An Expanded IPFM Model for Heart Rhythm Analysis

Detecting Atrial Fibrillation Using a Physiological Model

A. Kordes

TU Delft

An Expanded IPFM Model for Heart Rhythm Analysis

Detecting Atrial Fibrillation Using a Physiological Model

by

A. Kordes

to obtain the degree of Master of Science
at the Delft University of Technology

Student number:	4457560	
Project duration:	May 17, 2021 – December 16, 2021	
Thesis committee:	dr. ir. R. C. Hendriks,	TU Delft, supervisor
	dr. ir. C. Varon,	TU Delft, supervisor
	prof. dr. N. M. S. de Groot,	Erasmus Medical Center

TU Delft

Preface

The subject of this thesis is the modelling of a part of the electrical conduction system of the human heart. On the far horizon, an extensive physiological model of the heart may greatly improve diagnostics and detection of diseases. I hope this work will aid in this vast endeavor.

I would like to express my gratitude to my supervisors. To Richard Hendriks, for accepting me as a student with a different background than usual and for a shorter project time. Your optimism has always inspired me to make the most out of this project. To Carolina Varon, for your fast understanding of my ideas (even the strange ones) and of course your extensive knowledge on this topic. I would like to thank my friends and family for listening to me whenever I was trying to explain the exciting ideas I was working on. Without you, this thesis would not have been the same.

Arthur Kordes
Delft, December 2021

Abstract

Atrial Fibrillation affects millions of people worldwide. It is associated with an impaired quality of life and an increased risk of stroke, cardiac failure and mortality. Treatments exist, but early detection and treatment is crucial, due to the progressive nature of the disease. Algorithms can help with early detection. Machine learning algorithms are commonly trained to diagnose based on ECG data, but the interpretability is low. A physiological model that simulates the heart gives more insight into the situation of the patient. Current approaches, like the IPFM model, simulate only the SA node and generate RR intervals as output, while completely neglecting the interaction between the AV and SA node. By using an IPFM model and including the AV node as well, an extended and more accurate physiological model was built to more accurately detect Atrial Fibrillation.

The AV node model is able to estimate PR intervals when the P waves are annotated. This result shows that the model extension is able to capture information about the signal conduction. When the SA node model and the AV node model are cascaded and only the R peaks are considered, the classification accuracy does not improve compared to the SA node model alone. The R peaks alone do not contain sufficient information for accurate parameter estimation. The parameters governing the behavior of the AV node seem different for NSR compared to AF, but more data is needed to confirm this. The ability of the model to predict PR intervals gives hope that the inclusion of P wave data should improve the performance of the classification with the extended physiological model.

Contents

1 Introduction

Atrial Fibrillation (AF) affects millions of people worldwide. AF is associated with an impaired quality of life and an increased risk of stroke, cardiac failure and mortality [1]. Due to these risks, as well as the progressive nature of the disease, early detection accompanied with early treatment is of tremendous importance.

The Electrocardiogram (ECG) shows information about the activity of a heart. It measures the electrical activity of the heart in real time using electrodes on the skin. Trained specialists are able to make an evaluation of the patient and may be able to diagnose AF. However, accurate diagnosis requires expertise and a lot of time. Also, it is difficult for clinicians to distinguish between the progressive developmental stages of AF.

It would be ideal to extract vital information from the ECG using an algorithm and predict the AF development stage. Machine Learning (ML) algorithms are already able to make predictions about the risk of a patient's risk of AF [2]. They either take the entire ECG trace as input, or make decisions based on certain metrics. In this thesis, we will build a physiological model that is able to generate sequences that repesent the timing of the heart beats. When this model is fitted to a patient, the outcome should give a diagnosis in an interpretable manner.

The ventricles create a QRS complex, which is clearly visible on an ECG trace. The fluctuations in the timing of heart beats visible in the fluctuations of the QRS complexes is commonly known as Heart Rate Variability (HRV). Analysis of these intervals is often done by observing the sequence of RR intervals, also called a tachogram, as shown in Figure 1.1.

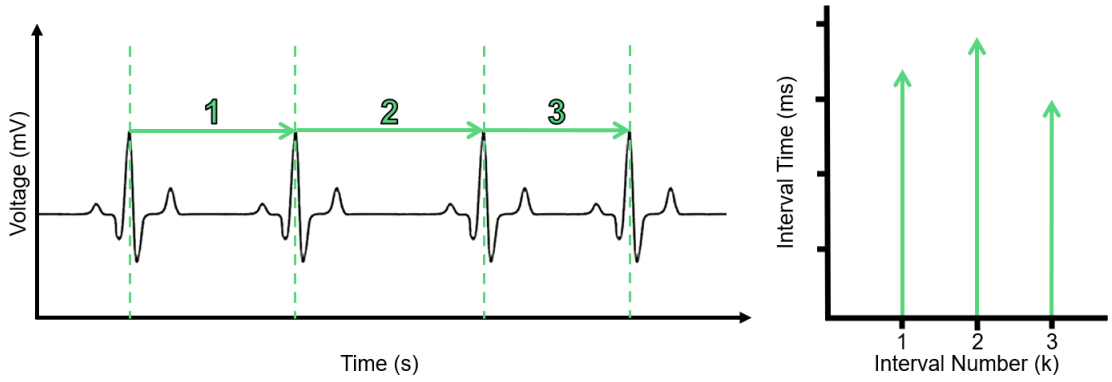


Figure 1.1: The generation of the RR tachogram from R peak detection.

We are interested in characterizing the tachogram in order to get insight into the health condition of a patient's heart. To do so, we first develop a physiological model that can generate realistic artificial tachograms. Then, at a later stage we use this model as an inverse model to extract the underlying physiological parameters of a patient based on real tachograms. Based on the extracted parameters we then obtain insight in the condition of the patient's heart.

In Chapter 2, the physiology of the heart rate will be explored. Both the sinoatrial (SA) node and the atrioventricular (AV) node are essential for regulating the pace of contraction. We will also take a look at AF and the possible differences compared to a healthy heart producing a Normal Sinus Rhythm (NSR).

In Chapter 3, the analysis of the ECG and specifically HRV is presented. We start with the delineation of the ECG. Given the RR or PP tachogram, there are methods of HRV analysis in three domains: time domain, frequency domain and geometric domain. Our model should create RR tachograms that resemble the original tachogram in all these domains. Lastly, we will analyze the PR interval.

In Chapter 4, a commonly used parametric model, the Integrated Pulse Frequency Modulation (IPFM) model, is described and expanded. This model is able to generate artificial impulses by an 'integrate and fire' model [3]. The IPFM model is generally used to describe the behaviour of the SA node, with the intention of modelling the RR intervals. The RR peaks however, are the result of contraction of the ventricles, which closely follows pacing of the AV node. We will include a second IPFM model, modelling the AV node. The two IPFM models are then cascaded. The first one will model the SA node. Its output will be a PP tachogram, which will act as the input for the AV node model. An evolutionary algorithm will aid in finding the model parameters.

The results section in Chapter 5 shows the results in realistic interval generation, the prediction of PR intervals, as well as parameter significance and classification. Both RR tachograms and PR tachograms are created. We will test which model parameters are different during AF. These parameters should be able to eventually classify and help with the diagnosis. We use Machine Learning (ML) algorithms to use these parameters to classifiy RR sequences originating from different heart conditions.

Finally, the discussion and future work is presented in Chapter 6 and the conclusions are considered in Chapter 7.

2

Physiology of the Heart

Knowledge of the physiology of the heart is fundamental in order to understand any physiological model. In Section 2.1 the basic functions, mechanisms and measurement techniques of the electrical activity of the human heart are examined. Next, in Section 2.2 we will focus more on the cardiac nodes. They make impulses propagate through the cardiac chambers, such that contraction directly follows their excitation. In Section 2.3 we investigate Atrial Fibrillation and some other cardiac arrhythmias. Finally, after discussing the physiology of the heart and existing physiological models, we identify that particular aspects of the physiology of the human heart are missing in existing models. This gives rise to extend these models as explained in Section 2.4, where we will state the research goal of this work.

2.1. Overview of the Heart

The human heart has the duty to pump blood around the body [4]. Blood is the transport medium that carries oxygen and nutrients as well as waste products. There are four compartments inside the heart. There is a right side and a left side and both sides consist of atria and ventricles. A frontal section of the human heart is shown in Figure 2.1. The right side of the heart pumps blood to the lungs and the left side brings it to the rest of the body. The split between atria and ventricles exists because of efficiency. The atria can fill up with more blood, while the ventricles contract. This split requires the heart to have a mechanism to control the timing of contraction of both the atria and the ventricles.

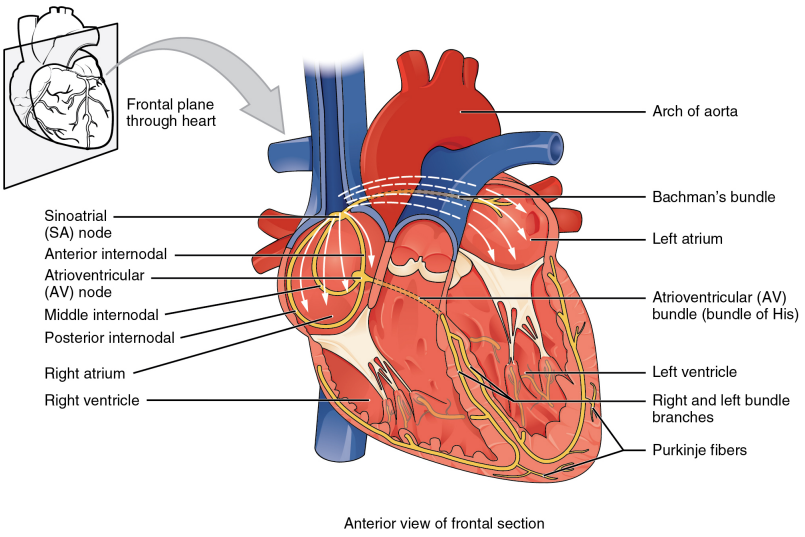
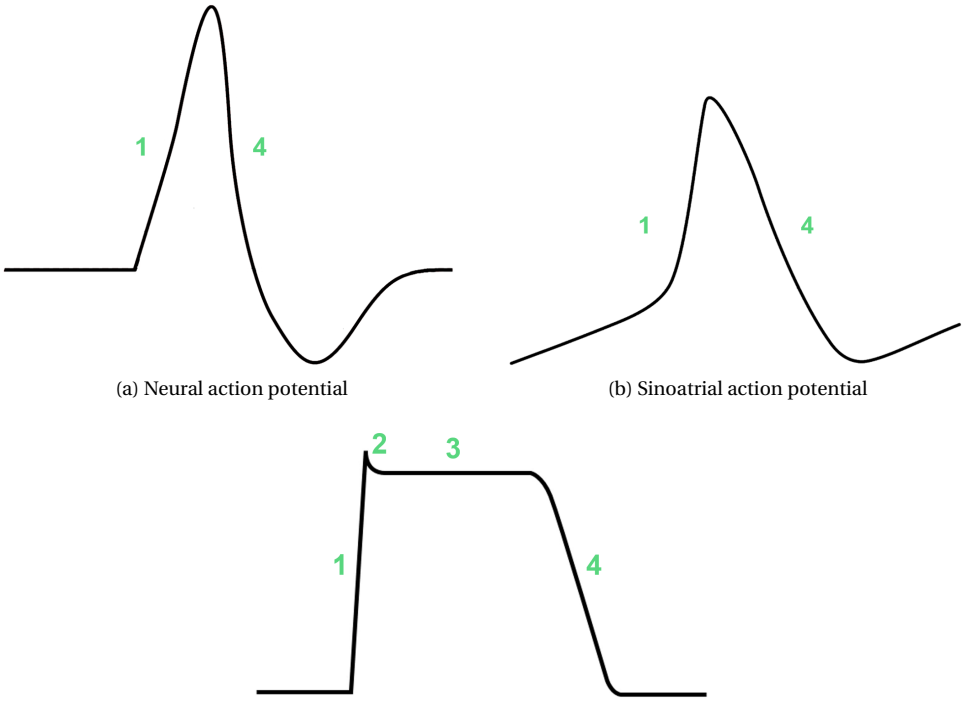


Figure 2.1: A schematic coronal view of the heart. The conduction system is highlighted (via openstax.org, 2013. Accessed 03-12-2021 [5]).



(a) Neural action potential

(b) Sinoatrial action potential

(c) Ventricular action potential

Figure 2.2: Some typical action potentials. 1: Depolarization 2: Efflux of K+ starts repolarization 3: Influx of Ca2+ balances the transmembrane potential 4: Ca2+ channels close and repolarization finishes. There are no obvious phases 2 and 3 in neural or pacemaker cells.

2.1.1. Electrical Conduction in the Heart

The cardiac cycle starts with the atria that contract. After some delay, the ventricles will contract as well. The timing of these contractions is organized by two regulatory nodes. These nodes are the Sinoatrial (SA) node and the Atrioventricular (AV) node. The bioelectrical signal starts its journey at the SA node. From here, it will travel through the atria, which will contract upon stimulation. The signal will then reach the AV node. After a delay of about 0.1 seconds it will then continue through the bundle of His towards the Purkinje fibres in the ventricles. In Figure 2.1, the signal path can be followed when starting from the SA node.

The signal that travels through the heart is not a flow of electrons, but rather a wave of depolarizing cells. Each myocardial cell produces an action potential (AP). An AP will trigger the neighboring cells using pores known as gap junctions. Just like in neurons and skeletal muscle cells, an AP is caused by the opening of voltage-gated ion channels. The membrane potential of a myocardial cell is negative at rest. When the potential crosses a threshold, a positive feedback loop is started. Voltage-gated Na+ channels will open causing more Na+ to enter the cell and increase the potential further. After depolarization, the efflux of K+ ions starts the repolarization. During this phase, there is a difference between cardiac myocytes and neurons in the brain. A neuron has a function as signal conductor, while a myocardial cell has is a signal conductor as well as a contracting unit. An influx of Ca2+ in the myocytes balances the flow of K+, with as a result a stabilized membrane potential. This allows for effective contraction. Typical cardiac AP's are shown in figure 2.2.

In figure 2.2 three typical AP's are shown. The neural AP only conveys information. The firing of a neuron sends a binary signal in the nervous system. Sinoatrial AP's are different in that they are able to fire without external stimulation. There is an upward slope visible, so the threshold will eventually be reached, even without external stimulation. The ventricular AP shows a clear plateau, where the ventricles are given the time to contract and pump out all its contents.

2.1.2. The Electrocardiogram

The electrocardiogram (ECG) is used to unobtrusively measure the electrical activity of the heart. It does so by using electrodes to measure the voltage on the skin of the patient. The combination of depolarizing cells creates an electrical field that is recorded. It will result in a plot where the voltage (often in millivolts) is plotted against time. The ECG is unobtrusive, it requires little preparation from the patient and it can be

performed in just a few minutes. These properties make it the most commonly used method for diagnosing cardiac problems. Experts are able to detect abnormalities in ECG traces. The 12-lead ECG method is the clinical standard, but even a single lead is able to provide useful information.

The characteristic waves of an ECG are shown in Figure 2.3. The contraction of the atria results in the P wave. The atria contain less muscle than the ventricles, which results in a lower amplitude. The ventricles create the so called QRS complex. Repolarization of the ventricles produces the T wave.

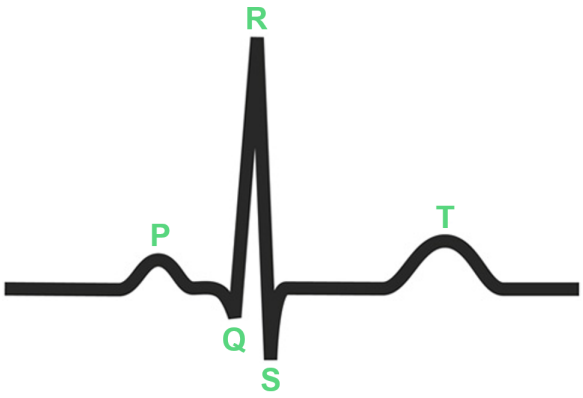


Figure 2.3: A typical shape of an electrocardiogram trace for normal sinus rhythm. The P wave, the QRS complex and the T wave are annotated.

2.2. Cardiac Nodes

2.2.1. The SA Node

The SA Node initiates the pulses that will travel through the heart. The SA node is autonomous, implying it is able to fire on its own [6]. However, they are influenced by the Autonomic Nervous System (ANS) via hormones. On the one hand, the sympathetic nervous system will increase the Heart Rate (HR). This is known as the 'fight or flight' system. On the other hand, the parasympathetic nervous system will decrease the HR and is known as the 'rest and digest' system. The firing of the SA node directly results in the contraction of the atria resulting in the P wave. A visualization of the influence of the ANS on the HR is shown in Figure 2.4.

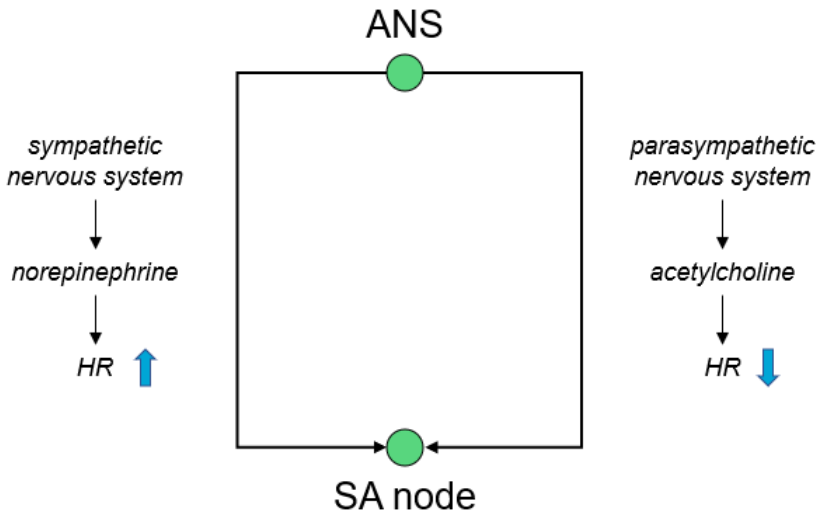


Figure 2.4: The Autonomic Nervous System uses hormones to modulate the Heart Rate.

Leaking ions allow the pacemaker node to fire autonomously. A steady stream of Na+ ions and K+ ions increases the membrane potential [7]. It is called the pacemaker current. The hormone norepinephrine increases this current [8]. The transmembrane voltage over time follows equation 2.1 for small changes.

$$\frac{dV_m}{dt} = \frac{I_L}{C_m} \tag{2.1}$$

Here V_m is the transmembrane voltage, I_L is the leak current and C_m is the membrane capacitance. The voltage increases when cations flow inward. Once the AP threshold is reached, this formula does not hold anymore, as the capacitor will be in parallel with by a low valued resistor when the ion channels open.

2.2.2. The AV Node

The AV node starts the sequence of stimulation that will lead to the contraction of the ventricles. The ventricles should contract when the atria have had the time to empty their contents. To guarantee this, 2the AV node introduces a time delay. Together with the time it takes for the cardiac activation to pass through the atria and the bundle of His, the PR interval is generally about 0.1 to 0.2 seconds. The interval is composed of a relatively fixed conduction delay and a variable part determined by the state of the AV node [9]. The R wave is produced when the ventricles contract.

The AV node is a highly complex structure. Although its functionality is partly understood, much is still unknown [10]. Inside a living heart, its behaviour is rather unpredictable due to the plethora of body signals and other noise present. The properties listed below are better distinguishable in perfused and isolated hearts, where the firing of the SA node is controlled. Their effects are still observable in living human hearts.

- Recovery: After firing, the AV node will slowly recover its ability to fire again. An impulse arriving from the SA node earlier than expected will increase the time it takes the impulse to travel through the AV node.
- Fatigue: An increased basic rate of impulse arrivals will also increase the AV nodal conduction time.
- Accommodation: The AV nodal conduction time settles when a new basic rate is applied.
- Concealed conduction: A premature impulse may be blocked in the AV node, but it will influence its response to the next impulse.

Another aspect of the AV node is the dual pathway conduction. Inside the AV node there is a fast pathway and a slow pathway. During NSR the fast pathway generally transduces impulses to the bundle of His. If an impulse arrives at the AV node when the fast pathway has not yet repolarized, it will follow the slow pathway. For this reason, the dual pathways often become apparent during arrhythmias.

Just like the SA node, the AV node is able to autonomously produce impulses. The base rate of the SA node is higher, so it will generally dictate the pace of the heart. The automaticity of the AV node is useful as a backup in case of a complete heart block as described in the following section. The structure of the AV node is complex, so we will construct the physiological model in such a way that it aims to model the general characteristics of the AV node behavior instead of the physical aspects.

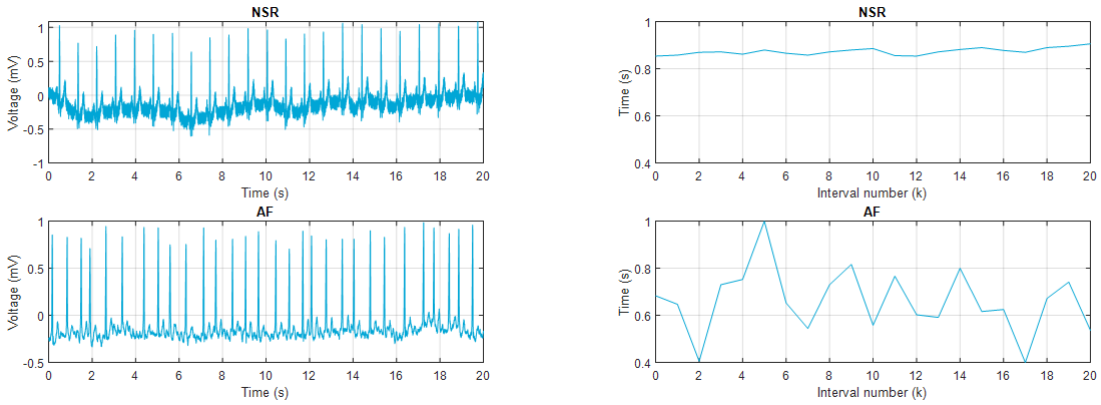
2.3. Atrial Fibrillation

A cardiac arrhythmia is an abnormality within either the rate, rhythm, sequence of conduction or origin of conduction [11]. When the timing of the impulses goes wrong, the heart will pump less efficiently. In severe cases the heart can get damaged. Atrial Fibrillation is the most common arrhythmia. During AF, the typical wave of cell depolarization in the atria is impaired.

2.3.1. Pathophysiology

There is no general consensus on the mechanisms behind AF. First, we have to distinguish between de novo post-operative AF (POAF) and persistent AF. POAF is a common complication developed after cardiac surgery [12]. It is generally short-lasting, while persistent AF is long-lasting. The mechanisms behind POAF clearly have to be related to the surgery. Inflammation and oxidative stress are probable triggers. Still, predicting its occurrence or prevention is difficult. Two possible mechanisms behind AF are foci and rotors. Foci are irritable ectopic areas that start producing impulses together with the SA node, but not in sync [13]. Rotors are regions where re-entry of an impulse is possible. This location will continually stimulate the cells around it.

AF creates a recognizable pattern on an ECG. There is no clear P wave, but smaller chaotic waves are present as visible in the ECG traces from Figure 2.5. Also, the RR intervals are irregular as visible in the



(a) ECG traces

(b) Tachograms

Figure 2.5: ECG traces (left) and the accompanying tachograms (right) of Normal Sinus Rhythm (top) and Atrial Fibrillation (bottom).

tachogram from Figure 2.5. After diagnosis, possible treatments include medication, ablation or the implantation of an artificial pacemaker [14]. Various drugs are used to control the ventricular rate to a safe level. When the origin is thought to be highly local, ablation can be applied. Here, the ectopic area is removed by local application of energy, like radiofrequency, laser or cryotherapy. A pacemaker may aid in restoring and maintaining the heart's normal rhythm. AF is a progressive disease, so early detection and treatment is crucial. In Section 2.4 we will describe how we will aim to improve detection methods.

2.3.2. Other Cardiac Arrhythmias

AF is not the only cardiac arrhythmia. A diagnosis should discriminate between AF and other complications. A similar arrhythmia to AF is Atrial Flutter (AFL). AFL is commonly caused by re-entrant circuits in the atria, often near the tricuspid valve. It has a more regular pattern of atrial contraction compared to AF. The differences between the two as seen on an ECG are shown in Figure 2.6. When untreated, AFL can often degenerate into AF.

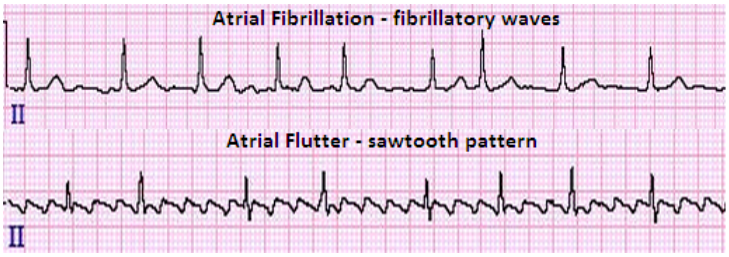


Figure 2.6: ECG of Atrial Fibrillation and Atrial Flutter (via https://www.healio.com/cardiology/learn-the-heart/cardiology-review/topic-reviews/atrial-fibrillation, 2021. Accessed 26-10-2021).

An abnormal heart rhythm in the ventricles is called Ventricular Fibrillation (VF). It is much more dangerous than AF, since the ventricles pump blood to the body. Malfunctioning ventricles result in cardiac arrest or even death [15]. In the case of AF, the AV node acts as a filter for the rapid atrial impulses. The result is a somewhat stable rhythm of ventricular contraction. However, when the origins of fibrillatory contraction are present in the ventricles, no filtering is possible. The ventricles will not be able to effectively pump blood, which results in critical situations.

During NSR, the mean HR lies between the 60 and 100 beats per minute. If the rate is higher, this is called tachycardia. If it is slower it is called bradycardia. A slow HR does not necessarily mean the heart is unhealthy. Vagal tone, sleep, athleticism all lower the HR. There could however be an underlying reason for the altered average HR, which may lead to more severe arrhythmias in the future if it remains untreated.

Re-entrant circuits can occur in the atria, but sometimes an accessory pathway will make re-entries on a larger scale possible. The bundle of Kent is an example of an unwanted conduction area. It connects the

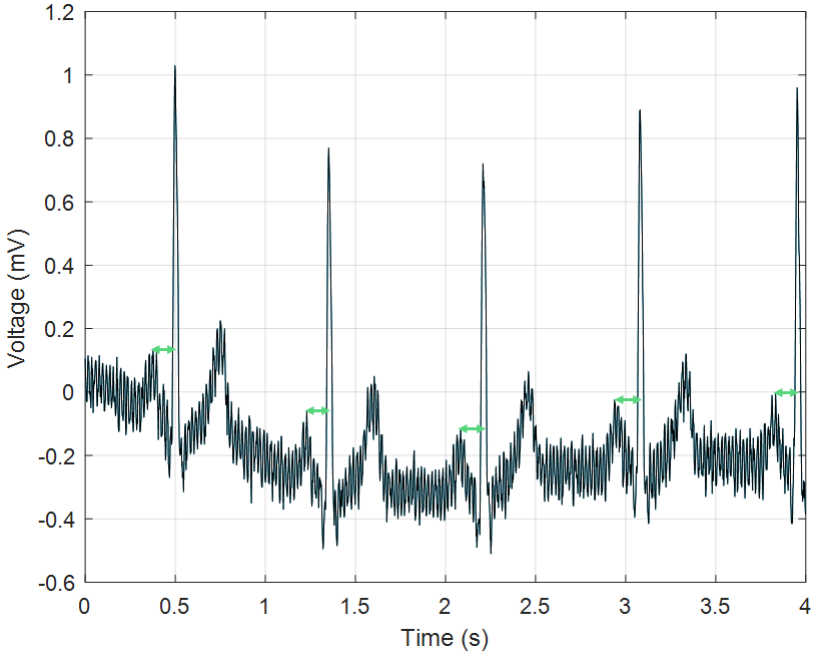
ventricles with the atria and allows impulses to travel in the opposite direction without control. A contraction of the ventricles will immediately stimulate the atria again. And a contraction of the atria will stimulate the ventricles without having to pass the AV node. The combination of AF and a Kent pathway dangerous, because the fast contractions of the atria will directly result in fast contractions of the ventricles.

When the electrical connection in the heart is inadequate, we speak of Heart Block (HB). Blockades can occur everywhere within the electrical conduction system, but they most commonly refer to AV node blocks. Three degrees exist for SA and AV node blocks, where a higher degree is more severe. For example, in a first-degree AV block the conduction through the AV node is slowed down. This results in a larger PR interval, but it does not require treatment. In a third-degree AV block the conduction is completely blocked. The AV node will fire autonomously, but the P waves and QRS complexes will have no correlation at all. An artificial pacemaker is often used to restore synergy between the impulses.

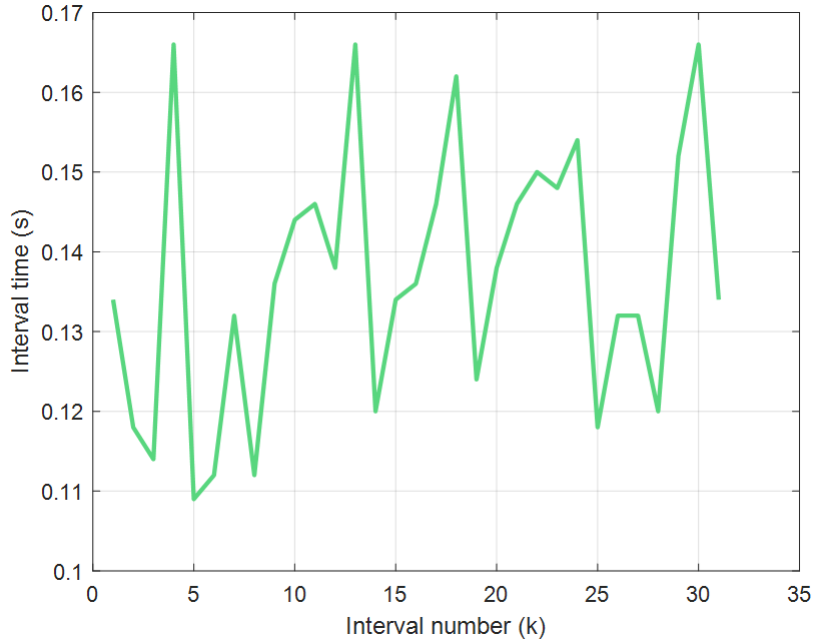
Re-entrant circuits can appear in one more situation. An impulse travelling from the SA node towards the AV node will take two pathways. Both the fast pathway and the short pathway as described in Section 2.2 will be stimulated. During NSR, the short one is dominant. However, during AF it often happens an impulse arrives while the short path is still repolarizing. The long pathway will now conduct the impulse towards the His bundle. A re-entrant circuit can appear when the signal goes back through the fast pathway once it has recovered during the passing of the signal through the slow pathway.

2.4. Research Goal

We hypothesize that parametric models can be used to predict the development stage of arrhythmias and the general condition of the heart. Given a parametric model, the model parameters are to be extracted from measured signals. An existing model is the Integrated Pulse Frequency Modulation (IPFM) model. The IPFM model models the SA node, but the model parameters are taken from the RR tachogram. The PR intervals are thus ignored and assumed constant. However, the intervals are not constant, as can be seen in Figure 2.7. The PR interval contains information about the atrial conduction and the AV node. We hypothesize that extracting the model parameters of this extended model will improve the detection of AF. Our research goal is to investigate whether we can expand the IPFM model, such that it combines the SA node with the AV node. The proposed cascaded IPFM model will produce pulse trains, resembling the PP and RR tachograms. We will investigate whether the proposed AV node model is able to model the AV node accurately and whether it could potentially increase classification accuracy.



(a) ECG trace



(b) PR Tachogram

Figure 2.7: The PR intervals taken from a fully annotated dataset show variability.

3

Analysis of Heart Rate Variability

The human heart rate is not constant. Aside from a significant low frequency variation due to changes in activity, the autonomic nervous system also modulates the pace of the heart. Due to this fact, HRV is a commonly used marker for the assessment of vagal and sympathetic activity. By analyzing the frequency content of the RR intervals, the activity of both the sympathetic and the parasympathetic ANS can be assessed.

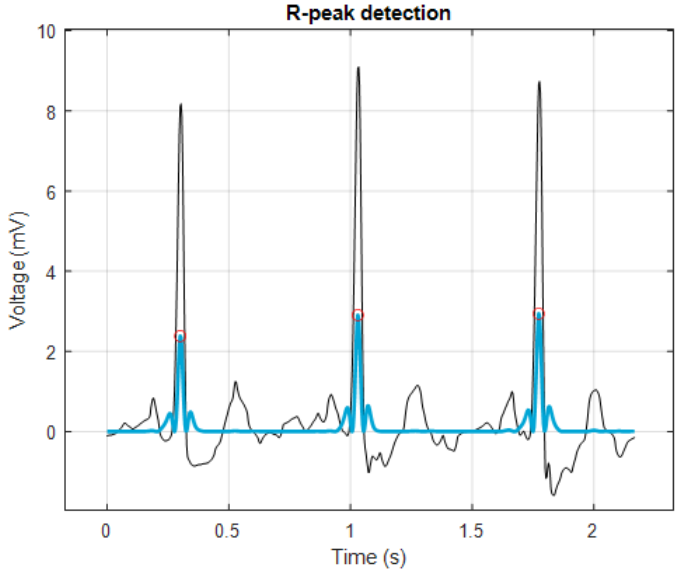


Figure 3.1: The R peaks of the ECG signal (black) can be robustly detected after a wavelet transformation is applied.

The goal of HRV analysis in our case is to compare the RR tachogram from measured data with the RR tachograms generated by our physiological model. The model should create a sequence that is not necessarily close to the data reference in absolute value, but the statistical properties should match. The 'Task Force of The European Society of Cardiology and The North American Society of Pacing and Electrophysiology' has written a guideline on the standards of measurement, physiological interpretation and clinical use of HRV [16]. They present various HRV metrics with their benefits and downsides. We have selected a subset of these metrics based on their recommendations and we will explain them in this chapter. We use Matlab 2020a for analysis. For HRV analysis, we start with the detection of the R peaks. The result of such a peak detection algorithm is shown in Figure 3.1. Next, we show the analysis of the RR tachograms. Lastly, we describe analysis of PR intervals.

3.1. ECG Delineation

In order to analyze the timing properties of the heart, we need some reference points. As HRV analysis is sensitive to outliers, the detection of these points must be robust. Simple peak detection algorithms often

make mistakes due to noise in the recording. Time-frequency analysis allows for the detection of specific frequency contents over the entire recording. Wavelets are commonly used in these types of analysis. A wavelet is a function used to transform into different frequency components. An example of a wavelet is the sym4 wavelet shown in Figure 3.2. This specific wavelet is frequently used for the detection R peaks in ECG data.

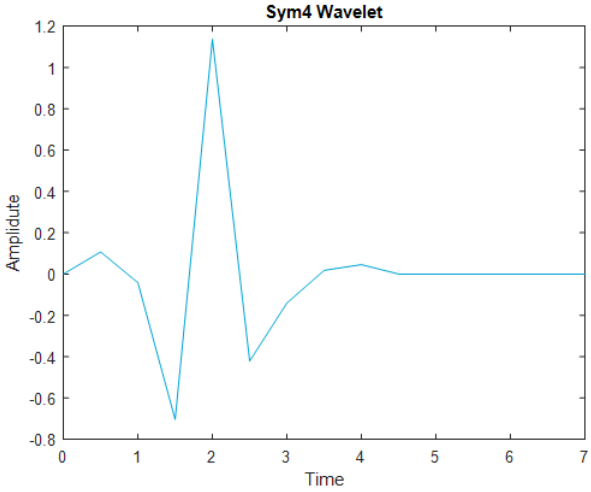


Figure 3.2: An approximation of the sym4 wavelet from the Matlab wavelet transform filter bank. The axes do not contain units, as the width and height of the wavelet can be adjusted freely.

Traditional HRV analysis uses the peak of the R wave as a marker. For HRV analysis to be accurate, it requires the ECG signal to have a high temporal resolution. A sampling frequency of 500 Hz is desirable, but lower frequencies may be sufficient if the data is clean and if detection of small shifts in RR interval variability is not required [17]. The R peak is detectable using an ECG with even one single lead. Distinct points on an ECG are the P wave, the QRS complex and the T wave, as described in Chapter 2. Delineation algorithms, as proposed by Laguna et al.[18], are able to detect these waves. The Physionet toolbox uses this algorithm, so we use it to detect both the P waves and the R peaks, such that PR interval analysis is possible. For just R-wave detection however, a simpler algorithm will suffice. Matlab can execute this less complex version easily, using just the sym4 wavelet and peak detection. The result of the algorithm used by Matlab is shown in figure 3.1.

3.2. Analysis of the RR tachogram

3.2.1. Time Domain Analysis

In this section we present several instrumental measures to analyze the tachogram.

- Sample Mean
  The sample mean is used as an estimator for the mean RR interval and is given by (3.1), where N denotes the total number of intervals and RR_n denotes the nth interval length.

$$\overline{RR} = \frac{1}{N}\sum_{n=1}^{N} RR_n. \tag{3.1}$$

- Sample Variance
  The variance of the RR tachogram reflects the deviation from the mean of the HRV. Also, it is equal to the total power of the spectral density function. It is estimated using the unbiased estimator given by (3.2):

$$\sigma^2 = \frac{1}{N-1}\sum_{n=1}^{N} (RR_n - \overline{RR})^2. \tag{3.2}$$

- RMSSD
  There are several measures describing interval differences. They are notably correlated, since they all

measure short-term variation. We will use the Root Mean Square of Successive Differences (RMSSD), which shows the beat-to-beat variance. The unbiased estimator is given by (3.3).

$$RMSSD = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N-1}(RR_{n+1} - RR_n)^2}. \tag{3.3}$$

3.2.2. Geometric Analysis

In geometric analysis, the data is plotted in a geometric pattern. The metrics are based on the resulting shapes and sizes of these patterns. This method of analysis is robust, because single mistakes in the measurements will not lead to large errors in the metric. However, relatively large numbers of data is required for reasonable estimation. The Poincaré plot is the only analysis we use from the geometric domain.

The Poincaré plot or Lorentz plot plots each interval against the previous one. A point cloud is formed of all consecutive intervals. An example is shown in Figure 3.3. The metrics are visual, which allows for easy interpretation. The standard deviation along the diagonal axes shows the short- and long term variability of the RR tachogram [19]. In Figure 3.3, the standard deviation along the X_1 axis (the width of the scatterplot) denotes the short term variation. The X_2 axis denotes the long term HRV.

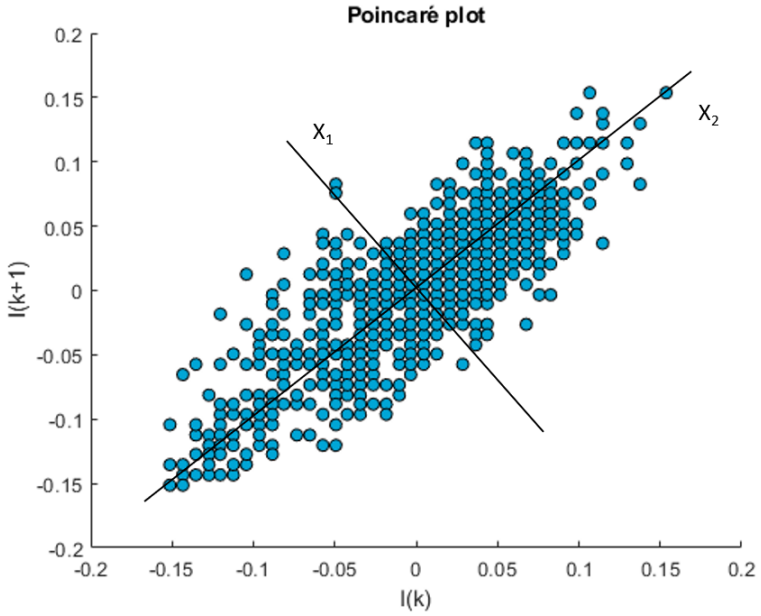


Figure 3.3: A Poincaré plot of an RR tachogram. Intervals I with interval number k are plotted against the following interval. The metrics $SD1$ and $SD2$ are computed by taking the standard deviation along the two axes.

3.2.3. Frequency Domain Analysis

The Power Spectral Density (PSD) denotes the power for all frequencies. For the RR tachogram, different frequency bands have a specific physiological origin. The High Frequency (HF) band ranges from 0.15 to 0.4 Hz and consists of the activity of the parasympathetic nervous system. This part of the ANS is modulated by breathing. Respiratory Sinus Arrhythmia (RSA) is a phenomenon where the breathing frequency is directly detectable in the PSD plot. The Low Frequency (LF) band ranges from 0.04 to 0.15 Hz and consists of the activity of the sympathetic nervous system. Furthermore there are the Very Low Frequency (VLF) band and the Ultra Low Frequency (ULF) band. These bands are only measurable when the length of the recording is long enough. The commonly used power bands are visible in Figure 3.4. The red line indicates the minimum frequency that is measurable. The value is calculated as follows: $f_{min} = \frac{2}{T}$, where f_{min} is the minimal frequency and T is the total measurement time.

PSD estimation for the RR tachogram is non-trivial, because the horizontal axis only expresses the index number, but the actual interval times are lost. Hence, the intervals times are plotted regularly, but the intervals are not constant. This means a simple uniform Fast Fourier Transform (FFT) will produce a distorted

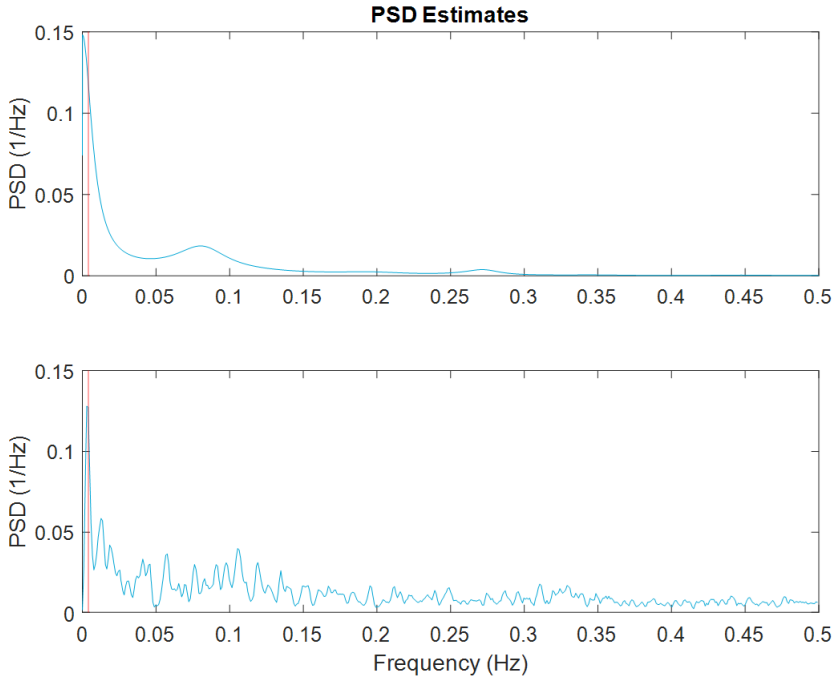


Figure 3.4: PSD estimates using a 15^{th} order AR model (top) and the Heart Timing method (bottom).

result. A multitude of solutions exist that aim to tackle this obstacle [20]. The Lomb-Scargle method is used to detect periodicity in unevenly sampled data. The Non-Uniform Discrete Fourier Transform (NDFT) allows for computation of the PSD of nonuniform sampling too.

Another strategy is the use of interpolation. When the signal is interpolated, methods that allow unevenly sampled data are no longer needed. Common approaches include spline interpolation, where the signal is fitted with a polynomial curve. After interpolation, a regular FFT or a PSD estimation using an Autoregressive (AR) model can be used. Both will produce smooth curves, but the results are still distorted and some assumptions have to be made. Spline interpolation assumes a polynomial fit and the AR method demands the choice for a suitable model order.

Laguna et al. introduced the Heart Timing method [21], which is a method that solves all shortcomings mentioned in the last paragraph. As shown in Figure 3.5 the signal is composed by subtracting the mean from the RR tachogram. The FFT of the interpolated result gives an estimate for the PSD without distortion. The Heart Timing method and the AR method are compared in Figure 3.4. The AR method produces smooth results, so the peaks are easier to detect, while the Heart Timing method is free of distortion. We use the latter for the calculation of power in the frequency bands, because the smoothness is of no relevance here, while distortion is. We use the AR method for visualization and comparison between different PSD's.

There is correlation between the metrics of the three domains. The RMSSD is correlated with SD1 and HF power, while SD2 is related to the LF power. Each domain has its benefits, so the combination will make the statistical evaluation more robust. There should be no imbalance between short- and long-term variability. The weights for all the individual error metrics are equal.

Some metrics will change depending on the duration of the length of the measurement. The total variance of HRV increases with recording length. The recording length also determines the smallest frequency power that can be accurately measured. Ectopic beats and distorted peak detections heavily influence most of the metrics. The removal of these beats prior to analysis will greatly improve performance.

3.3. Analysis of Atrial Conduction

The PR interval gives information about the atrial conduction and the AV node. Next to calculating the average and the standard deviation of the PR intervals, we can calculate the autocorrelation and we can inspect the histogram.

We can also analyze the atrial conduction without considering the PR interval. The Functional Refractory Period of atrioventricular Node (FRPN) denotes the shortest RR interval and by including the PP intervals in

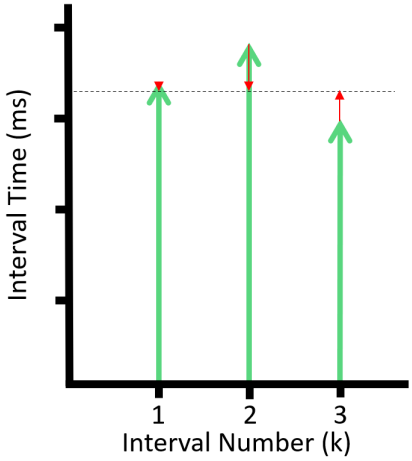


Figure 3.5: The Heart Timing signal shown in red is composed of the deviations from the mean interval length.

the analysis, we are in a position to calculate the Effective Refractory Period of atrioventricular Node (ERPN). This metric is equal to the shortest preceding PP interval that does not result in a response from the bundle of His [10].

Another geometric option includes a histogram of the PR intervals. In a histogram, it is possible to see if both pathways in the AV node are transducing signals. If this is the case, this would result in a bimodal histogram. In the case of Heart Block, some PR intervals may be nonexistent. If an impulse is blocked by the AV node, an R wave will not follow the P wave. Missing beats are expressed as ratios in Heart Block patients.

We can use the autocorrelation coefficient to plot the correlation between the current PR interval and previous PR intervals. The lower lags will be the most interesting, as the higher lags show less correlation. Some AV node characteristics may be noticeable here. The autocorrelation coefficient is the normalized autocorrelation function with the mean extracted. The formula for the consistent and unbiased estimator is given in 3.4.

$$r_{uu}(\tau) = E\left(\left(\frac{u(t+\tau)-\mu_u}{\sigma_u}\right)\left(\frac{u(t)-\mu_u}{\sigma_u}\right)\right) \tag{3.4}$$

Here, the signal u is multiplied with itself at various lags. The signal mean is denoted by μ_u and the signal standard deviation is denoted by σ_u. The result is r_{uu}, the autocorrelation coefficient as a function of the lag. An example is plotted in Figure 3.6. The sawtooth pattern is the result of the recovery property of the AV node, because a decrease in PR interval length is followed by an increase and vice versa.

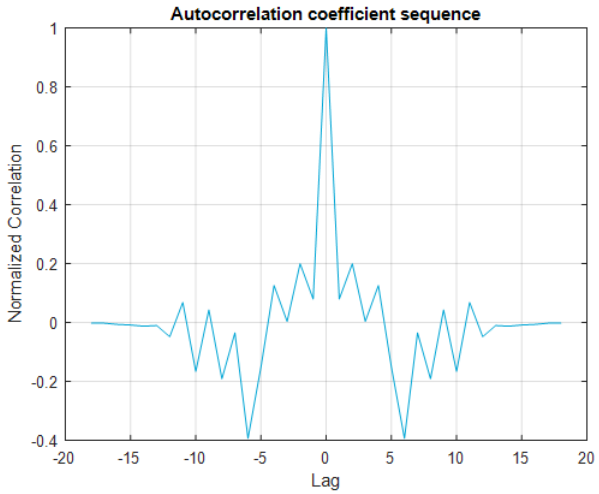


Figure 3.6: The autocorrelation coefficient sequence is plotted for PR intervals from ECG data.

4

The Extended Integrated Pulse Frequency Modulation Model

In this chapter we will describe the IPFM model in more detail and we will illustrate how we expand the classical SA node model with the AV part. We start with the IPFM model in general and its use in the modelling of the SA node in Section 4.1. Then we will describe how the same idea is translated to the AV node model in Section 4.2. Both sections end with an overview of the model parameters. The last step will then be the parameter estimation, for which we use an Evolutionary Algorithm (EA). The idea behind the estimation method and the consequences are outlined in Section 4.3.

4.1. The IPFM Model for the SA node

The Integrated Pulse Frequency Modulation (IPFM) model is the most commonly used physiological model for the generation of artificial RR intervals. The idea for this model was first described by Li and Jones [3]. The structure works as an "integrate and fire" system, which is inspired by neural communication. The diagram in Figure 4.1 shows the model that generates the R peaks.

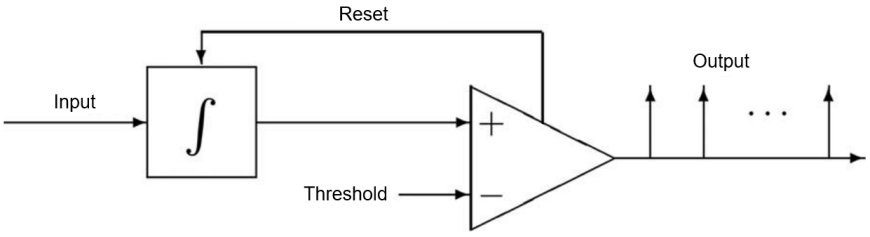


Figure 4.1: The diagram of the IPFM model. Image adjusted from [22].

The integrating unit corresponds to the SA node in this case. It will integrate the input until a threshold is reached. The input consists of a constant part and a variable part. The constant part is analogous to the ion leakage occurring in pacemaker cells. In case that this is the only input, the HR would be constant. The variable part of the input originates from the ANS. This input follows a sum of sinusoids, in this case two. One corresponds to the sympathetic nervous system, the other to the parasympathetic nervous system. The threshold corresponds to the membrane potential threshold. An action potential will occur when this threshold is reached. In the particular case of this model, the output is composed of a pulse train. Each pulse represents an R peak. When the peak is generated the integrated input will be reset. The formula showing the n impulses is as follows:

$$n = \int_0^{t_n} \frac{1+\mu(t)}{T1} dt. \tag{4.1}$$

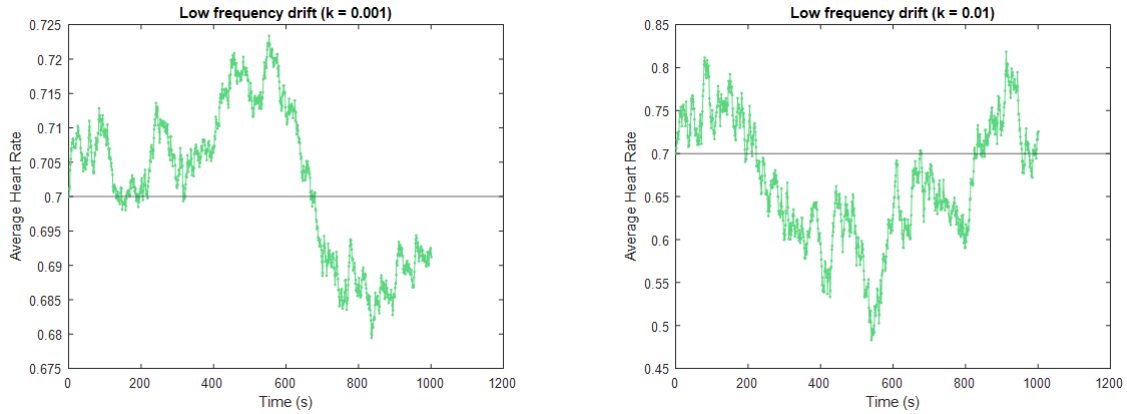
Here, n is the impulse number, $T1$ is the threshold value and $\mu(t)$ is the modulation signal, which is composed of two frequencies FL and FH with amplitudes AL and AH. We can divide the fraction into two

components. The part $\frac{1}{T_1}$ represents the near constant or low frequency factor. The part $\frac{\mu(t)}{T_1}$ represents the variable or high frequency part of the HR.

The low frequency factor $T1$ is not necessarily constant, especially over longer periods of time. For the hearts of living humans, changes in mood or surroundings have a low frequency impact on the HR. Bailón et al. [22] coined the idea of a time-varying threshold. Their study regarded the time varying threshold as a low frequency influence and they estimated its value in both a simulation study and a stress testing database. We take inspiration from this idea by including a time varying threshold in a forward model. We use a random walk to allow the variable $T1$ to wander. A central drift is included to prevent non physiological heart rates. Formula 4.2 expresses this behavior. Only one parameter, k, governs the nature of the drift.

$$x(n+1) = x(n) + k\left(V(n) + T_1 - x(n)\right), V(n) \sim N(0,1). \tag{4.2}$$

Here, the value of the threshold on moment n is given by $x(n)$. The space between successive values of n is $1/f_s$, where f_s is the sampling frequency. The middle term is White Gaussian Noise (WGN) with unit variance. The last term is the drift towards the estimated average threshold value of the SA node model. The effects of two extreme values for the parameter k are plotted in Figure 4.2. A very low value of the k parameter allows more drift, while a high value keeps the value closer to the mean.



(a) Low value

(b) High value

Figure 4.2: A low value of k results in lower variance compared to a high value.

The IPFM model is a well known model for NSR, but not necessarily for AF. During AF, aspects of the heart which are not modelled here get a more noticeable influence. This will result in a worse fit for the modelled tachogram, so lastly we add $noise1$, the final parameter. The model can use this parameter to compensate for the part that is not modelled by the rest of the model. The parameter itself determines the variance of the WGN that is added to the intervals, simulating the inherent stochasticity of the mechanisms that underlie the biological processes. An overview of the seven parameters of the SA node model are summarized in table 4.1. Lastly, the output of a model for the SA node is supposed to be a P wave. As described in the research goal, we will add a model for the AV node, which produces an R peak.

Table 4.1: The parameters for the SA node model

Parameter	Explanation
T_1	SA node threshold
k	Governs the low frequency drift
FL	Low modulation frequency value
FH	High modulation frequency value
AL	Low modulation frequency amplitude
AH	High modulation frequency amplitude
noise1	Variance of the added WGN

4.2. The IPFM Model for the AV node

The AV node is a highly complex structure. A physiological model should capture the most significant aspects of the structure. Similar to the SA node, we also model the AV node with an IPFM-based model. The AV node has automaticity. The ANS influences the conduction, but the input from the SA node is much more significant. The repolarization phase is crucial in the AV node. The possibility of a long recovery time allows it to block fibrillatory waves from the atria. The block diagram of a potential configuration of this version of the IPFM model is shown in Figure 4.3. The two model blocks for the SA node and the AV node are cascaded. The output of the SA node is modelled as a square wave and serves as the input of the AV node model. The output of the AV node model is the R peak. The AV node model can also be used on its own on a dataset where both P peaks and R peaks are annotated.

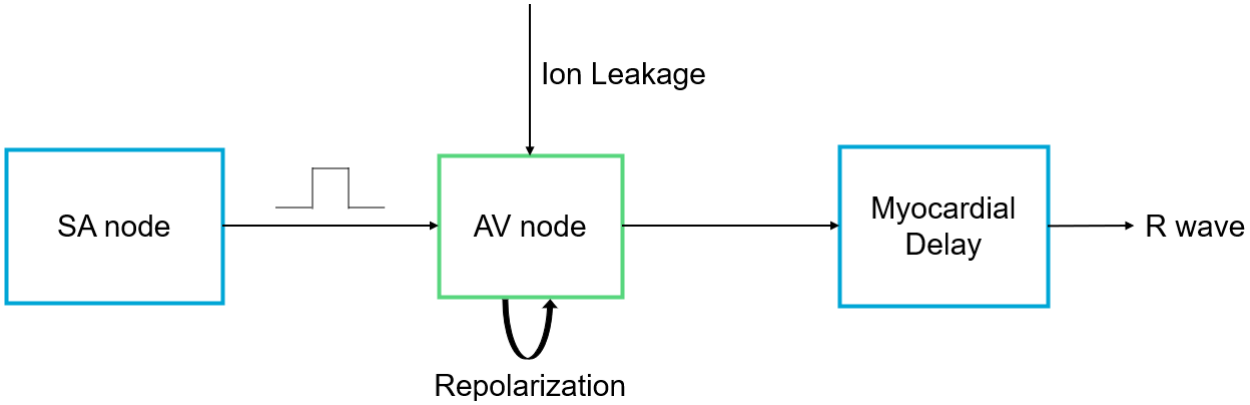


Figure 4.3: The schematic showing the IPFM model for the AV node.

The equation that characterizes the "integrate and fire" part of the model is similar to (4.1). The repolarization prevents the build-up of charge during this phase. The formula is given below:

$$\int_{t_{k-1}+r}^{t_k} \frac{1+\mu(t)}{T_2}\, dt = 1. \tag{4.3}$$

Here, the moment of output is denoted by t_k. The previous output pulse is t_{k-1}. The variable r describes the repolarization time, $\mu(t)$ the modulation signal and T_2 the threshold. The output of the SA node and thus the input of the AV node model is assumed to be a square wave. Two parameters model the shape: the height, and the width. The pulse will look like a narrow rectangular pulse for a healthy heart. For AF however, the pulse can model the combination of fibrillatory waves as a broad wave with a lower amplitude. The formula is given in (4.4).

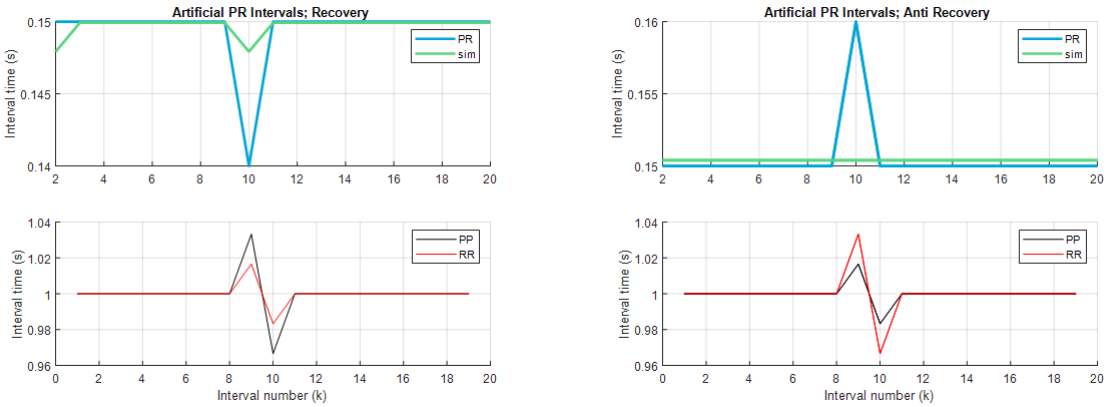
$$\mu(t) = \sum_{n=1}^{N} h(u(t-\tau_n) - u(t-w-\tau_n)), \quad w * f_s = \mathbb{Z}^+. \tag{4.4}$$

The modulation signal $\mu(t)$ resembles the sum of all P waves. The height of a P wave is denoted by h and its width is denoted by w. Also, $u(t)$ is the unit step function. The vector τ contains the time occurrences of the P waves. The signal is discrete, so the width of the pulse multiplied with the sampling frequency f_s must equal a positive integer.

Table 4.2: The parameters for the AV node model

Parameter	Explanation
T_2	AV node threshold
tw	Repolarization time
pw	Atrial pulse width
ph	Atrial pulse height
dmin	Controls repolarization fatigue
noise2	Variance of the added WGN

The fifth parameter is $dmin$. It adjusts the repolarization time, to allow the simulation of fatigue. As a last step we introduce a second noise parameter $noise2$, which simulates the variability in atrial conduction. The patterns of behavior of the AV node are described in Chapter 2. In Figure 4.4 we see the model is able to fit the shape of an artificial sequence if this sequence simulates a property of the AV node. When the opposite effect is simulated, the fit is a straight line.



(a) Low value (b) High value

Figure 4.4: The AV model will fit the shape of the left artificial sequence using recovery. It does not find a fit for the right sequence, where the oppostite of recovery is simulated.

When a sudden new pace is introduced by the SA node, the AV node will follow in steps following an exponential function [23]. This effect is seen when we introduce such a shift in pace with an artificial sequence as shown in Figure 4.5. These tests with artificial intervals show that the model structure contains specific information , which results in behavior that is typical to the AV node.

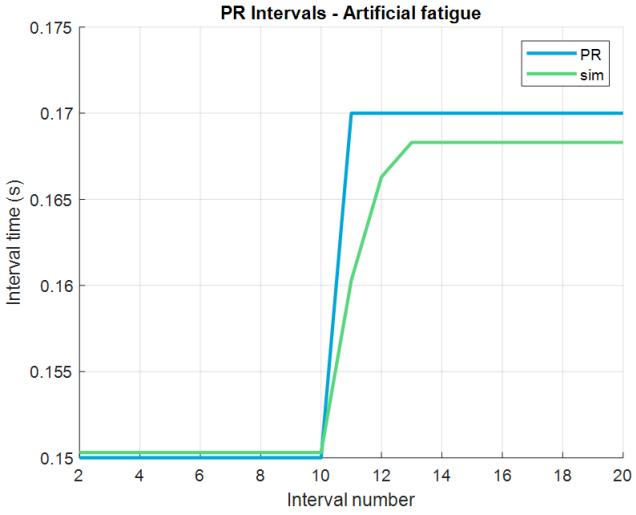


Figure 4.5: The model follows an exponential trajectory with the introduction of fatigue.

An overview of the six parameters in the IPFM model for the AV node is shown in Table 4.2. Some properties of the AV node are still not included in the model. There are for example no dual pathways present in the model. The absence of this structure will result in worse fits for AF, where both pathways are often used.

4.3. Parameter Estimation Using an Evolutionary Algorithm

4.3.1. Motivation

We use an evolutionary algorithm for the parameter estimation, because the solution space is non-convex. A technique like gradient descent, where the parameters are iteratively changed towards the steepest descent in error, is not possible here, since we would get stuck in local minima. In contrast, an evolutionary algorithm will find its way even without any knowledge of the system that is being optimized. Also, this algorithm is easily adaptable. An extension to the model will lengthen the process of parameter estimation, but it will not change the process or strategy.

4.3.2. The Algorithm

An evolutionary algorithm finds the model parameters by random mutation of the parameters and selection based on specified errors [24]. It is a method that is commonly used in optimization problems where the system is highly non-linear. Pseudo-code for a simple evolutionary or more specifically a genetic algorithm is shown in Figure 4.6. Here, the population P is a group of parameter sets. The survivors per generation compose S. They generate a new generation: $\mathcal{O}$, which has the same size as P. The perturbed set $\mathcal{O}'$ is obtained after mutation. Hyperparameters include the number of parameters, the number of survivors per generation, the number of children per survivor and the mutation rate.

Simple Genetic Algorithm

STEP 1. INITIALIZATION
Generate initial population $\mathcal{P}$ at random or with prior knowledge
STEP 2. FITNESS EVALUATION
Evaluate the fitness for all individuals in $\mathcal{P}$
STEP 3. SELECTION
Select a set of promising candidates $\mathcal{S}$ from $\mathcal{P}$
STEP 4. CROSSOVER
Apply crossover to the mating pool $\mathcal{S}$ for generating a set of offspring $\mathcal{O}$
STEP 5. MUTATION
Apply mutation to the offspring set $\mathcal{O}$ for obtaining its perturbed set $\mathcal{O}'$
STEP 6. REPLACEMENT
Replace the current population $\mathcal{P}$ with the set of offspring $\mathcal{O}'$
STEP 7. TERMINATION
If the termination criteria are not met, go to STEP 2

Figure 4.6: Pseudo-code for a basic evolutionary algorithm. Image taken from [24].

We perform the fitness evaluation and selection by picking the parameter sets with the smallest total error. This error is the summation of the relative errors of the metrics described in Chapter 3 compared to the data sequence. Crossover is skipped and the mutation rate is increased when the relative decrease in error does not drop. The mutation is relative to the current value of the parameter and all parameters are mutated at the same time. This approach prevents the algorithm from getting stuck in local minima. The two best performing sets will survive and compete with the new generation. The algorithm terminates when the error has not gotten smaller and the number of generations is more than twice the number of generations it took to find the last best performing parameter set.

4.3.3. Drawbacks and Possible Improvements

Due to the stochastic nature of the model, solutions will be slightly different each time. On the one hand, running the EA multiple times will give slightly different parameter sets. On the other hand, a parameter set will give a slightly different error each time they are used to generate an interval sequence. This variability evens out over large samples of data, so it does not pose a problem for training a dataset. When analyzing a single sequence it may be advisable to run the EA several times and take the average parameter values of the runs.

A higher number of parameters greatly increases the run time. This phenomenon can be attributed to the "curse of dimensionality". If we find the parameter sets for both the SA node and the AV node separately based on just R peak data, this method will greatly reduce the flexibility and the ability to alter the SA node parameters based on AV node behavior. The parameters of the AV node part would essentially be fitting the error of the SA node model. If we find all 13 parameters in one run, the parameter space will be very large. The algorithm might get stuck in a local minimum in the highly dimensional solution space.

We can take some extra steps to help the EA. By imposing physiological constraints on the parameters we can reduce the solution space and with it the run time. By imposing boundaries we also prevent non physical solutions. The mean of the RR tachogram directly relates to the threshold of the SA node, so we can constrict the boundaries for this parameter to the calculated mean ± 10%. We can also detect the peaks in the PSD of the tachogram data and constrict the boundaries of the modulation frequencies in the same way. If all goes well, the error steadily drops towards a stable value. The error per generation of a successful run is plotted in figure 4.7.

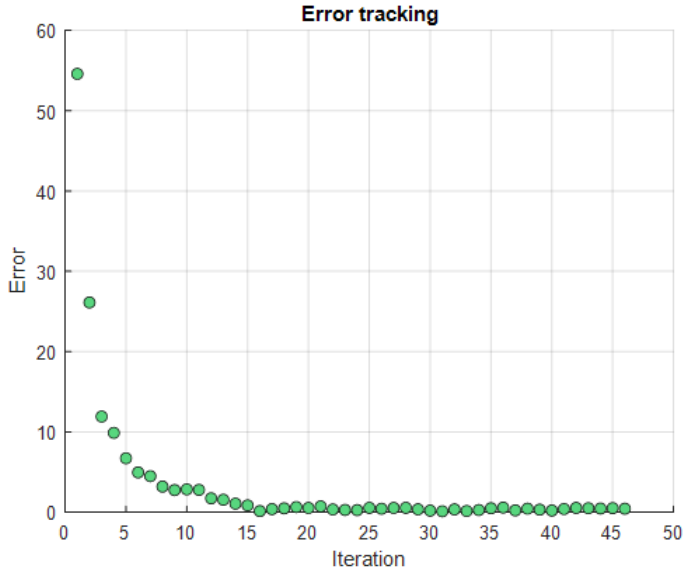


Figure 4.7: The smallest error of the population is plotted against the generation number.

5
Methods

In this chapter we describe the methods for testing the validity and the effectiveness of the model. We will first list the data that we use to then specify the tests we constructed. For all tests, either the SA node model, the AV node model or both models are fitted to a tachogram. The results of the tests are shown in the following chapter.

5.1. Datasets

We use datasets from PhysioNet databases [25]. PhysioNet is a public service of the PhysioNet Research Resource for Complex Physiologic Signals and they provide open-source data and software. For our tests we will use the locations of P waves and R peaks and we will need data for NSR as well as for AF. We can use ECG data and detect the peaks ourselves or we can use tachogram data to remove this delineation step.

5.1.1. Tachogram Data

Since all our tests are based on tachograms the simplest method is to use tachogram data. We use the dataset from the PhysioNet challenge "Is the Normal Heart Rate Chaotic?", which includes RR tachograms for both NSR and AF for tests where we just need the R peaks and these two categories. By using a dataset of intervals there is no risk of producing errors in the peak detection step. This "Chaos" dataset even provides filtered versions of the tachograms where outliers and ectopic beats have been removed.

5.1.2. ECG Data

We can use tachogram data, but we can also use annotated ECG data and compute the tachograms directly from the annotations. When there are no annotations, we can detect the relevant peaks from the ECG data with delineation algorithms. The use of ECG data adds more information and thus adds more possibilities, but there is a risk of errors in the peak detection, which would negatively impact the performance of the model.

The tachogram dataset only contains R peaks. We use the "ECG-ID" database for NSR signals where we need the P waves [26]. We use the delineation algorithm from the WFDB Software Package from the PhysioToolkit to detect the peaks of the P waves. The data for AF is taken from the Brno University of Technology ECG Signal Database with Annotations of P Wave or "BUT PDB" [27]. The P waves are only annotated in the absence of an episode. Lastly, we use the data for the PhysioNet Computing in Cardiology (CIC) Challenge 2017 as a dataset with ECG signals from NSR, AF and "other" arrhythmias. We do not detect P waves for this data, as they are ill-defined for AF.

5.2. Tests and Benchmarks

The tests described in this section aim to measure the ability of the model to simulate the behavior of a heart during NSR and AF. For all tests, we will fit parts of the physiological model to either RR tachograms or PR tachograms. We will first investigate the output of the model, which is the forward approach. Then we will investigate the parameters that underlie these outputs, which is the inverse approach. The output of the

model is a tachogram, so we will analyze both RR and PR tachograms. We will test the differences in the model parameters with statistical tests and use these parameters for classification.

5.2.1. Tachogram Generation

The model should be able to generate realistic RR tachograms, where the statistics as described in Chapter 3 are similar to the data. We fit the complete physiological model to data from the RR tachogram dataset. We compare the tachograms and their PSD's to check the similarity.

We can compare not only RR tachograms, but also PR tachograms. The prediction of PR intervals based on past PP intervals has been shown to be possible in isolated perfused rat hearts [23]. The same result will not be possible for a living human hearts, as it is much more complex. Still, it is worth trying to see if the past PP intervals help with the prediction. Next to the AV node model, we use two other methods. They will act as benchmarks, because there are no results in the literature to compare with. All three methods will be trained based on a training dataset and the error will be measured for a separate test dataset. We use the "ECG-ID" database for this test.

The first alternative method is the *histogram method*. We compose a histogram of PR intervals from the training set. New PR intervals for the test dataset can be generated based solely on PR intervals from this training data. If the new values are taken from the histogram of training PR intervals, there is no assumption needed about the underlying distribution of PR intervals. The only choice we have to make concerns the bin size. This *histogram method* will serve as the first benchmark.

The second alternative method will make use of an Artificial Neural Network (ANN). The ANN is trained on the training data with a certain number of PP intervals as input and the PR interval as output. We found 12 neurons in the hidden layer combined with 3 past PP intervals to have the best performance. Finally, the physiological model is trained on the training data where the P peaks serve as input and the R peaks serve as output.

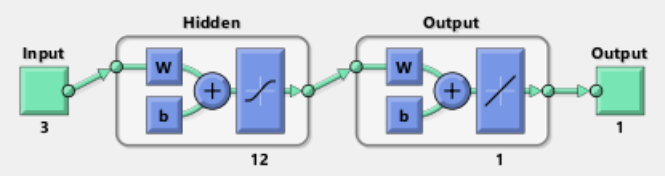


Figure 5.1: The neural network takes PP intervals as input and PR intervals as output. It has 12 neurons in the hidden layer.

We measure the performance of the methods using the summed mean squared error between the predicted PR interval and the PR interval from the test data. The mean squared error gives us information about the ability of the methods to predict PR intervals, but it does not show general behavior of the AV node. In addition to using the mean squared error as metric we can plot the autocorrelation coefficient, which is described in Chapter 3.

5.2.2. Parameter Significance and Classification

The model parameters should be significantly different for different pathologies, so the combination of the model parameters will then be able to classify ECG traces. We will first test the significance of the parameters individually. By using the Anderson-Darling test we conclude that the parameter values are not normally distributed ($P < 0.001$), so we use the Wilcoxon signed-rank test to check whether the parameters from different datasets are statistically different. It is a non-parametric method and it does not assume the data is normally distributed. The null hypothesis of the Wilcoxon signed-rank test states that the parameters are not significantly different. We use a P value of 0.001 for the rejection of the null hypothesis. We use the "Chaos" dataset for the complete physiological model and we use the "ECG-ID" dataset and the "BUT PDB" dataset to fit the AV node model on PR intervals. Lastly, the PhysioNet CIC Challenge of 2017 provides ECG traces from four different categories:

- Normal; Normal Sinus Rhythm
- AF; Atrial Fibrillation
- Other; Other arrhythmias. This category includes all kinds of pathologies, ranging from atial to ventricular problems.
- Noisy; The recording is too noisy

The challenge is to classify the ECG traces from the validation set into the correct categories. In order to compare the parameters for the CIC challenge we need a multicomparison test. This test relates a group of parameters to all the others.

Eventually, the physiological model should be used for interpretable classification. We will finish the tests with two classification challenges. We first use the "Chaos" dataset for binary classification, in order to test the basic functionality of discriminating between NSR and AF. Finallly, we try the PhysioNet CIC Challenge to test the ability of the model to distinguish between NSR, AF and other signals with ECG traces as input.

We classify by using the model parameters as input for an ML algorithm. An overview of the data pipeline is given in Figure 5.2.

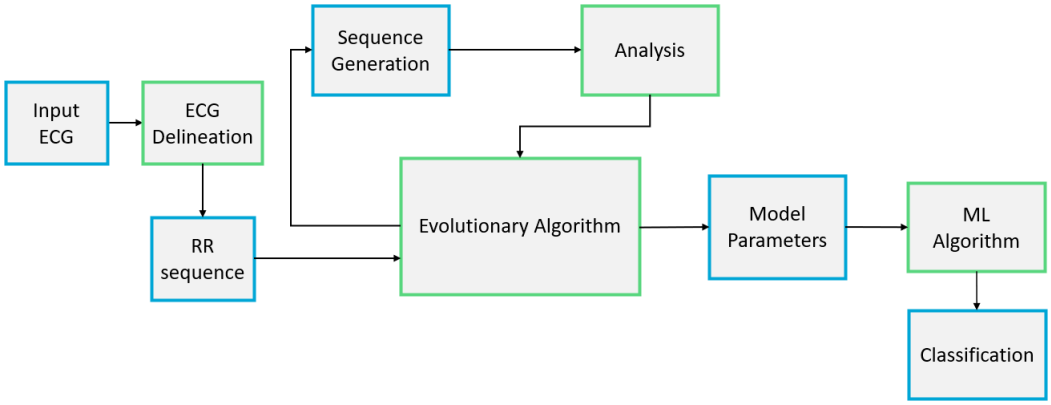


Figure 5.2: An overview of the scripts that give a classification result with an ECG trace as input.

The ML algorithm of choice is a Support Vector Machine (SVM) for binary classification and a Random Forest (RF) for the multiclass classification. We use the standard Matlab settings for both algorithms and the RF has 256 trees. We use 5-fold cross validation for the binary classification. For the CIC Challenge, we follow the approach from the challenge. We train the model on the available training data and the score is based on the performance on the validation dataset.

6
Results

In this chapter we will list the results of the experiments listed in the previous chapter. First of all, the model should generate realistic RR tachograms for both NSR and AF. The output test are split between RR tachograms and PR tachograms. Next, tests for the model parameters that are used to create these tachograms are split between statistical significance tests and classification tests. Next to presenting the results, we will also explain all result shortly.

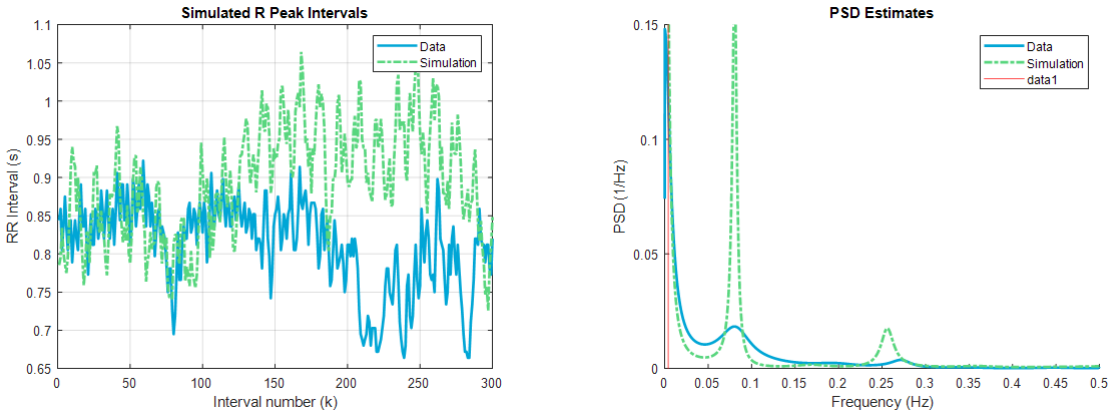
6.1. Realistic RR Tachogram Generation

First, we will aim to generate a realistic RR tachogram. A realistic tachogram should have similar statistical properties compared with a tachogram from real data. Tachograms from the "Chaos" dataset will serve as input for the Evolutionary Algorithm. The simulated output sequence is plotted together with the real data in figure 6.1 for NSR. The PSD's are plotted next to the tachograms. In Figure 6.3 the tachogram for AF is plotted.

The model generates RR tachograms that look realistic in the time domain. However, when looking at the PSD we see the peaks are much sharper for the simulated data. The error is calculated on power within the frequency bands. Since there are only two frequency modulators for the model, these frequencies have a high amplitude as a compensation. The modulating frequencies are constant during the simulation and they have to provide the power for the entire frequency band. In Figure 6.2 the modulating frequencies are allowed to wander just like the threshold. The resulting PSD looks much better. Unfortunately, the run time is much longer due to the iterative calculation method. Although the result looks more realistic, it makes no difference for the eventual parameter. There will be no improvement in classification for this reason.

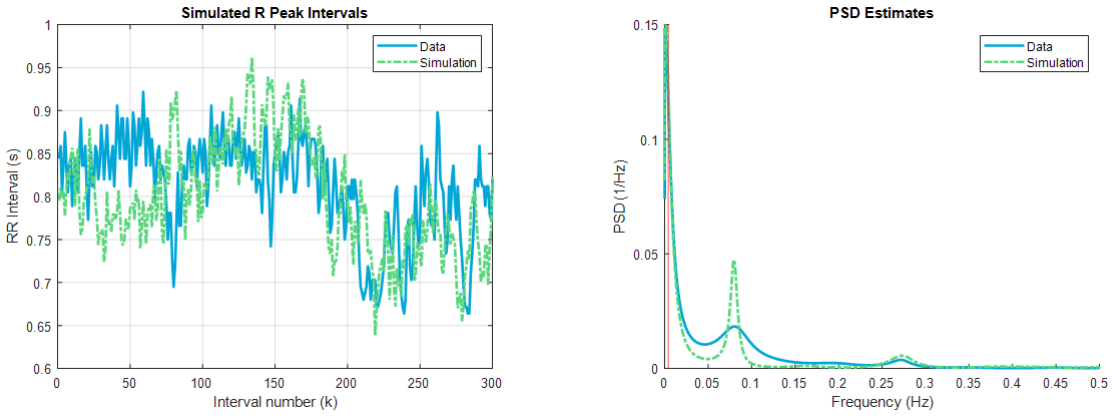
The tachogram is visibly different for AF. Both the data and the simulation give a PSD plot with less structure. The two peaks commonly seen in PSD's of NSR hearts are not clear in this plot. Instead, the average power is higher and more smaller peaks seem to appear. The model is able to raise the overall power with the noise parameter and there are more than two peaks, but the peaks are not always in the correct position.

We can conclude that the physiological model in combination with the evolutionary algorithm is able to generate tachograms that look realistic by eye, but are distinguishable from real data with analytical methods. Still, the sequences are significantly different for NSR and AF, which is the most important in the end.



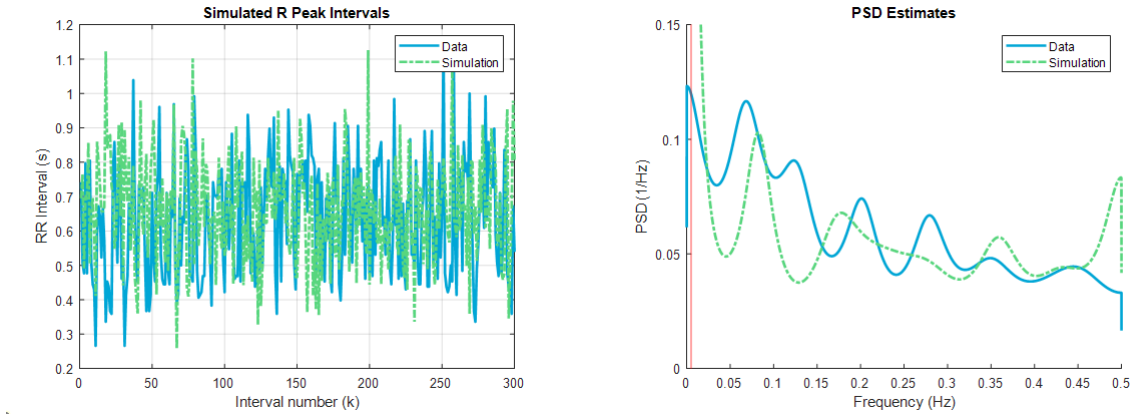
(a) Sequences (b) PSD

Figure 6.1: The real and simulated NSR RR tachograms are plotted together with their PSD's.



(a) Sequences (b) PSD

Figure 6.2: The real and simulated NSR RR tachograms with the variable modulation frequencies are plotted together with their PSD's.



(a) Sequences (b) PSD

Figure 6.3: The real and simulated AF RR tachograms are plotted together with their PSD's.

6.2. PR Interval Prediction

When we fit the AV node model to a PR tachogram, we get the intervals as shown in figure 6.4.

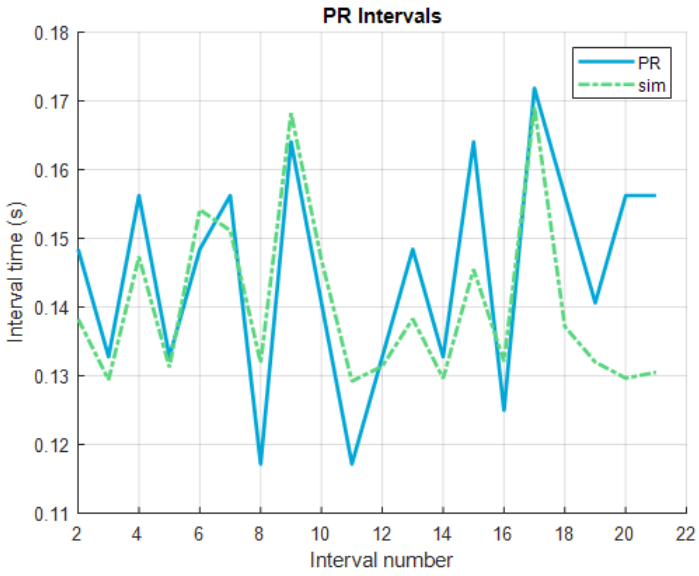


Figure 6.4: The PR interval sequences are plotted against the data.

The physiological model is able to capture the shape and characteristics of the intervals. In Chapter 4 we have shown that the model is able to fit individual characteristics emulated by artificial intervals. Here, we show the combination of these characteristics can be fitted in real data. Still, it does not mean the model is able to generalize well. We cannot check for overfitting with just this plot. So, we plot the validation tachogram together with the predictions of both the model and the alternative methods as described in the last chapter. We plot the PR interval sequences in Figure 6.5. Then, we run the methods for 100 times and plot the boxplot of the mean squard error in Figure 6.6. In this figure we also plot the autocorrelation coefficients of the data and one of the generated sequences for each method.

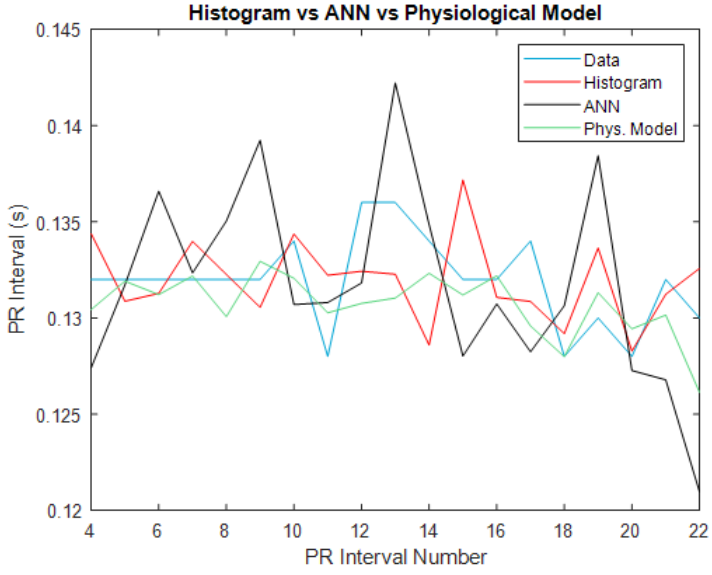
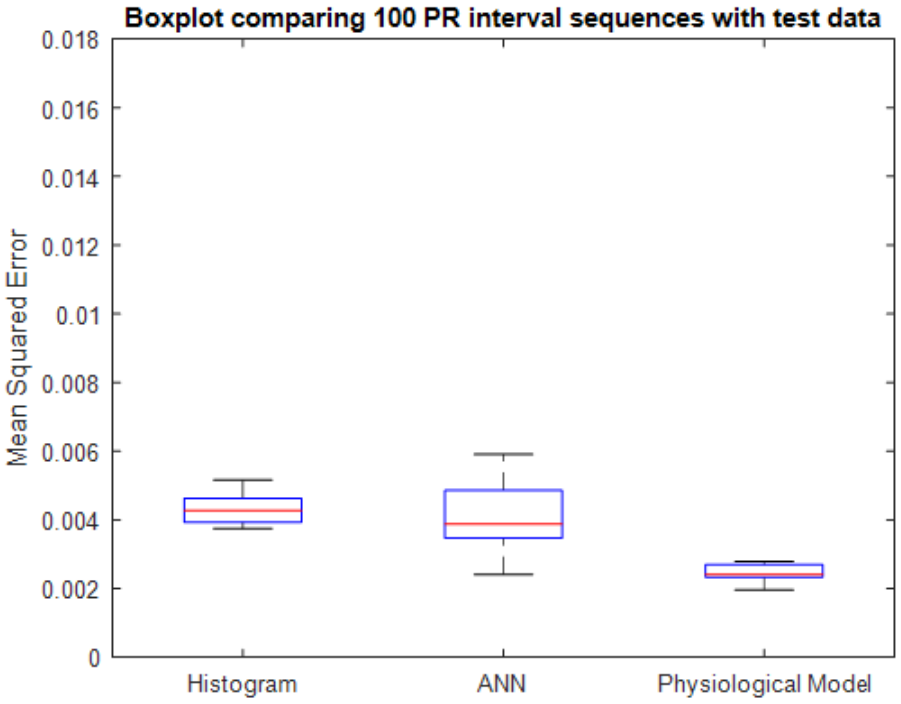
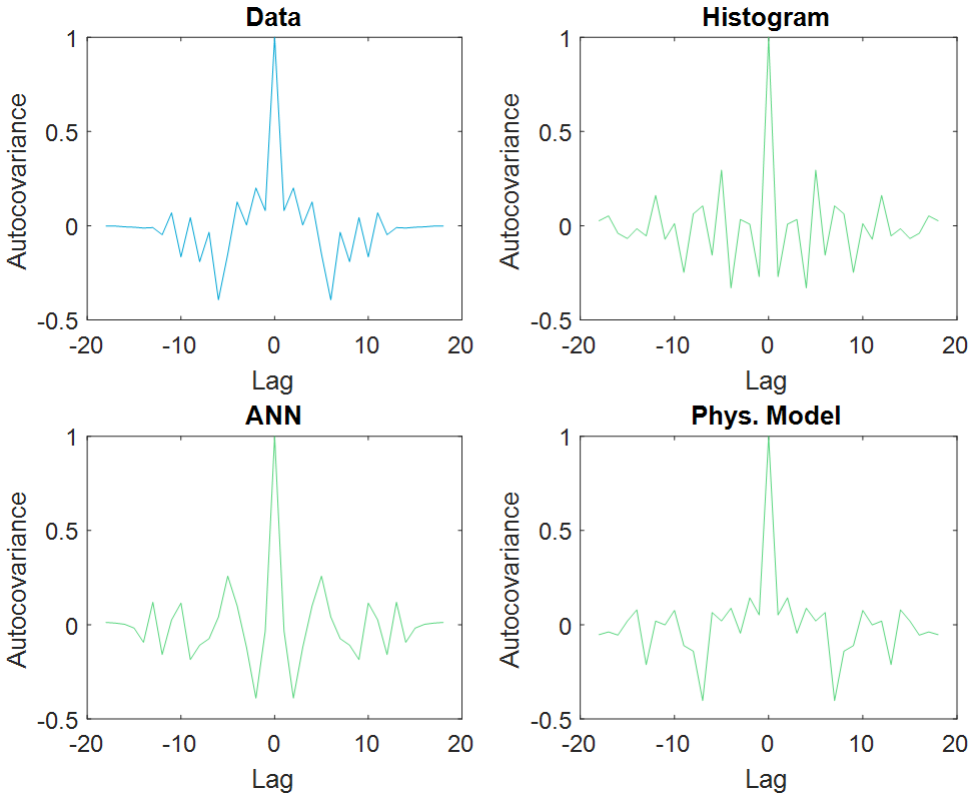


Figure 6.5: The PR interval sequences are plotted.



(a) The mean squared error of the predicted PR intervals versus the test data over 100 runs.



(b) Autocorrelation plots of the real data (upper left) and the simulated PR interval sequences.

Figure 6.6: Analysis of the PR interval predictions. a) Boxplot of the mean squared errors. b) Autocorrelation coefficient.

The physiological model performs better than the alternatives. Using the Wilcoxon signed-rank test, we calculate that the error of the physiological model is significantly lower compared to both the histogram method and the ANN ($P < 0.001$). The prediction is naturally not as accurate as the one achieved by Heethaar et al. with the isolated perfused rat hearts. A living human heart is a lot more complex and unpredictable, which is probably the reason the prediction of PR intervals has not been done before here. We can only compare our results to the alternative methods.

The physiological model performs better than the histogram method because it is able to capture some influence from the PP intervals. The reduction in mean squared error proofs there is useful information present in the P wave data. In theory, the ANN should outperform the other methods. It is able to use the information of the PP intervals and it allows for the most complexity. However, an ANN needs a lot of data. This neural network would probably perform much better with more data. Next to interpretability, data efficiency seems another advantage for the physiological model.

The sequence of the histogram method should not show significant correlation, because the intervals are picked randomly from the distribution with no correlation to the previous ones. The ANN has a different plot, but once again this is probably due to the lack of sufficient data. The physiological model matches the data quite well on the low lags. When the lags get higher, the correlation differs, as expected. The added noise over multiple lags will add up.

We can conclude that the AV node model is not only able to fit artificial sequences, but also real PR interval data. The model is able to generalize this result to a validation dataset by outperforming the alternative methods. There has to be information present in the location of the P waves for this improvement to be possible. From the autocorrelation coefficient we see the shape of the correlation for the lower lags is similar, so we can conclude the model captures some physiological characteristics of the AV node.

6.3. Significance of the Model Parameters

The model parameters must discriminate between different conditions of the heart. We will try to classify ECG's in the next section, but first we will test the significance of the parameters individually. The P values for the acceptance of the null-hypothesis of the Wilcoxon signed-rank test for the SA node model parameters are given in Table 6.1. The null hypothesis states that the parameters for the different categories come from the same distribution. This means that a lower P value stems from a more significant difference.

Table 6.1: P values for the Wilcoxon signed-rank test to determine whether the SA node model parameters come from the same distribution.

T_1	k	FL	FH	AL	AH	noise1
6.07e-22	6.28e-6	0.068	0.012	7.81e-9	1.44e-4	0.054

Five of the seven parameters are significant. The threshold value T_1 directly relates to the average interval length. The average HR is higher for AF patients, so this parameter is different. The parameter k is significant as well. It makes sense, because the parameter controls the low frequency drift. The drift is present during NSR, but when AF is present the control over the HR decreases. The parameter is now used by the model to find a better fit without much of a physiological equivalent. The modulating frequencies are somewhat significant. Lastly, the noise is able to compensate for the error. The model structure works best for NSR, so the noise is generally higher for AF sequences.

We will first show the result of the parameter significance test when the parameters from the two models are found iteratively. This means the parameter set for the SA node model is found first and the output serves as P wave data for the AV node model. The result is shown in Table 6.2.

Table 6.2: P values for the Wilcoxon signed-rank test to determine whether the AV node model parameters come from the same distribution.

T_2	RW	PW	PH	RMIN	noise2
0.26	0.92	0.11	0.93	0.35	0.066

None of the parameters are significant, because separate parameter estimation does not provide sufficient flexibility. When we start parameter estimation for the AV node model, the parameters for the SA node model will already be fixed. This means the AV node model will essentially be fitting the error of the SA node model.

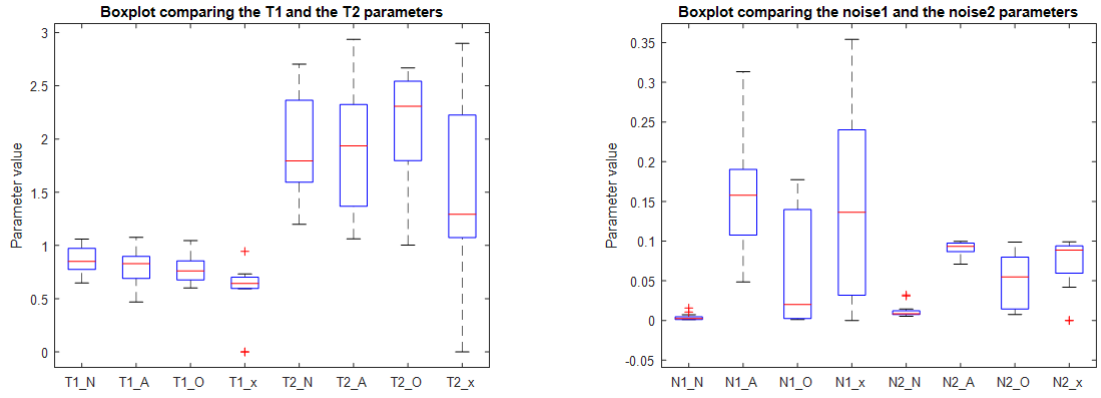
A better approach would be to fit the AV node model with the PR intervals for prediction. There is very little useful data, which means the P values will also be less significant. We plot the P values in Table 6.3.

Table 6.3: P values for the Wilcoxon signed-rank test to determine whether the AV node model parameters come from the same distribution.

T_2	RW	PW	PH	RMIN	noise2
0.026	0.053	0.80	0.55	0.71	1.0

We see the $T2$ and the RW parameters seem different. They are both related to the AV node. The data for the AF case is taken right after an episode, so the change is behavior might be a consequence of the episode. The combination of the relatively low values for the parameters related to the AV node in combination with a possible physiological explanation makes this result promising, but more data is needed to test if the difference is significant.

Lastly, we compare the parameters when there are four different categories with the PhysioNet CIC challenge. All parameters are estimated at the same time for this test. The boxplot is shown in Figure 6.7.



(a) Thresholds

(b) Noise parameters

Figure 6.7: These boxplots show the parameter values of $T1$, $T2$, $noise1$ and $noise2$ for Normal (N), AF (A), Other (O) and noisy (x) data.

The boxplot shows a lot of overlap between the parameter values for the different groups. The AV node model parameters provide little extra information. These results do not have a coherent explanation and the parameter sets do not look consistent. A consistency test will show whether it is possible to find consistent parameter sets when the SA node model and the AV node model are combined. The results are shown in Table 6.4.

Table 6.4: Coefficient of variance values for the parameter consistency test.

	T_1	k	FL	FH	AL	AH	noise1
SA	0.0018	0.0014	0.0154	0.0102	0.0182	0.0074	0.0039
SA+AV	0.747	0.0137	2.085	0.466	0.0779	0.388	1.038

The table shows a significant difference between the consistency of the SA node model parameters with and without the addition of the AV node model. We can conclude there is not enough information in the RR tachogram alone to find consistent parameter sets for both the SA node model and the AV node model. The combined model results in an overdetermined system.

6.4. Classification Results

After the parameter sets are found, a Support Vector Machine (SVM) is able to perform binary classification. The classification accuracy is 98.7% for just the SA node model alone. The inclusion of the AV node model does not improve the accuracy, but the significance of this result is difficult to assess, because the number of wrong classifications is small. A more extensive test is required to determine whether the AV node parameters add value.

The result of the PhysioNet CIC challenge is summarized in the confusion matrix shown in Table 6.5. The total accuracy is 64%. First we note the "Noisy" category is poorly recognized. There is no separate method for detecting noisy data, so the sequences are detected based on abnormal parameter sets. Both the sensitivity and the specificity are below 50%.

The algorithm reaches a sensitivity of 81% for NSR and there are no false positives for AF when the heart is healthy. The largest groups of mistakes are related to the "Other" category. This category contains various tachycardias, so some will be similar to NSR while others will be comparable with AF. For example, ventricular fibrillation will result in a similar parameter set as atrial fibrillation when the P wave data is absent. We can conclude that the addition of the AV node model does not provide an increase in prediction accuracy when there is no P wave data.

Table 6.5: Confusion Matrix

		Predicted Classification			
		Normal	AF	Other	Noisy
Reference Classification	Normal	122	0	19	9
	AF	8	20	16	6
	Other	23	6	36	5
	Noisy	3	7	6	14

7

Discussion and Future Work

The rationale behind the usage of a physiological model for diagnosing cardiac arrhythmias is the increase in interpretability. However, the approach may result in a drop in accuracy. If a deep neural network achieves a higher accuracy there will be a trade-off between accuracy and interpretability. However, an interpretable result in combination with further examination may lead to a better understanding and a better diagnosis in the end. A combination of the two methods is possible too. If the algorithms disagree, further testing could prove useful. Also, interpretability may not be the only benefit of a physiological model. The model outperformed the ANN for PR interval prediction. The ANN should outperform the physiological model when there is unlimited data, because the network will capture properties that are not modelled in the AV node model. We argue that a physiological model has a higher data efficiency, because there is already information present in the structure of the model, even before there is any input data. A more efficient model could result in shorter test times and shorter waiting times for patients.

Data snooping occurs in machine learning when decisions are made concerning the algorithm after looking at the data. The observer may for example choose a particular architecture based on a pattern seen in the data. This pattern however, might just be a statistical incident. When we use a physiological model, there is already a lot of information present in the architecture of the model. It is different from data snooping, because the patterns are well documented in medical literature. The structure of the heart contains valid information that can be used in a physiological model.

Our model assumes the output to be stochastic. However, whether the heart rate is stochastic or rather chaotic is a topic of debate [28]. This debate even resulted in another challenge by PhysioNet. A system is chaotic when small perturbations or differences in starting positions result in wildly different outcomes. Toker et al. [29] stated that the heart rate is not chaotic, but rather stochastic. Running their algorithm on the data used in this thesis returned the same result.

As described in Chapter 3, some metrics behave differently for varying recording lengths. The total variance of HRV for example increases with the recording length. The length also directly restricts the lowest relevant frequency. Our Evolutionary Algorithm takes the relative error of the power in the LF band with constant boundaries and the SD is not related to the recording length. Compensation for the recording length in the error calculation could improve accuracy when working with different lengths. However, the model works best for similar recording lengths, so the safest approach would be to fit the model on measurements with a similar length to the training data.

The Withings ScanWatch provides the option of an ECG measurement. It is single lead ECG and the resolution is low. Still, it is able to detect episodes of AF. Nowadays there are even more ways to measure HRV. The possibility of measuring HRV from photoplethysmography (PPG) is shown [30]. PPG uses infrared light to illuminate the skin. The light is more strongly absorbed by blood compared to the surrounding tissue, so amount of returning light reflects changes in blood flow. This find could allow the calculation of HRV metrics for a smart watch. The PPG technology may provide the same possibilities as the ScanWatch, but cheaper. Data security becomes an important aspect when health analysis with a smart watch becomes mainstream.

The parameter significance test for the AV node model fitted on PR interval data shows a promising result, but more data is needed to validate the AV node model and its ability to differentiate between NSR and AF. This result could be the result of different data types. The P waves for the NSR data are detected with a delineation algorithm, while they are annotated for the AF data.

Although the end goal of the model is classification, the goal of this thesis was to expand the model and test if the expansion is able to model AV node characteristics properly. We can conclude the model is able to do so, provided there is P wave data. The reason this is needed shows the difficulty in combining two imperfect stochastic models. If the SA node model would have been perfect and complete and the parameters are known, the RR tachogram would be sufficient. This is the case because the RR interval is composed of the part modelled by the SA node model and the part modelled by the AV node model. If we know the SA node perfectly, we would end up with the same information as provided by the PR intervals. Unfortunately, the model is not complete, especially for AF. Also, the parameters are not perfectly estimated due to the inherent stochasticity of the model. This means there is no ground thruth and the AV node model will model both the AV node part combined with the error of the SA node model.

In the PhysioNet CIC challenge, the algorithm by Datta et al. [31] achieved the highest score. Their approach was feature extraction and classification with an ML algorithm. Improvements to the ML learning step should improve accuracy. In order to increase the score further for the PhysioNet CIC challenge, a separate prior step should be taken to filter out the noisy data, because the resulting parameters are unpredictable. In order to increase the classification performance with respect to the "Other" category, more research is needed. We discuss further possible research below.

Future Work
Several small improvements are possible with respect to the current approach:

- *Adjustable weights for the calculation of the total error*: The error for the statistics of an RR tachogram related to the data is calculated as the sum of the error of each metric. However, some metrics might be more important than others. A set of weights may improve performance of the EA.
- *Include the modulating frequencies for the AV node model*: The modulating frequencies can influence the AV node model. The hormones norepinephrine and acetylcholine influence the atrial conductivity. No extra parameters are needed, but conflicts may arise between the two model parts.
- *Optimized EA*: Right now the evolutionary algorithm is written by manually, but standard packages exist. Self-written code is easy to adjust and interpret, but it is not optimized for speed. Packages should be used once the model structure is set to reduces both training time and classification time.

The most important result from this thesis is the AV node model fit for PR interval data. The model is able to find a fit and outperform alternative methods on a different dataset. The parameter significance test and the classification test show that P wave data is crucial for correct parameter estimation for the AV node model. Naturally, the most obvious way to go forward with the model is to detect the P waves and use them as input for the AV model.

When the model is fitted to real PR intervals from the data, the performance is much better, as seen with the PR interval prediction. However, the P waves are ill-defined for AF. The fibrillatory waves result in the absence of a clear peak. Still, the information here is usable. One possibility could be to include the shape of the P waves. This way, all information regarding the P wave would directly be imported from the ECG data to the AV node model. With this approach, detection of the peaks would not be necessary, the pulse width and pulse height parameters would not be needed anymore and abnormal behaviour of the atria would be directly incorporated into the model. The drawback would be the loss of synergy with the SA node model, because the computation of HRV metrics requires high resolution detection of peaks. This is not possible for the ill-defined P waves.

8

Conclusions

The newly introduced AV node model extension is able to estimate PR intervals with more accuracy than any other methods we have tried. The model outperformed both the histogram method, which generates PR intervals based on the histogram of past PR intervals. It also outperformed an ANN, which generated PR intervals based on past PP intervals. A better performance compared to the histogram method shows that the PP intervals hold information. The increase in performance compared to the ANN shows a gain in data efficiency. The structure of the AV node model already contains information. With the limited amount of data, the physiological model outperformed an ANN in the prediction of PR intervals. This result shows the validity of the model structure for the AV node during NSR.

It is possible to discriminate between NSR and AF with the parameters of the physiological model in most cases. However, discriminating between AF and other arrhythmias remains difficult. The classification accuracy is far below the best performing algorithms at the time of writing. When the two IPFM models are cascaded and fitted on the RR tachogram, the model parameters of the SA node model are discriminatory, but those of the AV node model do not provide an increase in performance. When the parameter sets are computed separately, there is not enough flexibility in this configuration. The AV node model will have to fit the error of the SA node model, instead of fitting the PR interval. On the other hand, when all parameters are found at the same time, the system is overdetermined, which prevents consistent parameter estimation. We can conclude that the RR tachogram alone does not provide sufficient information for an accurate use of both an SA model and an AV model.

Of the SA model, several parameters are significantly different for NSR and AF. The threshold value directly relates to the average HR, which is often higher in the case of AF. The k parameter describes the wander of the average heart rate. This effect is not present in the same way during AF, so the values are different. Lastly, the noise that is added for a good fit is different for AF. The IPFM model is intended for the SA node during NSR, so a small amount of noise is needed. For AF, the amount of noise is generally much higher. The modulating frequencies are often not significant.

The AV node model shows realistic behavior for NSR when P waves from the data are used as input. However, when only R peak data is available and the output of the SA node model is used as P wave input, the evolutionary algoritm is not able to find a consistent solution. The AV node parameters seem different when we use P wave data for both NSR and AF right after an episode, but more data is needed to confirm this. Future research should include information about the P wave into the parameter estimation process.

A

Abbreviations

AF	Atrial Fibrillation
AFL	Atrial Flutter
ANN	Artificial Neural Network
AP	Action Potential
AR	Autoregressive
AV	Atrioventricular
CIC	Computing in Cardiology
EA	Evolutionary Algorithm
ECG	Electrocardiography
FFT	Fast Fourier Transform
FRPN	Functional Refractory Period of atrioventricular Node
ERPN	Effective Refractory Period of atrioventricular Node
HB	Heart Block
HF	High Frequency
HRV	Heart Rate Variability
IPFM	Integrated Pulse Frequency Modulation
LF	Low Frequency
NSR	Normal Sinus Rhythm
PPG	Photoplethysmography
PSD	Power Spectral Density
RF	Random Forest
RMSSD	Root Mean Square of Successive Differences
RSA	Respiratory Sinus Arrhythmia
SA	Sinoatrial
SD	Standard Deviation
SVM	Support Vector Machine
ULF	Ultra Low Frequency
VF	Ventricular Fibrillation
VLF	Very Low Frequency
WGN	White Gaussian Noise

.

Bibliography

[1] L. Friberg and L. Bergfeldt. Atrial fibrillation prevalence revisited. *Journal of Internal Medicine*, 274(5):461–468, 2013.

[2] Ali Rizwan, Ahmed Zoha, Ismail Ben Mabrouk, Hani M. Sabbour, Ameena Saad Al-Sumaiti, Akram Alomainy, Muhammad Ali Imran, and Qammer H. Abbasi. A Review on the State of the Art in Atrial Fibrillation Detection Enabled by Machine Learning. *IEEE Reviews in Biomedical Engineering*, 14:219–239, 2021.

[3] C.C. Li and R.W. Jones. Integral pulse frequency modulated control systems. *IFAC Proceedings Volumes*, 1(2):186–195, 1963.

[4] Rob Oberman and Abhishek Bhardwaj. Physiology, cardiac. *StatPearls [Internet]*, 2021.

[5] J Gordon Betts, P Desaix, E Johnson, JE Johnson, O Korol, D Kruse, B Poe, J Wise, MD Womble, and KA Young. Openstax college & rice university. *Anatomy & physiology*, 2013.

[6] Kieran E. Brack, John H. Coote, and G. André Ng. Interaction between direct sympathetic and vagus nerve stimulation on heart rate in the isolated rabbit heart. *Experimental Physiology*, 89(1):128–139, 2004.

[7] Pan Li, Glenn T. Lines, Mary M. Maleckar, and Aslak Tveito. Mathematical models of cardiac pacemaking function. *Frontiers in Physics*, 1(October):1–13, 2013.

[8] William J Heinze, Barbara D Brown, and John M Davis. How does adrenaline accelerate the heart? *Nature*, 280(July):235–236, 1979.

[9] Fernand A. Roberge, Réginald A. Nadeau, and Thomas N. James. The nature of the PR interval. *Cardiovascular Research*, 2(1):19–30, 1968.

[10] Jacques Billette and Rafik Tadros. An integrated overview of AV node physiology. *PACE - Pacing and Clinical Electrophysiology*, (December 2018):805–820, 2019.

[11] Deeptankar DeMazumder and Gordon F. Tomaselli. Molecular and cellular mechanisms of cardiac arrhythmias. *Muscle*, 1:583–599, 2012.

[12] Yasushige Shingu, Suguru Kubota, Satoru Wakasa, Tomonori Ooka, Tsuyoshi Tachibana, and Yoshiro Matsui. Postoperative atrial fibrillation: Mechanism, prevention, and future perspective. *Surgery Today*, 42(9):819–824, 2012.

[13] Maurits A. Allessie, Natasja M.S. De Groot, Richard P.M. Houben, Ulrich Schotten, Eric Boersma, Joep L. Smeets, and Harry J. Crijns. Electropathological substrate of long-standing persistent atrial fibrillation in patients with structural heart disease longitudinal dissociation. *Circulation: Arrhythmia and Electrophysiology*, 3(6):606–615, 2010.

[14] Eric N. Prystowsky, Benzy J. Padanilam, and Richard I. Fogel. Treatment of atrial fibrillation. *JAMA - Journal of the American Medical Association*, 314(3):278–288, 2015.

[15] Key Words. V ENTRICULAR F IBRILLATION : Mechanisms of Initiation and Maintenance. *New York*, pages 25–50, 2000.

[16] Guidelines, The North American, and Guidelines. Guidelines Heart rate variability. *European Heart Journal*, 17:354–381, 1996.

[17] Fred Shaffer and J. P. Ginsberg. An Overview of Heart Rate Variability Metrics and Norms. *Frontiers in Public Health*, 5(September):1–17, 2017.

[18] Juan Pablo Martínez, Rute Almeida, Salvador Olmos, Ana Paula Rocha, and Pablo Laguna. A Wavelet-Based ECG Delineator Evaluation on Standard Databases. *IEEE Transactions on Biomedical Engineering*, 51(4):570–581, 2004.

[19] Michael Brennan, Marimuthu Palaniswami, and Peter Kamen. Do existing measures of poincare plot geometry reflect nonlinear features of heart rate variability? *IEEE transactions on Biomedical Engineering*, 48(11):1342–1347, 2001.

[20] Thorsten Schaffer, Bernhard Hensel, Christian Weigand, Jürgen Schüttler, and Christian Jeleazcov. Evaluation of techniques for estimating the power spectral density of RR-intervals under paced respiration conditions. *Journal of Clinical Monitoring and Computing*, 28(5):481–486, 2014.

[21] J. Mateo and P. Laguna. Improved heart rate variability signal analysis from the beat occurrence times according to the IPFM model. *IEEE Transactions on Biomedical Engineering*, 47(8):985–996, 2000.

[22] Raquel Bailón, Ghailen Laouini, César Grao, Michele Orini, Pablo Laguna, and Olivier Meste. The integral pulse frequency modulation model with time-varying threshold: Application to heart rate variability analysis during exercise stress testing. *IEEE Transactions on Biomedical Engineering*, 58(3 PART 1):642–652, 2011.

[23] Robert M. Heethaar, Ruud M. De Vos Burchart, Jan J. Denier Van Der Gon, and Frits L. Meijler. A mathematical model of A-V conduction in the rat heart. II. Quantification of concealed conduction. *Cardiovascular Research*, 7(4):542–556, 1973.

[24] Chang Wook Ahn. *Advances in evolutionary algorithms*. Springer, 2006.

[25] AL Goldberger, LAN Amaral, L Glass, JM Hausdorff, P Ch Ivanov, RG Mark, JE Mietus, GB Moody, CK Peng, and HE Stanley. Components of a new research resource for complex physiologic signals. *PhysioBank, PhysioToolkit, and Physionet*, 2000.

[26] Tatiana S Lugovaya. Biometric human identification based on ecg. *PhysioNet*, 2005.

[27] Lucie Maršánová, Radovan Smíšek, Andrea Němcová, Lukáš Smital, and Martin Vítek. Brno university of technology ecg signal database with annotations of p wave (but pdb). 2021.

[28] Leon Glass. Introduction to controversial topics in nonlinear science: Is the normal heart rate chaotic? *Chaos*, 19(2), 2009.

[29] Daniel Toker, Friedrich T. Sommer, and Mark D'Esposito. A simple method for detecting chaos in nature. *Communications Biology*, 3(1):1–13, 2020.

[30] N. Selvaraj, A. Jaryal, J. Santhosh, K. K. Deepak, and S. Anand. Assessment of heart rate variability derived from finger-tip photoplethysmography as compared to electrocardiography. *Journal of Medical Engineering and Technology*, 32(6):479–484, 2008.

[31] Shreyasi Datta, Chetanya Puri, Ayan Mukherjee, Rohan Banerjee, Anirban Dutta Choudhury, Rituraj Singh, Arijit Ukil, Soma Bandyopadhyay, Arpan Pal, and Sundeep Khandelwal. Identifying normal, af and other abnormal ecg rhythms using a cascaded binary classifier. In *2017 Computing in cardiology (cinc)*, pages 1–4. IEEE, 2017.