



TECHNISCHE
UNIVERSITÄT
WIEN

Vienna University of Technology

DIPLOMARBEIT

MODELS AND ALGORITHMS FOR MINIMIZING THE EFFECTS OF ECTOPIC BEATS ON HEART RATE VARIABILITY DETERMINATION

Ausgeführt am Institut für
Analysis und Scientific Computing
der Technischen Universität Wien

unter der Anleitung von

Ao.Univ.Prof. Dipl.-Ing. Dr.techn. Felix Breitenecker

durch

MARTIN FRANK

HAINBURGERSTRASSE, 36/17

1030

Ort, Datum

Unterschrift

Erklärung zur Verfassung der Arbeit

Martin Frank
Hainburgerstraße 36/17, 1030 Wien

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit - einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

(Ort, Datum)

(Unterschrift Verfasserin)

Acknowledgements

“Life is like a heart rate. If there aren’t some ups and downs, then you aren’t living.”

- unknown

During my diploma thesis I had many of these ups and downs. Luckily the ups compensated the downs every time I finished another part of my diploma thesis. Although some tasks were rather challenging, I am glad that I was able to manage all of them, thanks to many people.

First of all, I want to thank my two advisers at work, Martin Bachler and Christopher Mayer, for the possibility to carry this interesting topic out and for continuous feedback and suggestions during my work. Further I want to thank my supervisor, Felix Breitenecker, not only for supervision, but moreover for the possibility to take part at several social activities of the ARGESIM group, especially the summer school in Dolsko.

I want to thank my family, especially my parents, for their motivation and their financial support during my entire studies. In particular, I want to thank my girlfriend, Lina Bittner, for the special care during the finishing of this thesis.

I am grateful to my professors, who nourished my interests in scientific research, especially in the interdisciplinary field of biomedical engineering.

Last but not least, I want to thank my colleagues, who provided a pleasant working atmosphere. In the same way I want to thank my fellow students and friends for the enjoyable company during my entire studies.

Martin Frank

Abstract

Heart rate variability (HRV) is one of the most important predictors for the assessment of the autonomic nervous system and cardiovascular health. Artifacts in ECG signals, such as ectopic beats, disrupt HRV parameters. Their correction is thus crucial before HRV analysis can be performed. Several correction approaches, such as deletion, interpolation and model based ones are available to deal with ectopic beats. Although all methods are suitable in specific cases, there is currently no comparison available that accounts for a wide field of applications.

This thesis compares several common and new ectopic beat correction algorithms. Different test cases were generated by insertion of artificially generated ectopic beats. These test cases account for slightly ($< 2\%$ single ectopic beats) and heavily corrupted RR interval time series (containing tachycardia, bi- and trigeminy and sequences of couplets and triplets), as well as the determination of robustness against single and successive ectopic beats. The focus of this comparison lies on the recovery of the original HRV parameters, but also compares the root mean squared deviation (RMSD) of the RR intervals, sensitivity, positive predictive value (PPV), mean computation time and peak memory. One additional test set containing natural ectopic beats was used to get information about the performance on real data.

Basically, all correction approaches show a large recovery rate of the HRV parameters. The differences between the methods highly depend on the individual test cases. Slightly corrupted RR interval time series are most reliably corrected by interpolation methods or median filtering. In contrast, highly corrupted RR interval time series are best corrected by deletion of all erroneous RR intervals. This finding is in accordance with literature, since several studies suggested the removal of ectopic segments or even the rejection of the entire record, if error density is high. The reason lies in the impossibility to recover the original sinus rhythm. Models require a larger peak memory and longer computation time, without an increase in HRV recovery rate. Most of them are designed to correct just single ectopic beats with a compensatory pause. In this comparison ectopic beats with and without compensatory pause were used, as well as successive ectopic beats, which explains the weaker performance. However, the goal of this comparison is to capture the performance of the correction approaches in a wide range of ectopic beats and all of the used artifacts are common in nature. Additional tests and evaluations have shown that the performance of most correction approaches indeed depends largely on the type of artifact.

The main finding of this thesis is that removal of ectopic beats is advisable in highly erroneous signals, whereas interpolation may be preferred in slightly erroneous ones. Models do not seem to be advantageous in recovery of the original signal, if different kinds of artifacts are present. Nevertheless, they seem to be superior for specific ectopic beats, such as premature ventricular contractions (PVC). In a similar way, deletion is not always preferable in highly erroneous sig-

nals. As long as correct NN intervals are still present between ectopic beats, interpolation and adaptive median filtering seems to be superior. Hence, deletion is most likely to be only suitable to correct ectopic segments lasting for longer periods, such as a tachycardia. Some authors already suggested to use different correction approaches for specific HRV analysis and various types of errors. Further research has to be performed to verify this assumption.

Kurzfassung

Herzratenvariabilität (HRV) ist einer der wichtigsten Vorhersagewerte für die Beurteilung des autonomen Nervensystems und des Herz-Kreislaufsystems. Artefakte in EKG-Signalen, wie z.B. Extrasystolen, verändern HRV-Parameter und sollten vor der eigentlichen HRV-Analyse korrigiert werden. Dabei können Extrasystolen durch Entfernen, Interpolation oder modellbasierte Methoden korrigiert werden. Obwohl alle Korrekturmethode in speziellen Fällen geeignet sind, gibt es momentan keinen Vergleich, der ein breiteres Einsatzgebiet berücksichtigt.

Diese Arbeit vergleicht einige gängige und neue Korrekturmöglichkeiten für Extrasystolen. Dazu wurden verschiedene Tests generiert, in denen künstlich generierte Extrasystolen eingesetzt wurden. Die Tests berücksichtigen dabei schwach ($< 2\%$ einzelne Extrasystolen) und stark fehlerhafte RR-Intervall Zeitserien (enthalten Tachykardien, Bi- und Trigeminnie, und Sequenzen von Couplets und Triplets), wie auch die Bestimmung der Robustheit gegen einzelne und aufeinander folgende Extrasystolen. Der Fokus dieses Vergleichs liegt in der Wiederherstellbarkeit der originalen HRV-Parameter, aber auch auf dem „Root mean squared deviation (RMSD)“, Sensitivität, positiver Vorhersagewert, mittlere Berechnungszeit und Spitzenspeicherbedarf. Ein zusätzlicher Test, der nur natürliche Extrasystolen enthält, wurde verwendet um Informationen über reale Daten zu bekommen.

Im Wesentlichen zeigen alle Korrekturmethode eine hohe Wiederherstellungsrate der HRV-Parameter. Die Unterschiede zwischen den Methoden hängen dabei hauptsächlich von den individuellen Tests ab. Leicht fehlerhafte Signale sind am besten mit Interpolationsmethoden oder Median-Filtern zu korrigieren. Im Gegensatz dazu sollten stark fehlerhafte Signale am besten durch Entfernen aller inkorrekten RR-Intervalle bereinigt werden. Dieses Resultat ist im Einklang mit der Literatur, da einige Studien das Löschen von ektopischen Segmenten, bzw. sogar des ganzen Signals, vorschlagen, wenn die Fehlerdichte hoch ist. Der Grund dafür dürfte die mangelnde Wiederherstellbarkeit des ursprünglichen Sinusrhythmus sein. Modelle haben einen höheren Spitzenspeicherbedarf und eine längere Berechnungszeit, ohne eine Zunahme bei der HRV-Wiederherstellungsrate zu erreichen. Die meisten Modelle sind für eine Korrektur von einzelnen Extrasystolen mit kompensatorischer Pause vorgesehen. In diesem Vergleich wurden jedoch Extrasystolen mit und ohne kompensatorischer Pause, sowie aufeinanderfolgende Extrasystolen verwendet, womit das schlechtere Abschneiden erklärt werden könnte. Allerdings ist das Ziel dieses Vergleichs die Bewertung der Korrekturmethode in einem breiteren Bereich von Extrasystolen zu erfassen. Des Weiteren entsprechen alle künstlich erzeugten Artefakte natürlichen Krankheitsbildern. Zusätzlich durchgeführte Tests und Evaluierungen zeigten, dass die Korrektur der meisten Methoden hauptsächlich von der Art der Extrasystole abhängt.

Das Hauptresultat dieser Arbeit ist, dass das Entfernen von Extrasystolen in stark fehlerbehaft-

teten Signalen bevorzugt werden soll, während Interpolation in leicht verunreinigten Signalen besser ist. Modelle bringen keinen Vorteil in der Wiederherstellung des ursprünglichen Signales, sofern verschiedene Arten von Fehlern auftreten. Nichtsdestotrotz könnten sie für spezielle Artefakte, wie z.B. ventrikuläre Extrasystolen, besser geeignet sein. In ähnlicher Weise ist das Entfernen von Extrasystolen nicht zwangsweise die bessere Korrekturmethode in stark fehlerhaften Signalen. Solange genügend korrekte NN Intervalle zwischen Extrasystolen präsent sind, ist Interpolation oder der Einsatz von Median-Filtern vorzuziehen. Daher sollte das Entfernen von Artefakten nur im Fall von länger anhaltenden ektopischen Segmenten, wie z.B. einer Tachykardie, vorgezogen werden. Einige Autoren haben bereits vorgeschlagen, verschiedene Korrekturmöglichkeiten für spezielle HRV-Analysen und Fehler anzuwenden. Weitere Forschungsarbeit ist daher notwendig, um diese Vermutung zu belegen.

Contents

1	Introduction	1
1.1	Scope of Work	1
1.2	Methodical Approach	2
1.3	Structure of Thesis	2
2	Background	3
2.1	Motivation for the Usage of the HRV, its Definition and Complications	3
2.2	Measurement of heart rate variability (HRV)	4
2.3	HRV Analysis	6
2.4	Physiological Correlates of HRV	10
2.5	The Electrical Conduction System of the Heart	11
2.6	Artifacts in HRV	12
3	State of the Art of Inter Beat Interval Processing	15
3.1	Artifact Identification	16
3.2	Methods for HRV Preprocessing	18
3.3	Ectopic Beat Correction	21
3.4	Effects of Editing Methods on HRV Analysis	24
3.5	Suggestions for Artifact Handling in HRV Analysis	25
4	Methods	27
4.1	Ectopic Beat Correction	27
4.2	Test Cases	56
4.3	Data Source	58
4.4	Generation of Artificially Corrupted HRV signals	61
4.5	Evaluation of Test Results	63
5	Results	71
5.1	Test-1A: Correction of Single Ectopic Beats	72
5.2	Test-1B: Correction of Successive Ectopic Beats	87
5.3	Test-2: Robustness	101
5.4	Test-3: Correction of Natural Ectopic Beats	102
6	Discussion	107

6.1	Test-1A: Weakly Corrupted RR Interval Time Series	108
6.2	Test-1B: Strongly Corrupted RR Interval Time Series	109
6.3	Test-2: Robustness	110
6.4	Test-3: Naturally Corrupted RR Interval Time Series	111
6.5	Correction Approach Specific Results	112
6.6	Limitations	130
7	Conclusion	131
	List of Figures	133
	List of Tables	136
A	Acronyms	139
B	Detailed Results of Absolute Deviations of HRV Parameters	143
B.1	Specific Results of Test-1A, Single Ectopic Beats	143
B.2	Specific Results of Test-1B, Multiple Ectopic Beats	149
B.3	Specific Results of Test-2, Robustness	155
	Bibliography	157

Introduction

According to the newest report of the “American Heart Association”, heart diseases are still the leading cause of death in the United States [70]. Further, the “American Heart Association” mentions that reliable markers have to be established, as early action is the key of surviving heart failures.

In the last three decades it was recognized that there is a significant relationship between the autonomic nervous system and cardiovascular mortality [61]. One of the most promising markers for this relationship is the heart rate variability (HRV). For example, HRV is able to display the risk of mortality following an acute myocardial infarction (AMI), detect autonomic neuropathy in diabetic patients and can even act as a marker for sudden cardiac death [61]. However, severe artifacts, especially ectopic beats, disrupt HRV parameters. Sethuraman et al. observed that even one single ectopic beat can change the HRV immensely [74]. This results in hardly interpretable measures that are no longer suitable to predict cardiovascular health. Therefore, it is recommended to remove ectopic beats before performing HRV analysis.

1.1 Scope of Work

Although an immense number of different ectopic beat correction methods exists, there is no common agreement about the application of these approaches. Some reviews summarize some of the available methods [61, 42] but they do not offer an explicit comparison on hand that highlights the strengths and weaknesses of the different approaches. Scope of this thesis is to implement existing methods and compare them in different test cases to evaluate their performance on several parameters such as statistical parameters, reduction in the RMSD of the beat-to-beat intervals, computation time, and peak memory. Additionally, it is necessary to optimize all approaches regarding their computational complexity and computation time, since they must be able to run on an embedded system.

1.2 Methodical Approach

In the first step, a literature research was carried out to get an overview about the different methods for ectopic beat correction. Then, these already existing approaches were implemented and compared with respect to their ability to recover the original HRV parameters (described in the “Guidelines - heart rate variability” [61]) of artificial corrupted signals. Therefore, HRV signals free of ectopic beats and 5 min in duration were corrupted with an in-house-developed ectopic beat generator. This offers the advantage of not only knowing the erroneous beat occurrence times but also the correct HRV parameters of the original signals. Last but not least, a reduction of the parameter space was carried out with principal component analysis (PCA) and validation was performed with statistical methods such as box-plots and correlation coefficients. All of the computation was carried out in MATLAB[®], “The MathWorks, Inc”.

The available electrocardiography (ECG) recordings were obtained from the “PhysioNet” database (see section 4.3). This source offers the advantage of providing already annotated ECG signals. Accordingly, the focus of this thesis is solely on the ectopic beat correction, since the detection is already included in the data sets and detection algorithms already exist [7, 8].

1.3 Structure of Thesis

First, chapter “2 Background” gives a brief overview of the definition of *heart rate variability*, its measurement, the most common HRV parameters, and physiological correlations. Further, the basic principles of the cardiac conduction system, its excitation, and possible disturbances are described. The next section summarizes the most common arrhythmias, with a special focus on ectopic beats and other artifacts, which are relevant in HRV signals. Chapter “3 State of the Art of Inter Beat Interval Processing” summarizes available editing methods of RR interval time series, like artifact identification and pre-processing, with a special focus on ectopic beat correction. Most of these correction approaches were compared in several test cases and are described in chapter “4 Methods”. Moreover, this chapter describes the data source, the design of the test cases, and the used statistical methods for the evaluation of the tests. Chapter “5 Results” presents the outcomes of the tests, divided into the three test cases, signals containing single ectopic beats (Test-1A), multiple ectopic beats (Test-1B), robustness (Test-2) and natural ectopic beats (Test-3). Since the tests resulted in a huge amount of diagrams, only the most important and informative ones are shown. The specific behaviour of each correction class is discussed in chapter “6 Discussion”. Additionally, each method is discussed separately with respect to its strengths and weaknesses. Finally, chapter “7 Conclusion” provides a suggestion for the usage of the best correction approaches in specific types of artifacts.

Background

HRV is one of the most promising measures to investigate changes in the autonomic control of heart rate and further, an early predictor of cardiac failures [61]. HRV is not just a single measure, but is rather reflected by several different parameters. These are hard to interpret and have to be chosen carefully, depending on the desired study. Therefore, the “European Society of Cardiology” and the “North American Society of Pacing and Electrophysiology” constituted a Task Force to formulate standards of HRV. In more detail, they summarized the basic measures for HRV, their interpretation and the physiological correlations in “Guidelines - Heart Rate Variability” [61]. Unless otherwise voted, this chapter is based on this report, as it is an established and discordantly accepted standard when using HRV.

2.1 Motivation for the Usage of the HRV, its Definition and Complications

In the past, several studies indicated that there is a significant relationship between the autonomic nervous system and cardiovascular mortality [44, 10, 72]. Therefore, quantitative markers to assess the autonomic activity of the heart emerged. It was recognized that oscillations in the beat-to-beat intervals are well suited to reflect changes in the cardiac autonomic balance. Since these variations are also recognizable in the heart rate, *heart rate variability* (HRV) has become the conventionally accepted term. HRV can be determined by two different mathematical relationships: Basically, HRV is the time series of the beat-to-beat intervals, but these are further the reciprocals of the heart rate. Since the beat-to-beat intervals are mostly determined by the intervals between the R peaks in electrocardiography (ECG) signals, they are often named RR intervals.

The “Guidelines - heart rate variability” state that short-term recordings are more sensitive to artifacts compared to long-term recordings [61]. The more RR intervals are in the signal, the higher is the chance of preserving the original beat-to-beat variability, despite of the presence of false intervals. For example, a 24 h RR interval time series contains approximately 90000

RR intervals whereas a 5 min signal contains only about 300 intervals [66]. Thus, even a small amount of edited RR intervals may alter the original HRV indexes in short-term signals.

Salo et al. proved that in 24 h recordings even 50 % of the RR intervals can be edited without resulting in any major errors (< 5 %) in the HRV analysis [71]. In contrast, they mentioned that editing of less than 5 % of RR intervals in short-term recordings leads to errors in not only the spectrum analysis but also in the statistical parameters, such as the $pNN50$ and the $RMSSD$ (for a description of the statistical HRV parameters see table 2.1).

Sethuraman et al. observed that even one single ectopic beat can change several HRV parameters immensely [74]. They concluded that erroneous RR intervals introduce false frequency components into the power spectrum, but also time-domain measures, such as the $pNN50$, seem to be strongly affected by present ectopic beats. Therefore, it is always advisable to correct these artifacts before the evaluation of the HRV parameters is performed.

Although the most obvious approach is the deletion of erroneous intervals, it is often more advisable to replace the considered interval by means of interpolation or filtering [71]. Dependent on the occurrence frequency and the length of the ectopic segment different correction approaches can be selected. For example, single ectopic beats or short arrhythmic events can be rejected or replaced by different editing methods. However, if the occurrence frequency is high or if multiple successive ectopic beats occur, it is better to reject the whole erroneous segment instead of editing it [65].

Automatic RR interval editing methods are convenient to use, especially in the case of low beat-to-beat variability and distinct ectopic beats (like in healthy persons). On the contrary, many experts state that manual editing with visual verification can never be fully replaced by automatic methods [65]. Unfortunately, ECG annotations from experts are expensive and rather time consuming. Automatic editing methods must rely on high-quality R-peak detection and beat classification in order to perform a reliable HRV analysis [34]. New algorithms in ECG signal characterization seem to be promising tools for the automatic beat annotation. Another possibility to overcome this problem is to perform annotation by experts but to correct RR intervals automatically. No matter what approach is used, more comparative studies are needed to define standard recommendations for suitable editing of RR interval time series [65].

2.2 Measurement of HRV

Basically, *heart rate variability* can be obtained by a continued measurement of the heart rate. Attention has to be paid on the fact that heart rate is a discrete process, as there is only a detectable signal if a heart beat occurs. As a result, heart rate can be monitored continuously, while it is actually a time discrete signal.

Several approaches exist to measure heart rate, but all of them can be assigned to the two following phenomena: To the electrical activity of the heart, which results in voltage changes that can be detected in the whole body and to the pumping activity of the heart, resulting in blood flow and the generation of the pulse wave.

Measurement of the electrical activity of the heart is performed with surface electrodes (see

figure 2.1c) [88]. Thereby, either the whole cardiac cycle can be displayed, e.g. in ECG, or just the depolarization of the ventricles is observed, e.g. in pulse watches. The main difference between these two approaches is the purpose of the measurement. An ECG monitors the heart activity (important in cardiac health assessment), whereas pulse watches only view the current heart rate (for example when doing sport). From the point of view of HRV measurement, both measurements are equal, at least at the first sight, since both use the detection of the depolarization of the ventricles, namely the R-peak in the ECG-signal. Nevertheless, an ECG is recorded more accurately and is therefore always preferred in clinical studies. Attention has to be paid that the actual true marker for the heart rate is the P-wave onset, since it is directly related to the depolarization of the atria [20]. However, R-peaks are more easily detectable and are thus used for heart rate determination [65].

As already mentioned, the mechanical activity of the heart generates a pulse wave at each beat that travels through the arteries and is visible as pressure changes. Further, the blood flow is also maintained, modulated by heart beats and causes changes in the arterial blood volume. Both, volume and pressure changes, can either be monitored invasively with a catheter or non-invasively. Since catheters are only used for long-term measurements and also require a specific request, non-invasive methods are mostly preferred. Pulse wave measurement is assigned by a sphygmomanometer that can be placed at particular sites on the extremities where pulsation in the arteries can be detected, e.g. the arteria brachialis [87]. As shown in figure 2.1a an occlusive cuff is placed around an extremity and filled with air up to a pressure equivalent to the diastolic pressure. As a result, pressure changes inside the artery cause alterations in the pressure inside the cuff and thus can be monitored. In contrast, volume changes inside the arteries are mostly displayed with pulse oximeters. This technique uses light, which is sent into the arteries, gets absorbed, reflected, and scattered, and then is detected by a light sensor (see figure 2.1b) [89]. Primary, the detected light provides information about the blood oxygenation, but it can also display the volume changes inside the arteries due to differences in absorption during blood flow. Basically, two different approaches exist for pulse oximetry: Transmission and reflection. In transmission pulse oximetry only thin body parts can be used, like a finger or the earlobe, whereas in reflection pulse oximetry nearly every superficial artery can be used, since light is sent into the body and is reflected back to the sensor that is located next to the emitter. The obtained pressure and volume curves look quite similar, but are highly dependent on the location where they are measured (illustrated in figure 2.1a). This is related to the reflection waves and attenuation, caused by transmission of the waves through the arteries. However, both curves show a dominant peak that is related to the peak pressure (P-peak) or to the peak volume. As shown in figure 2.2, beat-to-beat alterations are reflected in different lengths of successive RR and PP intervals. Since blood volume and pulse waves can be determined in a similar manner, only the ECG signal and the pulse wave are displayed. Instead of RR or PP-interval, the term normal-to-normal (NN) interval is widely used when referring to normal heart beats (as opposed to ectopic beats). Other synonymous terms for RR interval are beat-to-beat interval and inter beat interval (IBI). IBI will be used as the standard term in this thesis when referring generally to HRV, and RR when ECG signals are used for HRV determination. These intervals form the discrete RR interval time series called the heart rate variability signal or the RR interval

tachogram [65].

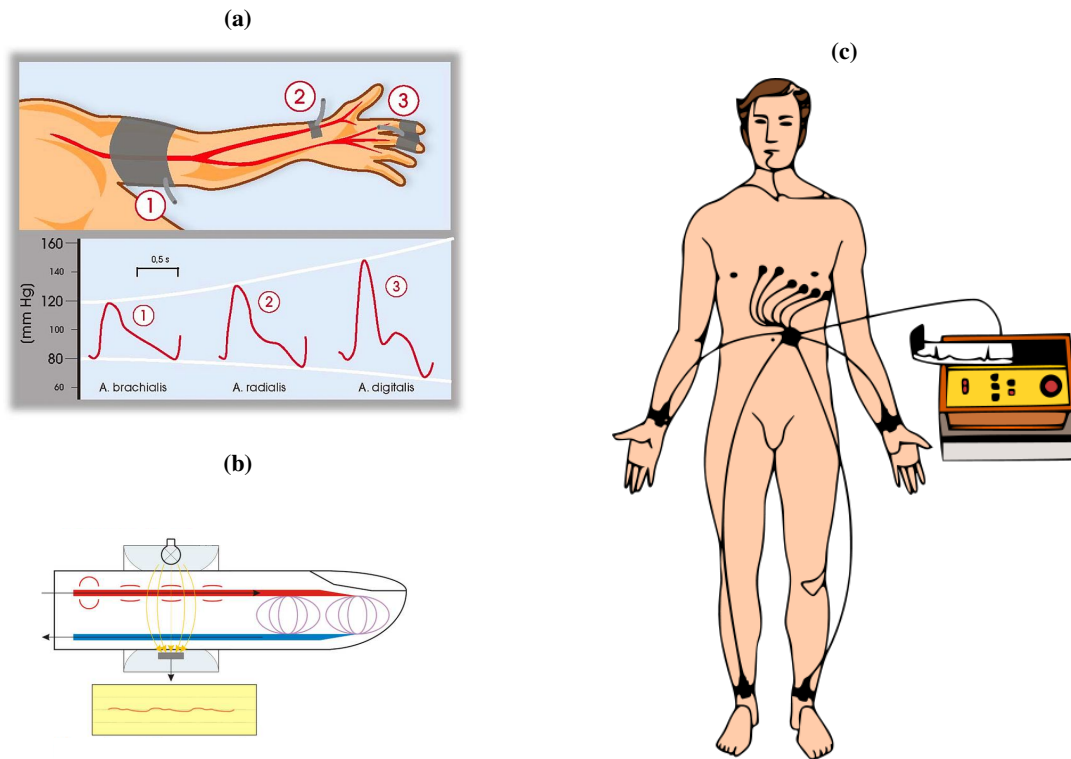


Figure 2.1: Kinds of measurement of heart rate: (a) Continuous measurement of blood pressure at different locations, (b) Measurement of blood volume with pulse oximetry and as an additional possibility, measurement of blood pressure with vascular unloading technique, (c) Electrocardiography, electrodes placed on extremities and directly on the chest next to the heart. Schemes (a) and (b) (modified) are obtained from [87], (c) from [88]

By now, ECG measurement of heart rate can be seen as the gold standard for heart rate determination. However, an ECG measurement requires special equipment and is thus not well-suited for daily life. Instead, pulse measurement is often assessed by pulse watches or electrodes, which are placed directly on the measurement device. This type of measurement is especially very common in online monitoring of the pulse, when doing sports.

2.3 HRV Analysis

HRV analysis can be divided into the following three categories: Time-domain, frequency-domain and non-linear analysis. Non-linear methods were not used in this thesis for HRV analysis and are only mentioned for the sake of completeness.

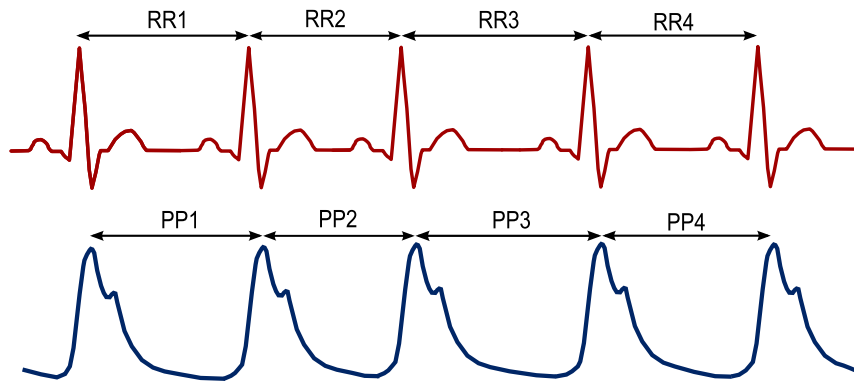


Figure 2.2: Scheme of ECG and Pulse Wave to determine HRV: Red: ECG showing differences in succeeding R-to-R-peak intervals. Blue: Pulse wave showing differences in successive Pulse Pressure Peaks (PP).

2.3.1 Time Domain Methods

Statistical methods are mostly used for the determination of HRV parameters in long-term ECG recordings, i.e. 24 h, although some parameters are also common in short-term recordings. Since only correct normal-to-normal (NN) intervals should be used for the determination of HRV parameters, the term NN will be used instead of RR, which would also encounter erroneous intervals. Table 2.1 gives an overview of the most common time domain measures. They can be divided into two classes: those derived directly from the ECG NN intervals and those derived from the differences between NN intervals. The simplest variable of the first class is the *SDNN*. Since variance is mathematically equivalent to the *total power* of spectral analysis, *SDNN* reflects the modulations of HRV, dependent on the recording length. The comparison of the *SDNN* of different signals only makes sense if they have the same recording length (standards are 5 min for short term and 24 h for long term recordings).

Measures based on the differences between NN intervals are the *RMSSD*, *NN50* and the *pNN50*. According to the “Guidelines - heart rate variability” [61] the *RMSSD* method is preferred to *pNN50* and *NN50* because it has better statistical properties.

The following equations display the calculation of the HRV parameters mentioned above (obtained from Masek [52]).

$$R\bar{R}_m = \frac{1}{N} \sum_{i=1}^N RR_i \quad (2.1)$$

$$SDNN = \sqrt{\frac{1}{N} \sum_{i=1}^N (RR_i - R\bar{R}_m)^2} \quad (2.2)$$

$$RMSSD = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (RR_{i+1} - RR_i)^2} \quad (2.3)$$

$$NN50 = \sum_{i=1}^N (|RR_{i+1} - RR_i| > 50 \text{ ms}) \quad (2.4)$$

$$pNN50 = \frac{NN50}{N} \cdot 100 \quad (2.5)$$

$$SDSD = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (|RR_i - RR_{i+1}| - R\bar{R}_d)^2}, \quad \text{where} \quad (2.6)$$

$$R\bar{R}_d = \frac{1}{N-1} \sum_{i=1}^{N-1} (|RR_i - RR_{i+1}|) \quad (2.7)$$

Geometrical methods are based on the density distribution of the NN intervals. Therefore, only long term recordings should be used. Two common geometrical measures exist: The HRV triangular index and the triangular interpolation of NN interval histogram (TINN). The HRV triangular index is the integral of the density distribution divided by the maximum height. The TINN is the baseline width of the distribution (measured as a base of a triangle, which approximates the NN interval distribution). Thereby, the minimum square difference is used to find such a triangle. For a more thorough description of both variables see figure 2.3.

These two measures of HRV can be calculated as follows:

$$HRV_{index} = \frac{N}{Y}, \quad (2.8)$$

$$TINN = M - N. \quad (2.9)$$

2.3.2 Frequency Domain Methods

In short-term recordings, basically two major spectral components can be distinguished: *low frequency (LF)* and *high frequency (HF)* components, whereas *very low frequency (VLF)* and *ultra low frequency (ULF)* components are additionally present in long-term recordings. The *total power* of the components and their ratio may vary due to changes in autonomic modulations of the heart period. Attention has to be paid on the fact that the lowest parts of the frequency components (*VLF* and *ULF*) can only be evaluated in long term recordings. Normally, the

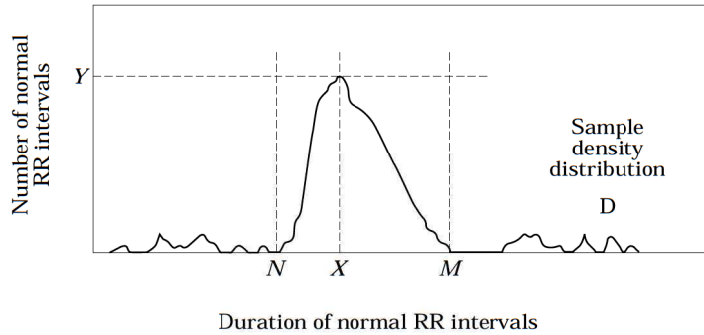


Figure 2.3: Geometrical measures of HRV: First, the histogram of all RR intervals is drawn. Then, the most frequent number of RR intervals (with duration X) are labeled with Y . The HRV index is defined by dividing the total number of all RR intervals by Y . Further, M and N are defined in the histogram as the points where the distribution intersects the x -axis, next to the maximum Y . The TINN is the baseline width of the distribution, restricted by M and N : $TINN = M - N$. Figure obtained from HRV Guidelines [61].

Variable	Unit	Description
SDNN	ms	standard deviation of all NN intervals
RMSSD	ms	root of the mean squared successive differences of NN intervals
SDSD	ms	standard deviation of differences between adjacent NN intervals
NN50 count	#	number of interval differences of successive NN intervals greater than 50 ms
pNN50	%	proportion derived by dividing NN50 by the total number of NN intervals

Table 2.1: Time-Domain measures for 5 min recordings, according to the “Guidelines - heart rate variability” [61].

different components are measured in absolute values of power [ms^2], but LF and HF may also be expressed in terms of normalized units (n.u.), which represent the relative value of each power component in proportion to the *total power* minus the VLF component. This representation can be seen as the reflection of the two branches of the autonomic nervous system, the sympathetic and the parasympathetic system. Table 2.2 provides an overview of these parameters.

Either time-domain or frequency-domain measures should be used for the characterization of HRV, depending on the recording duration. For short-term recordings, far more experience exists on the physiological interpretation of spectral measures whereas they are more or less unsuited for long-term recordings. The reason is a strong correlation of these frequency-domain measures because of mathematical and physiological relationships. In that case, time-domain measures seem to be more promising, as they reflect also long-term influences on HRV, like the day-and-night differences.

Variable	Units	Description
Total Power	ms ²	The variance of NN intervals over the temporal segment
LF norm	n.u.	LF power in normalized units
HF norm	n.u.	HF power in normalized units
LF/HF	-	Ratio LF/HF

Table 2.2: Frequency-Domain measures recommended for 5 min recordings, according to the “Guidelines - heart rate variability” [61].

2.3.3 Non-linear HRV analysis

Beside traditional time and frequency-domain HRV analysis, there are also other methods available that are based on non-linear system theory and beat-to-beat dynamics. For example, Poincaré plots, fractal scaling analysis (e.g. detrended fluctuation analysis) and approximate entropy are used in nonlinear analysis [65]. Further, beat-to-beat dynamics can be evaluated by means of the *heart rate turbulence (HRT)* after ventricular premature beats [75].

2.4 Physiological Correlates of HRV

The heart rate is determined by various pacemaker tissues, especially the sinoatrial (SA) node, which are mostly under the control of the autonomic nervous system (see figure 2.4) [63]. These influences can be divided into the two parts of the autonomic system, the parasympathetic and the sympathetic part.

The parasympathetic part regulates the heart rate via the release of acetylcholine by the vagus nerve [62]. In contrast, the sympathetic part regulates the heart rate by the release of epinephrine and norepinephrine [90]. In rest, vagal tone prevails and so variations in heart rate are mainly dependent on vagal modulation [62]. Additionally, parasympathetic effects impose sympathetic ones by the reduction of norepinephrine, which is released because of sympathetic activity and weakening of adrenergic stimuli [40].

Variations in RR intervals, which are observed under resting conditions, are related to beat-to-beat changes in autonomic activity [3]. Activation of either vagal or sympathetic nerves result in the inhibition of the counter part [40]. Furthermore, both systems are modulated by central (e.g. vasomotor and respiratory centers) and peripheral (e.g. oscillation in arterial pressure and respiratory movements) oscillators [46]. This results in short and long-term modulation of the heart period and can display state and function of the central oscillators, the sympathetic and vagal efferent activity, humoral factors, and the SA node [61].

These modulatory effects of the autonomous nervous system (ANS) can be observed by spectral analysis of HRV. The *HF* component is known to reflect efferent vagal activity [46, 67]. In contrast, the origins and clinical significance of the *LF* component have been discussed controversially. Some authors suggest that it only reflects sympathetic activity [46, 57], whereas others indicate that it is a marker of both, sympathetic and vagal modulations [4]. Consequently, in the

opinion of many researchers the LF/HF ratio seems to reflect the sympatho/vagal balance [61]. A recent review from Goldstein et al. sheds new light on this issue [24]. The authors highlighted many different studies that clearly underline the fact that the LF component of HRV is not a measure of cardiac sympathetic tone but of baroreflex function. However, they also mentioned that modulation of LF and LF/HF may be effected by baroreflexes. On the contrary, physiological correlations of the VLF and the ULF component still have to be evaluated in more detail, although some indications already exist [82, 25, 73].

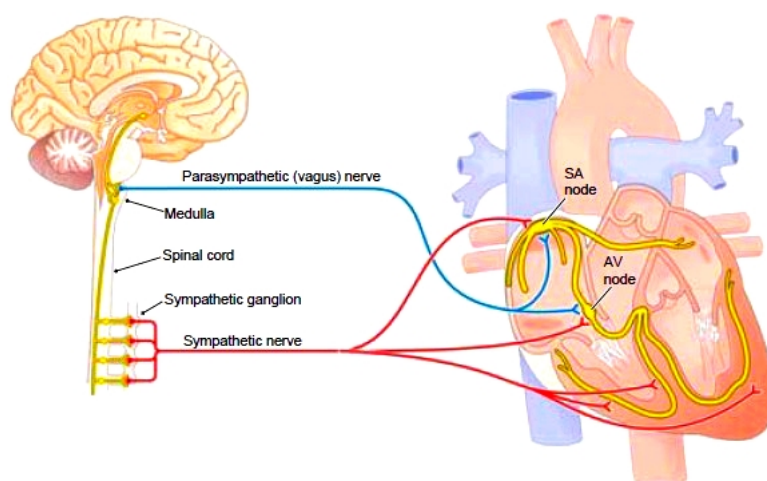


Figure 2.4: Autonomic control of the heart by the sympathetic (in red) and parasympathetic nervous system (in blue) [45]. Both nervous systems act on the SA node and the AV node, whereas only the sympathetic cardiac nerve is able to stimulate the myocardium directly.

2.5 The Electrical Conduction System of the Heart

Unless otherwise mentioned, this section is based on the book “Arrhythmia Recognition: The Art of Interpretation” by Garcia and Miller [21].

At every heart beat, the heart is excited electrically in order to pump blood through the vessels. Since this process has to be performed in a defined way, electrical impulses are transmitted through a specialized conduction system (see figure 2.5). This system has to ensure that the heart muscle cells contract in the right order to pump blood from the atria into the ventricles and then into the body. In a normal sinus rhythm, the pacemaker cells, which form the SA node, spontaneously depolarize and produce an electric impulse. This impulse spreads over the atria and results in contraction of the excited muscle cells. Then, the electric impulse is slowed down at the atrioventricular (AV) septum to guarantee that the ventricles are excited after the atria. Thereby, the AV node is the only electrical communication between the atria and the ventricles. After excitation of this node, the impulse is transmitted to the bundle of His, the purkinje fibres, and finally excites the ventricular myocardium.

All of the mentioned tissues do not only conduct the electrical impulse, which is generated by the sinus node, but rather generate their own electrical impulses. However, the farther these

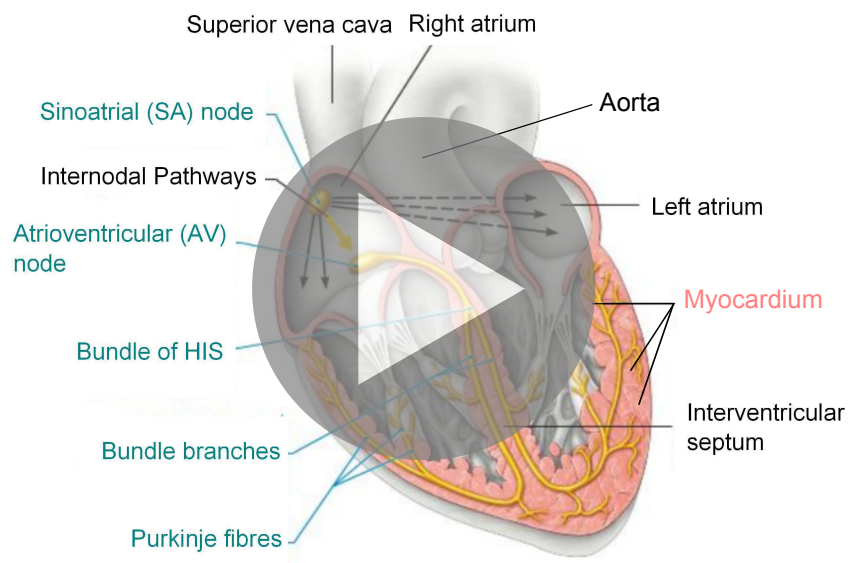


Figure 2.5: Cardiac conduction system: Pacemaker tissues are highlighted in yellow. The SA node automatically generates electrical impulses that are transmitted by the internodal fibres to the AV node, further to the bundle of His, then to the purkinje fibers and finally to the myocardium of the ventricles (pink). Figure modified from [47, p. 677]. Remark: An animation of the electrical excitation of the heart can be started by clicking on the image in the electronic version, obtained from the American Heart Association [6].

tissues are away from the SA node, the longer it takes until they are able to generate their own impulses. As a result, in healthy subjects these cells are forced by the impulses of the SA node to depolarize much earlier than they would do automatically. Therefore, the other pulse generators, like the AV node, the His-bundle and the purkinje fibres only act as a back-up system, if one of the former pacemakers fails to generate an electric impulse.

2.6 Artifacts in HRV

Missed or extra beats can alter the HRV immensely, especially the power spectral density (PSD) [53]. The most common causes of artifacts are arrhythmic events, for example ectopic beats and technical artifacts (caused by poorly fastened electrodes or movement and sweating of the patient) [65]. Since most HRV measurements are performed under defined conditions (like in bed at rest), movement artifacts play a minor role, compared to arrhythmias and ectopic beats. On the contrary, long-term ECG recordings of infants or uncooperative persons are very likely to result in these artifacts [65]. Since the focus of this thesis lies on ectopic beat correction, only this type of artifact was considered.

2.6.1 Types of Arrhythmias

The following chapter describes the definition of arrhythmia, its appearances and its differentiations, with a special focus on ectopic beats. Most of the literature is based on the book “Arrhythmia Recognition: The Art of Interpretation” [21].

According to the “American Heart Association”, the term *arrhythmia* refers to any alteration in the normal sequence of electrical impulses [6]. Arrhythmias can be divided regarding their time of occurrence or their origin. They can happen ahead of schedule (*premature complexes*) or too late (*escape complexes*). Premature Complexes can either occur single, at every second beat (referred as *bigeminy*), third beat (referred as *trigeminy*), fourth beat (referred as *quadrigeminy*), and so on. Common to all these beats is their sequential repetitive occurrence.

Escape complexes can only occur if the primary pacemaker fails, resulting in the take-over of the rhythm of the next successive pacemaker. These complexes can be single or even hold on for a longer period, as long as the primary pacemaker does not succeed to generate an electrical impulse. A special case of arrhythmia is cardiac ectopy. An ectopic beat does not only change the sinus rhythm but also originates from a different focus than the SA node.

Ectopic Beats

Since ectopic beats can even appear in healthy subjects, they are a frequent source of artifacts [61]. For example, *premature ventricular contractions (PVC)* occur in 1 % of clinically normal people and they were detected in 40 - 75 % of apparently healthy persons in long-term recordings (24 - 48 h) [59]. According to Sethuraman et al. even one single ectopic beat can change the HRV indexes immensely [74]. For example, ectopic beats result in an increase of the power spectrum of *HF* bands [65] and the standard deviation of the RR intervals [81]. Therefore, it is recommended to modify ectopic beats before evaluating HRV parameters. Basically, ectopic beats can be divided into two categories (see figure 2.6): Atrial and ventricular ectopic beats.

The most common kind of an atrial ectopic beat is the *premature atrial contraction (PAC)*. As its name states, it is a beat that occurs too early and has its origin in the atria. This kind of beat occurs when a small area inside one of the atria, mostly the left one, is more easily excitable than the rest of the atrial tissue. Since the electrical pathway does not change much, compared to a normal beat, and further the mechanical contraction order is maintained, PACs are mostly clinically insignificant. The only difference in the electrical conduction takes place in the atrium. In PACs, excitation starts from a different focus than the top of the right atrium (location of the SA node). This results in a downward P-wave in the ECG (see figure 2.6).

Another kind of multiple atrial ectopic beats is the *supra-ventricular tachycardia*: It is mostly caused by an SA node tachycardia induced by circulatory repetitive excitations (so called *re-entries*) [88].

A premature ventricular contraction (PVC) is a beat that occurs too early and has its origin in one of the ventricles. PVCs can occur either separately or sequentially. Thereby, two successive

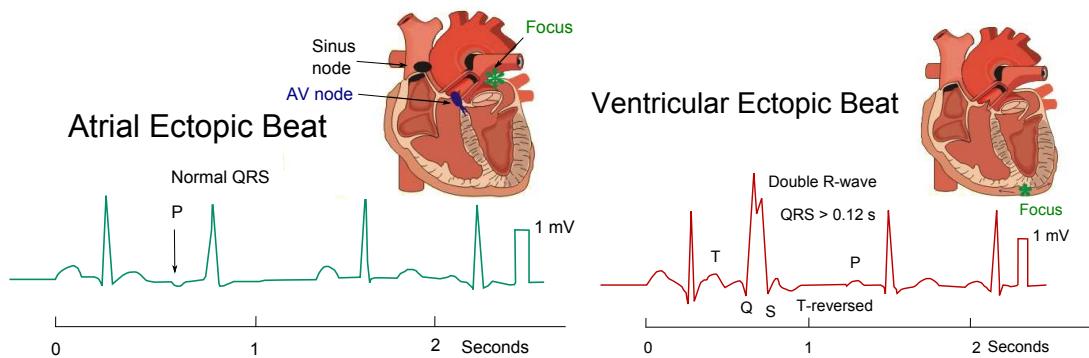


Figure 2.6: Single ectopic beats, dependent on the focus of origin, modified from [63, Chapter 11 - Figure 11-8]. Left: An atrial ectopic beat displays a normal ECG with a time shift. Right: A ventricular ectopic beat results in an abnormal ECG, often with a double R-wave.

PVCs are referred as *couplet*, three as *triplet* and more sequential PVCs are known as *salvo*. Such salvos can further be divided into *sustained* and *non-sustained ventricular tachycardias* [88]. In a ventricular tachycardia an area in one of the ventricles is irritated, resulting in dissociation from the atria [88]. PVCs always result in abnormal QRS complexes, since there is an abnormal excitation of the ventricular myocardium. Further, successive PVCs can have the same (unifocal) or different (multifocal) origins. As there are multiple electrical impulses interfering with each other, they are able to fuse and result in different ECG morphologies.

A typical characteristic of a PVC is the compensatory pause after the ectopic beat. In supra ventricular beats, the sinus rhythm is altered, whereas it stays unchanged in ventricular ectopic beats. Therefore, the impulse that originated from the SA node falls right into the reactionary phase of the myocardium. Only the following sinus beat can excite the ventricles again, resulting in the so called *compensatory pause*. Thereby, the combination of the RR interval before and after the ectopic beat corresponds to two normal RR intervals [88].

State of the Art of Inter Beat Interval Processing

The analysis of HRV is a widely used approach in several kinds of assessment of the circulatory system. Thereby, the used RR interval time series is very vulnerable to impulse noise, such as ectopic beats. In order to remove this noise, a huge amount of different processing and correction methods have emerged. Figure 3.1 illustrates these approaches by means of a mind-map, summarized into classes. These are reflected in sections “3.1 Artifact Identification”, “3.2 Methods for HRV Preprocessing”, and “3.3 Ectopic Beat Correction”. All of these methods have some effects on the different HRV parameters, depending on the carried out editing approach. As a result, several suggestions for the handling of artifacts are available (see section “3.5 Suggestions for Artifact Handling in HRV Analysis”). Jung et. al compared four different HRV analysis software for calculation of time- and frequency-domain HRV parameters and found large differences between those approaches [31]. They concluded that the differences are mainly a result of different editing methods, such as the incorrect deletion of NN intervals. They highlight the importance of standards for HRV processing and request a clear documentation of the used methods for HRV analysis to provide a base for comparative studies [31].

According to a very recent review from Peltola, comparative studies are needed to define standard recommendations for the suitable editing of RR-interval time series [65]. Most of the literature in this chapter is based on this review, since it convincingly highlights the role of *inter beat interval (IBI)* processing. As already mentioned in chapter 2, the terms *IBI* and *RR interval* are synonyms, but as most studies perform R-peak detection for heart rate determination, *RR interval* will be preferred.

The categories of RR interval processing are not distinct but rather contain the same methods in some cases. Some approaches perform detection and correction of artifacts in one algorithm, but most can be separated in either artifact detection or correction. Preprocessing can be divided into “Detection of Missed and Extra Beats” and “Trend Removal” and refers to every process that is performed after artifact identification but before the actual correction. Trend removal is an obvious separate process as it alters the whole RR time series instead of single RR intervals.

Processing of missed and extra beats is referred as a preprocessor, since it performs artifact identification and correction in one single step.

3.1 Artifact Identification

Ectopic beat detection can basically be seen as beat classification. In more detail, the classification can either be performed directly on the detected beats (R-peaks) or on the RR interval time series. Most IBI editing methods do not only classify beats but also reject detected ectopic beats in the same instance. Therefore, it is hardly possible to separate detection and correction completely, when comparing different studies. All detection algorithms just remove ectopic beats and thus can be summarized to a single correction approach, namely deletion.

3.1.1 Beat Classification

The following four approaches perform beat classification directly on the ECG signal (see “Beat classification” in the min-map in figure 3.1 in bright red).

Neuronal Network

One very recent approach was described by Mateo et al. [54]. They presented a method for ECG ectopic beat cancellation based on a radial basis function neuronal network. This method is able to distinguish six different types of beats: Normal beat, PVC, left and right bundle branch block, paced beat and ectopic beat. Although their focus was on *atrial fibrillation (AF)* recordings (with the corresponding beat types), they were also able to distinguish between normal beats and PVCs and thus mentioned that normal ECG recordings can also be used for beat classification.

ECG Morphology

Another example of ectopic beat elimination in the ECG signal was provided by Acar et al. [2]. Morphological changes in the QRS-complex and the P-wave are detected and compared to templates. Thereby, the matching determines if the beat has its origin in the SA node, the atria or the ventricles.

Impulse Rejection with Template Matching

Liu et al. developed a combination method of *impulse rejection* and *template matching* for beat classification [43]. They mentioned that impulse rejection filters are sensitive to identification of false positive (FP) and false negative (FN) beats and *supra ventricular ectopic beat (sVEB)* but only template matching is able to detect morphological changes induced by a *ventricular ectopic beat (VEB)* or noisy ECG recordings. They concluded that a combination of both methods is better suitable for identification of anomalous RR intervals.

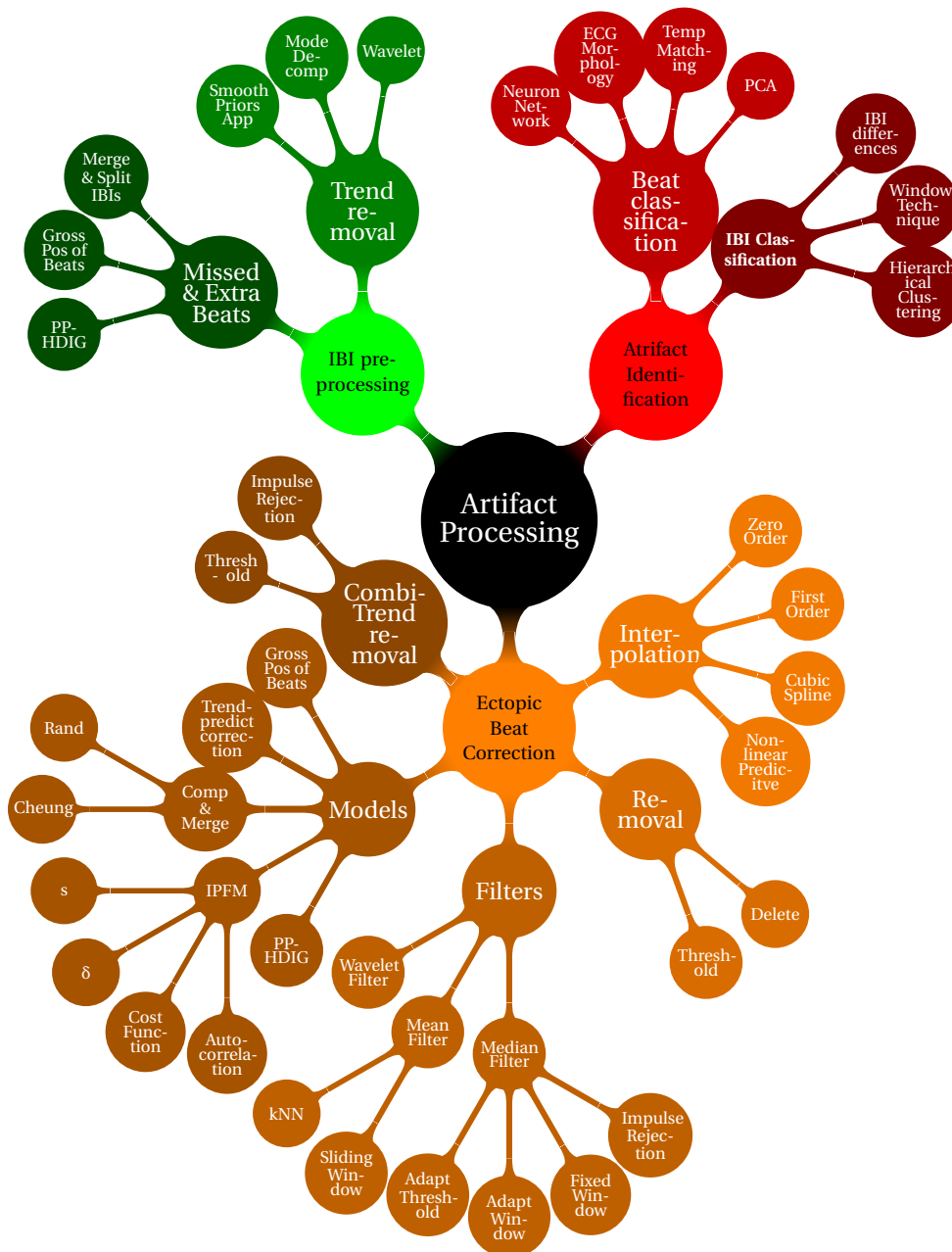


Figure 3.1: Mind-map of artifact processing methods: All approaches can be assigned to one of the following categories: “Artifact Identification” (red part), “Methods for HRV Preprocessing” (green part), and “Ectopic Beat Correction” (orange to ochre part). All ectopic beat correction methods, except the combination with trend removal, and all processes for missed and extra beats (in dark green) were implemented in this thesis. The reason is that all algorithms for missed and extra beats are also able to correct ectopic beats.

Principal Component Analysis

Martinez et al. highlighted the possibility of using *principal component analysis* for the detection of ectopic beats in AF recordings in several studies [50, 51, 48, 49]. They stated that using PCA offers the big advantage of a very low computational cost, compared to other techniques. In their latest paper they mentioned that their PCA-based template matching method can further be used in any other type of ECG recordings, not just in AF ones [49].

3.1.2 Inter Beat Interval Classification

Inter Beat Interval Classification Based on Beat-to-Beat Changes

As already mentioned, the RR interval time series can also be used directly to detect artifacts. Berntson et al. found out that beat-to-beat changes are relatively large in artifacts and may be therefore identified by the following two criteria: The *maximum expected difference* (for NN intervals) and the *minimal artifact difference* [13]. In contrast to the former mentioned approaches, this algorithm only performs detection of erroneous RR intervals, without any downstream correction.

Windowing Technique

Begum et al. described an IBI classification approach that is based on the standard deviation in relation to the mean of a window [11]. The RR interval time series is split into windows and the window containing the largest number of data points is used to determine an average RR interval. All RR intervals that deviate more than a specific threshold (e.g. two or three times the standard deviation of the actual window) from this average interval are classified as erroneous.

Hierarchical Clustering

The windowing technique was compared with another classification approach, *hierarchical clustering*, by Begum et al. [11]. They used built-in MATLAB[®] functions to determine the dendrogram in a bottom-up approach, meaning that in the beginning every RR interval forms a individual cluster. Then, the pairwise distance between observations is calculated, based on the euclidean distance. The hierarchical cluster tree is built up by using the linkage function with a weighted parameter (weighted average distance). At last, the cluster number is set to three in order to form the following clusters: The first cluster contains too short RR intervals, the median cluster the correct NN intervals and and the third one contains too long RR intervals. Although hierarchical clustering performed better than the windowing technique it does have a major drawback [11]: Clusters are always built, even if there is no artifact present.

3.2 Methods for HRV Preprocessing

Some approaches do not detect or correct ectopic beats, but rather act as kind of a pre-processor. These methods are well-suited to detect extra beats (corresponds to FP), missed beats (corresponds to FN) or movement artifacts. Especially in long-term signals trends can be detected

and removed. Since these artifacts are not necessarily related to ectopic beats, these correction methods are mentioned separately in this section. Figure 3.1, “IBI pre-processing” (green part) gives an overview about these methods.

3.2.1 Detection of Missed and Extra Beats

Merge and Split Inter Beat Intervals

One approach to remove extra beats or insert missed beats is to *merge* or *split RR intervals*. This method was first described by Cheung in 1981 [17]. His algorithm compares the actual RR interval with its predecessor and its successor. If the deviation is too large, the actual interval is considered as erroneous. However, if the comparison intervals are also erroneous the method is likely to fail.

Buffer with Combination and Split Rules

Rand et al. improved the “merge and split” method by considering more RR intervals before and after the actual interval [69]. They called this approach *buffer*, since many RR intervals are buffered for evaluation. In addition, they used a set of rules to not only combine and split two RR intervals but to use more intervals (up to three) and even split and combine them in different numbers (e.g. combine three RR intervals and split into two). Instead of only acting as a pre-processor, the buffer can also be used for real ectopic beat identification and correction.

Inter Beat Interval Classification with Template

In a very similar approach Widjaja et al. also used combination and splitting rules of successive RR-intervals in order to classify the type of artifact [86]. However, this method relies on a template that is calculated by adding the previous three RR intervals with different weights. Several rules are applied in a defined order to judge if the actual RR interval is erroneous and should be split or merged with the following interval.

Model Based Inter Beat Interval Classification

Two other approaches, *gross positioning of beats* and the *point process with history dependent inverse Gaussian distribution (PPHDIG)* are also well suited to detect FP and FN beats in HRV signals. Instead of using the RR interval time series these methods detect errors in the beat occurrence time series. Basically, the gross positioning method performs beat identification by calculation of an upper threshold of beat occurrence time variation [53]. FPs are simply removed and FNs are corrected by insertion of evenly spaced beats. The PPHDIG approach checks the correctness of every beat by calculating its probability of occurrence. This probability is also calculated for the preceding and following beats in order to judge if the actual beat is a FP or FN. Like the buffer, these mentioned methods are also used to detect and correct ectopic beats and are described in more detail in section 3.3.

All methods, which can be used for ectopic beat correction, are also able to correct FP and FN, since these beats induce a larger RR interval alteration than ectopic beats (for an overview of ectopic beat correction methods see the mind-map in figure 3.1, orange part). Nevertheless, these approaches cannot distinguish between missed, extra, and ectopic beats and are thus not as well-suited as pre-processors.

3.2.2 Trend Removal

Berntson et al. wrote a committee report, used as a reference for this section, where they mentioned the role of trend removal in HRV analysis [12].

Depending on the desired HRV analysis, it is sometimes recommended to perform prior detrending. Time-domain measures as well as spectral analysis require data that is at least stationary. *strict stationary* means that all moments are stable over time, whereas *weak stationary* just requires that the first and second moment (mean and variance) are stable [12]. However, the initial approach, before detrending, is always to guarantee for as stable conditions as possible during the entire recording period. Moreover, the usage of detrending methods always depends on the regarded frequency bands and thus, on the signal length. Trend removal is always associated with data loss and, especially in short-term recordings, it is not necessarily better to remove trends. Berntson et. al stated that the relation of trends from *ultra low frequency (ULF)* and *very low frequency (VLF)* rhythms to complex psychophysical mechanisms remains still unclear [12]. As a result, many researchers apply detrending methods prior to HRV analysis to get rid of these components [38]. The mind-map in figure 3.1, “Trend removal” (green part), illustrates some common detrending approaches.

Smoothness Priors Approach

One widely used trend removal method is based on the *smoothness priors approach* that acts like a time-varying *finite impulse response (FIR)* high pass filter [80]. Tarvainen et al. mentioned that the most dominant effect of detrending can be seen in the *SDNN*, whereas only little alterations are visible in the *RMSSD* and *pNN50* (in time-domain HRV analysis). They further stated that the *VLF* component distorts other components, especially the *LF* part of the power spectrum.

Wavelet Filter

Another approach to remove undesired trends is the application of *wavelet filters* [38, 81]. These filters outperform FIR filters in terms of memory allocation since they do not need to calculate the inverse matrix [38]. Wavelet filters are based on the discrete wavelet transform. The signal is decomposed into several sub-bands that are described by the wavelet coefficients. The coefficients in the lowest frequency band are set to zero and then the inverse discrete wavelet transform is applied to recompose the signal [81].

Empirical Mode Decomposition

Ji et. al tested the *empirical mode decomposition* as another approach to remove trends irrelevant for HRV analysis [29]. Thereby, they especially focused on trends that were caused by exercise or activity. This was important since they wanted to analyze the underlying rhythm without the interference from extra physical activity.

Effects of Detrending

Yoo et al. summarized the effects of detrending on HRV parameters [91]. They figured out that temporal statistical measures, such as the *SDNN*, *RMSSD*, *NN50*, and *pNN50* as well as spectral measures, such as the *LF* and *HF* power, do not show significant differences between raw and detrended HRV signals. In contrast, nonlinear HRV analysis, like detrended fluctuation analysis, did show an improvement when detrending was performed. Thus, they recommended to perform non-linear HRV analysis without detrending, but to remove trends in linear analysis [91].

3.3 Ectopic Beat Correction

According to the review of Peltola, RR interval editing may be divided into the following categories: Deletion, interpolation, and filtering [65]. However, some methods can neither be convincingly assigned to the category “deletion” and nor to “interpolation” or “filtering” but rather rely on more complex editing structures and are thus summarized in an additional category, namely “models” (for an overview see the mind-map in figure 3.1). Actually, a fifth category may be specified, namely combinations, but as these are filters that are just combined with trend removal, which actually is no ectopic beat correction method, this class is not further evaluated.

3.3.1 Deletion

The simplest way of RR interval editing is to delete false intervals by shifting the preceding NN intervals to replace the removed ones [65]. However, if a large amount of segments is deleted, this can lead to an unsustainable loss of information [71]. The decrease in the signal length might cause problems when analyzing the power spectrum or ECG recordings of short duration [65]. Detection of the erroneous RR intervals can be performed prior to the removal process, e.g. with one of the methods described in section 3.1.

3.3.2 Threshold filter

Some methods, like threshold filters, do not only detect ectopic intervals but also reject them in the same process. A very simple way is to apply hard thresholds. RR intervals that lie outside a fixed range, e.g. $RR_i < 0.4\text{ s}$ or $RR_i > 2.0\text{ s}$, are removed to exclude non-physiologic intervals [56]. This approach can also be used as a pre-processor before applying the actual filter [56]. Hard threshold limits can further be used to directly remove artifacts, no matter what their origin is [38]. Some authors suggested to use adaptive thresholds that depend on the surrounding RR intervals [56, 34].

3.3.3 Interpolation

In contrast to deletion, interpolation methods preserve the initial number of RR intervals. Common to all interpolation methods is the replacement of erroneous RR intervals with new interpolated RR intervals. Basically, interpolation methods can be seen as kind of low-pass filters [65]. The most common interpolation methods are: *Interpolation of degree zero*, of *degree one*, and *cubic spline interpolation*. Interpolation of degree zero uses the adjacent RR intervals to calculate a mean interval that replaces the actual IBI [64] and thus can also be seen as kind of a mean filter. Interpolation of degree one uses a straight line to interpolate an ectopic segment and therefore connects the NN interval prior to this segment with the following NN interval [42]. In contrast to linear interpolation, spline interpolation uses a higher order function that is defined piecewise by polynomials [64]. The big advantage over classical linear interpolation is the generation of smooth curves [64]. For example, Albrecht and Cohen recognized that spline interpolation should be used especially when evaluating the power spectrum [5].

3.3.4 Non-linear Predictive Interpolation

In a completely different approach Lippman et al. introduced a *non-linear interpolation* method that is based on chaos theory [41]. However, this method also accounts for the fact that heart rate is a predictive process. Their algorithm builds trajectories in phase space to determine similar structures in the RR interval time series that can even lie far away in time space. This segment is then used to replace the ectopic one.

3.3.5 Filtering

Filters can be seen as the “super-class” of several correction approaches, including interpolation and threshold filtering. Inherently, filtering can either occur in the time- or the frequency-domain. For example, *wavelet filters* are a very common filtering technique in the frequency domain. Keenan performed a comparison of wavelet hard thresholding with linear interpolation and figured out that wavelet filters are more suited for detection of ectopic beats than for correction [33]. He further stated that wavelet filters show the disadvantage of changing neighboring NN intervals and thus disrupt the power spectrum.

Time domain filters can be distinguished regarding the value that is used to replace the incorrect RR interval. Most of these filters use a window to determine either the mean or the median. Thereby, the window length can be fixed or adaptive. Kumaravel and Santhi compared these two approaches and concluded that adaptive windows perform better than fixed ones [36]. They even went one step further and designed an adaptive median filter with an adaptive error identification threshold, based on the last correctly accepted RR interval. In another study, Mc Names et al. suggested an impulse rejection filter to efficiently detect ectopic beats and replace them with the median of a fixed window [55]. All of the above mentioned time domain filters are based on a window that incorporates RR intervals before and after the actual interval. In another approach, Begum et al. determined the best fitting neighboring RR intervals based on the *k Nearest Neighbors (kNN)* method [11]. They replaced the erroneous RR interval with the mean of either the *k* previous or the *k* following RR intervals.

3.3.6 Models

Merge and Split Inter Beat Intervals

In addition to the different interpolation and filtering methods, more sophisticated RR interval editing methods have emerged. For example, Cheung proposed an algorithm that detects and corrects erroneous RR intervals by comparing and merging adjacent intervals [17]. Since this method still had some drawbacks, especially when dealing with successive erroneous intervals, Rand et al. developed an improved technique that considers more adjacent RR intervals [69]. Further, their algorithm does not only combine two adjacent RR intervals but rather perform a set of rules in a predefined order. The algorithm is also capable to combine and split RR intervals in different numbers, e.g combine two intervals and split the resulting interval in three new ones.

Gross Positioning of Beats

Another rudimentary approach is described by Mateo et al. as *gross positioning of beats* [53]. This method is mostly used to remove extra beats or to insert missed beats, but can also be used to correct whole ectopic segments by insertion of evenly spaced beats.

Trend Predict Correction

Wen et. al suggested to determine short-term trends in RR interval time series in order to detect and correct ectopic beats [85]. They called this method *trend predict correction* and calculated the trend based on two features, the actual trend that reflects the correlation of preceding sinus beats and the turbulence between adjacent RR intervals.

Integral Pulse Frequency Modulation Model

In the last decades, some physiological models have evolved that are able to effectively describe the modulation of the heart rate. One of the most prominent models is the *integral pulse frequency modulation (IPFM)* model. This approach integrates the input signal until a beat is generated and is then reset to zero [53]. Basically, this model is able to simulate the ANS activity on the SA node by illustrating the series of cardiac events as firings of the SA node [84]. This model is widely used to simulate heart rate variations [53, 76, 84, 5]. Mateo et al. described an approach where they used the s-parameter, which reflects the phase shift in the integration process by premature resting [53]. In a more computationally efficient method, Solem et al. used the fact that an ectopic beat shifts the occurrence times of the following normal heartbeats by the time δ , and determined several parameters based on this time shift [76].

Another approach that relies on the IPFM model was introduced by Albrecht et al. They used the autocorrelation of RR intervals to correct bad intervals [5]. Additionally, Brennan et al. developed an algorithm which is based on the same model, but uses a cost function to determine the lowest costs for the beat occurrence time [15].

Point Process with History Dependent Inverse Gaussian Distribution

Barbieri et al. recognized that the inverse Gaussian probability density, combined with a history-dependent point process, is a well-suited method to perform heart beat analysis [9]. Based on this model Citi et al. developed an algorithm that is able to detect and correct erroneous heart beats [19].

3.4 Effects of Editing Methods on HRV Analysis

Although a huge amount of different ectopic beat correction methods exist, there is a lack of detailed reviews about their impact on HRV analysis. However, some studies performed an evaluation of the effects of different correction methods on specific HRV parameters [5, 14, 42, 71, 66]. A comparison of these findings shows clearly that the results highly depend on the study setting. Nevertheless, all studies indicate that editing methods do have an effect on HRV analysis. As already mentioned, especially spectrum analysis is sensitive to any loss of RR intervals, signal length and discontinuities in the signal [65]. Therefore, several studies recommended to use replacement methods instead of deletion when performing HRV spectrum analysis [61, 71, 53]. Another drawback of deletion of RR intervals is the introduction of step-like changes in the RR interval time series, depending on the natural trend of the heart rate [65]. Especially when removing successive RR intervals, steep steps may arise that disturb these natural fluctuations. Despite of all these disadvantages, deletion is a feasible method for the analysis of time-domain parameters, such as the *SDNN* and *SDANN*, in 5 min signals [71]. In contrast, Salo et al. indicated in the same study that deletion should not be used as correction method when determining the *pNN50* and the *RMSSD*. Artifacts lasting for longer periods, like fibrillation, should be deleted instead of replaced [32].

In general, HRV analysis is highly dependent on the used methods for RR editing and the type and length of the RR interval time series. Lippman et al. made a comparison of ectopic beat removal with linear, cubic spline and non-linear predictive interpolation in short-term HRV signals [42]. They found out that the removal of ectopic beats and non-linear predictive interpolation result in better performance than linear or cubic spline interpolation, since both overestimated the *LF* power and underestimated the *HF* power. In another study, Albrecht and Cohen concluded that the linear interpolation method results in a more accurate power spectrum compared to the predictive auto-correlation method [5]. These two examples show that there are some comparisons available but most studies only compare their self developed algorithm with two or three already established ones. Unfortunately, they do not use the same study settings or even similar editing methods, making it nearly impossible to compare different HRV studies. Just the basic interpolation methods and deletion are compared in some studies. For example, Salo et al. compared the performance of RR interval editing by deletion with interpolation of degree one and zero [71]. They used short and long-term recordings of healthy subjects and patients with AMI and focused on the analysis of frequency parameters. In their findings both interpolation methods performed well and in some cases were superior to the deletion method. Nevertheless, interpolation always implies the possibility of introducing false shapes and trends

into the HRV signal. For instance, interpolation of degree zero leads to flat shapes when correcting successive RR intervals, since it replaces all edited RR intervals with the same mean. Interpolation of degree one results in slope-like shapes, as it simply links the last correct RR interval with the next correct one by a straight line. The longer the RR segments to be corrected is, the more dominant this phenomenon becomes. The introduction of these false shapes may explain the increase of the power of the *LF* and *VLF* components in the HRV spectrum [71, 14]. Therefore, interpolation methods should be avoided when editing longer RR interval segments. Peltola et al. also examined the influence of interpolation of degree zero, one, cubic spline interpolation, and deletion for detrended fluctuation analysis and concluded that interpolation is more suitable than deletion [66].

RR interval editing is not only related to alterations of RR intervals but also relies on artifact detection. Kemper et al. compared the effects of three different detection and two correction methods (removal and interpolation) for ectopic beat handling. They found out that ectopic beat detection is more prone to errors than correction [34]. In their study, two of three artifact identification methods falsely detected too many errors. As a result, correct NN intervals were either edited or deleted, but both correction methods resulted in nearly the same HRV parameters. Therefore, it can be concluded that the most critical step in HRV analysis is not the selection of an appropriate correction method but rather to ensure that all RR intervals are correctly classified.

3.5 Suggestions for Artifact Handling in HRV Analysis

Peltola clearly stated in her review that there is no consensus about how to best correct artifacts [65]. Further, no standards or recommendations for editing of RR interval time series before HRV analysis are available. There is only one common opinion when performing spectrum analysis: Interpolation seems to be superior to deletion. Moreover, the literature lacks information about the maximum amount of RR intervals that should be edited: Most studies suggest that at least 80 % should be NN intervals, especially when evaluating frequency parameters. However, other studies edit from 1 to 30 % of all RR intervals [65]. Moreover, Salo et al. mentioned that there is no universal RR interval editing method that can be used in time- and frequency-domain HRV analysis [71].

This inconsistency in editing of RR intervals leads to problems when comparing HRV parameters of different studies. Reasons are different numbers of edited intervals, HRV parameters, study population, IBI signal length or the used editing method. Therefore, it is recommended to provide an accurate description of the used editing methods [65, 34].

Methods

This chapter describes the methods used in context of this diploma thesis. Basically, the chapter can be divided into the following sections: Ectopic Beat Correction, Test Cases, Data Source, Generation of Artificially Corrupted HRV signals and Evaluation of Test Results. The first section, “4.1 Ectopic Beat Correction”, describes the different implemented correction methods in detail. A more general overview about artifact handling, i.e. identification and editing, is given in section “3 State of the Art of Inter Beat Interval Processing”.

Section “4.4 Generation of Artificially Corrupted HRV signals” illustrates the different ectopic beats which were considered to corrupt non-erroneous ECG signals. The next section explains the test cases that were used for the comparison of the different ectopic beat correction algorithms. The last section, “4.5 Evaluation of Test Results”, gives a short overview of the used methods for the analysis of the test results.

4.1 Ectopic Beat Correction

Ectopic beats can be handled by a variety of different approaches. Chapter “3 State of the Art of Inter Beat Interval Processing” explains the most prominent methods that have been used so far. Some of these mentioned approaches focus rather on ectopic beat detection than on correction. Further, most of the detection algorithms just reject erroneous RR intervals instead of editing. Since the goal of this thesis is the comparison of editing methods, detection algorithms are not considered. However, “deletion” is also a very common approach for RR interval editing, especially when dealing with few ectopic beats [65]. Ectopic beats are already identified in the used ECG signals as the beats are already annotated. Therefore, even the huge amount of ectopic beat identification algorithms can be summarized to only one correction method, namely “deletion”. All other approaches actually perform different kinds of editing methods, such as interpolation, replacement with mean or median, determination of beat occurrence probabilities and so on. Figure 4.1 illustrates all of these methods in terms of a classification with respect to their editing approach. These classes are: Removal, Interpolation, Filters and Models.

The descriptions of the individual methods in this section are illustrated by exemplary RR interval time series. Since already annotated ECG signals are used, only erroneous RR intervals are edited in most approaches. Nonetheless, some filters also edit more RR intervals, since they perform annotation on their own or are only able to alter two RR intervals instead of one (they shift beats instead of just changing single RR intervals). This behavior is always specified in the description of the distinct approaches. Unless otherwise mentioned, each approach only performs editing on the previously defined RR intervals.

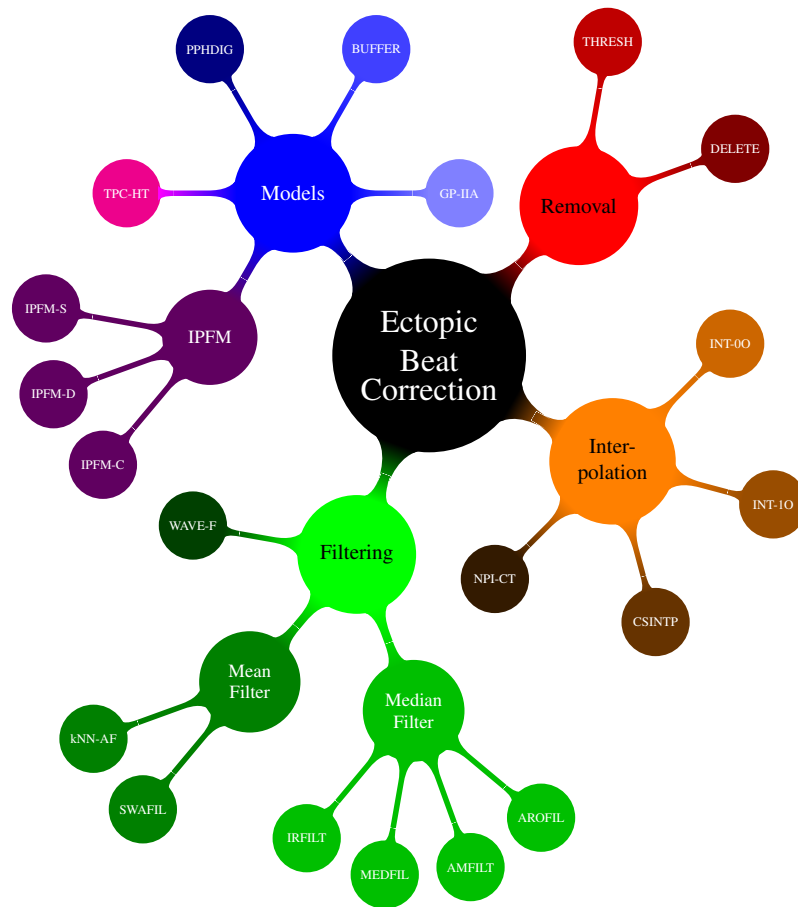


Figure 4.1: Overview of all methods implemented for ectopic beat correction: The methods are classified as removal (red), interpolation (orange), filtering (green), and models (blue).

Table 4.1 displays the acronyms of the implemented correction methods in this thesis. The table serves as a listing, since all methods will be labeled by their acronyms in the text for a consistent comparison of the diagrams in chapter “5 Results”. However, in this chapter, both, acronyms and original descriptions, will be used for a better understanding of the algorithms.

Acronym	Description
THRESH	threshold filter
IRFILT	impulse rejection filter
MEDFIL	median filter with fixed window length
AMFIL	adaptive median filter
AROFIL	adaptive rank order filter
SWAFIL	sliding window average filter
INT-00	interpolation of degree zero
INT-10	interpolation of degree one
CSINTP	cubic spline interpolation
kNN-AF	k nearest neighbors average filter
BUFFER	buffer
WAVE-F	wavelet filter
GP-IIA	gross positioning of beats, method IIA
IPFM-S	IPFM model with s-parameter
IPFM-D	IPFM model with δ -parameter
IPFM-C	IPFM model with cost function
TPC-HT	trend predict correction considering the heart turbulence
DELETE	deletion
NPI-CT	non-linear predictive interpolation based on chaos theory
PPHDIG	point process with history dependent inverse Gaussian distribution

Table 4.1: 6 digit acronyms for all implemented correction methods

4.1.1 Removal

Filters of the class “Removal” simply delete erroneous RR intervals. Thereby, they either delete only those intervals labeled as erroneous, or they perform error detection on their own. As mentioned previously, several approaches are represented by *deletion*, since this approach just deletes marked RR intervals. In contrast, *threshold filter* represents an approach that detects erroneous RR intervals on its own and sequentially deletes them.

Deletion (DELETE)

This approach deletes all RR intervals that are annotated as erroneous. Therefore, artifacts are always excluded. This is an enormous advantage over all other methods, since it is the only approach that really guarantees that only correct sinus beats are used for HRV analysis. On the contrary, information is always lost when deletion is used. This might play a minor role when only a few ectopic beats are present, but may lead to incorrect HRV parameters if too many intervals are removed. Figure 4.2 displays that there is nearly no change of the RR interval time series for a PAC (a) and for the first few PVCs (b), but as their number increases, the time shift between the original and the corrected signal increases enormously. This phenomenon becomes even worse if whole ectopic segments, such as sequential PVC couplets, are deleted

(see figure 4.2c).

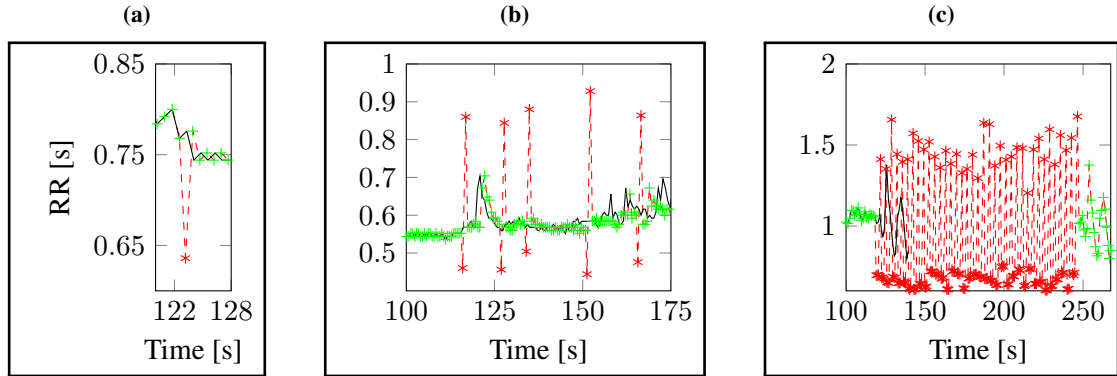


Figure 4.2: Deletion of ectopic beats at annotated points. A single PAC (a) and multiple PVCs (b) are reasonably corrected, as long as they occur not too frequent. In contrast, (c) shows that successive ectopic beats, such as sequential PVC couplets, result in an enormous signal loss. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown on the original signal: TPs are labeled as stars in red and TNs as plus signs in green.

Threshold Filter (THRESH)

Threshold filters can be used in a variety of different configurations. Basically, threshold filters can either exclude RR intervals that lie outside hard limits, e.g. 0.4 and 1.20 s [38], or delete values that deviate more than a relative value to a defined reference point [56, 34], e.g. 20 % from the mean [36]. Normally, this reference point is either the mean or the median of the whole signal or a local statistical value of a window. Since the mean is sensitive to outliers and signal modulations, it is mostly recommended to use the median instead. The filter implemented in this thesis deletes all RR intervals that deviate by more than 20 % from the median of the whole signal. This value is suitable according to the literature [36, 56] and additionally performs better than using the mean or other relative thresholds (tested for 10 % and 30 %) or even the hard limits specified by Lee et al. [38]. Although threshold filters may also be used as a windowing filter, they are mostly used as a fast outlier rejection filter that analyzes the whole signal at once, such as the implemented one in this thesis. Figure 4.3 illustrates the characteristics of this filter. Sharp impulses are always excluded if they lie outside the defined threshold, independent of their origin. This problem occurs mainly in signals that contain many fluctuations. Figure 4.3a shows a nearly non-fluctuating signal where ectopic beats are properly detected and corrected, whereas the signal in figure 4.3b clearly displays a trend that results in inappropriate outlier detection. Therefore, Lee et al. suggested to perform detrending in order to avoid this undesired behavior [38]. However, as the focus of this thesis lies on the comparison of correction methods of annotated RR interval time series without any additional pre-processing, detrending is not performed. Thus, threshold filters of this kind can only effectively remove erroneous RR inter-

vals in weakly fluctuating signals, at least if they do not rely on windowing. Precise rejection of errors is covered by the deletion approach and windowing is performed by every other filter of the filtering and interpolation class.

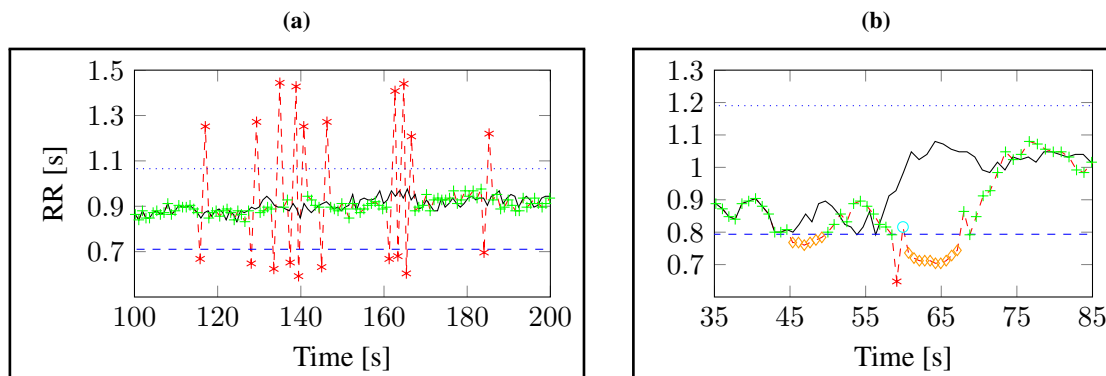


Figure 4.3: Threshold filter: (a) Detection of erroneous RR intervals with subsequent deletion is performed correctly in weakly fluctuating signals. (b) In contrast, signals containing trends cannot be processed reasonably. The original signal is displayed dashed in red, the corrected signal solid in black, the fixed thresholds pointed (upper) and dotted (lower) in blue. Annotation is shown on the original signal: TPs are labeled as stars in red, FPs as diamonds in orange, FNs as circles in cyan and TNs as plus signs in green.

4.1.2 Interpolation

In contrast to the removal approach, interpolation preserves the initial number of RR intervals by substituting erroneous intervals with interpolated values. Thereby, only correct NN intervals are used to calculate the new RR intervals. Figure 4.1, orange part, lists different interpolation methods: *Interpolation of Degree Zero (INT-00)*, *Interpolation of Degree One (INT-10)*, *Cubic Spline Interpolation (CSINTP)* and *Non-linear Predictive Interpolation (NPI-CT)*. The first three approaches have been widely used for a long time in several studies and are described in the PhD thesis from Peltola [64]. These approaches use the RR intervals next to the ectopic segment for interpolation. On the contrary, *Non-linear Predictive Interpolation (NPI-CT)* determines the RR intervals used for interpolation in a phase space. This method was first described by Lippman et al. [41] and is also compared to the former mentioned interpolation methods [42].

Interpolation of Degree Zero (INT-00)

Interpolation of degree zero can be seen as replacement of ectopic beats by a constant value. It uses the n neighboring correct NN intervals of a window, which includes previous and following RR intervals, to replace the erroneous RR intervals by the mean of the window:

$$RR_i = \frac{1}{n} \cdot \sum_{k=i-\frac{n-1}{2}}^{i+\frac{n-1}{2}} NN_k. \quad (4.1)$$

According to Peltola, the window length n should be set to a value from three to ten RR intervals [64] and was set to five RR intervals in this thesis. Interpolation of degree zero offers an easy and quick method for replacing occasional ectopic beats (see figure 4.4a). Large segments, on the other hand, such as tachycardia or sequential couplets and triplets, cannot be reasonably edited since this interpolation method produces flat shapes in the RR time series (see figure 4.4b). All erroneous RR intervals in one segment are replaced by the same value, the mean, resulting in loss of the variability. However, one advantage of this approach over traditional mean filters is that only correct NN intervals are used. Therefore, interpolated RR intervals are never biased by neighboring erroneous intervals.

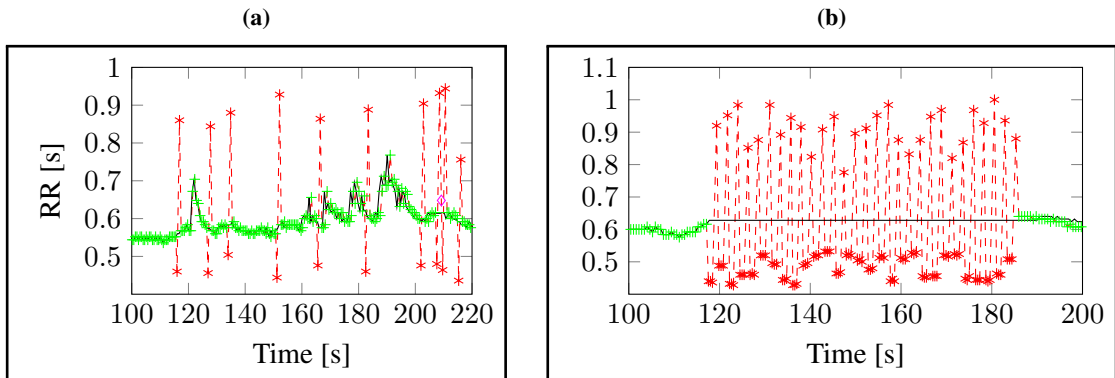


Figure 4.4: Interpolation of degree zero: (a) Single ectopic beats, such as PVCs or bigeminy, can be reasonably edited. (b) In contrast, ectopic segments, such as PVC triplets, are replaced by inserting the mean RR interval of the specified window at every place marked as erroneous, resulting in a flat shape. Since successive erroneous RR intervals are always shorter than correct NN intervals, the signal length increases (black line stretches over the green marked intervals). The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown on the original signal: TPs are labeled as stars in red and TNs as plus signs in green.

Interpolation of Degree One (INT-1O)

Basically, the two adjacent RR intervals of the ectopic segment are linked by a straight line. Then, the corresponding RR interval of the point on this line is inserted at every beat occurrence time in the ectopic segment. Like in interpolation of degree zero, this results in flat shapes when

interpolating over long segments (see figure 4.5b). The following equation states the calculation of the new RR interval RR by fitting a straight line through the points (t_0, NN_0) and (t_1, NN_1) :

$$RR(t) = NN_0 + \frac{NN_1 - NN_0}{t_1 - t_0} (t - t_0). \quad (4.2)$$

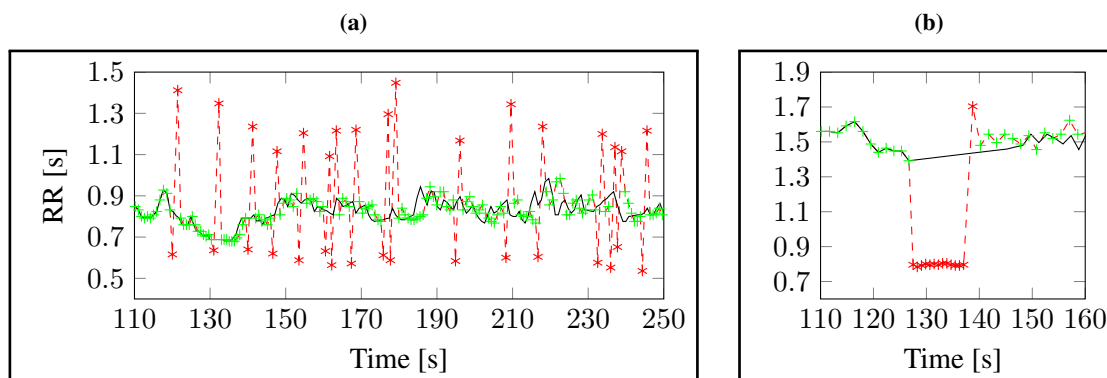


Figure 4.5: Interpolation of degree one: (a) Single ectopic beats, bi- and trigeminy are reasonably corrected. In contrast, successive ectopic beats, such as ventricular tachycardia result in a long straight line between the adjacent NN intervals that do not reflect the original physiological variation (b). Since successive erroneous RR intervals are always shorter than correct NN intervals, the signal length increases (black line stretches over the green marked intervals). The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown on the original signal: TPs are labeled as stars in red and TNs as plus signs in green.

Cubic Spline Interpolation (CSINTP)

This interpolation method uses splines to link correct NN intervals. The probably most frequently used approach is *cubic spline interpolation*, where a third order polynomial is fitted to the data points. The big advantage over linear interpolation is the usage of smooth curves for the interpolation, instead of straight lines. This effect can be best observed in the interpolation over long ectopic segments, as demonstrated in figure 4.6b. In contrast, interpolation over single ectopic beats does not show such huge differences to linear interpolation (see figure 4.6a).

The coefficients of cubic polynomials are calculated to bend the line in such a way that a continuous curve passes every data point. This piece-wise function can be expressed in the following form [68]:

$$S(x) = \begin{cases} S_1(x) & \text{if } x_1 \leq x \leq x_2 \\ S_2(x) & \text{if } x_2 \leq x \leq x_3 \\ \vdots & \\ S_{n-1}(x) & \text{if } x_{n-1} \leq x \leq x_n, \end{cases} \quad (4.3)$$

where $S_i(x)$ is a cubic polynomial that is defined by

$$S_i(x) = a_i + b_i \cdot (x - x_i) + c_i \cdot (x - x_i)^2 + d_i \cdot (x - x_i)^3. \quad (4.4)$$

$S(x)$ represents the spline function and the continuous piece-wise function must fulfill the following requirements:

$$S_i(x_i) = y_i, \quad (4.5)$$

$$S_i(x_{i+1}) = y_{i+1}. \quad (4.6)$$

In order to ensure the continuity of the piece-wise functions at their coincidence points, the first and second derivatives have to be equal, respectively:

$$S'_{i-1}(x_i) = S'_i(x_i), \quad (4.7)$$

$$S''_{i-1}(x_i) = S''_i(x_i). \quad (4.8)$$

Cubic splines are only required to be continuous in the first and second order, resulting in discontinuities in the higher derivatives. If the data is sensitive to smoothing effects that are higher than second order cubic spline interpolation results in inadequate splines.

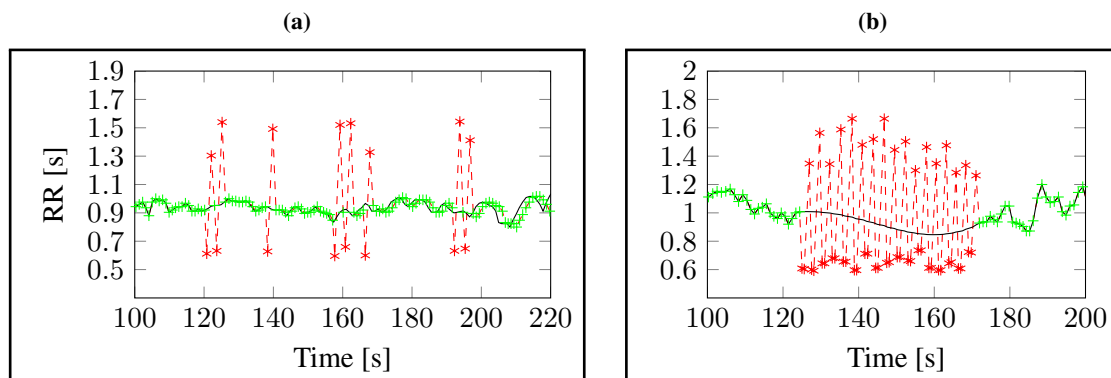


Figure 4.6: Cubic spline interpolation: (a) Single ectopic beats, such as PVCs, bi- and trigeminy are reasonably corrected. (b) Ectopic segments, such as PVC triplets, are replaced by a smooth line that better fits the data than a straight line but still does not reflect a physiological behavior. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown onto the original signal: TPs are labeled as stars in red and TNs as plus signs in green.

Non-linear Predictive Interpolation (NPI-CT)

Although this approach is based on chaos theory, it assumes that beat-to-beat changes in heart rate occur in a deterministic way. Hence, similar structures in the same RR interval time series, located even far away from the actual segment, are able to predict those RR intervals, which

would have been present if no ectopic beat had occurred. This method was first described by Lippman et al. and can be used to correct single ectopic beats as well as sequential ones [41]. It is different to cubic or linear interpolation since it does not assume any mathematical relationship between successive RR intervals. Before applying the algorithm, the number of erroneous RR intervals has to be known. As annotated signals are used, this algorithm only edits the determined RR intervals. Further, the user has to specify how many NN intervals have to be considered before and after the ectopic segment. In terms of this thesis, five previous and four following RR intervals were used, as suggested by Lippman et al. [41]. The algorithm searches in the phase space for the most similar ectopic free segment by determination of the smallest Euclidean distance to the test set. Then, this segment is used to replace the ectopic one (see figure 4.7b).

Although this method performs very well for few ectopic beats (see figure 4.7a), it cannot correct long ectopic segments reliably (see figure 4.7b). The longer the ectopic segment is, the fewer comparison segments are available that can be used to correct the ectopic segment. As a result, bad-matching segments are used to replace the whole interval.

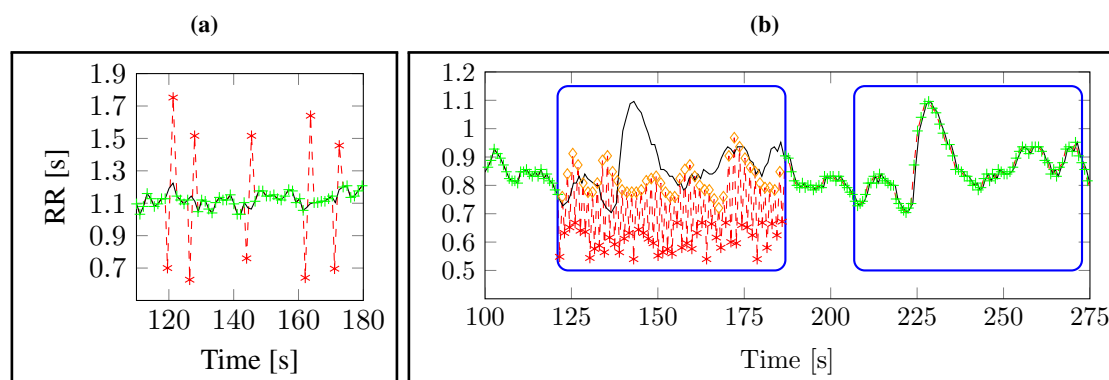


Figure 4.7: Non-linear predictive interpolation: (a) Single ectopic beats are reliably replaced, whereas long ectopic segments are completely replaced by similar structures in the signal (compare the two blue marked areas in (b)) that do not consider physiologic relations. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown on the original signal: TPs are labeled as stars in red, FPs as diamonds in orange and TNs as plus signs in green.

4.1.3 Filters

Filters edit the RR interval time series by replacing erroneous intervals with calculated values, in the same way as interpolation techniques do. Thereby, filters can either operate in the time- or in the frequency-domain, whereby some interpolation methods can also be assigned to time-domain filters. However, instead of determining a function that fits the data points, the filters calculate a statistical parameter, like the median or the mean, to replace the desired RR interval.

Interpolation of degree zero is an exceptional case, since it is also a mean filter with a fixed window length. Frequency domain filters, such as wavelet filters, do not change a specific RR interval but rather alter frequency bands and thus, always affect several RR intervals.

Median Filters

Median filters simply replace the erroneous RR interval with the median of a specified window, with the actual RR interval in the center. Further, the window length can either be fixed or adaptive.

Different types of median filters are widely used for the removal of impulse noise, like *salt-and-pepper noise*, in digital image processing [16]. Since ectopic beats result in sharp impulses, Kumaravel et al. considered median filters for the application of RR interval processing [36]. Although median filters are well-suited to process impulse noise, they fail if the noise density is too high. Therefore, Hwang et al. improved the median filter by adapting the window length dynamically when the median is false [28]. However, this filter still fails if the noise density is very high, as the processed values are reused again and again. As a result, Hsieh et al. proposed an *adaptive rank order filter* for denoising images [26]. This filter addresses two main issues: The window size is increased, if all values are noise (not only the median) and furthermore, it calculates an adaptive threshold for the determination of the correctness of the median. Kumaravel et al. adapted these two approaches for ectopic beat correction and compared them with a standard median filter with a fixed window length [36]. Every window-based filter can only correct values that lie within the following range, with respect to the window edges:

$$\nu = \frac{\text{window length}}{2} - 1. \quad (4.9)$$

Figure 4.8 displays this behavior on the basis of an *adaptive median filter* with a maximum window size of 25 values. The RR intervals at the beginning and the end of the ectopic segment can be reasonably edited whereas the intervals in the middle are too far away from the nearest correct NN intervals to be replaced by a correct median.

Median Filter with fixed Window Length (MEDFIL) is the simplest median filter. Erroneous RR intervals are replaced with the median of the current window. Of course, the window length has to be specified and normally ranges from three to eleven RR intervals [36]. In this thesis a window length of three intervals is used. The advantage of this filter is that it is a quick and easy approach. However, if adjacent RR intervals are erroneous, incorrect intervals are used to determine the median (see figure 4.9). This may result in inappropriate median values. Further, especially in long ectopic segments, the error propagates to the next RR intervals.

Adaptive Median Filter (AMFILT) overcomes the drawback of standard median filters to being unable to adapt to longer lasting artifacts. An AMFILT checks if the calculated median lies in a reasonable range and increases the window size until this criterion is satisfied [16, 36]. This offers great advantages over traditional filters. First, the filter provides information about the correctness of the median and second, it is able to automatically include more values to calculate

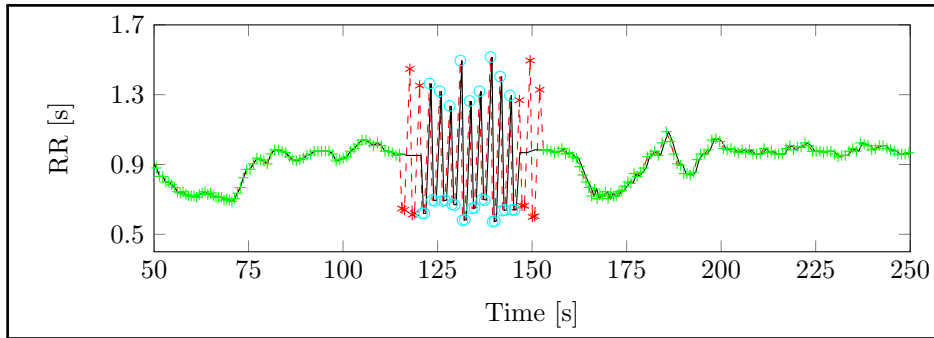


Figure 4.8: Window filter: Values that lie more than ν RR intervals away from the window edges cannot be processed. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown onto the original signal: TPs are labeled as stars in red, FN as circles in cyan and TNs as plus signs in green.

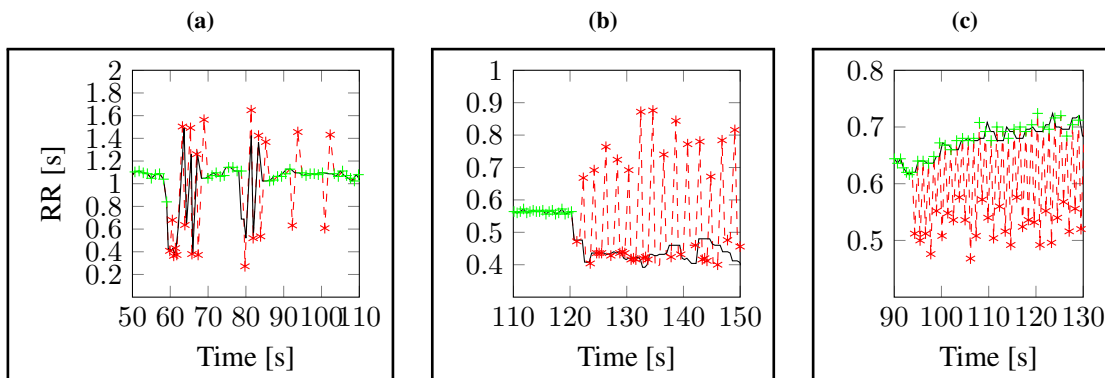


Figure 4.9: Median filter: (a) Only ectopic beats that fit into the fixed window can be processed. Single ectopic beats are always edited whereas longer segments cannot be corrected. (b) PVC couplets cannot be reasonably corrected, since the median is always too low. (c) Atrial bigeminy can be corrected completely, since the number of NN intervals is greater than the erroneous RR intervals. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown on the original signal: TPs are labeled as stars in red and TNs as plus signs in green.

a more appropriate median. In this thesis the criterion for median rejection is a combination of the two criteria described by Chan et al. [16] and Kumaravel et al. [36]. Chan et al. suggested the following requirement:

$$RR_{min} < RR_{med} < RR_{max}, \quad (4.10)$$

which means that the median must not be equal to the minimum or maximum [16]. Since this approach is reasonable in grey scale images with salt-and-pepper noise, it has to be adapted for ectopic beat processing. Therefore, the calculated median must also lie within 20 % deviation from the median of the whole signal [36]. In addition, the median must be different from the original value, since this value is erroneous.

The median filter starts with a window size of three values and enlarges the window up to 25 values. Inclusion of more RR intervals is not reasonable if the intervals are too far away to influence the actual value. Figure 4.10 illustrates the ability of the adaptive median filter (AMFILT) to not only correct single ectopic beats, bigeminy and trigeminy, but also longer ectopic segments. Nevertheless, long ectopic segments have the drawback that successive medians are often equal and, similar to in interpolation, this tends to flat out the RR interval time series (see figure 4.10c).

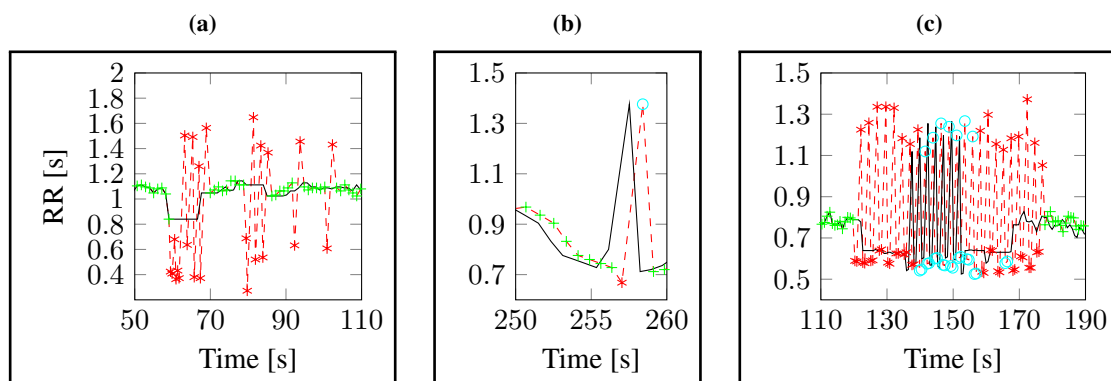


Figure 4.10: Adaptive median filter: (a) Single and sequential PVCs are replaced by the same median as long as the processed segment fits in the maximal window length. (b) A downward trend of the signal, starting at 250 s and ending at 255 s, results in a failure of the AMFILT to correct the ectopic beat. In the first step of the AMFILT algorithm, the erroneous RR interval is corrected, but then the new RR interval is compared to the median of the whole signal, which is much higher than the values inside the shown window, and the correction is not accepted. (c) Sequential ectopic beats, like PVC couplets, may only be corrected at the edges of the ectopic interval. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown on the original signal: TPs are labeled as stars in red, FNs as circles in cyan and TNs as plus signs in green.

Adaptive Rank Order Filter (AROFIL) was first suggested by Chang-Hsiung et al. for removal of high density noise in images [26]. This filter was adapted for processing highly

erroneous RR interval time series by Kumaravel et al. [36]. Identical to the AMFILT, this algorithm increases the window length as long as the median of the window is false. In addition, three improvements are made with respect to the AMFILT. Instead of using a fixed threshold, the AROFIL compares the actual median with the last accepted median for the evaluation of its correctness (20 % deviation are used in this thesis). This provides a great advantage when trends are present next to ectopic beats (compare figure 4.10b with 4.11b). Further, the window is also enlarged if all values inside are erroneous. In this thesis the initial window length was set to three and the maximum to 25, for the same reasons as mentioned for the AMFILT. In general, this filter performs very similar to the AMFILT, especially when dealing with moderate noise (compare figure 4.10a with 4.11a). However, when noise density increases, the AROFIL seems to outperform the AMFILT, at least according to the findings of Kumavarel et al. [36]. Nevertheless, large ectopic segments cannot be edited because of the limitation due to the maximal window length (as illustrated in figure 4.11c).

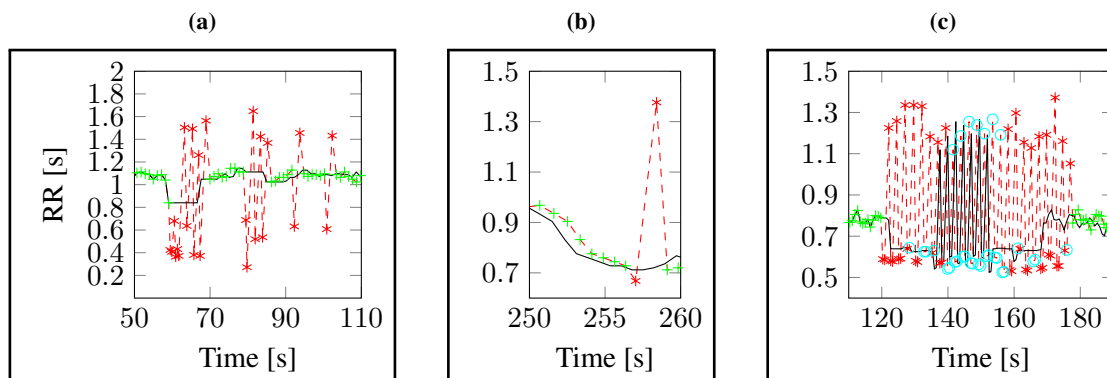


Figure 4.11: Adaptive rank order filter: (a) Single or sequential PVCs are edited as long as the ectopic segment fits into the window. (b) Since the threshold is also adaptive, RR intervals can even be edited when trends are present. (c) Sequential ectopic beats, such as PVC couplets, may only be corrected at the edges of the ectopic interval. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown on the original signal: TPs are labeled as stars in red, FNs as circles in cyan and TNs as plus signs in green.

Impulse Rejection Filter (IRFILT) have to detect ectopic beats on its own, in contrast to the other mentioned median filters. The filter was introduced by McNames et al. and is based on a simple impulse identification with subsequent replacement of this value by the median [55]. The absolute deviation of all RR intervals with respect to the median of the whole RR time series is computed. These values are set in relation to the median of the absolute deviation in order to detect sharp impulses. The following formula describes the calculation of the test statistic:

$$D(t) = \frac{|RR(t) - RR_m|}{1.483 \cdot \underset{i=1 \dots n}{\text{median}}(|RR_i - RR_m|)}, \quad (4.11)$$

where $\text{median}_{i=1\dots n}(|RR_i - RR_m|)$ is the median of the absolute deviations of all RR intervals with respect to the median RR interval (RR_m) of the entire RR time series. Thereby, the denominator is a robust estimate of the standard deviation of an equivalent Gaussian distribution [55].

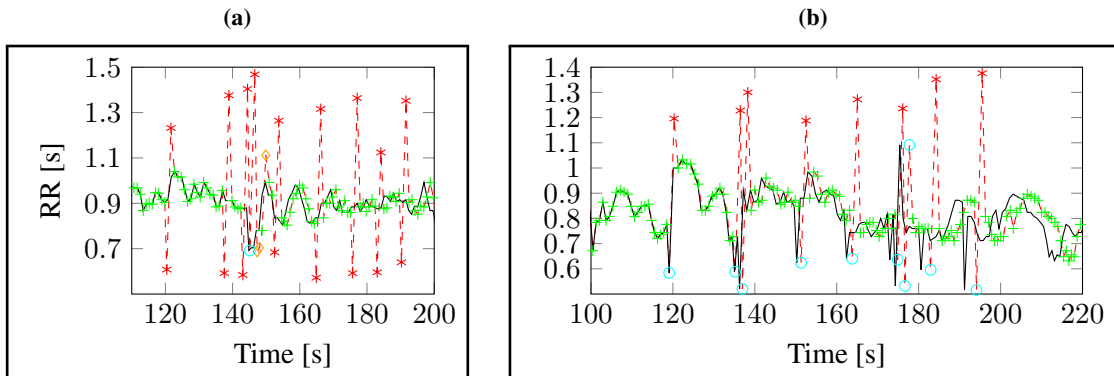


Figure 4.12: Impulse rejection filter: (a) Sharp impulses are always corrected, whether their focus of origin is ectopic or not. In weakly fluctuating signals PVCs are mostly corrected reasonably, but steep trends may also be falsely edited. (b) In more fluctuating signals, some ectopic beats are missed to be identified, because the detection is based on the median difference, which is large in such signals. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown on the original signal: TPs are labeled as stars in red, FNs as circles in cyan, FPs as diamonds in orange and TNs as plus signs in green.

Figure 4.12a illustrates that every sharp impulse, which deviates enough from the median, is replaced, no matter if its origin is ectopic or not. In contrast, in RR interval time series with high modulations some impulses do not deviate enough to be detected as errors (see figure 4.12b). Hence, ectopic beats are corrected only partly.

Mean Filters

Mean filters perform very similar to median filters, but are less frequently used when dealing with impulse noise, such as ectopic beats. The reason lies in the statistical characteristics of the two parameters. The mean is rather sensitive to outliers, whereas the median is not. Therefore, the median is better suited when dealing with distinct sharp impulse noise. Nevertheless, the mean can also be a reasonable choice, if it is guaranteed that no outlier is used for its calculation. A very easy approach is to apply an upstream threshold filter that deletes non physiologic RR intervals, as suggested by Mietus [56]. Begum et. al described a completely different method, based on the k nearest neighbors (kNN) of the actual RR interval [11]. Their algorithm searches for the k nearest RR intervals in the time series, calculates the mean and replaces the erroneous RR interval with it.

Sliding Window Average Filter with Upstream Threshold Filter (SWAFIL) is a combination of a threshold filter with a downstream window average filter. Since only non-physiologic RR intervals should be excluded, the threshold is set to 40 %. The second parameter is the window length (n), which stays fixed, and is set to five intervals.

$$RR_{corr} = \frac{1}{n} \cdot \sum_{k=i-\frac{n-1}{2}}^{i+\frac{n-1}{2}} RR_k, \quad (4.12)$$

where i denotes the original RR interval that is replaced by RR_{corr} .

Beside this correction ability, the filter is also suited to detect erroneous RR intervals. As can be seen in figure 4.13a, single ectopic beats and bigeminy or trigeminy are edited reasonably, although some NN intervals are falsely processed. However, figure 4.13b clearly shows the weaknesses of this filter. If there are multiple ectopic beats without a compensatory pause, the calculated mean RR interval is too short with respect to the previous and following NN intervals.

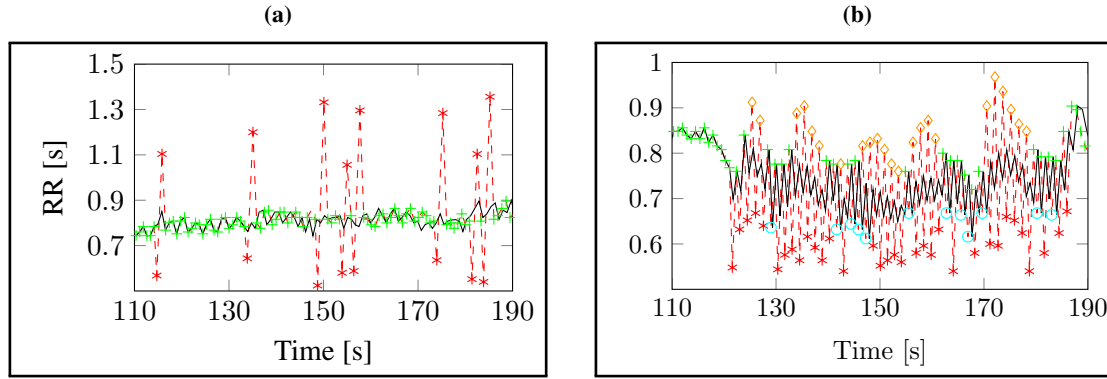


Figure 4.13: Sliding window average filter with upstream threshold filter: (a) Single PVCs, bi- and trigeminy are edited reasonably. (b) Ectopic segments where no compensatory pauses are present cannot be edited reasonably since the mean is too low with respect to the NN intervals. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown on the original signal: TPs are labeled as stars in red, FNs as circles in cyan, FPs as diamonds in orange and TNs as plus signs in green.

kNN Average Filter (kNN-AF) was suggested by Begum et al. for ectopic beat correction [11], whereby k denoted the number of adjacent considered RR intervals. First, the algorithm determines if the k previous or following beat occurrence times are closer to the beat occurrence time under consideration. Then, the mean of the k RR intervals, which are closer to the actual beat occurrence time, is calculated and replaces the erroneous interval. Since this filter is based on the mean, it has some drawbacks when the length of the compensatory pause differs from the previous RR interval. Figure 4.14 displays that single ectopic beats can be edited very well, as the first erroneous interval can be replaced by the mean of the previous RR intervals and

the second by the mean of the following ones. Thus, the mean does not depend on the length of the compensatory pause. However, if successive erroneous RR intervals occur (like in bi- or trigeminy) the mean is biased and so the impulses can only partly be reduced. The problem even becomes worse if multiple PVC couplets or triplets are present, since the mean is mainly based on too short RR intervals and therefore the filter is not able to reasonably replace the erroneous interval.

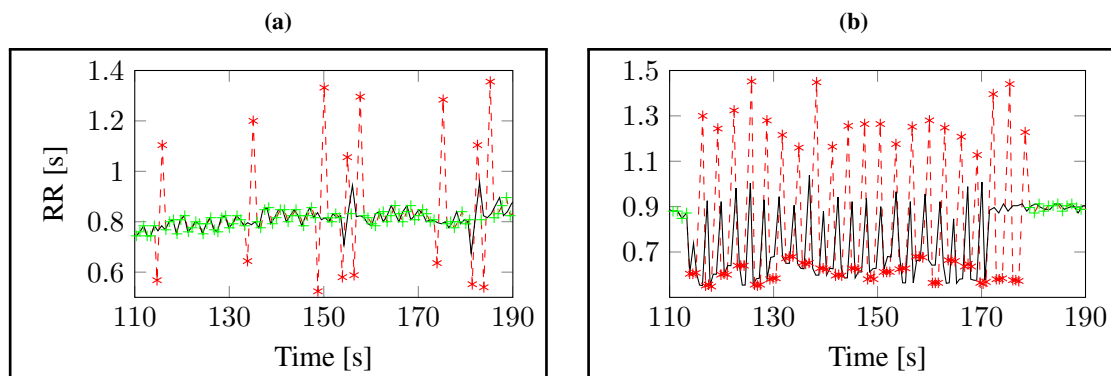


Figure 4.14: *k* nearest neighbors average filter: (a) Single ectopic beats, bi- and trigeminy are corrected rather well although there are still impulses present after the editing of trigeminy. (b) Successive couplets or triplets are edited but can only partly restore the original RR interval lengths. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown on the original signal: TPs are labeled as stars in red, FNs as circles in cyan, FPs as diamonds in orange and TNs as plus signs in green.

Wavelet Filter (WAVE-F)

In contrast to the mean and median filters, wavelet filters operate in the frequency domain. This offers a great advantage since sharp impulses, like from ectopic beats, can be clearly detected and corrected. Wavelet filters are also used for the detection and removal of trends, as performed by several detrending methods [38, 81]. Therefore, wavelet filters are more common in pre-processing, rather than in correction. However, Keenan suggested to use wavelet filters for the correction of ectopic beats [33]. He described a *db4 wavelet filter* that decomposes the RR interval time series into five sub-bands (see figure 4.15).

Ectopic beats can be clearly detected and may be corrected by two different approaches: Hard or soft thresholding. Hard thresholding considers the wavelet coefficients in the first two frequency bands, which correspond to the high frequency components. The coefficients $w(n)$ above a defined threshold T are reduced to zero, whereas all other coefficients are left unchanged:

$$w(n) = \begin{cases} 0 & \text{if } |w(n)| \geq T \\ w(n) & \text{if } |w(n)| < T, \end{cases} \quad (4.13)$$

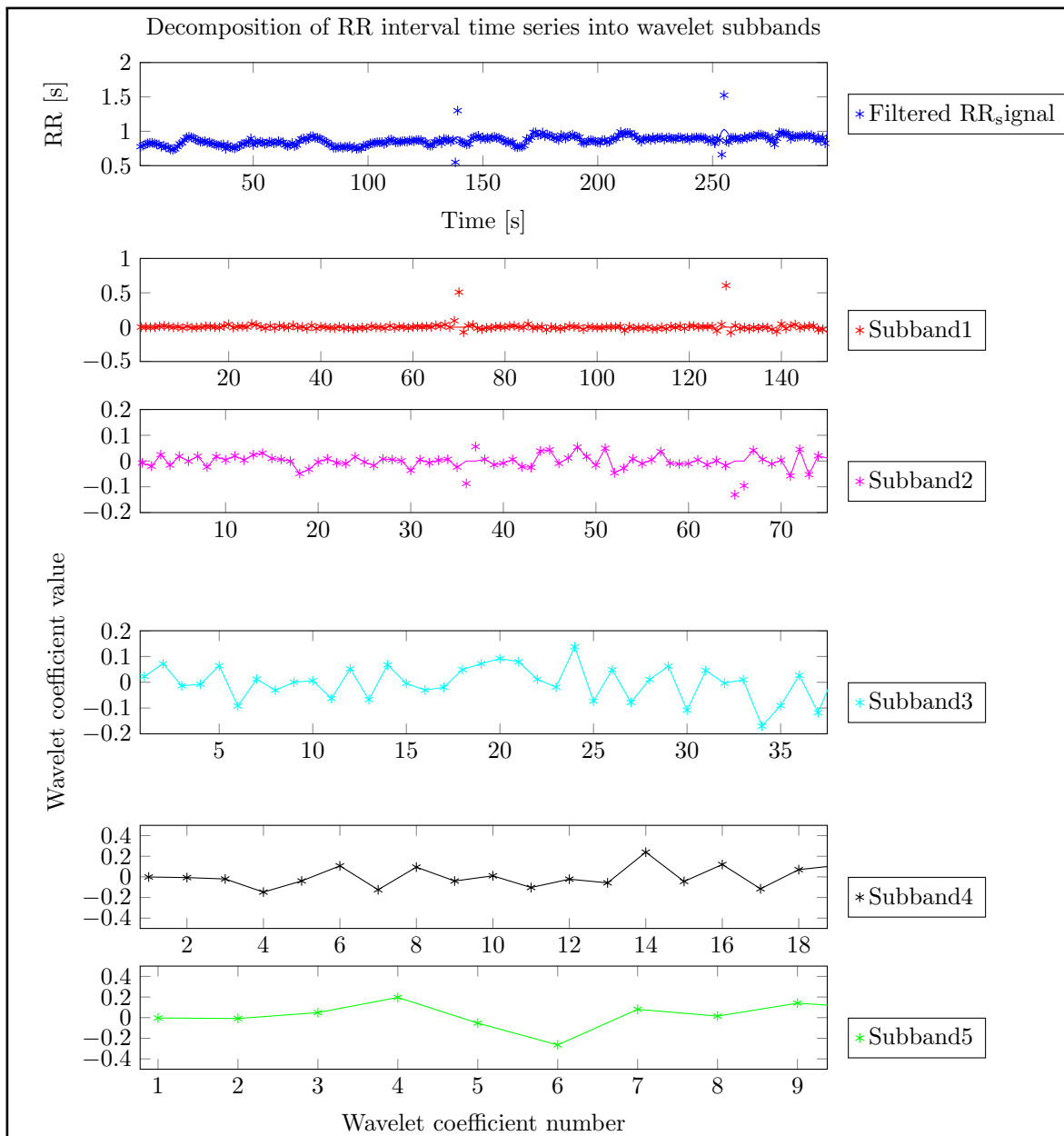


Figure 4.15: Wavelet filter: Decomposition of RR time series into five sub-bands. Highest frequency band on top (red), lowest on bottom (green). Stars denote original wavelet coefficients, the solid lines represent the coefficients after hard thresholding. Only the first two sub-bands are altered, where ectopic beats can be easily detected and removed.

where $n = 1, 2, 3, 4, \dots, N$ is the index of the RR intervals in the corresponding frequency band.

In soft thresholding the coefficients $w(n)$ of all sub-bands that are above the threshold T are also reduced to zero. Moreover, all other coefficients have the threshold subtracted from them so that they tend towards zero:

$$w(n) = \begin{cases} 0 & \text{if } |w(n)| > T \\ w(n) - T & \text{if } |w(n)| \leq T. \end{cases} \quad (4.14)$$

Despite of the successful removal of ectopic beats, wavelet filters always alter neighboring RR intervals as well. This results in a huge amount of FP intervals that change the frequency-domain HRV parameters immensely, especially in the case of single ectopic beats. The longer the ectopic segment is, the fewer additional RR intervals are falsely edited when considering relative values (compare figure 4.16a with 4.16b).

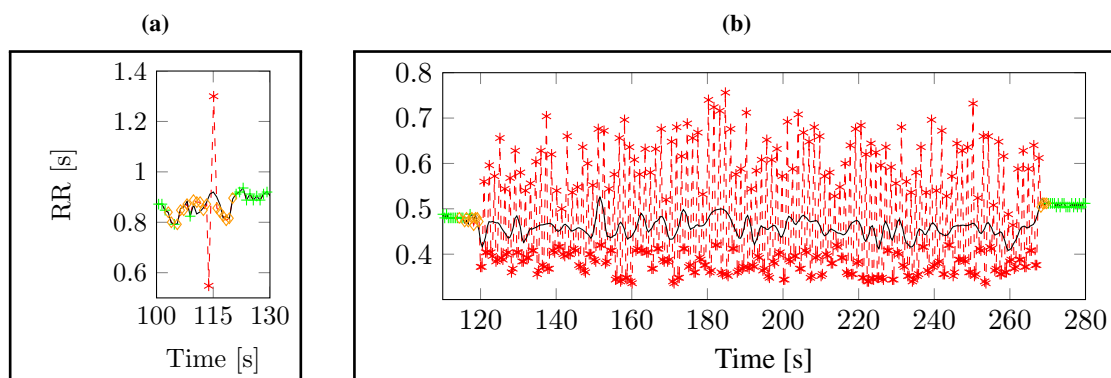


Figure 4.16: Wavelet filter: (a) Correction of a single PVC results in a huge amount of falsely edited RR intervals (FPs). (b) Correction of multiple ectopic beats results in far less FPs with respect to the total number of changes. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is performed onto the original signal: TPs are labeled as stars in red, FPs as diamonds in orange and TNs as plus signs in green.

4.1.4 Models

The algorithms mentioned in this section do not have mathematical common relationships and their complexity differs over a wide range. However, they cannot be assigned to removal, interpolation or filtering, as their correction algorithms are much more complex. Therefore, all of these methods are summarized as “Models” in this section.

Gross Positioning of Beats (GP-IIA)

This method was designed as an easy and fast correction approach. It was suggested by Mateo et al. as coarse positioning of beats for visualization before the actual correction algorithm

is applied [53]. This method does not only coarse position false beats, but rather performs a classification of all beats. According to the integral pulse frequency modulation (IPFM) model the SA node is band limited and thus, the variation of the instantaneous heart rate is band limited too [53]. With the help of Lagrange's interpolation formula the derivative of the instantaneous heart rate (r'_k) can be estimated at the k^{th} beat (at beat time t_k) by considering the previous and following beat occurrence times, $t_{k\pm 1}$ by

$$|r'_k| = 2 \left| \frac{t_{k-1} - 2t_k + t_{k+1}}{(t_{k-1} - t_k)(t_{k-1} - t_{k+1})(t_k - t_{k+1})} \right| < U, \quad (4.15)$$

where U is a predefined threshold that is calculated by

$$U = 4.3 \cdot std(r'_k), \quad (4.16)$$

and limited by 0.5.

Equation 4.15 is just able to determine if one of the beats at t_{k-1} , t_k or t_{k+1} is erroneous, but not exactly which one is ectopic. However, if t_{k-1} is erroneous, the condition in equation 4.15 would not be met in the previous step. Therefore, six different test cases are carried out by insertion and deletion of beats in order to determine the beat type of t_k and t_{k+1} . Single ectopic beats are corrected by shifting them to an intermediate position between the previous and following normal sinus beat. FPs are simply deleted and FNs are corrected by insertion of an evenly spaced beat, with respect to the adjacent beats. Consecutive beats are simply corrected by insertion of multiple beats at evenly spaced positions. Figure 4.17 illustrates the correction of some common ectopic beats. Despite of its detection ability this method can also be used for correction. Since the goal of this thesis is the comparison of different correction algorithms, only beats detected as anomalous were processed.

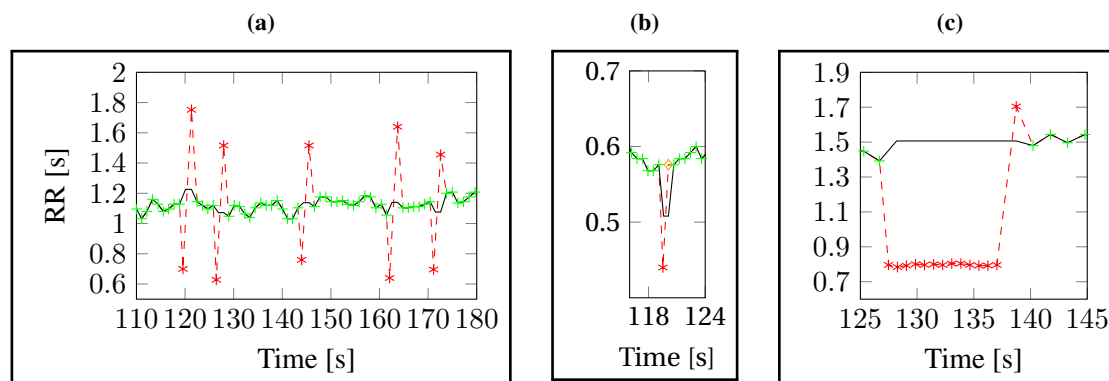


Figure 4.17: Gross Positioning of Beats: (a) Single PVCs are just replaced by two equally long RR intervals, (b) PACs are also replaced by two equally long RR intervals and thus result in two too short RR intervals, (c) Ectopic segments, such as ventricular tachycardia, are replaced by insertion of evenly spaced beats. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is showed on the original signal: TPs are labeled as stars in red, FPs as diamonds in orange and TNs as plus signs in green.

Correction	Description
Split	Missed heartbeat; divide IBI into two equal intervals
Split 3	Two missed heartbeats; Split IBI into three equal intervals
Combine	False trigger; combine two IBIs into one
Combine 2 / Split 2	Replace two IBIs with their average
Combine 3 / Split 2	Get two new IBIs as average of three
Combine 2 / Split 3	Get three new IBIs as average of two
Combine 3 / Split 3	Replace three IBI values with their average
Uncorrected	Could not apply any rule, but IBI appears faulty

Table 4.2: Order of combination rules for buffer

Buffer with Combination Rules (BUFFER)

Another fast correction algorithm was described by Rand et al. [69]. This approach applies several combination and/or splitting rules of adjacent RR intervals to correct erroneous intervals (see table 4.2).

In order to judge the usefulness of the correction, N previous RR intervals are maintained inside a buffer (see figure 4.18). Therefore, this approach implies a short time delay (about 3 to 6 s) to the processed intervals with respect to the original time series, similar to windowing filters and approaches that need more than one successive RR interval for signal processing. Additionally, a specific amount of RR intervals (the sum of them has to be at least 6 s) is also stored in a buffer past the actual interval to ensure that the corrected RR interval also fits to these values.

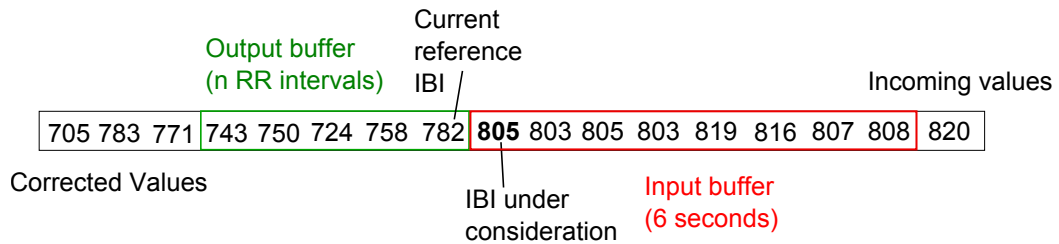


Figure 4.18: Structure and function of buffer: New values come from the right end and enter the input buffer, containing at least 6 s, including the IBI under consideration. This IBI is compared to statistical parameters of the IBIs maintained in the output buffer (5 IBIs in the original work of Rand et al. [69] and in this thesis).

Figure 4.19 shows the correction ability of this approach. This model is also able to perform detection on its own. Since pre-annotated signals were used in this thesis, only the correction ability was used.

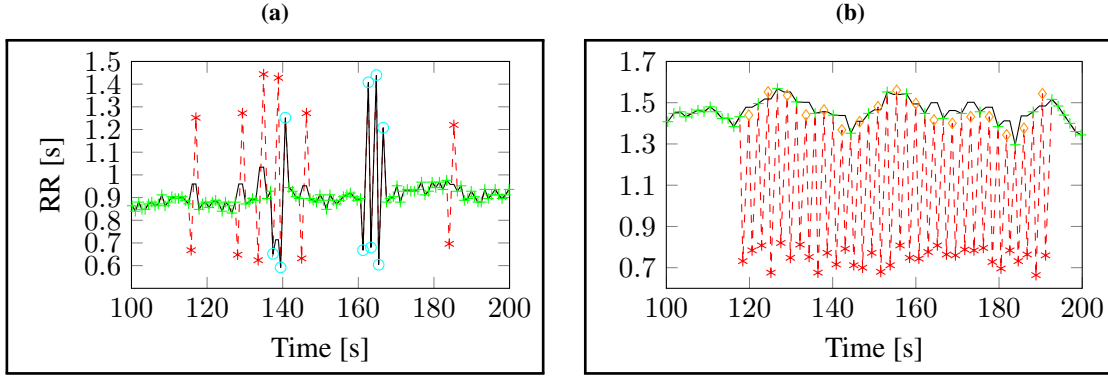


Figure 4.19: Buffer: (a) Single PVCs are corrected reasonably if the compensatory pause is nearly as long as the previous RR interval. However, the buffer cannot correct more than three successive RR intervals, resulting in FNs (cyan circles). (b) Atrial bigeminy cannot be corrected completely since the buffer has to correct at least two RR intervals and thus also changes the correct NN intervals between the erroneous intervals. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown on the original signal: TPs are labeled as stars in red, FPs as diamonds in orange, FPs as circles in cyan and TNs as plus signs in green.

Trend Predict Correction (TPC-HT)

Wen et al. developed a method that corrects ectopic beats based on trend correlation of the heart timing signal [85]. Basically, the predicted RR interval is composed of two parts, the trend and the turbulence:

$$RRI_{pred} = RRI_{trend} + RRI_{turb}. \quad (4.17)$$

The trend is simply calculated as the weighted mean of the n previous NN intervals:

$$RRI_{trend} = \sum_{t=t_k-n}^{t_k-1} w(t) \cdot RR(t), \quad (4.18)$$

where $w(t)$ is the time dependent weight. Since Wen et. al do not further specified the weighting function [85], exponential weights described by Citi et al. were used instead [19]. The turbulence can be seen as the slope of the previous NN intervals and is approximated by the following calculation, as described by Wen et al. [85]:

$$RRI_{turb} = I[t_k] \cdot E[t_k], \quad (4.19)$$

where $I[t_k]$ is just the sign of the turbulence and $E[t_k]$ is the quantity. $I[t_k]$ can be judged by the signs of the slopes at t_{k-1} and t_{k-2} :

$$I[t_k] = \frac{k_1 \cdot k_2 \cdot (k_1 + k_2)}{|k_1 \cdot k_2 \cdot (k_1 + k_2)|}. \quad (4.20)$$

The slopes are determined by the following calculation (in addition a tiny value is added to both slopes to avoid zero, not shown):

$$k_1 = \frac{y(t_{k-1}) - y(t_{k-2})}{y(t_{k-1}) + y(t_{k-2})} \quad \text{and} \quad (4.21)$$

$$k_2 = \frac{y(t_{k-2}) - y(t_{k-3})}{y(t_{k-2}) + y(t_{k-3})}. \quad (4.22)$$

The quantity of the turbulence ($E[t_k]$) is calculated by consideration of the two slopes k_1 and k_2 , RRI_{trend} and the standard deviation of the previous RR intervals SD_{RRI} :

$$E[t_k] = RRI_{trend} \cdot \frac{\sqrt{|k_1 \cdot k_2|}}{\frac{a+b}{SD_{RRI}}}. \quad (4.23)$$

The two coefficients a and b are not specified by Wen et al. [85] and thus were approximated by comparison of several values in different magnitudes.

Figure 4.20 illustrates that the correction performance of the trend predict correction method is very reasonable, both for single PVCs and for ectopic segments.

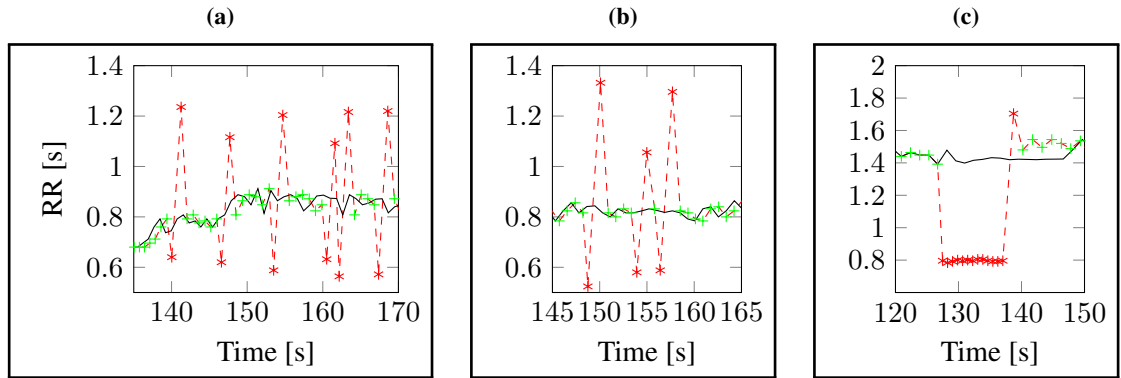


Figure 4.20: Trend predict correction: (a) Single ectopic beats, (b) bigeminy and trigeminy are reasonably corrected. (c) Longer ectopic segments, such as ventricular tachycardia, are corrected reasonably at least for the first RR intervals and then the RR interval time series flats out because of the diminishing modulating part. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown on the original signal: TPs are labeled as stars in red and TNs as plus signs in green.

Integral Pulse Frequency Modulation (IPFM)

The *IPFM model* is one of the few approaches that rely on a physiological relationship. The model is able to predict the ANS activity on the SA node by simulating the series of cardiac events as firings of the SA node [84]. Thereby, the IPFM integrates the input signal until a beat

is generated and is then reset to zero [53] (see figure 4.21a). The index of the k^{th} beat can be calculated as follows:

$$k = \int_0^{t_k} \frac{1 + m(t)}{T} dt. \quad (4.24)$$

Here, t_k denotes the occurrence time of the k^{th} beat, $m(t)$ is the modulating part of the heart rate and T is the mean RR interval length. Hence, the k^{th} beat can be interpreted as the integration of the instantaneous heart rate over the actual RR interval (from 0 to t_k , since the function is reset at every beat). The RR interval time series is therefore the difference of adjacent beat occurrence times. Further, Mateo et al. introduced the *heart timing (HT) signal* [53]:

$$ht(t) = \int_0^t m(\tau) d\tau. \quad (4.25)$$

Hence, the HT signal can be seen as the modulation of the IPFM model.

Integral Pulse Frequency Modulation Model with s-parameter (IPFM-S) was introduced first by Mateo et al. [53]. According to the IPFM model, the integrator is reset too early if an ectopic beat occurs (see figure 4.21a). This results in a lower integration value than expected and is denoted as s . Additionally, an indexing function of the beat occurrence times is introduced (see figure 4.21a). Once the index of the artifact is known, the indexing function is split into a forward and a backward function. The forward function is based on the normal sinus beats prior to the ectopic beat (see figure 4.21b blue line) and the backward function is based on the normal sinus beats following the ectopic beat (see figure 4.21b green line). Both functions are extrapolated to the neighboring beat until they overlay (see figure 4.21b red stars). The vertical difference of the two indexing functions is the s -parameter and can be calculated as follows:

$$\hat{s} = \frac{1}{t(k_e) - t(k_e - 1 + s)} \int_{t(k_e - 1 + s)}^{t(k_e)} (\hat{x}^f(t) - \hat{x}^b(t)) dt, \quad (4.26)$$

where k_e denotes the index of the ectopic beat, \hat{x}^f is the forward indexing function, and \hat{x}^b is the backward indexing function.

The correction of the artifact is carried out in the HT signal, which describes the modulating part of the RR interval time series, according to the IPFM model:

$$ht(t_k) = ht(t(k)) = kT - t(k) \quad (4.27)$$

$$= kT - t_k, \quad \text{for } k < k_e \quad (4.28)$$

$$ht(t_k) = ht(t(k - 1 + s)) \quad (4.29)$$

$$= (k - 1 + s)T - t(k - 1 + s) \quad (4.30)$$

$$= (k - 1 + s)T - t_k, \quad \text{for } k > k_e, \quad (4.31)$$

where T is the mean RR interval length. T is calculated as:

$$T = \frac{t_N}{N + \sum_j \hat{s}_j}, \quad (4.32)$$

where N is the final beat order, $\sum_j \hat{s}_j$ is the sum of all jumps, and t_N is the last beat occurrence time. Finally, the RR interval time series is obtained by subtraction of the differentiation of the HT signal from the mean RR interval length T :

$$RR(t) = T - ht'(t). \quad (4.33)$$

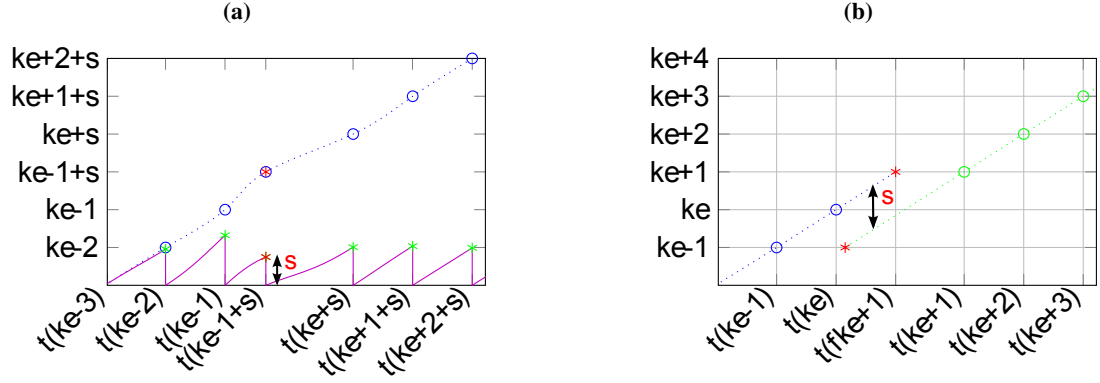


Figure 4.21: Determination of the s -parameter according to the IPFM model: (a) IPFM model showing the indexing function in dependence on the beat occurrence times t_k in blue with the ectopic beat marked in red. The integration function of the IPFM model is displayed in purple and reset too early at the ectopic beat. (b) Determination of the s -parameter by forward and backward extrapolation of the two indexing functions to the next index. The vertical difference is the s -parameter.

This approach is able to reliably correct single PVCs and bi- and trigeminy (see figure 4.22a). PACs can also be corrected but always result in one FP RR interval, since the IPFM model is based on beat occurrence times and thus, always the actual and the following RR interval are changed (see figure 4.22b). Sequential ectopic beats, like PVC couplets, result in a model related drift of the altered RR intervals (see figure 4.22c).

Integral Pulse Frequency Modulation Model with δ -parameter (IPFM-D) was suggested by Solem et al., since calculation of the s -parameter may be rather time consuming [76]. The δ -parameter corresponds to the time shift of the beat occurrence times followed by an ectopic beat and is related to the IPFM model as follows:

$$\int_0^{t_k} 1 + m(\tau) d\tau = kT_0 + \delta. \quad (4.34)$$

Similar to the s -parameter the correction is performed in the *heart timing signal*:

$$d_{HT\delta}(t_k) = \begin{cases} kT_0 - t(k) & \text{for } k \leq k_e \\ kT_0 - t(k) + \delta & \text{for } k > k_e. \end{cases} \quad (4.35)$$

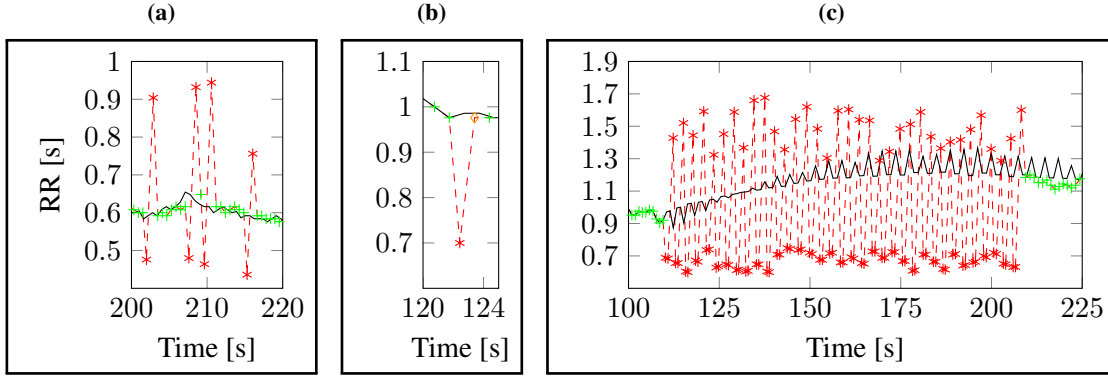


Figure 4.22: Correction of erroneous RR intervals with s -parameter according to the IPFM model: (a) Single ectopic beats and trigeminy are reasonably corrected. (b) A resetting beat is corrected by shifting the beat and thus two RR intervals are corrected instead of only one. (c) Interpolation over a longer period, like sequential PVC couplets, results in model-related drifts. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown on the original signal: TPs are labeled as stars in red, FPs as diamonds in orange and TNs as plus signs in green.

Solem et al. state that different δ -parameters may be used, dependent on how many beats prior to the ectopic beat are involved in the calculation. Equation 4.36 describes the determination of the N^{th} order δ estimator, based on Pascals triangle [76]:

$$\hat{\delta}_N = \sum_{l=0}^{N+1} (-1)^l \binom{N+1}{l} t_{ke+1-l} \quad N = 1, 2, \dots \quad (4.36)$$

Solem et al. compared the δ_1 , δ_2 and δ_3 parameter with each other and concluded that the consideration of more than one prior beat does not enhance the correction ability. Hence, the δ_1 -parameter was used in this thesis. In order to use the HT signal for correction of ectopic beats, the mean RR interval length \hat{T}_0 has to be calculated by:

$$\hat{T}_0 = \frac{t_k - \hat{\delta}_N}{K}. \quad (4.37)$$

As mentioned by Solem et al. the performance of the δ -parameter is nearly identical to that of the s -parameter. A comparison of figure 4.22c with figure 4.23c clearly illustrates the similar trend of the corrected signal (a bended line superimposed by spikes near original beats).

Integral Pulse Frequency Modulation Model with cost function (IPFM-C) was suggested by Brennan et al. [15]. This approach was designed to only account for single PVCs. Basically, the cost function calculates the quadratic deviation from the mean of the impulse height $E[\Psi]$ of

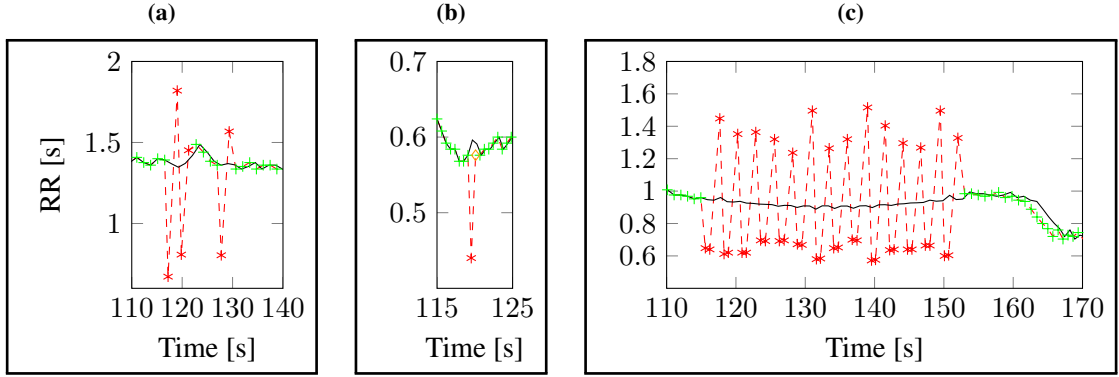


Figure 4.23: Correction of erroneous RR intervals with IPFM δ -parameter: (a) Bigeminy and single PVCs are reasonably corrected. (b) Similar to the s -parameter, a resetting beat is corrected by shifting the beat and thus two RR intervals are corrected instead of only one. (c) Ectopic segments, such as sequential PVC couplets, are corrected rather reasonably but also show the same drift effect as the correction with the s -parameter. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown on the original signal: TPs are labeled as stars in red, FPs as diamonds in orange and TNs as plus signs in green.

the integration function in dependence of the beat occurrence time points t_e . Thereby, just the $\pm M$ adjacent impulse heights Ψ , with respect to t_e , are considered:

$$C(t_e) = \sum_{S_k \in \Psi} (S_k - E[\Psi])^2, \quad (4.38)$$

where S_k denotes the height of the integration value at each beat occurrence time at the reset point, and can be calculated as follows:

$$S_k = \frac{1}{2\pi f_c} \sum_{j=k-M}^{k+1+M} Si(2\pi f_c(t_{k+1} - t_j)) - Si(2\pi f_c(t_k - t_j)), \quad (4.39)$$

where Si is the *sinc* function and f_c is calculated as the reciprocal of the doubled mean RR interval length \bar{I} :

$$f_c = \frac{1}{2 \cdot \bar{I}}. \quad (4.40)$$

Further, \bar{I} is the threshold for the integrate-to-threshold process (the integration function). The mean integration height $E[\Psi]$ may be approximated by the following equation:

$$E[\Psi] \simeq E[S_k] = \frac{1}{2M+2} \sum_{k=e-M-1}^{e+M} S_k. \quad (4.41)$$

The variable M has to be specified and denotes the number of adjacent beats considered for the calculations (the standard value is six). Figure 4.24 illustrates the behavior of the correction algorithm in two different cases: The upper row (figure 4.24a - 4.24c) displays a reasonable correction, since a clear minimum of the cost function can be determined (see figure 4.24b). In contrast, if the minimum is not obvious (see figure 4.24e), it is very likely that the correction does not work well (see figure 4.24d and 4.24f).

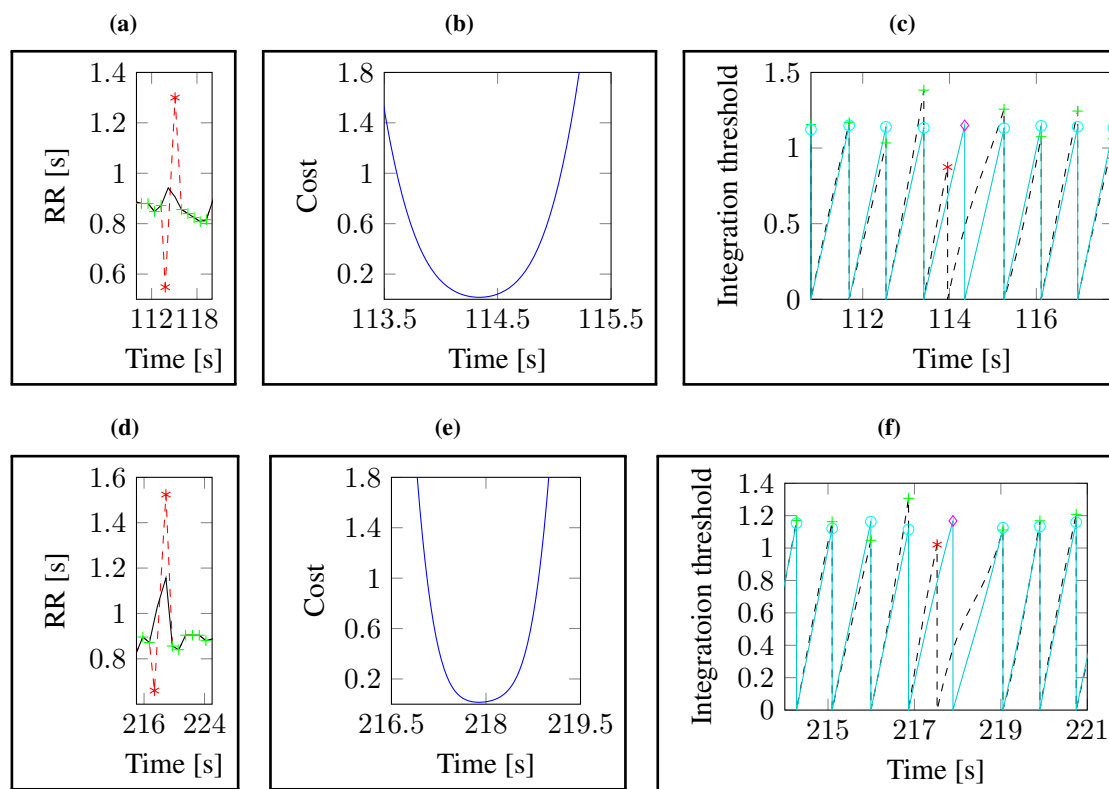


Figure 4.24: Correction of erroneous RR intervals with IPFM model and cost function: (a) Single PVC that is corrected reasonably, whereas an impulse still remains in (d). (b) Corresponding cost function of (a) with distinct minimum, (e) shows the broader, not distinct minimum of time series (d). The integration function of the upper row before correction is displayed in (c) dashed in black, the correct beats are labeled as a green plus, and the ectopic beat with a red star. After correction, the integration function is changed in a way that the integration heights are more uniform ((c) corrected integration function is displayed solid in turquoise, correct beats are labeled as cyan circles, and the ectopic beat with a purple diamond). (f) In the lower row the integration function is displayed in the same way.

Point Process with History Dependent Inverse Gaussian Distribution (PPHDIG)

Similar to the IPFM model this approach is based on a physiologically motivated model. Citi et al. mentioned that the Gaussian random walk model with drift is an elementary, stochastic integrate-and-fire model that is able to reflect afferences to the SA node [19]. These excitatory inputs are responsible for the basal cardiac rhythm and the influence of the autonomic nervous system through the sympathetic and para-sympathetic inputs. Hence, Citi et al. introduced a history-dependent, time-varying model based on the inverse Gaussian probability distribution of the waiting time until the next beat occurrence. The probability of the length of the next RR interval, $\tau - u_k$, is described at any beat event u_k by the probability density function (PDF):

$$f(\tau - u_k | \mu(H_k, \theta(t)), \lambda(\theta(t))) = \sqrt{\frac{\lambda(\theta(t))}{2\pi(\tau - u_k)^3}} \cdot \exp\left(-\frac{1}{2} \frac{\lambda(\theta(t))(\tau - u_k - \mu(H_k, \theta(t)))^2}{(\tau - u_k)\mu^2}\right). \quad (4.42)$$

H_k is the history vector and contains the P previous RR intervals (standard parameter $P = 5$). Further, λ denotes the shape parameter and μ the mean of the inverse Gaussian distribution. Both depend on the time varying parameters $\theta(t) = \theta_1(t), \dots, \theta_{P+1}(t)$, whereby $\lambda(\theta(t))$ is simply $\theta_{P+1}(t)$. The history dependent mean is a regression of the past P RR intervals with time-varying weights:

$$\mu(H_k, \theta(t)) = \sum_{i=1}^P \theta_i(t) w_{k-i}. \quad (4.43)$$

The unknown time-varying parameter set $\theta(t)$ is estimated by a local maximum likelihood method. At each time t , the parameter vector that maximizes the local log likelihood in a given observation interval $U_{m:n}$ is obtained. m denotes the index of the first beat in this interval and n the index of the last beat.

$$L(\theta(t) | U_{m:n}) = \sum_{k=m+P}^{n-1} \omega(t - u_{k+1}) \cdot \log[f(u_{k+1} - u_k | \mu(H_k, \theta(t)), \lambda(\theta(t)))], \quad (4.44)$$

where $\omega(\tau) = e^{-\alpha\tau}$ is an exponential weighting function for the local likelihood.

Then, the logarithmic probabilities of the following beat types are calculated as described by Citi et al. [19] to classify each beat as one of the following types: Extra beat, missed beat, misplaced beat, two misplaced beats and resetting beat (see figure 4.25).

The beat is only then classified as normal, if none of the hypotheses holds that are described in figure 4.25. According to the error type, beats are corrected by deletion, insertion or shifting of beats. Citi et al. stated that more than three consecutive ectopic beats are very unlikely and are therefore not considered in their approach [19]. Before acceptance of a correction, an improvement check is performed. This ensures that the new RR interval time series is always more reliable than the original one.

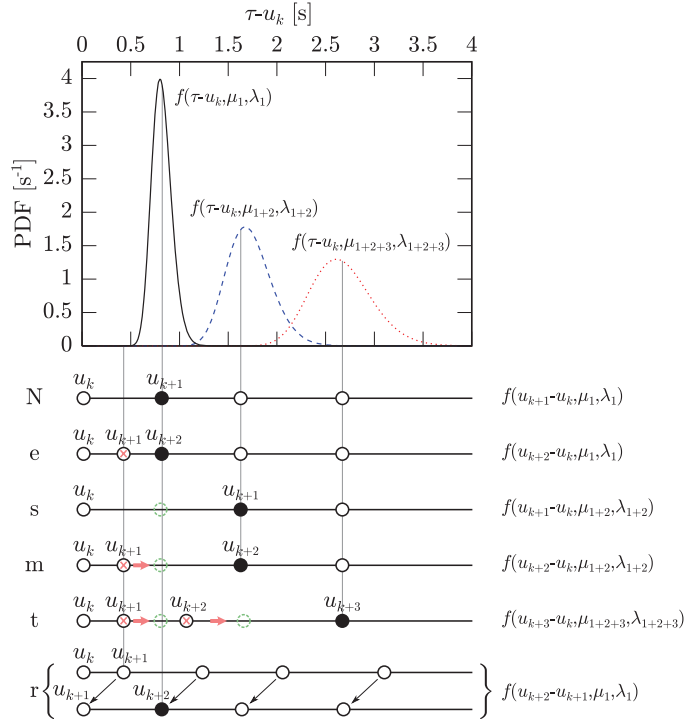


Figure 4.25: PDF of PPHDIG of three consecutive beats. The PDF of the first sequential beat is displayed solid in black, the second beat dashed in blue and the third beat dotted in red. If there is a beat next to every PDF maximum present and one additional beat occurs before the maximum of the first PDF the beat is denoted as an extra beat (e). If the first beat occurs next to the maximum of the second PDF it is likely that one beat was skipped (s). If the first beat occurs before the maximum of the first PDF and the second and third beat occur next to their expected time points, the first beat is misplaced (m). If only the third beat occurs next to the maximum of its PDF but the first and second beat both occur too early, then both are misplaced (t). If the first beat and every following beats occur too early, with respect to the maximum of their PDF, it is very likely that the first beat is a resetting beat (r). Only if all these hypotheses are wrong, a normal beat is present (n) and all beats should therefore occur next to the maximum of their specific PDF. Figure obtained from from Citi et al. [19].

Figure 4.26 displays the different correction abilities of the PPHDIG approach. Single PVCs, bigeminy (see figure 4.26a), trigeminy (see figure 4.26b) and PACs (see figure 4.26c) are reasonably corrected. Even two misplaced beats, referred as a single PVC couplet, are corrected well (see figure 4.26d). Since the method was not designed to edit longer ectopic segments, such as sequential PVC couplets and triplets, it fails to correct these types of arrhythmia (see figure 4.26e).

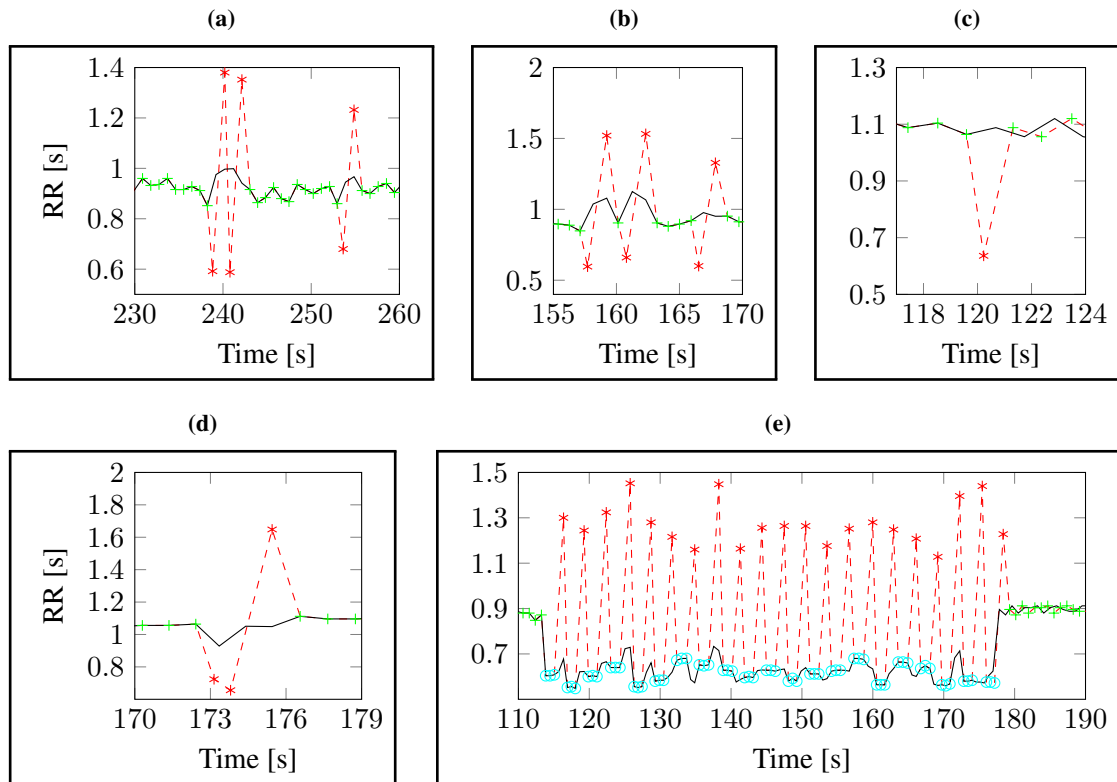


Figure 4.26: PPHDIG showing different cases of correction types: (a) Single PVCs and bigeminy, (b) trigeminy and (c) resetting beats are reasonably edited. (d) Even single PVC couplets can be corrected. (e) In contrast, ectopic segments, such as sequential PVC triplets cannot be edited correctly. The original signal is displayed dashed in red, the corrected signal solid in black. Annotation is shown on the original signal: TPs are labeled as stars in red, FPs as circles in cyan and TNs as plus signs in green.

4.2 Test Cases

Different test cases were designed to compare the behavior of each ectopic beat correction approach. Basically, three different tests were established, Test-1, Test-2 and Test-3. Test-1 and Test-2 are based on the same data set containing only correct NN intervals (termed *setN*), which

is artificially corrupted with ectopic beats. An in-house developed ectopic beat generator was used to artificially corrupt error-free RR interval time signals. This offers four great advantages. First, the right correction is already known since it is just the original signal. Second, the RR interval time series can be annotated correctly, as the type of each ectopic beat is known. This would not be necessarily the case if only the ECG signal is annotated without knowing the ectopic beat type. Third, this is the only possibility to ensure that all artifacts are known and correctly classified. Forth, the algorithm is able to simulate the most common single ectopic beats and heart disease associated arrhythmia in defined occurrence frequencies. This part is especially useful when comparing the different correction methods in the specific test cases. In all Test-1 cases, ten different test sets were created by application of the algorithm for ectopic beat generation ten times on *setN*. In contrast, Test-3 is based on a different data set (termed *setX*), which already contains ectopic beats. Hence, only one data set was used for Test-3.

Test-1: Correction of Single Ectopic Beats and Ectopic Segments

Test-1 is used to determine the influence of the different correction approaches on several HRV and statistical parameters. Therefore, error-free RR interval time signals were corrupted with an in-house developed ectopic beat generator (see section 4.4). Further, this test is split into weakly corrupted (Test-1A) and strongly corrupted (Test-1B) signals.

Test-1A contains only RR time series with single ectopic beats at a moderate density (about one to five ectopic beats with or without compensatory pause per 5 min signal).

Test-1B includes all of the ectopic beat types described in section 4.4 (PVC, PAC, PVC couplets and triplets, bi- and trigeminy, sustained and non-sustained ventricular tachycardia).

Test-2: Robustness

Test 2 is used to determine the robustness of each ectopic beat correction algorithm. This test also uses error-free RR interval time series and adds single PVCs and PACs at a ratio of 1:1 by application of the ectopic beat generator. Test-2 determines the robustness as shown in figure 4.27. The number of erroneous RR intervals is increased in all three test cases, but in a different manner. As a result, a step-wise increase in the RMSD of the HRV parameters and the RR intervals can be observed. Since this increase is not well detected in the statistical parameters, they were not considered for robustness determination. The test is split into the following three sub-tests:

Test-2A determines the maximal amount of single ectopic beats that can be processed by each method. The test starts to insert several PVCs or PACs, which are spaced 150 NN intervals, enough to not influence each other. Then, the space between them is narrowed by introducing more and more PVCs, until the error increases sharply.

Test-2B determines the robustness of each method regarding the maximal ectopic segment length. It starts to insert one single ectopic beat and then raises the amount of successive PVCs or PACs until the error shows a sharp increase.

Test-2C uses the result from Test-2A (the maximum amount of successive ectopic beats that can be processed) for each correction approach and then performs the same

procedure as in Test-2A, the narrowing of the space between ectopic segments. The result is the necessary amount of correct NN intervals between ectopic segments and a measure of the robustness against multiple ectopic beats (together with the results of Test-2B).

Test-3: Natural Ectopic Beats

In addition to the tests with the artificially corrupted data sets, one data set with real ectopic beats was used to evaluate the performance of the algorithms on real signals (Test-3). This offers the possibility to compare the findings of the artificially corrupted signals with natural ones. This data set is based on the databases used by several research groups that described the implemented algorithms. Thus the outcome of Test-3 can be used for reproducibility and comparison to results in literature.

4.3 Data Source

The entire set of ECG signals used in this thesis is based on the following three Physionet databases [23]:

The European ST-T Database [78]

Originally, this database was set up to evaluate algorithms for the analysis of ST and T-wave changes. It consists of 90 two-channel ambulatory ECG recordings that were annotated independently by two cardiologists. The signals were obtained from 79 subjects (70 men aged 30 to 84 and 8 women aged 55 to 71, 1 unknown). The primary selection criterion was diagnosis of myocardial ischemia. Several other selection criteria were further established to obtain a representative selection of ECG abnormalities in the database. However, each record is supplemented by a compact clinical record, which contains technical information. Each record is two hours in duration and the signals are sampled at 250 samples per second with 12-bit resolution over a nominal 20 mV input range.

The MIT-BIH Arrhythmia Database [58]

In 1980 this database was designed to provide a first reference material set to evaluate arrhythmia detectors. This database contains 48 half-hour excerpts of two-channel ambulatory ECG recordings which were annotated independently by two or more cardiologists. The signals were obtained from 47 subjects and are divided into two groups: 23 recordings were chosen at random and the other 25 were selected to include uncommon but clinically significant arrhythmias that would not be well-represented in a small random sample (two recordings were obtained from one subject). Recordings were obtained from 25 men, aged 32 to 89, and 23 women, aged 23 to 89; 60 % of the subjects were inpatients. The recordings were digitized at 360 samples per second per channel with 11-bit resolution over a 11 mV range.

The QT Database [37]

Intention of this database was to test QT detection algorithms on a variety of QRS and ST-T morphologies. The recordings are mostly obtained from existing databases including the

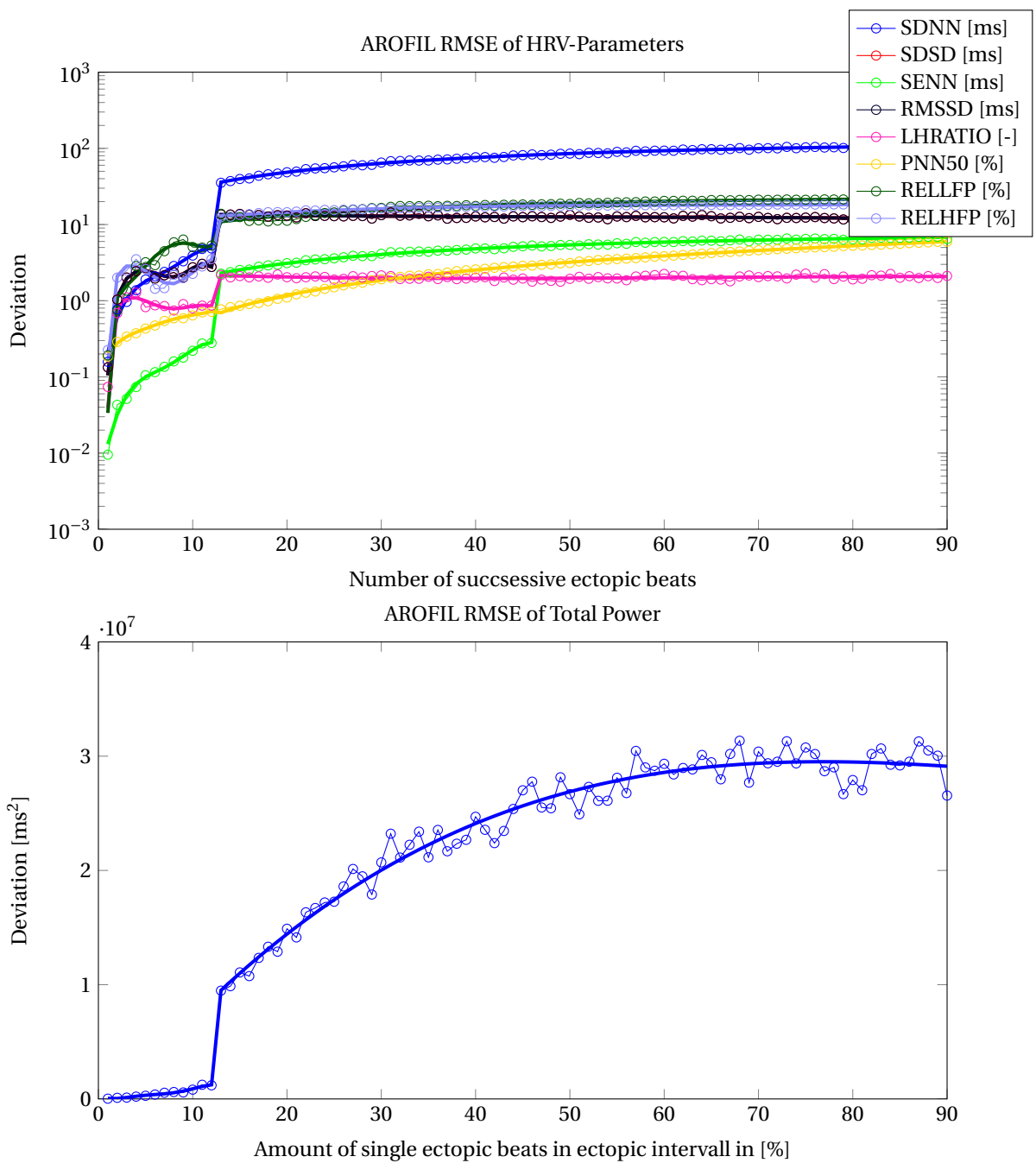


Figure 4.27: Determination of robustness by detection of the largest and steepest increase in RMSD of HRV parameters and RR intervals.

MIT-BIH Arrhythmia Database, the European Society of Cardiology ST-T Database, and several other ECG databases collected at Boston's Beth Israel Deaconess Medical Center. It contains 105 fifteen-minute excerpts of two-channel ECG recordings that were selected to avoid artefacts such as baseline drift. In contrast to the afore mentioned databases, only normal beats were annotated manually by cardiologists. All recordings were sampled at 250 samples per second. Those recordings that were not originally sampled with 250 Hz were converted by an application from the MIT Waveform Database Software Package.

The signals of these three databases were selected using Physionet tools to generate two data sets that were used in this diploma thesis, one with natural ectopic beats (*set X*) and one free of ectopic beat (*set N*). The first data-set contains 5 min parts that include at least one ectopic beat. The other data-set also contains 5 min parts, but only regions where no ectopic beats are present, since this set is used for artificial insertion of ectopic beats. Signals that do not show 5 min ectopic free regions were not included in this dataset. Usage of this data source offers the benefit of manually annotated ECG signals from experts. Moreover, the Physionet databases are widely used and thus comparisons with existing methods can be conducted easily since performance measures on this data source are available in many other studies.

The basic information about the two data sets are as follows:

Set N contains 151 recordings, obtained from 17 women, aged 24 to 84, 109 men, aged 30 to 84, and from 25 unknown subjects. All known subjects suffer from at least one of the following heart diseases: Myocardial infarction, coronary artery disease, coronary angiography, resting angina, effort angina, mixed angina or 1-,2-, or 3-vessel disease (see table 4.3). As already mentioned, only 5 min excerpts of ectopic free regions were used in this set.

Men/women	109/17
Age (years)	54.5 (11.7 SD)
Mixed angina	22 (17.5 %)
Resting angina	23 (18.3 %)
Effort angina	10 (7.9 %)
Coronary artery disease	43 (34.1 %)
Coronary angiography	27 (21.4 %)
Myocardial infarction	39 (31.0 %)
Hypertension	25 (19.8 %)
1-vessel disease	24 (19.0 %)
2-vessel disease	11 (8.7 %)
3-vessel disease	16 (12.7 %)

Table 4.3: Baseline clinical data of set N

Set X contains most of the signals from Set N (97 signals are identically). It contains 166 recordings, obtained from 34 women, aged 28 to 84, 110 men, aged 30 to 89, and from 22

unknown subjects. Cardiologists not only annotated the recordings of the subjects, but also described the PVC forms, such as couplets, triplets, bigeminy and trigeminy. 20 signals contained more than 50 % ectopic beats and were excluded from the test set.

4.4 Generation of Artificially Corrupted HRV signals

The generation of artificially ectopic beats was performed by two different algorithms. The signals used in Test-1 were created by simulation of natural ectopic beats, either single beats or successive ones. This beat generation is described in more detail in section 4.4.1 and 4.4.2. In contrast, the signals used in Test-2A just contained single ectopic beats, whereas those in Test-2B and 2C contained successive beats.

4.4.1 Single Ectopic Beats

Single ectopic beats may occur either with or without compensatory pause, depending on their focus of origin. A *premature ventricular contraction (PVC)* is an early beat, followed by a *compensatory pause* (see figure 4.28a and 4.28c) whereas a *premature atrial contraction (PAC)* is followed by a normal sinus beat without any delay (see figure 4.28b). If PVCs are only separated by one sinus beat they are called *bigeminy*, by two *trigeminy* and so on. A more detailed description of the physiology of ectopic beats is provided in section “2.6.1 Ectopic Beats”.

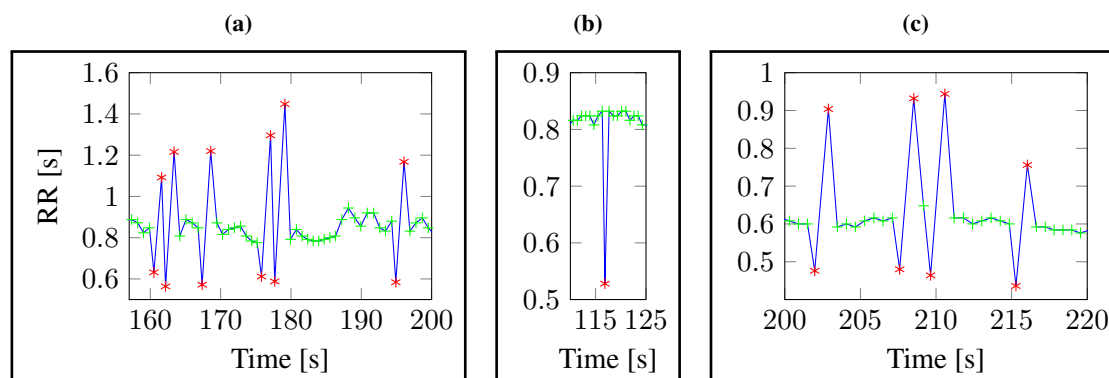


Figure 4.28: Cases of single ectopic beats: (a) Bigeminy (PVCs that are separated only by one sinus beat) and single PVCs; (b) Single PAC; (c) Trigeminy (PVCs that are separated by two sinus beats) and single PVCs. The RR interval time series is displayed solid in blue with normal RR intervals labeled as green plus signs and erroneous intervals with red stars.

4.4.2 Successive Ectopic Beats

Successive ectopic beats occur in different patterns and are mostly of ventricular origin. The two most common types of successive ectopic beats are *couplets* and *triplets* of PVCs. Couplets are

two successive ectopic beats that are followed by a compensatory pause and then again by two successive ectopic beats and so on (see figure 4.29a). Similar to couplets, triplets also occur in sequence, but here three ectopic beats are followed by a compensatory pause (see figure 4.29b).

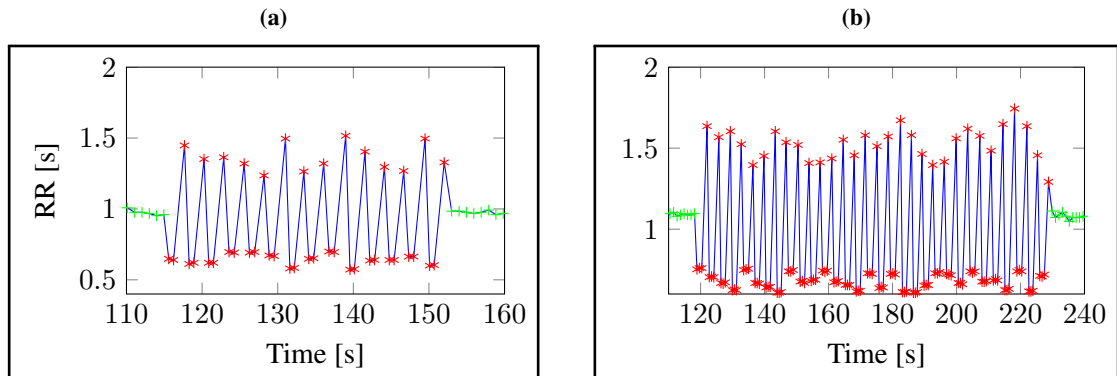


Figure 4.29: Most common types of successive ectopic beats: (a) Sequential PVC couplets: Sequence of two successive PVCs followed by a compensatory pause; (b) Sequential PVC triplets: Sequence of three successive PVCs followed by a compensatory pause. The RR interval time series is displayed solid in blue with normal RR intervals labeled as green plus signs and erroneous intervals with red stars.

If four or more ectopic beats occur in sequence without any compensatory pause between them, they are called *arrhythmia* (see section “2.6.1 Types of Arrhythmias” for more details). The origin of arrhythmia can either be inside the atrium or in the ventricle. *Ventricular arrhythmia* can further be divided into sustained (lasting for longer than 30 s) and non-sustained (shorter than 30 s) arrhythmia (see figure 4.30a and 4.30b).

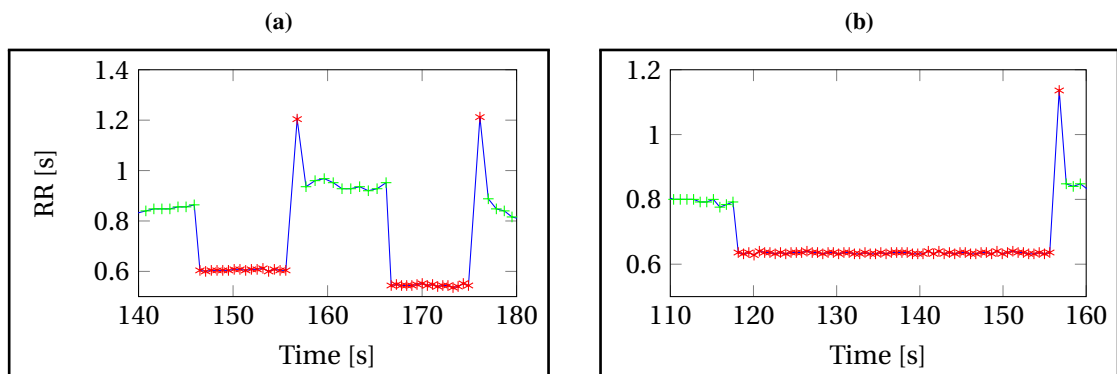


Figure 4.30: Cases of arrhythmia: (a) Non-sustained ventricular tachycardia and (b) sustained ventricular tachycardia. The RR interval time series is displayed solid in blue with normal RR intervals labeled as green plus signs and erroneous intervals with red stars.

4.5 Evaluation of Test Results

Test results are evaluated according to the specific test cases, as described in section 4.2. Test-1A and 1B are evaluated in the same procedure, based on the following six characteristics:

Computation Time, Peak Memory, Statistical Parameters, Time Parameters, Frequency Parameters and mean RMSD of all RR intervals in one data set. Peak memory and computation time are both obtained from the undocumented *profiler* function in MATLAB[®]. The peak memory is just a single value for all signals in one test case and can be used directly for the comparison of the correction approaches. In contrast, the computation time is actually the time of each method required for all signals in one test case, and thus has to be divided by the number of processed signals to get the mean computation time.

The changes in the RR interval time series after correction are reflected by the comparison of several HRV parameters with original ones. The deviations of the HRV parameters are described by their deviation and by means of the RMSD. Each test case consists of ten test sets, whereby each set contains 151 ECG signals. The RMSD of each parameter was calculated for all signals in one test set and for all signals in the entire test case (1510 signals).

$$RMSD = \sqrt{\frac{\sum_{i=1}^n (H\hat{R}V_i - HRV_i)^2}{n}}, \quad (4.45)$$

where $H\hat{R}V_i$ denotes the HRV parameter after correction and HRV_i the original (true) HRV parameter before corruption with ectopic beats.

Time- and frequency-domain HRV parameters are composed of the HRV parameters according to table 4.4 and 4.5 and are reduced to two single parameters by means of principal component analysis (PCA) (see section 4.5.4). Alternatively, each variable could be ranked before PCA is performed, which would allow the inclusion of all parameters in a single PCA. This approach is described by Korhonen et al. [35]. However, theoretical as well as practical issues arise: First, the differences between correction methods that attain similar results get separated in this approach and second, Peltola clearly states that the effects on time- and frequency domain HRV parameters of RR editing should be observed separately [65].

The statistical parameters also show many redundancies but cannot be reduced by means of PCA because of their relationship and distribution. Section 4.5.1 explains the used statistical parameters in detail, as well as useful combinations of them.

The RMSD of the RR intervals has to be adapted to compare signals with a different length (due to deletion or insertion of RR intervals). Section 4.5.2 describes the calculation of the adapted RMSD in detail.

Determination of the robustness in Test-2 results in the following three outcomes: Amount of single erroneous RR intervals that can be processed (Test-2A), number of successive ectopic beats (Test-2B) and necessary number of NN intervals between ectopic segments (Test-2C). The outcome of Test-2A describes the robustness against single ectopic beats, whereas Test-2B and Test-2C can be summarized to the robustness against successive ectopic beats, as described in section 4.5.5.

The results of Test-3 only contain the statistical parameters, computation time and peak memory. The HRV parameters, which would be present if no ectopic beat had occurred, are not available

for ECG signals containing natural ectopic beats. Therefore, this test does not contain information about the quality of the correction, but still enables a comparison of the other mentioned parameters to the corresponding ones of Test-1.

PCA time-domain HRV parameters				
RMSSD	SDSD	SENN	pNN50	SDNN

Table 4.4: HRV time-domain HRV parameters used for PCA

PCA frequency-domain HRV parameters			
HF ^{norm}	LF ^{norm}	LF/HF	Total Power

Table 4.5: HRV frequency-domain HRV parameters used for PCA

4.5.1 Statistical Parameters

Statistical Parameters are composed of sensitivity, specificity, accuracy, positive predictive value (PPV) and negative predictive value (NPV). These parameters describe the statistical behavior of the right correction of annotated RR intervals. Table 4.6 illustrates the relationship between these parameters and provides their calculation. Sensitivity determines the ability of each method to correct erroneous RR intervals. Specificity states the ability to leave correct NN intervals unchanged. The PPV determines the probability to only correct erroneous RR intervals but not correct NN intervals. The negative predictive value displays the relative amount of NN intervals that are correctly not changed. Accuracy reflects the combined results of either sensitivity and specificity or PPV and NPV.

		Condition		
		Positive	Negative	
Test outcome	Positive	TP	FP (type I error)	PPV = $\frac{\sum TP}{\sum (TP+FP)}$
	Negative	FN (type II error)	TN	NPV = $\frac{\sum TN}{\sum (TN+FN)}$
		Sensitivity $= \frac{\sum TP}{\sum (TP+FN)}$	Specificity $= \frac{\sum TN}{\sum (TN+FP)}$	Accuracy $= \frac{\sum (TN+TP)}{\sum (TN+TP+FP+FN)}$

Table 4.6: Contingency table of statistical parameters

4.5.2 Adapted Root Mean Squared Error

Statistical parameters only state if erroneous RR intervals are processed but not how well they were altered. In order to account for the time shift in the RR interval time series of unequally long signals an adapted form of the RMSD is introduced. For equally long signals this value is simply the standard RMSD:

$$RMSD = \sqrt{\frac{\sum_{i=1}^n (\hat{RR}_i - RR_i)^2}{n}}, \quad (4.46)$$

where \hat{RR}_i denotes the RR intervals after correction and RR_i the original (true) RR interval before corruption with ectopic beats.

This value is used for the evaluation of all signals in Test-1 and Test-2, but not in Test-3, since there is no knowledge about the original RR intervals. If RR intervals are either deleted or inserted, the standard RMSD cannot be used because of an unequal signal length. Hence, the adapted RMSD uses the last correct RR interval as reference for calculating the specific deviation of each following RR interval from this reference. This is best performed when using the difference matrix. The corrected RR intervals are assigned on the vertical and the original RR interval time series on the horizontal dimension (see figure 4.31). The original difference matrix without any performed correction is shown in figure 4.31a. The differences between all original and new RR intervals of the same index are thus displayed on the main diagonal (colored in figure 4.31a). If a correction is applied and the signal length is not changed, then the adapted RMSD is also determined using the diagonal of the difference array (see figure 4.31b, colored entries) and equals the standard RMSD. In case of an insertion or deletion of an RR interval there is a change in signal length, and thus also in the diagonal used for error calculations. Instead of using the entry on the original diagonal for determination of the RMSD, the entry on the new diagonal is used (see figure 4.31c and figure 4.31d, colored entries).

4.5.3 Box plots

Box plots are a statistical tool to visualize the distribution of cardinal scaled data [83]. Its goal is to give an impression where data points lie and how they are distributed. Therefore, box plots are well-suited to compare different methods and still consider the deviations inside one method. Box plots were created with a built-in MATLAB[®] function. The central mark is the median, the edges of the box are the 25th and 75th percentile, the whiskers mark the most extreme data points not considered as outliers and outliers are plotted individually. Hubert and Vandervieren have recognized that skewed distributions are not well reflected by standard box plots since the cutoff values for outliers are derived from the normal distribution [27]. They formulated a generalization of the box plot by inclusion of a robust measure of skewness in the determination of the whiskers [27]. Further, the adjusted box plot can be downloaded as part of the *LIBRA* library for MATLAB[®] from their homepage [39].

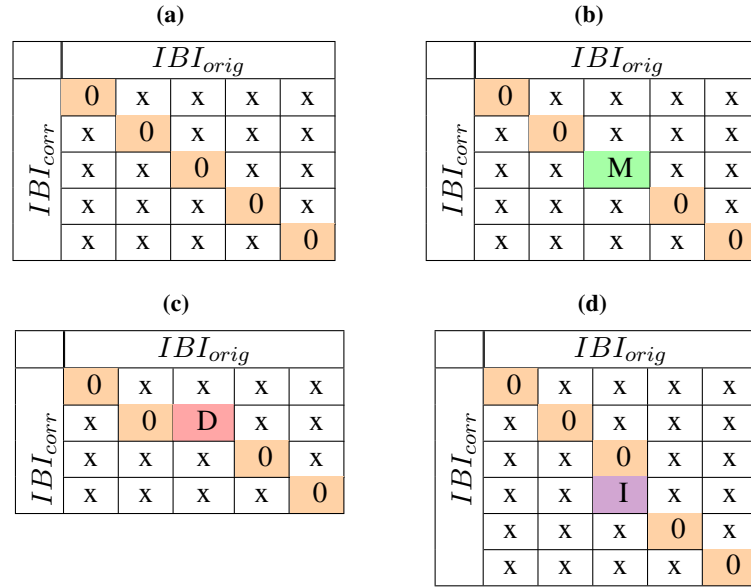


Figure 4.31: Determination of the adapted RMSD by means of the difference matrix: (a) If both IBI series are equally long and no correction is performed, the diagonal contains only zeros (marked in light orange, x denotes other, non-zero differences). (b) Both IBI series are equally long but the third IBI is changed and marked in light green (note that the diagonal is not changed). (c) The third IBI of the original time series is deleted during the correction (marked with light red) and thus, the new IBI series is shorter, resulting in a time shift (diagonal changes). (d) After the third IBI one additional IBI is inserted (marked in light violet). Therefore, the new time series contains one additional IBI and thus induces a time shift and changing of the diagonal.

4.5.4 Principal Component Analysis (PCA)

The huge amount of variables describing the correction ability of each method makes it necessary to perform a variable reduction. Basically, variable reduction may be either carried out by means of *principal component analysis (PCA)*, *factor analysis (FA)* or *independent component analysis (ICA)*. Suhr describes the differences between PCA and FA and states that FA is used to determine underlying constructs by means of factors, whereas PCA reduces the number of variables as long as they are correlated [77]. Jolliffe provides a detailed description of PCA and further states that ICA is preferably used to non-Gaussian data. ICA is more similar to FA than to PCA and can be seen as a generalization of PCA for non-normal data [30]. Nevertheless, PCA does not require any specific assumptions about the data, such as a Gaussian distribution [30] and is used in the context of this thesis for variable reduction.

The following brief description of PCA is based on the book “Principal Component Analysis” by Jolliffe [30] and the recent review of Nikolov [60]. PCA is a statistical method that uses an orthogonal transformation of observations of correlated

variables into a new set of linearly uncorrelated variables called *principal components (PC)*. The transformation is performed in a way that the first principal component (PC) contains the largest possible variance and each following PC covers the highest possible variance under the constraint that it is uncorrelated with the previous ones. However, the components are only guaranteed to be independent if the original data is jointly normal distributed. PCA calculates loadings that can be seen as weights of each variable for each PC. The scores are the coordinates of the data in the new variable space. Visualization of the new data is performed in so called score-score plots that show the new coordinates of the first two principal components. PCA can be performed by Eigenvalue decomposition of either the data covariance or correlation matrix. Before performing PCA, the variables of the data matrix have to be mean centered to ensure that the first PC contains the largest variance. Further, the variables may also be standardized (division of each variable by its standard deviation), if the variances in the variables differ by more than one order of magnitude. Mean centering can be accomplished by performing the PCA on the correlation matrix instead of the covariance matrix.

In this thesis PCA was performed using the correlation matrix. This offers the possibility to use the Kendall rank correlation instead of the Pearson correlation that assumes a linear relationship and requires normal distributed data [1]. Before the correlation matrix was computed, the logarithm was taken from the data, since the data showed a highly skewed distribution. According to Jolliffe, logarithmic scaling improves the usage of non-normal data [30] and comparison of PCA with non-logarithmically scaled data clearly showed a better distinction of the data points in the score-score plots, reflecting the new coordinates [30] (see figures 5.2, 5.3, 5.11 and 5.12 for the comparison of original data and data where the logarithm was taken).

Since the variables used in the data matrix describe similar properties (RMSD of HRV parameters), all coefficients of the first PC are positive. Jolliffe mentions that if all coefficients of the first PC are positive, this is a measure of size [30]. Thus, the first PC may be used to describe most of the other features of the original variables as long as the first Eigenvalue is much larger than the successive ones. The determination of the necessary number of PCs can be done in several ways. A very simple and effective technique is to plot the Eigenvalues with respect to their number to visualize the covered variance [30].

The correlations between the variables of the time and frequency domain were calculated by means of the Kendall's rank correlation. High correlations (over 0.9) between the different HRV variables can be seen in the time-domain HRV parameters of Test-1A and 1B (see tables 4.7 and 4.9, marked in red), but not in the frequency-domain parameters. The correlations clearly illustrate the possibility of variable reduction by means of PCA. Although the correlations for the frequency-domain HRV parameters in table 4.8 and 4.10 are low, compared to the time-domain correlations, the PCA still resulted in a good outcome, as illustrated by the large decrease in the Eigenvalue diagram after the first PC (see figure 5.1 for Test-1A and 5.10 for Test-1B).

4.5.5 Determination of Robustness

Robustness is addressed by three test cases, which all result in one parameter each: The amount of single PVCs that can be processed, the number of successive PVCs that can be processed (denoted as ectopic segment length) and the number of correct NN intervals necessary between ectopic segments. These parameters have to be summarized for a better comparison behavior.

	RMSSD	SDNN	SDSD	SENN	pNN50
RMSSD	1				
SDNN	0.63	1			
SDSD	1.00	0.63	1		
SENN	0.62	0.95	0.62	1	
pNN50	0.52	0.53	0.52	0.52	1

Table 4.7: Kendall's rank correlation of time-domain HRV parameters for Test-1A

	Total Power	LF/HF	LF ^{norm}	HF ^{norm}
Total Power	1			
LF/HF	0.58	1		
LF ^{norm}	0.55	0.54	1	
HF ^{norm}	0.64	0.64	0.62	1

Table 4.8: Kendall's rank correlation of frequency-domain HRV parameters for Test-1A

	RMSSD	SDNN	SDSD	SENN	pNN50
RMSSD	1				
SDNN	0.67	1			
SDSD	1.00	0.67	1		
SENN	0.62	0.93	0.62	1	
pNN50	0.63	0.50	0.63	0.47	1

Table 4.9: Kendall's rank correlation of time-domain HRV parameters for Test-1B

	Total Power	LF/HF	LF ^{norm}	HF ^{norm}
Total Power	1			
LF/HF	0.50	1		
LF ^{norm}	0.56	0.49	1	
HF ^{norm}	0.62	0.52	0.66	1

Table 4.10: Kendall's rank correlation of frequency-domain HRV parameters for Test-1B

First, a PCA was performed, but resulted in an insufficient outcome. Table 4.11 illustrates that the rank correlation of the three robustness parameters is extremely low. This is related to the high variations in nearly all methods in the specific variables.

	Multiple EB	NN Space	Single EB
Multiple EB	1		
NN Space	0.28	1	
Single EB	-0.24	0.01	1

Table 4.11: Kendall's rank correlation of robustness parameters

In a new approach, the robustness is divided into two cases, one for single ectopic beats and one for successive ones. Hence, the outcome of Test-2A, the amount of single PVCs that can be processed, is used as the robustness against single ectopic beats. The robustness against ectopic segments is calculated as the fraction of successive PVCs that can be processed in relation to the required number of correct NN intervals between those segments.

$$Rob_{seg} = \frac{Multiple_{EB}}{NN_{Space}} \quad (4.47)$$

This new variable takes the fact into count that the robustness against ectopic segments is only high if there are few NN intervals needed for correction.

4.5.6 Kivat-Diagrams

The final results are summarized in Kivat diagrams, also known as spider or net diagrams. They offer the ability to compare the results of multiple methods in several dimensions in one plot. The used variables are on different scales and have to be processed before proper illustration can be performed in a Kivat diagram.

The statistical parameters cannot be summarized by means of PCA due to their relations and data distributions. Nevertheless, they are highly correlated and show some redundancies. Visual comparison displayed that the usage of sensitivity and PPV is enough to cover the overall performance, since the other statistical parameters do not contain additional information. These two parameters are directly plotted on two axis of the Kivat diagram. All other variables were scaled to the range from 0 to 100, where 0 denotes the worst possible outcome and 100 the best one in all correction approaches:

$$V_{scaled}^i = \frac{max(V_{orig}) - V_{orig}^i}{max(max(V_{orig}) - V_{orig}^i)} \cdot 100, \quad V_{orig}^i \in V_{orig}, \quad (4.48)$$

where V_{orig} is the quantity of all scores.

The used variables are mean computation time, peak memory, RMSD of RR intervals, scores of PC 1 of HRV time-domain parameters and scores of PC 1 of HRV frequency-domain parameters. In case of Test-3 (naturally occurred ectopic beats), only the first two variables were used.

CHAPTER 5

Results

The results are described according to the test cases mentioned in section “4.2 Test Cases”. In short, Test-1A describes the behaviour of the correction ability of single ectopic beats, Test-1B the correction performance of multiple ectopic beats. Test-2 displays the robustness of all correction approaches with respect to the maximum number of single PVCs that can be processed, the maximal number of PVCs in one ectopic segment and the necessary amount of NN intervals between these ectopic segments. Test-3 displays the correction behavior of natural ectopic beats.

At the beginning of section 5.1 and 5.2, the results of the PCAs of the time- and frequency-domain HRV parameters are shown in score-score plots of the first two principal components (PC). The scores of the first PC, the RMSD of RR intervals, computation time and peak memory are illustrated in box plots. Box plots of the absolute deviation of the individual HRV parameters can be seen in appendix B. Sensitivity, PPV and the RMSD of the HRV parameters of all signals of the same test are illustrated in one table. At the end of both sections, the outcome of each test case is summarized in four Kivat diagrams (for a description see section “4.5.6 Kivat-Diagrams”). This segmentation is necessary to be able to distinguish the different methods and is performed according to the classification of ectopic beat correction methods as displayed in figure 4.1. Additionally, the best approaches of each class (determined by the highest mean score of all variables) is displayed in the fourth Kivat diagram. The variables used in the Kivat diagrams are abbreviated as described in table 5.1.

A table illustrates the reached scores of all correction approaches for Test-1A and -1B, starting with the best performing method on top. This tables are intended to give a quick overview of the performance of all correction approaches for each test specifically. At the end of these two sections, three tables show the distributions of the deviations of all HRV parameters by means of the median, the 2.5 and 97.5 percentile, the confidence interval and the Kendall’s correlation coefficient, compared to the original HRV parameters. The methods are clustered in the same way as the Kivat diagrams, according to figure 4.1.

Abbreviation	Description
Time	First PC of time-domain HRV parameters
Freq	First PC of frequency-domain HRV parameters
Mem	Peak Memory
CompT	Computation Time
RMSD	RMSD of RR intervals

Table 5.1: Abbreviations of variables used in the Kivat diagrams

Test-2 illustrates the robustness against single ectopic beats and ectopic segments in two boxplots. Thereby, the robustness against ectopic segments is the summarized result of Test-2B and -2C (see section 5.3). The individual results can be seen in section B.3.

Test-3 contains only information about sensitivity, PPV, computation time and peak memory of only one test set. Hence, the results are displayed in a 1-D point plot, instead of a box plot. The Kivat diagrams and the table illustrating the ranking of the performance of all correction approaches are the same as described previously.

5.1 Test-1A: Correction of Single Ectopic Beats

This section describes the outcome of the correction approaches to restore HRV signals disrupted with single PVCs and PACs at a ratio of 1:1. The detailed results of the absolute deviations of each time- and frequency-domain HRV parameter are illustrated in appendix B.1.1 and B.1.2. In the current paragraph, only the results of the PCA are described. Figure 5.1 shows the decrease of the Eigenvalue with respect to the number of PCs. Obviously, there is no appreciable increase in information after the first Eigenvalue. Thus, in time- and frequency-domain, only the first PC is considered for the comparison of the correction ability.

Figure 5.2 illustrates the score-score plot of the first and second PC after PCA on Kendall's rank correlation matrix of the frequency-domain HRV parameters. Figure 5.2a shows the score-score plot of the data after taking the logarithm, whereas figure 5.2b displays the same plot on the original data. Taking the logarithm of the original data obviously enhances the separation of the clusters, especially in the lower region. Nevertheless, the original data offers the possibility to use the *Zero Error* point as a reference for the performance of the correction approaches. Both plots clearly demonstrate that all approaches perform much better than no correction. The results of the PCA are described by means of figure 5.2a, since distinctions can be better made in the logarithmic scaled data. Low scores of PC 1 refer to a good correction ability, high scores to a bad one. The best correction methods are kNN-AF, INT-00, AROFIL and MEDFIL. As the variance is rather high in these clusters, no precise ranking can be done. Further, INT-10, IPFM-S, AMFIL and DELETE are also located very near to the former mentioned approaches. In contrast, approaches that yielded high scores in the first PC, like THRESH, WAVE-F, BUFFER, GP-IIA, IPFM-C, SWAFIL and IRFIL are clearly apart from all other methods and thus, perform worse. Additionally, these methods form more distinct clusters and

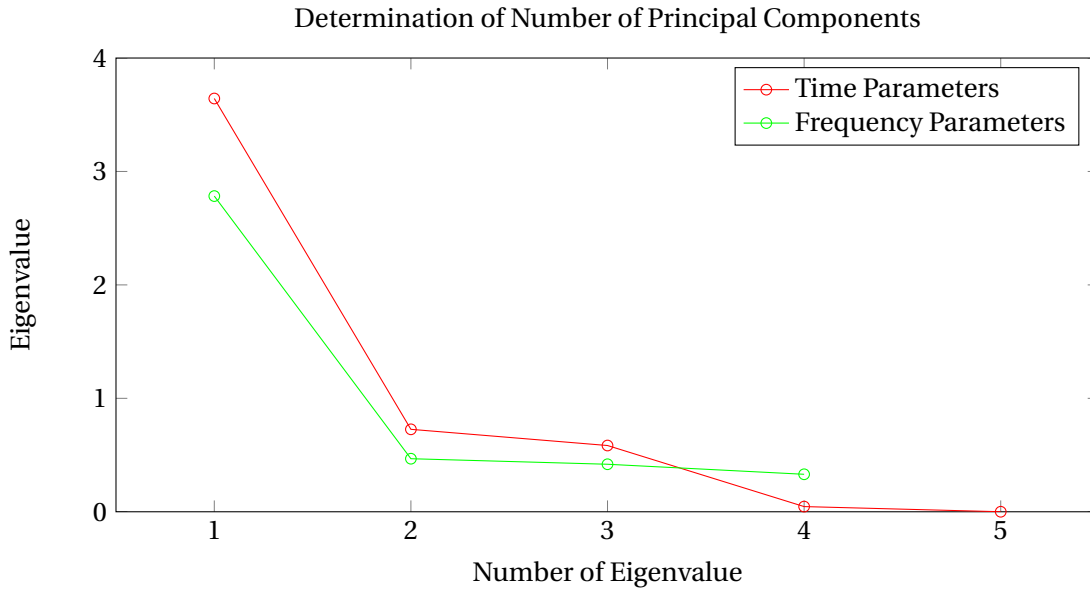


Figure 5.1: Eigenvalue plot of Test-1A: Time- and frequency-domain HRV parameters show a flattening decrease after the second eigenvalue. Hence, the first PC is sufficient for the representation of HRV parameters in both cases.

can be better separated from other approaches in the score-score plot. Figure 5.2b illustrates that well performing methods are located rather near to the *Zero Error* point compared to worse performing ones. The detailed results of each frequency-domain HRV variable are illustrated in appendix B.1.2.

Figure 5.3 illustrates the score-score plot of the first and second PC after PCA on Kendall's rank correlation matrix of the time-domain HRV parameters in the same manner as described by means of the frequency-domain parameters. Basically, the outcome is very similar to the frequency-domain. All approaches perform much better than no correction.

Again, the best correction methods are MEDFIL, INT-00, INT-10 and AROFIL. CSINTP, DELETE, kNN-AF, AMFILT and IPFM model with s-parameter (IPFM-S) are also located very near to the best approaches. Algorithms that yielded high scores in the first PC, like THRESH, SWAFIL, IPFM-C, GP-IIA, BUFFER, WAVE-F, PPHDIG and IRFILT are clearly separated from all other methods. Figure 5.3b illustrates that well performing methods are located rather near to the *Zero Error* point, with respect to the deviations from worse performing approaches. Detailed results of each time-domain HRV variable are illustrated in appendix B.1.1.

The box plots in figure 5.4 and 5.5 show the scores of the first principal component, based on the root mean squared deviation of the HRV parameters. The results are the same as described in the score-score plots previously. However, box plots are advantageous to visualize the variances in each method and to compare the median performance of them, but do not take the second PC into account. Both box plots demonstrate that most correction approaches perform

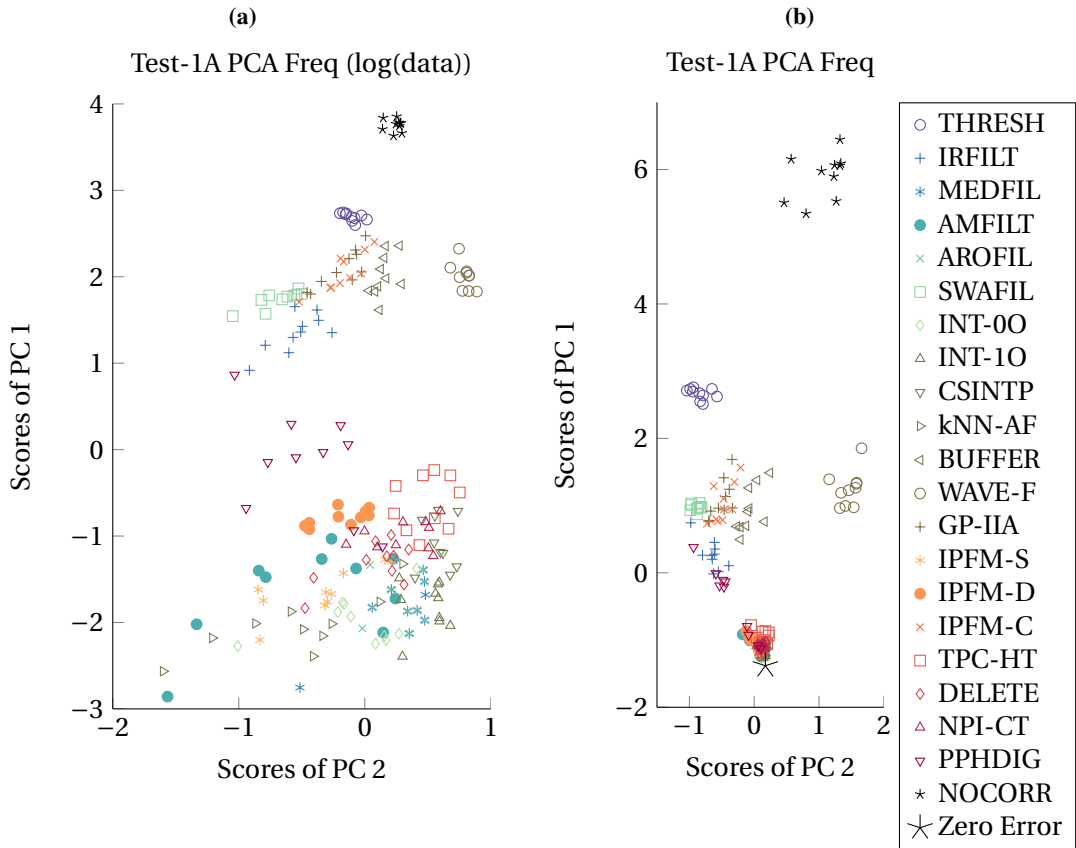


Figure 5.2: Test-1A score-score plots of frequency-domain HRV parameters: Low scores of PC 1 refer to a good correction behavior, high scores to a bad one (all variables have positive loadings for this component). In the case of PC 2 there is no linear relationship to the original data because of positive and negative loadings. (a) The logarithm was taken of the data before PCA was applied, resulting in a better separation of the clusters, especially in the case of lower scores. (b) Plot of original data and addition of the “Zero Error” point. The legend is valid for both diagrams.

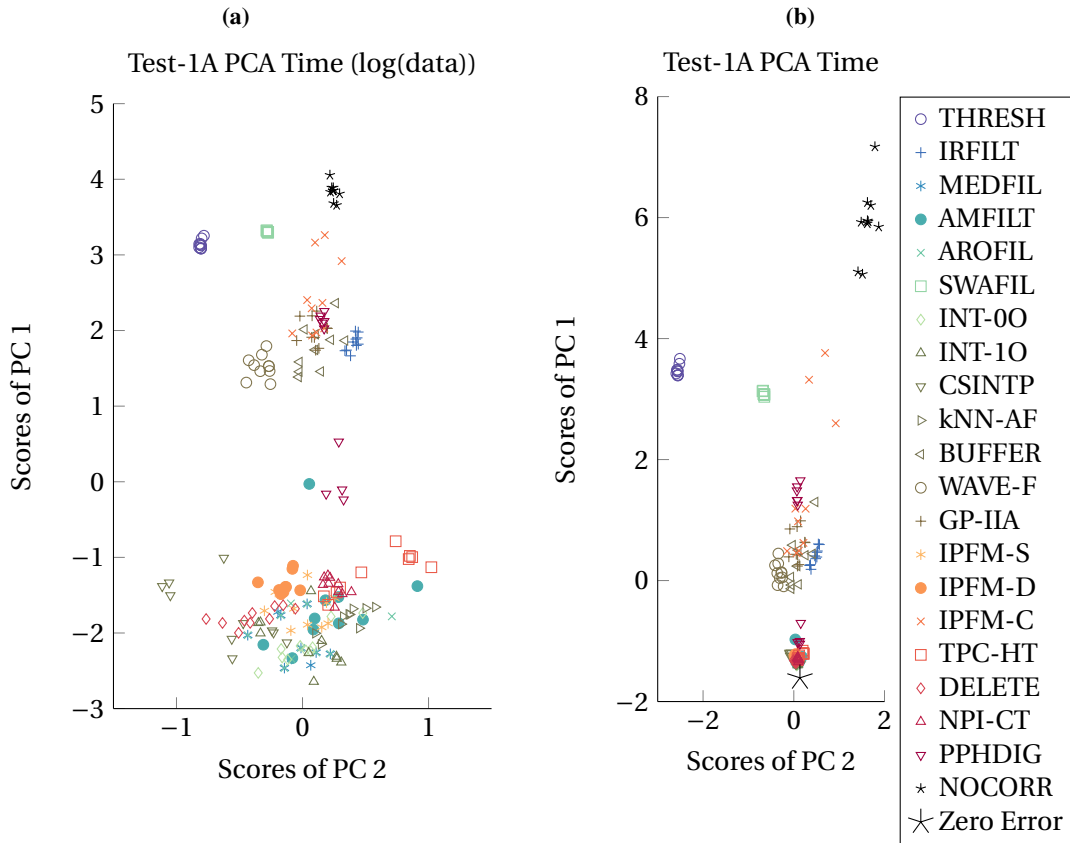


Figure 5.3: Test-1A score-score plots of time-domain HRV parameters: Low scores of PC 1 refer to a good correction behavior, high scores to a bad one (all variables have positive loadings for this component). In the case of PC 2 there is no linear relationship to the original data because of positive and negative loadings. (a) The logarithm was taken of the data before PCA was applied, resulting in a better separation of the clusters, especially in the case of lower scores. (b) Plot of original data and addition of the “Zero Error” point. The legend is valid for both diagrams.

very similar, except of the three methods that rely on self-detection of erroneous RR intervals (THRESH, IRFILT and SWAFIL) and models that cannot correct single ectopic beats without a compensatory pause reliably (see section 6.6 for more details).

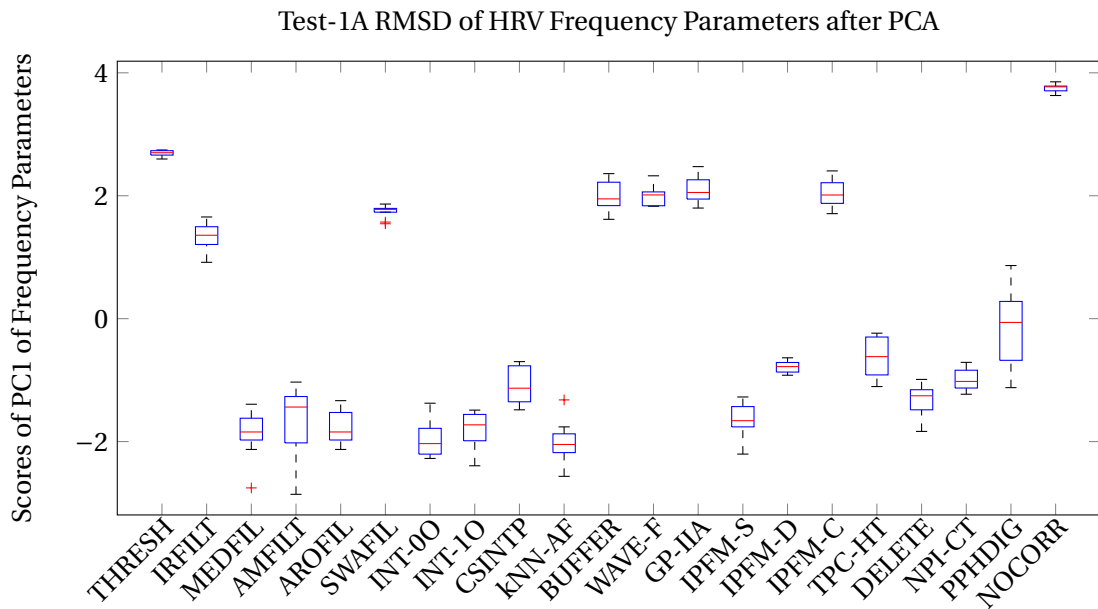


Figure 5.4: Test-1A box plot of HRV frequency-domain parameters. PCA was performed on the RMSD of the HRV parameters mentioned in table 4.8 for ten test sets containing 151 signals each. Hence, the box plot displays the scores of the first principal component with ten values for each method.

A comparison of the mean computation time clearly shows that eight methods perform very similar (figure 5.6 shows overlapping whiskers): All median filters (MEDFIL, AMFILT, AROFIL), IRFILT, THRESH, INT-00, kNN-AF and TPC-HT. Only DELETE obviously needs the shortest time, since it is the simplest correction ability. Methods that require a longer computation time are three physiologically motivated models (PPHDIG, IPFM-C and IPFM-S), sliding window average filter (SWAFIL), WAVE-F, CSINTP and NPI-CT.

The box plot of the peak memory shows that the memory is mostly constant in the iterative testing of the correction approaches (see figure 5.7). Exceptions are one to two outliers per correction approach. Not only complex approaches, such as models and adaptive filters (AMFILT and AROFIL), but also more simple interpolation methods (CSINTP, INT-10 and NPI-CT) have a higher peak memory. In contrast, both removal approaches (DELETE and THRESH), IRFILT, SWAFIL and INT-00 need the lowest memory.

The adjusted box plot in figure 5.8 illustrates the adapted root mean squared deviation (RMSD) of the RR intervals. Basically, the medians show the same relations as already de-

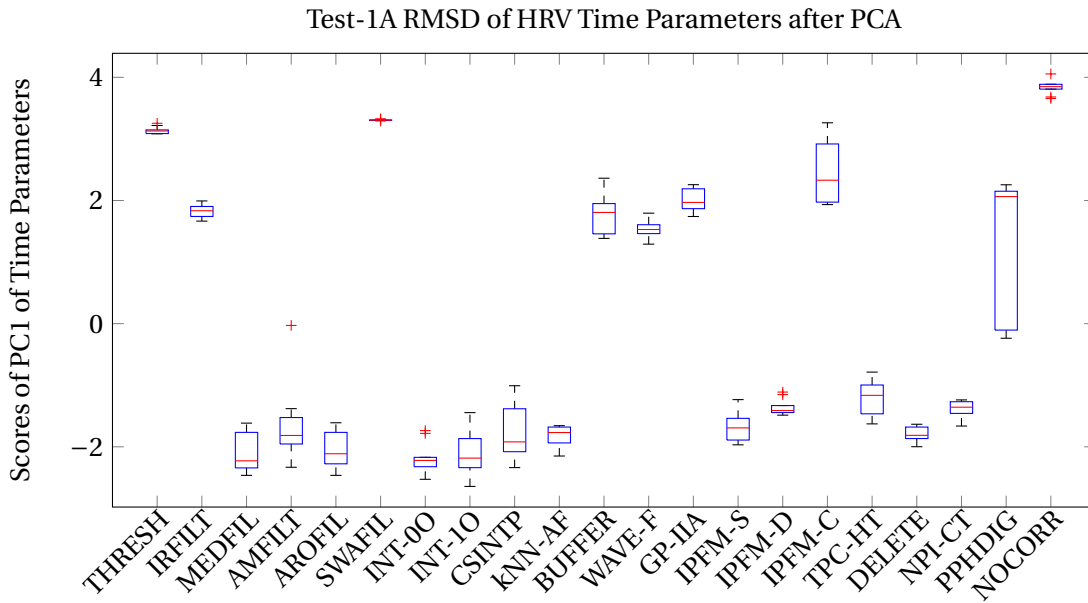


Figure 5.5: Test-1A box plot of HRV time-domain parameters. PCA was performed on the RMSD of the HRV parameters mentioned in table 4.7 for ten test sets containing 151 signals each. Hence, the box plot displays the scores of the first principal component with ten values for each method.

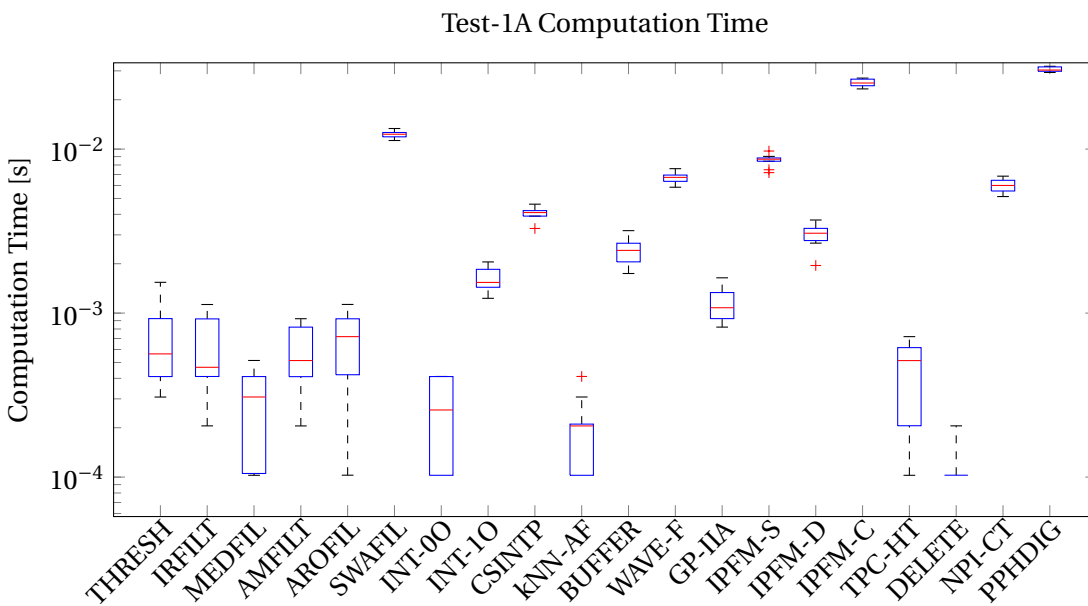


Figure 5.6: Test-1A box plot of mean computation time. Mean computation time was calculated for each of the ten test sets by dividing the total time for one test set by the 151 signals processed. Hence, the box plot displays ten values for each method.

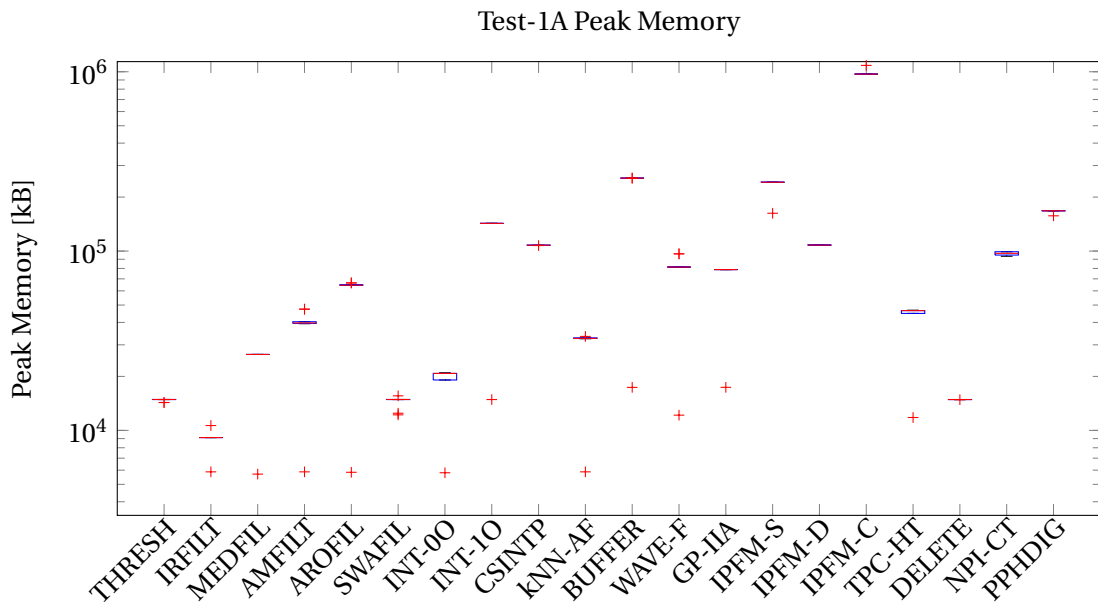


Figure 5.7: Test-1A box plot of peak memory. Peak memory was determined by means of the built-in “profiler” in MATLAB[®] for each of the ten test sets separately. Hence, the box plot displays ten values for each method.

scribed in the time- and frequency-domain HRV parameters. Methods that rely on self detection of erroneous RR intervals and models that cannot correct single ectopic beats without compensatory pause reliably perform worse than most other approaches. PPHDIG shows a rather low median RMSD, but several outliers towards higher values, which may describe the worse performance of this method in time-domain parameters, compared to frequency ones.

The sensitivity of all correction approaches is very high (between 97 and 100 %), except for IRFILT (76.4 %), THRESH (86.3 %) and GP-IIA (91.0 %), see table 5.2). IRFILT, THRESH and SWAFIL have to detect erroneous RR intervals on their own, whereas all other methods use already annotated data, resulting in a higher sensitivity.

The PPV is much lower than the sensitivity in most cases (see table 5.2), except for the median filters (MEDFIL, AMFIL and AROFIL), all kinds of interpolation (INT-00, INT-10, CSINTP and NPI-CT), DELETE and TPC-HT. These approaches can only fail to correct erroneous intervals, resulting in a loss of sensitivity. They can never falsely correct NN intervals, resulting in a PPV of 100 % in all test cases. Again, THRESH, IRFILT and SWAFIL show very low values, since they need to detect erroneous RR intervals on their own. The extremely low PPV of WAVE-F is related to the processing of erroneous RR intervals in the frequency domain. Changing of wavelet coefficients always results in alteration of multiple neighboring NN intervals. All models that can only correct ectopic beats with compensatory pause (GP-IIA, IPFM-S, IPFM-D, IPFM-C and BUFFER) reach a PPV below 75 %, since they always alter two RR intervals (shifting of beats), even if only one is erroneous.

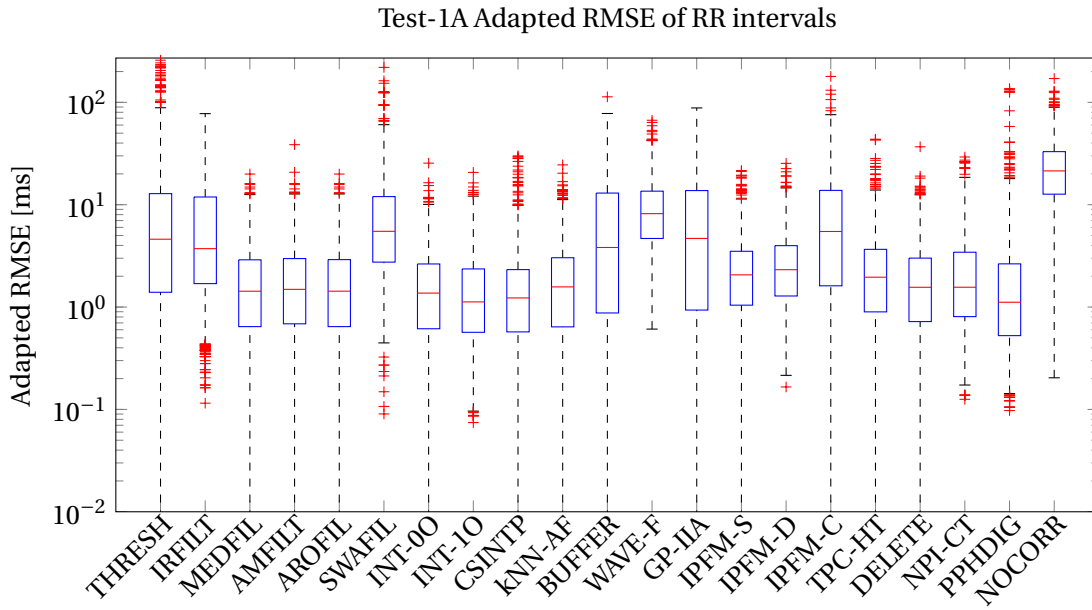


Figure 5.8: Test-1A adjusted box plot of adapted RMSD of RR intervals. The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1A cases, containing 151 signals each. Note: Whiskers that extend below 10^{-2} are zero and thus cannot be displayed on the logarithmic scale.

The RMSD of most HRV parameters is reduced remarkably, compared to no correction (see table 5.2). A comparison to the original variables indicates that the individual HRV parameters are affected in a different way by inclusion of ectopic beats. For example, in the time domain, *RMSSD* and *SDSD* are much more altered than *pNN50* and *SENN*. In the time domain, *HF* is more changed than *LF*.

RMSD between original and corrected RR interval time series of HRV parameters of all 1510 signals in Test-1A											
	Sens [%]	PPV [%]	RMSSD [ms]	SDSD [ms]	SDNN [ms]	SENN [ms]	pNNS0 [%]	LF/HF [-]	LF ^{norm} [nu.]	HF ^{norm} [nu.]	TP [s ²]
THRESH	86.03	11.91	7.42	7.44	17.34	0.87	1.25	1.28	5.70	3.43	3.29
IRFILT	76.42	43.31	7.53	7.55	2.53	0.15	0.85	1.32	1.27	2.26	0.41
MEDFIL	99.87	100.00	0.25	0.25	0.24	0.01	0.28	0.09	0.30	0.37	0.03
AMFIL	99.00	100.00	0.68	0.68	0.42	0.03	0.28	0.27	0.36	0.58	0.05
AKFIL	99.82	100.00	0.31	0.31	0.24	0.01	0.28	0.11	0.31	0.38	0.03
SWAFIL	96.84	29.86	10.45	10.47	9.09	0.42	2.65	0.64	1.86	3.71	2.02
INT-00	99.95	100.00	0.28	0.28	0.21	0.01	0.26	0.12	0.32	0.35	0.03
INT-10	100.00	100.00	0.26	0.26	0.23	0.01	0.31	0.09	0.28	0.37	0.03
CSNTP	100.00	100.00	0.23	0.23	0.62	0.04	0.26	0.10	0.45	0.65	0.08
KNN-AF	99.95	100.00	0.47	0.47	0.22	0.01	0.25	0.14	0.35	0.40	0.03
BUFFER	97.63	70.12	6.54	6.56	3.27	0.21	0.71	1.35	3.27	3.83	0.43
WAVE-F	100.00	8.49	2.72	2.72	2.59	0.17	1.33	0.95	4.60	3.56	0.32
GP-1A	91.07	67.26	7.04	7.05	3.86	0.25	0.74	1.50	2.71	4.77	0.52
IPFM-S	99.95	66.98	0.36	0.36	0.29	0.02	0.34	0.16	0.35	0.41	0.05
IPFM-D	99.97	65.66	0.40	0.40	0.42	0.02	0.34	0.36	0.44	0.44	0.07
IPFM-C	98.55	73.71	12.34	12.36	6.14	0.37	0.84	1.47	2.97	4.74	0.48
TPC-HT	99.89	100.00	1.03	1.04	0.29	0.02	0.30	0.30	0.61	0.77	0.05
DELETE	100.00	100.00	0.28	0.28	0.37	0.03	0.22	0.15	0.61	0.46	0.05
NPI-CT	99.95	100.00	0.58	0.58	0.33	0.02	0.27	0.16	0.53	0.72	0.05
PPHDIG	97.87	93.25	8.68	8.70	5.24	0.33	0.30	0.30	0.63	0.65	1.33
NOCORR	-	-	32.58	32.64	12.11	0.77	1.29	2.73	6.47	13.83	1.76
ORIGINAL			25.36 (7.06 95.32)	25.40 (7.07 95.49)	45.84 (11.13 130.30)	2.58 (0.49 7.86)	4.48 (0.00 60.59)	2.91 (0.15 9.96)	23.89 (3.52 56.61)	8.70 (1.00 57.98)	2.01 (0.18 16.51)

Table 5.2: Sensitivity and PPV are shown in the first two columns (the higher the better). The right part of the table shows the RMSD of the HRV parameters (the smaller the better) of all 1510 signals in Test-1A. The RMSD is calculated with respect to the original HRV parameters. These are illustrated in the last row by means of the median, the 2.5 percentile and the 97.5 percentile.

Figure 5.9 summarizes the results of Test-1A in four Kivat diagrams, whereby three show the results of each class, as described in section 4.1 and illustrated in figure 4.1. The fourth Kivat diagram compares the best representation of each class. As already mentioned in the outcome of the PCA of the HRV time- and frequency-domain parameters, distinctions in the performance of the correction methods can be made only partly. For example, all of the “Interpolation and Removal” approaches, except the THRESH, perform very similar. In the “Filtering” class, all median filters (MEDFIL, AMFIL and AROFIL) and kNN-AF perform best and reach similar scores as interpolation techniques. IRFILT, SWAFIL and WAVE-F show a clearly worse operating level. Comparison of the different models for ectopic beat correction shows that TPC-HT achieves the highest scores, except in the frequency-domain, where IPFM-S performs better. The IPFM-D approach performs nearly identical to the TPC-HT method in terms of the HRV parameters and the RMSD of the RR intervals. All other models accomplish much lower scores, especially in the case of the error of time- and frequency-domain HRV parameters. The comparison of the best methods points out that MEDFIL achieves similar results compared to INT-00, in all variables. Although TPC-HT performs worse in the error of the HRV parameters, it attains nearly the same scores in all other variables compared to the other two approaches.

Table 5.3 summarizes the findings of Test-1A by means of the achieved scores. The table is intended to give a quick overview of the mean performance of all correction approaches and is ordered in decreasing scores. Hence, methods on top of the table are highly recommended for correction of single ectopic beats whereas those on the bottom should be avoided in this test case.

The detailed deviations of all used frequency- and time-domain HRV parameters are summarized in three tables, according to the three clusters, as represented in the Kivat diagrams (see figure 5.9). The tables illustrate the median deviation of each HRV parameter with respect to the original ones. Further, the 2.5 and 97.5 percentile is given as a measure of deviation. The confidence interval and Kendall’s rank correlation coefficient compare the distributions of the new HRV parameters to the original ones.

Table 5.4 illustrates the deviations of the HRV parameters and the RMSD of the RR intervals of all removal and interpolation approaches. Most correlation coefficients are above 0.9 for all HRV parameters. Thus, all HRV parameters are very reasonably restored by all correction approaches, except of THRESH, which performs a little weaker in most parameters.

Table 5.5 demonstrates the deviations of the HRV parameters and the RMSD of the RR intervals for all filter methods. Similar to the interpolation and removal class, most correlation coefficients are above 0.9 for all HRV parameters. All HRV parameters are also very reasonably restored by all filter approaches. The median filters and the kNN-AF perform better than the IRFILT, SWAFIL and WAVE-F, since the first two approaches rely on self detection of erroneous RR intervals and the wavelet filter changes several adjacent correct NN intervals.

The correction performance of the BUFFER, GP-IIA and IPFM-C is weaker than the performance of the other models (see table 5.6). Additionally, the deviations of the HRV parameters and the RMSD of the RR intervals in these methods are higher than in the other approaches. Again, most correlation coefficients are above 0.9 for all HRV parameters. All HRV parameters are also very reasonable restored by IPFM-S, IPFM-D, TPC-HT and PPHDIG.

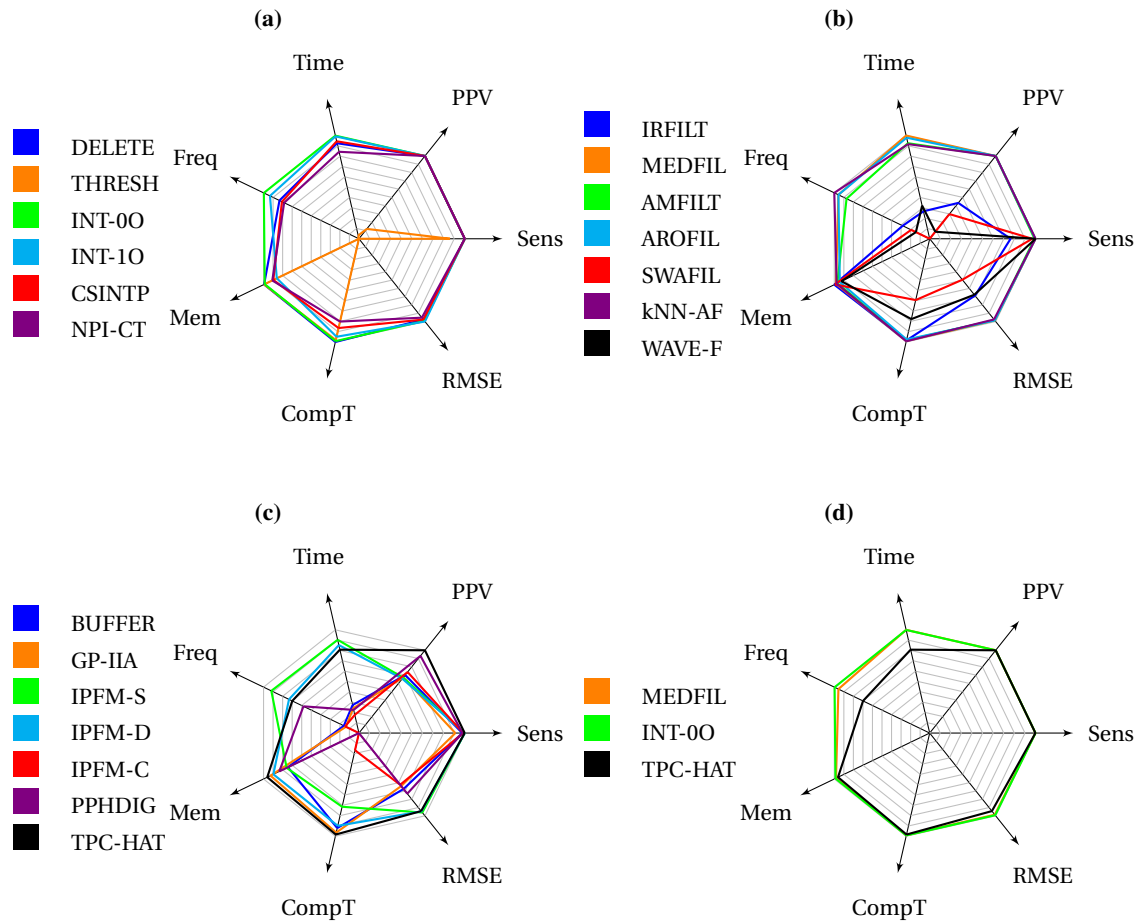


Figure 5.9: Test-1A Kivat Diagrams: Kivat diagrams are shown for the correction methods of each class separately, namely (a) “Interpolation and Removal”, (b) “Filtering” and (c) “Models”. The fourth diagram (d) compares the best methods of each class (determined by the highest mean scores of all variables) with each other. Abbreviations of the variables can be obtained from table 5.1.

	Sens	PPV	Time	Freq	Mem	CompT	RMSD
INT-00	99.95	100.00	99.89	99.69	98.79	99.15	99.26
MEDFIL	99.87	100.00	100.00	95.71	98.19	98.98	98.65
kNN-AF	99.95	100.00	91.70	100.00	97.56	99.32	97.47
AROFIL	99.82	100.00	97.93	95.71	94.24	97.63	98.65
INT-10	100.00	100.00	99.20	93.26	86.10	94.93	100.00
DELETE	100.00	100.00	92.53	83.31	99.40	100.00	97.61
AMFILT	99.00	100.00	92.55	87.16	96.82	98.31	98.28
CSINTP	100.00	100.00	94.44	80.70	89.73	86.48	97.61
TPC-HT	99.89	100.00	80.78	69.89	96.11	98.31	94.26
NPI-CT	99.95	100.00	84.26	78.38	90.84	80.22	95.33
IPFM-D	99.97	65.66	85.25	73.31	89.71	89.88	94.62
IPFM-S	99.95	66.98	90.33	91.85	75.77	71.40	96.01
GP-IIA	91.07	67.26	24.14	13.63	92.74	96.45	64.06
BUFFER	97.63	70.12	27.09	15.82	74.41	92.05	67.57
IRFILT	76.42	43.31	26.58	28.28	100.00	98.46	68.67
PPHDIG	97.87	93.25	22.43	58.16	83.50	0.00	73.25
WAVE-F	100.00	8.49	32.09	14.46	92.48	77.84	67.63
SWAFIL	96.84	29.86	0.00	19.44	99.40	59.40	49.50
THRESH	86.03	11.91	3.18	0.00	99.40	98.14	0.00
IPFM-C	98.55	73.71	17.62	14.49	0.00	16.65	62.46

Table 5.3: Summarized results of Test-1A: Correction methods are listed in decreasing order of their mean performance in all parameters. Abbreviations are: Sensitivity (Sens), positive predictive value (PPV), PC-1 of time-domain HRV parameters (Time), PC-1 of frequency-domain HRV parameters (Freq), Peak Memory (Mem), Computation Time (CompT) and RMSD of all RR intervals (RMSD). Sensitivity and PPV are given in % whereas all other parameters denote the score with respect to the best performing (100) and worst performing (0) approach (see section 4.5.6 for a more detailed description).

	DELETE	THRESH	INT400	INT10	CSINTP	NPICT	NOCORR
$RMSSD - RMSSD$	0.05 (-0.2110.60) [0.060.07]	0.02 (-1.337.85) [0.000.04]	-0.03 (-0.540.06) [-0.061.05]	-0.05 (-0.660.00) [-0.081.07]	-0.03 (-0.500.08) [-0.051.04]	0.02 (-0.291.43) [0.030.04]	15.63 (1.2382.96) [17.9519.94]
$R[RMSSD vs RMSSD]$	0.991	0.891	0.992	0.992	0.992	0.986	0.599
$SDNN - SDNN$	0.04 (-0.9210.49) [0.0410.05]	-0.06 (-57.731.33) [-3.21-2.41]	-0.01 (-0.5810.19) [-0.031.02]	-0.01 (-0.3710.26) [-0.011.01]	-0.00 (-0.2410.51) [-0.0010.01]	-0.00 (-0.7510.50) [-0.011.00]	4.40 (0.2832.97) [5.255.96]
$R[SDNN vs SDNN]$	0.992	0.861	0.994	0.995	0.994	0.992	0.836
$SDSD - SDSD$	0.05 (-0.2110.61) [0.0610.07]	0.02 (-1.337.88) [0.000.04]	-0.03 (-0.540.06) [-0.061.05]	-0.05 (-0.660.00) [-0.081.07]	-0.03 (-0.500.08) [-0.051.04]	0.02 (-0.291.44) [0.030.04]	15.66 (1.2383.12) [17.9819.97]
$R[SDSD vs SDSD]$	0.991	0.891	0.992	0.992	0.993	0.986	0.599
$SENN - SENN$	0.01 (-0.0410.08) [0.0110.01]	0.00 (-3.6910.12) [-0.111.07]	-0.00 (-0.0410.01) [-0.001.00]	-0.00 (-0.0210.01) [-0.001.00]	-0.00 (-0.0210.03) [-0.0010.00]	-0.00 (-0.0410.03) [-0.0010.00]	0.23 (0.0112.17) [0.2910.33]
$R[SENN vs SENN]$	0.993	0.904	0.994	0.994	0.993	0.994	0.846
$pNN50 - pNN50$	0.01 (-0.5010.47) [0.0110.02]	0.01 (-3.471.34) [0.0010.02]	0.00 (-0.7910.23) [0.0010.00]	0.00 (-0.9510.00) [0.0010.00]	0.00 (-0.7410.19) [0.0010.00]	0.00 (-0.5510.73) [0.0010.00]	0.93 (0.0012.83) [0.9911.05]
$R[pNN50 vs pNN50]$	0.993	0.911	0.995	0.995	0.995	0.987	0.912
$TP - TP$	0.00 (-0.1410.08) [0.0010.00]	-0.00 (-9.3410.09) [-0.391.25]	-0.00 (-0.1010.03) [-0.001.00]	-0.00 (-0.0510.04) [-0.001.00]	-0.00 (-0.0410.08) [-0.0010.00]	-0.00 (-0.1010.07) [-0.001.00]	0.53 (0.0314.97) [0.6410.74]
$R[TP vs TP]$	0.992	0.854	0.994	0.995	0.994	0.993	0.805
$LF^{norm} - LF^{norm}$	0.02 (-0.9911.31) [0.0110.04]	0.07 (-8.1818.09) [0.0610.19]	-0.00 (-0.5810.64) [-0.011.00]	0.01 (-0.3510.73) [0.0110.02]	0.01 (-0.3311.15) [0.0210.03]	-0.01 (-1.1210.62) [-0.021.01]	-1.91 (-18.1215.65) [-2.801.2.31]
$R[LF^{norm} vs LF^{norm}]$	0.978	0.809	0.989	0.989	0.985	0.984	0.765
$HF^{norm} - HF^{norm}$	-0.02 (-1.0810.84) [-0.061.03]	0.16 (-2.37110.36) [0.5910.85]	-0.03 (-0.8510.25) [-0.071.05]	-0.02 (-0.9310.17) [-0.061.04]	-0.00 (-1.0310.43) [-0.021.01]	0.00 (-0.6211.17) [0.0010.01]	5.13 (-2.24139.80) [6.3117.29]
$R[HF^{norm} vs HF^{norm}]$	0.982	0.831	0.986	0.988	0.986	0.980	0.625
$LF/HF - LF/HF$	0.01 (-0.2110.35) [0.0110.01]	-0.03 (-4.0310.61) [-0.201.13]	0.01 (-0.0810.26) [0.0110.01]	0.01 (-0.0710.23) [0.0110.01]	0.00 (-0.1510.17) [0.0010.00]	-0.00 (-0.3310.17) [-0.011.00]	-1.10 (-7.1410.06) [-1.601.40]
$R[LF/HF vs LF/HF]$	0.978	0.833	0.985	0.987	0.984	0.978	0.561
$RMSE of RR$	1.56 (0.0019.64)	4.60 (0.00139.71)	1.37 (0.1017.51)	1.12 (0.1116.82)	1.23 (0.1118.93)	1.56 (0.00110.85)	21.39 (4.09175.50)

Table 5.4: Deviations of HRV parameters from original, error-free, RR interval time series for removal and interpolation techniques in Test-1A. Deviations of HRV parameters are represented as follows: Median (2.5 percentile | 97.5 percentile) in the upper row and the confidence interval (computed by the difference between the population medians for matched pairs, according to [22]) in the lower row, for each HRV parameter. The correlation R was computed as the Kendall's rank correlation. The RMSD of the RR intervals is also represented by the median, the 2.5 and the 97.5 percentile. RMSSD, SDNN, SDSD and SENN are given in ms, pNN50 in %, LF^{norm} and HF^{norm} in n.u., TP in s^2 and LF/HF unit-less.

	IRFILT	MEDFIL	AMFILT	AROFIL	SWAFIL	kNN-AF	WAVE-F	NOCORR
$RMSSD - RMSSD$	-0.05 (-11.18 24.38) [-0.07 -0.02] 0.835	-0.02 (-0.48 0.35) [-0.03 -0.02] 0.991	-0.02 (-0.48 0.49) [-0.03 -0.02] 0.988	-0.02 (-0.48 0.36) [-0.03 -0.02] 0.990	0.18 (-29.07 2.56) [-0.24 0.08] 0.921	-0.01 (-0.40 1.15) [-0.02 -0.01] 0.989	0.62 (-0.97 7.11) [0.76 0.91] 0.929	15.63 (1.23 82.96) [17.95 19.94] 0.599
$\mathbf{R}[RMSSDvsRMSSD]$								
$SDNN - SDNN$	-0.04 (-7.11 3.81) [-0.14 -0.08] 0.963	-0.00 (-0.38 0.41) [-0.00 0.00] 0.995	-0.00 (-0.57 0.38) [-0.01 -0.00] 0.994	-0.00 (-0.38 0.41) [-0.00 0.00] 0.995	0.01 (-26.22 20.72) [-0.39 -0.18] 0.950	-0.00 (-0.44 0.38) [-0.01 -0.00] 0.994	0.43 (-0.90 7.70) [0.60 0.73] 0.956	4.40 (0.28 32.97) [5.25 5.96] 0.836
$\mathbf{R}[SDNNvsSDNN]$								
$SDSD - SDSD$	-0.05 (-11.20 24.42) [-0.07 -0.02] 0.835	-0.02 (-0.48 0.35) [-0.03 -0.02] 0.991	-0.02 (-0.48 0.49) [-0.03 -0.02] 0.988	-0.02 (-0.48 0.36) [-0.03 -0.02] 0.991	0.18 (-29.12 2.57) [-0.24 0.08] 0.921	-0.01 (-0.40 1.15) [-0.02 -0.01] 0.989	0.62 (-0.98 7.12) [0.76 0.91] 0.929	15.66 (1.23 83.12) [17.98 19.97] 0.599
$\mathbf{R}[SDSDvsSDSD]$								
$SENN - SENN$	-0.00 (-0.48 0.22) [-0.01 -0.00] 0.967	-0.00 (-0.03 0.02) [-0.00 0.00] 0.994	-0.00 (-0.03 0.02) [-0.00 -0.00] 0.994	-0.00 (-0.03 0.02) [-0.00 0.00] 0.994	0.00 (-1.68 0.07) [-0.01 -0.00] 0.955	-0.00 (-0.03 0.02) [-0.00 -0.00] 0.994	0.02 (-0.05 0.51) [0.03 0.04] 0.956	0.23 (0.01 2.17) [0.29 0.33] 0.846
$\mathbf{R}[SENNvsSENN]$								
$pNN50 - pNN50$	0.00 (-2.52 1.48) [0.00 0.00] 0.923	0.00 (-0.80 0.32) [0.00 0.00] 0.994	0.00 (-0.80 0.36) [0.00 0.00] 0.991	0.00 (-0.80 0.32) [0.00 0.00] 0.994	0.00 (-7.62 1.77) [0.04 0.15] 0.915	0.00 (-0.68 0.35) [0.00 0.00] 0.994	0.31 (-1.48 3.54) [0.44 0.54] 0.913	0.93 (0.00 2.83) [0.99 1.05] 0.912
$\mathbf{R}[pNN50vspNN50]$								
$TP - TP$	-0.00 (-0.89 0.88) [-0.01 -0.00] 0.954	-0.00 (-0.05 0.05) [-0.00 0.00] 0.995	-0.00 (-0.09 0.05) [-0.00 -0.00] 0.994	-0.00 (-0.05 0.05) [-0.00 0.00] 0.995	0.00 (-4.36 0.06) [-0.05 -0.02] 0.942	-0.00 (-0.06 0.04) [-0.00 -0.00] 0.994	0.04 (-0.16 0.87) [0.06 0.07] 0.952	0.53 (0.03 4.97) [0.64 0.74] 0.805
$\mathbf{R}[TPvsTP]$								
$LF^{norm} - LF^{norm}$	-0.03 (-3.09 2.37) [-0.11 -0.05] 0.956	0.01 (-0.49 0.72) [0.01 0.02] 0.988	0.01 (-0.55 0.72) [0.01 0.02] 0.986	0.01 (-0.50 0.72) [0.01 0.02] 0.987	-0.00 (-2.85 6.89) [-0.03 0.04] 0.933	0.00 (-0.64 0.77) [-0.01 0.00] 0.988	0.65 (-1.80 5.15) [0.95 1.17] 0.854	-1.91 (-18.12 5.65) [-2.80 -2.31] 0.765
$\mathbf{R}[LF^{norm}vsLF^{norm}]$								
$HF^{norm} - HF^{norm}$	-0.02 (-5.97 4.91) [-0.04 0.00] 0.873	-0.01 (-0.78 0.36) [-0.03 -0.02] 0.986	-0.01 (-0.78 0.51) [-0.03 -0.02] 0.983	-0.01 (-0.78 0.37) [-0.03 -0.02] 0.986	0.04 (-10.96 1.90) [-0.18 -0.05] 0.905	-0.02 (-0.79 0.46) [-0.05 -0.03] 0.985	0.19 (-5.38 7.03) [0.22 0.33] 0.886	5.13 (-2.24 39.80) [6.31 7.29] 0.625
$\mathbf{R}[HF^{norm}vsHF^{norm}]$								
$LF/HF - LF/HF$	0.00 (-3.84 0.87) [-0.02 -0.00] 0.839	0.00 (-0.17 0.20) [0.00 0.01] 0.986	0.00 (-0.23 0.20) [0.00 0.00] 0.982	0.00 (-0.19 0.20) [0.00 0.01] 0.985	-0.01 (-1.60 1.78) [-0.04 -0.01] 0.891	0.00 (-0.19 0.22) [0.00 0.01] 0.984	-0.01 (-2.61 0.85) [-0.13 -0.07] 0.862	-1.10 (-7.14 0.06) [-1.60 -1.40] 0.561
$\mathbf{R}[LF/HFvsLF/HF]$								
$RMS\ of\ RR\ of\ RR$	3.72 (0.23 55.21)	1.43 (0.00 8.32)	1.49 (0.00 8.95)	1.43 (0.00 8.32)	5.50 (0.83 51.90)	1.57 (0.14 10.11)	8.17 (1.59 32.81)	21.39 (4.09 75.50)

Table 5.5: Deviations of HRV parameters from original, error-free, RR interval time series for filter techniques in Test-1A. Deviations of HRV parameters are represented as follows: Median (2.5 percentile | 97.5 percentile) in the upper row and the confidence interval (computed by the difference between the population medians for matched pairs, according to [22]) in the lower row, for each HRV parameter. The correlation R was computed as the Kendall's rank correlation. The RMSD of the RR intervals is also represented by the median, the 2.5 and the 97.5 percentile. RMSSD, SDNN, SDSD and SENN are given in ms, pNN50 in %, LF^{norm} and HF^{norm} in n.u., TP in s² and LF/HF unit-less.

	BUFFER	GP-HA	IPFM-S	IPFM-D	IPFM-C	TPC-HT	PPHDIG	NOCORR
$RMSSD - RMSSD$	0.16 (-0.33 13.34) [0.94 1.21]	0.63 (-0.26 21.40) [1.5 11.93]	-0.03 (-0.68 0.68) [-0.04 -0.03]	-0.01 (-0.66 0.90) [-0.02 -0.00]	0.87 (-0.14 19.94) [1.49 1.84]	0.01 (-0.27 1.53) [0.03 0.05]	0.01 (-0.32 2.85) [0.02 0.03]	15.63 (1.23 82.96) [17.95 19.94]
$R[RMSSD_{vs}RMSSD]$	0.899	0.859	0.988	0.984	0.859	0.986	0.985	0.599
$SDNN - SDNN$	0.00 (-0.28 10.25) [0.48 0.63]	0.12 (-0.23 12.97) [0.60 0.80]	-0.01 (-0.75 0.38) [-0.03 -0.02]	-0.08 (-1.11 0.02) [-0.13 -0.11]	0.25 (-0.17 12.31) [0.64 0.83]	-0.01 (-0.75 0.32) [-0.02 -0.01]	-0.00 (-0.42 2.12) [0.01 0.02]	4.40 (0.28 32.97) [5.25 5.96]
$R[SDNN_{vs}SDNN]$	0.953	0.941	0.993	0.992	0.935	0.994	0.990	0.836
$SDSD - SDSD$	0.16 (-0.33 13.37) [0.94 1.21]	0.63 (-0.26 21.45) [1.5 11.93]	-0.03 (-0.68 0.68) [-0.04 -0.03]	-0.01 (-0.66 0.91) [-0.02 -0.00]	0.87 (-0.14 19.98) [1.49 1.84]	0.01 (-0.27 1.54) [0.03 0.05]	0.01 (-0.32 2.85) [0.02 0.03]	15.66 (1.23 83.12) [17.98 19.97]
$R[SDSD_{vs}SDSD]$	0.899	0.860	0.988	0.984	0.859	0.986	0.985	0.599
$SENN - SENN$	0.00 (-0.02 0.66) [0.02 0.03]	0.01 (-0.01 0.83) [0.03 0.04]	-0.00 (-0.05 0.02) [-0.00 -0.00]	-0.00 (-0.07 0.00) [-0.01 -0.01]	0.01 (-0.01 0.82) [0.03 0.04]	-0.00 (-0.04 0.02) [-0.00 -0.00]	0.00 (-0.02 0.13) [0.00 0.00]	0.23 (0.01 2.17) [0.29 0.35]
$R[SENN_{vs}SENN]$	0.956	0.946	0.994	0.993	0.941	0.994	0.990	0.846
$pNN50 - pNN50$	0.00 (-0.55 1.85) [0.28 0.32]	0.21 (-0.52 1.91) [0.31 0.36]	0.00 (-0.92 0.51) [0.00 0.00]	0.00 (-0.80 0.68) [0.00 0.00]	0.39 (-0.38 2.13) [0.41 0.48]	0.00 (-0.52 0.80) [0.00 0.00]	0.00 (-0.52 0.79) [0.00 0.00]	0.93 (0.00 2.83) [0.99 1.05]
$R[pNN50_{vs}pNN50]$	0.920	0.918	0.990	0.982	0.915	0.986	0.987	0.912
$TP - TP$	0.00 (-0.04 1.17) [0.05 0.07]	0.02 (-0.02 1.57) [0.07 0.09]	-0.00 (-0.10 0.06) [-0.00 -0.00]	-0.01 (-0.23 0.01) [-0.01 -0.01]	0.03 (-0.01 1.46) [0.08 0.10]	-0.00 (-0.14 0.05) [-0.00 -0.00]	-0.00 (-0.05 0.41) [0.00 0.00]	0.53 (0.03 4.97) [0.64 0.74]
$R[TP_{vs}TP]$	0.946	0.929	0.993	0.990	0.930	0.993	0.989	0.805
$LF^{norm} - LF^{norm}$	0.01 (-3.17 9.73) [0.05 0.13]	0.00 (-3.52 8.76) [0.02 0.05]	-0.00 (-0.66 0.85) [-0.00 0.01]	0.02 (-0.92 1.11) [0.03 0.05]	0.00 (-4.57 8.94) [0.01 0.05]	-0.03 (-1.41 0.41) [-0.08 -0.05]	0.00 (-0.81 1.26) [0.01 0.02]	-1.91 (-18.12 5.65) [-2.80 -2.31]
$R[LF^{norm}_{vs}LF^{norm}]$	0.900	0.910	0.985	0.980	0.900	0.982	0.983	0.765
$HF^{norm} - HF^{norm}$	0.02 (-1.65 10.20) [0.47 0.71]	0.07 (-1.41 13.92) [0.84 1.23]	-0.01 (-0.92 0.60) [-0.04 -0.02]	0.02 (-0.78 1.22) [0.02 0.04]	0.31 (-0.95 13.22) [0.94 1.29]	0.00 (-0.67 1.23) [0.00 0.01]	-0.00 (-0.89 1.47) [-0.02 -0.01]	5.13 (-2.24 39.80) [6.31 7.29]
$R[HF^{norm}_{vs}HF^{norm}]$	0.845	0.813	0.982	0.975	0.817	0.979	0.977	0.625
$LF/HF - LF/HF$	-0.00 (-4.59 0.26) [-0.15 -0.06]	-0.01 (-4.79 0.16) [-0.29 -0.19]	0.00 (-0.33 0.19) [-0.00 0.00]	-0.00 (-0.71 0.15) [-0.03 -0.02]	-0.08 (-4.65 0.16) [-0.33 -0.23]	-0.00 (-0.64 0.14) [-0.02 -0.01]	0.00 (-0.50 0.18) [0.00 0.00]	-1.10 (-7.14 0.06) [-1.60 -1.40]
$R[LF/HF_{vs}LF/HF]$	0.830	0.799	0.978	0.965	0.800	0.971	0.972	0.561
$RMS_{Eo}fRR$	3.82 (0.00 35.02)	4.68 (0.00 39.99)	2.06 (0.23 10.00)	2.32 (0.41 10.63)	5.47 (0.13 39.69)	1.96 (0.10 12.00)	1.11 (0.00 16.58)	21.39 (4.09 75.50)

Table 5.6: Deviations of HRV parameters from original, error-free, RR interval time series for model based techniques in Test-1A. Deviations of HRV parameters are represented as follows: Median (2.5 percentile | 97.5 percentile) in the upper row and the confidence interval (computed by the difference between the population medians for matched pairs, according to [22]) in the lower row, for each HRV parameter. The correlation R was computed as the Kendall's rank correlation. The RMSD of the RR intervals is also represented by the median, the 2.5 and the 97.5 percentile. RMSSD, SDNN, SDDSD and SENN are given in ms, pNN50 in %, LF^{norm} and HF^{norm} in n.u., TP in s^2 and LF/HF unit-less.

5.2 Test-1B: Correction of Successive Ectopic Beats

This section is structured in the same way as the previous section (see section “5.1 Test-1A: Correction of Single Ectopic Beats”). However, Test-1B considers successive ectopic beats, like PVC couplets and triplets, bigeminy, trigeminy and sustained and non-sustained ventricular tachycardia in addition to single PVCs and PACs. Hence, Test-1B can be seen as testing for a high noise level and Test-1A for a moderate one. As in Test-1A, a PCA was performed on the RMSD of the time- and frequency-domain HRV parameters. The detailed results of each time- and frequency-domain HRV variable are illustrated in appendix B.2.1 and B.2.2. In the current section, only the results of the PCA are described. Figure 5.10 illustrates the Eigenvalue diagram for the determination of the number of necessary principal components. The figure shows a flattening decrease after the second Eigenvalue, meaning that the first principal component is sufficient to describe most of the variation in time- and frequency-domain HRV parameters.

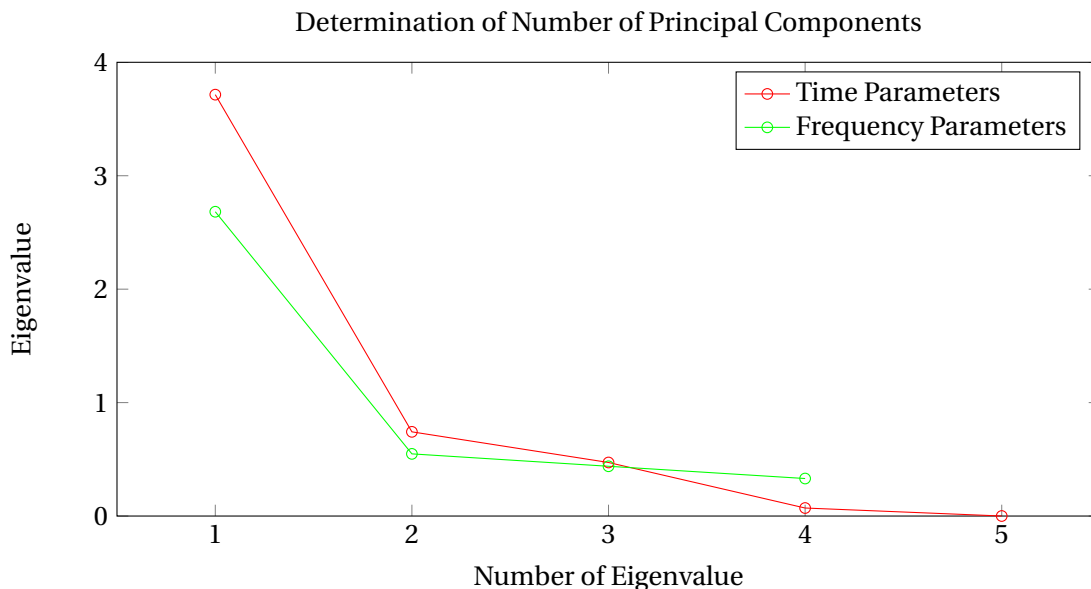


Figure 5.10: Eigenvalue plot of Test-1B: Time- and frequency-domain HRV parameters show a flattening decrease after the second Eigenvalue. Hence, the first PC is sufficient for the representation of the HRV parameters in both cases.

Figure 5.11 displays the score-score plot of the first and second principal component of the HRV frequency-domain parameters. Similar to Test-1A, PCA on the original data makes it hard to distinguish the correction methods, whereas taking the logarithm of the data enhances the separation. Low scores in PC 1 refer to a good correction ability, high scores to a bad one. Especially well performing correction approaches result in more compact clusters, compared to the same methods in Test-1A (compare to figure 5.2). As a result, DELETE performs obviously best in the frequency domain. Further, INT-1O and INT-0O are both located near to DELETE and are followed by IPFM-S in terms of performance. TPC-HT, THRESH, GP-IIA and IPFM-D

also accomplish reasonable scores but all other methods are slightly set apart. Most of them do not form distinct clusters, except of CSINTP, which is located far away from the other algorithms and performs even worse than no correction. Therefore, cubic spline interpolation seems to introduce many highly erroneous RR intervals, instead of replacing the original erroneous ones with better estimates. BUFFER, IPFM-C, IRFILT, MEDFIL AMFILT, AROFIL, kNN-AF and PPHDIG form one big cluster that performs rather bad. However, all correction approaches, except CSINTP, still perform better than no correction. Interestingly, all methods form a more compact cluster than in Test-1A. Therefore, good correction approaches are not well separated from worse ones, compared to Test-1A (compare figure 5.11b with 5.2b). Even the distance to the *Zero Error* point is enlarged for the best correction approaches.

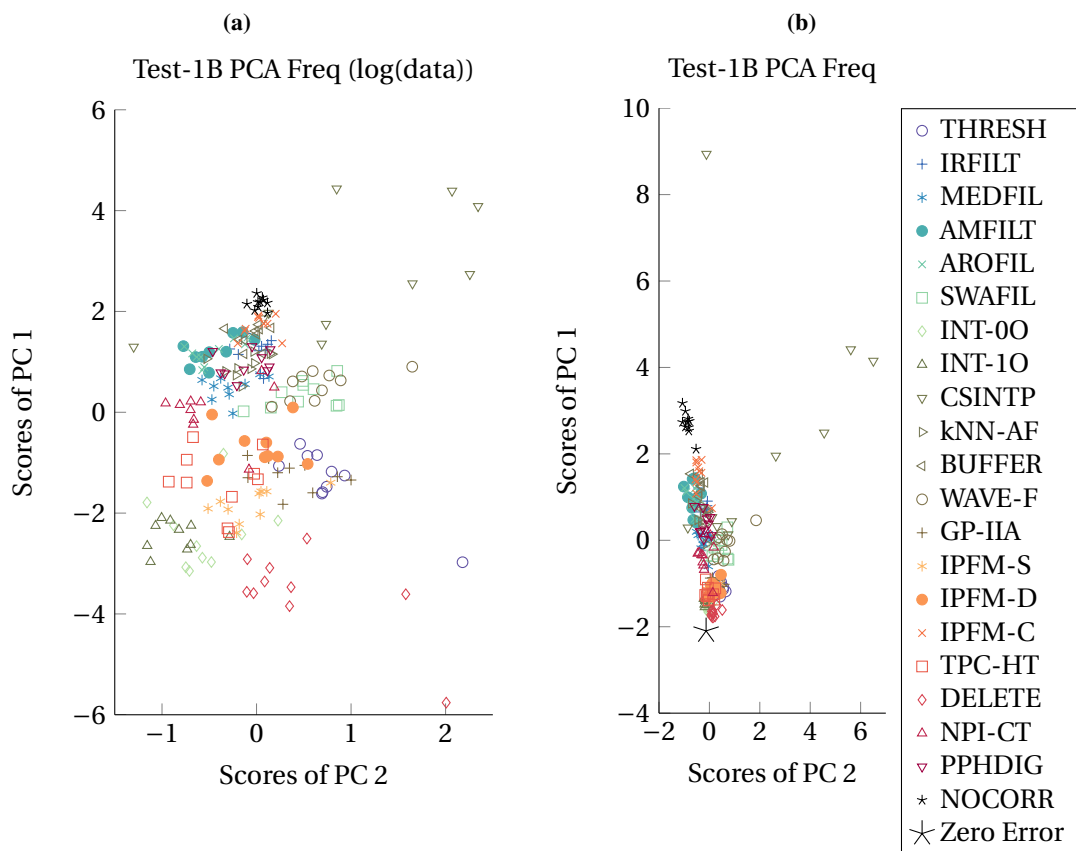


Figure 5.11: Test-1B score-score plots of frequency-domain HRV parameters: Low scores of PC 1 refer to a good correction behavior, high scores to a bad one (all variables have positive loadings for this component). In the case of PC 2 there is no linear relationship to the original data because of positive and negative loadings. (a) The logarithm was taken from the data before PCA was applied, resulting in a better separation of the clusters, especially in the case of lower scores. (b) Plot of original data and addition of the “Zero Error” point. The legend is valid for both diagrams.

The score-score plot of the time-domain HRV parameters shows that taking the logarithm of the data results in a very good separation of most correction approaches, especially in lower scores (see figure 5.12a). Therefore, DELETE is clearly the best correction approach in terms of correcting time-domain HRV parameters. INT-00, INT-10 and IPFM-S also accomplish a good performance. IPFM-C and BUFFER perform worst, but CSINTP shows the largest variance. THRESH and GP-IIA perform rather well and are still clearly set apart from the cluster of worse performing approaches. There is a linear increase in the scores of PC-1 in the plot of the original data (see figure 5.12b) from the *Zerro Error* point towards the worst performing method. This trend is even more extreme than in the frequency domain (compare to figure 5.11b).

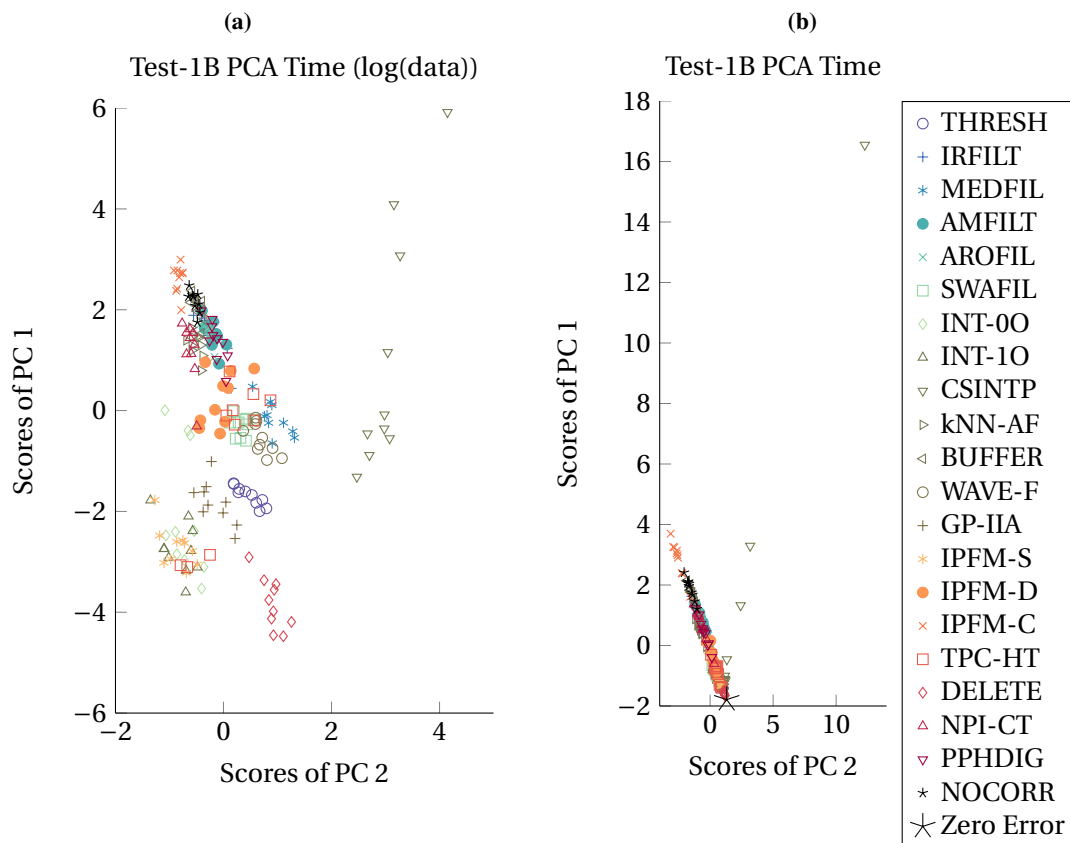


Figure 5.12: Test-1B score-score plots of time-domain HRV parameters: Low scores of PC 1 refer to a good correction behavior, high scores to a bad one (all variables have positive loadings for this component). In the case of PC 2 there is no linear relationship to the original data, because of positive and negative loadings. (a) The logarithm was taken from the data before PCA was applied. This results in a better separation of the clusters, especially in case of lower scores. (b) Plot of original data and addition of the “Zero Error” point. The legend is valid for both diagrams.

Box plots of time- and frequency-domain HRV parameters illustrate the behavior of the cor-

rection methods by means of scores of the first principal component (see figure 5.13 and 5.14). Both figures show the same relations between the correction approaches, except of CSINTP which performs a little better in the time domain than in the frequency domain. CSINTP, INT-00 and TPC-HT show a larger variation in time-domain parameters compared to frequency ones. The description of the results can be obtained from that of the score-score plots.

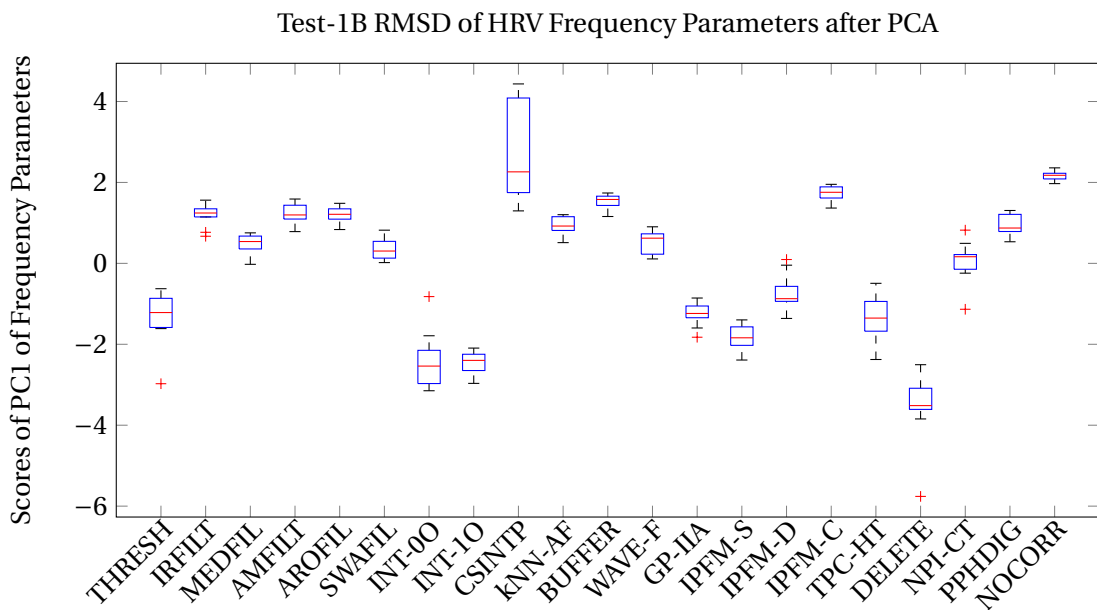


Figure 5.13: Test-1B box plot of HRV frequency-domain parameters. PCA was performed on the RMSD of the HRV parameters mentioned in table 4.8 for ten test sets containing 151 signals, each. Hence, the box plot displays the scores of the first principal component with ten values for each method.

Like in Test-1A, the RMSD of the RR intervals is also very similar to the time- and frequency-domain HRV parameters (compare figure 5.15 with figure 5.13 and 5.14). The RMSD is lowest for DELETE, INT-10, INT-00 and for IPFM-S. Further, it is highest for CSINTP in terms of the outliers.

The computation time increased a lot for many methods compared to Test-1A (compare figure 5.6 with figure 5.16). For example, the computation time of the two adaptive median filters, AROFIL and AMFILT, increased almost by two orders of magnitude. In contrast, the best performing methods, DELETE, INT-00 and THRESH require nearly the same computation time as in the low error case (DELETE has a median computation time of zero, which may be related to the inability of the *profiler* in MATLAB[®] to display very small values). DELETE and INT-00 are clearly set apart from the majority of the other correction algorithms by one order of magnitude. The IPFM model with cost function (IPFM-C) and PPHDIG need by far the highest computation time (more than 200 ms, whereby all other methods require less than 60 ms on the

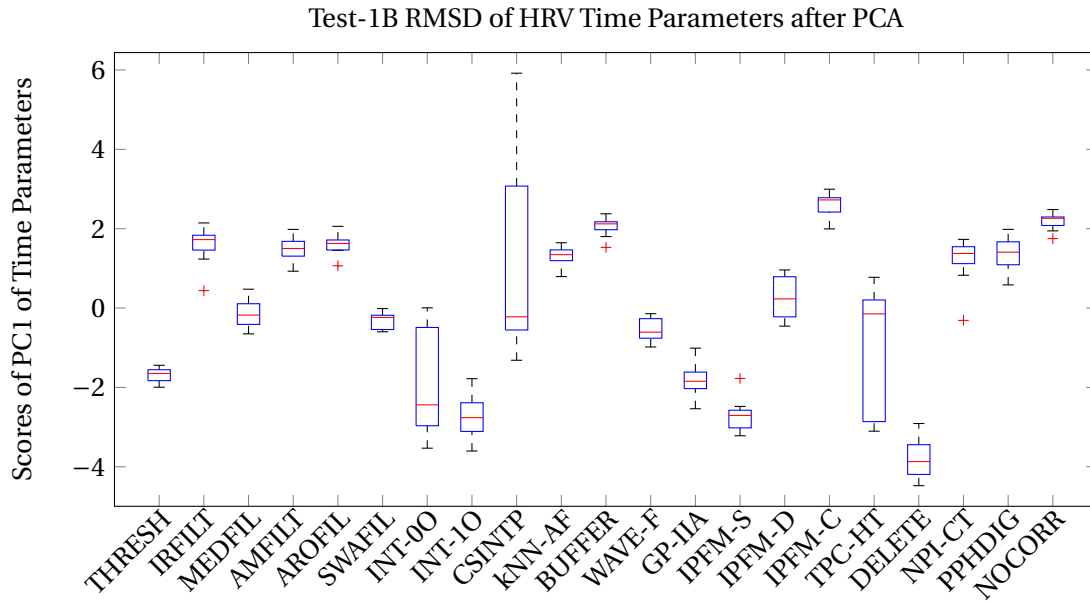


Figure 5.14: Test-1B box plot of HRV time-domain parameters. PCA was performed on the RMSD of the HRV parameters mentioned in table 4.7 for ten test sets containing 151 signals, each. Hence, the box plot displays the scores of the first principal component with ten values for each method.

test system).

The peak memory in Test-1B differs over two orders of magnitude between the correction approaches, like in Test-1A (compare figure 5.7 with figure 5.17). Both test cases even show nearly identical values for each correction approach, just the outliers are sometimes different. The peak memory is lowest for IRFILT and highest for IPFM-C. In general, more complex methods such as models demand a higher peak memory, whereas simpler ones, such as threshold and mean filters, require less memory.

The comparison between Test-1B and Test-1A showed that sensitivity decreased by a factor of ten for the BUFFER, by a factor of five for PPHDIG, by a factor of three for IRFILT, AROFIL and SWAFIL and almost by a factor of two for MEDFIL, AMFIL and NPI-CT (compare table 5.2 with 5.7). All of these methods yield a sensitivity lower than 50 %, except of the MEDFIL, AMFIL and NPI-CT. In contrast, the best performing approaches, INT-10, CSINTP, kNN-AF, DELETE, TPC-HT, IPFM-S and IPFM-D, reach a sensitivity of almost 100 %. PPV increased for most approaches, except for INT-00 and NPI-CT, compared to Test-1A (compare table 5.2 with 5.7). Especially, the PPV of WAVE-F and THRESH has increased by a factor of six and IRFILT and SWAFIL doubled their PPV. This huge increase is related to the much higher amount of erroneous RR intervals (more positive values for correction). All other methods that already yielded a PPV of 100 % in Test-1A also reached this value in Test-1B: All

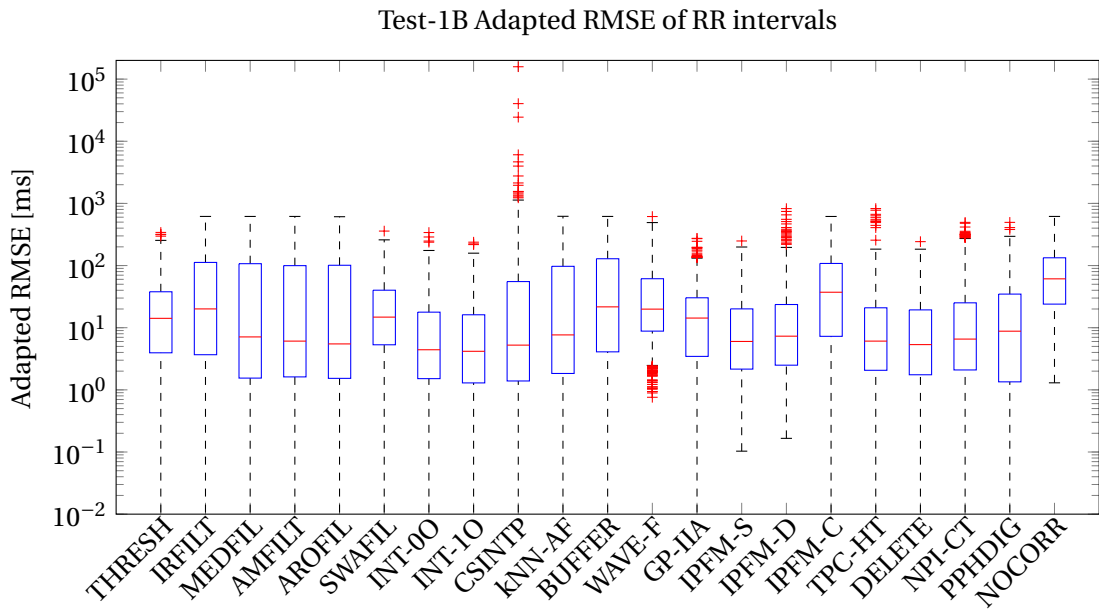


Figure 5.15: Test-1B adjusted box plot of adapted RMSD of RR intervals. The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1A cases, containing 151 signals each. Note: Whiskers that extend below 10^{-2} are zero and thus cannot be displayed on the logarithmic scale.

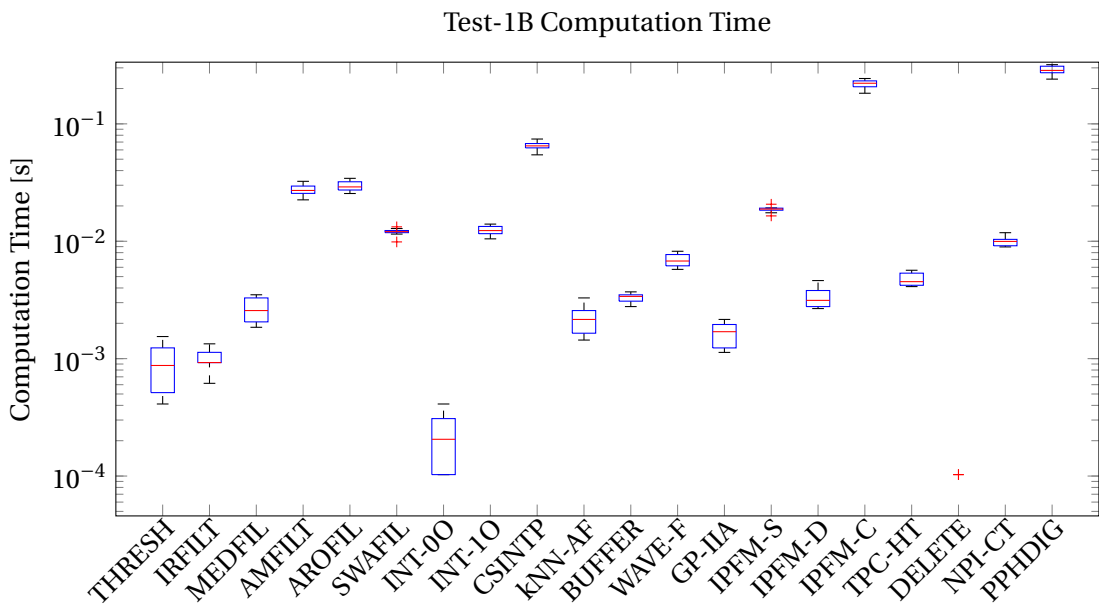


Figure 5.16: Test-1B box plot of mean computation time. Mean computation time was calculated for each of the ten test sets by dividing the total time by the 151 signals processed. Hence the box plot displays ten values.

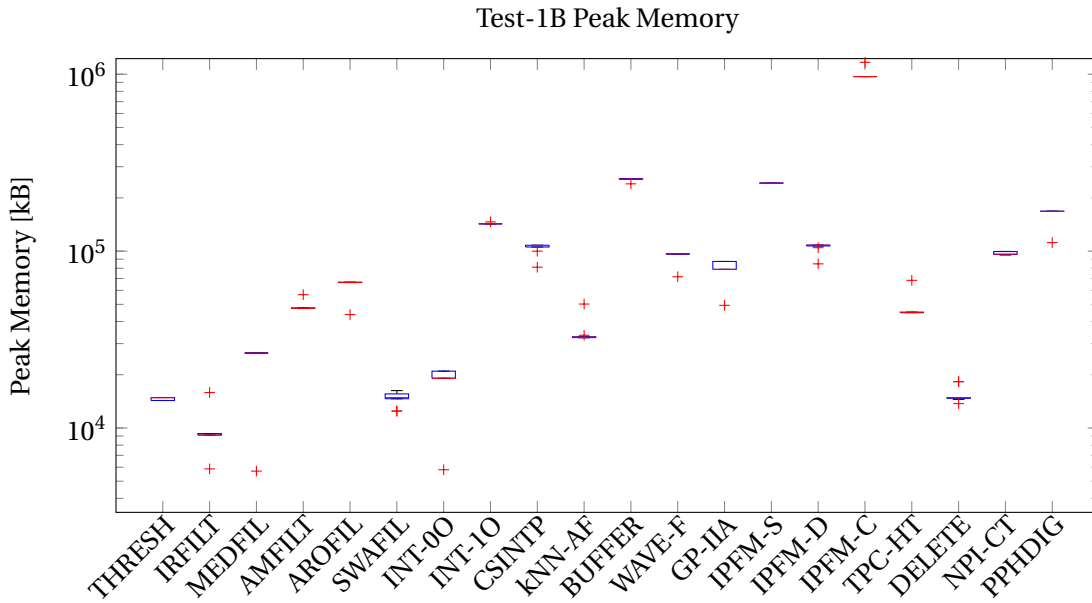


Figure 5.17: Test-1B box plot of peak memory. Peak memory was determined by means of the built-in “profiler” in MATLAB® for each of the ten test sets separately. Hence, the box plot displays ten values for each method.

median filters, INT-10, CSINTP, kNN-AF, TPC-HT and DELETE.

The RMSD of most HRV parameters increased, due to the much higher error density in Test-1B, compared to Test-1A. This can be seen best in the large errors in the *total power* and the *SDNN* for cubic spline interpolation. Most approaches result in a remarkable reduction of the RMSD in all HRV parameters, compared to no correction. Nevertheless, some correction approaches induce more errors than no correction, although there are differences between the errors in different HRV parameters.

RMSD between original and corrected RR interval time series of HRV parameters of all 1510 signals in Test-1B												
	Sens [%]	PPV [%]	RMSSD [ms]	SDSD [ms]	SDNN [ms]	SENN [ms]	pNN50 [%]	LF/HF [-]	LF ^{norm} [n.u.]	HF ^{norm} [n.u.]	TP [s ²]	
THRESH	79.85	72.67	8.81	8.82	19.23	1.05	2.35	1.90	8.51	5.92	3.40	
IRFILT	23.78	76.08	89.58	89.73	68.03	4.07	7.25	2.63	13.43	17.73	21.62	
MEDFIL	62.40	100.00	15.42	15.45	57.87	3.55	3.75	1.92	12.84	11.16	15.05	
AMPFLT	50.75	100.00	93.23	93.40	65.46	3.92	5.97	2.16	12.70	21.28	19.99	
AKOFL	33.14	100.00	98.40	98.56	65.95	3.90	6.50	2.11	12.76	21.91	19.90	
SWAFIL	36.25	74.45	12.69	12.71	40.40	2.42	5.29	2.79	11.80	10.17	7.92	
INT-00	98.93	99.80	19.47	19.50	11.13	0.63	4.24	0.81	4.19	3.40	3.88	
INT-10	100.00	100.00	5.00	5.01	5.85	0.34	4.18	0.66	4.71	3.19	0.93	
CSINTP	100.00	100.00	91.67	91.82	3186.03	182.57	3.85	9.29	12.34	9.90	86408.92	
KNN-AF	99.98	100.00	53.57	53.65	54.11	3.30	9.75	2.35	13.32	16.18	13.70	
BUFFER	8.83	90.21	120.27	120.46	76.85	4.51	8.65	2.77	14.62	25.07	24.65	
WAVE-F	92.85	54.00	8.18	8.20	44.37	2.74	5.06	3.32	13.27	10.61	11.32	
GP-IIA	99.05	98.53	8.14	8.15	12.06	0.82	3.65	1.68	6.69	7.50	1.46	
IPFM-S	99.98	98.49	5.09	5.10	5.63	0.32	4.16	1.25	5.27	3.39	0.90	
IPFM-D	100.00	98.41	48.08	48.13	44.15	2.20	4.06	1.53	5.31	6.48	51.45	
IPFM-C	84.43	99.17	154.77	155.01	90.65	5.23	13.64	3.11	14.82	22.31	27.37	
TPC-HT	99.85	100.00	25.63	25.68	38.47	2.35	3.57	1.10	4.70	4.72	12.19	
DELETE	100.00	100.00	2.45	2.46	4.49	0.79	1.11	0.94	3.67	2.24	0.69	
NPI-CT	66.70	98.90	86.10	86.23	47.49	2.68	6.91	1.55	8.24	13.82	16.44	
PPHDIG	21.78	97.54	74.51	74.64	66.41	3.96	6.61	2.35	13.23	14.92	20.67	
NOCORR	-	-	130.60	130.81	79.43	4.67	9.25	3.56	16.43	29.28	24.94	
ORIGINAL			25.36 (7.06 95.32)	25.40 (7.07 95.49)	45.84 (11.13 130.30)	2.58 (0.49 7.86)	4.48 (0.00 60.59)	2.91 (0.15 9.96)	23.89 (3.52 56.61)	8.70 (1.00 57.98)	2.01 (0.18 16.51)	

Table 5.7: Sensitivity and PPV are shown in the first two columns (the higher the better). The right part of the table shows the RMSD of the HRV parameters (the smaller the better) of all 1510 signals in Test-1B. The RMSD is calculated with respect to the original HRV parameters. These are illustrated in the last row by means of the median, the 2.5 percentile and the 97.5 percentile.

The summarized results of Test-1B in the Kivat diagrams illustrate that the performance of the methods differs more than in Test-1A (compare figure 5.9 with figure 5.18). Therefore, the methods can be better separated, allowing an easier detection of the best methods, except in the model class. In the “Interpolation and Removal” class, DELETE clearly performs best, whereas NPI-CT and CSINTP perform worst. In the “Filtering” class all methods yield very low scores in time- (below 50) and frequency-domain HRV parameters (below 30). However, the kNN-AF still reaches the highest possible scores in all other variables and thus, performs best in this class. The weakest performing filter is IRFILT, showing the lowest scores in all parameters, except PPV and peak memory. Comparison of the correction approaches in the model class shows that IPFM-S achieves the highest scores, except in peak memory and computation time. GP-IIA performs nearly similar and reaches scores that just are ten points lower in the error of the time- and frequency-domain HRV parameters. A comparison of the best methods of each class clearly displays that deletion is the best approach to handle huge amounts of errors. IPFM-S still performs reliably, whereas kNN-AF is clearly outperformed.

Table 5.8 summarizes the scores of all approaches in Test-1B. The table is intended to give a quick overview of the mean performance of all correction approaches and is ordered in decreasing scores. Hence, methods on top of the table are highly recommended for correction of successive ectopic beats in this test case.

The detailed deviations of all frequency- and time-domain HRV parameters used are summarized in the same way as described in section section “5.1 Test-1A: Correction of Single Ectopic Beats”.

Table 5.9 illustrates the deviations of the HRV parameters and the RMSD of the RR intervals of all removal and interpolation approaches. The correlation coefficients are lower, compared to Test-1A and also the 2.5 and 97.5 percentile show a much larger deviation from the median. Interestingly, the median itself did not change in most of the HRV parameters remarkably. The RMSD of the RR intervals also increased for all approaches. In general, the relations between the performances of the individual methods are similar to Test-1A. However, DELETE performs a little better than all other approaches and CSINTP shows the largest deviation in most HRV parameters.

Table 5.10 demonstrates that the performance of all filter approaches decreased tremendously, compared to Test-1A. The correlation coefficients are much lower, and the 2.5 and 97.5 percentile show an enormous increase in the deviation from the median. Surprisingly, the median itself still is very small in most parameters for all filter methods. The RMSD of the RR intervals increased for all approaches at least by a factor of two, compared to Test-1A. A comparison of table 5.10 with 5.9 clearly illustrates that filters cannot compete with interpolation and removal approaches, when dealing with successive ectopic beats.

Like the filter methods, most models perform much weaker when dealing with ectopic segments and successive ectopic beats (see table 5.11). Most approaches show a high decrease in the correlation coefficients and an increase in the deviations of the 2.5 and 97.5 percentile. A comparison of table 5.11 with 5.9 clearly illustrates that also models cannot compete with interpolation and removal approaches when dealing with successive ectopic beats.

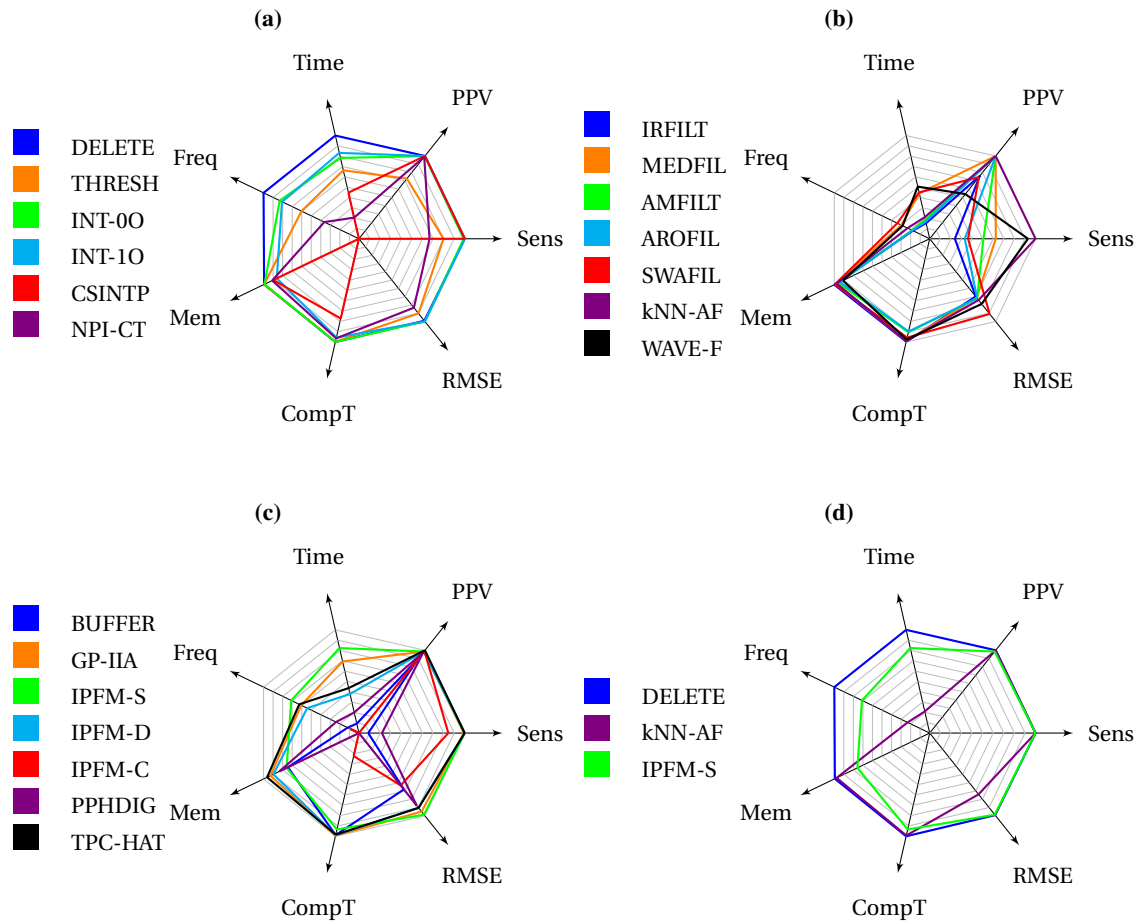


Figure 5.18: Test-1B Kivat Diagrams: Kivat diagrams are shown for the correction methods of each class separately, namely (a) “Interpolation and Removal”, (b) “Filtering” and (c) “Models”. The fourth diagram (d) compares the best methods of each class (determined by the highest mean scores of all variables) with each other. Abbreviations of the variables can be obtained from table 5.1.

	Sens	PPV	Time	Freq	Mem	CompT	RMSD
DELETE	100.00	100.00	100.00	100.00	99.40	100.00	99.05
INT-00	98.93	99.80	78.32	83.09	98.96	99.93	99.92
INT-10	100.00	100.00	83.21	80.66	86.10	95.68	100.00
IPFM-S	99.98	98.49	82.36	70.99	75.77	93.41	98.64
GP-IIA	99.05	98.53	69.31	60.58	92.74	99.41	94.93
TPC-HT	99.85	100.00	43.55	62.56	96.28	98.41	90.42
IPFM-D	100.00	98.41	37.82	54.26	89.71	98.90	90.14
THRESH	79.85	72.67	66.31	60.19	99.41	99.69	90.23
kNN-AF	99.98	100.00	20.92	23.18	97.55	99.24	74.05
MEDFIL	62.40	100.00	44.01	29.81	98.19	99.10	71.29
NPI-CT	66.70	98.90	20.46	36.31	90.94	96.51	83.28
WAVE-F	92.85	54.00	50.53	28.37	90.91	97.62	78.71
SWAFIL	36.25	74.45	44.92	33.88	99.40	95.77	90.84
AMFILT	50.75	100.00	18.56	18.43	96.01	90.51	71.86
AROFIL	33.14	100.00	16.65	18.16	94.04	89.82	71.99
CSINTP	100.00	100.00	44.67	0.00	89.76	77.24	0.00
IRFILT	23.78	76.08	15.16	17.61	100.00	99.68	69.79
BUFFER	8.83	90.21	9.15	11.81	74.40	98.81	68.08
PPHDIG	21.78	97.54	19.96	24.03	83.50	0.00	87.36
IPFM-C	84.43	99.17	0.00	8.74	0.00	22.25	63.31

Table 5.8: Summarized results of Test-1B: Correction methods are order decreasing by their mean performance in all parameters. Abbreviations are: Sensitivity (Sens), positive predictive value (PPV), PC-1 of time-domain HRV parameters (Time), PC-1 of frequency-domain HRV parameters (Freq), Peak Memory (Mem), Computation Time (CompT) and RMSD of all RR intervals (RMSD). Sensitivity and PPV are given in % whereas all other parameters denote the score with respect to the best performing (100) and worst performing (0) approach (see section 4.5.6 for a more detailed description).

	DELETE	THRESH	INT-00	INT-10	CSINTP	NPI-CT	NOCORR
$RMSSD - RMSSD$	0.13 (-2.105,76) [0.200,27]	0.15 (-14.5116,01) [0.350,72]	-0.31 (-11.920,04) [-1.08,-0.75]	-0.45 (-13.64,-0.00) [-1.31,-0.98]	-0.21 (-9.411,37) [-0.62,-0.46]	0.10 (-2.303,47,57) [0.230,33]	42.23 (2.293,62,96) [56.306,4,73]
$R[RMSSD_{vs}RMSSD]$	0.960	0.797	0.906	0.903	0.922	0.805	0.291
$SDNN - SDNN$	0.07 (-6.619,80) [0.090,16]	-0.01 (-56.273,0,22) [-1.99,-0.90]	-0.14 (-15.510,63) [-0.76,-0.47]	-0.07 (-14.133,29) [-0.34,-0.20]	0.10 (-0.735,93,08) [2.43,6,67]	0.00 (-6.241,86,62) [0.000,0,4]	27.80 (0.49121,2,24) [37.214,4,12]
$R[SDNN_{vs}SDNN]$	0.951	0.719	0.913	0.929	0.609	0.807	0.394
$SDSD - SDSD$	0.13 (-2.065,78) [0.200,27]	0.15 (-14.481,6,14) [0.360,73]	-0.31 (-11.940,0,4) [-1.08,-0.76]	-0.45 (-13.67,-0.00) [-1.32,-0.98]	-0.21 (-9.421,3,7) [-0.63,-0.46]	0.10 (-2.313,48,05) [0.230,33]	42.31 (2.293,63,54) [56.396,4,83]
$R[SDSD_{vs}SDSD]$	0.960	0.797	0.906	0.903	0.922	0.805	0.291
$SENN - SENN$	0.05 (-0.052,34) [0.110,14]	0.02 (-2.832,59) [0.030,0,7]	-0.01 (-0.890,0,4) [-0.04,-0.03]	-0.00 (-0.810,0,18) [-0.02,-0.01]	0.01 (-0.043,1,27) [0.120,3,4]	0.00 (-0.259,8,2) [0.000,0,1]	1.51 (0.021,3,28) [2.032,3,8]
$R[SENN_{vs}SENN]$	0.891	0.734	0.920	0.934	0.625	0.812	0.453
$pNN50 - pNN50$	0.02 (-1.572,66) [0.060,10]	0.07 (-4.174,0,1) [0.130,22]	0.00 (-12.330,3,31) [-0.30,-0.19]	0.00 (-12.840,0,0) [-0.42,-0.28]	0.00 (-9.080,7,4) [-0.25,-0.16]	0.00 (-2.282,7,69) [0.070,1,6]	1.55 (-5.951,2,7,69) [2.332,9,8]
$R[pNN50_{vs}pNN50]$	0.953	0.845	0.933	0.935	0.934	0.848	0.625
$TP - TP$	0.00 (-1.091,22) [0.000,0,1]	-0.00 (-8.623,4,4) [-0.23,-0.12]	-0.02 (-2.400,0,3) [-0.08,-0.05]	-0.01 (-2.000,1,5) [-0.05,-0.03]	0.01 (-0.151,8,2,7,40) [0.230,6,8]	0.00 (-0.796,1,29) [0.000,0,1]	4.33 (0.067,6,7,3) [6.868,8,9]
$R[TP_{vs}TP]$	0.947	0.688	0.904	0.919	0.608	0.806	0.327
$LF^{norm} - LF^{norm}$	0.03 (-7.788,93) [0.010,12]	0.07 (-17.502,0,29) [0.040,40]	-0.03 (-12.274,5,3) [-0.17,-0.08]	-0.01 (-15.063,0,0) [-0.20,-0.09]	-0.01 (-38.917,4,8) [-1.34,-0.20]	-0.02 (-29.207,1,8) [-0.18,-0.07]	-5.90 (-41.411,6,5,1) [-8.37,-7,0,3]
$R[LF^{norm}_{vs}LF^{norm}]$	0.873	0.658	0.879	0.864	0.566	0.762	0.349
$HF^{norm} - HF^{norm}$	0.01 (-3.364,74) [0.010,0,6]	0.44 (-8.801,2,98) [0.821,1,9]	-0.10 (-5.372,7,3) [-0.25,-0.18]	-0.15 (-8.701,0,7) [-0.41,-0.30]	-0.17 (-36.371,5,4) [-1.24,-0.75]	0.03 (-4.316,3,49) [0.050,1,2]	7.22 (-22.717,8,5,5) [10.241,2,8,0]
$R[HF^{norm}_{vs}HF^{norm}]$	0.920	0.709	0.920	0.914	0.599	0.828	0.307
$LF/HF - LF/HF$	0.00 (-2.141,27) [-0.02,-0.00]	-0.09 (-5.682,0,3) [-0.32,-0.22]	0.01 (-1.541,4,0) [0.010,0,2]	0.01 (-1.391,3,5) [0.020,0,3]	0.01 (-2.320,7,9) [0.020,0,4]	-0.01 (-4.270,9,7) [-0.08,-0.04]	-1.62 (-8.521,0,6) [-2.09,-1,8,5]
$R[LF/HF_{vs}LF/HF]$	0.892	0.708	0.904	0.909	0.778	0.779	0.270
$RMSE_{of}RR$	5.35 (0.367,3,7,1)	14.15 (0.401,4,3,1,0)	4.42 (0.226,2,9,2)	4.17 (0.196,2,0,9)	5.24 (0.218,0,5,1,7)	6.56 (0.252,3,8,6,6)	61.13 (6.163,1,7,5,2)

Table 5.9: Deviations of HRV parameters from original, error-free, RR interval time series for removal and interpolation techniques in Test-1B. Deviations of HRV parameters are represented as follows: Median (2.5 percentile | 97.5 percentile) in the upper row and the confidence interval (computed by the difference between the population medians for matched pairs, according to [22]) in the lower row, for each HRV parameter. The correlation R was computed as the Kendall's rank correlation. The RMSD of the RR intervals is also represented by the the median, the 2.5 and the 97.5 percentile. RMSSD, SDNN, SDSD and SENN are given in ms, pNN50 in %, LF^{norm} and HF^{norm} in n.u., TP in s^2 and LF/HF unit-less.

	IRFLT	MEDFL	AMFLT	AROFL	SWAFL	KNN-AF	WAVE-F	NOCORR
$RMSSD - RMSSD$	2.92 (-1.01336.13) [6.548.87]	0.02 (-3.8241.63) [2.103.15]	0.01 (-1.37333.44) [1.835.86]	0.01 (-3.16335.87) [1.172.99]	0.59 (-2.9221.47) [0.631.06]	0.16 (-1.77162.64) [5.148.42]	1.08 (-8.9218.96) [1.501.94]	42.23 (2.29362.96) [56.3064.73]
$RMSSD_{vs}RMSSD$	0.551	0.728	0.609	0.615	0.742	0.552	0.802	0.291
$SDNN - SDNN$	1.09 (-5.88206.87) [19.6025.87]	0.17 (-0.43167.44) [18.0223.09]	0.07 (-0.85203.54) [13.3020.56]	0.08 (-0.54197.52) [14.8621.65]	0.29 (-2.18111.82) [8.1414.55]	0.14 (-0.49161.88) [16.7621.94]	1.93 (-4.65135.43) [4.296.22]	27.80 (0.49212.24) [37.2144.12]
$SDNN_{vs}SDNN$	0.516	0.534	0.546	0.542	0.536	0.543	0.644	0.394
$SDSD - SDSD$	2.92 (-1.03336.53) [6.558.88]	0.02 (-3.8341.70) [2.103.16]	0.01 (-3.18333.98) [1.833.87]	0.01 (-3.17336.66) [1.182.99]	0.59 (-2.9721.51) [0.631.07]	0.16 (-1.77162.90) [5.158.43]	1.08 (-8.9318.98) [1.501.94]	42.31 (2.29363.54) [56.3964.83]
$SDSD_{vs}SDSD$	0.552	0.728	0.610	0.615	0.743	0.552	0.802	0.291
$SENN - SENN$	0.06 (-0.3712.98) [0.9811.32]	0.01 (-0.0310.66) [0.8911.15]	0.00 (-0.0512.66) [0.6611.01]	0.00 (-0.0311.94) [0.7211.06]	0.03 (-1.316.87) [0.530.82]	0.01 (-0.0310.51) [0.8411.09]	0.10 (-0.258.70) [0.240.35]	1.51 (0.0213.28) [2.032.38]
$SENN_{vs}SENN$	0.557	0.578	0.586	0.581	0.575	0.586	0.680	0.453
$pNN50 - pNN50$	0.00 (-7.5124.66) [0.260.43]	0.00 (-8.844.89) [0.000.08]	0.00 (-7.6219.05) [0.000.14]	0.00 (-7.4121.18) [0.000.13]	0.62 (-10.2710.89) [0.7511.00]	0.00 (-5.8432.49) [0.280.53]	0.62 (-9.6310.12) [0.740.94]	1.55 (-5.9527.69) [2.332.98]
$pNN50_{vs}pNN50$	0.733	0.840	0.789	0.778	0.744	0.690	0.788	0.625
$TP - TP$	0.17 (-0.63171.72) [3.284.54]	0.02 (-0.0645.73) [2.763.83]	0.01 (-0.1466.82) [1.803.22]	0.01 (-0.0765.48) [2.113.45]	0.03 (-4.1821.75) [1.262.26]	0.02 (-0.0641.57) [2.523.68]	0.19 (-0.8433.71) [0.400.58]	4.33 (0.0676.73) [6.868.59]
$TP_{vs}TP$	0.476	0.492	0.511	0.507	0.493	0.505	0.609	0.327
$LF^{norm} - LF^{norm}$	-0.31 (-38.0519.02) [-2.291.25]	-0.02 (-35.6817.73) [-1.6010.36]	-0.02 (-40.468.25) [-1.3710.36]	-0.02 (-39.3010.37) [-1.2010.26]	-0.17 (-33.6016.40) [-0.8310.37]	-0.06 (-37.4915.79) [-1.6810.54]	1.00 (-29.3636.02) [1.352.00]	-5.90 (-41.4116.51) [-8.3717.03]
$LF^{norm}_{vs}LF^{norm}$	0.508	0.545	0.536	0.537	0.552	0.532	0.517	0.349
$HF^{norm} - HF^{norm}$	-0.06 (-35.1464.97) [-0.2310.02]	-0.07 (-38.0919.21) [-0.7210.30]	-0.01 (-24.2175.34) [-0.0501.01]	-0.01 (-24.3275.08) [-0.0501.00]	-0.17 (-35.0015.70) [-1.5210.89]	-0.02 (-31.0150.70) [-0.0710.01]	-0.00 (-33.8311.00) [-0.6510.20]	7.22 (22.7178.55) [10.2412.80]
$HF^{norm}_{vs}HF^{norm}$	0.456	0.569	0.570	0.569	0.559	0.539	0.532	0.307
$LF/HF - LF/HF$	-0.01 (-6.6713.21) [-0.3610.16]	0.00 (-5.3113.03) [-0.0101.00]	-0.00 (-6.2712.25) [-0.0910.03]	-0.00 (-6.0912.47) [-0.0610.02]	0.00 (-4.9417.47) [-0.0301.04]	-0.00 (-6.1914.12) [-0.0710.02]	0.06 (-4.6417.20) [0.0710.19]	-1.62 (-8.5211.06) [-2.091.85]
$LF/HF_{vs}LF/HF$	0.462	0.647	0.592	0.596	0.587	0.553	0.570	0.270
$RMSE$	20.06 (0.42317.25)	7.11 (0.35310.38)	6.09 (0.35298.26)	5.49 (0.35294.08)	14.80 (0.92149.51)	7.66 (0.20280.96)	19.91 (2.20259.42)	61.13 (6.16317.52)

Table 5.10: Deviations of HRV parameters from original, error-free, RR interval time series for filtering techniques in Test-1B. Deviations of HRV parameters are represented as follows: Median (2.5 percentile | 97.5 percentile) in the upper row and the confidence interval (computed by the difference between the population medians for matched pairs, according to [22]) in the lower row, for each HRV parameter. The correlation R was computed as the Kendall's rank correlation. The RMSD of the RR intervals is also represented by the the median, the 2.5 and the 97.5 percentile. RMSSD, SDNN, SDSD and SENN are given in ms, pNN50 in %, LF^{norm} and HF^{norm} in n.u., TP in s^2 and LF/HF unit-less.

	BUFFER	GP-HIA	IPFM-S	IPFM-D	IPFM-C	TPC-HT	PPHDIG	NOCORR
$RMSSD - RMSSD$	6.70 (-0.33362.96) [19.0426.35]	-0.01 (-9.3618.61) [0.130.53]	-0.08 (-10.646.26) [-0.330.17]	-0.12 (-1.2977.71) [-0.510.29]	17.71 (-0.11538.91) [27.1232.94]	-0.05 (-10.0712.39) [-0.440.22]	0.55 (-2.83308.19) [3.8715.26]	42.23 (2.29162.96) [56.3064.73]
$R[RMSSD vs RMSSD]$	0.398	0.806	0.884	0.822	0.336	0.893	0.608	0.291
$SDNN - SDNN$	4.99 (-0.40212.24) [26.4734.14]	-0.00 (-14.6031.84) [0.220.55]	-0.02 (-13.849.98) [-0.090.03]	-0.41 (-17.7107.16) [-1.110.78]	9.97 (-0.091284.60) [23.7228.94]	-0.07 (-14.7511.39) [-0.350.21]	0.67 (-0.531202.53) [20.06126.61]	27.80 (0.491212.24) [37.2144.12]
$R[SDNN vs SDNN]$	0.449	0.835	0.919	0.853	0.439	0.904	0.507	0.394
$SDSD - SDSD$	6.71 (-0.33363.54) [19.0726.38]	-0.01 (-9.37118.65) [0.130.54]	-0.08 (-10.666.27) [-0.340.17]	-0.12 (-1.31177.79) [-0.520.29]	17.74 (-0.11539.84) [27.1632.99]	-0.05 (-10.0812.41) [-0.450.22]	0.55 (-2.83308.68) [3.8715.26]	42.31 (2.29163.54) [56.39164.83]
$R[SDSD vs SDSD]$	0.398	0.806	0.884	0.822	0.336	0.893	0.608	0.291
$SENN - SENN$	0.28 (-0.02113.28) [1.3511.78]	-0.00 (-0.7712.14) [0.020.04]	-0.00 (-0.770.55) [-0.010.00]	-0.02 (-0.983.93) [-0.060.04]	0.54 (-0.01116.37) [1.240.54]	-0.00 (-0.810.64) [-0.020.01]	0.04 (-0.0312.54) [0.9911.30]	1.51 (0.02113.28) [2.032.38]
$R[SENN vs SENN]$	0.494	0.849	0.925	0.861	0.487	0.913	0.554	0.453
$pNN50 - pNN50$	0.63 (-5.95127.69) [0.9211.22]	0.00 (-9.8913.98) [0.000.14]	0.00 (-11.1916.56) [-0.130.00]	0.00 (-11.502.65) [-0.1510.00]	2.34 (-0.4038.94) [4.5115.54]	0.00 (-11.3911.22) [-0.1310.00]	0.02 (-7.2319.27) [0.3710.69]	1.55 (-5.95127.69) [2.3312.98]
$R[pNN50 vs pNN50]$	0.671	0.854	0.894	0.893	0.541	0.922	0.726	0.625
$TP - TP$	0.56 (-0.04176.73) [4.7716.72]	-0.00 (-2.3612.98) [0.010.04]	-0.00 (-1.980.86) [-0.010.01]	-0.03 (-2.3714.70) [-0.100.07]	1.37 (-0.01100.37) [3.844.80]	-0.01 (-2.0811.26) [-0.040.03]	0.08 (-0.05167.00) [3.384.41]	4.33 (0.0676.73) [6.8618.59]
$R[TP vs TP]$	0.401	0.826	0.916	0.844	0.391	0.895	0.467	0.327
$LF^{norm} - LF^{norm}$	-0.01 (-41.1921.40) [-2.2110.78]	-0.00 (-17.9711.29) [-0.190.01]	-0.05 (-16.743.21) [-0.380.21]	0.00 (-14.5716.84) [-0.120.01]	-0.72 (-40.3515.97) [-4.930.343]	-0.14 (-12.2713.56) [-0.590.36]	-0.02 (-37.4614.94) [-1.640.51]	-5.90 (-41.4116.51) [-8.370.703]
$R[LF^{norm} vs LF^{norm}]$	0.423	0.748	0.835	0.834	0.411	0.862	0.512	0.349
$HF^{norm} - HF^{norm}$	0.68 (-21.9178.55) [2.002.98]	0.00 (-15.0316.04) [0.060.30]	-0.06 (-9.4413.14) [-0.220.12]	0.01 (-6.9917.72) [-0.030.03]	1.01 (-26.0071.64) [2.563.88]	-0.00 (-6.734.90) [-0.050.00]	-0.01 (-30.74152.07) [-0.080.01]	7.22 (-22.7178.55) [10.2412.80]
$R[HF^{norm} vs HF^{norm}]$	0.446	0.706	0.887	0.844	0.386	0.895	0.551	0.307
$LF/HF - LF/HF$	-0.26 (-7.5711.18) [-0.950.74]	-0.01 (-4.8111.31) [-0.2110.12]	-0.00 (-2.750.98) [-0.040.02]	-0.01 (-3.6511.28) [-0.090.05]	-0.46 (-7.564.01) [-1.050.82]	-0.01 (-2.8210.91) [-0.070.04]	-0.00 (-6.2813.19) [-0.110.03]	-1.62 (-8.5211.06) [-2.090.1.85]
$R[LF/HF vs LF/HF]$	0.429	0.739	0.848	0.800	0.363	0.867	0.564	0.270
$RMSE$	21.65 (0.26317.52)	14.29 (0.27107.89)	6.01 (0.5271.86)	7.31 (0.521125.22)	37.23 (0.32152.71)	6.09 (0.2590.23)	8.79 (0.14182.03)	61.13 (6.16317.52)

Table 5.11: Deviations of HRV parameters from original, error-free, RR interval time series for model based techniques in Test-IB. Deviations of HRV parameters are represented as follows: Median (2.5 percentile | 97.5 percentile) in the upper row and the confidence interval (computed by the difference between the population medians for matched pairs, according to [22]) in the lower row, for each HRV parameter. The correlation R was computed as the Kendall's rank correlation. The RMSE and SENN are given in ms, pNN50 is also represented by the the median, the 2.5 and the 97.5 percentile. RMSSD, SDNN, SDSD and SENN are given in ms, pNN50 in %, LF^{norm} and HF^{norm} in n.u., TP in s^2 and LF/HF unit-less.

5.3 Test-2: Robustness

This section shows the results of the robustness estimation, after variable reduction, as defined in section 4.5.5. The detailed outcome of all three Test-2 cases can be seen in appendix B.3. Figure 5.19 illustrates the robustness against single ectopic beats. The best performing methods reach a robustness of 50% single ectopic beats, the highest possible value. Thus, these approaches are even suitable to correct long segments of bigeminy (every second beat is ectopic). The best approaches are: THRESH, GP-IIA, INT-1O, WAVE-F and DELETE. The deviations of INT-1O and WAVE-F are rather high and their whiskers extend over the medians of weaker performing methods. Therefore, distinctions between approaches which reach a robustness of 25 to 33% cannot be made pervasively. The two least robust methods are IPFM-C and IPFM-S.

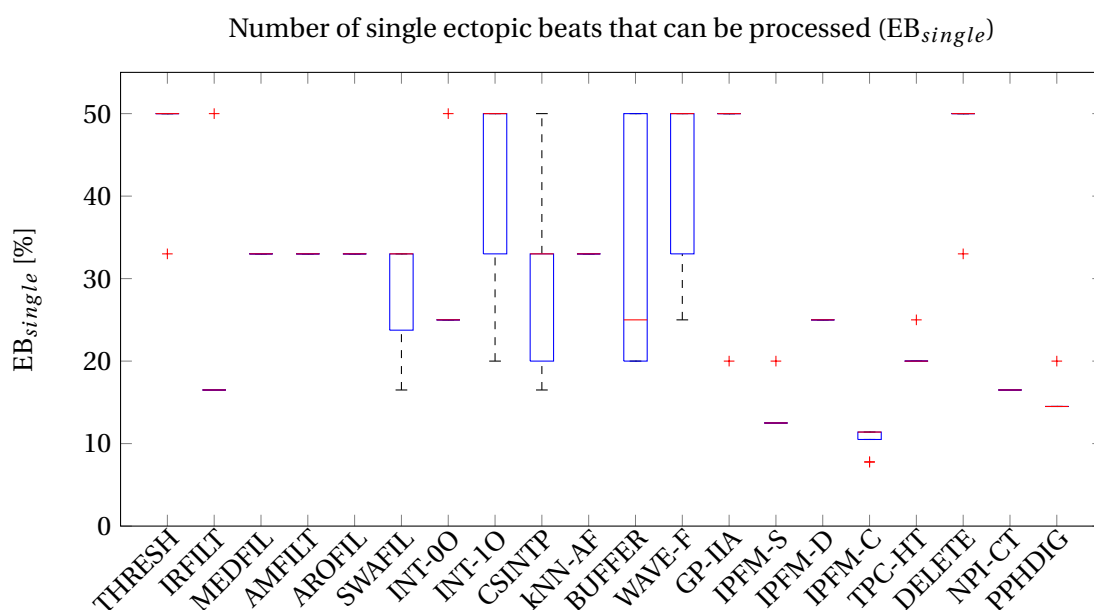


Figure 5.19: Test-2, box plot of robustness against single ectopic beats. Robustness was determined for each of the nine HRV parameters and the adjusted RMSD of the RR intervals separately, resulting in ten data points per method.

Figure 5.20 illustrates the ratio of successive ectopic beats that can be processed to the necessary number of correct NN intervals between those segments. THRESH and GP-IIA clearly outperform all other approaches. Most of the other methods show overlapping boxes due to their high deviation, and make any ranking impossible. Several methods show a ratio about one, meaning that processing of successive ectopic beats requires the same amount of correct NN intervals between ectopic segments. The IPFM-C, IPFM-D, PPHDIG model and the kNN-AF method show the lowest ratio. All of these approaches require more correct NN intervals between ectopic segments than successive ectopic beats can be reliably corrected.

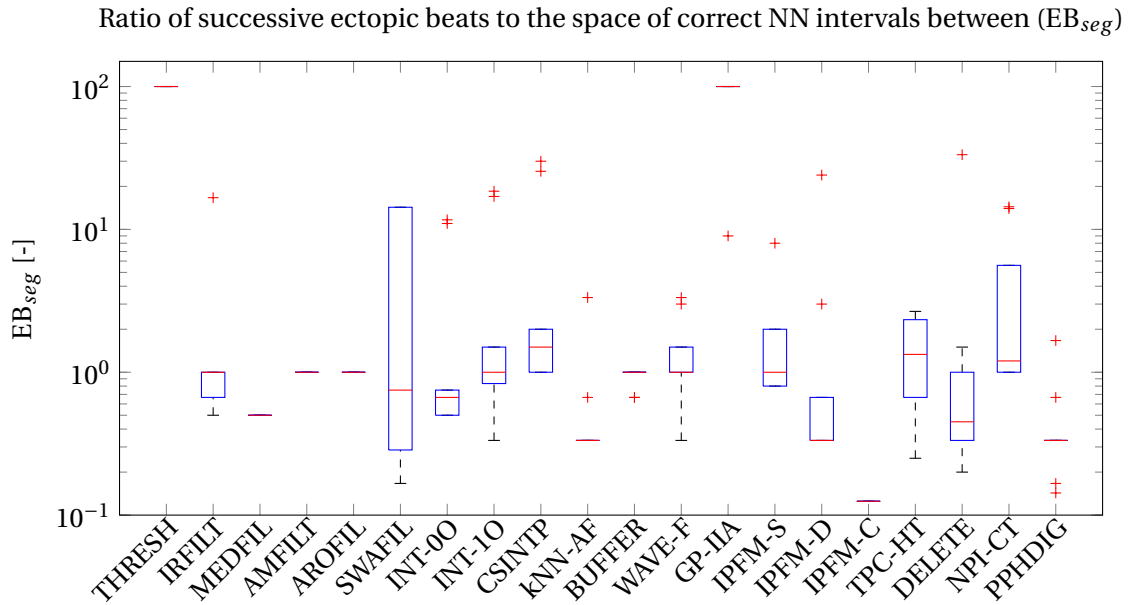


Figure 5.20: Test-2, box plot of robustness against ectopic segments. Robustness was determined for each HRV parameter and the adjusted RMSD of RR intervals, resulting in ten data points per method.

5.4 Test-3: Correction of Natural Ectopic Beats

Test-3 offers the possibility to compare the results of artificially corrupted RR interval time series to naturally erroneous signals. A comparison can only be made of statistical parameters, computation time and peak memory, but not of the deviations from the original HRV parameters since these are unknown. Further, Test-3 enabled a comparison of the correction ability of each method between artificial and natural ectopic beats. This data is not shown, since it does not provide additional information about the comparison between the different correction algorithms.

The sensitivity in Test-3 is lowest for PPHDIG, IPFM-C and IRFILT, and highest for DELETE, INT-10 and CSINTP (100 % in all three cases; see figure 5.21). The results in the sensitivity are very similar to those in Test-1B, except for BUFFER and IPFM-C (see table 5.7). Comparison with Test-1A shows a much higher deviation from Test-3 (see table 5.2). Since sensitivity is also very different between Test-1A and -1B, and Test-3 contains more complex ectopic segments, these findings just illustrate that the correction behavior is very similar between Test-3 and Test-1B.

The PPV in Test-3 is also very similar to Test-1B: Same methods yield 100 % PPV in both tests (all median filters, INT-10, CSINTP, k nearest neighbors average filter (kNN-AF), IPFM-C, TPC-HT and DELETE; see figure 5.21). Further, weak performing methods, like WAVE-F, THRESH and SWAFIL are the same as in Test-1B, but the absolute values are much

lower than in Test-1B. All the other approaches achieve a very similar PPV in Test-1B and Test-3.

Peak memory is highest for IPFM-C, followed by BUFFER, IPFM-S and PPHDIG. IRFILT, SWAFIL, THRESH and DELETE require the lowest peak memory (see figure 5.21). Test-1B shows very similar results for all approaches. The same relations are valid for the comparison of Test-3 with Test-1A (see figure 5.7). Therefore, the peak memory seems to be largely independent of the used data set for correction.

Comparison of the computation time between Test-3 with Test-1A and 1B shows that the relationship is the same in all test cases (compare figure 5.21 with figure 5.16 and 5.6). The values of the computation time in Test-3 lie between those of Test-1A and -1B, but much closer to Test-1B. DELETE, followed by THRESH and IRFILT, is the fastest algorithm, IPFM-C the slowest, in all test cases.

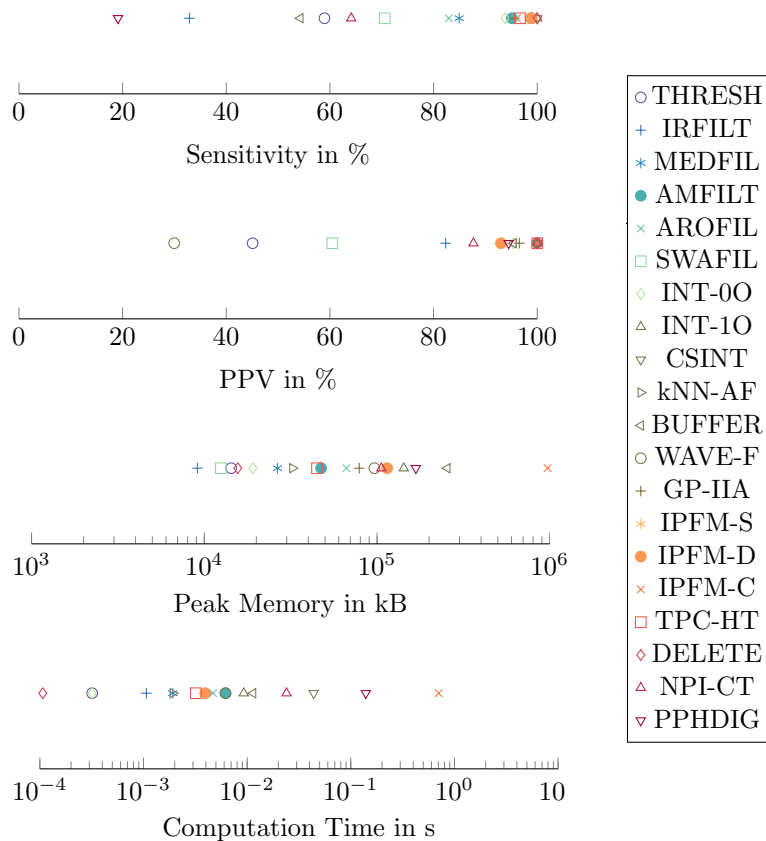


Figure 5.21: Test-3, 1-D point plot: Sensitivity, PPV, peak memory and computation time are compared by single values since only one data set with natural ectopic beats was available.

The Kivat diagrams of Test-3 show a very similar behaviour to those of Test-1B (compare figure 5.22 with figure 5.18). “Interpolation and Removal” approaches achieve the same relationships, but sensitivity and PPV are lower for THRESH in Test-3. The comparison of the “filtering” class also displays the same relationships as in Test-1B, but not in terms of absolute statistical values. All median filters accomplish a much higher sensitivity. The models perform also very similar to Test-1B despite the sensitivity of BUFFER and IPFM-C.

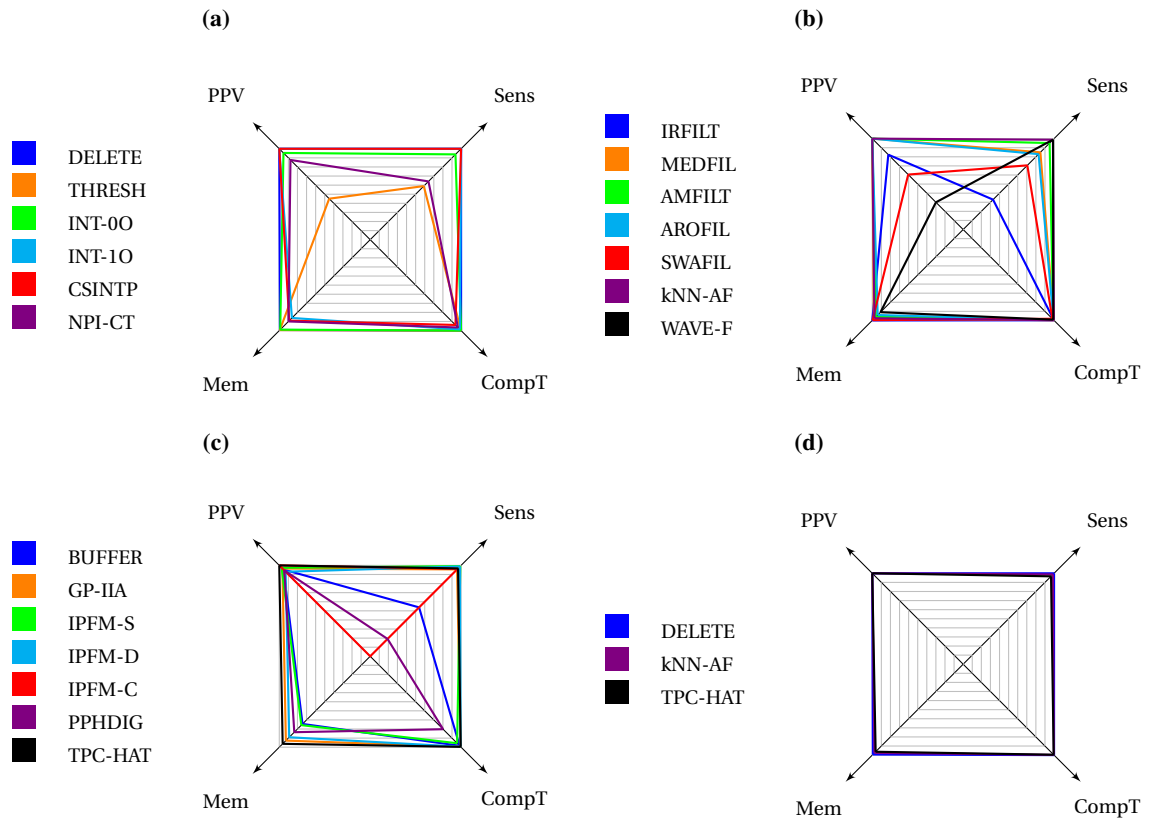


Figure 5.22: Test-3 Kivat Diagrams: Kivat diagrams are shown for the correction methods of each class separately, namely (a) “Interpolation and Removal”, (b) “Filtering” and (c) “Models”. The fourth diagram (d) compares the best methods of each class (determined by the highest mean scores of all variables) with each other. Abbreviations of the variables can be obtained from table 5.1.

Table 5.12 illustrates the reached scores of each correction algorithm in Test-3. The table is intended to give a quick overview of the mean performance of all correction approaches and is ordered in decreasing scores.

	Sens	PPV	Mem	CompT
DELETE	100.00	100.00	99.32	100.00
kNN-AF	98.75	100.00	97.58	99.74
TPC-HT	96.67	100.00	96.28	99.56
AMFILT	95.11	100.00	96.01	99.14
INT-00	93.84	95.39	98.96	99.97
INT-10	100.00	100.00	86.10	98.70
GP-IIA	95.58	96.55	92.74	99.44
CSINTP	100.00	100.00	89.08	93.80
MEDFIL	84.95	100.00	98.19	99.74
IPFM-D	98.79	93.00	89.06	99.46
AROFIL	82.94	100.00	94.04	99.35
IPFM-S	99.13	97.43	75.73	95.64
NPI-CT	64.11	87.69	89.93	96.61
SWAFIL	70.59	60.46	99.65	98.14
BUFFER	53.80	95.44	74.41	98.40
WAVE-F	98.88	29.96	90.91	99.14
IRFILT	32.89	82.34	100.00	99.86
THRESH	58.95	45.09	99.46	99.97
PPHDIG	19.00	94.45	83.50	80.20
IPFM-C	24.90	100.00	0.00	0.00

Table 5.12: Summarized results of Test-3: Correction methods are ordered in decreasing scores by their mean performance in all parameters. Abbreviations are: Sensitivity (*Sens*), positive predictive value (*PPV*), Peak Memory (*Mem*) and Computation Time (*CompT*). Sensitivity and *PPV* are given in % whereas the two other parameters denote the score with respect to the best performing (100) and worst performing (0) approach (see section 4.5.6 for a more detailed description).

Discussion

The results of all test cases have demonstrated the effectiveness of correction approaches in restoration of the original HRV parameters. Especially in RR interval time series containing only single ectopic beats, the error reduction is about one order of magnitude for most HRV parameters. Therefore, all correction methods have their justification, at least in specific error cases. This finding enhances the recommendation of the “Task Force of ESC & NASPE” that only RR interval time series free of ectopic beats should be used for HRV analysis [61]. Mietus demonstrated that unfiltered RR interval time series exhibit enormous errors in the frequency domain, especially in the *HF* (error of 93 000 %), but also in the time domain, such as the *RMSSD* (error of 5907 %) [56]. In contrast, *pNN50* and *LF/HF* seem to be only weakly affected (error of 35 % and error of 987 %) [56]. A comparison with Test-1A (about 2 % ectopic beats) shows that there is an error of 72 392 % for *HF* and 737 % for *RMSSD*. However, there is no information about the used signals or artifacts of Mietus [56]. Both studies clearly highlight the necessary correction of ectopic beats. RR interval time series containing multiple types of ectopic beats and even long lasting tachycardias are not anymore reliably corrected by all approaches, due to the inclusion of false trends. Therefore, just robust approaches or removal corrections are suitable to deal with these kinds of errors. Kamath and Fallen already suggested that long ectopic segments should be deleted, instead of replaced [32].

In the following, the results are discussed according to the test cases Test-1A, Test-1B, Test-2 and Test-3. This provides an overview of well and weakly performing correction approaches for each test specifically, with a focus on the comparison between the different approaches. Such a huge comparative study of ectopic beat correction algorithms in HRV analysis has not been performed yet, at least to my knowledge. Therefore, there is no information about the performance between the different correction classes, interpolation, filtering and models. Only two reviews compare different correction approaches, but mainly those of one class, like the one by Lippman et al. [42] and Peltola [65]. These two reviews serve as a reference to literature and are mentioned in the comparisons of the different test cases.

Section 6.5 describes each method separately in more detail regarding its advantages and disadvantages in all application possibilities and compares the outcome with the findings of the

original papers, where each method was introduced. If the results were not comparable, the correction algorithms were compared in additional tests as described in the original papers. The results are illustrated as suggested in these papers to get an easy comparison of the performance with the original algorithms. Moreover, these comparisons show that the algorithms perform very differently, depending on the specific error type.

6.1 Test-1A: Weakly Corrupted RR Interval Time Series

The comparison of all correction methods, as illustrated in the Kivat diagrams in figure 5.9, demonstrated that interpolation of degree zero and the standard median filter perform best in weakly corrupted RR interval time series. Interpolation of degree one achieves also high scores in all variables, except of peak memory. This approach uses the built-in MATLAB[®] function *interp1*, with the option “linear”. Hence, it can only be assumed that the calculation of linear interpolation is internally more complex than the calculation of the mean or the median. The two adaptive median filters, AROFIL and AMFILT, and the kNN-AF also perform very similar to the standard median filter. However, it seems that both adaptive median filters use a worse fitting median when they increase their window size, compared to the fixed window size of the standard median filter. Hence, it seems that there is no better performance in these more complex filters, but the differences to the standard median filter are very small. Since interpolation of degree zero and kNN-AF are both mean filters, it seems that there is no remarkable difference between mean and median filtering in weakly corrupted RR interval time series.

Lippman et al. compared four approaches, deletion, interpolation of degree one, cubic spline interpolation and non-linear predictive interpolation, for correction of single PVCs in 5 min RR interval time series [42]. They found out that deletion and non-linear predictive interpolation performed better than interpolation of degree one and cubic spline interpolation, since both overestimated low frequency power and underestimated high frequency power. A comparison of the frequency bands after correction of the RR interval time series in Test-1A with the same correction approaches shows no difference between any of these methods. In this thesis, the deviations of the relative high frequency power and the relative low frequency power show overlapping distributions for all four approaches (see table 5.4). However, Lippman et al. used 150 RR time interval series [42], whereas 1510 signals were used in this thesis. Further, Lippman et al. only used single PVCs, while both, PVCs and PACs were used in this thesis. These two implications may account for the observed differences between the results of Lippman et al. and the findings in this thesis. A rather similar RMSD of RR intervals (see figure 5.8 and table 5.4) in all these approaches enhances this assumption.

Approaches that are not recommended for the correction of single ectopic beats occurring at a low frequency are the threshold filter, impulse rejection filter, wavelet filter, sliding window average filter, gross positioning of beats, buffer, IPFM model with cost function and the PPHDIG model. All these methods yield scores lower than 33 % with respect to the best performing method in time- and frequency-domain HRV parameter deviations (except of the PPHDIG model that yields a score of 58 % in the frequency domain). The reason for the low performance of the threshold filter, impulse rejection filter and sliding window average filter lies in the necessary self-detection of erroneous RR intervals. This results in a relatively low PPV of these methods,

displayed in table 5.2, and causes the high deviations in all HRV parameters (see figure 5.5 for time-domain parameters and figure 5.4 for frequency-domain parameters). Further, the wavelet filter achieves the lowest PPV at all, since it changes several adjacent RR intervals when correcting just a single ectopic beat. As already mentioned, tiny changes in the frequency domain always result in a much larger alteration of the RR intervals in the time domain. Although the models, which perform very weak (gross positioning of beats, buffer, IPFM model with cost function and PPHDIG model), do not yield a very low PPV (all around 70 %), they still fail to perform a reliable correction of the RR intervals. It seems that these algorithms are not suited to correct this type of weakly corrupted RR interval time series due to the underlying assumptions. The buffer and gross positioning of beats simply replace the RR intervals adjacent to the single ectopic beat with their mean. Since Test-1A also involves ectopic beats without a compensatory pause, many correct NN intervals are falsely changed. This is also one of the reasons of the low performance of the IPFM model with cost function. The IPFM model is based on the beat occurrence times and thus can only shift beats, resulting in a change of the previous and the following RR interval. Additionally, the cost function cannot always reliably detect the minimum of the lowest cost for the beat insertion position even in the case of a single PVC (see figure 4.24). This effect can only be seen if the PVC is asymmetric, meaning that not only one beat is shifted, but all beats following the ectopic one. The performance of the PPHDIG model is more complicated to judge. On the one hand, it yields a very good PPV of 93.3 % and a very high sensitivity of 97.9 %, but on the other hand the deviations of the HRV parameters are larger than in several other approaches. These discrepancies can be best observed in table 5.6 (by a comparison of the 2.5 and 97.5 % percentiles of PPHDIG with IPFM-S) and the box plots in section B.1.1 and B.1.2. The median deviations of the PPHDIG model are rather low, but there are more outliers present towards higher deviations, compared to the IPFM-S model. This can be caused by one of the following two implications: The model corrects erroneous RR intervals in the first step of the algorithm, leading to the high statistical performance, but does not accept the corrections in the review of the second step, resulting in the high error. Alternatively, the corrections could be accepted in the second step, but the failure in a reliable correction is related to the model or its used parameters. Visual inspection of the RR interval time series after correction revealed that most corrections are reliable, but sometimes the impulse induced by a single ectopic beat is just reduced but not entirely removed.

6.2 Test-1B: Strongly Corrupted RR Interval Time Series

The comparison of the correction approaches, as illustrated in the Kivat diagrams in figure 5.18, clearly demonstrates the superiority of simple deletion of erroneous RR intervals when dealing with longer ectopic segments or RR intervals disrupted by several error types. This result is also in accordance with the literature. Peltola states that ectopic segments should be deleted instead of replaced by other RR intervals, because of the inability to recover the original heart rhythm by RR interval replacement [65]. Lippman et al. also concluded that deletion is preferable to more complicated methods when comparing different interpolation approaches [42].

Birket et al. compared the spectra of RR interval time series containing ectopic beats that were either deleted (whole 1 min segment that contains the ectopic beat) or interpolated (by means

of linear and cubic spline interpolation) [14]. They found out that interpolation increases very low and low frequency power, when compared to deletion. This phenomenon was only present in highly erroneous RR interval time series ($> 5\%$ ectopic beats). An evaluation of the same comparison of the results of Test-1B of cubic spline interpolation, interpolation of degree one and deletion results in a similar outcome. The difference in the low frequency band between cubic spline interpolation and deletion is 12.4 s^2 and -0.1 s^2 between linear interpolation and deletion. Birket et al. found a difference of about 4 s^2 for both interpolation techniques [14]. The high frequency domain cannot be compared entirely as Birket et al. used a range of 0.2-0.5 Hz [14], whereas the standard recommendation for this domain, as used in this thesis, is 0.15-0.4 Hz [61]. Birket et al. found a difference in the high frequency of about 0.5 s^2 [14], whereas it is 12.4 s^2 for cubic spline interpolation and -0.1 s^2 for linear interpolation in Test-1B. Therefore, it can be concluded that the same trends are observable for cubic spline interpolation in both studies, but not for linear interpolation. However, Birket et al. used long-term recordings (24 to 48 h recordings) [14], whereas short-term recordings (5 min) were used in this thesis. Further, Birket et al. deleted whole 1 min segments [14], instead of single RR intervals. Nevertheless, both studies found that the mean difference between interpolation (both linear and cubic spline) and deletion is almost zero for signals containing few ectopic beats ($< 1\%$), equivalent to Test-1A.

Interpolation of degree one and zero still perform very well and are the second and third best approach to correct successive ectopic beats. This relatively high performance can be related to the usage of only correct NN intervals before and after the ectopic segment for all interpolation techniques. In contrast, all filters only use direct adjacent RR intervals, with respect to the considered RR interval, no matter if they are erroneous or not. This results in the high deviations of the HRV parameters of all filters. Adaptive median filters check the suggested correction before their application and therefore discard most of them in the case of highly erroneous RR interval time series, leading to a low sensitivity. The relation of the performance of the models is similar to Test-1A, except for gross positioning of beats. Most methods were designed to deal with no more than three consecutive ectopic beats, such as the buffer and the PPHDIG model, and are therefore not suited for the error types present in Test-1B. Gross positioning of beats yields the lowest deviations in the HRV parameters after the IPFM model with the s-parameter in the model class. This is caused by the fact that the algorithm of gross positioning uses two different approaches when dealing with ectopic beats. Single PVCs and PACs are just replaced by two equally long RR intervals, whereas ectopic segments are replaced by insertion of multiple evenly spaced RR intervals, based on the mean RR interval length (see figure 4.17).

6.3 Test-2: Robustness

The robustness was divided into two separate parameters, the robustness against single ectopic beats (see figure 5.19) and the one against ectopic segments (see figure 5.20). It has to be denoted that the robustness is described with respect to the increase in error for each method specifically, in both parameters. Therefore, weak performing approaches may still have a high robustness if the error is very high but does not increase tremendously if more ectopic beats are processed. The two most robust approaches, threshold filtering and gross positioning of beats, perform

rather bad when dealing with single ectopic beats (as in Test-1A), but are much more suitable to correct highly erroneous signals (as in Test-1B). Thus, the outcome of Test-2 for the determination of the robustness is well reflected by the increase in performance of both methods in Test-1B, compared to Test-1A. Methods that show a much lower robustness than most other approaches are: kNN average filter, IPFM model with cost function and s-parameter, nonlinear predictive interpolation, impulse rejection filter and the PPHDIG model. All of these approaches also yield low scores in Test-1B (except IPFM-S), confirming their weak performance of correcting ectopic segments. The kNN-filter just uses two neighboring RR intervals for correction and is therefore unsuited to correct successive ectopic beats. The IPFM model with cost function assumes only one minimum and thus, cannot replace more than one beat. Nonlinear predictive interpolation can only correct RR time series reasonably if enough correct NN intervals are present for replacement (see figure 4.7). Further, the PPHDIG model was designed to correct no more than two consecutive ectopic beats [19]. The impulse rejection filter is designed to detect sharp impulses and cannot classify long ectopic segments as errors. All other approaches reach a very similar robustness, both against single ectopic beats and segments and cannot be ranked reliably. As a result, the robustness of those approaches is discussed individually in section 6.5.

6.4 Test-3: Naturally Corrupted RR Interval Time Series

The Kivat diagrams in figure 5.22 demonstrate that 11 approaches perform very similar (all nearly reach one of the highest possible scores in PPV, sensitivity, peak memory and computation time). Therefore, only weak performing methods are discussed. As already mentioned in section 5.4 the results of Test-3 lie between Test-1A and Test-1B, since both single and successive ectopic beats are occurring in this test set. Further, Test-3 only contains single results, in contrast to Test-1A and -1B, which consists of ten test sets each. As a result, figure 5.22 shows single values instead of the median of ten test sets. Especially in the model class it can be observed that the IPFM model with the cost function requires the highest peak memory and the longest computation time. These findings are in accordance with the results of Test-1A and -1B. For a comparison of peak memory see figure 5.7 for Test-1A and 5.17 for Test-1B, and for a comparison of computation time see figure 5.6 for Test-1A and 5.16 for Test-1B. The same holds for the relatively low performance of the PPHDIG model in peak performance and computation time. Further, the relative relations of the test approaches to each other are the same as in Test-1. The threshold filter achieves the worst statistical performance in the interpolation and removal class, whereas cubic spline interpolation and nonlinear predictive interpolation require the longest computation time. In the filtering class, the wavelet filter achieves the lowest PPV and the impulse rejection the lowest sensitivity. A comparison of the sensitivity in Test-1A (see table 5.2) and Test-1B (see table 5.7) shows similar findings. Equally, the PPV in Test-1A and Test-1B is lowest for the wavelet filter. In the IPFM model, the δ -parameter requires much less peak memory than the s-parameter (nearly half of it) and the computation time is reduced by almost one order of magnitude, which is in consensus with the literature [76].

6.5 Correction Approach Specific Results

This section describes the specific behavior of each of the 20 correction approaches separately and highlights their strengths and weaknesses. Moreover, linkages to the literature are made to compare findings in the correction behavior with both theoretical assumptions and test outcomes. In those cases where a direct comparison of the results is hardly possible, or when large differences are observed, additional tests and evaluations of the implemented methods were accomplished, according to the original papers.

Deletion offers a simple and fast possibility to remove annotated ectopic beats correctly. Additionally, it requires one of the lowest peak memories in all test cases and is also the fastest correction approach. Correction of mildly corrupted RR interval time series causes one of the smallest errors in both time and frequency-domain HRV parameters (6th best method). Salo et al. found out that interpolation (of degree zero and one) is preferable to deletion, especially when considering the frequency domain [71]. This finding is in accordance with the outcome of Test-1A, where deletion is outperformed by both interpolation techniques in the frequency domain (see figure 5.4) and performs slightly worse in the time domain (see figure 5.5). However, Test-1A contained about 2 % ectopic beats, whereas Salo et al. used about 5 % erroneous beats [71]. They mentioned that the error in the low and high frequency domain was > 5 % when using the deletion approach. In contrast, the error was 0.5 % in Test-1A. The reason could be the difference in the error density or in the study population. Salo et al. used ten healthy subjects and ten patients with a previous myocardial infarction [71], whereas 151 patients with different heart diseases, of which just 25.8 % suffered from myocardial infarction, were used in this thesis. These limitations may also explain the differences in the individual HRV parameters, although the overall relations between the interpolation and deletion methods are the same in both studies.

The comparison of the correction behavior in RR interval time series containing successive ectopic beats, such as PVC couplets or tachycardia, clearly showed that deletion outperforms all other methods. These findings are consistent with the review of Peltola, who stated that it is always recommended to delete long ectopic segments or even reject the whole signal if the error density is too high [65]. Additionally, Kamath and Fallen mentioned that deletion is the most common method for artifact removal in case of error lasting for longer periods, such as atrial fibrillation and successive ectopic beats [32]. Similarly, Lippman et al. concluded that deletion may be preferred to more complex interpolation techniques when dealing with highly erroneous RR interval time series [42].

All these findings are in accordance with the high robustness of 50 % against single ectopic beats, which means that deletion can even correct sustained bigeminy. In contrast, deletion shows one of the lowest robustness against ectopic segments. Thus, there are more correct NN intervals required between ectopic segments than successive ectopic beats can be processed. Since the error is already very low when processing single ectopic beats, this rather low robustness still does not result in a weak performance if error density is high. Peltola already concluded that deletion results in a higher error in the frequency domain HRV parameters than interpolation, if error density is high [64]. The loss of information about the original rhythm seems to be more

severe than the induction of false trends, as induced by interpolation techniques. Salo et al. already observed the same drawback of deletion and thus recommended interpolation in most cases [71]. To verify this assumption, two individual test sets of Test-2A were compared. One set contained 5 % error (every 20th beat is ectopic) and the other one 25 % error (every 4th beat is erroneous). Figure 6.1b illustrates the outcome of this comparison. As expected, deletion performs worse than interpolation techniques, especially at high error densities (25 % error means a sustained quadrigeminy, since every fourth beat is ectopic). As a result, it seems that deletion does not always perform better than interpolation techniques if error density is high. As long as enough correct NN intervals are present between the ectopic beats, interpolation still seems to be superior. In conclusion, deletion just seems to be preferable in the case of ectopic beats lasting for a longer period. However, this assumption should be verified in more specific test cases.

Threshold filtering is also a quick ectopic beat removal approach, but does not rely on annotated RR interval time series. Therefore, it has to detect errors on its own and is thus highly dependent on the used parameters for the rejection criterion. The most common rejection criterion is a fixed threshold, as suggested in various studies [56, 43, 34]. However, threshold filters that use a strict criterion may delete too many NN intervals, whereas a weak criterion may fail to delete erroneous RR intervals at all. Especially fluctuating signals are challenging to correct if a fixed threshold is used instead of a window filter. All these reasons result in a rather bad performance of the threshold filter in terms of the RMSD of the HRV parameters. On the contrary, the threshold filter is one of the best correction methods when strongly corrupted RR interval time series are considered. This result is in consistence with the findings that deletion results in a much lower error than interpolation. Both removal methods are better suited to correct highly erroneous RR interval time series, suggesting that replacement correction techniques are not able to recover the original time series well enough. Threshold filtering is also the most robust correction approach, since it shows no tremendous increase in error in any HRV parameter if error density is increased. This phenomenon could be related to the rather high error in slightly erroneous signals and is compensated if error density is high, since the probability to detect errors is also increased.

One of the drawbacks of threshold filtering is the necessary self detection of erroneous RR intervals. Kemper et. al compared three different detection methods for threshold filtering, fixed threshold, percentage change and standard deviation change with two different editing approaches, deletion and cubic spline interpolation [34]. They concluded that detection is more crucial than the choice of the correction method. This is in accordance with the outcome of this thesis that deletion performs much better than threshold filtering despite of the same correction approach.

Interpolation of degree zero and one are the simplest interpolation techniques. Interpolation of degree zero just replaces all erroneous RR intervals in one segment with the same value (the average of the adjacent NN intervals), whereas interpolation of degree one connects these adjacent NN intervals with a straight line (by replacement of the RR intervals at the original beat occurrence times). Both techniques are among the best correction approaches in all test cases,

just outperformed by deletion in the case of successive ectopic beats. Kamath and Fallen also mentioned that it is recommended to use interpolation techniques when dealing with occasional ectopic beats [32]. The extremely good performance may be mainly related to the usage of only correct NN intervals for interpolation, in contrast to most other filters and models that do not incorporate this implication. Interpolation of degree zero performs slightly better than interpolation of degree one in weakly corrupted RR interval time series and is further the best method in terms of error reduction. Interpolation of degree one requires far more peak memory and a longer computation time, both in weakly and strongly corrupted signals. When correcting RR interval time series containing ectopic segments, relations change and interpolation of degree one slightly performs better than interpolation of degree zero. The reason could be that interpolation of degree one maintains parts of the original trends, since it still preserves the slope-like rise over long ectopic segments. On the contrary, interpolation of degree zero just inserts the same RR interval over the whole ectopic segment.

Both interpolation techniques show a robustness against ectopic segments of nearly one (see figure 5.20). The median robustness against single ectopic beats is 25 % for interpolation of degree zero and 50 % for interpolation of degree one. Interpolation of degree zero uses a window to calculate the mean and relies therefore on a minimum amount of correct NN intervals. In contrast, interpolation of degree one just needs one correct NN interval before and one after the ectopic segment to connect both with a straight line. Despite the rather low robustness against ectopic segments, both methods perform very reasonable in Test-1B. The reason is that they already showed a very low error in Test-1A and even a large increase in error still does not result in the same high error as in most other approaches.

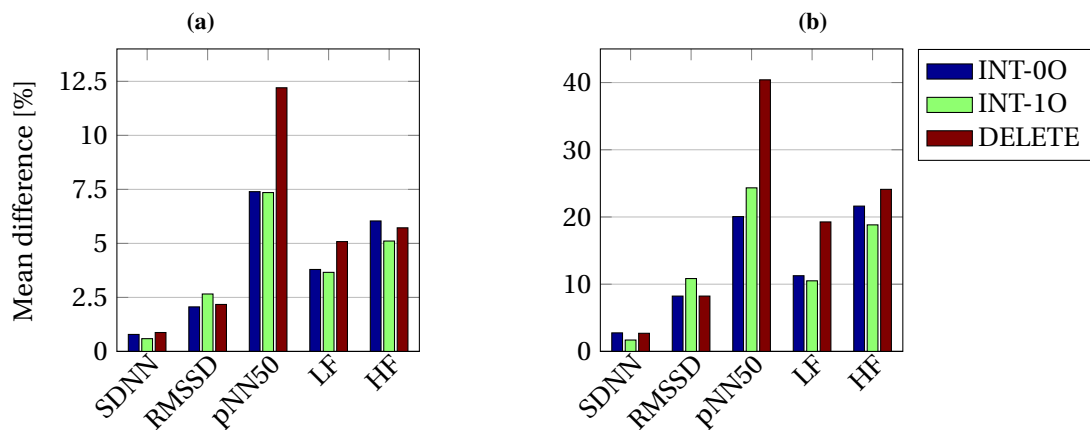


Figure 6.1: Mean difference in HRV parameters by means of interpolation of degree zero, one and deletion. (a) 5 % of the considered RR intervals were edited. Deletion performs worse in the pNN50 parameter, whereas it performs similarly in all other ones. (b) Editing of 25 % of the considered RR intervals results in a nearly four fold increase in error of all parameters in all methods. Deletion performs especially worse in the LF domain, compared to 5 % error.

Peltola found out that different RR interval editing methods have to be used for different HRV analysis, when comparing interpolation of degree one, zero and deletion [64]. Moreover, she mentions that there are differences between healthy subjects and patients suffering from acute myocardial infarction (AMI). The test set used in this thesis contains both subjects suffering from acute myocardial infarction (AMI) and those who were not. Hence, only trends can be compared. Probably this difference in the subjects is one of the reasons why there are some deviations in the different HRV parameters observed between the two studies. Additionally, the error density is very different in both test cases (Test-1A and -1B), compared to the used signals of Peltola (error density was increased in 5 % steps from 5 to 50 % [64]). She inserted errors at random sites and different densities, leading to mainly single erroneous RR intervals or couplets. Neither Test-1A, with a error density of about 2 %, nor Test-1B, containing long lasting tachycardias, can be compared to her test cases. As a result, the outcome of two Test-2A cases (one containing approximately 5 % erroneous beats and one with 25 %) was used for a more appropriate comparison. As in the thesis of Peltola, the mean difference (with respect to the original values) of five HRV parameters was calculated (see figure 6.1). Interpolation of degree zero and one perform almost identical, whereas deletion performs worse in the pNN50 parameter at a low and high error density and in the *LF* parameter just worse in case of a high error density. Peltola also found out that all methods perform rather similar in the *RMSSD* and *SDNN* parameter, especially in patients that previously suffered from AMI. She also observed a much higher deviation of the pNN50 for deletion, but also for interpolation of degree zero. Patients showed a larger error in the low frequency band, healthy subjects in the high frequency band (in both interpolation methods). In this thesis, just a higher difference in *LF* could be observed in highly erroneous signals (see figure 6.1b). As already mentioned by Peltola, differences between the study population (healthy or not) account for large variations in the outcome of different correction approaches for the same HRV parameter [64]. Nevertheless, trends between HRV parameters and the correction approaches are identical in several cases.

Cubic spline interpolation results in a slightly larger error of the HRV parameters compared to interpolation of degree zero and one, resulting in the 8th best correction approach in weakly corrupted RR interval time series. The method requires a much longer computation time and a similar peak memory, compared to interpolation of degree one. Signals containing long ectopic segments cannot be reasonably corrected anymore by cubic spline interpolation. It is the worst performing method, regarding the error of the HRV frequency-domain parameters and only yields average scores in the time-domain. Splines are smooth curves that can reasonably interpolate RR intervals over short distances but tend to overshoot physiologic fluctuations of RR interval time series when interpolating over long ectopic segments.

Kemper et al. compared cubic spline interpolation with deletion in weakly corrupted RR interval time series (less than 2 % ectopic beats) and found out that both methods perform nearly identical [34]. The outcome of the median differences in the HRV parameters in Test-1A also displays a very similar behavior in both time and frequency-domain HRV parameters. The only difference is the higher variation of most HRV parameters in cubic spline interpolation, compared to deletion. Kemper et al. used recordings from 10 newborns, 33 pediatric oncology patients and 15 healthy older adults, representing a very diverse set of subjects compared to the recordings

used in this study (see section 4.2).

Begum et al. compared “Kubios HRV”, a software that uses cubic spline interpolation for artifact correction [79], with kNN-filtering and concluded that interpolation results in a lower mean absolute deviation (MAD) than kNN-filtering in the *RMSSD* and *pNN50* [11]. This result is in accordance with the outcome of Test-1A that demonstrates the better performance of cubic spline interpolation in these two parameters (see figure B.1 and B.5).

Cubic spline interpolation performs average in terms of robustness. However, this method can only reliably correct three successive erroneous RR intervals (see figure B.19), whereas sustained tachycardias (lasting for longer than 30 s) were also used in Test-1B. Hence, cubic spline interpolation is especially error prone to long ectopic segments, but still seems to be a very reasonable correction approach in highly erroneous signals, if no more than three consecutive ectopic beats (single triplets) are processed.

Non-linear predictive interpolation performs slightly worse than cubic spline interpolation in terms of the errors of time- and frequency-domain HRV parameters in signals containing only single ectopic beats. A comparison of the box plots of the time- and frequency-domain HRV parameters demonstrates that non-linear predictive interpolation shows overlapping whiskers with interpolation of degree one and deletion (see figure 5.4 and 5.5). Lippman et al. compared all these interpolation techniques with deletion and concluded that deletion and non-linear predictive interpolation perform slightly better than interpolation of degree one and cubic spline interpolation [42]. As already mentioned, they used a much smaller set of RR interval time series, paired with the overlapping deviations in the findings of this thesis, both results are comparable. This assumption is further supported by the comparison of the correction of only single PVCs in a similar test set as described by Lippman et al. [42], where every 100th beat is a PVC. Figure 6.5 demonstrates that all interpolation methods and deletion perform very similar in all HRV parameters, except in the *HF* parameter, where non-linear interpolation performs worse. Peak memory lies in the same range as of cubic spline interpolation, whereas computation time is nearly one order of magnitude lower. Correction of successive ectopic beats results in an excessive increase in error of all HRV parameters. This finding is in accordance with the statement of Lippman et al. that non-linear predictive interpolation may only be useful if enough correct NN intervals are present and if interpolation is performed over short distances [41]. The lack of appropriate NN interval segments makes it impossible to correct long erroneous segments reliably (see figure 4.7b). Nevertheless, the error is still lower than in cubic spline interpolation. The relatively low robustness against single ectopic beats further enhances the assumption that non-linear predictive interpolation requires a specific amount of correct NN intervals to correct even single ectopic beats reliably.

Median filters may either be used with a fixed or an adaptive window size. The median filter with a fixed window length is the fastest and most computationally efficient median filter. Moreover it is the second best correction approach in weakly corrupted RR interval time series. adaptive median filter (AMFILT) and adaptive rank order filter (AROFIL) perform very similar, especially in the time domain. Correction of strongly corrupted signals results in a much weaker performance of the two adaptive filters, compared to the standard median filter.

The AROFIL performs nearly identical to the standard median filter in slightly corrupted RR interval time series, whereas the AMFILT performs a little bit worse. The reason is maybe the slightly lower sensitivity of the AMFILT compared to the other two median filters. The median filter always replaces the erroneous RR interval with the median, whereas the AMFILT and the AROFIL check the reliability of the median before they use it. If the value does not fit, they expand their window and try again. If the maximum window size is reached and no value was found, they do not apply any correction at all. Kumavarel and Santhi compared AROFIL with AMFILT, a standard median filter and a wavelet filter [36]. They concluded that the AROFIL performs best, followed by the AMFILT, especially at a high error density (tested for 10 to 60 % in 10 % steps). Test-1A shows that just the wavelet filter performs worse than the median filters, which all perform very similar. However, Test-1A contains only a error density of 2 %. To get a more reliable comparison of the filters at a high error density, the outcome of one Test-2A case (containing a sustained PVC trigeminy, meaning a error density of 33.3 %) was evaluated in the same way as described by Kumavarel and Santhi [36] (see figure 6.2a). The error is displayed as the normalized RMSD, which is the division of the RMSD of the HRV parameters by the mean of the original parameter. The wavelet filter performs worst, except in the *RMSD* parameter, whereas median filter and AROFIL perform very similar. The AMFILT performs clearly worse than the two other median filters. The weak performance of the wavelet filter is in accordance with the findings of Kumavarel and Santhi [36]. The adaptive median filter performed much better in their study and the standard median filter much worse. The reason of this behavior might be related to the used signal. The median filter always corrects the erroneous RR intervals, without any verification of the new values. In contrast, both adaptive median filters check the reliability of the new values. However, the AMFILT just uses the median if it fits and does not perform any correction otherwise. The AROFIL replaces the erroneous RR interval with the next best fitting value inside the window if the median does not fit. Therefore it seems that the AMFILT simply does not perform any correction in many cases, whereas the other two filters do. Moreover, the median filter performs better than in the study of Kumavarel and Santhi, since in a sustained trigeminy there are always two correct NN intervals (for supra ventricular trigeminy) or one correct NN intervals (for ventricular trigeminy) between the ectopic beats. In contrast, Kumavarel and Santhi mainly used sequences of ventricular bigeminy, where no correct NN intervals are present between the erroneous ones. Hence, a new test set was used that contained only bigeminy at random sites and at a random length (a modification of Test-1B, which originally contained this kind of error in addition to several other kinds of ectopic beats). Figure 6.2b illustrates the outcome of the comparison in the same way as described previously by means of figure 6.2a. As expected, the median filter performs much worse, compared to as sustained trigeminy. Further, the wavelet filter and the two adaptive median filters result in a lower normalized RMSD, especially in the high frequency power. The AMFILT performs much better when dealing with bigeminy, but is still outperformed by the AROFIL. The outcome of this test is in accordance with the results of Kumavarel and Santhi [36] and highlights the differences in the correction behavior between different error types.

The results of Test-1B further explain the better performance of the median filter with a fixed window length in signals containing ectopic segments. AMFILT and AROFIL cannot replace

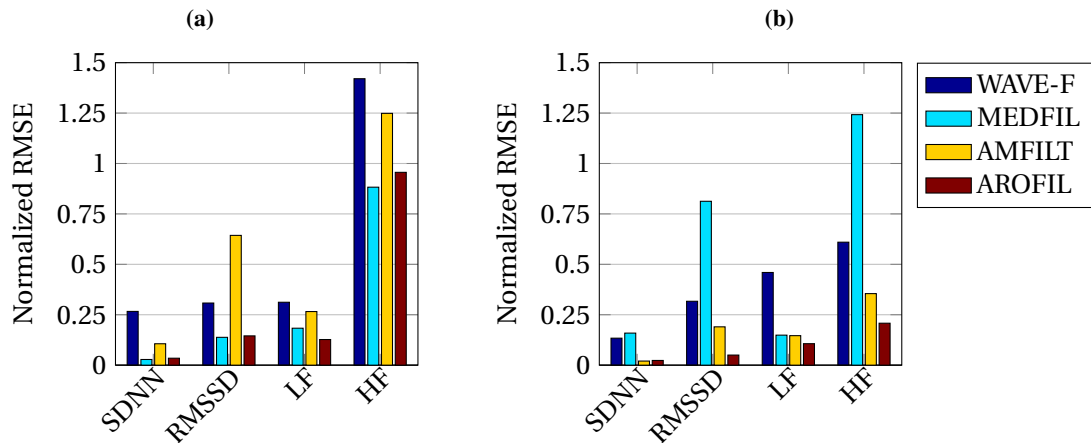


Figure 6.2: Normalized RMSD of HRV parameters in sustained trigeminy and stochastic bigeminy. (a) Sustained trigeminy: Every third beat is ectopic in the entire signal. (b) Stochastic bigeminy: Every second beat is ectopic, occurring at random sites and lasting for a random duration.

erroneous RR intervals that are located more than half the window size from the next correct NN interval (see figure 4.8). This implication shows that adaptive filtering is not necessarily always better than filtering with fixed parameters. The fixed median filter always replaces erroneous RR intervals, whereas adaptive filters only perform a correction if it is reliable. Moreover, the better performance of the AROFIL in weakly corrupted RR interval time series, compared to the AMFILT, demonstrates that an adaptive threshold (comparison to a local median) is more important than an adaptive window.

All kinds of median filters show a robustness against single ectopic beats of 33.3%, meaning that trigeminy can be still reliably corrected (robustness is compared to the error of single ectopic beats, explaining the same robustness of the AMFILT, although it shows a larger error in figure 6.2a). The limitation of the window size, paired with the implication that more correct NN intervals than erroneous RR intervals have to be present inside the window, result in the rather low robustness of median filters against ectopic segments (see figure 5.20).

In conclusion, adaptive median filtering is superior in highly erroneous signals containing ventricular bigeminy, but disadvantageous if longer lasting ectopic segments are present. Signals containing only few ectopic beats (< 2%) are reliably corrected by all three types of median filters.

Impulse rejection filter is effective in rejecting distinct impulse noise, but may fail if signal fluctuations are high or if ectopic beats do not result in large errors (see figure 4.12). This is related to the dependence of the rejection criterion on the median deviation of all RR intervals from the median RR interval of the entire record (see equation 4.11) [55]. The filter is a very fast correction algorithm, which further requires the lowest peak memory. However, these advantages are faced with a rather high RMSD of the RR intervals and the resulting high deviations

in all HRV parameters. All these findings are related to the necessary self detection of erroneous RR intervals that is highly dependent on the fluctuation of the underlying RR interval time series. The impulse rejection filter displays the lowest sensitivity (76 %) in weakly corrupted RR interval time series. In the case of a PVC, the impulse rejection filter can often only correct the RR interval before the ectopic beat or the one afterwards, if the signal shows high fluctuations. This in accordance with the findings of Liu et al., who mentioned that this filter is unable to identify PVCs, due to a too small change in the RR interval [43]. They suggested to combine the impulse rejection filter with a template matching approach to identify all common types of ectopic beats, resulting in an increase in sensitivity from 86 to 98.5 % [43]. In further studies, it seems to be advantageous to combine these two techniques for an improved performance of this filter.

The correction itself is just a median filter with a fixed window length (5 RR intervals were used in this thesis, as suggested by McNames et al. [55]). The failure to correctly identify erroneous RR intervals can also be seen in the PPV. If the error density is low, the impulse rejection filter often falsely detects correct NN intervals as erroneous, resulting in the low PPV of 43.3 %. In contrast, a high error density results in a rather low sensitivity (23.8 % in Test-1B), meaning that several erroneous RR intervals are missed in the detection process. In naturally occurring ectopic beats (where both, single and successive ectopic beats are present), sensitivity is much lower than PPV (32.9 % compared to 82.3 % PPV). A combination of these findings suggests that the impulse rejection filter should rather use an adaptive threshold for detection of ectopic beats instead of the fixed threshold as suggested by McNames et al. [55]. Otherwise, this approach is not competitive with other techniques, such as adaptive median filters or interpolation of degree zero or one. The robustness against single ectopic beats is rather low (15.5 %), since the detection of erroneous RR intervals relies on the median difference of all RR intervals with respect to the median RR interval of the entire record [55]. Further, the impulse rejection filter cannot correct more than two consecutive ectopic beats (see figure B.19). Hence, it is just slightly more robust than the standard median filter, which can only correct single ectopic beats reliably. On the contrary, the standard median filter is more robust against single ectopic beats (see figure 5.19). McNames et al. came to the same conclusion and mentioned that the impulse rejection filter obviously fails to detect several impulses, which may be modified by using an alternative threshold [55].

Sliding window average filter is one of the weaker performing methods when correcting single ectopic beats, since it has to detect errors on its own. Although this method uses an upstream threshold filter to remove outliers, it is not guaranteed that no outlier is included in the calculation of the mean. Since the mean is more error-prone to outliers than the median, this may result in the bad performance. Interpolation of degree zero is basically the best possible average filter, as it only uses correct NN intervals for mean calculation. As a result, the mean can never be biased by outliers.

Mietus suggested several other approaches to adapt standard mean filters, such as using the standard deviation or an adaptive threshold for error detection [56]. Since only the kind of filter described in section 4.1.3 was implemented, there is no information available about the other kinds mentioned by Mietus [56]. However, a median filter with a fixed window length was also

implemented in this thesis and thus, can serve as a comparison.

The sliding window average filter detects erroneous RR intervals on its own, representing an additional error source. Kemper et. al found out that detection is more crucial than correction of ectopic beats [34]. Hence, this filter faces the same problems as the threshold filter and the impulse rejection filter.

Correction of successive ectopic beats is performed much better than single ectopic beats. Although the mean filter just reaches approximately average scores (about 13th out of 20), it outperforms both adaptive median filters and reaches similar results as the standard median filter. There are two reasons that could explain this behavior: First, the filter always replaces erroneous RR intervals by the mean, in contrast to adaptive median filters, resulting in a error reduction, although the new RR intervals are still erroneous (see figure 4.13b). Second, the upstream threshold filter may already removed many erroneous RR intervals and as already mentioned, removal is much better than replacement when dealing with successive ectopic beats.

In terms of the robustness against single ectopic beats, this mean filter performs very similar to the other median filters (median robustness of 33.3 %, equivalent to a sustained trigeminy). This is related to the necessary amount of correct NN intervals in all window based filters. The number of successive ectopic beats that can be processed differs extremely between the different HRV parameters (see figure B.19). The two parts of the filter may be the reason for this performance. The threshold filter is almost independent of the number of successive ectopic beats, whereas a similar mean filter (interpolation of degree zero) can only correct two successive ectopic beats, limited by its window length.

kNN average filter is the 3th best correction method when dealing with single ectopic beats. It achieves the highest scores in the frequency-domain parameters and very high scores in the time-domain ones. In contrast, the standard median filter performs best in the time-domain parameters and reaches very high scores in the frequency domain ones. This result fits with the set-up of these two filters. The kNN filter just considers two previous or following RR intervals and replaces the erroneous RR interval with the better fitting mean. The median filter just replaces the erroneous RR interval with the median of the window, which is, at a window length of three RR intervals, either the previous or the following RR interval. Begum et al. determined the values shown in table 6.1 for the mean absolute deviation (MAD) in the HRV parameters for the kNN-AF [11]. They used three 2-minute ECG recordings, where artifacts were generated artificially by movement or shouting and generated 30 data sets, of which 27 were used for the comparison (three signals resulted in a very large error) [11]. The error density was rather low and can be compared to Test-1A in this thesis. The MAD of the same HRV parameters were calculated for just one Test-1A case (151 signals to use at least a similar amount of signals) and are also illustrated in table 6.1. Begum et al. also used self-detection of errors and thus the error is much lower (about one order of magnitude) in this thesis, compared to their results. Further, Begum et al. compared kNN-filtering with “Kubios HRV”, a software that uses cubic spline interpolation [79]. They found out that cubic spline interpolation performs better than the kNN-filter by means of a lower MAD in the *RMSSD* and *SDNN* [11]. Table 6.1 shows that cubic spline interpolation performed better or at least equal, compared to kNN filtering, in the used Test-1A set. Again, the MAD is much lower than in the study of Begum et al., due to the lack

of error detection.

	MAD_{kNN} [11]	MAD_{kNN-AF}	MAD_{Kubios} [11]	MAD_{CSINTP}
SDNN [ms]	6.567	0.105	0.778	0.106
pNN50 [%]	1.244	0.108	2.0864	0.0896
RMSSD [ms]	13.489	0.182	2.683	0.109
LF norm [n.u.]	0.019	0.002	-	0.002
HF norm [n.u.]	0.017	0.002	-	0.002
LF/HF ratio [-]	0.231	0.059	-	0.044

Table 6.1: Comparison of the MAD of HRV parameters, according to Begum et. al [11]. MAD_{kNN} and MAD_{Kubios} are the results of Begum et al. for the kNN filter and cubic spline interpolation (software “Kubios”). MAD_{kNN-AF} and MAD_{CSINTP} are the results of the Test-1A case used in this thesis.

When dealing with successive ectopic beats, the kNN filter still performs very well in statistical parameters, computation time and peak memory, but is one of the worst performing methods with respect to the changes in the HRV parameters. The reason is the set-up of the algorithm to only use two neighboring RR intervals and to not check the correction for reliability. Therefore, this filter always corrects just erroneous RR intervals, but if ectopic segments are present, it replaces erroneous intervals with the mean of them. This implication also explains the relative low robustness against ectopic segments, because of the lack to incorporate correct NN intervals (see figure 5.20). Figure 5.19 demonstrates that sustained trigeminy (two correct NN intervals between ectopic beats) can still be reliably corrected. This result is in accordance with the set-up of the filter, since two adjacent NN intervals are needed for the computation of the mean.

Wavelet filter is one of the weaker performing methods in all test cases. However, a comparison of the Kivat diagram of RR interval time series containing single ectopic beats (see figure 5.9) with the Kivat diagram of signals containing successive ectopic beats (see figure 5.18) demonstrates that this filter achieves much higher scores when dealing with successive ectopic beats. This finding is also in accordance with the behavior shown in figure 4.16, where the relative error decreases with increasing amount of erroneous RR intervals. The wavelet filter is one of the most robust approaches against single ectopic beats (see figure 5.19) and average robust against ectopic segments (see figure 5.20). This filter always yields a very low PPV, due to the huge amount of false positive corrections, in all test cases. Combined with a relatively high peak memory, because of the decomposition of the signal into wavelets, this filter cannot compete with other approaches. Keenan already came to the same conclusion when he suggested this filter for ectopic beat correction [33]. The main reason of this weak performance seems to be the false positive corrections, which emerge due to the correction in the phase space instead of the time domain. Keenan further mentioned that linear interpolation is more suitable for correction of ectopic beats than wavelet filtering [33]. The outcome of the test cases resulted in the same conclusion, both in weakly and strongly corrupted RR interval time series. Kumavarel

and Santhi compared a *db4 wavelet filter* with an adaptive rank order filter, adaptive median filter and a standard median filter [36]. Especially the adaptive rank order filter outperformed the wavelet filter in terms of a lower RMSD, the same as in Test-1A in this thesis. However, they found out that the RMSD in weakly corrupted RR interval time series (10 % artifacts) is about 4 ms, whereas it is about 8 ms in Test-1A. Since the generation of error is different in both test sets (Kumavarel and Santhi used a very prominent impulse error, inserted at random sites) and they do not specify the kind of RR interval editing used in wavelet filtering [36], the results are not entirely comparable. Therefore, an additional test set, as already described in the paragraph “Median filters”, was used to compare the wavelet filter with different types of median filters (see figure 6.2b). Although the wavelet filter is still outperformed by the two adaptive median filters, it still performs much better than the standard median filter (except in the *LF* parameter), when dealing with stochastic bigeminy. Kumavarel and Santhi found the same relationships in their study [36].

Limitations of models: Models perform weaker since they are designed to correct just specific beats types. Therefore, they rely on assumptions about the underlying signal, in contrast to more simple methods. Most models just shift beats and thus cannot correct single erroneous RR intervals, which are e.g. caused by PACs. Models are also not well suited to correct very fluctuating signals, since this fluctuations influence the assumed beat occurrence times. Further, if the sinus beat after the compensatory pause is also shifted, then the sinus rhythm is changed. Models can correct the ectopic beat partly and maintain a short increase in the RR interval time series (see figure 6.3).

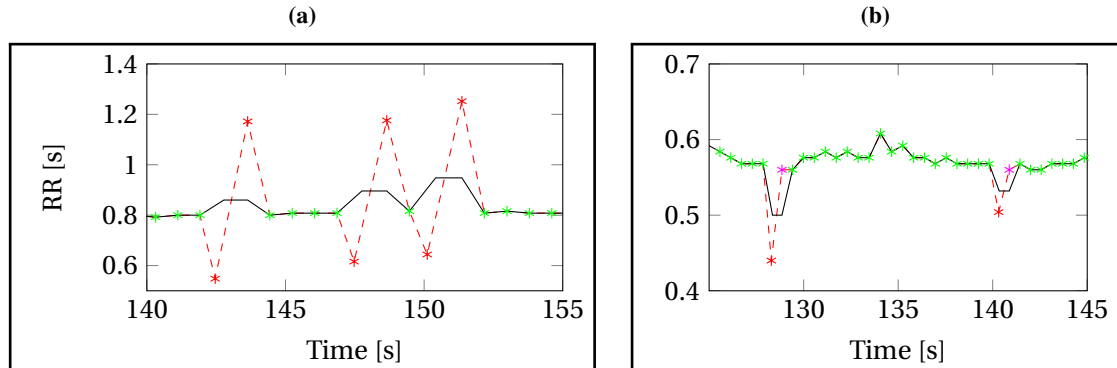


Figure 6.3: Limitations of models: (a) Ectopic beats, where the compensatory pause is longer or shorter than the forward shift of the previous beat also change the beat occurrence times of all succeeding beats. Since most models can only alter one beat, the corrected RR interval time series still shows a short increase after correction. (b) PACs exhibit the same problem, since two RR intervals are changed by most models instead of one.

Gross positioning of beats is not well suited to correct single ectopic beats, since it causes one of the largest deviations in the HRV parameters, paired with a rather high PPV (see the

Kivat diagram of Test-1A in figure 5.9). This high PPV, together with the high error in the RMSD of the RR intervals, is caused by the inability to correct PACs (see figure 6.3). Beats are shifted to the intermediate position between the previous and following sinus beat. Hence, the erroneous RR interval and the successive one are both too short. In contrast, GP-IIA is the 5th best correction approach when dealing with successive ectopic beats (see the Kivat diagram of Test-1B in figure 5.18). The reason is the alternative algorithm if more than three consecutive RR intervals are processed. Instead of shifting the erroneous beat to an intermediate position between the adjacent beats, it determines the number of RR intervals that should be inserted, based on the mean RR interval of the whole signal. Then, it inserts evenly spaced RR intervals over the whole ectopic segment (see figure 4.17). Further, figures 5.19 and 5.20 demonstrate the high robustness of this algorithm against both, single and successive ectopic beats. Gross positioning of beats is an easy and fast approach to correct whole ectopic segments, but still cannot reach the performance of interpolation techniques or deletion. However, gross positioning was suggested by Mateo et al. just for visualization of beat positioning before the actual algorithm, the IPFM model with the s-parameter, performs the correction [53]. The Kivat diagrams in figure 5.9 (Test-1A) and 5.18 (Test-1B) illustrate that this algorithm performs much better in terms of the RMSD of RR intervals and the error in the HRV parameters, but at cost of a much higher peak memory.

Citi et al. used an approach similar to gross positioning of beats: A new beat is inserted at an intermediate position between the adjacent beat times [19]. The new beat time \hat{u}_k^n is calculated as follows:

$$\hat{u}_k^n = \frac{u_{k+1} + u_{k-1}}{2} \quad (6.1)$$

Two new test sets were created to get comparable results to those of Citi et al. [19]. In one test set, every 100th beat was deleted, whereas every 100th beat was misplaced in the other set. Table 6.2 shows the comparison of the findings of Citi et al. with those of this thesis. The algorithm described by equation 6.1 resulted in a RMSD of the RR intervals of 19.69 ms in the study of Citi et al. [19]. GP-IIA showed a RMSD of 4.406 ms in this thesis for misplaced beats. Missed beats could be almost perfectly restored by the GP-IIA method. The differences in the value of the RMSD between both studies may be related to different test signals. However, the relation between the different algorithms is mainly the same in both studies. Gross positioning of beats is superior to the IPFM model with δ -parameter, but performs worse than the PPHDIG model.

A comparison of the sensitivity and PPV of gross positioning of beats with the findings of Citi et al. is not possible, since the detection feature of this algorithm was not used in this thesis. Instead, the algorithm used pre-annotated RR interval time series.

Buffer with combination rules performs nearly identical to gross positioning of beats in weakly corrupted RR interval time series (see figure 5.9), since they both rely on a similar algorithm. The only difference in Test-1A is the larger peak memory required by the buffer (see figure 5.7). In PACs, the same implications as already mentioned in gross positioning of beats arise (see figure 6.3). The buffer is designed to deal with a maximum amount of three consecutive erroneous RR intervals, resulting in a very bad performance when dealing with successive

ectopic beats (see figure 5.18). This effect can be best seen in the very low sensitivity of 8.8 % in Test-1B, meaning that most erroneous RR intervals are not corrected at all. The Kivat diagram in figure 5.22 displays the same behavior for naturally occurring ectopic beats. The buffer is one of the most robust correction methods against single ectopic beats, but performs only average on ectopic segments. Since the error is already rather high if few ectopic beats are present, this result does not mean that the buffer is suited for correcting a large amount of ectopic beats. Originally, this approach was designed to correct for movement artifacts and single ectopic beats at a high heart rate by online monitoring the HRV of soldiers in the field [69]. Although the parameters suggested by Rand et al. [69] were used and compared with each other for the best achievement, it seems that either other parameters have to be used (like locally adaptive ones instead of fixed ones) or this method is simply not usable for resting ECG recordings containing different types of artifacts. Rand et al. used a set of ECG recordings that contained about 10 % artifacts and found out that their algorithm only performs the same correction approach, as indicated by two human graders, in 49 % of all erroneous RR intervals [69]. This result may also explain the weaker performance of the buffer with respect to other correction approaches. Citi et al. compared an approach similar to the buffer, where the two beats before and after the ectopic beat are used to calculate the new beat time \hat{u}_k^n as follows [19]:

$$\hat{u}_k^n = u_{k+1} + u_{k-1} - \frac{u_{k-2} + u_{k+2}}{2} \quad (6.2)$$

The buffer was compared in the test set, described in the previous paragraph, to the findings of Citi et al. [19]. The algorithm in equation 6.2 resulted in a RMSD of the RR intervals of 16.79 ms in the study of Citi et al. [19]. In contrast, the buffer restored misplaced and missed beats almost perfectly in this thesis (RMSD < 0.000 ms). In the study of Citi et al. the PPHDIG model performed better than the algorithm described in equation 6.2. However, the algorithm of the buffer is just similar to this equation, but not identical. Therefore, the results cannot be compared entirely.

Like in gross positioning of beats, a comparison of the sensitivity and PPV of the buffer with the findings of Citi et al. is not possible because of the usage of already annotated data instead of the self detection algorithm.

Trend predict correction is the best correction model when dealing with single ectopic beats (see figure 5.9), and it is the only model that includes assumptions about the heart turbulence. The IPFM model with the s -parameter and the one with δ -parameter perform slightly better, in terms of lower RMSD of the RR intervals and a lower error in the HRV parameters (see figure 5.4 for frequency-domain HRV parameters, 5.5 for time-domain HRV parameters and 5.8 for the RMSD of RR intervals). However, trend predict correction still performs worse than the median filter and interpolation of degree zero (see figure 5.9, best methods of each class). Trend predict correction is the 3rd best performing model when correcting successive ectopic beats, outperformed by the IPFM-S model and GP-IIA (see figure 5.18). TPC-HT shows a rather low robustness and thus, seems to be outperformed by the more robust GP-IIA method. Nevertheless, a comparison with approaches of the “filtering” class shows that trend predict correction is much better suited for correction of highly erroneous RR time series than all of these approaches. In contrast, deletion and interpolation of degree zero and one clearly outperform this

approach. Test-3 illustrates further that it is the best correction model (see figure 5.22). All these findings demonstrate that trend predict correction is one of the best correction models, but never performs as well as simple interpolation or deletion techniques.

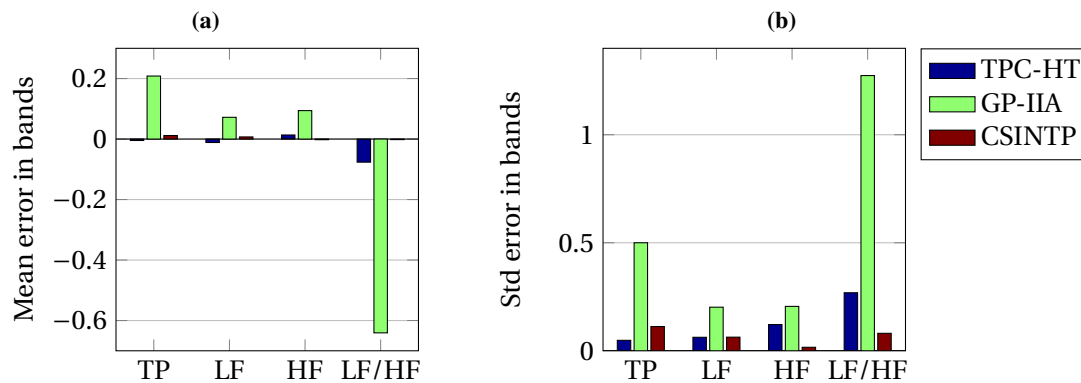


Figure 6.4: Error in frequency bands after correction of single ectopic beats (one Test-1A set). (a) shows the mean error, (b) the standard deviation. GP-IIA clearly performs worst, whereas cubic spline interpolation (CSINTP) performs slightly better than trend predict correction (TPC-HT).

Trend predict correction was compared by Wen and He with cubic spline interpolation and mean value replacement (MVR) [85] with respect to deviations in the frequency domain HRV parameters. They used one PVC and one PAC per signal as artifacts. In this thesis, GP-IIA was used for a comparison with MVR, since both techniques are identical in shifting the ectopic beat to an intermediate position between the two adjacent beats. Further, one Test-1A case was used for a comparison with their study, to use a similar amount of signals. But, two implications are observable between their set up and the one of Test-1A in this thesis. They used a simulated RR interval time series, in contrast to real data and used just one PVC and one PAC instead of several ones (about three per signal in this thesis). Despite of these differences, their data also suggested that trend predict correction performs similar to cubic spline interpolation and remarkably better than mean filtering, which is the same finding as the evaluation of the test in this thesis showed(see figure 6.4). GP-IIA (MVR in the study of Wen et al.) performed worse, since PACs cannot be reliably corrected at all.

Integral pulse frequency modulation model can be used in three different forms, with the s -parameter, the δ -parameter or with a cost function. Although these three approaches rely on the same model, they use very different methods to correct ectopic beats. Therefore, the results of them are just similar with respect to sensitivity and PPV, in all test cases (see tables 5.2, 5.7 and 5.21). The IPFM model always needs annotated beat time series, since it can only correct ectopic beats but it cannot detect them itself. Therefore, all approaches that rely on this model achieve nearly 100 % sensitivity. Since this model determines the beat occurrence times instead of the length of single RR intervals, two RR intervals are always changed if one ectopic beat

is shifted. Thus, single ectopic beats without a compensatory pause, e.g. PACs, cause a false positive correction on the following RR interval. Table 5.2 demonstrates that all three IPFM approaches yield a PPV lower than 75 % in Test-1A, since half of the ectopic beats are PACs and thus one quarter of all corrections are false positive. All other parameters differ distinctly from each other.

The cost function was the first suggested method that relied on the IPFM model. It was described by Brennan et al. and was suggested for the correction of single ectopic beats with a compensatory pause [15]. This approach is the worst performing method in all test cases because of several reasons. First, it cannot account for PACs because of the IPFM model. Second, the cost function is based on *sinc* functions that require a long computation time. Although a look-up table was generated, containing the previously computed values of the sinc functions in $\frac{1}{10000}$ steps, the computation time was still more than one order of magnitude longer than most other approaches. Third, the cost function often does not have a distinct minimum, resulting in a poor correction behavior (see figure 4.24). This is especially the case, if the compensatory pause is so long that all beats following the ectopic beat are also shifted. Fourth, the cost function is designed as a quadratic function and thus only contains one minimum. For correction of successive ectopic beats, a cost function of higher order had to be specified [15]. To get a more appropriate comparison of the different correction algorithms in a similar way as described by Brennan et al. [15], a new test set was generated. This test set contained just single PVCs at every 100th beat. Interestingly, the IPFM-C model outperforms all other approaches in almost all displayed HRV parameters (see figure 6.5). This result is in accordance with the outcome of the study of Brennan et. al, who also concluded that their algorithm performs much better than interpolation and deletion [15]. Moreover, it demonstrates that the IPFM-C method can correct single PVCs even better than interpolation of degree zero, but should not be used to correct signals where the type of artifact is unknown.

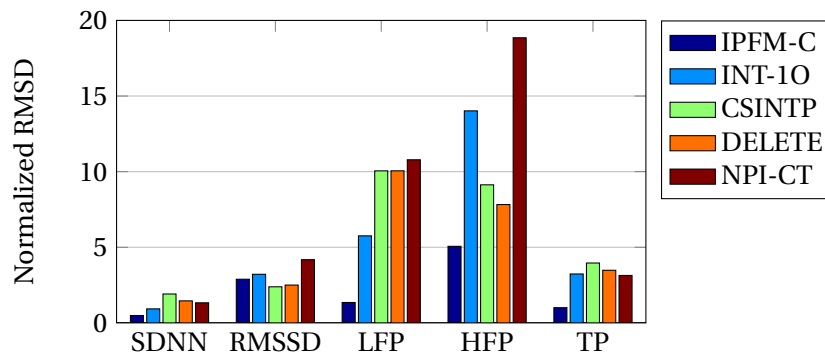


Figure 6.5: Normalized RMSD of HRV parameters of RR interval time series, where every 100th beat is a PVC. IPFM-C obviously performs best, whereas NPI-CT performs worst in HF but similar to interpolation and deletion in all other parameters.

Mateo et al. suggested the s-parameter for correction of ectopic beats and performed the correction not directly on the RR interval time series but rather on the heart timing signal [53].

They also considered the presence of ectopic segments in their parameter. The only demand on the RR time series is that there are enough correct NN intervals adjacent to the ectopic segment to perform an interpolation to calculate the s-parameter (see figure 4.21). The s-parameter is the best performing approach of all models with respect to the RMSD of the RR intervals and the HRV parameters (see figure 5.4, 5.5 and 5.8 for Test-1A and figure 5.13, 5.14 and 5.15 for Test-1B). On the other hand, it requires one of the largest peak memories in all test cases.

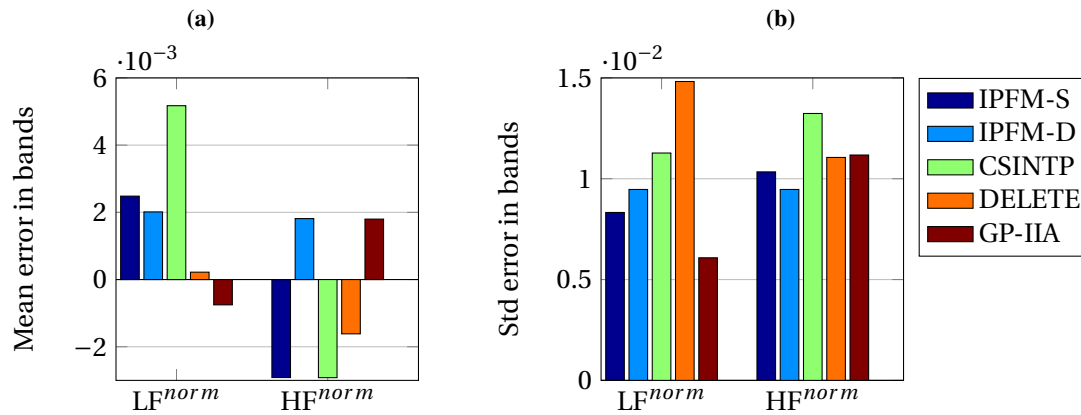


Figure 6.6: Error in frequency bands after correction of single ectopic beats (Test-1A). (a) shows the mean error, (b) the standard deviation. IPFM-S and IPFM-D perform very similar and outperform cubic spline interpolation (CSINTP). Deletion and GP-IIA perform nearly identical to the two IPFM approaches in the HF domain, but are better in the LF domain.

Solem et. al invented the δ -parameter to overcome the problem of the high memory demand of the s-parameter [76]. This new parameter resulted in a huge decrease in both, peak memory and computation time, making the method more computationally efficient. Although Solem et al. state that there is nearly no loss in the correction efficiency [76], this phenomenon could not be observed in the test cases described in this thesis. However, Solem et al. just described a comparison of spectral differences in the following three segments: An ectopic free segment before the ectopic beat(s), a segment containing the ectopic beat(s) and an ectopic free segment afterwards. Test-1A is comparable to their test set up, since only single ectopic beats are used. The difference in the RMSD of the RR intervals is rather small in this case (see figure 5.8), with respect to the deviation to the other approaches. A comparison of the deviation in the HRV parameters illustrates a much higher difference between the s- and the δ -parameter, but the box plots still indicate overlapping whiskers (see figure 5.4 and 5.5). The standard deviations of the spectral differences described by Solem et al. is also very high (see table I in [76]) and hence the deviation in the outcome may just rely on the different signals and parameters used. Test-1B demonstrates that the differences between the s- and δ -parameter increase dramatically when successive ectopic beats are corrected. Since Solem et. al just tested their parameter for single ectopic beats (at a maximum of ten ectopic beats per signal [76]), there is no compari-

son available. Moreover, Solem et al. mention that the δ_1 -parameter (the same parameter used in this thesis), just relies on the previous RR-interval and thus cannot account for long ectopic segments [76]. Taking the findings of Solem et al. and this thesis together, the δ -parameter is indeed computationally more efficient but at cost of a weaker correction ability, especially when dealing with successive ectopic beats.

To verify the statement that the δ -parameter performs similar to the s-parameter, both methods were also tested on signals containing one to two single ectopic beats that are located in the middle of the signal, as described by Solem et al. [76]. Figure 6.6 display the outcome of this test set in the same way as described by Mateo et. al [53]. Figure 6.6a displays the mean error in the relative high and low frequency power, figure 6.6b illustrates the corresponding standard deviations. Cubic spline interpolation performs worse than both IPFM models, which perform very similar, but show a contrary deviation in the mean error of the high frequency power. Further, deletion and gross positioning of beats (equally to mean value replacement used in the study of Mateo et. al) clearly show the lowest mean error in the low frequency band and a slightly lower one in the high frequency one. These findings are only partly in accordance with those of Mateo et. al. However, they did not only use different ectopic beat correction approaches, but also different methods for the computation of the power spectra [53]. Therefore, the observed differences between these two studies might be related to spectra computation. Nevertheless, it could be proofed that both IPFM approaches perform very similar if single PVCs are corrected. Citi et al. found out that the PPHDIG model performed better than the IPFM model with δ -parameter (IPFM-D) approach [19]. The comparison of these two models in the two tests described by Citi et al. resulted in the same outcome. Citi et al. calculated a RMSD of 34.59 ms for the IPFM-D model, whereas it was 3.633 ms for missed beats and 6.171 ms for misplaced beats. As mentioned previously, the much lower RMSD in this thesis is likely to be related to differences in the used signals, since the error is lower in all four correction approaches.

Point process with history dependent inverse Gaussian distribution achieves a rather low median RMSD of the RR intervals in Test-1A (see figure 5.8). However, the RMSD shows several outliers, resulting in a rather high deviation of the HRV parameters, compared to the other correction approaches. Citi et al. suggested this model for ectopic beat correction of up to two misplaced beats, but only considered missed and misplaced beats as error types in their study [19]. In contrast, all kinds of errors described in their paper were used in addition to artifacts lasting for a longer period, such as tachycardias (Test-1B in this thesis). Additionally, Citi et al. either shifted or deleted every 100^{th} beat, whereas the space between ectopic beats was smaller in this thesis (dependent on the random generator one to six ectopic beats per 200 RR intervals for Test-1A). Citi et al. used nine signals of the fantasia database for insertion of artificial errors, while 151 RR interval time series, in ten test sets (resulting in 1510 corrupted signals), were used in this thesis for both test sets.

These various differences in the set-up of the test cases may account for the different performance of the PPHDIG model. For example, Citi et. al mentioned that their model performs significantly better than the IPFM-D model [19], whereas the opposite was observed in all test cases in this thesis. To overcome these differences, two new test sets were created in the same way as described by Citi et al. [19]. Therefore, the error free signals of set N (see section 4.3,

which was used for the generation of all other test cases) were artificially altered. In one test set every 100^{th} beat was deleted and in the other set every 100^{th} beat was shifted. Then, the root mean squared deviation (RMSD) of the difference between the original beat times and the new determined ones was calculated (see table 6.2). As described previously, Citi et al. used two approaches for a comparison with their model, which are very similar to two methods used in this thesis (comparable methods are shown in the same row). The other two models, the PPHDIG and IPFM-D are identical to the approaches used by Citi et al. [19].

	RMSD [ms]			
	BUFFER	GP-IIA	IPFM-D	PPHDIG
missed	0.000	0.000	3.633	0.920
misplaced	0.000	4.046	6.171	0.000
	$\hat{u}_k^H - u_k$	$\hat{u}_k^{S'} - u_k$	$\hat{u}_k^S - u_k$	$\hat{u}_k^{\hat{P}P} - u_k$
Citi et al. [19]	18.79	19.69	34.59	14.98

Table 6.2: Comparison of estimation of unknown beat times, as described by Citi et al. [19].

A comparison of the RMSD of the new beat occurrence times to the original ones shows that the PPHDIG performs much better in this thesis. Further, all other algorithms have a much lower RMSD compared to the study of Citi et al. [19]. This is likely to be caused by the different signals used. Despite of the remarkable good performance of the buffer, the relations between the errors are similar to the findings of Citi et al. The PPHDIG model obviously performs better than the IPFM-D model, which performs worst in both studies. Further, there is a difference in the performance between the correction of misplaced beats, compared to missed beats observable. The GP-IIA approach is more suited to correct missed beats, whereas the PPHDIG model is best suited to correct misplaced beats.

Especially the better performance of the PPHDIG model, compared to the IPFM-D model rises the assumption that the PPHDIG highly depends on the kind of artifact. Visual inspection of the processed RR interval time series showed that nearly all erroneous RR interval time series are reasonably corrected by the PPHDIG model, but that three errors occur. First, if the impulse induced by a single PVC is strongly asymmetric, the beat is corrected but the RR time series shows a plateau towards the asymmetry (see figure 4.26b). Second, single PVCs are sometimes detected as PVC couplets, resulting in some false positive corrections. However, the relatively high PPV of 88 % indicates that the second phenomenon accounts for much less error than the first one. Further, it seems that the correction is sometimes rejected after verification. Signals used in this thesis were uploaded at the testing page of Citi et al. [18] to detect if specific signals result in the same correction, such as the plateaus. Indeed, the corrections showed the same trends in all correction types (for an overview of the correction types see figure 4.26).

Figure 5.18 displays that the correction behavior becomes even worse when dealing with successive ectopic beats. This is in accordance with the set-up of the algorithm, since it assumes that no more than three consecutive ectopic beats occur. This assumption is further corroborated by the low robustness, as illustrated in figures 5.19 and 5.20.

6.6 Limitations

Several limitations arose in this comparison. All algorithms were computed as described in the original papers. Although the methods were computed as efficiently as possible, they were not optimized with respect to peak memory and computation time. Further, the *profiler*, which was used as a function to determine these two parameters, is only mentioned in the undocumented help of MATLAB[®].

Many authors mentioned the parameters of their algorithms they found best performing, but some did not. No parameter study was performed in this thesis, but instead parameters were estimated by a comparison of several values with respect to the lowest RMSD of the RR intervals. This offers an possible error source. Nevertheless, the findings of all approaches are comparable to the original papers, suggesting that the parameters lie in a reasonable range.

Conclusion

Single ectopic beats are reliably corrected by all approaches used in this thesis, since all result in very good restoration of the original HRV parameters. Based on the defined test cases one can say that correction of ectopic beats is best performed by interpolation and removal approaches. Single ectopic beats, such as PVCs and PACs, are most reasonably corrected by means of interpolation of degree zero. Median filtering and kNN average filtering also achieve a very reasonable correction, but show a slightly higher error in either time- or frequency-domain HRV parameters. In the model class, just the IPFM model with the s -parameter achieves a comparable correction behavior. The reason of this weaker performance of many models lies in their design to correct only specific artifacts. A comparison of the correction behavior in signals containing only single PVCs demonstrated that the IPFM model with cost function reduces the error in HRV parameters much more than interpolation or deletion.

These differences may be caused by the fact that removal and interpolation methods do not make any specific assumptions, neither about the underlying RR interval time series, nor about the kinds of ectopic beats. Additionally, all interpolation techniques just use correct NN intervals for the calculation of the new RR intervals used for replacement, offering a big advantage over most other correction approaches.

Correction of successive ectopic beats, such as PVC couplets and tachycardia, is best performed by deletion of all erroneous RR intervals. This conclusion is in accordance with the review of Peltola [65]. Again, interpolation of degree zero and one are also very good correction approaches, whereby interpolation of degree one requires a higher peak memory, in relation to interpolation of degree zero. All filters are not suited to correct these kinds of artifacts, due to the incorporation of erroneous RR intervals in the calculation of the new RR intervals. The same is true for the models, since most of them rely on the assumption that no more than three consecutive erroneous RR intervals occur, such as the buffer and the PPHDIG model.

An evaluation of additional tests showed that deletion is not always the best correction approach if error density is high. Interpolation techniques seem to perform better if there are still correct NN intervals present between ectopic beats. Similar, the adaptive rank order filter is also very well suited for the correction of bi- and trigeminy at a high frequency. Therefore, deletion is

maybe just the best correction approach for ectopic segments, like tachycardias, couplets and triplets. To verify these assumptions additional tests should be carried out to determine the best correction approach for each specific kind of artifact.

Although Peltola et al. mentioned that there are differences in the deviations of time- and frequency-domain HRV parameters [65], this effect was not observed in any test case in this thesis. According to the findings of this thesis, the performance in the frequency- and the time-domain are always positively correlated.

It could be observed that detection of ectopic beats is more important than correction, since some approaches rely on self detection of errors, but perform the same correction method as other approaches used in this thesis. This conclusion is also in accordance with the findings of Kemper et al. [34].

To conclude, single ectopic beats should be always corrected by means of interpolation of degree zero or median filtering and ectopic segments or longer lasting successive ectopic beats should be just deleted.

List of Figures

2.1	Kinds of measurement of heart rate	6
2.2	Scheme of ECG and Pulse Wave to determine HRV	7
2.3	Geometrical measures of HRV	9
2.4	Autonomic control of the heart	11
2.5	Cardiac conduction system	12
2.6	Single ectopic beats	14
3.1	Mind-map of artifact processing methods	17
4.1	Overview of all methods implemented for ectopic beat correction	28
4.2	Deletion of ectopic beats at annotated points	30
4.3	Threshold filter	31
4.4	Interpolation of degree zero	32
4.5	Interpolation of degree one	33
4.6	Cubic spline interpolation	34
4.7	Non-linear predictive interpolation	35
4.8	Window filter	37
4.9	Median filter	37
4.10	Adaptive median filter	38
4.11	Adaptive rank order filter	39
4.12	Impulse rejection filter	40
4.13	Sliding window average filter with upstream threshold filter	41
4.14	k nearest neighbors average filter	42
4.15	Wavelet filter: Decomposition of RR time series into sub-bands	43
4.16	Wavelet filter: Correction of ectopic beats	44
4.17	Gross Positioning of Beats	45
4.18	Structure and function of buffer	46
4.19	Buffer	47
4.20	Trend predict correction	48
4.21	Determination of the s-parameter according to the IPFM model	50
4.22	Correction of erroneous RR intervals with s-parameter according to the IPFM model	51
4.23	Correction of erroneous RR intervals with IPFM δ -parameter	52
4.24	Correction of erroneous RR intervals with IPFM model and cost function	53

4.25	PDF of PPHDIG of three consecutive beats	55
4.26	PPHDIG showing different cases of correction types	56
4.27	Determination of robustness	59
4.28	Cases of single ectopic beats	61
4.29	Most common types of successive ectopic beats	62
4.30	Cases of arrhythmia	62
4.31	Determination of the adapted RMSD by means of the difference matrix	66
5.1	Eigenvalue plot of Test-1A	73
5.2	Test-1A score-score plots of frequency-domain HRV parameters	74
5.3	Test-1A score-score plots of time-domain HRV parameters	75
5.4	Test-1A box plot of HRV frequency-domain parameters	76
5.5	Test-1A box plot of HRV time-domain parameters	77
5.6	Test-1A box plot of mean computation time	77
5.7	Test-1A box plot of peak memory	78
5.8	Test-1A adjusted box plot of adapted RMSD of RR intervals	79
5.9	Test-1A Kivat Diagrams	82
5.10	Eigenvalue plot of Test-1B	87
5.11	Test-1B score-score plots of frequency-domain HRV parameters	88
5.12	Test-1B score-score plots of time-domain HRV parameters	89
5.13	Test-1B box plot of HRV frequency-domain parameters	90
5.14	Test-1B box plot of HRV time-domain parameters	91
5.15	Test-1B adjusted box plot of adapted RMSD of RR intervals	92
5.16	Test-1B box plot of mean computation time	92
5.17	Test-1B box plot of peak memory	93
5.18	Test-1B Kivat Diagrams	96
5.19	Test-2, box plot of robustness against single ectopic beats	101
5.20	Test-2, box plot of robustness against ectopic segments	102
5.21	Test-3, 1-D point plot	103
5.22	Test-3 Kivat Diagrams	104
6.1	Mean difference in HRV parameters by means of interpolation of degree zero, one and deletion	114
6.2	Normalized RMSD of HRV parameters in sustained trigeminy and stochastic bigeminy	118
6.3	Limitations of models	122
6.4	Error in frequency bands after correction of single ectopic beats (one Test-1A set) .	125
6.5	Normalized RMSD of HRV parameters of RR interval time series, where every 100 th beat is a PVC	126
6.6	Error in frequency bands after correction of single PVCs	127
B.1	Test-1A adjusted box plot of the absolute differences of <i>RMSSD</i>	144
B.2	Test-1A adjusted box plot of the absolute differences of <i>SDSD</i>	145
B.3	Test-1A adjusted box plot of the absolute differences of <i>SENN</i>	145
B.4	Test-1A adjusted box plot of the absolute differences of <i>SDNN</i>	146

B.5	Test-1A adjusted box plot of the absolute differences of $pNN50$	146
B.6	Test-1A adjusted box plot of the absolute differences of LF/HF	147
B.7	Test-1A adjusted box plot of the absolute differences of HF^{norm}	148
B.8	Test-1A adjusted box plot of the absolute differences of LF^{norm}	148
B.9	Test-1A adjusted box plot of the absolute differences of <i>total power</i>	149
B.10	Test-1B adjusted box plot of the absolute differences of <i>root of the mean squared successive differences of NN intervals (RMSSD)</i>	150
B.11	Test-1B adjusted box plot of the absolute differences of <i>SDSD</i>	150
B.12	Test-1B adjusted box plot of the absolute differences of <i>SENN</i>	151
B.13	Test-1B adjusted box plot of the absolute differences of <i>SDNN</i>	152
B.14	Test-1B adjusted box plot of the absolute differences of $pNN50$	152
B.15	Test-1B adjusted box plot of the absolute differences of LF/HF	153
B.16	Test-1B adjusted box plot of the absolute differences of HF^{norm}	153
B.17	Test-1B adjusted box plot of the absolute differences of LF^{norm}	154
B.18	Test-1B adjusted box plot of the absolute differences of <i>total power</i>	154
B.19	Box plot of robustness against successive ectopic beats	155
B.20	Box plot of necessary correct NN intervals between ectopic segments	156

List of Tables

2.1	Time-Domain measures recommended for 5 min recordings	9
2.2	Frequency-Domain measures recommended for 5 min Recordings	10
4.1	6 digit acronyms for all implemented correction methods	29
4.2	Order of combination rules for buffer	46
4.3	Baseline clinical data of set N	60
4.4	HRV time-domain HRV parameters used for PCA	64
4.5	HRV frequency-domain HRV parameters used for PCA	64
4.6	Contingency table of statistical parameters	64
4.7	Kendall’s rank correlation of time-domain HRV parameters for Test-1A	68
4.8	Kendall’s rank correlation of frequency-domain HRV parameters for Test-1A	68
4.9	Kendall’s rank correlation of time-domain HRV parameters for Test-1B	68
4.10	Kendall’s rank correlation of frequency-domain HRV parameters for Test-1B	68
4.11	Kendall’s rank correlation of robustness parameters	69
5.1	Abbreviations of variables used in the Kivat diagrams	72
5.2	Sensitivity, PPV and RMSD of HRV parameters of all 1510 signals in Test-1A	80
5.3	Summarized results of Test-1A	83
5.4	Deviations of HRV parameters from original, error-free, RR interval time series for removal and interpolation techniques in Test-1A.	84
5.5	Deviations of HRV parameters from original, error-free, RR interval time series for filter techniques in Test-1A.	85
5.6	Deviations of HRV parameters from original, error-free, RR interval time series for model based techniques in Test-1A.	86
5.7	Sensitivity, PPV and RMSD of HRV parameters of all 1510 signals in Test-1B	94
5.8	Summarized results of Test-1B	97
5.9	Deviations of HRV parameters from original, error-free, RR interval time series for removal and interpolation techniques in Test-1B.	98
5.10	Deviations of HRV parameters from original, error-free, RR interval time series for filtering techniques in Test-1B.	99
5.11	Deviations of HRV parameters from original, error-free, RR interval time series for model based techniques in Test-1B.	100
5.12	Summarized results of Test-3	105

6.1	Comparison of the MAD of HRV parameters, according to Begum et. al [11]	121
6.2	Comparison of estimation of unknown beat times, as described by Citi et al. [19].	129

Acronyms

AF	atrial fibrillation
AMFILT	adaptive median filter
AMI	acute myocardial infarction
ANS	autonomous nervous system
AROFIL	adaptive rank order filter
AV	atrioventricular
BUFFER	buffer
CSINTP	cubic spline interpolation
DELETE	deletion
ECG	electrocardiography (or electrocardiogram)
FA	factor analysis
FIR	finite impulse response
FN	false negative
FP	false positive
GP-IIA	gross positioning of beats, method IIA
HF	high frequency
HRT	heart rate turbulence

HRV heart rate variability

HT heart timing

IBI inter beat interval

ICA independent component analysis

INT-00 interpolation of degree zero

INT-10 interpolation of degree one

IPFM integral pulse frequency modulation

IPFM-C IPFM model with cost function

IPFM-D IPFM model with δ -parameter

IPFM-S IPFM model with s-parameter

IRFILT impulse rejection filter

kNN-AF k nearest neighbors average filter

LF low frequency

MAD mean absolute deviation

MEDFIL median filter with fixed window length

NN normal-to-normal

NN50 number of interval differences of successive NN intervals greater than 50 ms

NPI-CT non-linear predictive interpolation based on chaos theory

NPV negative predictive value

PAC premature atrial contraction

PC principal component

PCA principal component analysis

PDF probability density function

pNN50 proportion derived by dividing NN50 by the total number of NN intervals

PPHDIG point process with history dependent inverse Gaussian distribution

PPV positive predictive value

PSD power spectral density

PVC premature ventricular contraction

RMSD root mean squared deviation

RMSSD root of the mean squared successive differences of NN intervals

SA sinoatrial

SDANN standard deviation of the average NN interval calculated over short periods

SDNN standard deviation of all NN intervals

SDSD standard deviation of differences between adjacent NN intervals

SENN standard error of the mean

sVEB supra ventricular ectopic beat

SWAFIL sliding window average filter

THRESH threshold filter

TINN triangular interpolation of NN interval histogram

TN true negative

TP true positive

TPC-HT trend predict correction considering the heart turbulence

ULF ultra low frequency

VEB ventricular ectopic beat

VLF very low frequency

WAVE-F wavelet filter

Detailed Results of Absolute Deviations of HRV Parameters

This appendix contains the detailed results of the absolute deviations of the time- and frequency-domain HRV parameters of Test-1A and Test-1B. In addition, the outcome of Test-2B and Test-2C is also shown in this chapter. The diagrams are displayed for a comparison of the new variables (PC scores), which are used in chapter 5 for result description, with the original ones. Since the statistical parameters specificity, negative predictive value and accuracy are not used for the comparison of the correction approaches, they are not shown in this appendix. At the beginning of each section there is a short description of the summarized results of all diagrams in the actual section (either time or frequency domain or robustness) but the detailed evaluation of the results is given separately.

B.1 Specific Results of Test-1A, Single Ectopic Beats

The specific results of Test-1A contain all time- and frequency-domain HRV parameters mentioned in table 4.4 and 4.5.

B.1.1 Time-domain HRV Parameters

Basically, all time-domain HRV parameters show the same relationships, except of the $pNN50$ parameter. This result is maybe related to the similarities in the calculations of these four time-domain parameters. All are based on the error and/or the standard deviation of all RR intervals. In contrast, the $pNN50$ is calculated by the number of RR intervals that deviate more than 50 ms from the previous one by division of the total number of intervals.

Figure B.1 demonstrates that all interpolation methods, median filters and deletion perform best by means of the lowest absolute deviation in the $RMSSD$. Many other correction approaches yield in a median error about one order of magnitude higher, but the deviations inside every correction approach are rather high (in most cases the 25 and 75 percentile differ by more than

one order of magnitude). Therefore, it is not possible to make a precise ranking in average performing methods. Weak performing methods, like THRESH, IRFILT, IPFM-C, SWAFIL, buffer (BUFFER), gross positioning of beats, method IIA (GP-IIA) and wavelet filter (WAVE-F) result in a nearly 10-fold error.

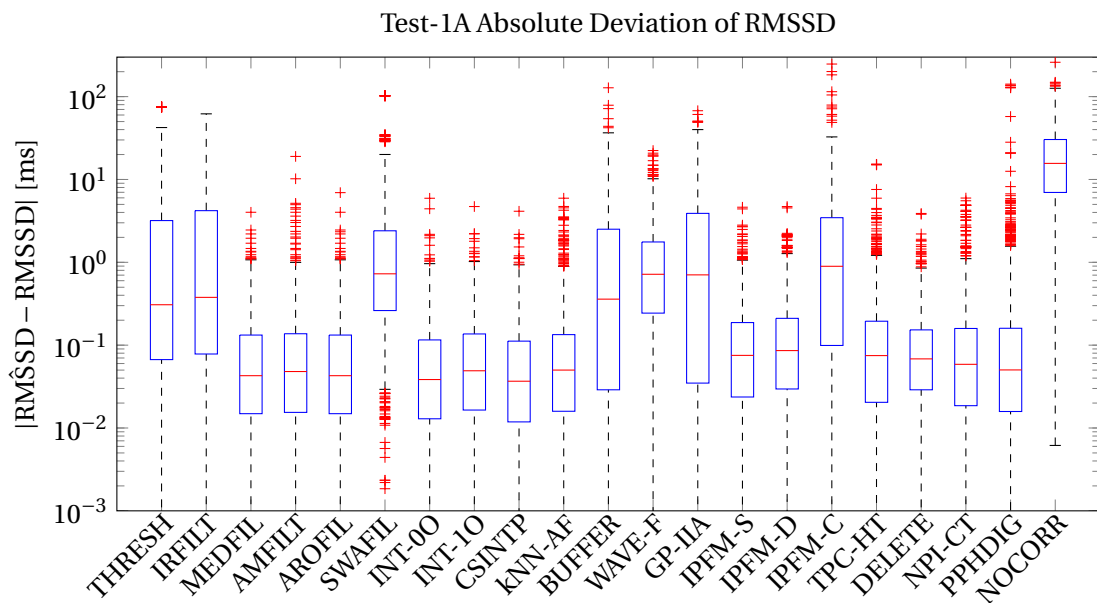


Figure B.1: Test-1A adjusted box plot of the absolute differences of RMSSD. The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1A cases, containing 151 signals each.

The error of the *SDSD* parameter is basically the same as in *RMSSD* (compare figure B.2 to B.1).

The absolute deviation of the *SENN* parameter is also very similar to the one in the *RMSSD* and the *SDSD* parameter, but there are some alterations (see figure B.3). First, the overall error in the *SENN* is about one order of magnitude smaller, compared to the other two parameters. This is related to the normalization of the *SENN* ($SENN = \frac{SDNN}{\sqrt{N}}$, where N is the number of NN intervals). Second, deletion and threshold filtering cause a larger error in the *SENN*, with respect to the other correction approaches.

The relations in the error of the *SDNN* parameter are the same as in the *SENN* case (see figure B.4). However the error is one order of magnitude larger and thus in the same range as in the *RMSSD* and *SDSD* parameter.

The error of the *pNN50* parameter is different from the other time-domain HRV parameters (see figure B.5). Although most relations look similar to that of the *RMSSD* and *SDSD* parameters, there are less deviations between the correction methods present. This is the only parameter, where nearly all correction methods show an error in the same order of magnitude. Further, deletion results in the lowest error and SWAFIL in the largest error.

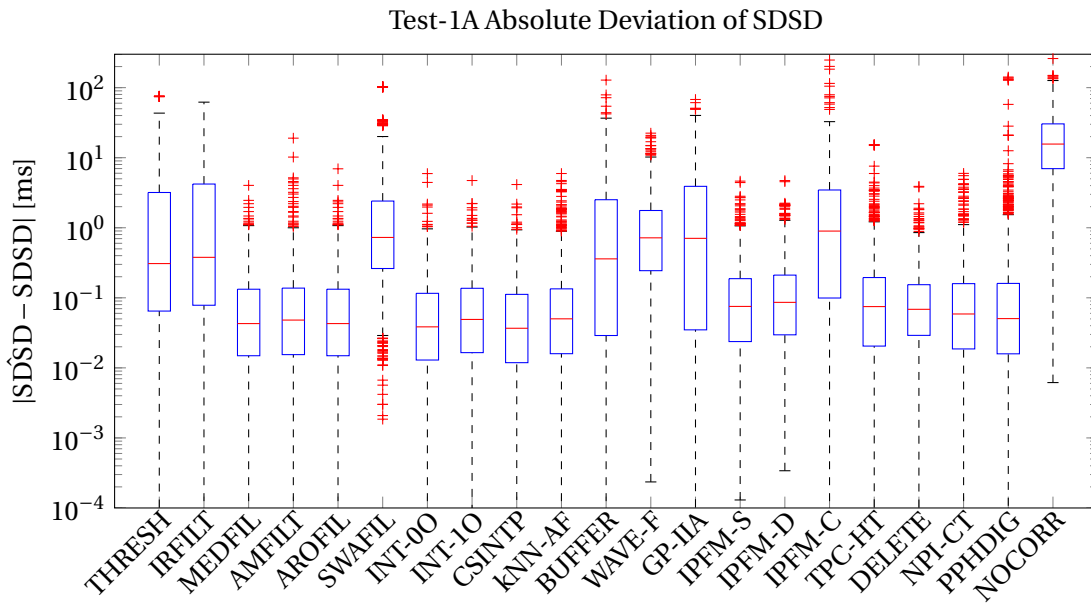


Figure B.2: Test-1A adjusted box plot of the absolute differences of SDDS. The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1A cases, containing 151 signals each.

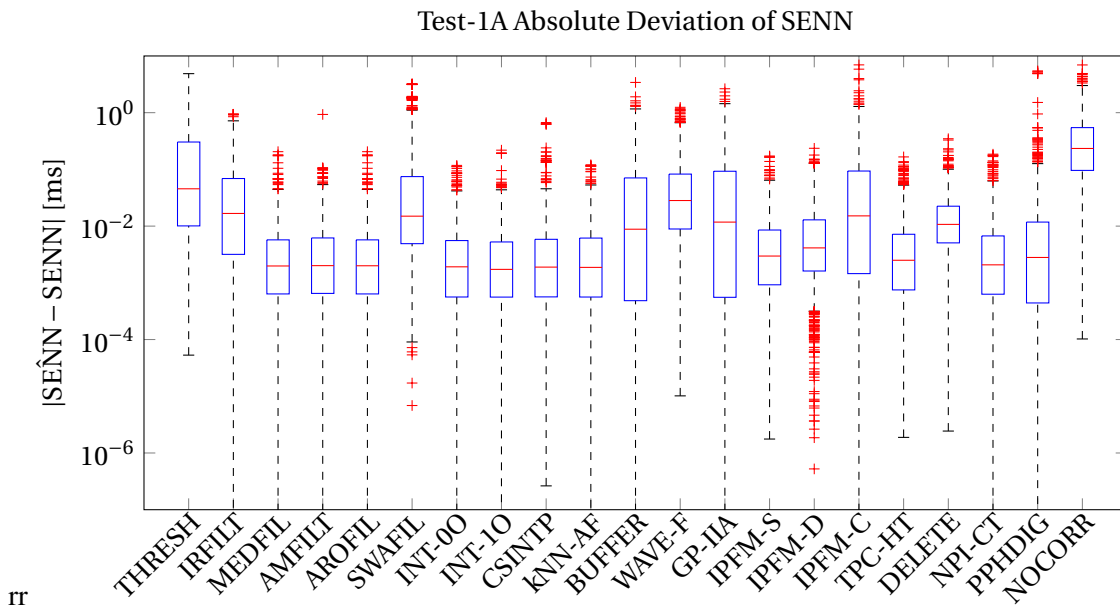


Figure B.3: Test-1A adjusted box plot of the absolute differences of SENN. The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1A cases, containing 151 signals each.

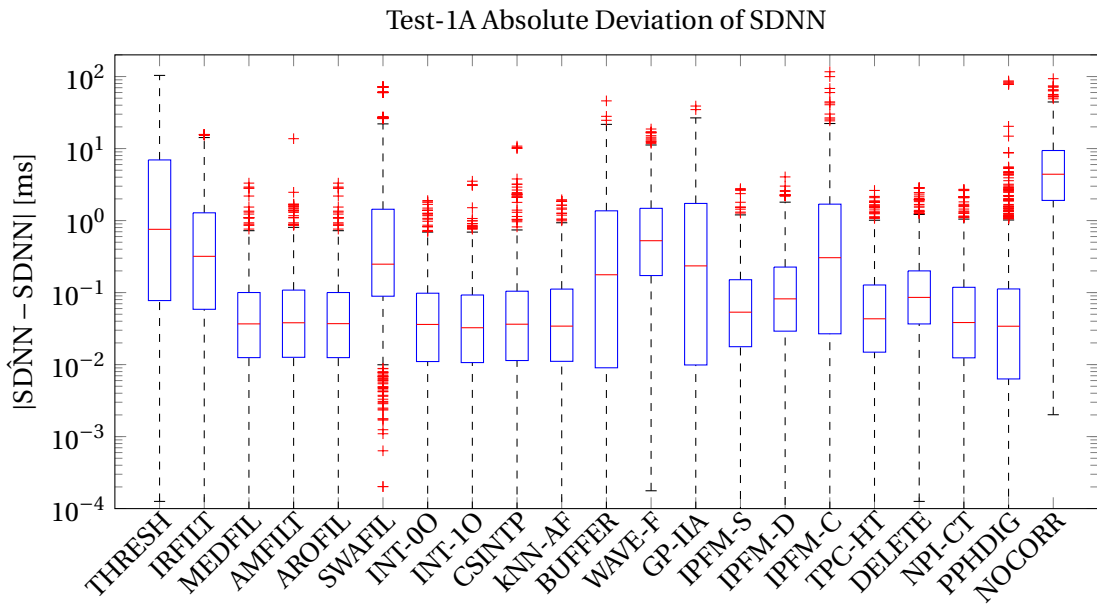


Figure B.4: Test-1A adjusted box plot of the absolute differences of SDNN. The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1A cases, containing 151 signals each.

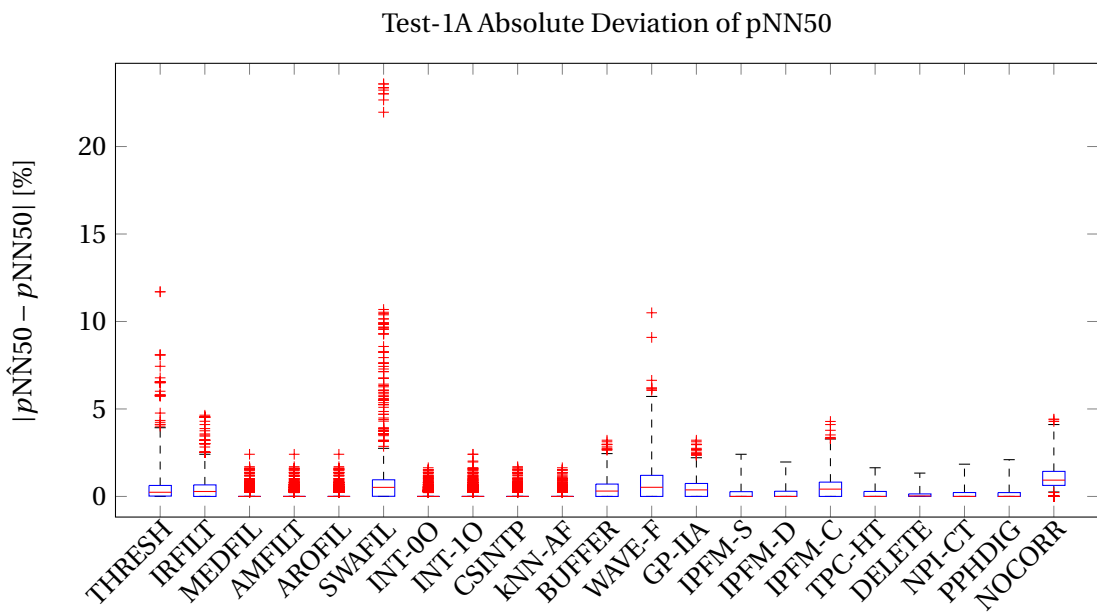


Figure B.5: Test-1A adjusted box plot of the absolute differences of pNN50. The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1A cases, containing 151 signals each.

B.1.2 Frequency-domain HRV Parameters

The errors in the frequency-domain HRV parameters do not show the same relations, as in the time-domain parameters, but the trends are very similar. Again, all interpolation techniques, all median filters and deletion cause the lowest error in all parameters, whereas IPFM-C, GP-IIA, WAVE-F, SWAFIL, THRESH, IRFILT and BUFFER are always among the weakest performing methods.

The error of the low to high frequency ratio is very high, compared to all other HRV frequency-domain parameters (see figure B.6). Interpolation of degree zero and one perform best, whereas IPFM-C, GP-IIA, wavelet filter and buffer show the largest error. Most other approaches demonstrate largely overlapping areas and thus cannot be ranked.

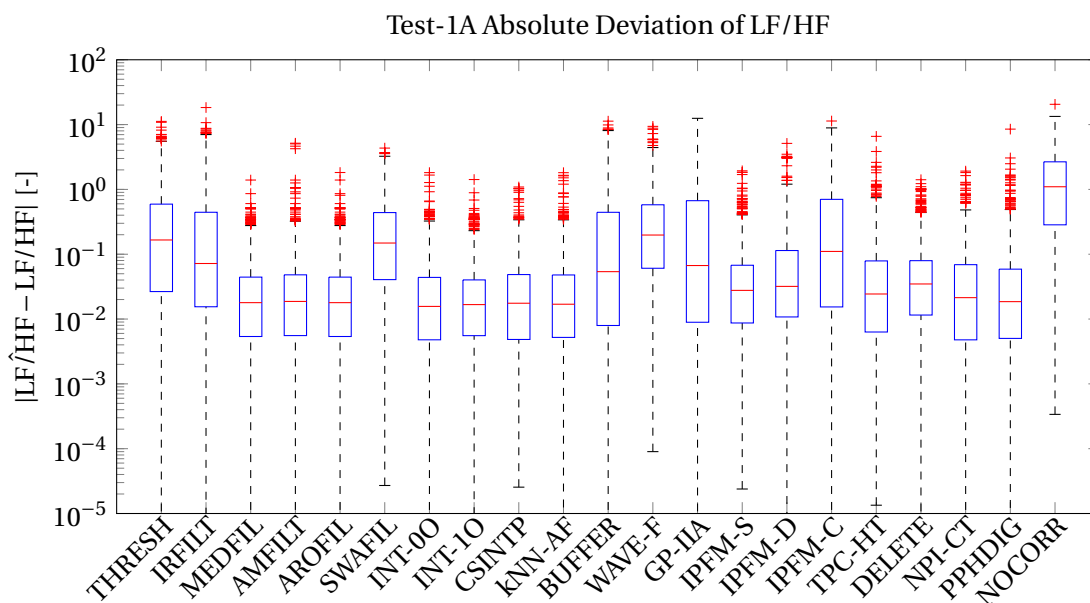


Figure B.6: Test-1A adjusted box plot of the absolute differences of LF/HF. The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1A cases, containing 151 signals each.

Most relations in the absolute deviations of the relative *HF* power are similar to the those in the *LF/HF* ratio (see figure B.7). However, the overall error is about one order of magnitude smaller.

The deviations between the different correction approaches is slightly smaller in the case of the absolute deviation of the relative *LF* power (see figure B.8) and are hence sometimes different to those of the high frequency power. Interestingly, the wavelet filter performs worst, even worse than the IPFM-C model, which performs worst in all other frequency-domain parameters. This is in accordance with the finding that the wavelet filter changes a huge amount of FP NN intervals, resulting in alterations in the low frequency domain.

The absolute deviation of the *total power* shows a similar relationship to the *LF/HF* ratio and the relative high frequency power (see figure B.9). Threshold filtering, sliding window average

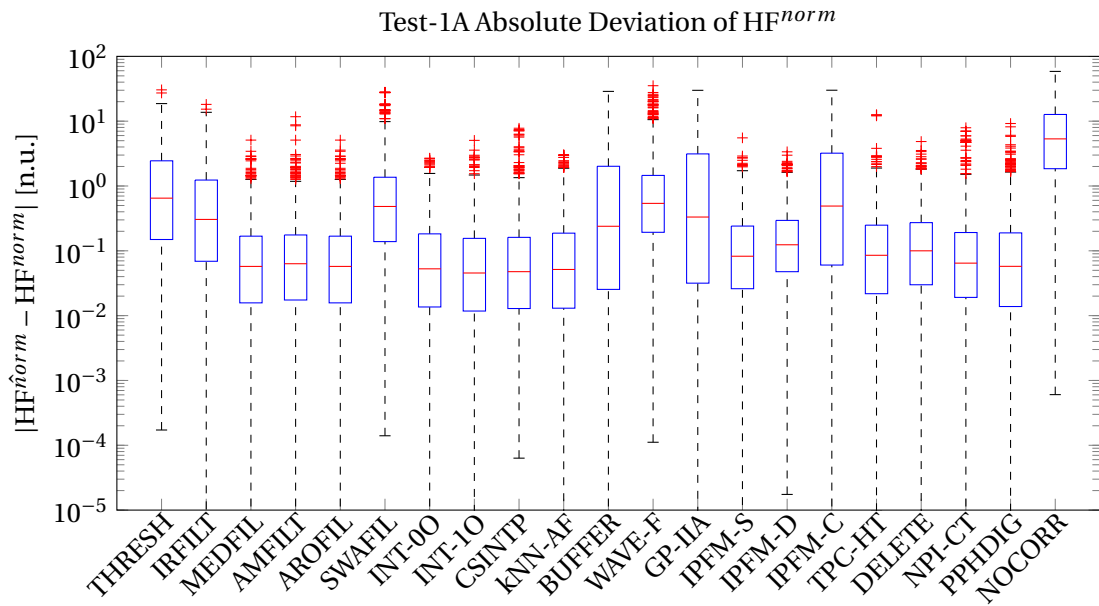


Figure B.7: Test-1A adjusted box plot of the absolute differences of HF^{norm} . The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1A cases, containing 151 signals each.

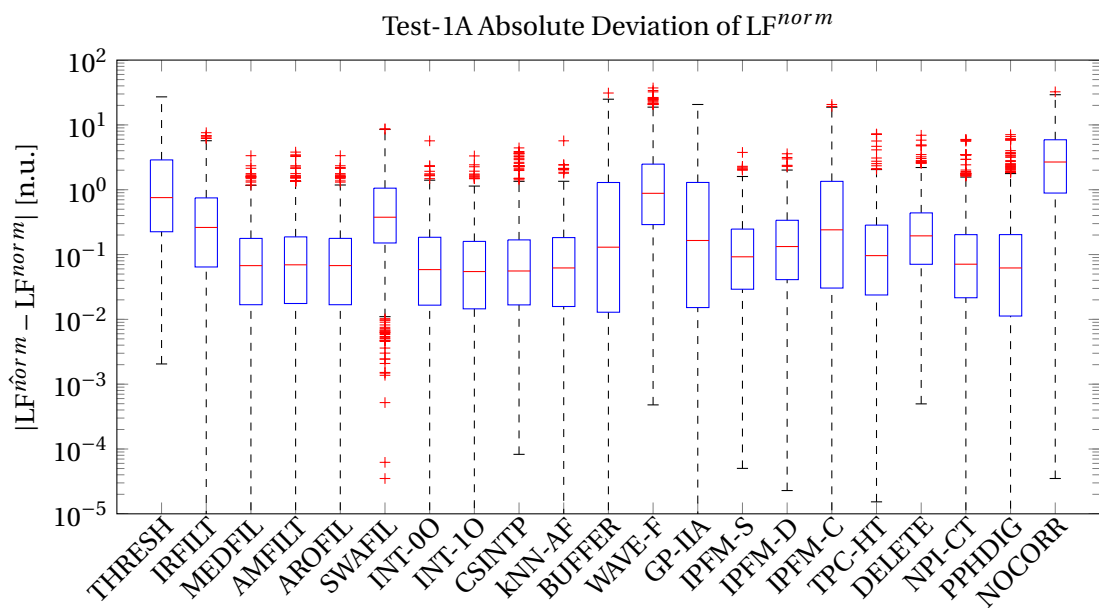


Figure B.8: Test-1A adjusted box plot of the absolute differences of LF^{norm} . The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1A cases, containing 151 signals each.

filter and the PPHDIG model perform worst. Several approaches yield very similar low errors in the *total power* deviation and thus cannot be ranked.

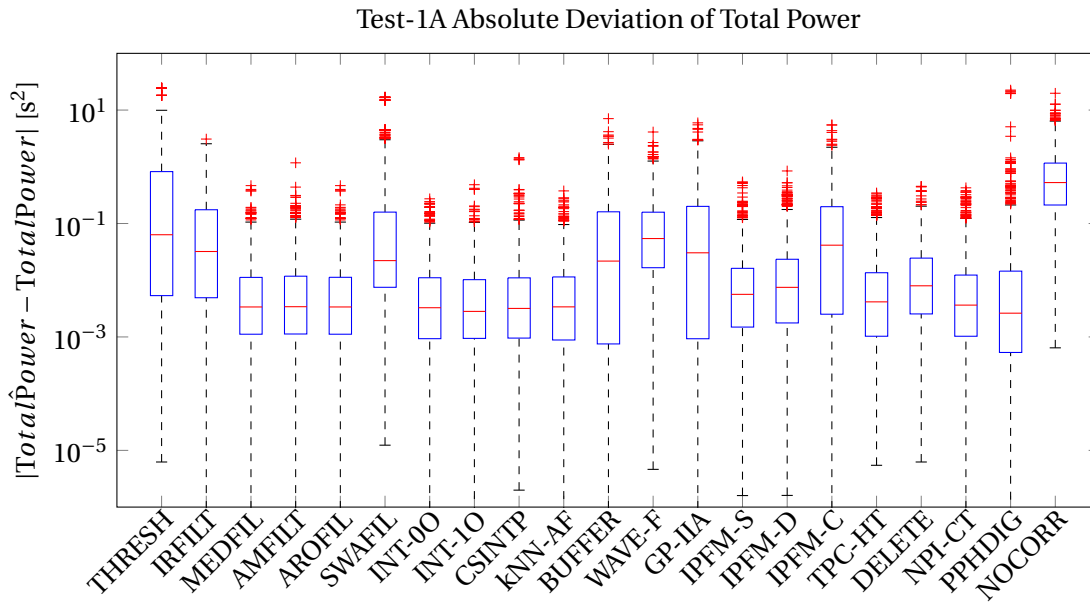


Figure B.9: Test-1A adjusted box plot of the absolute differences of total power. The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1A cases, containing 151 signals each.

B.2 Specific Results of Test-1B, Multiple Ectopic Beats

The specific results of Test-1B contain all time- and frequency-domain HRV parameters mentioned in table 4.4 and 4.5.

B.2.1 Time-domain HRV Parameters

Like in Test-1A, the absolute deviation of the *RMSSD* and the *SDSD* are almost identical (see figure B.10 and B.11). Deletion outperforms all other approaches but all interpolation techniques still perform very well. The same holds for the IPFM-S model. Interestingly, also the wavelet filter and the gross positioning of beats approach achieve a much lower error than in Test-1A, respectively.

Figure B.12 illustrates a quite different behavior in the absolute deviation of the *SENN* parameter, compared to the *RMSSD* and *SDSD*. Interpolation of degree zero and one and IPFM-S are the best approaches. Although there are again several methods that cannot be distinguished from each other, cubic spline interpolation shows the largest error due to several outliers. This is maybe related to the overshoot of some splines if error density is too high. Figure B.13 demonstrates the same effect in the case of the *SDNN* parameter. Both parameters consider the

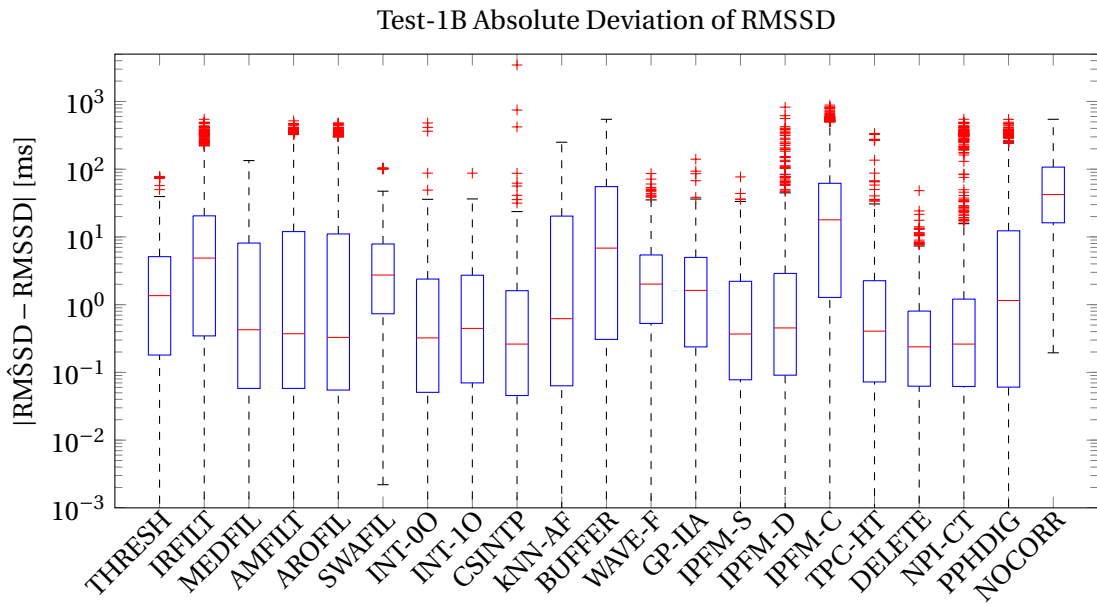


Figure B.10: Test-1B adjusted box plot of the absolute differences of RMSSD. The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1B cases, containing 151 signals each.

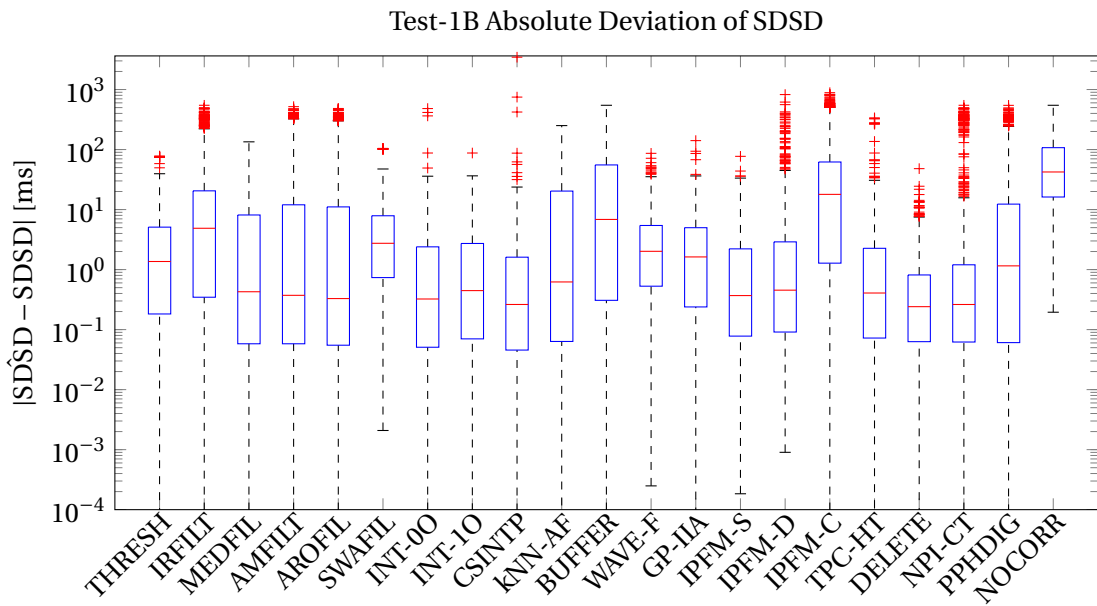


Figure B.11: Test-1B adjusted box plot of the absolute differences of SDDSD. The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1B cases, containing 151 signals each.

standard deviation of all RR intervals , whereas the $RMSSD$ and the $SDSD$ are based on the standard deviation of the error and the standard deviation of the standard deviations.

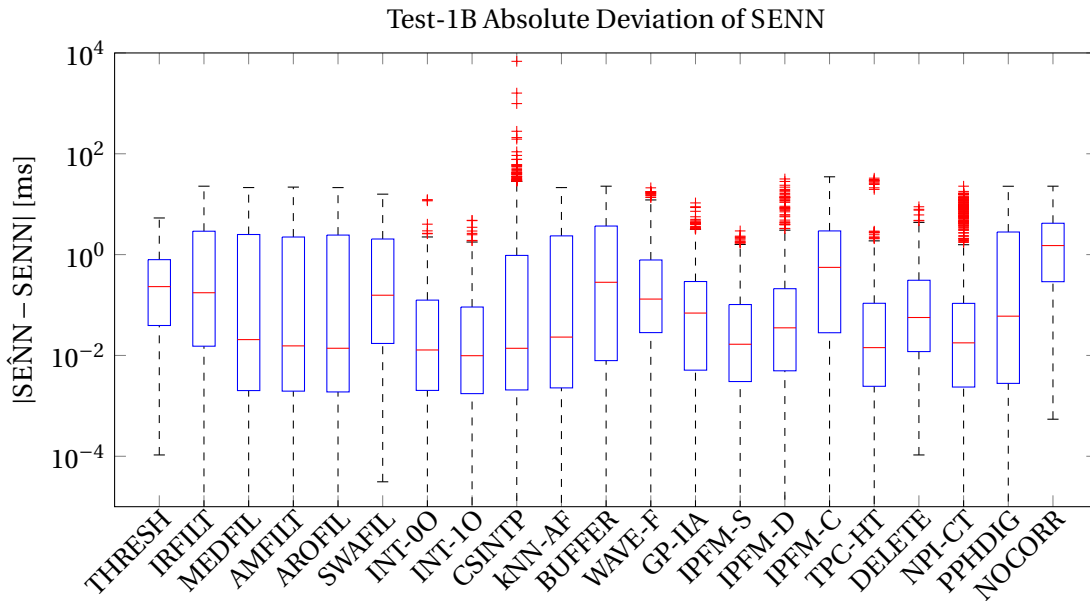


Figure B.12: Test-1B adjusted box plot of the absolute differences of SENN. The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1B cases, containing 151 signals each.

Figure B.14 displays that the error of the $pNN50$ parameter is very similar in several correction approaches. However, deletion clearly results in the lowest error, while the IPFM-C model causes the largest one.

B.2.2 Frequency-domain HRV Parameters

Most correction approaches perform very similar to the time-domain. For example, the LF/HF parameter shows very similar results for more than half of the approaches (see figure B.15). Just interpolation of degree zero and one perform better than all other methods. Cubic spline interpolation results in the largest error, because of several outliers.

The differences between the correction approaches are slightly larger for HF^{norm} , compared to LF/HF (see figure B.16). Therefore, it is hard to rank the best performing methods. Deletion, interpolation of degree zero and one and the IPFM-S model achieve the lowest median deviations, which all are one order of magnitude lower than the worst performing methods.

In contrast to HF^{norm} , the differences between the correction methods are slightly smaller in LF^{norm} (see figure B.17).

Figure B.18 demonstrates that the majority of the correction approaches perform similarly for the *total power* in Test-1B. Cubic spline interpolation clearly results in the largest error, whereas deletion, interpolation of degree zero and one, the IPFM-S model and GP-IIA have the lowest error.

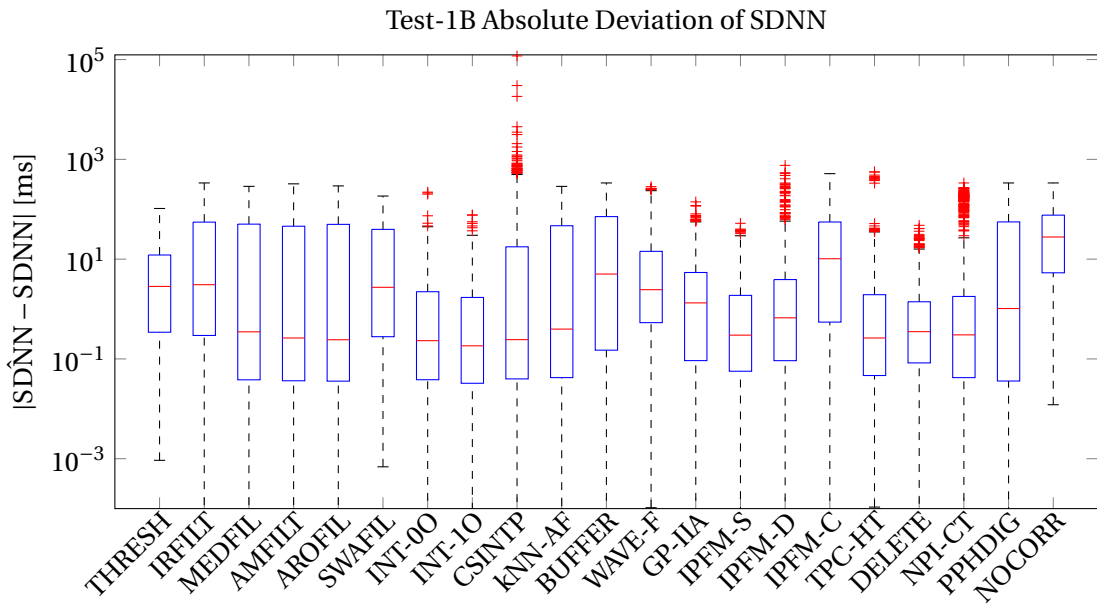


Figure B.13: Test-1B adjusted box plot of the absolute differences of SDNN. The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1B cases, containing 151 signals each.

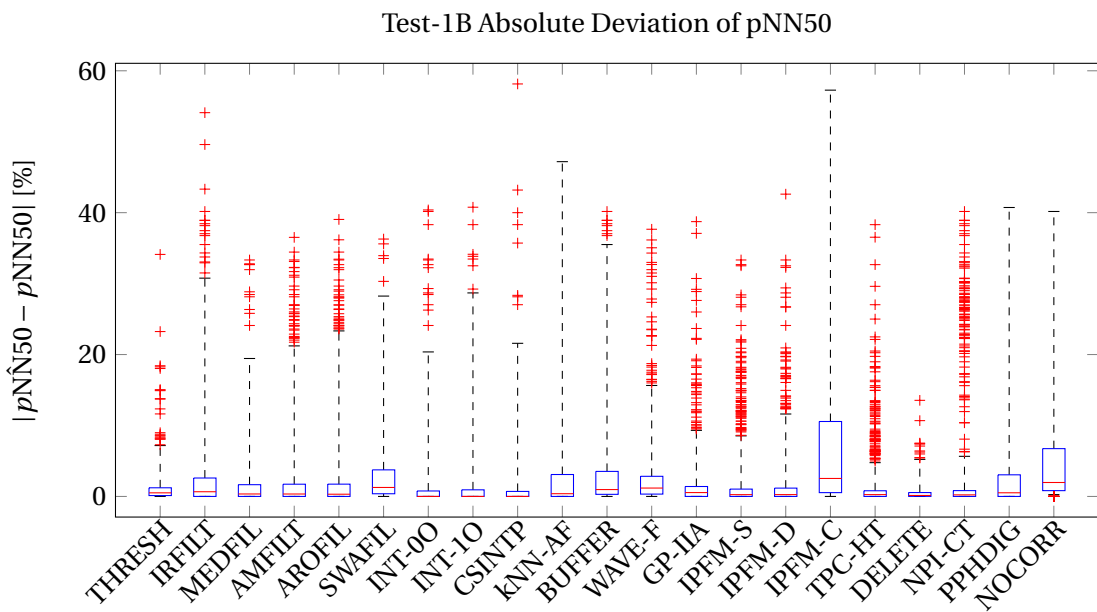


Figure B.14: Test-1B adjusted box plot of the absolute differences of pNN50. The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1B cases, containing 151 signals each.

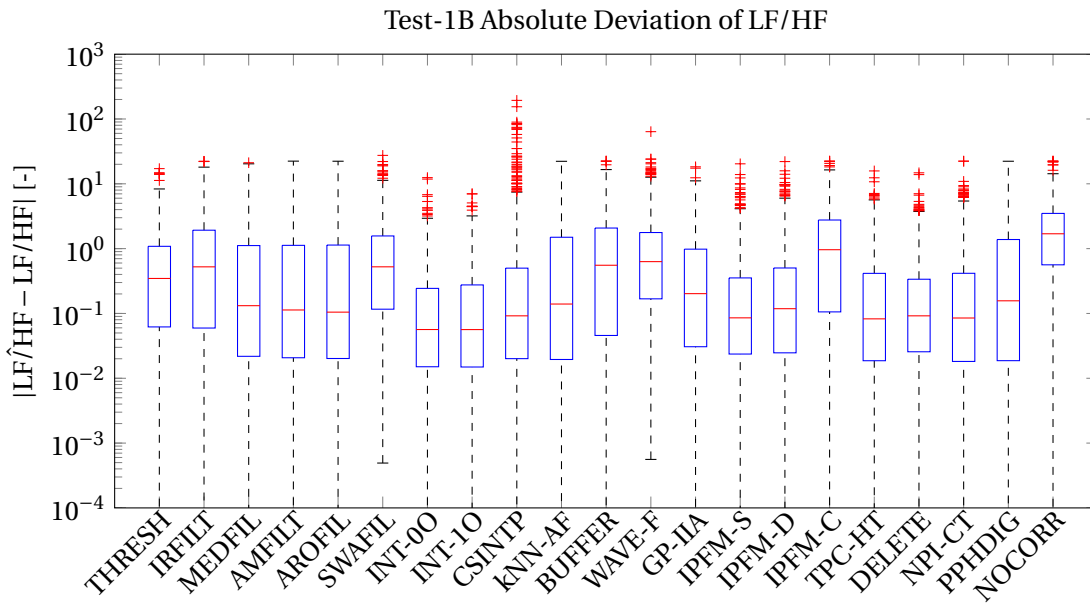


Figure B.15: Test-1B adjusted box plot of the absolute differences of LF/HF. The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1B cases, containing 151 signals each.

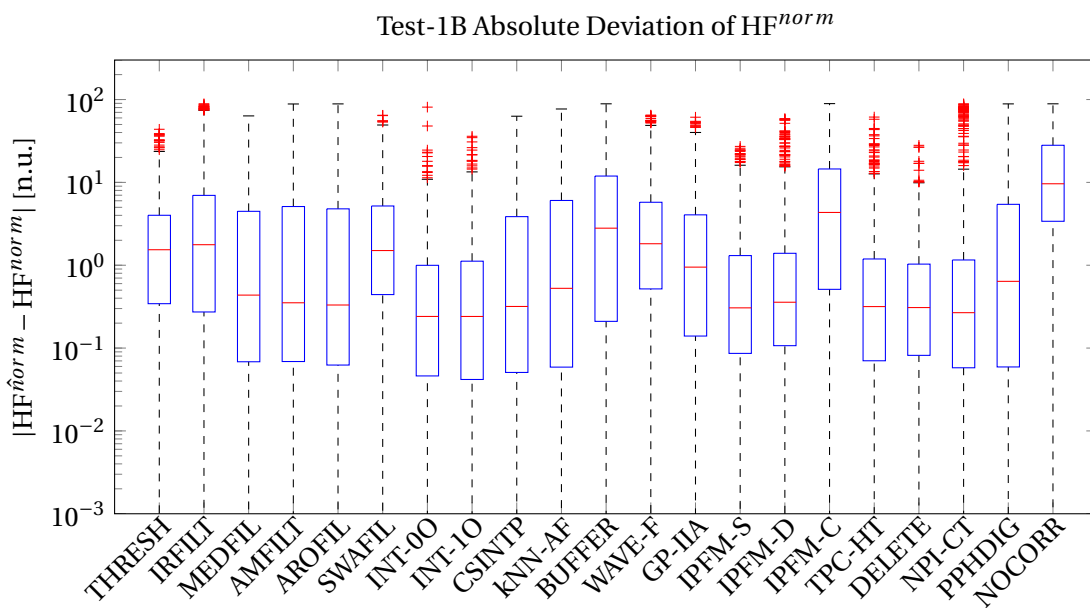


Figure B.16: Test-1B adjusted box plot of the absolute differences of HF^{norm} . The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1B cases, containing 151 signals each.

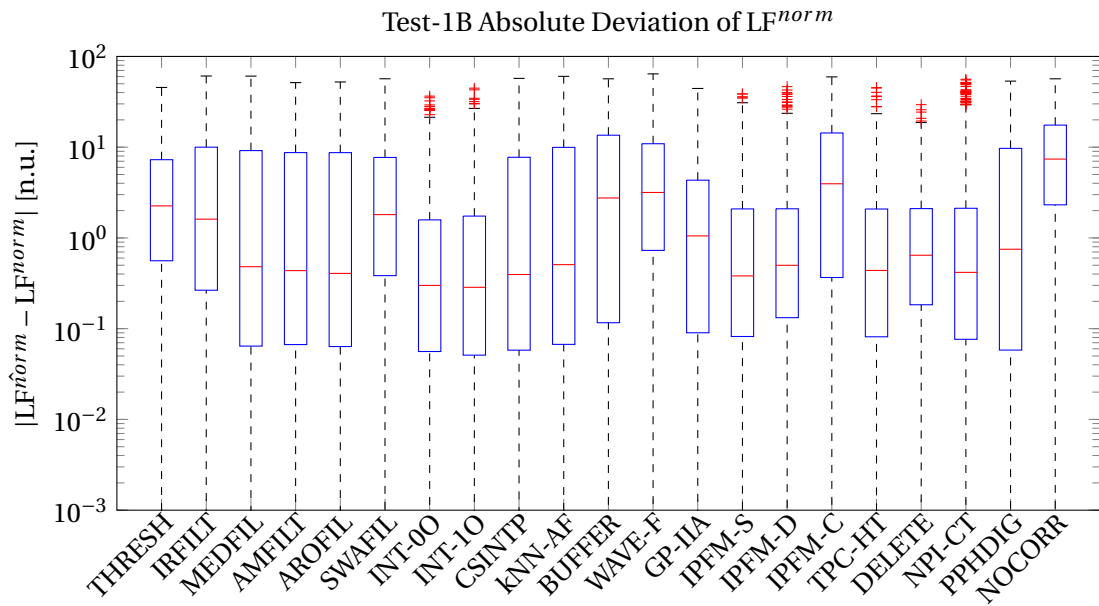


Figure B.17: Test-1B adjusted box plot of the absolute differences of LF^{norm} . The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1B cases, containing 151 signals each.

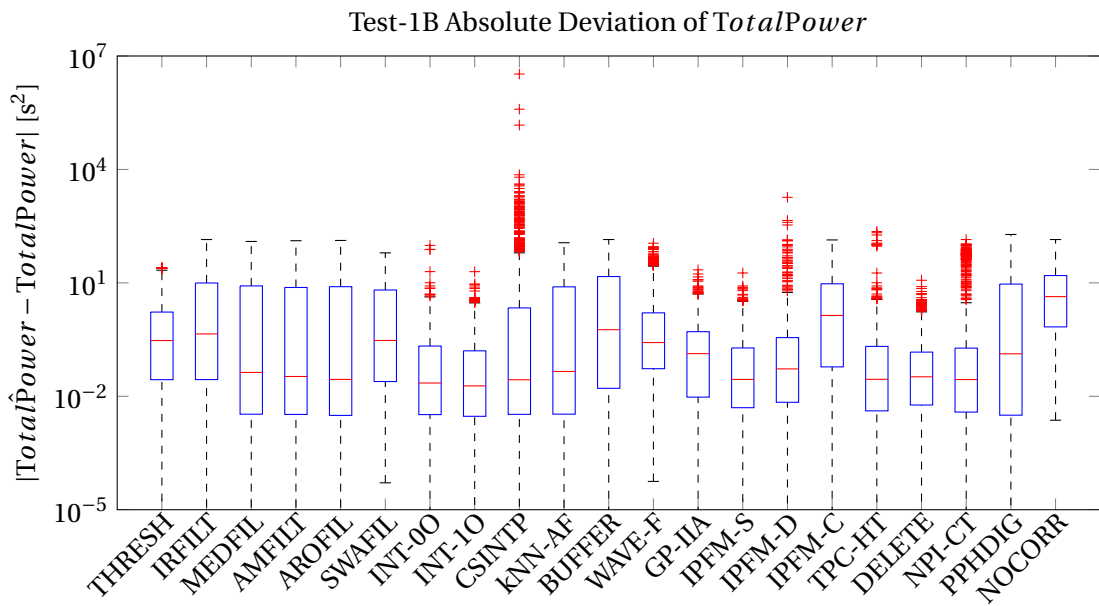


Figure B.18: Test-1B adjusted box plot of the absolute differences of total power. The adjusted box plot (according to Hubert and Vandervieren [27]) displays all 1510 recordings of the ten Test-1B cases, containing 151 signals each.

B.3 Specific Results of Test-2, Robustness

The number of successive ectopic beats that can be processed is about one to three for most approaches (see figure B.19). This finding is in accordance with the design of those approaches, since several methods just assume single ectopic beats, such as IRFILT and IPFM-C, or no more than three consecutive ectopic beats, such as the BUFFER and the PPHDIG model. Threshold filtering and gross positioning of beats can handle the largest number of multiple ectopic beats. Moreover, they show the same robustness in all HRV parameters. The two adaptive median filters, AMFILT and AROFIL, are both limited by half of their maximum window size, namely 12 successive ectopic beats.

Correction of successive ectopic beats is determined by the initial error (one single ectopic beat present). The best performing methods for single ectopic beats, such as deletion and interpolation approaches, reach a plateau, where error does not increase remarkable anymore. Therefore, they still perform a similarly reliable correction in highly erroneous signals, although they are not robust at all. In contrast, the threshold filter does not show any increase in error if the ectopic segment is enlarged. The GP-IIA approach shows a low increase at 12 ectopic beats and only a minor increase in error afterwards. Therefore, it seems to be nearly equivalent if successive ectopic beats are deleted or interpolated by several different approaches.

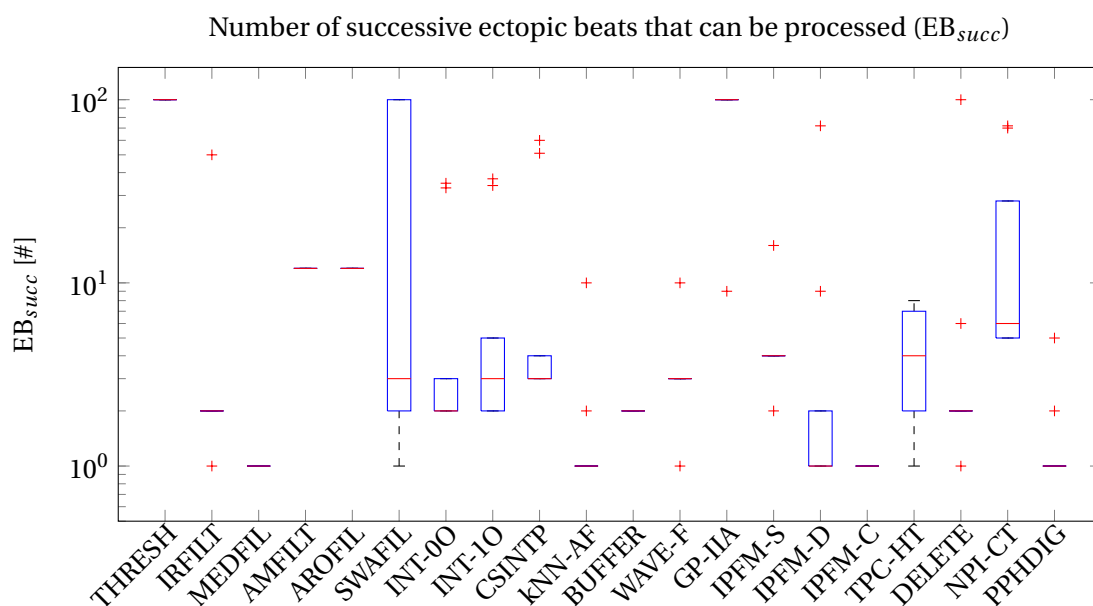


Figure B.19: Box plot of robustness against successive ectopic beats. Robustness was determined for each HRV parameter and the adjusted RMSD of RR intervals, resulting in ten data points per method.

Figure B.20 demonstrates that most approaches just need two to three correct NN intervals for a reliable correction. Further, most methods show a rather small variance. Threshold filtering and gross positioning of beats require only one correct NN interval and are thus best suited to

correct whole ectopic segments. AMFILT and AROFIL require 12 NN intervals between ectopic segments and are the least robust approaches in this case. As they were the third and fourth most robust methods against successive ectopic beats, it is obviously that the relation of these two parameters reflects the correction ability against multiple ectopic beats much better than the single parameters.

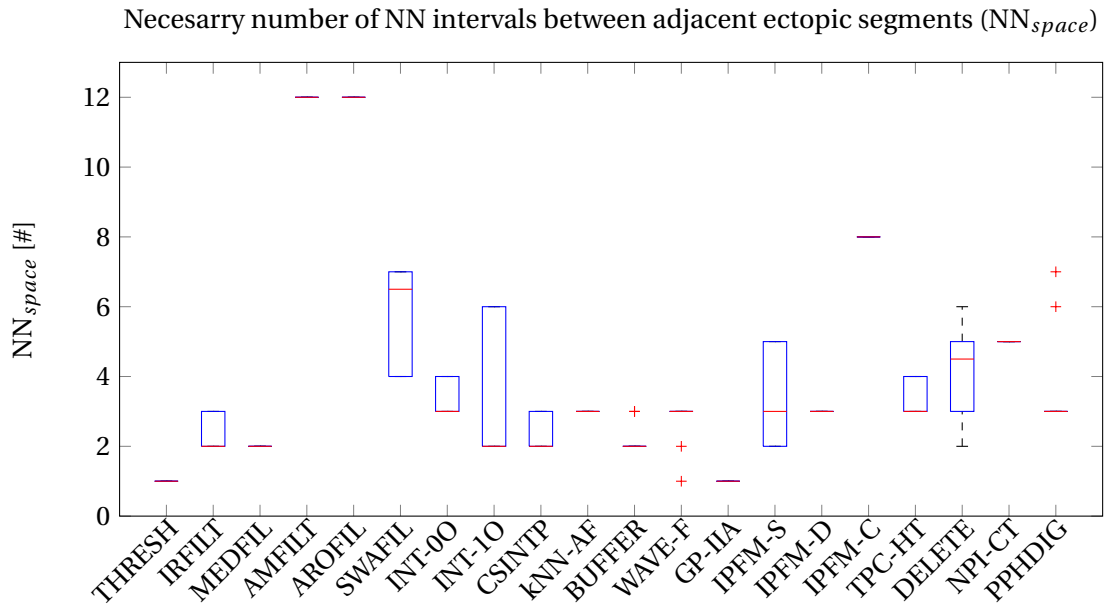


Figure B.20: Box plot of necessary correct NN intervals between ectopic segments. The number of intervals was determined for each HRV parameter and the adjusted RMSD of RR intervals, resulting in ten data points per method.

Bibliography

- [1] H. Abdi. The kendall rank correlation coefficient. *Encyclopedia of Measurement and Statistics*. Sage, Thousand Oaks, CA, pages 508–510, 2007.
- [2] B. Acar, I. Savelieva, H. Hemingway, and M. Malik. Automatic ectopic beat elimination in short-term heart rate variability measurement. *Comput Methods Programs Biomed*, 63(2):123–131, Oct 2000.
- [3] S. Akselrod, D. Gordon, J. B. Madwed, N. C. Snidman, D. C. Shannon, and R. J. Cohen. Hemodynamic regulation: investigation by spectral analysis. *Am. J. Physiol.*, 249(4 Pt 2):H867–875, Oct 1985.
- [4] S. Akselrod, D. Gordon, F. A. Ubel, D. C. Shannon, A.C. Berger, and R. J. Cohen. Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *Science*, 213(4504):220–222, 1981.
- [5] P. Albrecht and R.J. Cohen. Estimation of heart rate power spectrum bands from real-world data: dealing with ectopic beats and noisy data. In *Computers in Cardiology, 1988. Proceedings.*, pages 311–314, 1988.
- [6] American Heart Association. About arrhythmia. http://www.heart.org/HEARTORG/Conditions/Arrhythmia/AboutArrhythmia/About-Arrhythmia_UCM_002010_Article.jsp, 2013. [Online; accessed 11-November-2013].
- [7] M. Bachler. Automatische Detektion von QRS-Komplex, P- und T-Welle im Elektrokardiogramm in Echtzeit. Master’s thesis, Technische Universität Wien, Karlsplatz 13, 1040 Wien, Österreich, 2011.
- [8] M. Bachler, C. Mayer, B. Hametner, S. Wassertheurer, and A. Holzinger. Online and offline determination of QT and PR interval and QRS duration in electrocardiography. *Pervasive Computing and the Networked World*, 7719:1–15, 2013.
- [9] R. Barbieri and E.N. Brown. Analysis of heartbeat dynamics by point process adaptive filtering. *Biomedical Engineering, IEEE Transactions on*, 53(1):4–12, 2006.
- [10] H.V. Barron and M.D. Lesh. Autonomic nervous system and sudden cardiac death. *Journal of the American College of Cardiology*, 27(5):1053 – 1060, 1996.

- [11] S. Begum, M. S. Islam, M.U. Ahmed, and P. Funk. K-nn based interpolation to handle artifacts for heart rate variability analysis. In *Signal Processing and Information Technology (ISSPIT), 2011 IEEE International Symposium on*, pages 387–392, 2011.
- [12] G. G. Berntson, J. T. Bigger, D. L. Eckberg, P. Grossman, P. G. Kaufmann, M. Malik, H. N. Nagaraja, S. W. Porges, J. P. Saul, P. H. Stone, and M. W. van der Molen. Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology*, 34(6):623–648, Nov 1997.
- [13] G. G. Berntson, K. S. Quigley, J. F. Jang, and S. T. Boysen. An approach to artifact identification: application to heart period data. *Psychophysiology*, 27(5):586–598, Sep 1990.
- [14] C.L. Birkett, M.G. Kienzle, and G.A. Myers. Interpolation over ectopic beats increases low frequency power in heart rate variability spectra. In *Computers in Cardiology 1991, Proceedings.*, pages 257–259, 1991.
- [15] M. Brennan, M. Palaniswami, and P. Kamen. A new model-based ectopic beat correction algorithm for heart rate variability. In *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*, volume 1, pages 567–570 vol.1, 2001.
- [16] R. H. Chan, C. W. Ho, and M. Nikolova. Salt-and-Pepper noise removal by median-type noise detectors and detail-preserving regularization. *IEEE Trans Image Process*, 14(10):1479–1485, Oct 2005.
- [17] M. N. Cheung. Detection of and recovery from errors in cardiac interbeat intervals. *Psychophysiology*, 18(3):341–346, May 1981.
- [18] L. Citi, E. N. Brown, and R. Barbieri. Real-time automated point process method for detection and correction of erroneous and ectopic heartbeats. <http://neurostat-mit.appspot.com>, 2010-2013.
- [19] L. Citi, E. N. Brown, and R. Barbieri. A real-time automated point-process method for the detection and correction of erroneous and ectopic heartbeats. *IEEE Trans Biomed Eng*, 59(10):2828–2837, Oct 2012.
- [20] G. D. Clifford. Ecg statistics, noise, artifacts, and missing data. *Advanced Methods and Tools for ECG Data Analysis*, pages 55–99, 2006.
- [21] T. Garcia and G. Miller. *Arrhythmia Recognition: The Art of Interpretation*. Jones & Bartlett Learning, 2004.
- [22] M. J. Gardner and D. G. Altman. *Statistics with confidence: confidence intervals and statistical guidelines*. Bmj Books, 1989.

- [23] A. L. Goldberger, L. A.N. Amaral, L. Glass, J. M. Hausdorff, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley. Physiobank, physiotookit, and physionet components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000.
- [24] D. S. Goldstein, O. Benthoo, M. Y. Park, and Y. Sharabi. Low-frequency power of heart rate variability is not a measure of cardiac sympathetic tone but may be a measure of modulation of cardiac autonomic outflows by baroreflexes. *Exp. Physiol.*, 96(12):1255–1261, Dec 2011.
- [25] O.V. Grishin, V.G. Grishin, D. Y. Uryumtsev, S.V. Smirnov, and I.G. Jilina. Metabolic rate variability impact on very low-frequency of heart rate variability. *World Applied Sciences Journal*, 19(8):1133–1139, 2012.
- [26] C.-H. Hsieh and P.-C. Huang. Adaptive rank order filter for image noise removal. In *Computer Science and Information Engineering, 2009 WRI World Congress on*, volume 7, pages 90–94, 2009.
- [27] M. Hubert and E. Vandervieren. An adjusted boxplot for skewed distributions. *Computational Statistics & Data Analysis*, 52(12):5186–5201, 2008.
- [28] H. Hwang and R. Haddad. Adaptive median filters: new algorithms and results. *Image Processing, IEEE Transactions on*, 4(4):499–502, 1995.
- [29] L.Y. Ji, Y.J. Yang, A. G. Li, S.F. Wang, and J.K. Wu. Robust time series processing for heart rate variability analysis in daily life. In *Computing in Cardiology, 2011*, pages 301–304, 2011.
- [30] I. Jolliffe. *Principal Component Analysis*. Wiley Online Library, 2005.
- [31] J. Jung, A. Heisel, D. Tscholl, R. Fries, H. Schieffer, and C. Özbek. Assessment of heart rate variability by using different commercially available systems. *The American Journal of Cardiology*, 78(1):118–120, 1996.
- [32] M. V. Kamath and E. L. Fallen. *Heart Rate Variability*, chapter “Correction of heart rate variability signal for ectopics and missing beats”, pages 75–85. Armonk, NY: Futura, 1995.
- [33] D. B Keenan. Detection and correction of ectopic beats for hrv analysis applying discrete wavelet transforms. *Int. J. Inf. Technol*, 2:54–60, 2005.
- [34] K. J. Kemper, C. Hamilton, and M. Atkinson. Heart rate variability: impact of differences in outlier identification and management strategies on common measures in three clinical populations. *Pediatric research*, 62(3):337–342, 2007.
- [35] P. Korhonen and A. Siljamäki. Ordinal principal component analysis theory and an application. *Computational Statistics & Data Analysis*, 26(4):411–424, 1998.

- [36] N. Kumaravel and C. Santhi. Nonlinear filters for preprocessing heart rate variability signals. *International Journal of Computer Science and Network Security*, 10:250–254, 2010.
- [37] P. Laguna, R. G. Mark, A. Goldberg, and G. B. Moody. A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG. In *Computers in Cardiology 1997*, pages 673–676. IEEE, 1997.
- [38] M.-Y. Lee and S.-N. Yu. Improving discriminability in heart rate variability analysis using simple artifact and trend removal preprocessors. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 4574–4577, 2010.
- [39] K. U. Leuven. Libra. <https://wis.kuleuven.be/stat/robust/Programs/LIBRA/download-links>, 02 2012. [Online; accessed 28-February-2014].
- [40] M. N. Levy. Sympathetic-parasympathetic interactions in the heart. *Circ. Res.*, 29(5):437–445, Nov 1971.
- [41] N. Lippman, K. M. Stein, and B. B. Lerman. Nonlinear predictive interpolation. a new method for the correction of ectopic beats for heart rate variability analysis. *J. Electrocardiol.*, 26(Suppl.:14–19, 1993.
- [42] N. Lippman, K. M. Stein, and B. B. Lerman. Comparison of methods for removal of ectopy in measurement of heart rate variability. *Am. J. Physiol.*, 267(1 Pt 2):H411–418, Jul 1994.
- [43] C. Liu, L. Li, L. Zhao, D. Zheng, L. Li, and C. Liu. A combination method of improved impulse rejection filter and template matching for identification of anomalous intervals in RR sequences. *Journal of Medical and Biological Engineering*, 32(4):245–250, 2012.
- [44] B. Lown and R. L. Verrier. Neural activity and ventricular fibrillation. *N. Engl. J. Med.*, 294(21):1165–1170, May 1976.
- [45] Lubopitko. Function of the heart. http://encyclopedia.lubopitko-bg.com/Function_of_the_Heart.html. [Online; accessed 19-November-2013].
- [46] A. Malliani, M. Pagani, F. Lombardi, and S. Cerutti. Cardiovascular neural regulation explored in the frequency domain. *Circulation*, 84(2):482–492, 1991.
- [47] E.N. Marieb and K. Hoehn. *Human Anatomy and Physiology*. Human Anatomy & Physiology. Benjamin-Cummings Publishing Company, 2010.
- [48] A. Martinez, R. Alcaraz, and J. J. Rieta. Detection and removal of ventricular ectopic beats in atrial fibrillation recordings via principal component analysis. *Conf Proc IEEE Eng Med Biol Soc*, 2011:4693–4696, 2011.
- [49] A. Martinez, R. Alcaraz, and J. J. Rieta. Ventricular activity morphological characterization: ectopic beats removal in long term atrial fibrillation recordings. *Comput Methods Programs Biomed*, 109(3):283–292, Mar 2013.

- [50] A. Martinez, R. Alcaraz, and J.J. Rieta. Ectopic beats canceler for improved atrial activity extraction from holter recordings of atrial fibrillation. In *Computing in Cardiology, 2010*, pages 1015–1018, 2010.
- [51] A. Martinez, R. Alcaraz, and J.J. Rieta. Optimal cancellation template analysis for ectopic beats removal in atrial fibrillation recordings. In *Computing in Cardiology (CinC), 2012*, pages 801–804, 2012.
- [52] O. Masek. Heart rate variability analysis. Master’s thesis, Czech Technical University in Prague, Faculty of Electrical Engineering, Department of Cybernetics, 2009.
- [53] J. Mateo and P. Laguna. Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal. *IEEE Trans Biomed Eng*, 50(3):334–343, Mar 2003.
- [54] J. Mateo, A. Torres, and J.J. Rieta. An efficient method for ectopic beats cancellation based on radial basis function. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 6947–6950, 2011.
- [55] J. McNames, T. Thong, and M. Aboy. Impulse rejection filter for artifact removal in spectral analysis of biomedical signals. *Conf Proc IEEE Eng Med Biol Soc*, 1:145–148, 2004.
- [56] J. E. Mietus. Time domain measures: from variance to pnnx. <http://physionet.org/events/hrv-2006/mietus-1.pdf>, 2006. [Online; accessed 25-May-2013].
- [57] N. Montano, T. G. Ruscone, A. Porta, F. Lombardi, M. Pagani, and A. Malliani. Power spectrum analysis of heart rate variability to assess the changes in sympathovagal balance during graded orthostatic tilt. *Circulation*, 90(4):1826–1831, Oct 1994.
- [58] G.B. Moody and R.G. Mark. The impact of the MIT-BIH arrhythmia database. *Engineering in Medicine and Biology Magazine, IEEE*, 20(3):45–50, May 2001.
- [59] G. A. Ng. Treating patients with ventricular ectopic beats. *Heart*, 92(11):1707–1712, Nov 2006.
- [60] Stanislav Nikolov. Principal component analysis: Review and extensions. *Dept. of Electrical Engineering & Computer Science, Massachusetts Institute of Technology*, 2010.
- [61] Task Force of ESC and NASPE. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Circulation*, 93(5):1043–1065, Mar 1996.
- [62] B. Olshansky, H. N. Sabbah, P. J. Hauptman, and W. S. Colucci. Parasympathetic nervous system and heart failure: pathophysiology and potential implications for therapy. *Circulation*, 118(8):863–871, Aug 2008.

- [63] P.-E. Paulev and G. Zubieta. *New Human Physiology*. Copenhagen Medical Publishers, 2nd edition edition, 2004.
- [64] M. A. Peltola. *Analysis of heart rate variability from 24-hour ambulatory electrocardiographic recordings*. PhD thesis, University of Oulu, 2010.
- [65] M. A. Peltola. Role of editing of R-R intervals in the analysis of heart rate variability. *Front Physiol*, 3:148, 2012.
- [66] M. A. Peltola, T. Seppanen, T. H. Makikallio, and H. V. Huikuri. Effects and significance of premature beats on fractal correlation properties of R-R interval dynamics. *Ann Noninvasive Electrocardiol*, 9(2):127–135, Apr 2004.
- [67] B. Pomeranz, R. J. Macaulay, M. A. Caudill, I. Kutz, D. Adam, D. Gordon, K. M. Kilborn, A. C. Barger, D. C. Shannon, and R. J. Cohen. Assessment of autonomic function in humans by heart rate spectral analysis. *Am. J. Physiol.*, 248(1 Pt 2):H151–153, Jan 1985.
- [68] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. Numerical recipes in c: The art of scientific computing 2nd edition. *Cambridge (Mass.): Cambridge Univ. Press, Cambridge*, 1992.
- [69] J. Rand, A. Hoover, S. Fishel, J. Moss, J. Pappas, and E. Muth. Real-time correction of heart interbeat intervals. *IEEE Trans Biomed Eng*, 54(5):946–950, May 2007.
- [70] V. L. Roger, A. S. Go, D. M. Lloyd-Jones, E. J. Benjamin, J. D. Berry, W. B. Borden, D. M. Bravata, S. Dai, E. S. Ford, C. S. Fox, et al. Heart disease and stroke statistics 2012 update a report from the american heart association. *Circulation*, 125(1):e2–e220, 2012.
- [71] M. A. Salo, H. V. Huikuri, and T. Seppanen. Ectopic beats in heart rate variability analysis: effects of editing on time and frequency domain measures. *Ann Noninvasive Electrocardiol*, 6(1):5–17, Jan 2001.
- [72] P. J. Schwartz, M. T. La Rovere, and E. Vanoli. Autonomic nervous system and sudden cardiac death. Experimental basis and clinical observations for post-myocardial infarction risk stratification. *Circulation*, 85(1 Suppl):77–91, Jan 1992.
- [73] J. M. Serrador, H. C. Finlayson, and R. L. Hughson. Physical activity is a major contributor to the ultra low frequency components of heart rate variability. *Heart*, 82(6):e9, Dec 1999.
- [74] G. Sethuraman, K. L. Ryan, C. A. Rickards, and V. A. Convertino. Ectopy in trauma patients: cautions for use of heart period variability in medical monitoring. *Aviation, space, and environmental medicine*, 81(2):125–129, 2010.
- [75] K. Solem, P. Laguna, J.P. Martinez, and L. Sornmo. Model-based detection of heart rate turbulence. *Biomedical Engineering, IEEE Transactions on*, 55(12):2712–2722, 2008.
- [76] K. Solem, P. Laguna, and L. Sornmo. An efficient method for handling ectopic beats using the heart timing signal. *IEEE Trans Biomed Eng*, 53(1):13–20, Jan 2006.

- [77] D. D. Suhr. Principal component analysis vs. exploratory factor analysis. *SUGI 30 Proceedings*, pages 203–230, 2005.
- [78] A. Taddei, G. Distanti, M. Emdin, P. Pisani, G. B. Moody, C. Zeelenberg, and C. Marchesi. The European ST-T database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography. *Eur. Heart J.*, 13(9):1164–1172, Sep 1992.
- [79] M. P. Tarvainen, J.-P. Niskanen, J.A. Lipponen, P.O. Ranta-Aho, and P.A. Karjalainen. Kubios HRV - a software for advanced heart rate variability analysis. In *4th European Conference of the International Federation for Medical and Biological Engineering*, pages 1022–1025. Springer, 2009.
- [80] M.P. Tarvainen, P.O. Ranta-aho, and P.A. Karjalainen. An advanced detrending method with application to hrv analysis. *Biomedical Engineering, IEEE Transactions on*, 49(2):172–175, 2002.
- [81] R. A. Thuraisingham. Preprocessing RR interval time series for heart rate variability analysis and estimates of standard deviation of RR intervals. *Comput Methods Programs Biomed*, 83(1):78–82, Jul 2006.
- [82] K. K. Tripathi. Very low frequency oscillations in the power spectra of heart rate variability during dry supine immersion and exposure to non-hypoxic hypobaria. *Physiol Meas*, 32(6):717–729, Jun 2011.
- [83] J. W. Tukey. Exploratory data analysis. *Reading, Ma*, 231, 1977.
- [84] S. Ward, C. Heneghan, and P. Nolan. An integrate-and-fire based model of PP and PR variability in the human electrocardiogram. In *Neural Engineering, 2003. Conference Proceedings. First International IEEE EMBS Conference on*, pages 297–300, 2003.
- [85] F. Wen and F. T. He. An efficient method of addressing ectopic beats: new insight into data preprocessing of heart rate variability analysis. *J Zhejiang Univ Sci B*, 12(12):976–982, Dec 2011.
- [86] D. Widjaja, S. Vandeput, J. Taelman, M.A.K.A. Braeken, R.A. Otte, B.R.H. Van den Bergh, and S. Van Huffel. Accurate R peak detection and advanced preprocessing of normal ECG for heart rate variability analysis. In *Computing in Cardiology, 2010*, pages 533–536, 2010.
- [87] Wikipedia. Continuous noninvasive arterial pressure — wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=Continuous_noninvasive_arterial_pressure&oldid=565663464, 2013. [Online; accessed 11-November-2013].
- [88] Wikipedia. Electrocardiography — wikipedia, the free encyclopedia. <http://en.wikipedia.org/w/index.php?title=Electrocardiography&oldid=580042246>, 2013. [Online; accessed 11-November-2013].

- [89] Wikipedia. Pulse oximetry — Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Pulse_oximetry, 2013. [Online; accessed 04-December-2023].
- [90] M. N. Yang, T. & Levy. Effects of intense antecedent sympathetic stimulation on sympathetic neurotransmission in the heart. *Circ. Res.*, 72(1):137–144, Jan 1993.
- [91] C. S. Yoo and S. H. Yi. Effects of detrending for analysis of heart rate variability and applications to the estimation of depth of anesthesia. *Journal of the Korean Physical Society*, 44(3):561–568, March 2004.