certain limitations, such as difficulties in determining stopping criteria and interpolation problems. To overcome these challenges and improve analysis accuracy, we have chosen to adopt an advanced variant of EMD: graph-based EMD (GEMD). Although the various results of this technique have not been fully completed within the scope of this thesis, preliminary work is promising.

This technique mainly concerns the graphical representation of ECG signals based on graph theory. A new data interpolation technique is also used in the decomposition process to improve the performance of our proposed method. After decomposition, our filtering method is applied to eliminate the various noises. Then, the signal is reconstructed to obtain a clean and smoother signal.

However, it is important to note that these preliminary results require further confirmation and validation, and will not be detailed in this thesis. This work will be the subject of future research.

IV.4.4. Comparative Study

Based on the key performance indicators (MSE, MAE, and SNR) presented earlier, we compare the performance of the two proposed methods. The outcomes demonstrate the superior efficacy of the EMD-based approach over the DWT-based approach in several criteria for the applications under investigation.

Noise Intensity	***Criteria***	***EMD***	***DWT***
5db	***MSE***	0.000007	0.016769
	MAE	0.002245	0.103217
	SNR	13 dB	7 dB
10db	***MSE***	0.000004	0.004391
	MAE	0.001655	0.052856
	SNR	18 dB	12 dB
15db	***MSE***	0.000002	0.001467
	MAE	0.001338	0.0330502
	SNR	22 dB	16 dB

Table IV.4 – Performance comparison of EMD and DWT across different noise intensities.

From the above table, it is evident that the EMD displays lower MSE and MAE values, indicating better reconstruction accuracy and less reconstruction error compared to the DWT.

With its higher SNR, the EMD appears to have better signal quality compared to noise for DWT.

IV.5. Conclusion

In this chapter, we have explored an exciting topic: the pre-processing of electrocardiographic signals, which is crucial to ensuring accurate results in the diagnostic process. Several approaches have been developed for filtering these signals. The first approach is based on the use of EMD and the Average Wavelet Coefficient Method and the notion of the Hurst exponent. The simulation outcomes show that this approach is computationally efficient and performs well under diverse noise power circumstances without changing the shape of the ECG signal. The second approach developed was based on wavelets, wavelet decomposition and the Hurst exponent. The results obtained demonstrated the effectiveness of this method in improving signal quality. A third approach has also been mentioned, but will be the subject of future work: graph-based EMD. In what follows, we present the general conclusion of this thesis and propose directions for further research.

CONCLUSION

Through this thesis our objective is to propose a new framework for the intelligent modeling and identification of dynamical systems and time series. As well as, the processing and analysis of time series, more specifically the denoising of ECG signals. In the course of this research, several significant contributions have been made by combining existing techniques with new innovative structures. This work has focused on the modeling and identification of dynamical systems and time series, as well as the processing and analysis of these series, in particular the denoising of electrocardiogram (ECG) signals. Comparative studies were carried out to illustrate the efficacy and superiority of the approach suggested over reference techniques.

The first chapter is divided into two parts. The first part provides a general concept of dynamical systems and time series. Examples of time series are also presented, in particular the ECG time series, its principle, and the various noises that can affect ECG signals and distort diagnosis. The second part presents the state of the art of the various studies carried out in these fields.

In the second chapter, the theoretical and fundamental aspects of the various techniques used are detailed. This includes the basic notions of artificial neural networks (ANN), as well as various metaheuristic algorithms such as IWO, PSO, ICA and CEMA-ES and ECG signal decomposition and filtering techniques such as EMD and its variants, wavelets and Kalman filter and its extension.

ANN and metaheuristic algorithms are used to create a novel framework for modeling and identifying dynamical systems and are presented in the third chapter. This structure consists of two interconnected sub-models arranged in parallel. The primary model, which is the first sub-model, is made for the dynamical system or time series that is being studied. It has lower resolution and standard parameters. A second sub-model, known as the error model, is created in order to get around the resolution constraint and increase accuracy. The disparities between the primary model and the time series data or non-linear dynamical system are captured by this error model. Through the use of metaheuristic algorithms, the parameters of this structure have been optimized, showcasing its superiority and efficiency in the identification and modeling of time series and dynamical systems.

Chapter four focuses on the preprocessing of ECG signals. A first contribution proposes a technique based on EMD. A second contribution is based on the use of wavelets, while the third contribution will be the focus of our future work.

Future Work

As mentioned above in chapter (04), in addition to the two filtering techniques we have explored in our study, a promising new approach could be the application of graph theory to ECG signal processing. Graph theory offers an innovative way of modeling and analyzing data, and could provide significant advantages for ECG signal filtering, to which end we have developed a new filtering technique based on the use of a graph-adaptive EMD variant named GEMD defined in chapter (02). This method could offer significant advantages in terms of noise suppression and signal quality improvement. The outcomes of this technique will be presented in our future work.

Our future work will not be limited to this technique alone. It would be beneficial to investigate various additional denoising techniques and merge different methods to further improve ECG signal quality. The exploration of advanced techniques, such as deep neural networks and complex statistical models, could also contribute to improving these processes.

REFERENCES

[1] Astrom, K. J., & Murray, R. (2021). *Feedback systems: an introduction for scientists and engineers*. Princeton university press.

[2] Shumway, R. H., Stoffer, D. S., & Stoffer, D. S. (2000). *Time series analysis and its applications* (Vol. 3, p. 4). New-York: springer.

[3] Irwin, M., & Wang, Z. (2017). Dynamic systems modeling. *The International Encyclopedia of Communication Research Methods*, 1-12.

[4] Upadhyay, R. K., & Iyengar, S. R. (2013). *Introduction to mathematical modeling and chaotic dynamics*. CRC press.

[5] Castillo, O., & Melin, P. (2012). *Soft computing for control of non-linear dynamical systems* (Vol. 63). Physica.

[6] Quaranta, G., Lacarbonara, W., & Masri, S. F. (2020). A review on computational intelligence for identification of nonlinear dynamical systems. Nonlinear Dynamics, 99(2), 1709-1761.

[7] Yegnanarayana, B. (2009). Artificial neural networks. PHI Learning Pvt. Ltd.

[8] Qiao, J. F., & Han, H. G. (2012). Identification and modeling of nonlinear dynamical systems using a novel self-organizing RBF-based approach. Automatica, vol. 48, no 8, p. 1729-1734.

[9] YU, H., XIE, T., Paszczynski, S., & Wilamowski, B. M. (2011). Advantages of radial basis function networks for dynamic system design. IEEE Transactions on Industrial Electronics, vol. 58, n° 12, p. 5438-5450.

[10] Jinfeng, L., Fucai, L., & Yaxue, R. (2020). Fuzzy identification of nonlinear dynamic system based on input variable selection and particle swarm optimization parameter optimization. IEEE Access, vol. 8, p. 220557- 220569.

[11] Rajasekaran, S. & Pai, G. V. (2011). Neural networks, Fuzzy logic and Genetic algorithms, PHI Learning Private Limited. IUPESM WC2015 Full Paper Submission Template.

[12] Abraham, A. (2001). Neuro fuzzy systems: State-of-the-art modeling techniques. In: Connectionist Models of Neurons, Learning Processes, and Artificial Intelligence: 6th International Work-Conference on Artificial and Natural Neural Networks, IWANN 2001 Granada, Spain, June 13–15, 2001 Proceedings, Part 1 6. Springer Berlin Heidelberg, p. 269-276.

[13] Biagetti, G., Crippa, P., Falaschetti, L., & Turchetti, C. (2019). A Machine Learning Approach to the Identification of Dynamical Nonlinear Systems. In : 2019 27th European Signal Processing Conference (EUSIPCO). IEEE, p. 1-5.

[14] Rajendra, P. & Brahmajirao, V. (2000). Modeling of dynamical systems through deep learning. *Biophysical Reviews*, 12(6), 1311-1320.

[15] Chatterjee, S., Thakur, R. S., Yadav, R. N., Gupta, L., & Raghuvanshi, D. K. (2020). Review of noise removal techniques in ECG signals. *IET Signal Processing*, 14(9), 569-590.

[16] Friesen, G. M., Jannett, T. C., Jadallah, M. A., Yates, S. L., Quint, S. R., & Nagle, H. T. (1990). A comparison of the noise sensitivity of nine QRS detection algorithms. *IEEE Transactions on biomedical engineering*, 37(1), 85-98.

[17] Kluever, C. A. (2020). *Dynamic systems: modeling, simulation, and control*. John Wiley & Sons.

[18] Strogatz, S. H. (1996). Nonlinear dynamics and chaos.

[19] Arrowsmith, D. K., & Place, C. M. (1990). *An introduction to dynamical systems*. Cambridge university press.

[20] Luenberger, D. G. (1979). *Dynamic Systems* (p. 0007). J. Wiley Sons.

[21] Pavé, A. (2012). Modélisation des systèmes vivants. *Lavoisier (Hermes Science Publications),* 633 p.

[22] Close, C. M., Frederick, D. K., & Newell, J. C. (2001). *Modeling and analysis of dynamic systems*. John Wiley & Sons.

[23] Sokolowski, J. A., & Banks, C. M. (2010). *Modeling and simulation fundamentals: theoretical underpinnings and practical domains*. John Wiley & Sons.

[24] A. Borrie (1992). « Stochastic System for Engineering, Modelling, Estimation and Control », Prentice hall.

[25] Shumway, R. H., Stoffer, D. S., & Stoffer, D. S. (2000). Time series analysis and its applications (Vol. 3). New-York: springer.

[26] Musa, Y. & Jibrin, S. (2021). Time Series Analysis An Introduction.

[27] Van Bellegem, S. (2014). Introduction à l'économétrie (Dehon-Van Bellegem). *STAT-S, 308*.

[28] Cryer, J. D., & Chan, K. S. (2008). Fundamental concepts. *Time Series Analysis: With Applications in R*, 11-26.

[29] Loudjani, N. (2022). *Développement d'algorithmes pour l'analyse des séries temporelles des données de production d'eau potable* (Doctoral dissertation, Faculté des Sciences et de la technologie).

[30] Ondo, J.-C. (2002). Étude Comparative des Tests de Stationnarité. Ph.D. Thesis, Québec University, Québec, QC, Canada, p. 288.

[31] Hamilton, J. D. (2020). *Time series analysis*. Princeton university press.

[32] Avram, F. Series temporelles, Régression, Interpolation and Géostatistique.

[33] Hamilton, J. D. (1994). Time Series Analysis, volume 3. Princeton Univ Pr.

[34] Box, G. E. P., Jenkins, G. M. et Reinsel, G. C. (1994). Time Series Analysis : Forecasting and Control (Wiley Series in Probability and Statistics). Wiley.

[35] Chan, K. S., & Cryer, J. D. (2008). *Time series analysis with applications in R*. springer publication.

[36] Brockwell, P. J., & Davis, R. A. (Eds.). (2002). *Introduction to time series and forecasting*. New York, NY: Springer New-York.

[37] Chatfield, C. (2000). *Time-series forecasting*. Chapman and Hall/CRC.

[38] Jacques, J. (2013). Introduction aux séries temporelles. *Polytech Lille*, 26.

[39] Rhoades, R. A., & Bell, D. R. (Eds.). (2012). *Medical phisiology: Principles for clinical medicine*. Lippincott Williams & Wilkins.

[40] Williams, P.L., Warwick, R. (eds.) (1989). *Gray's Anatomy*, 37th ed., 1598 pp. Churchill Livingstone, Edinburgh.

[41] https://quizlet.com/122412654/pbs-41-heart-structure-study-guide-flash-cards/?x=1jqt (Figure of Heart Structure)

[42] Berne, R. M., & Levy, M. (1992). Electrical activity of the heart. *Cardiovascular physiology*, 40.

[43] Antzelevitch, C., & Burashnikov, A. (2011). Overview of basic mechanisms of cardiac arrhythmia. Cardiac electrophysiology clinics, 3(1), 23-45.

[44] Malmivuo J. & Plonsey R. (1995). Bioelectromagnetism Principles and Applications of Bioelectric and Biomagnetic Fields pp 207-208.

[45] Wasimuddin, M., Elleithy, K., Abuzneid, A. S., Faezipour, M., & Abuzaghleh, O. (2020). Stages-based ECG signal analysis from traditional signal processing to machine learning approaches: A survey. IEEE Access, 8, 177782-177803.

[46] Wagner, G. S. (2001). *Marriott's practical electrocardiography*. Lippincott Williams & Wilkins.

[47] Garcia, T. B. (2015). *12-lead ECG: The art of interpretation*. Jones & Bartlett Learning.

[48] Kusumoto, F. (2020). *ECG interpretation: from pathophysiology to clinical application*. Springer Nature.

[49] Lilly, L. S. (2012). *Pathophysiology of heart disease: a collaborative project of medical students and faculty*. Lippincott Williams & Wilkins.

[50] Tada, Y., Amano, Y., Sato, T., Saito, S. & Inoue, M. (2015). A smart shirt made with conductive ink and conductive foam for the measure-ment of electrocardiogram signals with unipolar precordial leads. Fiber 3, 463-477.

[51] Rangayyan, R. M. (2015). *Biomedical signal analysis*. John Wiley & Sons.

[52] Tanji Jr, A.K., de Brito M.A., Alves M.G., Garcia R.C., Chen G.L., Ama N. (Jan 2021). Improved Noise Cancelling Algorithm for Electrocardiogram Based on Moving Average Adaptive Filter. Electronics. 10(19) : 2366.

[53] Kannathal, N. U. Acharya, R., Joseph, P., Min, L. C., Suri, J. S. (2007). Analysis of electrocardiograms, Advances in cardiac signal processing, (eds). U. R. Acharya, J. S. Suri, J. A. E. Spaan, and S. M. Krishnan, springer, Berlin, pp. 55-81.

[54] D’Aloia, M., Longo, A., & Rizzi, M. (2019). Noisy ECG signal analysis for automatic peak detection. Information 10. 2 : 35.

[55] Lee, C. H., & Teng, C. C. (2000). Identification and control of dynamic systems using recurrent fuzzy neural networks. IEEE Transactions on fuzzy systems, 8(4), 349-366.

[56] Kang, S. J., Woo, C. H., Hwang, H. S., & Woo, K. B. (2000). Evolutionary design of fuzzy rule base for nonlinear system modeling and control. *IEEE Transactions on Fuzzy Systems*, *8*(1), 37-45.

[57] Juang, C. F. (2002). A TSK-type recurrent fuzzy network for dynamic systems processing by neural network and genetic algorithms. IEEE Transactions on Fuzzy Systems, 10(2), 155-170.

[58] Kim, J., Suga, Y., & Won, S. (2006). A new approach to fuzzy modeling of nonlinear dynamic systems with noise: Relevance vector learning mechanism. IEEE Transactions on Fuzzy Systems, 14 (2), 222-231.

[59] Chafaa, K., Ghanai, M., & Benmahammed, K. (2007). Fuzzy modelling using Kalman filter. *IET Control Theory & Applications*, *1*(1), 58-64.

[60] Singh, M., Srivastava, S., Gupta, J. R. P., & Handmandlu, M. (2007). Identification and control of a nonlinear system using neural networks by extracting the system dynamics. *IETE journal of research*, 53 (1), 43-50.

[61] Zhao, H., & Zhang, J. (2009). Nonlinear dynamic system identification using pipelined functional link artificial recurrent neural network. Neurocomputing, 72(13-15), 3046-3054.

[62] Juang, C. F., & Hsieh, C. D. (2010). A locally recurrent fuzzy neural network with support vector regression for dynamic-system modeling. IEEE Transactions on Fuzzy Systems, 18(2) : 261-273.

[63] Lin, Y. Y., Chang, J. Y., & Lin, C. T. (2012). Identification and prediction of dynamic systems using an interactively recurrent self-evolving fuzzy neural network. IEEE Transactions on Neural Networks and Learning Systems, 24 (2), 310-321.

[64] Cavuslu, M. A., Karakuzu, C., & Karakaya, F. (2012). Neural identification of dynamic systems on FPGA with improved PSO learning. *Applied Soft Computing*, 12(9), 2707-2718.

[65] Han, H., Wu, X. L., & Qiao, J. F. (2013). Nonlinear systems modeling based on self-organizing fuzzy-neural-network with adaptive computation algorithm. *IEEE transactions on cybernetics*, 44 (4), 554-564.

[66] Rios, R. A., & De Mello, R. F. (2013). Improving time series modeling by decomposing and analyzing stochastic and deterministic influences. Signal Processing, 93(11), 3001-3013.

[67] Yilmaz, S., & Oysal, Y. (2014). Nonlinear system modeling with dynamic adaptive neuro-fuzzy inference system. In *2014 IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA) Proceedings* (pp. 205-211). IEEE.

[68] Kayacan, E., Kayacan, E., & Khanesar, M. A. (2014). Identification of nonlinear dynamic systems using type-2 fuzzy neural networks—A novel learning algorithm and a comparative study. *IEEE Transactions on Industrial Electronics*, 62(3), 1716-1724.

[69] Honggui, H., Xiao-Long, W., & Jun-Fei, Q. (2014). Nonlinear systems modeling based on self-organizing fuzzy-neural-network with adaptive computation algorithm. IEEE transactions on cybernetics, 44(4): 554-564.

[70] Loussifi, H., Nouri, K., & Braiek, N. B. (2016). A new efficient hybrid intelligent method for nonlinear dynamical systems identification: The Wavelet Kernel Fuzzy Neural Network. *Communications in Nonlinear Science and Numerical Simulation*, 32, 10-30.

[71] Han, H.G., Lin, Z.L. & Qiao, J.F. (2017). Modeling of nonlinear systems using the self-organizing fuzzy neural network with adaptive gradient algorithm, Neurocomputing, Vol. 266, pp. 566-578.

[72] Wang, Y. (2017). A new concept using LSTM Neural Networks for dynamic system identification. In *2017 American control conference (ACC)* (pp. 5324-5329). IEEE.

[73] Marzi, H., Haj Darwish, A., & Helfawi, H. (2017). Training ANFIS using the enhanced bees algorithm and least squares estimation. Intelligent Automation & Soft Computing, 23(2) : 227-234.

[74] Han, H. G., Lin, Z. L., & Qiao J. F. (2017). Modeling of nonlinear systems using the self-organizing fuzzy neural network with adaptive gradient algorithm. Neurocomputing.

[75] Li, Q., Lin, Y., Lin, R. C., & Meng, H. F. (2019). Using self-constructing recurrent fuzzy neural networks for identification of nonlinear dynamic systems. *International Journal of Modelling, Identification and Control*, 33 (4), 378-386.

[76] Dass, A., & Srivastava, S. (2019). Identification and control of dynamical systems using different architectures of recurrent fuzzy system. *ISA transactions*, 85, 107-118.

[77] Samanta, S., Suresh, S., Senthilnath, J., & Sundararajan, N. (2019). A new neuro-fuzzy inference system with dynamic neurons (nfis-dn) for system identification and time series forecasting. *Applied Soft Computing*, 82, 105567.

[78] Han, R., Wang, R., & Zeng, G. (2020). Identification of dynamical systems using a broad neural network and particle swarm optimization. *IEEE Access*, *8*, 132592-132602.

[79] Niven, R. K., Mohammad-Djafari, A., Cordier, L., Abel, M., & Quade, M. (2020). Bayesian identification of dynamical systems. Multidisciplinary Digital Publishing Institute Proceedings, 33 (1), 33.

[80] Mao, W. L., Suprapto, Hung, C. W., & Chang, T. W. (2021). Nonlinear system identification using BBO-based multilayer perceptron network method. Microsystem Technologies, 27, 1497-1506.

[81] Ahmed, S. M., Al-Shrouf, A., & Abo-Zahhad, M. (2000). ECG data compression using optimal non-orthogonal wavelet transform. Medical engineering & physics, 22(1), 39-46.

[82] Batista, L. V., Melcher, E. U. K., & Carvalho, L. C. (2001). Compression of ECG signals by optimized quantization of discrete cosine transform coefficients. *Medical engineering & physics*, 23 (2), 127-134.

[83] Wei, J. J., Chang, C. J., Chou, N. K., & Jan, G. J. (2001). ECG data compression using truncated singular value decomposition. *IEEE Transactions on Information Technology in Biomedicine*, 5 (4), 290-299.

[84] Rajoub, B. A. (2002). An efficient coding algorithm for the compression of ECG signals using the wavelet transform. IEEE transactions on biomedical engineering, 49 (4), 355-362.

[85] Al-Shrouf, A., Abo-Zahhad, M., & Ahmed, S. M. (2003). A novel compression algorithm for electrocardiogram signals based on the linear prediction of the wavelet coefficients. Digital Signal Processing, 13 (4), 604-622.

[86] Blanco-Velasco, M., Cruz-Roldán, F., López-Ferreras, F., Bravo-Santos, A., & Martinez-Munoz, D. (2004). A low computational complexity algorithm for ECG signal compression. Medical engineering & physics, 26 (7), 553-568.

[87] Borsali, R., Naït-Ali, A., & Lemoine, J. (2004). ECG compression using an ensemble polynomial modeling: Comparison with the DCT based technique. Cardiovascular Engineering, 4 (3), 237-244.

[88] Ku, C. T., Wang, H. S., Hung, K. C., & Hung, Y. S. (2006). A novel ECG data compression method based on nonrecursive discrete periodized wavelet transform. *IEEE Transactions on Biomedical Engineering*, 53 (12), 2577-2583.

[89] Ouamri, A., & Naït-Ali, A. (2007). ECG compression method using Lorentzian functions model. *Digital Signal Processing*, 17 (1), 319-326.

[90] Sayadi, O., & Shamsollahi, M. B. (2008). ECG denoising and compression using a modified extended Kalman filter structure. IEEE transactions on biomedical engineering, 55 (9), 2240-2248.

[91] Ahmed, S. M., Al-Ajlouni, A. F., Abo-Zahhad, M., & Harb, B. (2009). ECG signal compression using combined modified discrete cosine and discrete wavelet transforms. *Journal of medical engineering & technology*, 33 (1), 1-8.

[92] Abo-Zahhad, M., Ahmed, S. M., & Zakaria, A. (2011). ECG signal compression technique based on discrete wavelet transform and QRS-complex estimation. *Signal Processing–An International Journal (SPIJ)*, 4 (2), 138.

[93] Bendifallah, A., Benzid, R., & Boulemden, M. (2011). Improved ECG compression method using discrete cosine transform. Electronics letters, 47 (2), 1.

[94] Ranjeet, K., Kuamr, A., & Pandey, R. K. (2012). ECG signal compression using optimum wavelet filter bank based on kaiser window. *Procedia engineering*, 38, 2889-2902.

[95] Abo-Zahhad, M., Al-Ajlouni, A. F., Ahmed, S. M., & Schilling, R. J. (2013). A new algorithm for the compression of ECG signals based on mother wavelet parameterization and best-threshold levels selection. Digital Signal Processing, 23 (3), 1002-1011.

[96] Kanhe, R. K., & Hamde, S. T. (2014). Wavelet-based compression of ECG signals. *International Journal of Biomedical Engineering and Technology*, 14 (4), 297-314.

[97] Kumar, R., Kumar, A., & Singh, G. K. (2015). Electrocardiogram signal compression based on singular value decomposition (SVD) and adaptive scanning wavelet difference reduction (ASWDR) technique. AEU-International Journal of Electronics and Communications, 69 (12), 1810-1822.

[98] Wang, X., Chen, Z., Luo, J., Meng, J., & Xu, Y. (2016). ECG compression based on combining of EMD and wavelet transform. Electronics Letters, 52 (19), 1588-1590.

[99] Soussi, I., & Ouslim, M. (2016). A New compression scheme based on adaptive vector quantization and singular value decomposition. *International review on computers and software*, 11 (5), 446-455.

[100] Fathi, A., & Faraji-kheirabadi, F. (2016). ECG compression method based on adaptive quantization of main wavelet packet subbands. *Signal, Image and Video Processing*, 10, 1433-1440.

[101] Bensegueni, S., & Bennia, A. (2016). ECG signal compression using a sinusoidal transformation of principal components. International Journal of Software Engineering and Its Applications, 10 (1), 59-68.

[102] Zhang, B., Zhao, J., Chen, X., & Wu, J. (2017). ECG data compression using a neural network model based on multi-objective optimization. *PloS one*, 12 (10), e 0182500.

[103] Jha, C. K., & Kolekar, M. H. (2021). Empirical mode decomposition and wavelet transform based ECG data compression scheme. *Irbm*, 42 (1), 65-72.

[104] Hassan, A. M. A, Alzaidi, M. S., Ghoneim, S. S., & El Nahal, W. (2022). Efficient Data Compression of ECG Signal Based on Modified Discrete Cosine Transform.

[105] Verma, A. R., Gupta, A., Kumar, P., & Kumar, Y. (2022). A Novel Approach of ECG Signal Compression using MRA Analysis based on WPD. *Journal of Scientific Research*, 66 (3).

[106] Singhai, P., Kumar, A., Ateek, A., Ansari, I. A., Singh, G. K., & Lee, H. N. (2023). ECG signal compression based on optimization of wavelet parameters and threshold levels using evolutionary techniques. *Circuits, Systems, and Signal Processing*, 42 (6), 3509-3537.

[107] Mohebbian, M. R., & Wahid, K. A. (2023). ECG compression using optimized B-spline. *Multimedia Tools and Applications*, 82 (14), 21071-21083.

[108] Novak, D., Frau, D. C., Eck, V., Pérez-Cortés, J. C., & Andreu-García, G. (2000). Denoising electrocardiogram signal using adaptive wavelets. *Energy*, 200, 250.

[109] Nikolaev, N., Gotchev, A., Egiazarian, K., & Nikolov, Z. (2001). Suppression of electromyogram interference on the electrocardiogram by transform domain denoising. *Medical and Biological Engineering and Computing*, 39, 649-655.

[110] Ziarani, A. K., & Konrad, A. (2002). A nonlinear adaptive method of elimination of power line interference in ECG signals, *IEEE Transactions on Biomedical Engineering*, vol. 49, n° 6, pp. 540-547.

[111] Esposito, A., & D’Andria, P. (2003). An adaptive learning algorithm for ECG noise and baseline drift removal. In *Italian Workshop on Neural Nets* (pp. 139-147). Berlin, Heidelberg: Springer Berlin Heidelberg.

[112] Ercelebi, E. (2004). Electrocardiogram signals de-noising using lifting-based discrete wavelet transform. *Computers in Biology and Medicine*, 34 (6), 479-493

[113] Chouakri, S. A., Bereksi-Reguig, F., Ahmaidi, S., & Fokapu, O. (2005). Wavelet denoising of the electrocardiogram signal based on the corrupted noise estimation. In *Computers in Cardiology*, 2005 (pp. 1021-1024). IEEE

[114] Łęski, J. M., & Henzel, N. (2005). ECG baseline wander and powerline interference reduction using nonlinear filter bank. *Signal processing*, 85 (4), 781-793.

[115] He, T., Clifford, G., & Tarassenko, L. (2006). Application of independent component analysis in removing artefacts from the electrocardiogram. Neural Comput Appl 15: 105-116.

[116] Sayadi, O., & Shamsollahi, M. B. (2007). Multiadaptive Bionic Wavelet Transform : Application to ECG Denoising and Baseline Wandering Reduction, EURASIP Journal on Advances in Signal Processing, vol. 2007, n° 1, pp. 1-11.

[117] Sameni, R., Shamsollahi, M.B., Jutten, C., & Clifford, G. D. (2007). A nonlinear Bayesian filtering framework for ECG denoising. IEEE Transactions on Biomedical Engineering.; 54 (12): 2172-85.

[118] Blanco-Velasco, M., Weng, B., & Barner, K. E. (2008). ECG signal denoising and baseline wander correction based on the empirical mode decomposition. *Computers in biology and medicine*, 38 (1), 1-13.

[119] Kaur, M., & Singh, B. (2009). Powerline interference reduction in ECG using combination of MA method and IIR notch. International Journal of recent trends in Engineering, 2 (6), 125.

[120] Al-Qawasmi, A. R., & Daqrouq, K. (2010). ECG signal enhancement using wavelet transform. *WSEAS Trans on Biology and Biomedecine,* 7 (2), 62-72

[121] El-Dahshan, E. S. A. (2011). Genetic algorithm and wavelet hybrid scheme for ECG signal denoising, Telecommunication Systems, vol. 46, n° 3, pp. 209-215.

[122] Smital, L., Vitek, M., Kozumplík, J., & Provaznik, I. (2012). Adaptive wavelet wiener filtering of ECG signals. *IEEE transactions on biomedical engineering*, 60 (2), 437-445.

[123] Kabir, M. A., & Shahnaz, C. (2012). Denoising of ECG signals based on noise reduction algorithms in EMD and wavelet domains, Biomedical Signal Processing and Control, vol. 7, n° 5, pp. 481-489.

[124] Chandrakar, C., & Kowar, M. K. (2012). Denoising ECG signals using adaptive filter algorithm. International Journal of Soft Computing and Engineering (IJSCE), 2 (1), 120-123.

[125] Yang, Y., & Wei, Y. (2013). Random interpolation average for ECG signal denoising using multiple wavelet bases. Biomedical Engineering: Applications, Basis and Communications, 25 (04), 1350042.

[126] Ari, S., Das, M. K., & Chacko, A. (2013). ECG signal enhancement using S-Transform, Computers in Biology and Medicine, vol. 43, n° 6, pp. 649-660.

[127] Talbi, M. (2014). Electrocardiogram de-noising based on forward wavelet transform translation invariant application in bionic wavelet domain. *Sadhana*, *39*, 921-937.

[128] Subbiah, S., Patro, R., & Rajendran, K. (2014). Reduction of noises in ECG signal by various filters. *International Journal of Engineering Research & Technology*, 3 (1), 656-660.

[129] Biswas, U., & Maniruzzaman, M. (2015). Power line interference removal from ECG signal using notch filter, adaptive filter and wavelet packet transform. American International Journal of Contemporary Scientific Research, 2 (6), 17-23.

[130] Salih, S. K., Aljunid, S. A., Aljunid, S. M., & Maskon, O. (2015). Adaptive filtering approach for denoising electrocardiogram signal using moving average filter. *Journal of Medical Imaging and Health Informatics*, 5 (5), 1065-1069.

[131] Wang, Z., Wan, F., Wong, C. M., & Zhang, L. (2016). Adaptive Fourier decomposition based ECG denoising. *Computers in Biology and Medicine*, 77, 195-205[e26] Nguyen, P., & Kim, J. M. (2016). Adaptive ECG denoising using genetic algorithm-based thresholding and ensemble empirical mode decomposition. *Information sciences*, 373, 499-511.

[132] Nguyen, P., & Kim, J. M. (2016). Adaptive ECG denoising using genetic algorithm-based thresholding and ensemble empirical mode decomposition. *Information sciences*, 373, 499-511.

[133] Panigrahy, D., & Sahu, P. K. (2016). Extended Kalman smoother with differential evolution technique for denoising of ECG signal. *Australasian physical & engineering sciences in medicine*, 39 (3), 783-795.

[134] Jenkal, W., Latif, R., Toumanari, A., Dliou, A., El B'charri, O., & Maoulainine, F. M. (2016). An efficient algorithm of ECG signal denoising using the adaptive dual threshold filter and the discrete wavelet transform. *Biocybernetics and Biomedical Engineering*, 36 (3), 499-508.

[135] Jain, S., Bajaj, V., & Kumar, A. (2017). Riemann Liouvelle fractional integral based empirical mode decomposition for ECG denoising. IEEE journal of biomedical and health informatics, 22 (4), 1133-1139.

[136] Singh, O., & Sunkaria, R. K. (2017). ECG signal denoising via empirical wavelet transform. *Australasian physical & engineering sciences in medicine*, 40, 219-229.

[137] Rakshit, M., & Das, S. (2018). An efficient ECG denoising methodology using empirical mode decomposition and adaptive switching mean filter, Biomedical Signal Processing and Control, vol. 40, pp. 140-148.

[138] Ahmad, A. S. S., Matti, M. S., Alhabib, O. A. M., & Shaikhow, S. (2018). Denoising of Arrhythmia ECG Signals. International Journal of Medical Research Health Sciences, vol. 7, n° 3, pp. 83-93.

[139] Bahaz, M., & Benzid, R. (2018). Efficient algorithm for baseline wander and powerline noise removal from ECG signals based on discrete Fourier series. Australasian Physical and Engineering Sciences in Medicine, 41(1), 143-160.

[140] Singh, P., & Pradhan, G. (2018). Variational mode decomposition based ECG denoising using non-local means and wavelet domain filtering. *Australasian physical & engineering sciences in medicine*, 41, 891-904.

[141] Kumar, S., Panigrahy, D., & Sahu, P. K. (2018). Denoising of Electrocardiogram (ECG) signal by using empirical mode decomposition (EMD) with non-local mean (NLM) technique. biocybernetics and biomedical engineering, 38 (2), 297-312.

[142] Zhang, D., Wang, S., Li, F., Wang, J., Sangaiah, A. K., Sheng, V. S., & Ding, X. (2019). An ECG Signal De-Noising Approach Based on Wavelet Energy and Sub-Band Smoothing Filter, Applied Sciences, vol. 9, n° 22, pp. 4968.

[143] Wang, Z., Zhu, J., Yan, T., & Yang, L. (2019). A new modified wavelet-based ECG denoising. *Computer Assisted Surgery*, 24 (sup1), 174-183.

[144] Asma, T., Mouna, G., Assam, O. M., & Kheireddine, C. (2019). Efficient filtering framework for electrocardiogram denoising. *International Journal Bioautomation*, 23 (4), 403.

[145] Ouali, M. A., Ghanai, M., & Chafaa, K. (2020). TLBO Optimization Algorithm Based-Type2 Fuzzy Adaptive Filter for ECG Signals Denoising. *Traitement du Signal*, 37 (4).

[146] Zhang, D., Wang, S., Li, F., Tian, S., Wang, J., Ding, X., & Gong, R. (2020). An efficient ECG denoising method based on empirical mode decomposition, sample entropy, and improved threshold function. *Wireless Communications and Mobile Computing*, 1-11.

[147] Houamed, I., Saidi, L., & Srairi, F. (2020). ECG signal denoising by fractional wavelet transform thresholding. *Research on Biomedical Engineering*, 36, 349-360.

[148] Zhang, M., & Wei, G. (2020). An integrated EMD adaptive threshold denoising method for reduction of noise in ECG. *Plos one, 15* (7), e 0235330.

[149] Talbi, M. (2021). A new ECG denoising technique based on LWT and TVM. *Circuits, Systems, and Signal Processing*, 40 (12), 6284-6300.

[150] Dwivedi, K., Ranjan, H., Menon, A., & Periasamy, P. (2021). Noise Reduction in ECG Signal Using Combined Ensemble Empirical Mode Decomposition Method with Stationary Wavelet Transform. Circuits, Systems, and Signal Processing 40, n° 2 : 827-844.

[151] Boda, S., Mahadevappa, M., & Dutta, P. K. (2021). A hybrid method for removal of power line interference and baseline wander in ECG signals using EMD and EWT. Biomedical Signal Processing and Control, 67, 102466.

[152] Mohguen, W., & Bouguezel, S. (2021). Denoising the ECG signal using ensemble empirical mode decomposition. Engineering, Technology & Applied Science Research, 11 (5), 7536-7541.

[153] Subbiah, B. (2022). ECG Signal Denoising Using Optimized Adaptive Hybrid Filter With Empirical Wavelet Transform.

[154] Hussein, A., Mohammed, W. R., Musa Jaber, M., & Ibrahim Khalaf, O. (2022). An Adaptive ECG Noise Removal Process Based on Empirical Mode Decomposition (EMD). Contrast Media & Molecular Imaging.

[155] Talbi, M., & Bouhlel, M. S. (2022). A novel technique of noise cancellation based on stationary bionic wavelet transform and WATV: application for ECG denoising. Int. Arab J. Inf. Technol., 19 (3), 381-387.

[156] Das, M., & Sahana, B. C. (2022). Optimized Orthogonal Wavelet-Based Filtering Method for Electrocardiogram Signal Denoising. *Journal of The Institution of Engineers (India): Series B*, 1-14.

[157] Malik, S. A., Parah, S. A., Aljuaid, H., & Malik, B. A. (2023). An Iterative Filtering Based ECG Denoising Using Lifting Wavelet Transform Technique. *Electronics*, 12 (2), 387.

[158] Yue, Y., Chen, C., Wu, X., & Zhou, X. (2023). An effective electrocardiogram segments denoising method combined with ensemble empirical mode decomposition, empirical mode decomposition, and wavelet packet. *IET Signal Processing*, 17 (6), e 12232.

[159] Prasad, V. V. K. D. V., Nagasirisha, B., Amulya, K. B., Sarada, K., & Sirisha, G. N. (2023). ECG SIGNAL DE-NOISING USING DIGITAL FILTER TECHNIQUES.

[160] Peterson, C., & Rögnvaldsson, T. S. (1991). *An introduction to artificial neural networks* (No. LU-TP-91-23). CERN.

[161] Sorensen, K., Sevaux, M., & Glover, F. (2017). A history of metaheuristics. *arXiv preprint arXiv:1704.00853*.

[162] McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5, 115-133.

[163] Hebb, D. O. (2005). *The organization of behavior: A neuropsychological theory*. Psychology press.

[164] Finger, S. (1994). Origins of Neuroscience, Oxford University Press, New York.

[165] Da Silva, I.N., Hernane Spatti, D., Andrade Flauzino, R., Liboni, L.H.B., & dos Reis Alves, S.F. (2017). Artificial Neural Networks. Springer International Publishing.

[166] Lin, J. W. (2017). Artificial neural network related to biological neuron network: a review. *Advanced Studies in Medical Sciences*, *5* (1), 55-62.

[167] Sharma, S., Sharma, S., & Athaiya, A. (2017). Activation functions in neural networks. *Towards Data Sci*, 6 (12), 310-316.

[168] Ding, B., Qian, H., & Zhou, J. (2018). Activation functions and their characteristics in deep neural networks. In *2018 Chinese control and decision conference (CCDC)* (pp. 1836-1841). IEEE.

[169] Goyal, M., Goyal, R., Venkatappa Reddy, P., & Lall, B. (2020). Activation Functions. In: Pedrycz, W., Chen, SM. (eds) Deep Learning: Algorithms and Applications. Studies in Computational Intelligence, vol 865. Springer, Cham.

[170] Gupta, T. K., & Raza, K. (2019). Optimization of ANN architecture: a review on nature-inspired techniques. Machine learning in bio-signal analysis and diagnostic imaging, 159-182.

[171] Djukanovic, M., Babic, B., Sobajic, D. J., & Pao, Y. H. (1993). Unsupervised/ supervised learning concept for 24-hour load forecasting. In IEE Proceedings C (Generation, Transmission and Distribution) (Vol. 140, N° 4, pp. 311-318). IET Digital Library.

[172] Dike, H. U., Zhou, Y., Deveerasetty, K. K., & Wu, Q. (2018). Unsupervised learning based on artificial neural network: A review. In 2018 IEEE International Conference on Cyborg and Bionic Systems (CBS) (pp. 322-327). IEEE.

[173] Dayan, P., Sahani, M., & Deback, G. (1999). Unsupervised learning. The MIT encyclopedia of the cognitive sciences, 857-859.

[174] Lange, K. (2013). Optimization (Vol. 95). Springer Science & Business Media.

[175] Collete, Y., Siarry. P. (2002). Optimisation multiobjectif. Eyrolles.

[176] Boussaïd, I. (2013). Perfectionnement de Metaheuristiques Pour l'Optimisation Continue. Ph. D. Thesis, University of Science and Technology Houari Boumediene, Bab Ezzouar, Algeria.

[177] Rao, S. S. (2019). *Engineering optimization: theory and practice*. John Wiley & Sons.

[178] Simon, D. (2013). *Evolutionary optimization algorithms*. John Wiley & Sons.

[179] Loubiere, P. (2016). *Amélioration des métaheuristiques d'optimisation à l'aide de l'analyse de sensibilité*. Ph.D. Thesis, University Paris Est.

[180] Talbi, E. G. (2009). Metaheuristics: from design to implementation. John Wiley & Sons. Verifier et modifier M1.

[181] Glover, F. (1986). Future paths for integer programming and links to artificial intelligence. Computers and Operations Research 13, 533-549.

[182] Betka, A. (2019). *Estimation de mouvement par les techniques métaheuristiques*. Ph. D. Thesis, University Mohamed Khider-Biskra.

[183] Chong, H. Y., Yap, H. J., Tan, S. C., Yap, K. S., & Wong, S. Y. (2021). Advances of metaheuristic algorithms in training neural networks for industrial applications. Soft Computing, 25 (16), 11209-11233.

[184] Mehrabian, A. R., & Lucas, C. (2006). A novel numerical optimization algorithm inspired from weed colonization. Ecological informatics, 1 (4), 355-366.

[185] Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In Proceedings of ICNN'95-international conference on neural networks (Vol. 4, pp. 1942-1948). IEEE.

[186] Hansen, N., & Ostermeier, A. (2001). Completely derandomized self-adaptation in evolution strategies. Evolutionary computation, 9 (2), 159-195.

[187] Bracewell, R. N. (1986). The Fourier Transform and Its Applications, McGraw-Hill Book Company.

[188] Andreaus, U., Baragatti, P., Casini, P., & Iacoviello, D. (2017). Experimental damage evaluation of open and fatigue cracks of multicracked beams by using wavelet transform of static response via image analysis, Structural Control and Health Monitoring, vol. 24, n° 4, pp. 1-16.

[189] Tu, J. H. (2013). *Dynamic mode decomposition: Theory and applications* (Doctoral dissertation, Princeton University).

[190] Mallat, S. (1998). A Wavelet Tour of Signal Processing, Academic Press, London.

[191] Huang, N. E., Shen, Z., Long, S. R., & Wu, M. C., Shih, H. H., Zheng, Q.,… & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proceedings of the Royal Society of London A, vol. 454, n° 1971, pp. 903-995.

[192] Liu, B., Riemenschneider, S., & Xu, Y. (2006). Gearbox fault diagnosis using empirical mode decomposition and Hilbert spectrum. *Mechanical Systems and Signal Processing, 20* (3), 718-734.

[193] Alfaouri, M., Daqrouq, K. (2008). ECG signal denoising by wavelet transform thresholding, *Ame. J. Appl. Sci.*, 5, (3), pp. 276-281.

[194] Poornachandra, S., Kumaravel, N. (2008). A novel method for the elimination of power line frequency in ECG signal using hyper shrinkage function, *J. Digital Signal Process.*, 18, pp. 116-126.

[195] Tremblay, N., Borgnat, P., & Flandrin, P. (2014, September). Graph empirical mode decomposition. In 2014 22nd European Signal Processing Conference (EUSIPCO) (pp. 2350-2354). IEEE.

[196] Chen, X., Li, B., Jia, H., Feng, F., Duan, F., Sun, Z., & Solé-Casals, J. (2022). Graph Empirical Mode Decomposition-Based Data Augmentation Applied to Gifted Children MRI Analysis. *Frontiers in Neuroscience*, 16, 866735.

[197] In, F., & Kim, S. (2013). An introduction to wavelet theory in finance: a wavelet multiscale approach‟. World scientific.

[198] Charles, C. (2011). Introduction aux ondelettes. Notes de Statistique et d'Informatique.

[199] Strang, G. (1993). Wavelet transforms versus Fourier transforms. *Bulletin of the American Mathematical Society*, 28 (2), 288-305.

[200] Kaiser, G. (2010). A friendly guide to wavelets. Springer Science & Business Media.

[201] Conraria, L. A. & Soares, M. J. (2011). The continuous wavelet transform: A primer‟, NIPE Working Paper. Universidade do Minho. Núcleo de Investigação em Políticas Económicas (NIPE), 16, pp. 1-43.

[202] Daubechies, I. (1992). Ten lectures on wavelets. Siam.

[203] Liu, C.-L. (2010). A tutorial of the wavelet transform‟, NTUEE, Taiwan.

[204] Olkkonou, H. (2011). Discrete wavelet transforms: Algorithms and applications “Intech”.

[205] Van Fleet, P. J. (2011). Discrete wavelet transformations: An elementary approach with applications. John Wiley & Sons.

[206] Stark, H. G. (2005). Wavelets and signal processing, Springer.

[207] Chen, S. H., Jakeman, A. J., & Norton, J. P. (2008). Artificial intelligence techniques: an introduction to their use for modelling environmental systems. Mathematics and computers in simulation, vol. 78, n° 2-3, p. 379-400.

[208] Chen, S., Billings, S. A., & Grant, P. M. (1990). Non-linear system identification using neural networks. International journal of control, vol. 51, n° 6, p. 1191-1214.

[209] Feng, X., & Schulteis, J. (1993). Identification of high noise time series signals using hybrid ARMA modeling and neural network approach. In: IEEE International Conference on Neural Networks. IEEE, p. 1780-1785.

[210] Narendra, K. S. & Parthasarathy, K. (1992). Neural networks and dynamical systems. International Journal of Approximate Reasoning, vol. 6, n° 2, p. 109-131.

[211] The MIT-BIH Arrhythmia Database (2000). [online] https://www.physionet.org/physiobank/database/mitdb/.

[212] Mcsharry, P. E., Clifford, G. D., Tarassenko, L., & Smith, L. A. (2003). A dynamical model for generating synthetic electrocardiogram signals. IEEE transactions on biomedical engineering, vol. 50, n° 3, p. 289-294.

[213] Moody, G. B., & Mark, R. G. (2001). The impact of the MIT-BIH Arrhythmia Database, IEEE Engineering in Medicine and Biology Magazine, vol 20, n° 3, pp. 45-50. https: //www.physionet.org/content/mitdb/1.0.0

[214] Dandapat, S., Sharma, L. N., & Tripathy, R. K. (2015). Quantification of diagnostic information from electrocardiogram signal: A review, *Advances in Communication and Computing*, vol. 347, pp. 17-39.

[215] Sundar, A., Pahwa, V., Das, C., Deshmukh, M., & Robinson, N. (2016). A comprehensive assessment of the performance of modern algorithms for enhancement of digital volume pulse signals. *International Journal of Pharma Medicine and Biological Sciences*, 5 (1), 91.

[216] Simonsen, A. Hansen, & Nes, O. M. (1998). Determination of the Hurst exponent by use of wavelet transforms. Physical Review E, vol 58, pp. 2779-2787.

[217] Hurst, H.E. (1951). Long term storage capacity of reservoirs. Trans. Am. Soc. Civil Eng. 116: 776-808.