Alignment of R–R Interval Signals using the Circadian Heart Rate Rhythm

Η ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ ΕΞΕΙΔΙΚΕΥΣΗΣ

υποβάλλεται στην

ορισθείσα από την Γενική Συνέλευση Ειδικής Σύνθεσης

του Τμήματος Πληροφορικής Εξεταστική Επιτροπή

από την

Nathalie Thérèse Hélène Gayraud

ως μέρος των Υποχρεώσεων για τη λήψη του

ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΣΤΗΝ ΠΛΗΡΟΦΟΡΙΚΗ

ΜΕ ΕΞΕΙΔΙΚΕΥΣΗ

ΣΤΙΣ ΤΕΧΝΟΛΟΓΙΕΣ – ΕΦΑΡΜΟΓΕΣ

Ιουνιος 2015

Dedication

To my parents, who brought me and my brother into this world.
To my family and friends, who see it through...

Acknowledgements

In all sincerity, I would like to thank my supervisor George Manis, for he provided me with the most noble armor in the world: knowledge, patience, objective thinking, self–confidence.
I would also like to aknowledge my dear friend Marina Argyri for her companionship in our many hours of hard work.
Thank you both for for providing me with inspiration and constant support.

Table of Contents

List of Figures

List of Algorithms

Περιληψη

Ναταλί Τ.Ε. Γκαιρώ του Πασκάλ και της Καλυψούς. MSc, Τμήμα Πληροφορικής, Πανεπιστήμιο Ιωαννίνων, Ιούνιος, 2015. Συγχρονισμός R–R interval σημάτων βάση του Κιρκάδιου Καρδιακού Ρυθμού. Επιβλέποντας: Γεώργιος Μανής.

Οι μελέτες που εστιάζουν σε R–R interval σήματα, τα οποία προέρχονται από διαφορετικά άτομα, τείνουν να ευθυγραμμίζουν και να υπολογίζουν το μέσο όρο των σημάτων αυτών βάση του ωρολογιακού χρόνου εκκίνησης του σήματος. Στην παρούσα εργασία, στόχος μας είναι να δείξουμε ότι η χρήση του ωρολογιακού χρόνου αποτελεί λανθασμένο κριτήριο ευθυγράμμισης, ενώ παρέχουμε μια νέα μέθοδο ευθυγράμμισης.

Τις μεθόδους μας ενέπνευσε ο κιρκάδιος κύκλος που ακολουθείται από τον ανθρώπινο καρδιακό ρυθμό. Το μοτίβο που προκύπτει δύναται να ποικίλει μεταξύ διαφόρων ανθρώπων. Οι μέθοδοι ευθυγράμμισης που προτείνουμε χρησιμοποιούν την εν λόγω παρατήρηση και δημιουργούν νέα ορόσημα για την αυτόματη ευθυγράμμιση R - R interval σημάτων.

Προτείνουμε δύο πρωτότυπους αλγόριθμους ευθυγράμμισης οι οποίοι λαμβάνουν υπ'όψιν τους τον ανθρώπινο κιρκαδικό ρυθμό: τον *Puzzle Piece Alignment αλγόριθμο*, και τον *Event Based Alignment αλγόριθμο*. Αρχικά, κάθε σήμα R–R μετατρέπεται σε μία σειρά από παράθυρα και υπολογίζεται ο μέσος καρδιακός ρυθμός ανά παράθυρο. Στη συνέχεια, οι αλγόριθμοι μας προσπαθούν να ταιριάξουν τα μοτίβα που δημιουργούνται από τον κάθε κιρκάδιο κύκλο, παράγοντας σχετικά ορόσημα για την ευθυγράμμιση των σημάτων. Στα πειράματα που διεξάγουμε, χρησιμοποιούμε R–R interval σήματα τα οποία προέρχονται από τέσσερις βάσεις δεδομένων της Τράπεζα δεδομένων Physionet. Και οι δύο αλγόριθμοι καταφέρνουν να ευθυγραμμίσουν τα σήματα ως προς τον κιρκάδιο καρδιακό ρυθμό. Επιπλέον, τα ευρήματά μας επιβεβαιώνουν την παρουσία περισσότερων από ένα μοτίβων στο κιρκαδικό καρδιακό ρυθμό, ένα ενδιαφέρον αποτέλεσμα που παρέχει έδαφος για περαιτέρω ανάλυση. Προτείνουμε μια αυτόματη ταξινόμηση των σημάτων σύμφωνα με τα τρια σημαντικότερα μοτίβα.

Τα αποτελέσματά μας δείχνουν ότι ο αλγόριθμος μας ευθυγραμμίζει σωστά τα σήματα σύμφωνα με τον ανθρώπινο κιρκαδικό καρδιακό ρυθμό. Μάλιστα, η επιτυχία στην ευθυγράμμιση των σημάτων προκύπτει ακόμη και όταν τα σήματα εισόδου ξεκινούν από τυχαία σημεία στο χρόνο. Τέλος, ο αλγόριθμος EBA κατορθώνει να διακρίνει την ύπαρξη διαφορετικών μοντέλων και να ταξινομήσει ανάλογα τα σήματα.

Abstract

Nathalie T.H. Gayraud. Msc, Computer Science Department, University of Ioannina, Greece. June, 2015. Alignment of R–R Interval Signals using the Circadian Heart Rate Rhythm. Thesis Supervisor: George Manis.

Studies that focus on R–R interval signals coming from different subjects tend to be aligned and averaged according to the horological time of the recordings. This work argues that the horological time can be an erroneous alignment criterion and provide a new alignment method to prove our argument.

What inspired our methods is the circadian cycle followed by the human heart rate rhythm. Since its pattern can vary among different classes of people, we suggest alignment methods that use that observation and establish new landmarks to automatically align R–R interval signals.

We propose two novel alignment algorithms that consider the HR circadian rhythm: the *Puzzle Piece Alignment Algorithm*, and the *Event Based Alignment Algorithm*. First, each R-R interval signal is converted into a series of time windows and the mean Heart Rate per window is calculated. Then our algorithms search for matching circadian patterns, and use relevant landmarks to align the signals. We conduct experiments using R-R interval signals extracted from four databases in the Physionet Data Bank. Both algorithms are able to align the signals with respect to the circadian rhythmicity of HR. Furthermore, our findings confirm the presence of more than one pattern in the circadian HR rhythm, an interesting result providing the ground for further analysis. We suggest an automatic classification of signals according to the three most prominent patterns.

Our results show that our algorithm are able to correctly align the signals in accordance to the human circadian Heart Rate rhythm. They succeed in aligning the signal even when the input signals start from random points in time. In addition, the EBA algorithm manages to discern the existence of different patterns.

Chapter 1

Introduction

1.1 Objective

Bioelectrical signals have been used in medicine for a long time, a fact that is mostly due to their non invasive acquisition nature versus the amount of information they provide. In this research, we focus on the the ECG; more specifically, we study R–R interval signals and their derived measurements. An R–R interval denotes the time between two consecutive heart beats; R–R interval signals consist of successive R–R intervals.

The increasing number of 24 hour R–R interval recordings available in databases such as the Physionet Data Bank [1] is of great value for researchers in the field of Biomedicine. However, not all of them contain horological information as to when the recordings begin. This means that a large number of datasets contained in these growing databases will not be suitable for researches that study bioelectrical signal evolution. What initially motivated this work is precisely that unnecessary loss of information. We unveiled a pressing need for a strong alignment criterion to synchronize these asynchronous recordings.

In this research, we present numerous cases that validate the existence of a circadian Heart Rate rhythm, closely correlated to the human sleep cycle. Generally, the HR has a tendency to augment when a person is awake and decrease when one is asleep. Naturally, other daily phenomena contribute to increase and decrease the heart rate, even unrelated to the human circadian cycle. Nevertheless, an average HR signal that is obtained by a large enough number of patients should follow the circadian pattern.

Hence, since we can identify it, it is natural that we should use that circadian rhythmicity to align 24 hour recordings. It is yet true that, not all humans have the same

sleep/wake cycle. Moreover, the human circadian cycle can start at different times for different people.

As a result of these observation, it becomes clear that even aligning signals using available horological information can be a faulty method. An alignment criterion must be defined when working with 24 hour recordings. This work aims to study the circadian rhythmicity of the Heart Rate, and use it as a standard to provide reliable alignment methods.

1.2 Contribution

In our research, we first investigate the circadian rhythmicity of relevant measurement of the Heart Rate Variability, and provide with computational methods that can be used for studying their evolution.

Our most essential contribution is two R–R interval alignment methods that can be used to synchronize 24 hour signals belonging from different subjects: The *Puzzle Piece Alignment* method and the *Event Based Alignment* method.

The first alignment method is a simple one; it aligns the input signals against each other. The second alignment method offers a higher level of sophistication. It uses the circadian rhythmicity of the heart rate to establish landmarks, who then serve two purposes. They are used as synchronization points, but the number of detected landmarks also serves as a classification label, so that in the output is a categorization of the input signals with respect to each subject's circadian cycle.

We develop and test these algorithms on 32 combinations of datasets and parameters. Our results confirm the efficiency of the proposed methods. Both are successfully in aligning signals, regardless to their original alignment. Furthermore, the classification algorithm uncovers the existence of different starting points for each subject's cycle, as well as the presence of more than one circadian patterns.

Finally, we demonstrate an example of the effect of signal synchronization on 24-hour Heart Rate Variability measurements derived by R–R interval signals.

1.3 Structure

The thesis is structured as follows. Chapter 2 starts by presenting the human cardiac electrophysiology and proceeds by explaining the basics on bioelectrical signals, Heart Rate Variability and the circadian rhythmicity of the human heart. In this chapter we cite a number of researches that use Heart Rate Variability measurements in the prognosis and/or diagnosis of several conditions. We also show the studies focused on the circadian Heart Rate rhythm.

In chapter 3 we start by introducing computational methods that can be used for analysing the evolution of Heart Rate Variability measurements. Section 3.1 gives an

introduction on the necessary signal pre-processing, whereas section 3.2 focuses on processing Heart Rate Variability measurement signals. We outline the importance of signal alignment, and present our two methods in section 3.3.

The last two chapters display the results we obtained, discussion, and future work.

Chapter 2

Biomedical Signals and Rhythmicity

2.1 The Human Heart

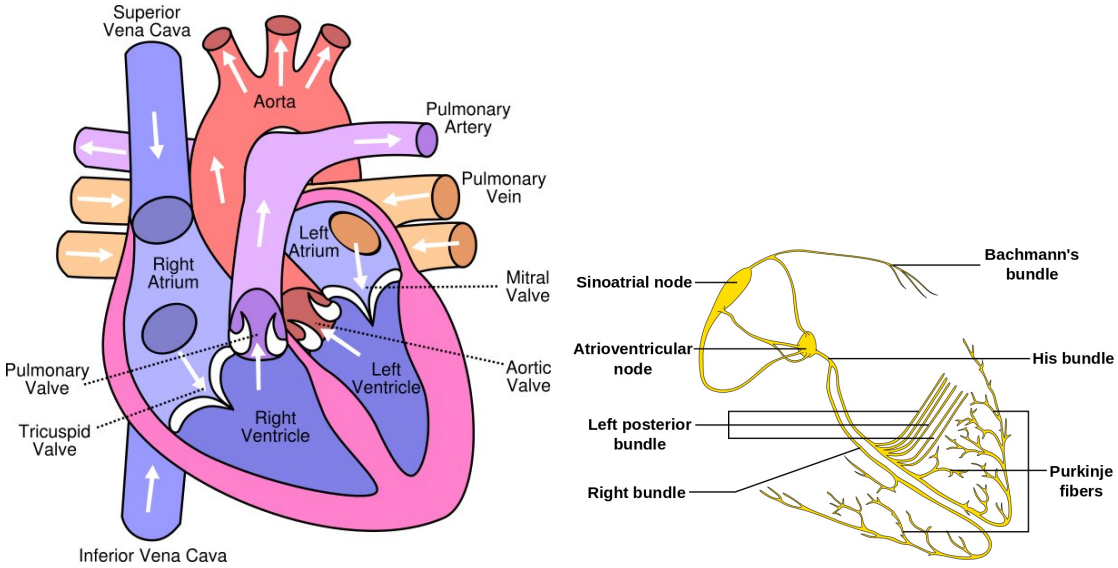
The heart represents one of the most important organs of the human body. It is in charge of pumping blood though the circulatory system to every cell in our body, supplying oxygen and nutrients to the tissues and removing carbon dioxide and other wastes.

The anatomy of the human heart is depicted on figure 2.1a [1]. As illustrated by [2],[2] four chambers compose the human heart; the upper left and right chambers called the atria, and the lower left and right chambers called the ventricles. The atria are thin-walled structures, acting primarily as holding chambers for the blood, and as anterooms to the ventricles. The ventricles are larger and thick-walled chambers, in charge of pumping the blood to the rest of the body. The left atria and ventricle are responsible for providing oxygenated blood from the lungs to the rest of the body, while the right atria and ventricle send the deoxygenated blood back to the lungs to be reoxygenated again. The left and right side of the heart are separated by the septum, a wall of tissue that prevents blood to pass from one side to the other.

Each chamber is connected to one or two *great vessels of the heart*. The great vessels of the heart and the respective chambers they are connected to are:

[1] https://en.wikipedia.org/wiki/File:Diagram_of_the_human_heart_(cropped).svg

[2] https://en.wikipedia.org/wiki/File:ConductionsystemoftheheartwithouttheHeart-en.svg



(a) The anatomy of the human heart.

(b) The electrical conduction system of the human heart.

Figure 2.1: Figures representing the anatomy and electrical conduction of the human heart.

- The superior and inferior vena cavae are connected to the right atrium. They empty all the deoxygenated blood from the body into the heart.
- The pulmonary artery is connected to the right ventricle. It carries that deoxygenated blood from the heart to the lungs.
- The pulmonary veins are connected to the left atrium. They transport the oxygenated blood from the lungs to the heart.
- The aorta, also called great artery, is connected to the left ventricle. It brings the oxygenated blood from the heart to the rest of the body.

Four sets of valves help the blood flow in the right direction when the heart is beating; the tricuspid valve, the pulmonic valve, the mitral valve and the aortic valve. Furhtermore, to pump the blood, the heart contracts and expands rhythmically; the contractions is called *systole* and the expansion *diastole*. These two pumping functions run simultaneously in the course of a cardiac beat.

Zarzoso et al. [3] describe the *Normal Sinus Rhythm*, i.e. the beating process corresponding to a healthy electrophysiological state, in the following way. While at rest, a difference in the potential across the myocardial membrane causes the cells to become polarized. In consequence, the heart's pacemaker, called the sinoatrial node and located in the top right atrium, initiates a beat. It causes a depolarization wavefront to propagate from the top right to the bottom left atrium, which in its turn is responsible for a synchronized atrial contraction and the efficient passage of blood towards the ventricles through the mitral and tricuspid valves.

The atrio-ventricular node, located between the atria and the ventricles, delays the subsequent depolarization wavefront to allow the blood to effectively flow from the atria to the ventricles. Afterwards, the wavefront promptly propagates through the bundle of His, the bundle branches and the Purkinje fibers (muscle cells and fibers specialized for electrical conduction) towards the bottom of the ventricular myocardium. The electrical activation ultimately propagates from the apex to the base of the ventricles, who contract upwards, ejecting the blood towards the lungs and the rest of the body. Figure 2.1b displays the heat's electrical circuit.

2.2 Bioelectrical Signals

Biomedical signal analysis has always been an crucial part of healthcare. In its most primitive and early state, it consists of listening to the sounds reproduced by the human body. From the dawn of medicine, physicians lean on patients to listen to the sound of their breathing and analyse the rhythmicity of their heart beats. In this technique lies the most ancient biomedical signal application.

Remarkably this method is still used today, by reason of its invasiveness versus acquired information ratio. This ratio is the reason why bioelectrical signal analysis is such and exciting and important research field. Through constantly evolving processing techniques, it can prove an indispensable asset to physicians, allowing them to drastically improve disease prevention, diagnosis, and treatment.

From a cellular perspective, bioelectrical signals are linked with ionic processes. These processes arise as a result of the electrochemical activity of a special group of cells. These *excitable* cells are driven by the same mechanisms, regardless of whether they are located in the heart, the brain or the muscles [4]. [5, 6] provide a thorough description of the cellular processes that induce bioelectrical signals.

On the body surface, the cumulative electrical activity of those cells can be noninvasively measured, as the tissue surrounding them is a conductive medium, i.e. a volume conductor [7]. The simplest way of recording a bioelectrical signal is to attach two electrodes to the body surface; the *exploring* electrode, that should be placed close to the electrical source, and the *indifferent electrode*, placed somewhere else [8]. In practice, multiple electrodes are used for the description of a bioelectrical phenomenon.

The following biosignals describe the activity of various organs in the human body. The process used to retrieve their properties is, naturally, signal specific, thus each biosignal has its own acquisition process, treatment, and description.

2.2.1 Electrocardiogram

An electrocardiogram (ECG) reflects the electrical activity of the heart. This bioelectrical signal is acquired through the non-invasive process of placing electrodes on the chest, arms and legs. Each heartbeat causes an impulse to travel through the heart, which allows for

the determination of its rate and rhythm. The heartbeat is manifested in the ECG by a series of waves, whose morphology and timing provide information used to diagnose diseases that are reflected by disturbances in the electrical activity of the heart, thus making it a standard clinical procedure for the investigation of heart diseases.

An electrocardiogram is acquired by attaching a set of electrodes on the skin, arranged in such a way that they reflect the spatiotemporal variations of the cardiac electrical field [9]. The two most common recording systems today are the standard twelve lead ECG, and the orthogonal lead system.

2.2.2 Electroencephalogram

The electroencephalogram (EEG) echoes the electrical activity produced by the brain. The acquisition process consists of placing several electrodes on the scalp, an instrument that is often referred to as an electroencephalograph [4]. The EEG is widely used for the diagnosis and evaluation of a significant number of brain disorders, such as, epileptic seizures and sleep disorders, as well as the diagnosis of conditions such as coma or brain death [10].

2.2.3 Other Bioelectrical Signals

While the two previously mentioned bioelectrical signals are the most well known ones, other types exist as well, which we briefly name here.

On the domain of nerves and muscles, we find the electromyogram (EMG) and the electroneurogram (ENG) . The EMG measures the electrical activity of skeletal muscles. It is mainly used in the detection of abnormalities in muscular activity, potentially caused by injuries or muscular diseases. The ENG measures the nerve response to electrical stimuli. More specifically, the electrogastogram (EGG) records the electrical activity of stomach muscles.

Electrical stimuli coming from the human eye can be measured by the electroretinogram (ERG) and the electrooculogram(EOG) The ERG is used to examine the electrical activity of the retina during light stimulation, while the EOG records the steady corneal-retinal potential, that corresponds to eye movement. Note that the EOG is closely related to the study of sleep disorders.

2.3 Heart Rate Variability Analysis

In section 2.2.1 we explained the basics of the ECG. Among the significant characteristics that can be extracted from an ECG, we focus on the R–R interval. To describe what is the R–R interval, we first need to introduce the QRS complex.

The QRS complex is a wave resulting from the depolarization of the left and right ventricles. The Q wave is the first negative deflection, followed by the R wave which is

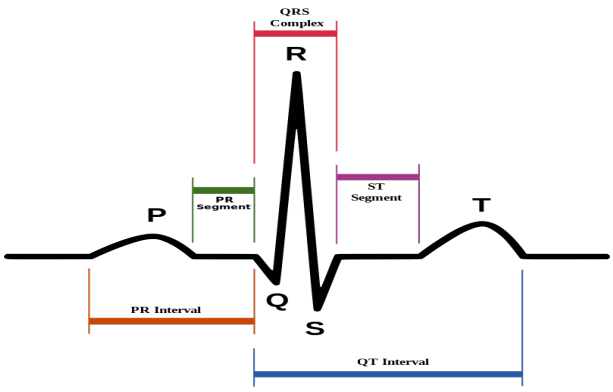


Figure 2.2: The QRS complex.

the first positive reflection. The last, negative reflection, is the S wave. Although the morphology of the QRS can show significant variations depending on the origin of the heartbeat, a typical wave can be seen in figure 2.2 [3]

The R–R interval measures the length between two successive R waves, which represents the length of a ventricular cardiac cycle. When it comes to ECG analysis, it is the fundamental rhythm quantity; it is the basic measurement used in Heart Rate Variability analysis [9]. Heart Rate Variability (HRV), also referred to as cycle length variability, or R–R variability, is the term used to describe the variation in the R–R interval.

2.3.1 Acquisition and Conditioning

The acquisition of an R–R interval signal, that is, a series of consecutive R–R intervals, has the sampling rate as its main condition. If the signal is crudely discretized, there is a risk of loosing small but important beat-to-beat variations. A good sampling frequency ranges between 200 - 225 Hz, whereas the commonly used range of around 128 Hz can be deemed adequate. [9]

An R–R interval signal can be easily acquired with the use of a Holter monitor, which is an ambulatory electrocardiography device used for continuous monitoring of the electrical activity of the cardiovascular system. Holter recordings typically use a sampling frequency that ranges between 100 and 128 Hz.

Nevertheless, signals that have been acquired from Holter recorders present artefacts, called non normal beats, that are mainly the result of noise bursts. To make the recording more reliable, the signal needs to be filtered and modified so that non normal intervals are excluded from it. Because of that, the resulting signal is commonly referred to as a normal-to-normal, i.e., N–N interval signal. The annotation of non normal beats can be done manually, it is however a time consuming process, hence sometimes annotation algorithms and filters may be used to weed out the artefacts.

In this work, the signals we refer to as R–R interval signals are actually N–N interval

[3] "SinusRhythmLabel" by Created by Agateller (Anthony Atkielski), converted to svg by atom. en:Image:SinusRhythmLabels.png. Licensed under Public Domain via Wikimedia Commons https://commons.wikimedia.org/wiki/File:SinusRhythmLabels.svg#/media/File:SinusRhythmLabels.svg.

signals, that is, signals who have been left with normal-to-normal beats.

2.3.2 Time Domain Measurements

The simplest Heart Rate Variability evaluation methods to perform are, apparently, the time domain methods. Time domain methods can be separated into thee subcategories [11]; *simple time domain variables*, *statistical methods* and *geometrical methods*. Time domain methods are either measured over the whole signal, or over smaller periods to evaluate their evolution and relate it to clinical conclusions. Note that, time domain methods are less affected by the presence of artefacts in the signal.

Simple Variables

Simple time domain variables include the *mean R–R* (or *mean N–N*) interval, the mean Heart Rate (HR) and the difference between the longest and the shortest R–R interval. These variables are easy to obtain and analyse, while they convey significant information towards the prognosis and diagnosis of a number of heart conditions. Consequently, they have been frequently used in studies, demonstrating variations between healthy an non-healthy patients.

The mean R–R interval has been used quite extensively to evaluate the relation between autonomic neuropathy and diabetes. In a research conducted by Murray et al. in 1975 [12], where 42 male diabetics are examined to assess whether they suffer from autonomic neuropathy, it is demonstrated that diabetic patients exhibit a smaller mean R–R, which leads to the conclusion that diabetes may induce autonomic nerve damage. Since the condition is symptom-free, measuring the R–R interval could help anticipate it. Later, Bennett et al. [13] further investigate this issue and conclude that the mean R–R interval should be measured during deep breathing since it is only during those tests that they observe a difference in the measurement between patients; the subjects of their research are two groups of diabetic patients, one group having diagnosed autonomic neuropathy, and the other not presenting the condition. Later, Ewing et al. in [14] and [15] use the difference and rate between the longest and the shortest R–R interval among other measurements, in small, 1-minute recordings in an attempt to evaluate the correctness of measurements used to detect autonomic disorders in diabetic patients. Recently, in 2008, Ziegler et al. [16] also showed that the difference between longest and shortest R–R interval is related to autonomic neuropathy in diabetic patients, in a research that lasted nine years and included 1720 patients out of which 160 were diabetic. Indeed, he observed that a low difference suggests a high mortality risk.

Naturally, other studies have used these simple HRV variables as well. A study conducted by Algra et al. [17] demonstrates that abnormal mean Heart Rates could be related to sudden death. Taylor et al. [18] use the mean R–R interval to study the possibility that sympathetic and vagal mechanisms are related to the survival of patients who have suffered a myocardial infarction (MI) . In their research, Hogue et al. [19] show that the

risk of a patient developing atrial fibrillation after coronary artery bypass graft surgery is closely related to the patient's Heart Rate, a conclusion reached by Huikuri et al. [20] as well.

A research led by Ortak et al. in 2005 [21] shows that modern treatment of post myocardial infarction patient has improved, an conclusion that can be made by observing the stability of the mean Heart Rate versus other HRV markers. The research conducted by Matt et al. [22] gives a different context to the aforementioned simple variable. In their work, they investigate the how the quality of a dual-source CT coronary angiography is affected by the mean HR. A very recent research by Shalev et al. [23] investigates the effect of post traumatic stress disorder (PTSD) on a patient's mean Heart Rate, recording higher readings for those who present it.

Statistical Methods

Statistical methods are typically considered over long signals of the order of 24 hours. The most straightforward and popular measurement is the standard deviation of the R–R (or N–N) interval (SDNN) . It can be computed over the whole signal, yet it conveys more information when calculated over shorter periods.

Two other important statistical HRV measurements are the standard deviation of the average interval (SDANN) and the SDNN index (SDNNi). The SDANN serves the purpose of investigating changes in the average interval, calculated over short periods (usually 5-minute time windows). The average interval is calculated over those small time widows, and the standard deviation is calculated to measure changes occurring in longer cycles. The SDNNi is instead calculated to asses the variability in shorter cycles The SDNN is evaluated over 5-minute periods, ensued by the calculation of their average. Note that these markers are mostly computed on 24-hour records, yet it is still debatable whether this is the most appropriate way of extracting every piece of useful information possible. As stated by the Task Force [11], the SDNN highly depends on the length of the recording.

Statistical measurements, especially the SDNN, are present in almost every study that analyses HRV markers. The SDNN is measured in [12, 13, 14, 15] as another HRV marker related to the development of asymptotic diabetic autonomic neuropathy. It is also used by [18, 19, 20, 24] in their respective researches. Both [17] and [21] compute and compare the SDANN and SDNNi in their research.

Congestive heart failure (CHF) is a condition where the heart empties itself inadequately, an event that leads to a decrease in the blood supply of organs. High pressure forms in the veins as a result, which in its turn causes fluid to leak from capillary blob vessels [25]. A research to evaluate whether one can reproduce the outcome of heart rate variability measures, i.e. SDANN and SDNNi, conducted by Hoogenhuyze et al. [26], showed that the results obtained by examining these markers are indeed replicable. The study included two groups of patients, one group of asymptotic patients and a group of patients with Congestive Heart Failure (CHF). In their research, Bilchick et al. show that

HRV can be used as a prognostic tool for CHF [24], namely, using the SDNN measurement.

The SDNNi marker is shown by Liu et al. [27] to be an important predictor of sudden cardiac death and cardiovascular death in patients who have survived an acute MI. Antelmi et al. [28] study a number of time and frequency domain measurements on 653 healthy patients to establish possible connections with their age, gender, body mass index and functional capacity. Among others, they compare the SDNN, SDANN and SDNNi of the subjects. They observed that their values decrease with age, however, only SDNNi varies among different genders.

Another family of statistical method, derived from the differences between consecutive R–R intervals, include the *square root of the mean squared differences of successive N–N (R–R) intervals* (RMSSD), and the number of consecutive N–N (R–R) intervals differences that exceed 50 milliseconds (NN50). Regarding the NN50, what is commonly used instead of the number, is the NN50 divided by the total number of intervals, (pNN50). These methods are highly corelated and demonstrate high frequency heart rate variations.

Previously mentioned studies that also use these markers include [14, 19, 28, 21]. They are also used in the research project led by Tuji et al. [29], a large study conducted to asses the effect of reduced HRV on risk for a cardiac event.

Geometrical Methods

The last class of time domain measurements includes geometrical methods. Geometrical methods can be measured over geometric patterns that occur from the conversion of the R–R interval signal. Some conversion methods include the sample density distribution of R–R intervals the sample density distribution of differences between consecutive R–R intervals. Upon conversion, the analysis of the geometric pattern can be performed using one of the following three approaches.

- 1^{st} approach: Convert a basic measurement of the geometrical pattern into an HRV measure. This approach includes the two most well-known geometrical measurements, i.e. the HRV triangular index and the Triangular Interpolation of R–R (N–N) interval histogram (TINN).
- 2^{nd} approach: Interpolate the geometrical pattern with a mathematically defined shape, and use the parameters of the said shape as an HRV measure.
- 3^{rd} approach: Classify the geometrical pattern according to the occurring geometrical shape.

2.3.3 Other Measurements

Other HRV analysis measurements include Frequency domain methods, Rhythm Pattern analysis, and Non-Linear methods.

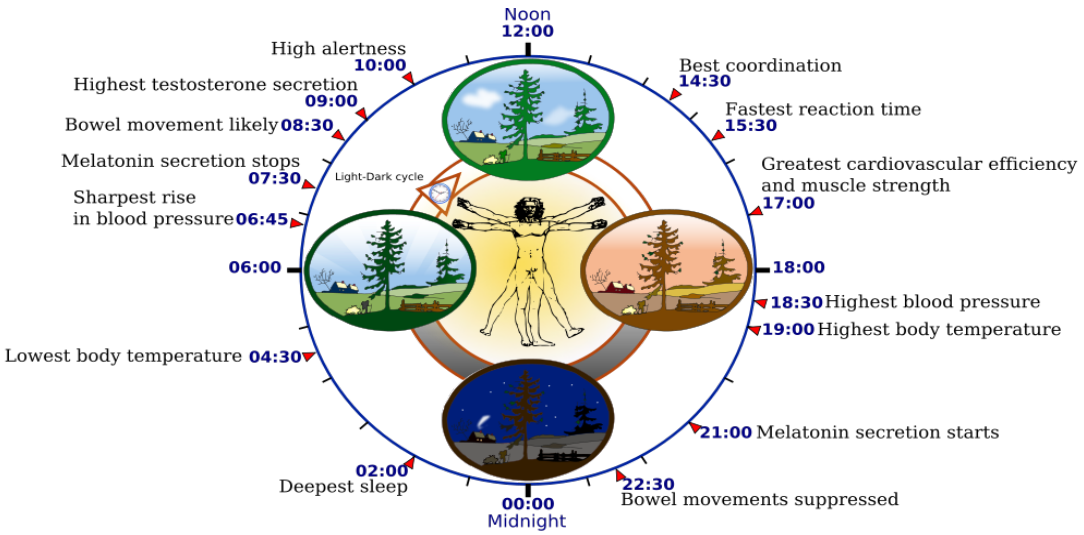


Figure 2.3: The typical circadian patterns of someone who rises early in morning, eats lunch around noon, and sleeps at night (10 p.m.).

2.4 The Human Circadian Rhythm

A circadian rhythm can be defined as a biological process that demonstrates an endogenous oscillation lasting approximately 24 hours, capable of being entrained, or brought into a specific rhythm.

Androsthenes, a ship captain serving under Alexander the Great, was the first to observe a circadian process when he described diurnal leaf movements of the tamarind tree. Jean-Jacques d'Ortous de Mairan in 1729 prodiced the first recorded observation of an endogenous circadian oscillation. He noted that 24-hour patterns in the movement of the leaves of the plant Mimosa pudica continued even when the plants were kept in constant darkness, in the first experiment to attempt to distinguish an endogenous clock from responses to daily stimuli [30]. These 24-hour rhythms have ever since broadly observed in plants, animals, fungi, and cyanobacteria.

Although circadian rhythms are endogenous, they are affected by stimuli that come from the environement including light, temperature and others. Figure 2.3 [4] demonstrates an example of the typical circadian human rhythm. According to [31], a biological rhythm is called circadian when it meets the three following criteria.

1. It has an endogenous free-running period lasting around 24 hours.
2. It can be reset by exposure to external stimuli.
3. It is able to maintain circadian periodicity over a range of physiological temperatures.

[4] "Biological clock human" by NoNameGYassineMrabetTalk fixed by Addicted04 - The work was done with Inkscape by YassineMrabet. Informations were provided from"The Body Clock Guide to Better Health" by Michael Smolensky and Lynne Lamberg; Henry Holt and Company, Publishers (2000). Landscape was sampled from Open Clip Art Library (Ryan, Public domain). Vitruvian Man and the clock were sampled from Image:P human body.svg (GNU licence) and Image:Nuvola apps clock.png, respectively. Licensed under CC BY-SA 3.0 via Wikimedia Commons – https://commons.wikimedia.org/wiki/File:Biological_clock_human.svg#/media/File:Biological_clock_human.svg.

2.4.1 The Circadian Rhythmiciy of HRV measurements

The circadian rhythmicity is observed in many human physiological variables, one of them being the Heart Rate. Huikuri et al. and Ekholm et al. [32, 33] were among the first to discover alterations to the circadian rhythm of cardiac neural regulation in patients with coronary artery disease and pregnant women respectively. Ewing et al. and Nakagawa et al [34, 35] noted that HRV markers such as, the f-QRS duration, HF, the LF/HF ratio, the hourly HR and R–R counts follow a circadian cycle, whereas Massin et al. [36] found that this circadian variation is present from late infancy or early childhood. Vandewalle et al. [37] indicate in their research that the autonomic control of HR is influenced by an endogenous cardiac rhythmicity. In their study, Yoshizaki et al. showed that HRV variables exhibit different patterns among different s leep-wake cycles when they studied rotating shift workers [38].

When averaging a set of time evolving signals, one must assume that their behavior is synchronized. All but one of the previously mentioned researches make that assumption, and they align and average HRV markers with respect to the horological time. Duffy et al. [39] are the only ones to assume different circadian patterns. They study *morning* and *evening* types of people, regarding their preference, and then manually align those signals using a *relative* time. Nonetheless, while it is the only research we found that considers different cycles, the alignment is manually performed according to each subject's preference, and no quantitative marker is provided for future development.

Chapter 3

Alignment and Classification Algorithms

3.1 Pre-processing the R–R Interval Signal

We consider an R–R interval signal corresponding to the i^{th} subject as a discrete signal $s_i(t)$, $t \in 1, \ldots, N$, where $N+1$ is the number of recorded R waves. Each value $s_i(1), s_i(2), \ldots, s_i(N)$ is measured in seconds and represents the interval between two consecutive R waves. We denote L_i the signal length in seconds, $L_i = \sum_{t=1}^{N_i} s_i(t)$. For the remainder of this section, the subscript i will be omitted since the preprocessing is the same for all subjects.

Figure 3.1 illustrates an R–R interval signal. Here, the signal length is $L = 81736$ sec and the number of intervals is equal to $N = 118700$.

Each signal undergoes a pre-processing procedure before analysis, composed of a denoising and a segmentation step. Those steps are further described in the next subsections. After the pre-processing is complete, the result is a set S, an ordered set of signals and appointed labels $S = \{(s^1(t), l^1(t)), (s^2(t), l^2(t)), \ldots, (s^W(t), l^W(t))\}$. The labels $l^j(t), j = 1, \ldots, W$ denote whether interval $s^j(t)$ is *noise* or not. The lengths L^j of $s^j(t)$ are equal. Pairs $(s^j(t), l^j(t))$ are also referred to as *time windows*.

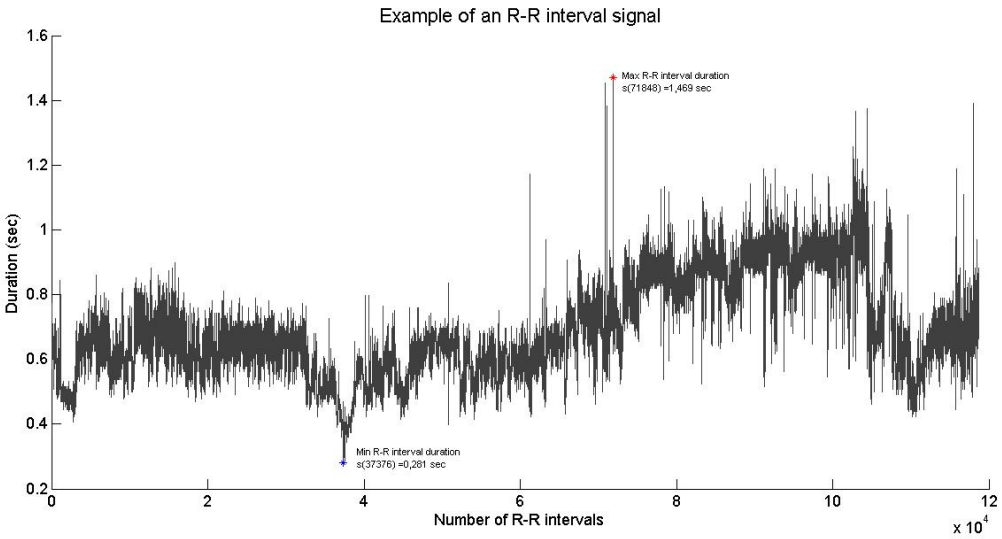


Figure 3.1: Example of an R–R interval signal. This signal comes from subject number 52 of the NSR database (reference), described in section (reference).

3.1.1 Noise

The R–R intervals used in this work are assumed to have already been annotated by experts so that they are clear of detected artefacts. However, to limit the noise even further, two filters are applied to each signal $s(t)$ aiming to mark noisy records. Each record $s(t)$ is appointed a label $l(t)$ that may take two values; thus, $l(t) = 1$ if the corresponding interval is labelled *not noise*, and $l(t) = 0$ if it is labelled *noise*. Initially all labels are labelled *not noise*, $l(t) = 1$.

The first filter is a simple one; it consists of marking as noise every interval superior than 5 seconds and every record inferior to 0.2 seconds. Intervals that exceed these limits are assumed not to be natural intervals between heartbeats. This margin is in fact a much wide one. Indeed, a 5 and 0.2 second interval correspond to a heart rate of 12 and 300 beats per minute respectively. Those two limits are far from the normal human heart rate, which is usually between 50-100 bpm. Cases of *bradycardia* have been observed when the respiratory rate decreases, but they have not been shown to drop below 25 bpm [9].

The second one is a filter that marks as noise every interval that is different that the previous or the next one by more than 10%. Thus, $l(t)$ and $l(t-1)$ are set equal to *noise* if $s(t)/s(t-1) < 0.90$ or $s(t)/s(t-1) > 1.10$. This filter is applied to make sure that the remaining intervals are normal heartbeats (thus not ectopic heartbeats).

After applying these filters, the total length of the signal remains unchanged. The intervals are not deleted from the signal, so that the actual length is preserved. An additional processing step takes place to facilitate the segmentation process; since it is important that all the time windows have an equal length after the division into time windows, every interval marked as noise is processed so that it does not exceed 1 second. Noisy intervals that exceed that margin are separated into *dummy* intervals that follow

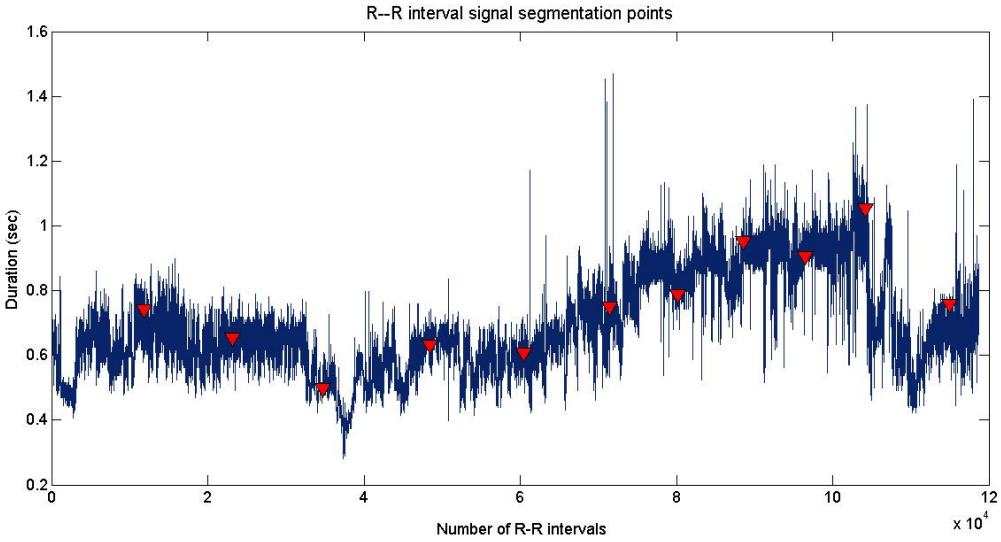


Figure 3.2: Example of an R–R interval signal segmentation, using the signal from Figure 3.1. Each triangle denotes a segmentation points. Here, the signal is divided into twelve 2–hour time windows.

that requirement. Naturally, only intervals labelled as noise go through that procedure.

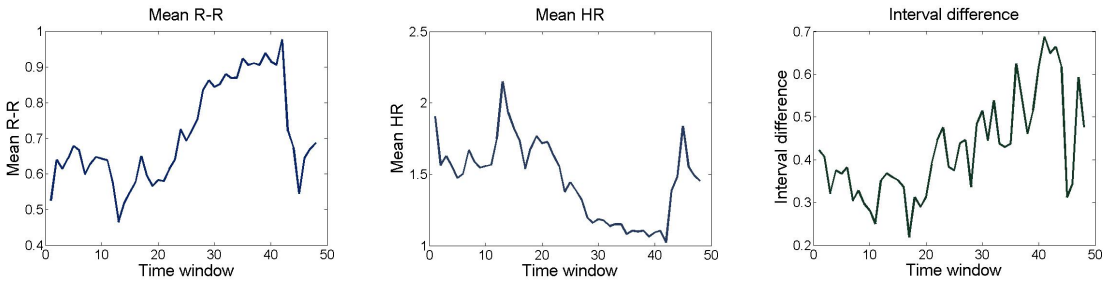
For the next section, signal $s(t)$ and labels $l(t)$ are assumed to be the outcomes of that operation.

3.1.2 The segmentation process

For the purpose of computing and analysing the HRV measurements described in section 3.2, each signal is divided into time windows of equal length. This section provides basic notations and definitions that will be used to describe the R–R interval signal segmentation. Moreover, we describe the segmenting process in a general way to account for various time widow lengths. Note that this step does not change the values of the labels that denote whether an interval is noise or not.

The R–R interval signal is segmented into a number of time windows of equal length. The time window length T (measured in seconds) defines the number $W = L/T$ of time windows $s(t)$ and $l(t)$ are divided into. $s^j(t)$ and $l^j(t)$ designate the signal and labels that corresponds to the j^{th} time window, $j = 1, \ldots, W$. It follows that $\sum_{t=1}^{N^j} s^j(t) = T$ and $\sum_{j=1}^{W} N^j = N$.

When the division product of L/T is not an integer, W is the quotient of the division plus one. In that case, every time window has an equal length except for the last one, thus, $\sum_{t=1}^{N^j-1} s^j(t) = T$ and $\sum_{t=1}^{N^W} s^W(t) = T^W < T$. The reason we divide the signal in such a way is that we want to have equal time windows even when processing different signals. In this way, when working with R–R interval signals obtained from different subjects and possibly having different lengths, we know that the segmentation process will output windows of equal length. Furthermore, as explained in the beginning of this



(a) Mean R–R (b) Mean HR (c) Interval Difference

Figure 3.3: Examples of simple variable measurements for 49 time windows, each one having a 30 minute length.

section, the last window is also viewed as a segment whose length is equal to T with the difference $T - T^W$ being eventually treated as noise.

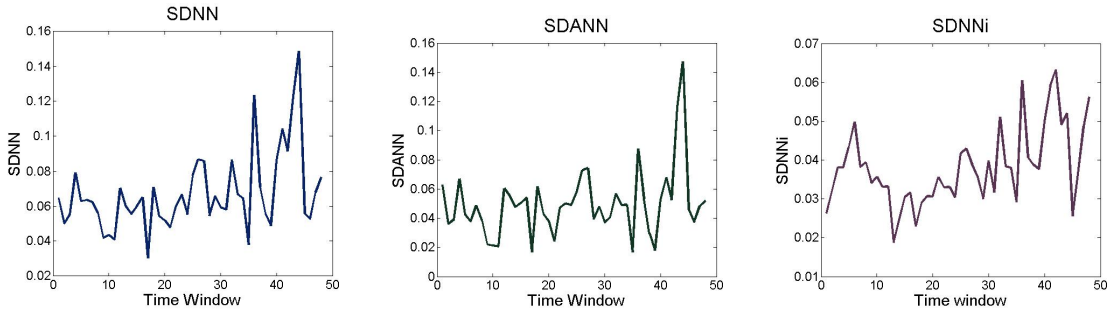
Figure 3.2 depicts an example of a signal segmentation. In this example, the R–R interval signal from figure 3.1 is segmented into $W = 12$ time windows, each having a length of than $T = 7200$ seconds except for the last one, whose length is $T^W = 2536$.

3.2 Calculating the HRV Measurements

In chapter 2, section 2.3, we introduced some known HRV measurements. In this section we present those who belong to the class of Time Domain methods, as described in section 2.3.2. In particular, we present the computation of simple variables and statistical methods (sections 2.3.2 and 2.3.2). We compute each measurement over the ordered set S that contains the time windows obtained by processing the signal as described in the previous section. The output of each computations is governed by one user defined parameter; the length of the time window T.

Therefore, given an R–R interval signal $s(t)$ and a time window length T, the first two steps are the pre processing steps, where $s(t)$ is transformed into the ordered set S. The following nine measurements can be computed over S, resulting to nine new signals, that will de generally referred to as *measurement signals*, so that they are not confused with $s(t)$, i.e. the R–R interval signal.

- The mean R–R, $M(x)$.
- The mean Heart Rate, $H(x)$.
- The difference between largest and shortest interval, $D(x)$.
- The SDNN, $V(x)$.
- The SDANN, $A(x)$.
- The SDNNi, $I(x)$.



(a) Standard Deviation (SDNN)

(b) Standard deviation of the average interval (SDANN)

(c) Average of the standard deviation (SDNNi)

Figure 3.4: Examples of statistical measurements for 49 time windows, each one having a 30 minute length. For the SDANN and the SDNNi, the small windows have a 5-minute length.

- The RMSSD, $R(x)$.
- The NN50 and pNN50, $F(x)$ and $P(x)$.

Note that we have decided to measure the evolution of all these markers, which implies that measurements such as the SDANN, SDNNi, RMSSD, usually computed on a 24-hour recording, are calculated over much smaller time windows. Our motivation behind this choice, as explained in the introduction, is to convey whether we can extract new information using different variations of parameters, such as the length of the time window.

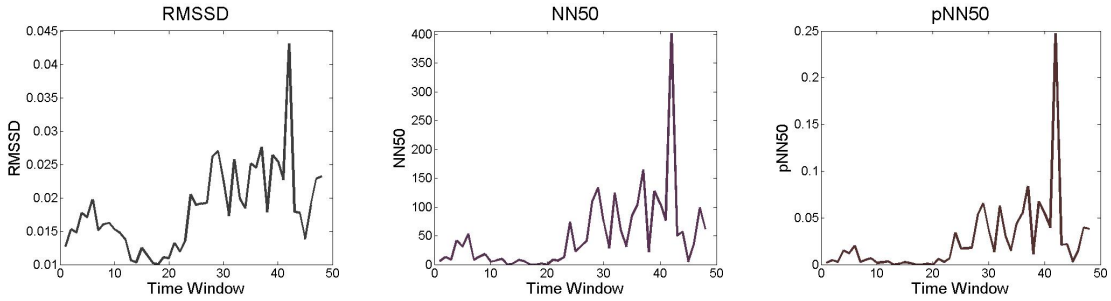
To illustrate each resulting measurement signal, we provide an exemplary ordered set S', that corresponds to the pre-processing of the R–R interval signal of subject number 52 of the NSR database (reference), the same one that was used in figures 3.1 and 3.2. The set S contains $W = 49$ time windows whose length is $T = 1800$ sec (30 minutes). Chapter 4 includes all sets of experiments.

We can see the evolution of the simple time domain variables on figure 3.3. Figure 3.3a shows the mean R–R interval signal $M(x)$ that was computed for S'. In figure 3.3b we can observe the evolution of the mean HR $H(x)$, while figure 3.3c illustrates $I(x)$, i.e. the difference between the largest and the shortest interval for each time widow. An early and obvious remark that can be made, is that the mean R–R is the inverse of the mean HR.

Statistical measurements are displayed on figure 3.4. We can observe the SDNN signal $V(x)$, the SDANN signal $A(x)$ and the SDNNi signal $I(x)$. The first two signals seem to be similar, while the range of the SDNNi signal is smaller than the other two.

Figure 3.5 shows statistical measurements that obtained from interval differences. Again, this class of signals shares a common pattern. As expected, the NN50 and pNN50 signals $F(x)$ and $P(x)$ convey the exact same information, which is why only $P(x)$ is refferenced in the results.

The following subsections describe in a more detailed way how each measurement was



(a) Root of the Mean Squared Successive interval Differences (RMSSD)

(b) Number of successive interval differences greater than 50ms (NN50)

(c) Percentage of successive interval differences greater than 50ms (pNN50)

Figure 3.5: Examples of statistical measurements based on successive interval differences for 49 time windows, each one having a 30 minute length.

computed. Recall that N^j is the size of the j^{th} time window (the number of intervals) and W is the cardinality of S (the number ot time windows).

3.2.1 Mean R–R

The mean R–R consists of finding the mean R–R interval over a discrete and continuous time window. Given an R–R interval signal set S, we compute the mean R–R over each segment $s^j(t)$ in S to obtain $M(x)$.

$$M(x) = \sum_{t=1}^{N^x} s^x(t) l^x(t), \text{ for each } x = 1, \ldots, W \tag{3.1}$$

3.2.2 Mean HR

The Heart Rate and its computation has been defined and analysed in chapter (reference), section (reference). Accordingly, we obtain the mean HR in by calculating the mean Heart Rate that corresponds to each time window in S. To do so, we use the previously computed mean R–R interval, $M(x)$ as presented in the following equation.

$$H(x) = \frac{M(x)}{N^x}, \text{ for each } x = 1, \ldots, W \tag{3.2}$$

3.2.3 Difference Between Largest and Shortest Interval

Calculating this measurement is fairly straightforward, as seen in equation 3.3. It amounts to finding the largest interval that is *not noise* and subtracting from that value the shortest interval that is *not noise*. Naturally, it is computed for each time window.

$$D(x) = \max\left(s^x(t) l^x(t)\right) - \min\left(s^x(t) l^x(t)\right) \text{ for each } x = 1, \ldots, W \tag{3.3}$$

3.2.4 SDNN

Even though the standard deviation of the R–R interval signal is usually a single number calculated over the whole signal, we compute $V(x)$ fore each time window. Let $N^{x'} = \sum_{t=1}^{N^x} l^x(t)$ be the number of intervals that are not noise, and $\mu^x = \frac{1}{N^{x'}} \sum_{t=1}^{N^x} s^x(t) l^x(t)$, their average, the SDNN signal V(x) is calculated in the following manner.

$$V(x) = \sqrt{\frac{1}{N^{x'} - 1} \sum_{t=1}^{N^x} \left| \left(s^x(t) - \mu^x\right) l^x(t) \right|^2} \tag{3.4}$$

3.2.5 SDANN

As described in subsection 2.3.2, the SDANN describes a standard deviation value. In particular, the R–R interval signal is divided into 5-minute segments, the average value per segments is calculated, and the SDANN is the standard deviation of these average values. Usually the signals are large, 24-hour recordings, whereas in our signals we calculate the SDANN for each time window. Thus, each $s^j(t)$ is segmented into smaller time windows, in the same way described in subsection 3.1.2.

This means that each time window is in its turn is transformed into a set S^j of smaller time windows. The new composition of the ordered set of signals S is described by equations 3.5 and 3.6.

$$S^j = \{(s^{j^1}(t) l^{j^1}(t)), (s^{j^2}(t) l^{j^2}(t)), \ldots, (s^{j^{W^j}}(t) l^{j^{W^j}}(t))\} \tag{3.5}$$

$$S = \{S^1, S^2, \ldots, S^W\} \tag{3.6}$$

The next step is to compute the mean R–R, $M^j(y)$ as described by equation 3.1, for each set S^j, hence, equation 3.6 becomes

$$S = \{M^1(y), M^2(y), \ldots, M^W(y)\}. \tag{3.7}$$

Now, W^j denotes the length of each mean R–R signal M^j, $j = 1, 2, \ldots, W$. To obtain A(x), we need to calculate the standard deviation of each M^j. Let $\mu^x = \frac{1}{W^x} \sum_{y=1}^{W^x} M^x(y)$ be the mean of each M^j. Equation 3.8 describes the final calculation of A(x).

$$A(x) = \sqrt{\frac{1}{W^x - 1} \sum_{y=1}^{W^x} |M^x(y) - \mu^x|^2} \tag{3.8}$$

3.2.6 SDNNi

The SDNNi is the average of a series of standard deviation measurements (over small windows), thus it is a measurement that also requires a second segmentation per time window. Proportionally to the computation of the SDANN described in subsection 3.2.5,

the standard deviation per small segment is calculated, and the SDNNi is the average of these values.

Consequently, each signal $s^j(t)$ is segmented into smaller segments as previously described by equations 3.5 and 3.6. For each ordered set S^j, the SDNN signal $V^j(y)$ is calculated as in equation 3.4, and S becomes:

$$S = \{V^1(y), V^2(y), \ldots, V^W(y)\}. \tag{3.9}$$

Once more, W^j denotes the length of each SDNN signal V^j, $j = 1, 2, \ldots, W$. The SDNNi signal I(x) is given by equation 3.10.

$$I(x) = \frac{1}{W^x} \sum_{y=1}^{W^x} V^x(y) \tag{3.10}$$

3.2.7 RMSSD

The RMSSD is a statistical measurement that describes the square root of the mean squared differences of successive N–N (R–R) intervals. Much like other statistical measurements, it is computed over 24-hour recordings.

Let $N^{x'} = \sum_{t=2}^{N^X} l(t)l(t-1)$ denote the number of differences of successive R–R intervals that do not include any interval labelled as noise. We compute the RMSSD signal $R(x)$ over each time window in the following way.

$$R(x) = \sqrt{\frac{1}{N^{x'}} \sum_{t=2}^{N^x} \Big[\big(s^x(t) - s^x(t-1)\big) l(t)l(t-1) \Big]^2} \tag{3.11}$$

3.2.8 NN50 and pNN50

The NN50 and pNN50 are grouped together, since their result conveys the same information, their only difference being the measurement unit, which is a number for NN50 and a percentage for pNN50. In chapter 4, only the pNN50 signal $P(x)$ is presented.

The NN50 marker measures the number of interval differences having a value greater than 50ms. We compute this number as follows.

Consider the function $p^j(t)$ over a time window $(s^j(t), l^j(t))$.

$$p^j(t) = \begin{cases} 1 & \text{if } \big(s^j(t) - s^j(t-1)\big) l^j(t)l^j(t-1) > 0.05 \\ 0 & \text{if } n \text{ otherwise} \end{cases}$$

Hence, p is equal to one if the difference between successive R–R intervals that are labeled not noise is greater than 50ms. Consequently, if $N^{x'} = \sum_{t=2}^{N^X} l(t)l(t-1)$ is the number of differences of successive R–R intervals that do not include any interval labelled as noise, $F(x)$ is computed by equation 3.12, and $P(x)$ by equation 3.13.

$$F(x) = \sum_{t=2}^{N^x} p^x(t) \tag{3.12}$$

$$P(x) = \frac{1}{N^{x'}} \sum_{t=2}^{N^x} p^x(t) \tag{3.13}$$

Remark 3.2.1. We calculate each measurement in a way that makes it a series of successive values, which is why we choose to represent it as a signal. The length of the segments (as well as the length of the smaller segments described in subsections 3.2.5 and 3.2.6) are parameters of our computations.

Remark 3.2.2. We chose to represent the noise in a way that keeps the time continuity of the R–R interval signal, so that it has the same length in seconds before and after the pre-processing. On the other hand, we use the label values to exclude it from computations, as if it did not exist in the first place. What we mean to express is that noisy records may be included in the segmentation process so that one R–R interval (or measurement) signal can be compared to another, however the noisy intervals must not interfere with the HRV marker values. Imagine the case where, while a subject is asleep, he or she moves in a way that disrupts the recording. The R–R interval signal portion corresponding to that period is not to be taken into account. Yet, if two events have occurred, one before and one after that disruption, they should have the correct amount of seconds between them.

3.3 Alignment Methods

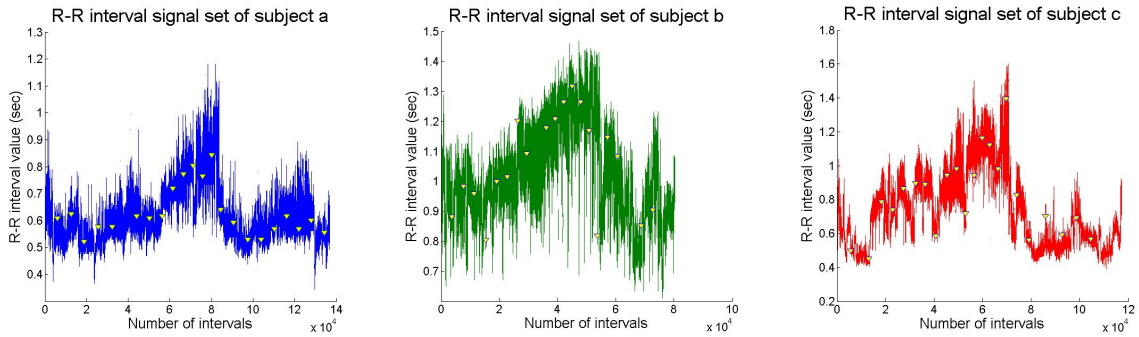
In the preceding sections we described how each R–R interval signal is pre-processed (section 3.1), which Time Domain measurements are studied in this work and how they are derived (section 3.2). For each $s_i(t)$ we extract nine discrete measurement signals.

In this section we examine the case where we have a set of signals $\{s_1(t), s_2(t), \ldots, s_I(t)\}$, that come from the same dataset. We also assume that these R–R interval signals correspond to at least 24-hour recordings.

3.3.1 The Importance of Alignment:Usual Methods

Cases that include R–R interval signals or measurement signals coming from 24-hour recordings are abundant in clinical researches, and a few examples, such as **(REFERENCES)** have already been discussed in previous chapters. It is thus safe to say that when such a set is to be used in a research, either the average signal will be computed, or the set will be split into subsets that share common characteristics and the average measurement signal will be computed for each group.

An issue that arises upon calculating the average measurement signal, or even the average R–R interval signal, is the definition of a starting point for each signal. The signals we are studying are recorded from live subjects, hence, when we seek to compute



(a) Signal set S_a (b) Signal set S_b (c) Signal set S_c

Figure 3.6: R–R interval signals coming from three subjects, a, b and c. The signals have been pre-processed, and the displayed signal is without noise. Yellow triangles denote the segmentation point. The time window length is $T = 3600$ sec

the average, we have to assume that all the recordings start at the same time of the day. Moreover, it would be normal to assume that the synchronization needs not always to be based on the horological time, but rather, either on a common event, or according to a common pattern.

Suppose a simple example including three subjects: subjects a, b and c and their corresponding R–R interval signals, $s_a(t)$, $s_a(t)$ and $s_c(t)$. Figure 3.6 show the pre-processed ordered signal sets $S_a(t)$, $S_b(t)$ and $S_c(t)$. For each signal, we compute the Mean HR signal $H_i(x)$ $(i \in \{a, b, c\})$.

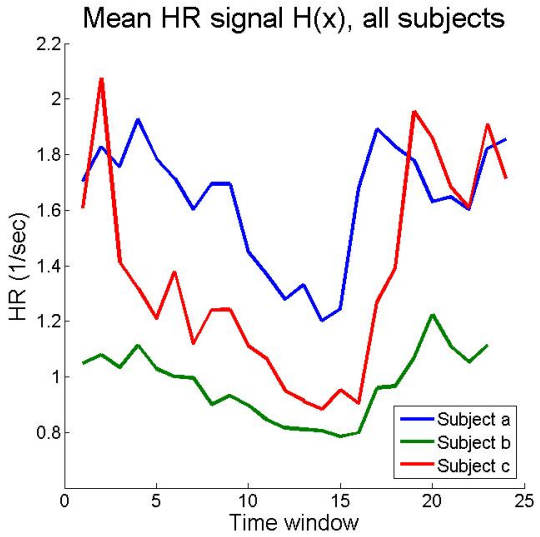
The resulting measurement signals can be seen on figure 3.7a. We can already remark that subjects a and c exhibit a similar pattern after the 9^{th} time window, however they seem to be out of synch by about two time windows. The mean HR values of subject b are lower than the other two, but we can observe an increase that lasts 3 time windows, from the 18^{th} to the 20^{th}. The same increase occurs for subject a between time windows 15 and 17, and for subject c between time windows 17 and 19.

For these three measurement signals, we compute the corresponding average measurement signal, $H(x)$, in the following simple way.

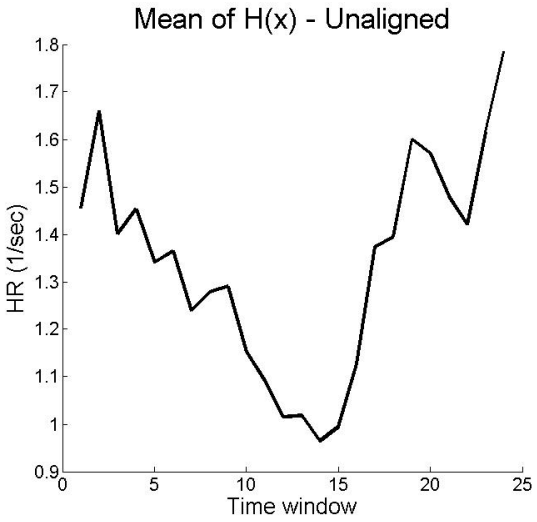
$$H(x) = \sum_{i=1}^{I} H_i(x) \tag{3.14}$$

In this example, $I = 3$. We can see the resulting average signal $H(x)$ on figure 3.7b. A first examination of this figure provides us with a noticeable remark. We can observe that the average value of the last time window is much different from the average value of the first, something that we do not see in figure 3.7a, where the value of the last time window is close to the value of the first one. This is due to the fact that the duration of the recording corresponding to subject b is shorter than the other two recordings, hence the last time window of S_b is empty, i.e. all labels $l_b^{j=24}(t)$ are noise, $l_b^{j=24}(t) = 0 \; \forall \; t \in \{1, 2, \ldots, N^{24}\}$.

To solve this issue, one might not consider the last time window at all, which would cause the resulting signal to end close to where it starts. Yet this means discarding



(a) Measurement signals $H^a(x)$, $H^b(x)$, $H^c(x)$



(b) The average, $H(x)$ without alignment

Figure 3.7: Mean HR signals computed from S_a, S_b and S_c. The average signal $H(x)$ has been computed with no alignment.

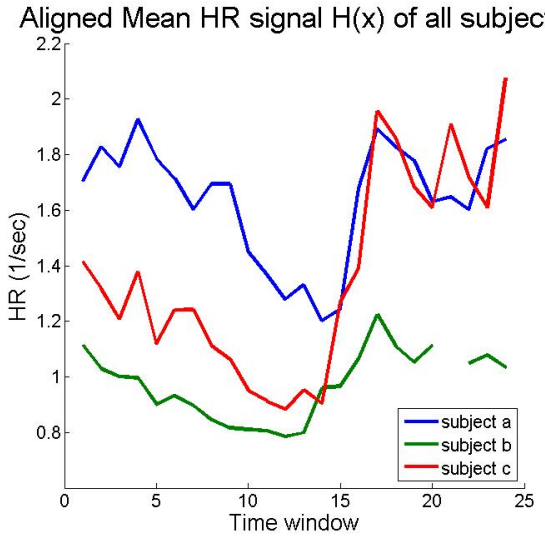
information from subjects a and c. Moreover, if we generalize this example to a larger set of signals, the amount of discarded information could be higher.

Another comment that can be made is that, in this example, the slight misalignment of the signals does not have a significant effect on the outcome of the averaging. Nevertheless, in an extreme case where the recordings would be completely misaligned, the average measurement signal could turn out to be completely different from the individual signal patterns. In that case, the researcher is faced with one of two solution. Either perform some manual averaging of the data, or discard the dataset and obtain a new one, if available.

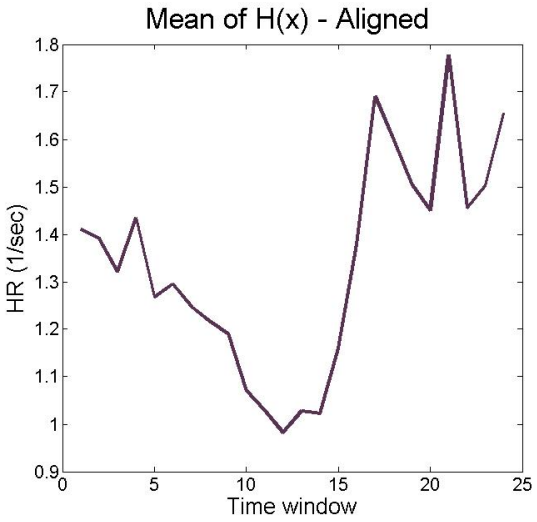
To avoid removing any information, we have manually aligned the three Mean HR signals, using the edges corresponding to the previously mentioned increases in the value of the measurement signal. The three aligned measurement signals are depicted in figure 3.8a. The three measurement signals now appear to be synchronised with respect to their patterns. The increase is much more noticeable in the resulting average signal $H(x)$, seen on figure 3.8b, and the difference between the last and the first time window is now less pronounced.

We can thus reach the following initial conclusion. For a set of R–R interval signals that correspond to different subjects, the alignment must have a reference point. It is safe to assume that, this reference point can depend on the nature of the data, and the signals may not always be aligned.

The aim of this work is to propose automated alignment methods that respect the pattern of the given R–R interval signals. We therefore present the two following methods; The *Puzzle Piece Alignment Method*, and the *Event Based Alignment Method*.



(a) Measurement signals $H^a(x)$, $H^b(x)$, $H^c(x)$



(b) The average, $H(x)$ with alignment

Figure 3.8: Manual alignment of the Mean HR signals H_a, H_b and H_c, with respect to the three visible edges between the 15^{th} and 20^{th} time windows seen in figure 3.7a.

3.3.2 The Puzzle Piece Alignment Method

A first approach when trying to manually align two signals, is to treat them like puzzle pieces; move them until they match. This initial idea resulted in the *Puzzle Piece Alignment Method* (PPA). A simple way to sum up the PPA method, is to say that each signal is treated like a puzzle piece: one by one, signals are added into the pool of aligned signals, so that in the end of the process, they all match.

Analysis

To describe the way we fit two signals, we need to describe what we perceive as the minimum distance between them. Assume that we have two discrete finite signals of equal length, $m^1(x)$ and $m^2(x), x \in [0, W-1]$. We define as *pointwise distance* between $m^1(x)$ and $m^2(x)$ the average pointwise absolute difference between those two signals, $d = \frac{1}{W}\sum_{x=0}^{W-1} |m^1(x) - m^2(x)|$. We can thus define a pointwise distance function with respect to $n \in [0, W-2]$ as follows:

$$d(n) = \frac{1}{W}\sum_{x=0}^{W-1} |m^1(x) - m^2((x+n) mod(W)| \tag{3.15}$$

We have defined our pointwise distance function using the euclidean distance between two points as our basis. To establish whether this function has a minimum or not, we proceed by analysing its derivative, to see under which circumstances there exists a value of n for which $\frac{\partial d(n)}{\partial n} = 0$.

The pointwise distance function described in 3.15 is minimized with respect to n when

its derivative is equal to zero:

$$\frac{\partial d(n)}{\partial n} = 0 \tag{3.16}$$

Upon deriving $d(n)$ and substituting the result to 3.16 we obtain

$$\frac{\partial}{\partial n} \sum_{x=0}^{W-1} |m^1(x) - m^2((x+n) mod(W)| = 0 \tag{3.17}$$

Instead of writing the modulo W operation, we assume without loss of generality that the signals repeats themselves infinitely every W values, which means that $m^i(x) = m^i(x+W)$, $i \in \{1,2\}$, $\forall x \in \mathbb{Z}$. We define the following function:

$$a(x,n) = \begin{cases} |m^1(x) - m^2(x+n)| & \text{when } m^1(x) \neq m^2(x+n) \\ 0 & \text{when } m^1(x) = m^2(x+n) \end{cases}$$

This allows us to write 3.17 in the following way.

$$\sum_{x=0}^{W-1} \frac{\partial a(x,n)}{\partial n} = 0 \tag{3.18}$$

We study the partial derivative of $a(x,n)$.

$$\frac{\partial a(x,n)}{\partial n} = \begin{cases} \frac{m^1(x)-m^2(x+n)}{m^1(x)-m^2(x+n)} = 1 & \text{when } m^1(x) > m^2(x+n) \\ \frac{m^2(x+n)-m^1(x)}{m^1(x)-m^2(x+n)} = -1 & \text{when } m^1(x) < m^2(x+n) \\ 0 & \text{when } m^1(x) = m^2(x+n) \end{cases} \tag{3.19}$$

Observe that the left part of equation 3.18 becomes a summation of terms that are now equal to 1,-1 or 0, and can be rewritten as:

$$1 + 1 + \cdots - 1 - 1 \cdots + 0 + 0 \cdots = \mathcal{W} \tag{3.20}$$

Furthermore, for the pointwise minimum distance function $d(n)$ to have a minimum, we need $\mathcal{W}$ to be equal to 0. If this is true, then equation 3.20 becomes:

$$W^+ \cdot 1 + W^- \cdot (-1) + W^0 \cdot 0 = 0 \Rightarrow W^+ = W^-$$

What we can conclude from the above analysis, is the following. Let $m^1(x)$ and $m^2(x)$ be two discrete finite signals and let n^* be such that:

$$m^1(x) = m^2\big((x+n^*) mod(W)\big) \quad \forall\ x \in [0, W-1]$$

In this case, the pointwise distance function has a minimum at n^*, which is equal to zero, since:

$$\frac{\partial a(x, n*)}{\partial n} = 0 \quad \forall\ x \in [0, W-1]$$

Furthermore, this means that the two signals have also the same mean.

Algorithm: PPA

Input: A number I of mean HR signals $M^1(x), \ldots, M^I(x)$ of equal size W.
Output: The aligned mean HR signals $\mathbf{M}^1(x), \ldots, \mathbf{M}^I(x)$

for $i = 1, \ldots, I$ **do**
  $\mu = \mathbf{avg}\{M^i(x)\}$
  $\tilde{M}^i(x) \leftarrow M^i(x) - \mu$
$\mathbf{M}^1(x) \leftarrow M^1(x)$
$\overline{\mathbf{M}}(x) \leftarrow \tilde{M}^1(x)$
for $i = 2, \ldots, I$ **do**
  $m(x) \leftarrow \tilde{M}^i(x)$
  $D = \infty$, $n = 0$
  for $j = 1, \ldots, W - 1$ **do**
    $d = \mathbf{avg}|\overline{\mathbf{M}}(x) - m\big((x + j) \ \mathbf{mod}\ (W)\big)|$
    if $d < D$ **then**
      $D = d$
      $n = j$
  $\overline{\mathbf{M}}(x) \leftarrow \mathrm{avg}|\overline{\mathbf{M}}(x) - \tilde{M}^i\big((x + n) \ \mathbf{mod}\ (W)\big)|$
  $\mathbf{M}^i(x) \leftarrow M^i\big((x + n) \ \mathbf{mod}\ (W)\big)$
return $\mathbf{M}^1(x), \ldots, \mathbf{M}^I(x)$.

Algorithm 1: The *Puzzle Piece Alignment* Algorithm

Remark 3.3.1. In our work, we assume for the PPA algorithm that the processed signals have the same pattern, thus if the pointwise minimum distance between two such signals is to be found, they need to have the same mean.

The Algorithm

The PPA algorithm accepts a set of mean HR signals $M^i(x)$ that correspond to I subjects as its input. For each $M^i(x)$, we first calculate its mean μ^i and extract it, so that their mean is the same and equal to zero. Then, starting from the first normalized signal $\tilde{M}^1(x)$, we create an average normalized signal $\overline{\mathbf{M}}(x)$, and for each normalized signal $\tilde{M}^i(x), i \in \{2, \ldots, I\}$ we calculate the index n^i that generates the minimum pointwise distance D^i between $\overline{\mathbf{M}}(x)$ and $\tilde{M}^i(x)$. The index n^i is used to generate the aligned HR signal, $\mathbf{M}^i(x) = M^i(x + n^i)$, which is the output of the PPA algorithm. Note that, for $i = 1$, $n^1 = 0$, since the first signal serves as a reference point.

The PPA method is presented in algorithm 1. Our algorithm thus performs a greedy search to find the global minimum of the pointwise distance function between the averaged curve and each new mean HR signals. If no global minimum exists, it returns the first local minima it finds. The complexity of PPA is O(IW).

3.3.3 The Event Based Alignment Method

The PPA algorithm assumes that all signals have the same, or, a very similar behaviour. As we have already stated in chapter 2, section 2.4 this is not always the case. Often, the inherent circadian cycle differs between two people. Furthermore, the signal could be disturbed by phenomena that cause an acceleration or a deceleration of the heart beat, like, among others, physical exercise, anxiety, and undetected heart conditions. The Event Based Alignment (EBA) algorithm was developed to categorize and align signals like that, by detecting features in each signal and categorizing them accordingly.

Analysis

We name *events* the features we have selected to use for the categorization process. An event can be informally described as a short-timed and important increase of the mean HR signal, in such a way that it can be detected as an edge by an edge detector.

In this work, we use the *Optimal Edge Detector* to detect events in the signal. The Optimal Edge Detector was developed by Deriche in 1987 [40]. While it was designed for two dimensional signals, i.e. 2D images, it can be used to detect edges in 1D signals as well.

The algorithm uses a filter $f(x)$ to detect the edges in the signal, while smoothing it at the same time. For each signal $M^i(x)$, the filter $f(x)$ is convolved with $M^i(x)$. The product of that convolution is a the first derivative of a smoothed $M^i(x)$. The filter has the following form:

$$f(x) = Sxe^{-\alpha|x|} \tag{3.21}$$

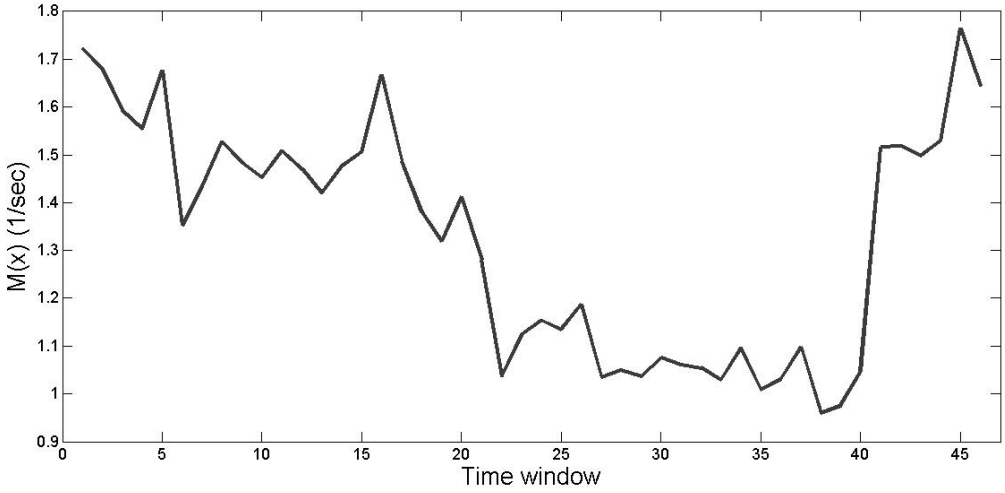
Note that $f(x)$ approximates the first derivative of a Gaussian function, which is, a smoothing function. Let $[-C, C]$ be the convolution window. The S variable is computed as a function of α as follows:

$$S = -\frac{1}{\sum_{k=-C}^{C} k^2 e^{-\alpha|k|}} \tag{3.22}$$

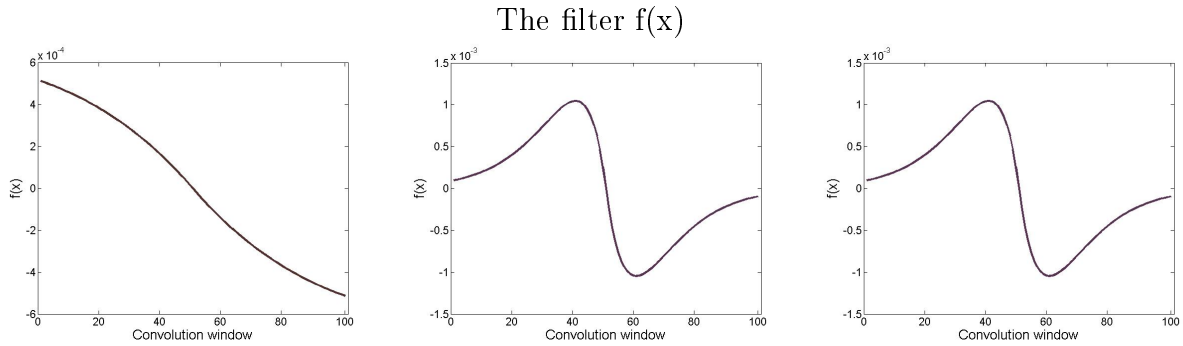
Thus, we detect events by convolving the signal $M^i(x)$ with the filter $f(x)$, and count the number of peaks in the convolved signal. Let E^i define the number of events detected in signal $M^i(x)$. E is computed in the following way.

$$E^i = \text{number of local maxima in: } (M^i * f)(x),\ x \in [0, W-1] \tag{3.23}$$

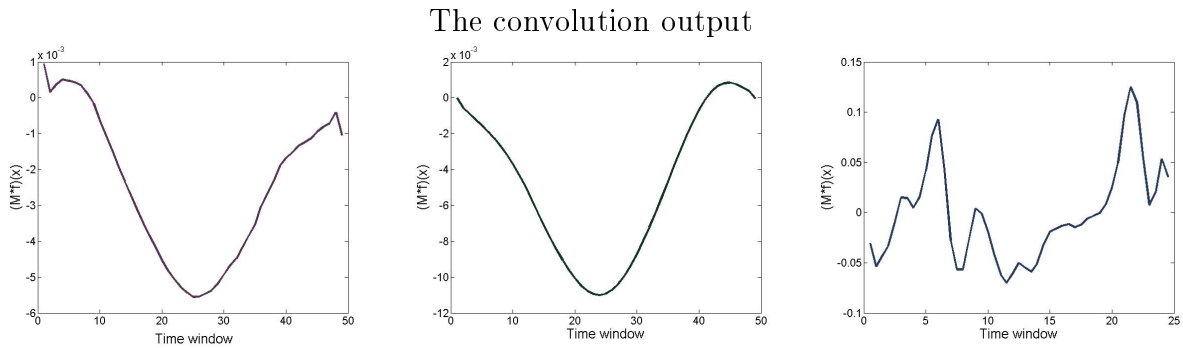
The only parameter of the filter, parameter α, reflects the localization ability and accuracy of the edge detector. Generally, when α has a small value, the real edges in the signal are accurately detected and separated from noise, but the resulting derivative has less pronounced peaks. Larger values of α should give a derivative with more and sharper peaks, that however may contain noise.



(a) Mean R–R signal. Time window length: T=1800 sec



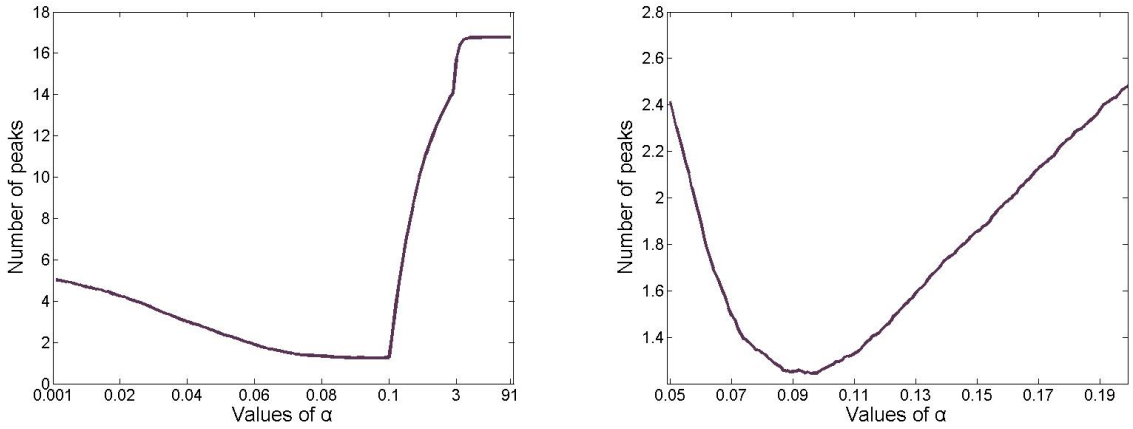
(b) $\alpha = 0.01$ (c) $\alpha = 0.1$ (d) $\alpha = 1$



(e) $\alpha = 0.01$ (f) $\alpha = 0.1$ (g) $\alpha = 1$

Figure 3.9: The result of convolving the signal of figure 3.9a with the filter $f(x)$, using different values for α

An example is presented in figure 3.9. Figure 3.9a shows the original signal, which corresponds to the 52^{nd} subject of the Normal Sinus Rhythm database (described in chapter 4.1). The signal is convolved with three different instances of the filter, corresponding



(a) Range: $\alpha \in [0.001, 91]$

(b) Range: $\alpha \in [0.05, 0.2]$

Figure 3.10: The mean of the number of detected peaks for all signals in our database with respect to the parameter α. The two figures show two sets of ranges for α; a large range on the left, and a smaller one on the right.

to three different values. We can observe that a very small value of α causes a disturbance in the extremities of the generated curve, which translates in a small increase in the number of detected peaks. That phenomenon can be attributed to the shape of the convolution filter for such small values (figure 3.9b). For an intermediate value of α, we can see in figure 3.9f that the generated output yields one peak in the end of the curve. That peak corresponds to the sharp increase, i.e. the event, observed in the end of the signal (figure 3.9a). This would be the desired output of the edge detecting subroutine of the algorithm. The last filter in figure 3.9d produces a larger number of peaks; that number is, nevertheless, sensitive to small increases as well.

While it is common to set the parameter α to be equal to some value and use it for all convolutions, we assign a value α^i for each signal $M^i(x)$. To compute the value α^i, we take advantage of the fact that very small values of α cause a few peaks to appear at the extremities of the output of the convolution, and use a wide range of possible values to finally select the one that generate the minimum amount of peaks. Thus the value of a^i is simply computed with the use of equation 3.23.

To obtain a visual evaluation of this selection method, we have plotted the mean E^i with respect to the values of α, for all the signals in our database (chapter 4, section 4.1). Figure 3.10a presents that curve for a wide range of values between $\alpha = 0.001$ and $\alpha = 91$. As expected, the minimum number of peaks is located in the range presented in figure 3.10b.

In consequence, each signal has its own filter $f^i(x)$, and equations 3.23 3.21 and 3.22 become:

$$E^i = \text{number of local maxima in: } (M^i * f^i)(x),\ x \in [0, W-1] \tag{3.24}$$

$$f(x) = S^i x e^{-\alpha^i |x|} \tag{3.25}$$

$$S^i = -\frac{1}{\sum_{k=-C}^{C} k^2 e^{-\alpha^i |k|}} \tag{3.26}$$

We can now formally define events and their corresponding values.

Definition 3.1. An *event* is defined as a point ϵ_k^i in the signal $M^i(x)$, $i \in [1, \ldots, I]$, $k = [1, \ldots, E^i]$, such that $(f^i * M^i)(\epsilon_k^i)$ is a local maximum. If $(f^i * M^i)(x)$ has no local maximum, then $\epsilon_k^i = 0$

Definition 3.2. The *value* v_k^i of an event ϵ_k^i is defined as $v_k^i = (f^i * M^i)(\epsilon_k^i)$

Remark 3.3.2. The value v_k^i of an event is directly related to the length of the corresponding edge in the signal. Hence, a larger value denotes a larger increase, whereas a smaller value denotes a smaller increase.

Remark 3.3.3. The value of E^i is used to differentiate patterns in a set of signals, to allow the EBA algorithm to be as flexible as possible in comparison to PPA. The events are used in two ways; the number of detected events serves as a classification feature, while the largest value v_k^i denotes the event that will serve as the starting point for the alignment of the signal. Note that the events are ordered in a decreasing way with respect to their value, hence ϵ_1^i corresponds to the maximum value v_1^i, $v_1^i = \mathbf{max}\{v_k^i\}$.

Remark 3.3.4. Even though we imply the events to correspond to transition between the awake state and the asleep state, it is not explicitly stated in the context of the present research. There is however no reason to believe that the event with the highest value is not related to act of waking up in the morning (with respect to each individual's sleep cycle).

The Algorithm

The EBA algorithm accepts a set of mean HR signals $\mathcal{M} = \{M^1(x), \ldots, M^I(x)\}$ as its input and returns four sets $\mathcal{M}^\ell, \ell \in \{1, 2, 3, 4\}$, with $\bigcap_{\ell=1}^4 \mathcal{M}^\ell = \emptyset$ and $\bigcup_{\ell=1}^4 \mathcal{M}^\ell = \mathbf{M}$. We denote by $\mathbf{M}$ the set $\mathbf{M} = \{\mathbf{M}^1(x), \ldots, \mathbf{M}^I(x)\}$, where $\mathbf{M}^i(x) = M^i(x + \epsilon_1^i)$ the aligned mean HR signals. The three first set correspond to the three more frequent values of E^i, whereas the fourth set contains the remaining signals. Recall that E^i denotes the number of detected events for each signal.

The EBA algorithm is presented in algorithm 2, and its subroutines in 3. It proceed in two steps.

Algorithm: EBA

Input: A set $\mathcal{M}$ of mean HR signals $\mathcal{M} = M^1(x), \ldots, M^I(x)$ of equal size W.
Output: Four sets $\mathcal{M}^\ell$ of aligned mean HR signals $\mathbf{M}^i(x)$, $\ell \in \{1, 2, 3, 4\}$

Step 1: The Classification Process
$E := \{\emptyset\}$
for each $M^i(x)$ **in** $\mathcal{M}$ **do**
  $[E^i, \epsilon_1^i] = Classify(M^i(x))$
  $E := E + \{E^i\}$
$L^\ell = RecoverLabels(\mathrm{E})$
Step 2: The Circadian Alignment Process
for $\ell = 1, \ldots, 4$ **do**
  $\mathcal{M}^\ell := \{\emptyset\}$
for each $M^i(x)$ **in** $\mathcal{M}$ **do**
  $\mathbf{M}(x) := \emptyset$
  for $\ell = 1, \ldots, 3$ **do**
    if $E^i = L^\ell$ **then**
      $\mathbf{M}(x) = M^i(x + \epsilon_1^i)$
      $\mathcal{M}^\ell := \{\mathbf{M}(x)\}$
  if $\mathbf{M}(x) = \emptyset$ **then**
    $\mathbf{M}(x) = M^i(x + \epsilon_1^i)$
    $\mathcal{M}^4 := \{\mathbf{M}(x)\}$
return $\mathcal{M}^1, \ldots, \mathcal{M}^4$.

Algorithm 2: The *Event Based Alignment* Algorithm

- **Step 1.** The Classification process
- **Step 2.** The Circadian Alignment process

During the Classification process, each mean HR signal $M^i(x)$ is assigned an α^i value, and according to that, the values of E^i, ϵ_1^i, and v_1^i, are resolved for each signal. Finally, a label L^ℓ, $\ell \in \{1, 2, 3, 4\}$ is associated to each signal, where the three first labels correspond to the three most frequent E^i values in descending order, and the last label to the al the rest. Hence, if we denote by E^ℓ, $\ell \in \{1, 2, 3\}$ those three values, the algorithm classifies mean HR signals as follows. For each signal, if $E^i = E^\ell, \ell \in \{1, 2, 3\}$, then $L^\ell = E^\ell$ and $\mathcal{L}^\ell := \mathcal{L}^\ell + M^i$. If $E^i \neq E^\ell \ \forall \ \ell \in \{1, 2, 3\}$, then $L^\ell = -1$ $\mathcal{L}^4 := \mathcal{L}^4 + M^i$

The Circadian Alignment Process aligns each signal according to the event ϵ_k^i that yields the highest value v_k^i. As previously established $v_1^i = \mathbf{max}\{v_k^i\}$. Therefore, for each $M^i(x)$, $\mathbf{M}^i(x) = M(x + \epsilon_1^i)$.

Subroutine: Classify

Input: A signal $M(x)$

Output: The number E^i of detected events, the maximum value event e_1^i

for each $M^i(x) \in \mathcal{M}$ **do**
  $\alpha^i = 0.01$
  $a = \inf$
  for $\alpha = 0.05, 0.051, 0.052, \ldots, 0.2$ **do**
    lmax(α)=**number of local maxima** $\{(f * M)(x)\}$
    if lmax(α)<a **then**
      a=lmax(α)
      $\alpha^i = \alpha$
for $i = 1, \ldots, I$ **do**
  $\alpha = \alpha^i$
  $E^i =$ **number of local maxima** $\{(f * M)(x)\}$
  $e_1^i = x^{max}$, where $v_1^i =$ **max**$\{(f * M)(x)\}$ and $(f * M)(x^{max}) = v_k^i$
return E^i, e_1^i.

Subroutine: RecoverLabels

Input: A set $E = \{E^1, \ldots, E^I\}$ of number of detected events.

Output: Four labels L^1, L^2, L^3, L^4

Count the frequency of appearance of each E^i
$L^{1,2,3}$ = The three most frequent E^i values in descending order
L^4 = -1
return L^1, L^2, L^3, L^4.

Algorithm 3: Subroutines for the *Event Based Alignment* Algorithm

Chapter 4

Experimental Results

4.1 Database Description

To run our experiments, we use three different databases that contain annotated R–R interval signals from long term ECG recordings. All signals are annotated to indicate whether a record is a normal beat, i.e. consistent with *Normal Sinus Rhythm*. All three databases are included in the Physionet Data Bank archives [1] [1]. The first one includes healthy, asymptotic patients, the second includes subjects that suffer from Congestive Heart Failure (CHF), and the last one contains subjects from the Cardiac Arrhythmia Suppression Trial (CAST) , a trial that included post MI patients. (see section 2.1), or not. The three databases are:

- Normal Sinus Rhythm RR Interval Database (NSR) R–R Interval Database.
- The Congestive Heart Failure R–R Interval Database (CHF) R–R Interval Database.
- The CAST R–R Interval Sub-Study Database [41].

[1] `http://physionet.org`

	#rec	Avg	min	Max	Age	Time
NSR	54	23,69	21,37	24,22	28,5 to 76	No
CHF	29	22,18	17,02	24,02	34 to 79	No
CAST RR Ea	262	23,51	10,39	26,48	-	Yes
CAST RR Eb	262	23,43	8,74	26,48	-	Yes
CAST RR Fa	207	23,36	18,20	26,24	-	Yes
CAST RR Fb	207	23,57	19,12	26,49	-	Yes
CAST RR Ma	266	23,43	0,99	25,96	-	Yes
CAST RR Mb	266	23,18	10,14	25,96	-	Yes

Table 4.1: Statistics for each database.
#rec = *Number of recordings,* **Avg** = *Average length (h),* **min** = *Minimum length (h),* **Max** = *Maximum length (h),* **Age** = *Range of age for subject,* **Time** = *Available horological time for the beginning of recordings.*

4.1.1 The Normal Sinus Rhythm R–R Interval Database

The NSR database includes annotated R–R intervals from long term ECG recordings that were obtained from 54 subjects. The recordings were digitized at 128 samples per second and an automated analysis along with manual review and correction was performed to obtain the beat annotations. The starting times in this database are not specified. Only the normal beats are taken into consideration for each subject.

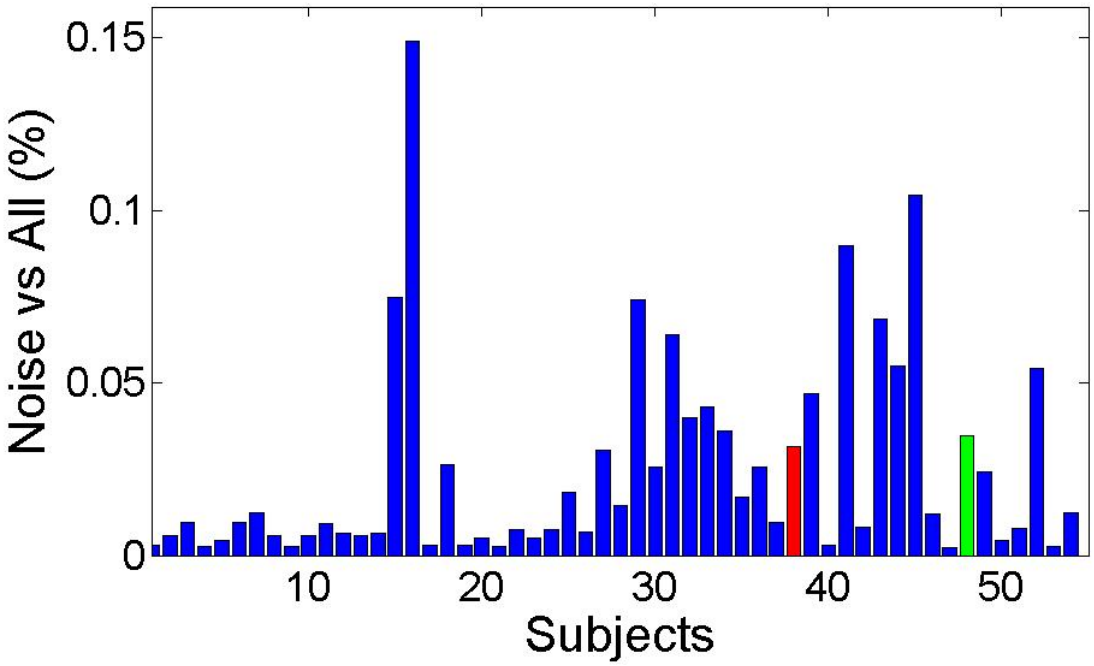
As we can see in table 4.1, the minimum length and maximum length of a record is $21,37$ hours and $24,22$ hours respectively, that is, this dataset has sufficiently long recordings to establish circadian heart rhythm. Figure 4.1a displays a bar graph showing the percentage of noisy R–R intervals versus all the R–R intervals in the record, for each record. The largest amount of noise in a recording corresponds to $14,9\%$. The average noise in the recordings is $2,48\%$.

The following researches have also made use of the NSR database [42, 43, 44, 45].

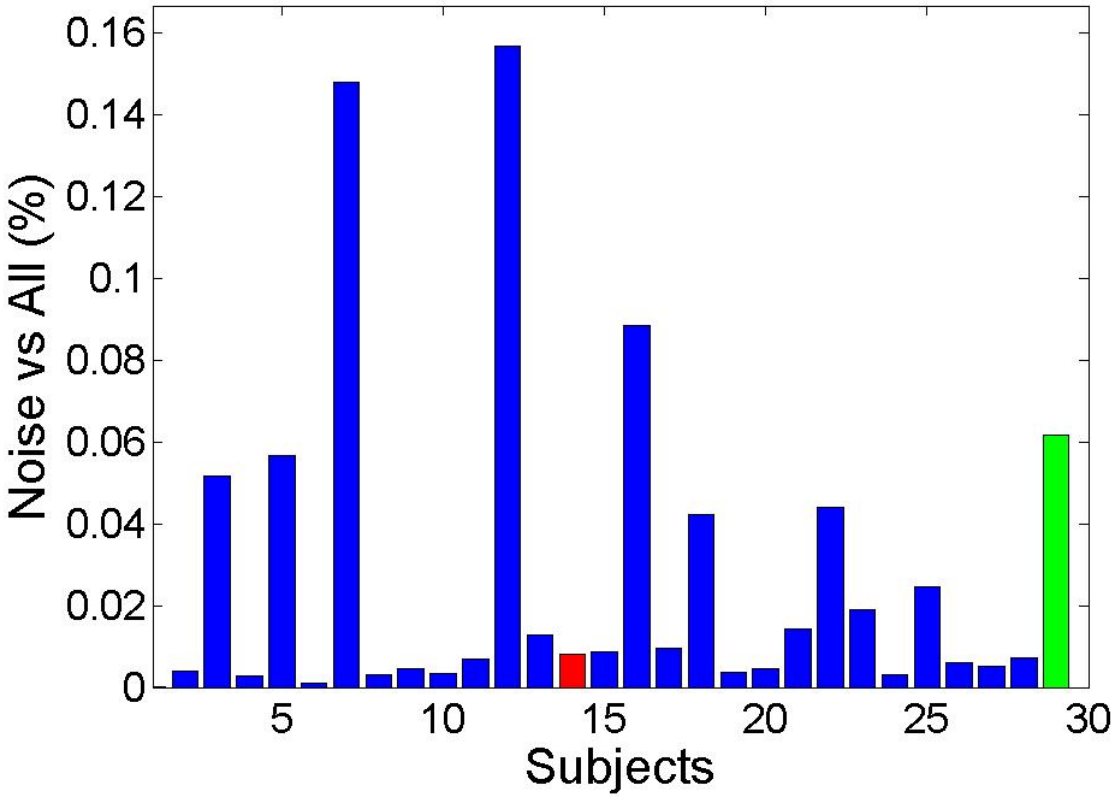
4.1.2 The Congestive Heart Failure R–R Interval Database

The CHF database contains 29 R–R interval signals obtained from long-term ECG recordings. The subjects in this database all suffer from CHF (New York Heart Association classes I, II, and III). Note that, as mentioned in chapter 2, section 2.3, HRV markers, namely the SDNN, can be used as a tool in the prognosis of CHF. The recordings were digitized at 128 samples per second and an automated analysis along with manual review and correction was performed to obtain the beat annotations.

The largest recording has a length of 24,02 hours whereas the smallest has a length of 17,02 hours. This is an important difference, and we will see later how the algorithm performs when there are recordings in the dataset that have a smaller length than the others. The average noise in the signals is $2,8\%$, whereas the largest amount of noise is equal to $15,65\%$. Figure 4.1b shows the bar graph of noise per recording.



(a) NSR database.



(b) CHF database.

Figure 4.1: Bar graphs showing the percentage of noise in the signal for each subject in the NSR and CHF databases. The green and red coloured bars respectively indicate the largest and smallest recording in length.

Researches that have also made use of the CHF database include [46, 47, 45].

4.1.3 The CAST R–R Interval Sub-Study Database

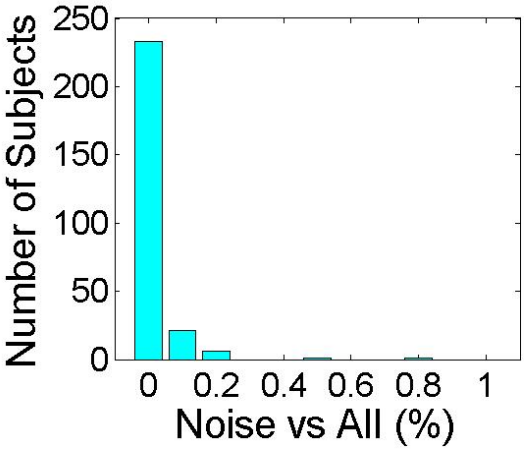
The CAST R–R database was obtained from a study designed to test whether the death rate of MI survivors would improve upon suppression of asymptomatic or mildly symptomatic ventricular premature complexes (PVCs). These patients were grouped into four categories and given three different drugs to induce PVC suppression, Encainide, Flecainide, and Moricizine, and a placebo drug. The CAST RR Interval Sub-Study Database consists of RR interval time series from the pre-treatment and on-therapy recordings from the patients in the three first groups. Hence, this databased is separated into six subcategories.

- Subjects who received Encainide prior to receiving the drug (Ea Database)
- Subjects who received Encainide after receiving the drug (Eb Database)
- Subjects who received Flecainide prior to receiving the drug (Fa Database)
- Subjects who received Flecainide after receiving the drug (Fb Database)
- Subjects who received Moricizine prior to receiving the drug (Ma Database)
- Subjects who received Moricizine after receiving the drug (Mb Database)

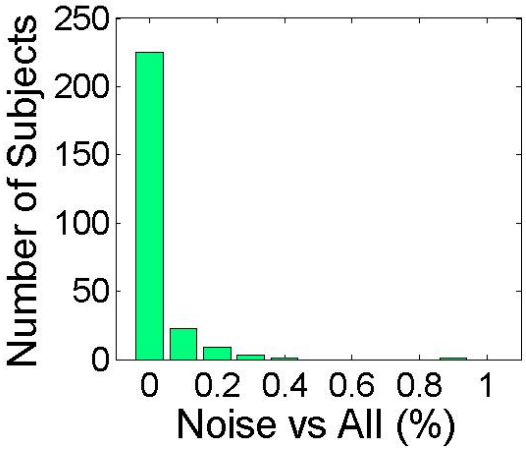
The databased has been scanned by an experienced arrhythmia analyst, who scanned all the recordings using standard Holter analysis procedures. A second edit of the annotation files was also made to identify and annotate all the improperly measured RR intervals. The original ECG recordings are not available. For this research, only the beats annotated as normal beats are taken into consideration. Furthermore, we kept only the subjects whose recordings are available both before and after drug administration.

As we can see in table 4.1, all CAST RR datasets are large, about 5 and 10 times larger than the CHF ans NSR datasets respectively, and present more irregularities. This is precisely the reason why we have also included these signals in our research, that is, in an attempt to test the effectiveness of the algorithm when for example among all signals, there is one that only lasts an hour. This is the case for the Ma database (see table 4.1). Of course, the average length of all signals is about 23,50 hours. For this dataset we also have the time information available, i.e. the starting time for each recording, hence, we will be able to compare and assess the accuracy of our algorithms.

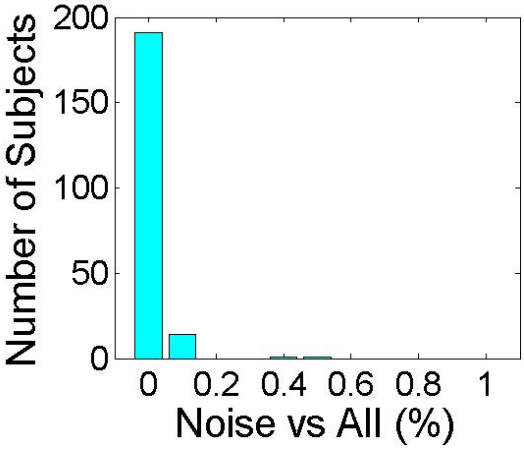
Figure 4.2 demonstrates one histogram per dataset, where for each bin we have counted the number of subjects that have the equivalent average noise in the signal. Observe that the only cases where there are really noisy signals in the dataset are the Ea and Eb datasets, where we can see in figures 4.2a and 4.2b that in very few cases, the noise in the signal exceeds 50%.



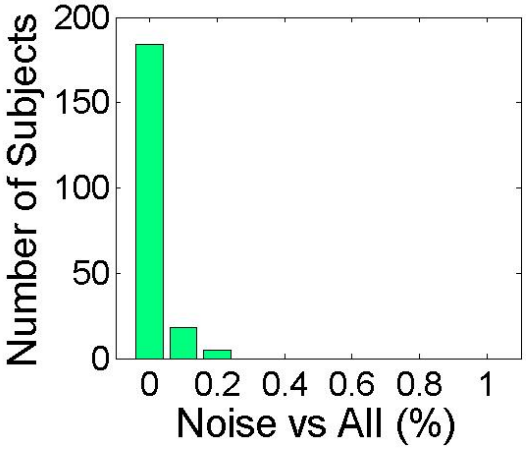
(a) Ea database.



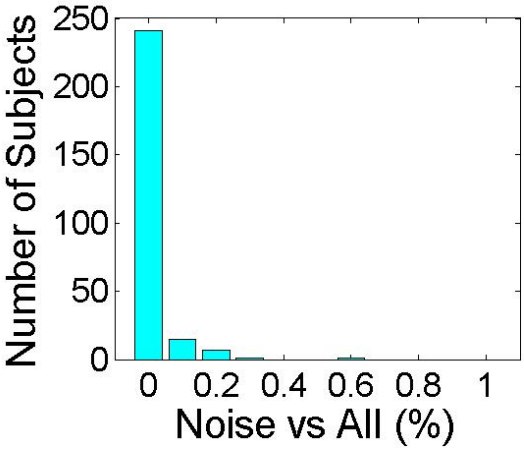
(b) Eb database.



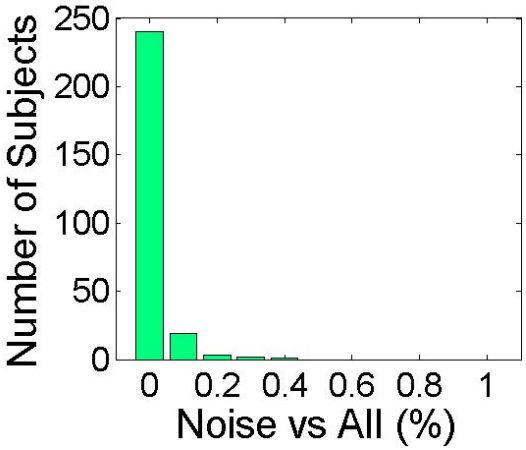
(c) Fa database.



(d) Fb database.



(e) Ma database.



(f) Mb database.

Figure 4.2: Histograms showing the percentage of noise in the signal for each subject in the CAST R–R database. The green and red coloured bars respectively indicate the largest and smallest recording in length.

4.2 Framework

All experiments were performed on an Intel Core i5 processor. Every run of the algorithm lasted at most 5min (this time was mainly observed when using the CAST RR datasets, which include a larger number of subjects).

For every run of an algorithm, the input of is the mean HR signal $H^i(x)$, computed according to sections 3.1 and 3.2.2 for each subject in the database. The only parameter that is use defined in both our algorithms, is the length of the time window T. Thus, for each dataset and for each algorithm, we used four values for T:

- T=10min
- T=20min
- T=30min
- T=60min

Recall that all other parameters are automatically defined. The total number of experiments is: **Number of datasets** $\times$ **Number of values for T** $= 8 \times 4 = 32$

After computing the starting points for each signal, we compute the average $H(x)$ signal according to equation 3.14. For the EBA algorithm, we compute the average per category. For each experiment, we present the following results:

- A random signal alignment, where the starting points for each signal in the datasets are randomly selected.
- The alignment of the signals as they are. (no alignment criterion)
- The alignment using the horological time, when that information is available.
- The aligned signal using the PPA algorithm
- The aligned signal for all categories using the EBA algorithm

In the next section we report the above first set of results and discuss the effectiveness of our algorithm with respect to the alignment they propose. We show a selected set of experiments; the rest of the experiments are presented in the appendix (REFERENCE).

In section 4.4 we discuss the categorization of the signals as proposed by the EBA algorithm.

In section 4.5 we present the effect of the alignments proposed by PPA and EBA for the NSR database on the time domain measurements presented in chapter 3, section 3.2.

4.3 The Alignment Process

The following cases demonstrate the performance of the PPA and EBA methods with respect to other alignment criterion. We first present two cases where the PPA and EBA methods were applied on a random alignment of the signals, created by a shuffle of the starting points of the signals, to demonstrate both the effectiveness of the algorithms and the importance of the alignment process in an signal averaging process.

Next, we evaluate the results of PPA and EBA using the available information in the database. The NSR and CHF databases provide no horological information as to the time of each recording, nevertheless, we make the assumption that they all start at the same time. The CAST R–R database provides horological information for all signals, which we use to manually align the signals. Our results show that we can achieve signal alignment that is as good as the horological alignment. In some cases, our alignment algorithm achieve even better results with respect to the circadian heart rhythm.

The observations we make with respect to the decrease and increase of the average heart are based on the circadian rhythmicity of the heart rate, and the phenomena we expect to see are those described in section 2.4. Thus, we assume that the *average patient*, i.e. the averaged mean HR signal $H(x)$ shows an important decrease in the sleep state/awake state transition, and an important increase in the awake state/sleep state transition.

Although such increases and increases can indeed occur due to different phenomena, we assume that they are individual and rare cases, and that the average signal of a database remains unaffected. Nevertheless, if such phenomena should occur on a larger scale, our results show that the EBA algorithm is be able to detect them and classify the signals accordingly.

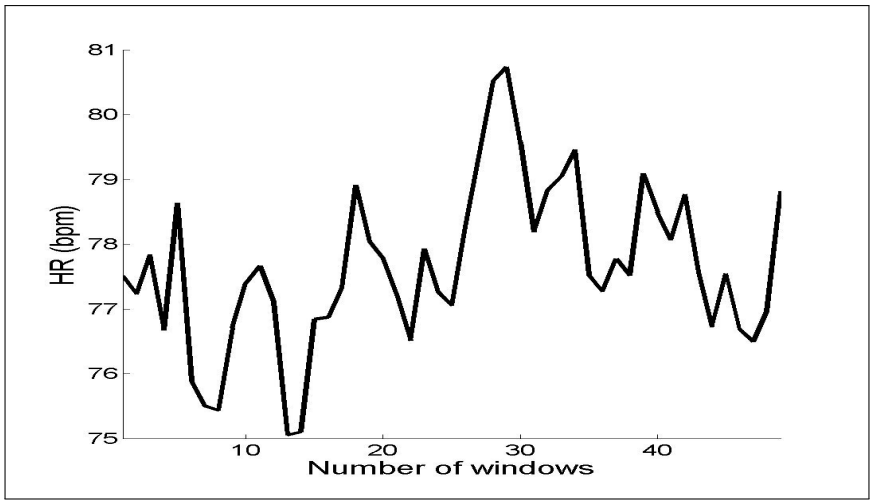
For the remainder of our comments, we also refer to the last part of the signals, where the mean HR value decreases to increase again towards the end of the signal, as *night time*. Specifically, for the results of the EBA algorithm, we refer to the beginning of the signal as *wakeup*.

4.3.1 Comparing The Results of our Algorithms to a Random Alignment

In this subsection we present two sets of results that demonstrate the alignment ability of both our algorithms. For each dataset and for each signal in the dataset, we shuffle the starting points so that the new starting point of each signal is different. Thus, for each $H^i(x)$, the new random signal is computed as follows:

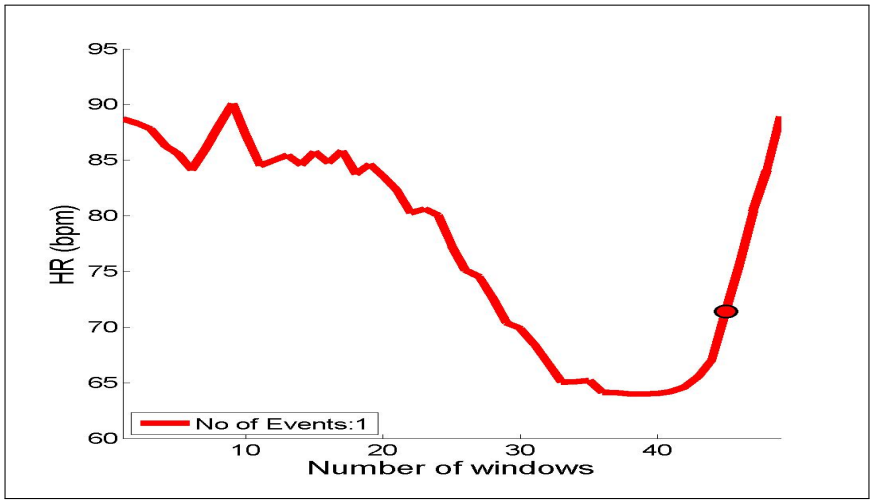
$$H^i_R(x) = H^i(x + r^i) \text{ where } r^i \text{ is randomly selected, } r^i \in [0, W-1]$$

The mean HR is computed from the resulting signals. Naturally, there is no time correlation between signals with respect to any criterion in the mean HR shuffled signal. This random shuffle is meant to simulate the case where a research is to be conducted with



(a) Random Alignment

(b) PPA result



(c) EBA result

Figure 4.3: **Parameters:** Database = NSR, $T = 30$ minutes. The first image represents an alignment of the subjects according to a random shuffle of the starting points. The second and third figures show the alignments proposed by the PPA and EBA algorithm respectively, when their input is the set of unsynchronized signals produced by that shuffling process.

24 hour HR of R–R interval signals of which, either the recording hours are unknown, or they are known and they vary, possibly making the manual alignment time consuming. Note that all our selected database fall into one of those categories. The recording times for the NSR and CHF database are unknown, whereas the manual alignment of the CAST R–R database with respect to horological time was a lengthy process.

The first set of results we present comes from the NSR database, for a time window of $T = 30$ minutes.

The random shuffle of the signal is presented in figure 4.3a. Observe that these is absolutely no circadian pattern that can be resolved from this signal. The increases and decreases of the heart rate are completely random.

We have given the shuffled $H_R(x)$ signals seen in figure 4.3a as an input to the PPA algorithm. The result can be seen in figure 4.3b. The difference between the two signals is striking; The mean HR signal resulting from the PPA follows a clear circadian rhythm, where the night time lasts 20 time windows, i.e. 10 hours since each time window equals 30 minutes.

Upon giving the same input to the EBA algorithm, we obtain the result depicted in figure 4.3c. Again, EBA manages to align the signal with respect to the circadian rhythmicity of the heart rate. Night time, same as with PPA, lasts approximately 20 time windows. One of the great advantages of the EBA algorithm is that it automatically aligns the HR signals with respect to wakeup time, which we consider to be the beginning of the day.

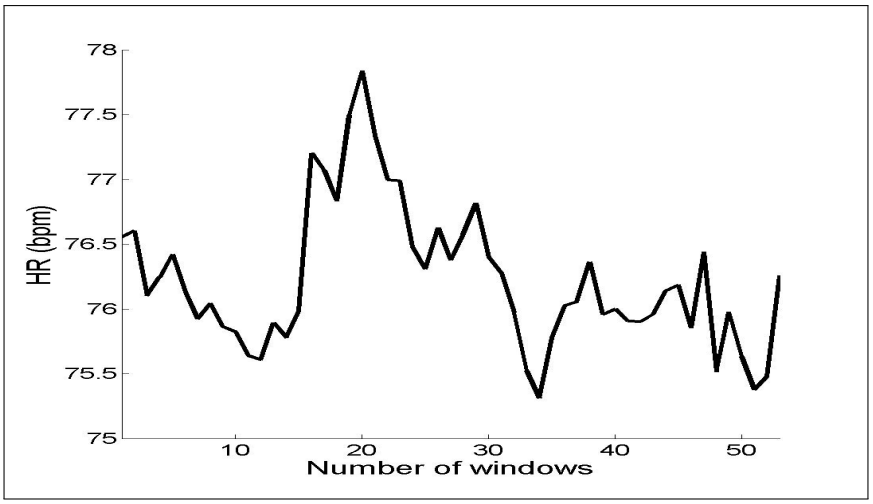
The second set of results corresponds to the Ea algorithm for $T = 30$ minutes.

We repeated the same process as before; a random shuffle of the mean HR signals $H_R^i(x)$ is performed to obtained an average mean HR signal. The result is seen in figure 4.4a. Note again that the individual subjects signals are now completely uncorrelated. The resulting average signal shows no circadian pattern.

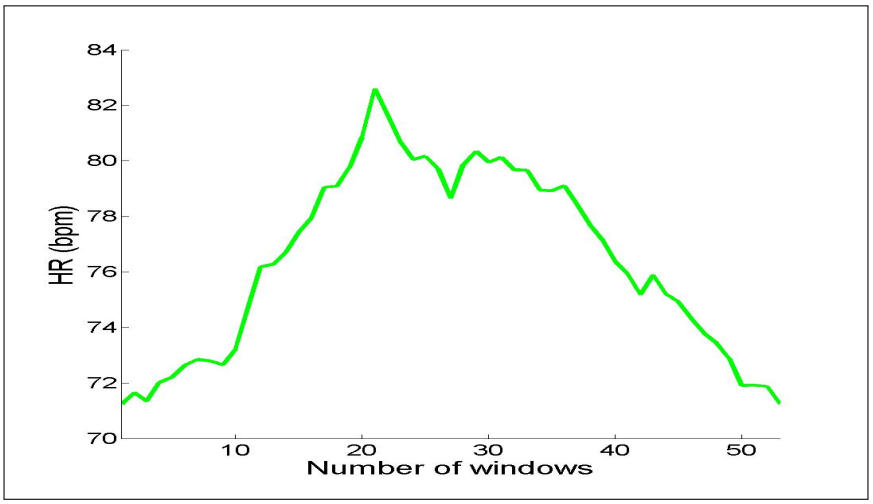
We provide these shuffled mean HR signals $H_R^i(x)$ as an input to both the PPA and EBA algorithms. The results are shown in figures 4.4b and 4.4c.

The PPA algorithm seems to have a harder time to align the signals. There is an emerging circadian pattern present, however the increase and decrease of the mean HR are not as steep as we expect them to be. Again, the PPA algorithm provides us with no automatic way of determining a starting point that would correspond to the beginning of the day.

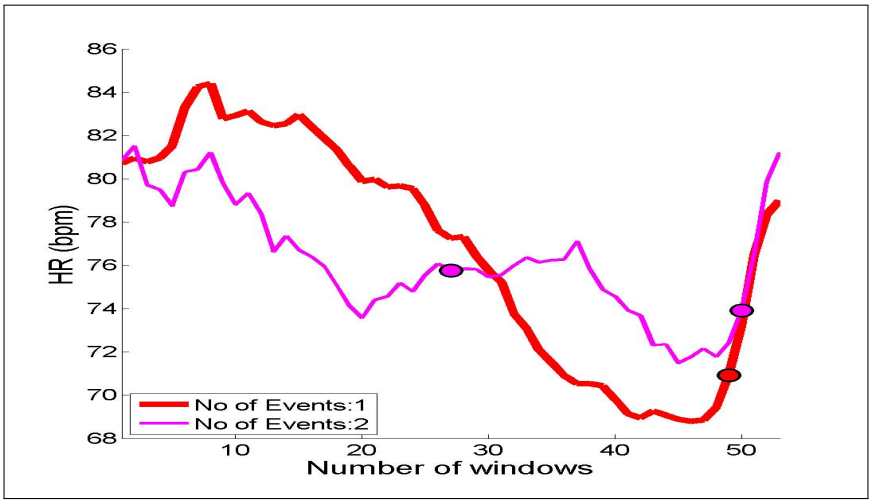
The results of the EBA algorithm are much more realistic and in accordance with circadian rhythmicity. EBA has detected two categories of subjects with respect to the detected events. For the first category, one event was detected, whereas for the second category two events occur for each mean HR signal. The circle markers denote the approximate time of detection. Upon observing the average signal that corresponds the first category (one event), the circadian pattern is clear. Night time lasts around 20 time windows, that is, 10 hours. Furthermore, night time begins about 27 time windows, i.e. 13,5 hours after wake up time.



(a) Random Alignment



(b) PPA Alignment



(c) EBA Alignment

Figure 4.4: **Parameters:** Database = Ea, $T = 30$ minutes. The first image represents an alignment of the subjects according to a random shuffle of the starting points. The second and third figures show the alignments proposed by the PPA and EBA algorithm respectively, when their input is the set of unsynchronized signals produced by that shuffling process.

For the second category (two events), it seems that there are two periods that could be characterised as night time, the second (that ends when the signal ends) being more pronounced. We argue that it is this set of subjects that causes the irregular behaviour of the PPA algorithm. This category clearly demonstrates a different circadian pattern than the first one. Nevertheless, EBA aligns the signals regardless of the presence of irregular circadian patterns among the original (and shuffled) mean HR signals.

4.3.2 Comparing The Results of our Algorithms to the Horological Alignment

The first set of results we use come from the CHF database, and the time window for the plots shown in figure 4.5 is equal to $T = 10$ min.

Figure 4.5a shows the result of aligning the signals using the information available in the database. For the CHF database, it is assumed in all researches that the recordings start at the same time, although the exact horological time is unknown. We can make the following two observations.

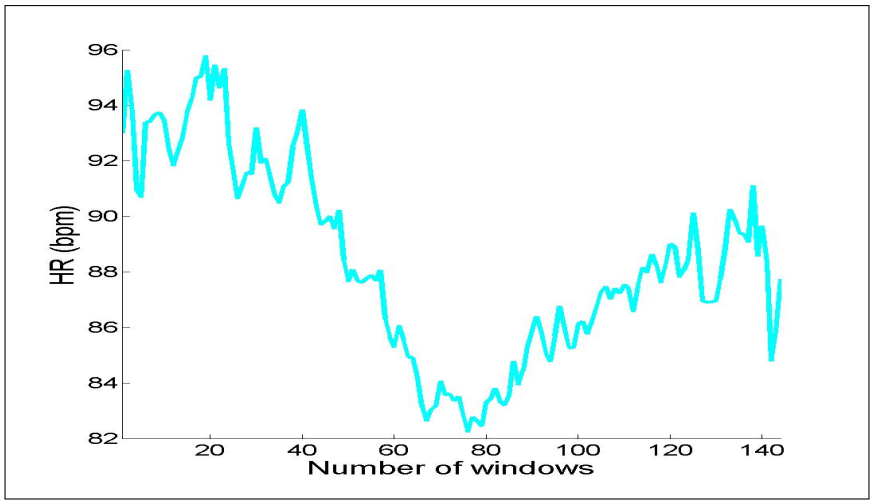
As stated in section 2.4, we expect to see a circadian rhythmicity, outlined by a sharp increase in the averaged signal that would demonstrate the transition between sleep state and wake state. Such an edge is indeed seen between the 80^{th} and the 138^{th} window. Since each time window is equal to 10 minutes, the 80^{th} time window is approximately 13^{th} hours after the beginning of the day and the 128^{th} is the 23^{rd}.

Hence, this implies that the average patient needs ten hours to wake up. Furthermore, if we consider that the signals starts dropping from the 40^{th} time window, that is, the 7^{th} hour from the beginning of the day, then for the average patient, the transition in and out of the the sleep state has lasted 16 hours, which does not seem like a realistic result.

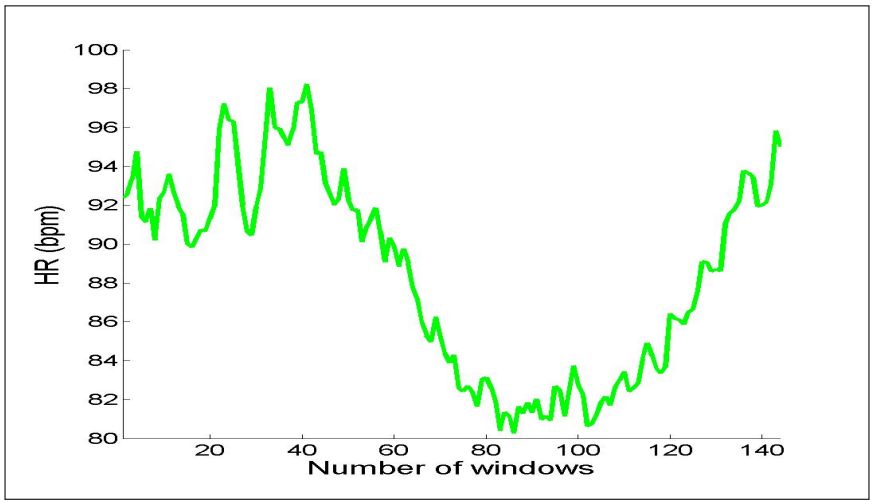
The second observation that can be made is that the end of the signal shows a great variance, and does not correspond to the beginning of the signal. In a circadian cycle we generally expect the last window to not be far from the first one. What is happening in our case is that the last time windows are sparse, as most of the recordings stop after the 22^{nd} hour (see table 4.1). In particular, the last 60 minutes (6 time windows) of the signal include less than 10 signals per time window.

In figure 4.5b we can observe the result of the PPA algorithm. With respect to the previous observations, we can see that the second issue is resolved without loss of information. The PPA algorithm aligns the signals using all the information available, and it seems that the signals are now distributed in such a way so that the end of the signal does not contain any less recordings. The beginning and ending of the sleep cycle is more apparent. Before the sleep cycle, the heart beats at a rhythm of 94 bpm, and in the end of the sleep cycle we again reach the same level. During the sleep state, the mean HR is ≈ 82 bpm. Nevertheless, the PPA algorithm still gives a large window for the sleep phase; we can see that it still begins at the 40^{th} time window and ends at the 140^{th} time window.

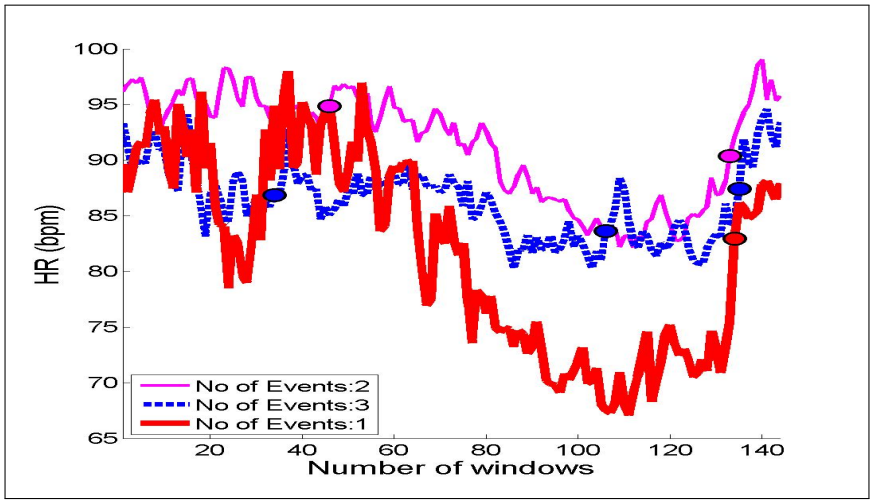
Upon applying the EBA algorithm, we obtain four categories of signals based on event



(a) No Alignment



(b) PPA result



(c) EBA result

Figure 4.5: **Parameters:** Database = CHF, $T = 10$ minutes. The first image represents an alignment of the subjects according to how each signal starts in the database. The second and third figures show the alignments proposed by the PPA and EBA algorithm respectively.

Number of events	2	1	3	-
Number of subjects	17	7	4	1

Table 4.2: Number of events and number of subjects per category for the CHF database, where $T = 10$ min

count. In figure 4.5c we present the three first categories and the corresponding number of events detected. The fourth category has been omitted since it contains one subject whose signal reported 5 events. Table 4.2 reports the number of subject per category.

We can see that, for all categories, the transition in and out of the sleep state is distinguishable for the categories where 1 or 2 events have been detected, and it last less time. When 2 events are detected, the descent seems to start at the 80^{th} time window, and ends when the signal ends, which corresponds to 10 hours. That transition lasts longer when only 1 event is detected; indeed, it corresponds to the result we observed on the PPA algorithm, i.e. a transition that lasts 13 hours. Finally, when three events are detected, we can see that the mean HR does not vary much between the awake state and the sleep state. The usual circadian pattern is not apparent either.

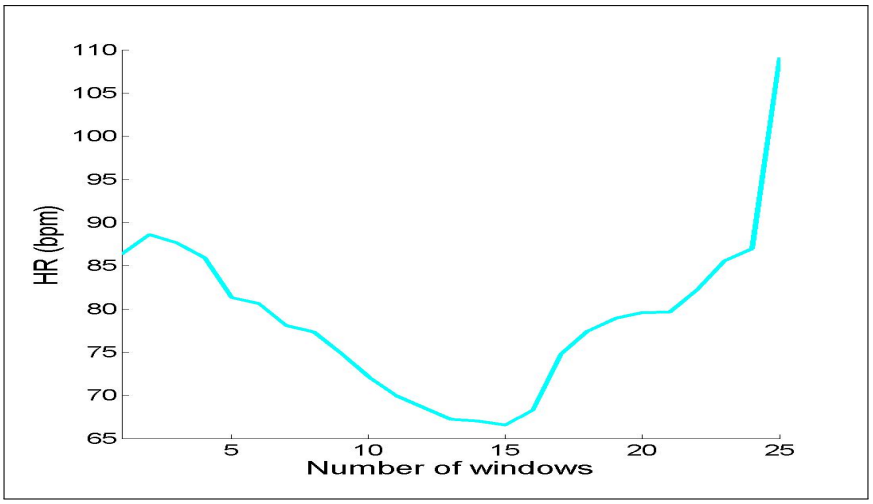
The next set of results presented comes from the NSR database. The time window T is equal to 1 hour.

The NSR database, same as the CHF database, does not include any information as to what time the recording begins. Nevertheless, we again assume that all the signals in the database start at the same time. In figure 4.6a we can see the result of averaging the mean HR signals as they are found in the database. Observe that there is no clear circadian HR rhythm, as we would expect to see. There is no clear differentiation between the asleep and awake state. Moreover, the last time window contains information coming from only one subject, which causes the last time window to have a much higher value than the rest.

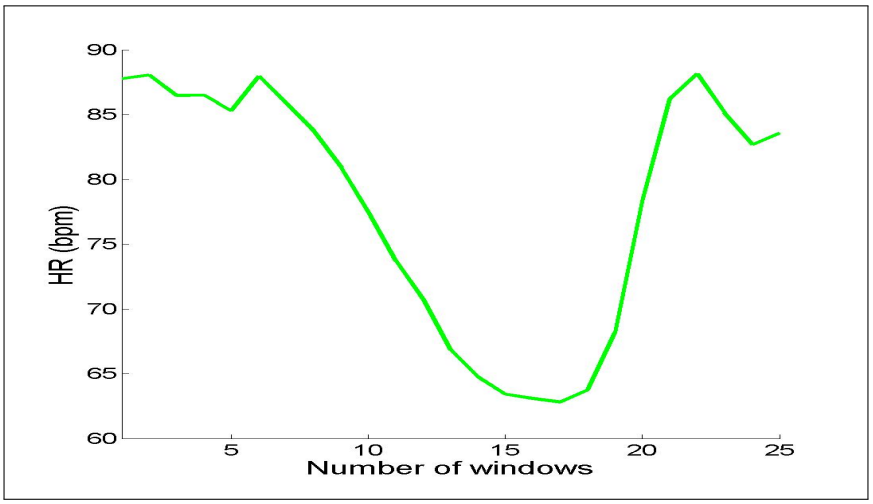
The result of the PPA algorithm can be seen in figure 4.6b. The expected circadian pattern is observed; there is a clear decrease and increase of the mean HR from the 8^{th} time window to the 21^{st} time window. Furthermore, we can conclude that the algorithm has efficiently used the information from the last time window of figure 4.6a, since there is no apparent irregularity in the averaged signal.

A similar result can be observed in figure 4.6c. The EBA algorithm has detected one category, where the decrease starts around the 12^{th} time window and the increase at the end of the signal. We make the remark here that the EBA algorithm not only synchronizes the signals with respect to the circadian HR rhythm, but it also holds a reference point, i.e. the most significant event. Thus, we can set the starting point automatically. This is not possible with the PPA algorithm, where the starting point is random and the beginning of the day must be manually annotated.

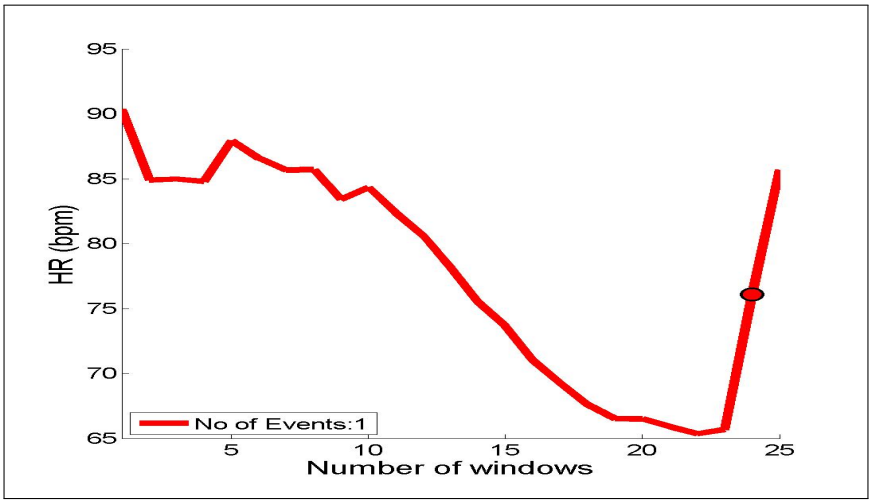
The third set of results we present come from the Fb database. The time window here is set to $T = 30$ minutes.



(a) No Alignment



(b) PPA result



(c) EBA result

Figure 4.6: **Parameters:** Database = NSR, $T = 60$ minutes. The first image represents an alignment of the subjects according to how each signal starts in the database. The second and third figures show the alignments proposed by the PPA and EBA algorithm respectively.

Number of events	2	1	3
Number of subjects	136	69	2

Table 4.3: Number of events and number of subjects per category for the Fb database, where $T = 30$ min

Number of events	1	2	-
Number of subjects	194	63	8

Table 4.4: Number of events and number of subjects per category for the Ma database, where $T = 20$ min

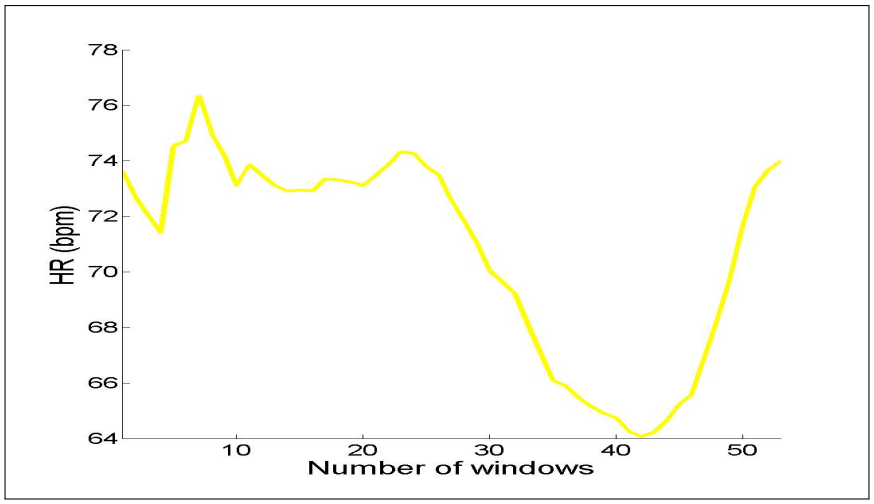
The horological time of the recordings is available for all subsets of the the CAST R–R database. We have used that information and aligned the signals according to the horological time. The result of the alignment for the Fb database can be seen in figure 4.7a. Here, the mean HR signal follows the human HR rhythmicity. The mean HR is around 74 bmp during the awake state, which lasts 25 windows, i.e. 12,5 hours, and during the sleep state we see it drop to 64 bpm.

The result of the PPA algorithm is displayed in figure 4.7b. We can observe that PPA does not capture the circadian rhythmicity as we see it when aligning the signals using the horological time. However, the same observations can be made as to the values of the mean HR during sleep state and awake state, which are bpm and 74 bpm respectively. What differs is the duration of the transition in and out of the sleep state; it is shown to last 35 time windows, that is, 17m,5 hours. However, the increase between sleep and awake state is not as steep.

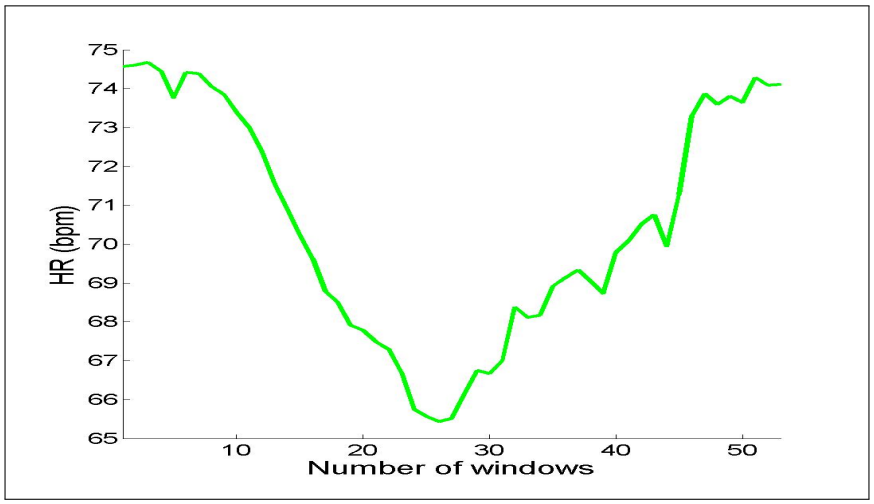
The alignment proposed by the EBA algorithm is much closer to the horological time based alignment. EBA discovers three categories according two the number of events, which can be seen on table 4.3. The category that contains the most subjects is the one where 2 events were detected. Observe in figure 4.7c that the second event is located approximately between the first and the beginning of the day. There is a small decrease at the 10^{th} and a small increase at the 20^{th} time window. At that time, the mean HR is about 72 bpm. The mean HR drops to about 66 bpm after the 35^{th} time window. We can see that the circadian pattern here is different than the one we usually observe. The second category contains 69 subjects with 1 detected event. The pattern is very similar to the horological alignment. The third category is not displayed, since it only contains two subjects.

The fourth set of results belongs to the Ma database, where the time window is set to $T = 10$ minutes. In figure 4.8 we can observe the horological alignment (figure 4.8a), the alignment proposed by the PPA algorithm (figure 4.8b), and the alignment proposed by the EBA algorithm (figure 4.8c).

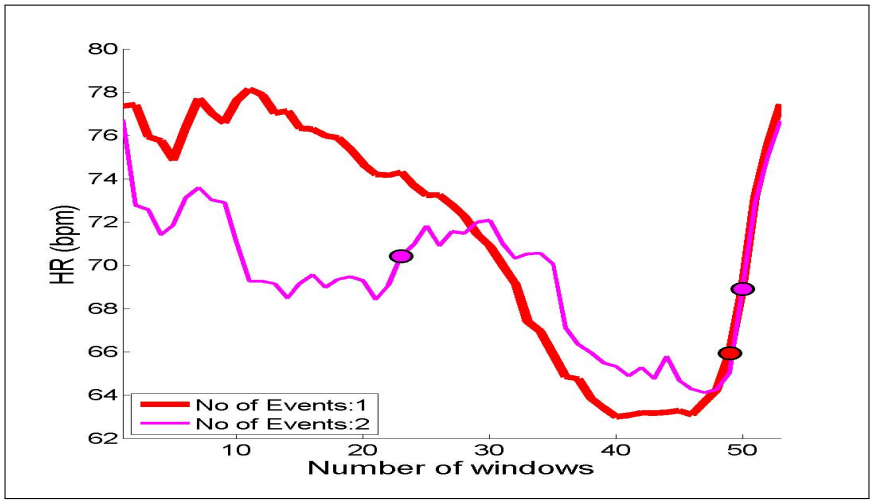
We have selected to present this last case so as to demonstrate our results when the mean HR signal that results from horological alignment does not follow an expected



(a) Horological Alignment

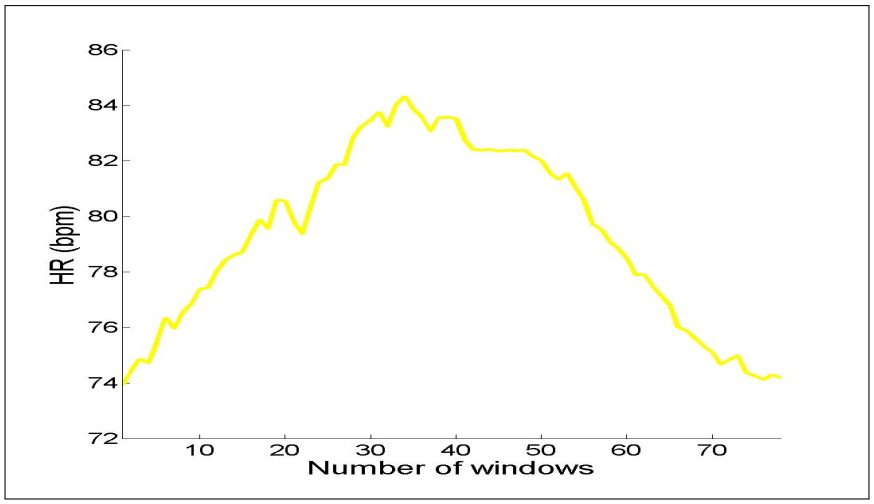


(b) PPA Alignment

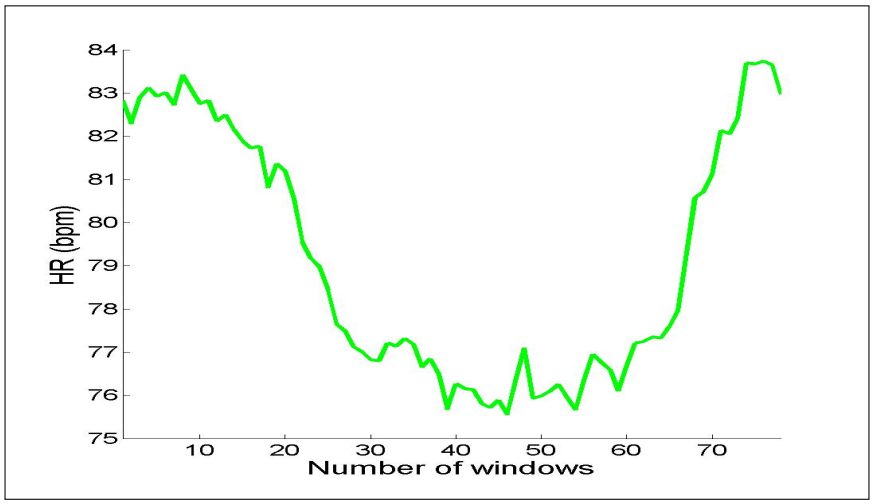


(c) EBA Alignment

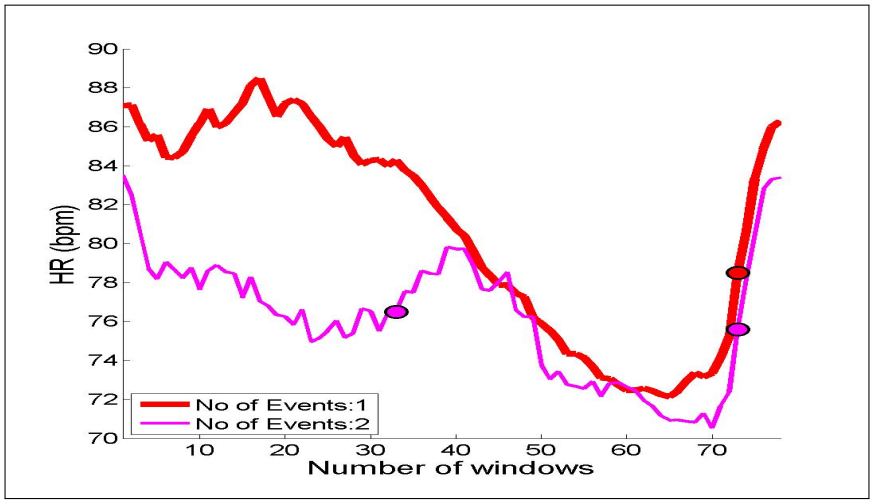
Figure 4.7: **Parameters:** Database = Fb, $T = 30$ minutes. The first image represents an alignment of the signals according to horological time; all signals are manually set to start at 8:30am. The second and third figures show the alignments proposed by the PPA and EBA algorithm respectively.



(a) Horological Alignment



(b) PPA Alignment



(c) EBA Alignment

Figure 4.8: **Parameters:** Database = Ma, $T = 20$ minutes. The first image represents an alignment of the signals according to horological time; all signals are manually set to start at 8:30am. The second and third figures show the alignments proposed by the PPA and EBA algorithm respectively.

circadian pattern. Indeed, note that the mean HR signal in figure 4.8a, where all HR signals have been manually aligned to start at 8:30 am, shows an increase until the 30^{th} time window, that is 6:30 pm and then remains somewhat stationary for 30 time windows, i.e. until midnight. Afterwards it is seen to decrease again. Such a behaviour can mean that either the subjects have a phase shift in their circadian cycle, or that the horological information is erroneous.

The PPA algorithm, seen in figure 4.8b displays a shifted but similar behaviour. Here, the apparent duration of the sleep phase is quite large as well, lasting about 50 time windows, i.e. 13,5 hours.

The EBA algorithm divides the subjects of the database into 3 categories, listed on table 4.4. The mean HR signals that corresponds to the first and second categories are depicted in figure 4.8b. In both cases, our algorithm has successfully managed to align the signals with respect to the circadian HR rhythm. Furthermore, in both cases, the signal apparently starts upon entering the awake state.

4.4 The Classification Process

Database	NSR	CHF	CHF	Ea	Ea	Ea	Eb	Eb	Eb
T (min)	10	10	20	10	20	30	10	20	30
Number	3	3	2	3	2	2	3	3	2

Database	Fa	Fa	Fa	Fb	Fb	Fb	Ma	Ma	Ma	Mb	Mb	Mb
T (min)	10	20	30	10	20	30	10	20	30	10	20	30
Number	3	3	2	3	2	2	3	2	2	3	3	2

Table 4.5: **Cases where the EBA algorithm has detected more than one category.** *Number* denotes the number of categories detected for each database and value of the parameter T (time length of window).

In this section we study the classification proposed by the EBA algorithm. As mentioned in the previous section, we have selected not to include in our remarks the last category produced by EBA. Recall that this category includes the uncategorised signals, that is, signals that have a different detected number of events than the other three categories. In all cases, only a very few number of signals belong to that category, a fact that could imply that these signals are either too noisy, or the subject exhibits an unusual circadian pattern.

Table 4.5 displays all the cases where EBA classifies the signals into more than one category. Note than, in our results, the cases where EBA unveils only one category correspond to lengthier time windows. Furthermore, in those cases, only one event is detected for each signals. Since we have inferred that this event corresponds to the transition between asleep state and awake state, it is a somewhat expected outcome, one

	NSR	CHF	Ea	Eb	Fa	Fb	Ma	Mb
Detected events: 1	185,24	194,15	199,70	187,41	199,13	187,67	217,53	217,43
Detected events: 2	188,94	219,04	196,60	197,01	194,88	183,84	203,04	199,90
Detected events: 3	184,76	205,98	206,29	200,51	194,59	184,69	198,71	207,09

Table 4.6: Areas below the curves of each signal, for $T = 10$ min

that reinforces our theory that the highest valued event is correlated with the waking up of the subject.

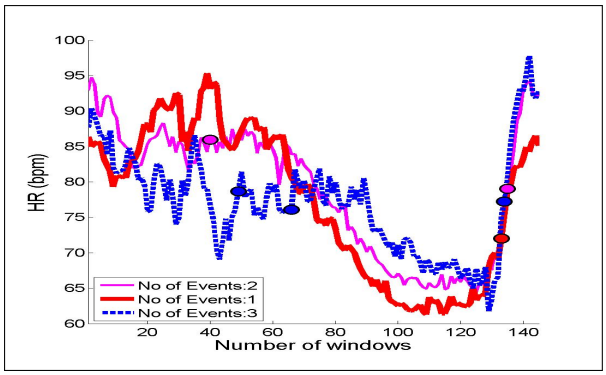
When the time window T is smaller, the EBA algorithm creates more classes. As we can observe in table 4.5, when $T = 10$ min, EBA detects three main categories; one, two or three events has been detected for each category. The circle marker on each line denotes the average moment when the event was detected, recovered by applying the edge detector on the mean HR signal of each category. Naturally, one of the detected events is found near the end of the day.

In figure 4.9, we can see the results of EBA for all databases when $T = 10$ minutes. One important observation that we can make here is that on five out of eight cases, during night time, the mean HR signal corresponding to one event (red lines in figure 4.9 is lower than the mean HR signal associated with two events (pink line in figure 4.9), which is in its turn lower than the one corresponding to three events (blue line in 4.9). Moreover, the one event signals are lower than all in 6 out of 8 cases, and as we will later see, this also stands for most of the other combinations of database and parameters.

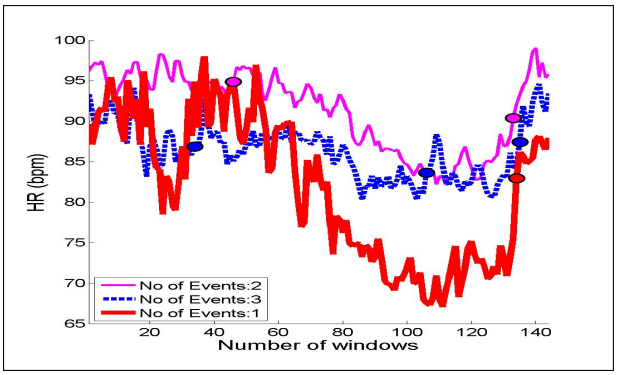
Recall that we refer to the beginning of the signal as *wakeup*. In general, when the EBA algorithm detects two or three events, the second is generally detected between 6.5 and 10 hours after wakeup, whereas the location of the third event is more arbitrary. For example, in figure 4.9b the third event was detected in the middle of night time. Nevertheless, the third evens seems to never be far from the second; in five out of six cases it is detected before night time.

In table 4.6 we can have calculated the area under the curves from all three categories. What we can see here is that generally, the areas do not differ much. This is an encouraging result; it would signify that the total energy spent by the heart for twenty four hours is unaffected by the number of detected events. Regardless of the increases and decreases in the heart rate, the area below that curve is preserved.

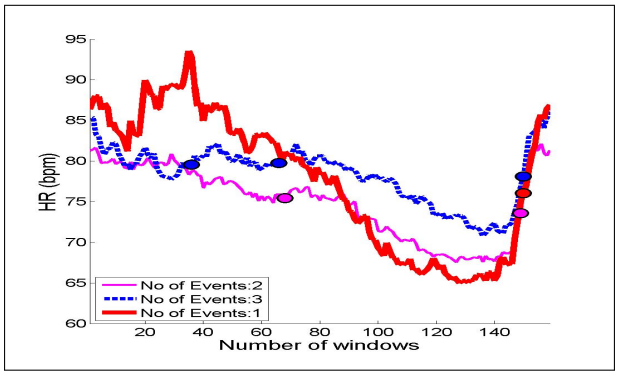
Our results show that these remarks also stand for larger time windows stand as well. However, the number of categories is inverse to the time length of T. As T increases, less categories are detected. This can be attributed to the sensitivity of the algorithm with respect to the information contained in each signal. When the time window is smaller, each signal contains more variation; thus, our algorithm detects edges that fade away when T is larger. Yet, the alignment capability of the algorithm is clearly not affected by the length of T. The corresponding plots for the results of the EBA algorithm when $T = 20, 30$ and 60 minutes can be found in the appendix.



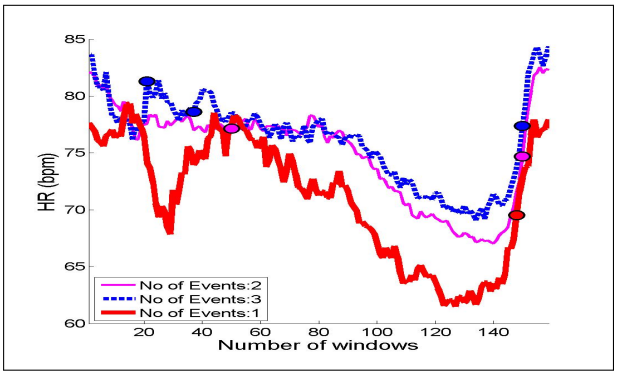
(a) NSR



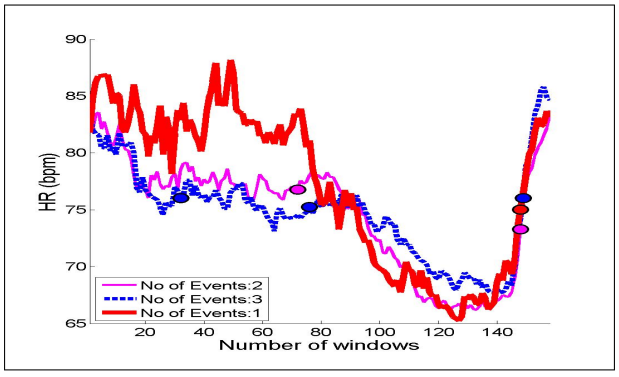
(b) CHF



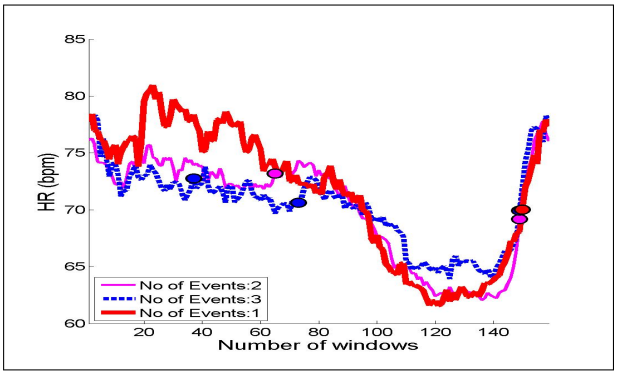
(c) Ea



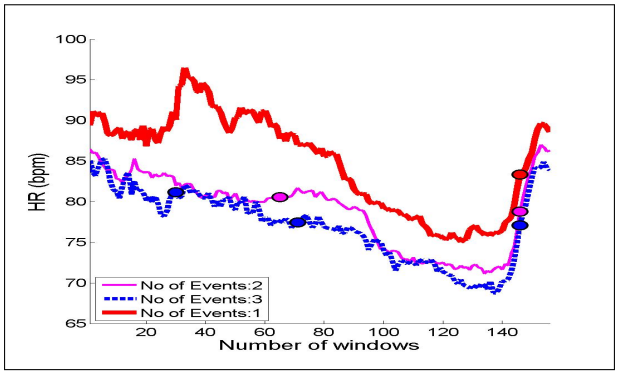
(d) Eb



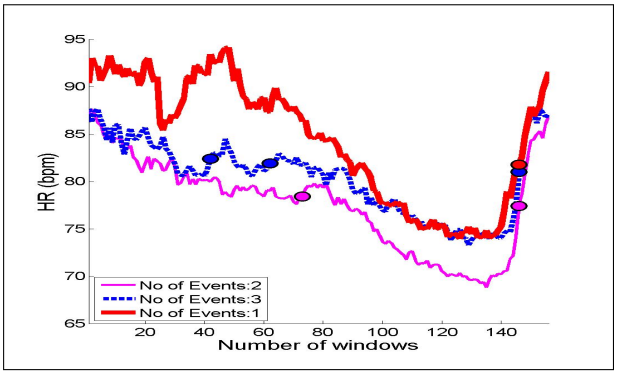
(e) Fa



(f) Fb

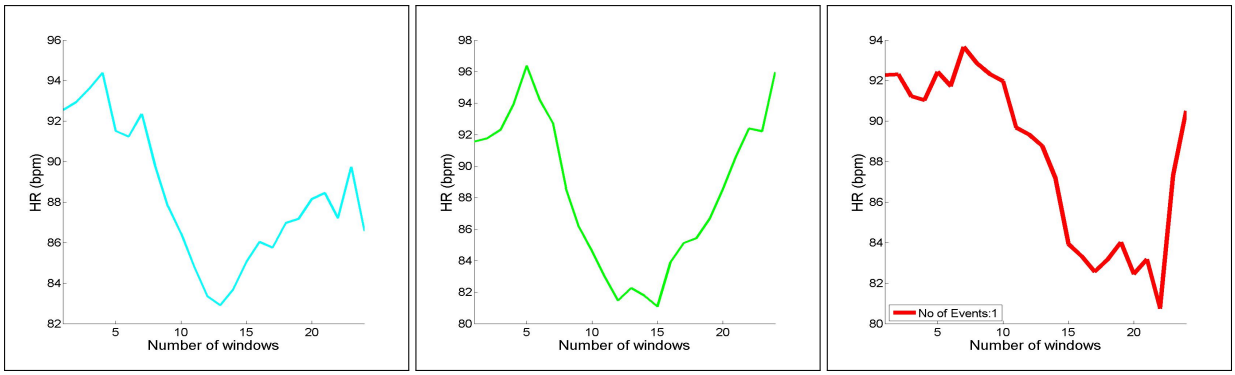


(g) Ma



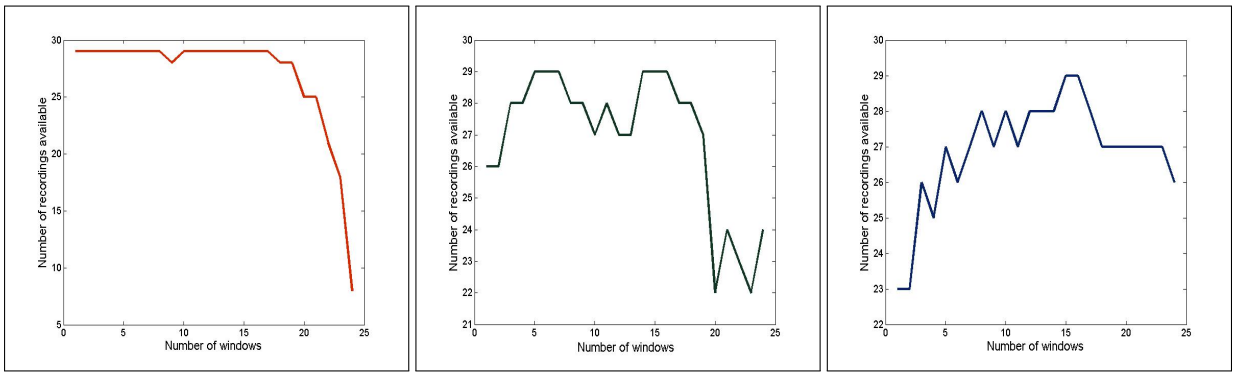
(h) Mb

Figure 4.9: Result of the EBA algorithm for all databases, where $T = 10$ minutes. In all cases, EBA has recovered three main categories according to the detected number of events.



(a) No alignment (b) PPA (c) EBA

Figure 4.10: Average Mean HR($H(x)$) signals with and without alignment.



(a) No alignment (b) PPA (c) EBA

Figure 4.11: Number of recordings available for each time window, with and without alignment.

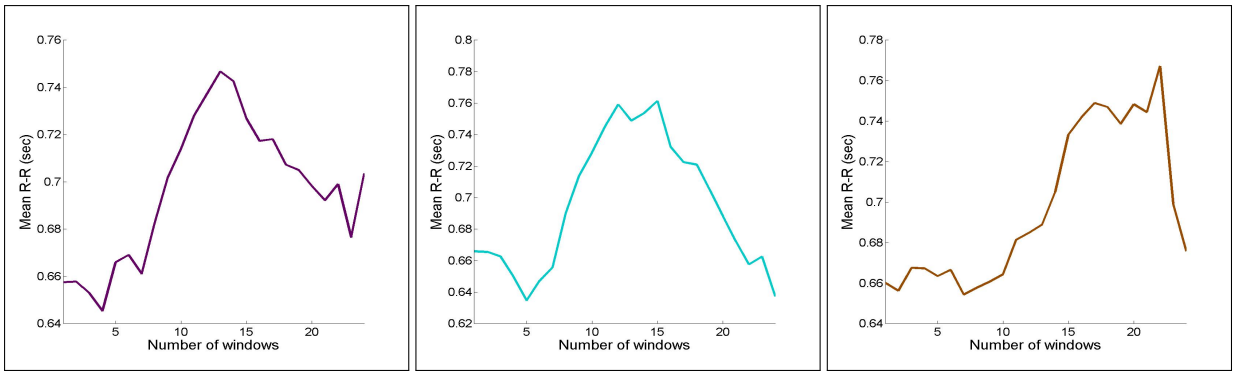
4.5 Aligning HRV Signals

After having aligned the R–R interval signals of a database, the next step is to use the starting points generated by each algorithm (in our case, PPA and EBA) to produce measurement signals, as explained in chapter 3 section 3.2.

Since the purpose of this work is R–R interval signal alignment, we do not comment on the nature of the results per se. In chapter 5, section 5.2, we discuss using the alignment algorithms to correctly align and interpret measurement signals as a proposition for future research.

As an example, we have selected to present and briefly comment the aligned measurement signals for the CHF dataset, for a time window of length $T = 60$ minutes. Recall that the CHF database does not contain any specific horological information, nevertheless all studies that have used this dataset quietly assume that they all start at the same time.

Figure 4.10 shows the signal with no alignment, averaged as presented by the database, aligned by the PPA algorithm, and aligned by the EBA algorithm. In figure 4.11 we can see how many recordings have been used for each time window. In other words, since some of the recordings last less than others, we can see their contribution. When no alignment



(a) No alignment (b) PPA (c) EBA

Figure 4.12: Average Mean R–R ($M(x)$) signals with and without alignment.

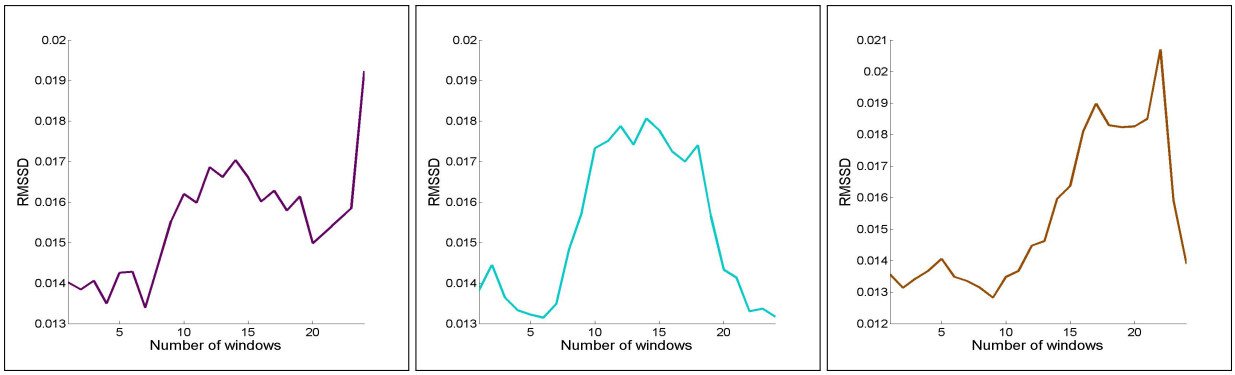
criterion is used, when we approach the end of the signal, only five recordings provide information to the average (figure 4.11a). Since this causes irregularities to appear at the end of most of the average signal measurements, it is important to note that our algorithms aligns the signals in such a way that, even though towards the end of the day we have less recordings, there are never less than 22 (figure 4.11b) and 23 (figure 4.11c) subjects contributing to the averaged signal.

Our results show that all signals demonstrate a circadian rhythmicity at some level. The application of both our methods modifies and corrects the measurement signals, whereas the aligned signals show some interesting characteristics.

4.5.1 Aligning the mean R–R

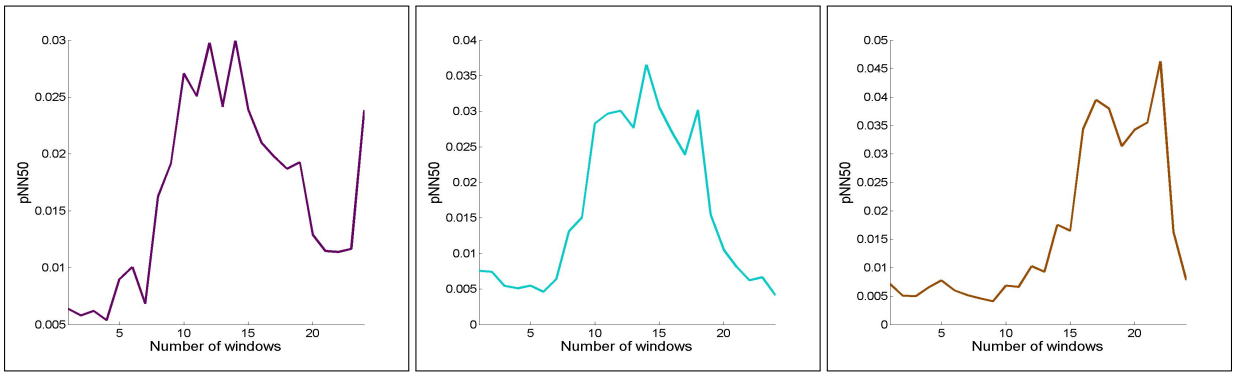
The mean R–R signal without alignment is presented in figure 4.12a. This particular signal can be seen as the inverse of the mean HR signal, since it depicts the average time between two consecutive beats. What we hence expect to see is an increase starting somewhere around the middle of the signal, which physically means that the heart beats at a slower rhythm, hence the subject is falling asleep. In the end we expect to se a sharp decrease, signifying that the heart starts beating at a normal pace again. While we can see the increase in the mean R–R interval, the decrease is not as steep as we expect, and in the end the signal does not end where it started; the mean R–R interval is 0,66 seconds in the beginning and around 0.71 seconds in the end.

The alignment proposed by the PPA algorithm gives the mean R–R interval signal a circadian pattern. In figure 4.12b, observe that the signal now stops where it has started, i.e. the mean R–R interval is between 0.66 and 0.64 seconds. We also observe the anticipated increase in the middle of the signal, between the 10^{th} and the 18^{th} time window. The EBA algorithm chooses a starting point that should coincide with the subjects waking up; indeed we can observe in figure 4.12c that for the first ten hours of the day, the mean R–R interval is relatively small; it increases between the 11^{th} and the 15^{th} hour, and drops in the last two hours. Mark that the decrease here is steep.



(a) No alignment (b) PPA (c) EBA

Figure 4.13: Average RMSSD ($R(x)$) signals with and without alignment.



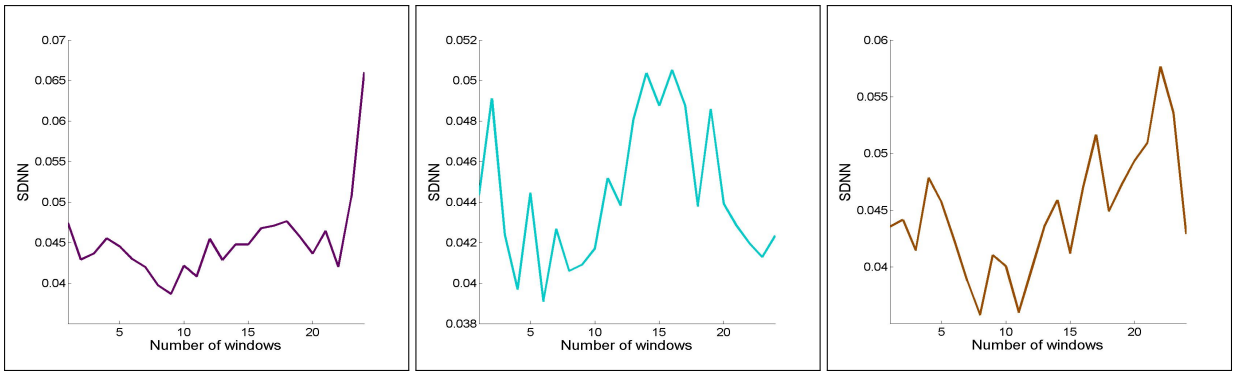
(a) No alignment (b) PPA (c) EBA

Figure 4.14: Average pNN50 ($P(x)$) signals with and without alignment.

4.5.2 Aligning the RMSSD

Figure 4.13a depicts the average RMSSD before applying any alignment criterion. We see here that the signal starts to increase at the 7^{th} time window up to the 10^{th}. It seems to decreases again after the 15^{th} time window, yet in the last hour the behaviour seems irregular. This is due to the fact that we most of the recordings stop earlier than that; what we see towards the end does not represent the whole dataset.

Our alignment algorithms both solve this issue. A we can observe in figures 4.13b and 4.13c, the last hour of the average RMSSD coincides with its beginning. In both signals, we see an increase in the average value, followed by a decrease, which occurs 10 and 8 hours later in the signal generated by the PPA and EBA algorithm respectively. Note in figure 4.13c, that the decrease is steeper, and that the maximum value is larger than the PPA generated signal. Since now the values of all windows are the average coming from more than 70% of the signal, this could mean that the RMSSD reaches its maximum value before the transition between the asleep and awake state.



(a) No alignment (b) PPA (c) EBA

Figure 4.15: Average SDNN ($S(x)$) signals with and without alignment.

4.5.3 Aligning the pNN50

The pNN50 seems to bear a strong similarity with the RMSSD. When comparing figures 4.13 and 4.14, we can see that their pattern is similar for all the alignment criteria.

With no alignment, the signal seen in figure 4.14a starts to increase at the 7^{th} hour, its value remains high for about 5 hours, which is when it starts decreasing again. Once more, there is an irregularity at the end of the signal, caused by lack of recordings.

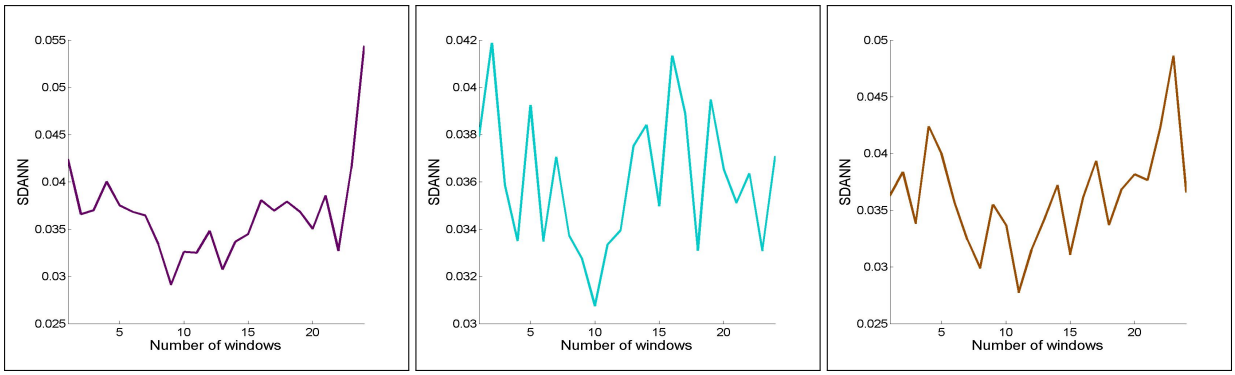
The average pNN50 given by the PPA algorithm displays a periodical behaviour. We can still observe in figure 4.14b an increase after the 6^{th} hour, followed by a somewhat steady series of values that last around 8 time windows, and then the value decreases again. This goes in line with the nature of the measurement. Recall that the pNN50 is the percentage of the successive interval differences greater than 50ms. We thus expect that number to increase during night time.

The EBA algorithm aligns the signal with respect to the transition between asleep and awake state. It is thus not surprising in figure 4.14c that the value starts low and then increases come night time. We observe same as in the previous measurement, that is the RMSSD, that the maximum value is observed around two hours before the end of the signal.

4.5.4 Aligning the SDNN

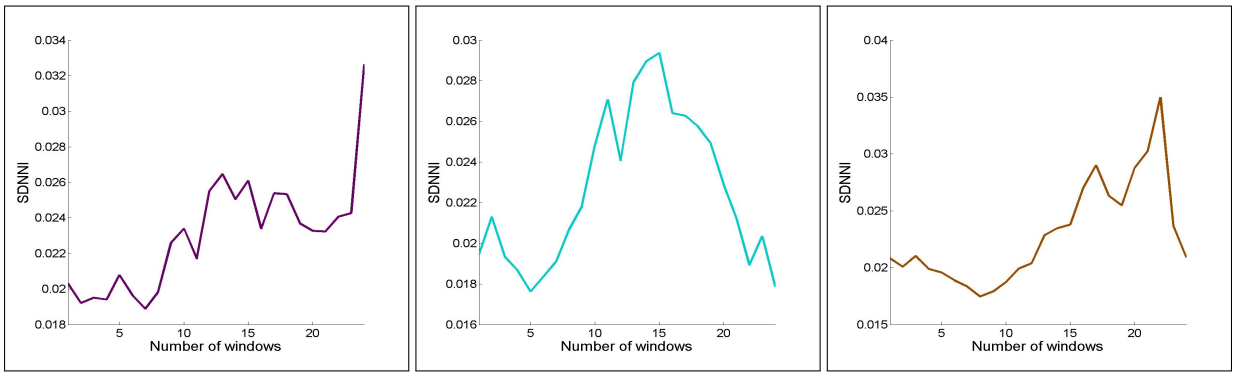
The SDNN signal before applying our alignment algorithms can be seen in figure 4.15a. The prevailing remark upon observing the signal is that the value in the last hour is much higher than the others. This prevents from observing the behaviour of the SDNN. An possible manual solution would be to remove that last value. This would enable us to have a better view of the circadian rhythmicity of the SDNN. It would however concern a 23-hour signal, and we would dismiss viable signal information.

The alignments proposed by the PPA and EBA algorithm are shown in figures 4.15b and 4.15c respectively. The SDNN signal is generally not an indicator of night time, since the standard deviation of the R–R interval signal can have high values even during the



(a) No alignment (b) PPA (c) EBA

Figure 4.16: Average SDANN ($A(x)$) signals with and without alignment.



(a) No alignment (b) PPA (c) EBA

Figure 4.17: Average SDNNI ($I(x)$) signals with and without alignment.

day. We can however observe that these values tend to increase during the night and decrease during the day.

4.5.5 Aligning the SDANN

The behaviour of the SDANN signals resemble that of the SDNN signals, which is nevertheless an expected observation. Again, very little conclusion can be drawn upon observing the unaligned signal from figure 4.16a. The last hour is again different from the entire signal, and the pattern is hard to resolve and associate with any human activity.

The alignments proposed by the PPA and EBA algorithms, seen in figures 4.16b and 4.16c respectively, provide us with the following important remarks. The PPA generated signal shows that the SDANN signal reaches its lowest point right at the begining of night time, whereas the one generated by EBA presents a peak right at its end. This could be further investigated, as these two points could be marking the transition in and out of sleep state.

4.5.6 Aligning the SDNNI

The average SDNNI without alignment is shown in figure 4.17a. Same as all the previously presented signals, it is showing an irregularity in the last hour. Moreover, the measurement signal does not seem to be periodic, its pattern being seemingly uncorrelated to the human circadian rhythmicity.

The alignments proposed by PPA and EBA once more give a more realistic shape to the average curve. In figure 4.17b, we can observe on the PPA proposed alignment that the SDANN increases during night time and decreases afterwards. This behaviour is a little less pronounced in figure 4.17c, yet the value one more time reaches its peak before the end of the signal.

Chapter 5

Conclusion

5.1 Discussion

The purpose of our work was to present new alignment methods and a classification algorithm for R–R interval signals. We have seen that HRV measurements signals are widely used in the bibliography, yet we are among the first to rise the issue of aligning intra subject R–R interval recordings.

We have developed and presented two methods that can be used to automatically align R–R interval signals and their respective measurement signals: the PPA method, and the EBA method. Our results show that both methods succeed in aligning such signals, even without any information on the time of the recording. The alignment is based on the circadian pattern observed on the mean HR signal; note that, the said pattern can vary among different subjects.

The PPA algorithm makes the assumption that all signals follow a similar circadian pattern. A common pattern that is generally observed for the HR rhythm is, a higher heart rate during the time when a person is awake, and lower values when one is asleep. Thus, when the time window is wide enough, i.e. larger than 20 minutes, the smaller fluctuations in these values tend to fade out, and the PPA algorithm achieves to align the signals with respect to that common pattern in a quick and efficient way.

While the PPA method is easy to implement and quite efficient, the drawbacks of the algorithm cannot be overlooked. First of all, the PPA algorithm aligns the signals and provides starting points that are calibrated with respect to the first mean HR signal in the signal set. Hence, it requires visual inspection to localize what can be perceived as the beginning of the daily cycle. Secondly, the most important drawback is the assumption

that all patterns are similar. This is unfortunately not a realistic assumption, which is why in some cases the PPA algorithm might give a distorted night time in the signal, or in some rare cases, a bad alignment.

These drawbacks have led to the development of the EBA method, which we have seen to remedy both these matters. The strong point of the EBA method is its ability to differentiate circadian patterns. EBA uses an edge detector to find increases in the heart rhythm, and the optimal edge detector that we chose to use [40] has the capacity to smooth the signal at the same time. In consequence, it detects only those variations that cause the HR value to have a lasting change. For example, a change in the value that only lasts a time window will remain undetected, whereas a change that persists will likely be identified as an event.

It is those persistent changes identified as events that are used to classify each signal into a category, whose label indicates precisely the number of events. Our results then clearly demonstrate that it is mostly the signals who present one event that follow the typical circadian pattern, the one that is associated with about 16 hours of active state, followed by 8 hours of sleep.

The EBA algorithm detects in many cases more than one pattern. In particular, when the time window has a smaller length, EBA discovers other patterns, and the number of events can be equal to two, three ore more events per mean HR signal. As we have seen, for each class, the algorithm successfully aligns all the signals and permits us to automatically set a relevant starting point; we set it to start two hours after the most significant increase, which should coincide with the transition between sleep and awake state. In most cases, the length of the time window does not seem to affect the result of PPA.

If we observe closely every output of the PPA algorithm, we will note that its outcome is in most cases closely correlated to the result of the EBA method. When the categories detected by the EBA algorithm are significantly diverse, the outcome of the PPA is not the anticipated pattern.

We can establish how crucial the correct alignment of R–R interval is by looking at the last part of our results. After using the starting points generated by our algorithms, every measurement result presents a circadian pattern, without loss of information. Both our methods manage to turn an set of R–R interval signals that are uncorrelated with respect to the dimension of time, into a set that has a common time reference, so that the evolution of each measurement can be studied in a more reliable manner.

We can conclude that both algorithms are successful in aligning R–R interval signals. We have seen the importance of the time window length in the alignment process, where small time windows lead the EBA algorithm to differentiate the R–R interval signals. The PPA is a faster method and can be used when one knows that all patterns are similar. The potential of the EBA algorithm needs to be further explore better; for the moment, we can safely point out its outstanding alignment capability.

5.2 Future Work

The future direction of our work branches into different directions, out of which three mai courses emerge.

The most obvious continuation is the further analysis and improvement of the algorithm themselves. There is still some work that can be done towards to enforce the robustness of our algorithms. They might be improved int becoming less input sensitive, as for example the PPA algorithm, whose output is closely connected to the shape and cycle of the first signal in the input set.

The natural course of our work is to apply the alignments on other measurements, using different datasets and time window lengths. Another interesting approach could focus on examining the use of other measurements that also present a circadian rhythmicity, and modify the algorithms so that they either use other measurements as their input, or a combination of relevant measurements.

One of the most important results generated by the EBA algorithm is the existence of different circadian patterns. The classification provided by the EBA algorithm is the most interesting output of this work. Since the classifying process is fast, there is much research that can continue over the resulting classes. An example would be the determination of the nature of events, whether they are indeed related only to transitions between the awake and asleep state, or not. Another idea would be to examine other measurement with respect to the categories created by the algorithm.

We conclude this research by stating that the result of this work should be further investigated, since it is our sincere belief that it can unveil a lot of information, valuable information to be used in medical research towards the prognosis and diagnosis of cardiovascular related conditions.

Bibliography

[1] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[2] B. Phibbs, *The Human Heart: A Basic Guide to Heart Disease*. Lippincott Williams & Wilkins, a Wolters Kluwer business, 2nd ed., 2007.

[3] V. Zarzoso, O. Meste, P. Comon, *et al.*, *Noninvasive Cardiac Signal Analysis Using Data Decomposition Techniques*. Springer, 2013.

[4] L. Sornnmo and P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applictions*, ch. 1. Elsevier Academic Press, 2005.

[5] J. E. Hall, *Guyton and Hall textbook of medical physiology*. Elsevier Health Sciences, 2010.

[6] E. P. Widmaier, H. Raff, and K. T. Strang, *Vander's human physiology: the mechanisms of body function*. McGraw-Hill Higher Education, 2011.

[7] J. Malmivuo and R. Plonsey, *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. Oxford University Press, 1995.

[8] M. R. Neuman, "Biopotential electrodes," in *Medical instrumentation: application and design* (J. Webster, ed.), ch. 5, pp. 183–232, John Wiley & Sons, 2009.

[9] L. Sornnmo and P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applictions*, ch. 6, pp. 419–432. Elsevier Academic Press, 2005.

[10] E. Niedermeyer and F. L. da Silva, *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.

[11] T. F. of the European Society of Cardiology, T. N. A. S. of Pacing & Electrophysiology, *et al.*, "Heart rate variability: standards of measurement, physiological interpretation and clinical use," *Circulation*, vol. 93, no. 5, pp. 1043–1065, 1996.

[12] A. Murray, D. Ewing, I. Campbell, J. Neilson, and B. Clarke, "Rr interval variations in young male diabetics.," *British heart journal*, vol. 37, no. 8, pp. 882–885, 1975.

[13] T. Bennett, P. Fentem, D. Fitton, J. Hampton, D. Hosking, and P. Riggott, "Assessment of vagal control of the heart in diabetes. measures of rr interval variation under different conditions.," *British heart journal*, vol. 39, no. 1, pp. 25–28, 1977.

[14] D. Ewing, D. Borsey, F. Bellavere, and B. Clarke, "Cardiac autonomic neuropathy in diabetes: comparison of measures of rr interval variation," *Diabetologia*, vol. 21, no. 1, pp. 18–24, 1981.

[15] D. Ewing and B. Clarke, "Diagnosis and management of diabetic autonomic neuropathy.," *British medical journal (Clinical research ed.)*, vol. 285, no. 6346, p. 916, 1982.

[16] D. Ziegler, C. P. Zentai, S. Perz, W. Rathmann, B. Haastert, A. Döring, and C. Meisinger, "Prediction of mortality using measures of cardiac autonomic dysfunction in the diabetic and nondiabetic population the monica/kora augsburg cohort study," *Diabetes care*, vol. 31, no. 3, pp. 556–561, 2008.

[17] A. Algra, J. Tijssen, J. Roelandt, J. Pool, and J. Lubsen, "Heart rate variability from 24-hour electrocardiography and the 2-year risk for sudden death.," *Circulation*, vol. 88, no. 1, pp. 180–185, 1993.

[18] J. A. Taylor, D. L. Carr, C. W. Myers, and D. L. Eckberg, "Mechanisms underlying very-low-frequency rr-interval oscillations in humans," *Circulation*, vol. 98, no. 6, pp. 547–555, 1998.

[19] C. W. Hogue, P. P. Domitrovich, P. K. Stein, G. D. Despotis, L. Re, R. B. Schuessler, R. E. Kleiger, and J. N. Rottman, "Rr interval dynamics before atrial fibrillation in patients after coronary artery bypass graft surgery," *Circulation*, vol. 98, no. 5, pp. 429–434, 1998.

[20] H. V. Huikuri, T. H. Mäkikallio, C.-K. Peng, A. L. Goldberger, U. Hintze, M. Møller, *et al.*, "Fractal correlation properties of rr interval dynamics and mortality in patients with depressed left ventricular function after an acute myocardial infarction," *Circulation*, vol. 101, no. 1, pp. 47–53, 2000.

[21] J. Ortak, G. Weitz, U. K. Wiegand, F. Bode, F. Eberhardt, H. A. Katus, G. Richardt, H. Schunkert, and H. Bonnemeier, "Changes in heart rate, heart rate variability, and heart rate turbulence during evolving reperfused myocardial infarction," *Pacing and clinical electrophysiology*, vol. 28, no. s1, pp. S227–S232, 2005.

[22] D. Matt, H. Scheffel, S. Leschka, T. G. Flohr, B. Marincek, P. A. Kaufmann, and H. Alkadhi, "Dual-source ct coronary angiography: image quality, mean heart rate, and heart rate variability," *American Journal of Roentgenology*, vol. 189, no. 3, pp. 567–573, 2007.

[23] A. Y. Shalev, S. Freedman, T. Peri, D. Brandes, T. Sahar, S. P. Orr, and R. K. Pitman, "Prospective study of posttraumatic stress disorder and depression following trauma," 2014.

[24] K. C. Bilchick, B. Fetics, R. Djoukeng, S. G. Fisher, R. D. Fletcher, S. N. Singh, E. Nevo, and R. D. Berger, "Prognostic value of heart rate variability in chronic congestive heart failure (veterans affairsβêŭ survival trial of antiarrhythmic therapy in congestive heart failure)," *The American journal of cardiology*, vol. 90, no. 1, pp. 24–28, 2002.

[25] J. D. Keith, "Congestive heart failure," *Pediatrics*, vol. 18, no. 3, pp. 491–500, 1956.

[26] D. Van Hoogenhuyze, N. Weinstein, G. J. Martin, J. S. Weiss, J. W. Schaad, X. N. Sahyouni, D. Fintel, W. J. Remme, and D. H. Singer, "Reproducibility and relation to mean heart rate of heart rate variability in normal subjects and in patients with congestive heart failure secondary to coronary artery disease," *The American journal of cardiology*, vol. 68, no. 17, pp. 1668–1676, 1991.

[27] P.-Y. Liu, W.-C. Tsai, L.-J. Lin, Y.-H. Li, T.-H. Chao, L.-M. Tsai, and J.-H. Chen, "Time domain heart rate variability as a predictor of long-term prognosis after acute myocardial infarction.," *Journal of the Formosan Medical Association= Taiwan yi zhi*, vol. 102, no. 7, pp. 474–479, 2003.

[28] I. Antelmi, R. S. De Paula, A. R. Shinzato, C. A. Peres, A. J. Mansur, and C. J. Grupi, "Influence of age, gender, body mass index, and functional capacity on heart rate variability in a cohort of subjects without heart disease," *The American journal of cardiology*, vol. 93, no. 3, pp. 381–385, 2004.

[29] H. Tsuji, M. G. Larson, F. J. Venditti, E. S. Manders, J. C. Evans, C. L. Feldman, and D. Levy, "Impact of reduced heart rate variability on risk for cardiac events the framingham heart study," *Circulation*, vol. 94, no. 11, pp. 2850–2855, 1996.

[30] J. de Mairan, "Observation botanique," *Hist. Acad. Roy. Sci*, vol. 35, p. 36, 1729.

[31] J. C. Dunlap, J. J. Loros, and P. J. DeCoursey, *Chronobiology: biological timekeeping.* Sinauer Associates, 2004.

[32] H. V. Huikuri, M. Niemelä, S. Ojala, A. Rantala, M. Ikäheimo, and K. Airaksinen, "Circadian rhythms of frequency domain measures of heart rate variability in healthy subjects and patients with coronary artery disease. effects of arousal and upright posture.," *Circulation*, vol. 90, no. 1, pp. 121–126, 1994.

[33] E. M. Ekholm, J. Hartiala, and H. V. Huikuri, "Circadian rhythm of frequency-domain measures of heart rate variability in pregnancy," *BJOG: An International Journal of Obstetrics & Gynaecology*, vol. 104, no. 7, pp. 825–828, 1997.

[34] D. Ewing, J. Neilson, C. Shapiro, J. Stewart, and W. Reid, "Twenty four hour heart rate variability: effects of posture, sleep, and time of day in healthy controls and comparison with bedside tests of autonomic function in diabetic patients.," *British heart journal*, vol. 65, no. 5, pp. 239–244, 1991.

[35] M. Nakagawa, T. Iwao, S. Ishida, H. Yonemochi, T. Fujino, T. Saikawa, and M. Ito, "Circadian rhythm of the signal averaged electrocardiogram and its relation to heart rate variability in healthy subjects," *Heart*, vol. 79, no. 5, pp. 493–496, 1998.

[36] M. M. Massin, K. Maeyns, N. Withofs, F. Ravet, and P. Gérard, "Circadian rhythm of heart rate and heart rate variability," *Archives of Disease in Childhood*, vol. 83, no. 2, pp. 179–182, 2000.

[37] G. Vandewalle, B. Middleton, S. M. Rajaratnam, B. M. Stone, B. Thorleifsdottir, J. Arendt, and D.-J. DIJK, "Robust circadian rhythm in heart rate and its variability: influence of exogenous melatonin and photoperiod," *Journal of sleep research*, vol. 16, no. 2, pp. 148–155, 2007.

[38] T. Yoshizaki, Y. Kawano, Y. Tada, A. Hida, T. Midorikawa, K. Hasegawa, T. Mitani, T. Komatsu, and F. Togo, "Diurnal 24-hour rhythm in ambulatory heart rate variability during the day shift in rotating shift workers," *Journal of biological rhythms*, vol. 28, no. 3, pp. 227–236, 2013.

[39] J. Duffy, D. Dijk, E. Hall, and C. Czeisler, "Relationship of endogenous circadian melatonin and temperature rhythms to self-reported preference for morning or evening activity in young and older people.," *Journal of investigative medicine: the official publication of the American Federation for Clinical Research*, vol. 47, no. 3, pp. 141–150, 1999.

[40] R. Deriche, "Using canny's criteria to derive a recursively implemented optimal edge detector," *International journal of computer vision*, vol. 1, no. 2, pp. 167–187, 1987.

[41] P. K. Stein, R. E. Kleiger, P. P. Domitrovich, K. B. Schechtman, and J. N. Rottman, "Clinical and demographic determinants of heart rate variability in patients post myocardial infarction: insights from the cardiac arrhythmia suppression trial (cast)," *Clinical cardiology*, vol. 23, no. 3, pp. 187–194, 2000.

[42] R. L. Goldsmith, J. T. Bigger, R. C. Steinman, and J. L. Fleiss, "Comparison of 24-hour parasympathetic activity in endurance-trained and untrained young men," *Journal of the American College of Cardiology*, vol. 20, no. 3, pp. 552–558, 1992.

[43] J. T. Bigger, J. L. Fleiss, R. C. Steinman, L. M. Rolnitzky, W. J. Schneider, and P. K. Stein, "Rr variability in healthy, middle-aged persons compared with patients with chronic coronary heart disease or recent acute myocardial infarction," *Circulation*, vol. 91, no. 7, pp. 1936–1943, 1995.

[44] P. K. Stein, A. A. Ehsani, P. P. Domitrovich, R. E. Kleiger, and J. N. Rottman, "Effect of exercise training on heart rate variability in healthy older adults," *American heart journal*, vol. 138, no. 3, pp. 567–576, 1999.

[45] J. Mietus, C. Peng, I. Henry, R. Goldsmith, and A. Goldberger, "The pnnx files: re-examining a widely used heart rate variability measure," *Heart*, vol. 88, no. 4, pp. 378–380, 2002.

[46] H. Krum, T. Bigger, R. L. Goldsmith, and M. Packer, "Effect of long-term digoxin therapy on autonomic function in patients with chronic heart failure," *Journal of the American College of Cardiology*, vol. 25, no. 2, pp. 289–294, 1995.

[47] R. L. Goldsmith, J. T. Bigger, D. M. Bloomfield, H. Krum, R. C. Steinman, J. Sackner-Bernstein, and M. Packer, "Long-term carvedilol therapy increases parasympathetic nervous system activity in chronic congestive heart failure," *The American journal of cardiology*, vol. 80, no. 8, pp. 1101–1104, 1997.

Appendix

This appendix contains all the results for each experiment run on a combination of dataset and parameter, as described in chapter 4. We thereby present all the outputs generated by running our algorithms (see section 3.3) with every one of the eight datasets in our research, and each one of the four selected values of the parameter describing the length of the time window, $T = 10, 20, 30$ and 60 minutes.

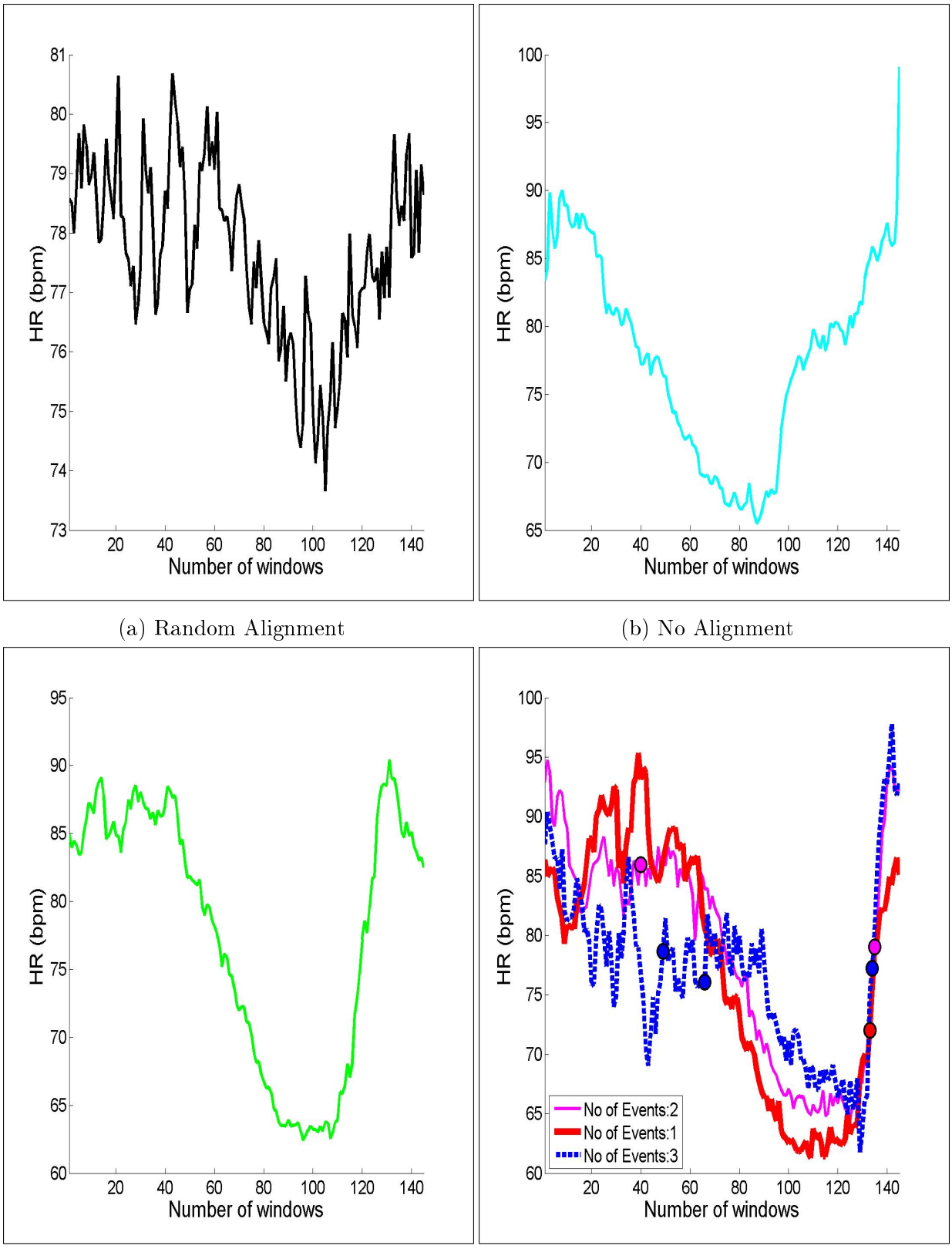
For each experiment, the depicted figures show the average mean HR signal. The following four diagrams are shown per combination.

a. A random shuffle of the starting points of the signals; the shuffling process is explained in section 4.3.1.

b. The average mean HR signal when all signals start at the same horological time.

c. The result of the PPA algorithm.

d. The result of the EBA algorithm.



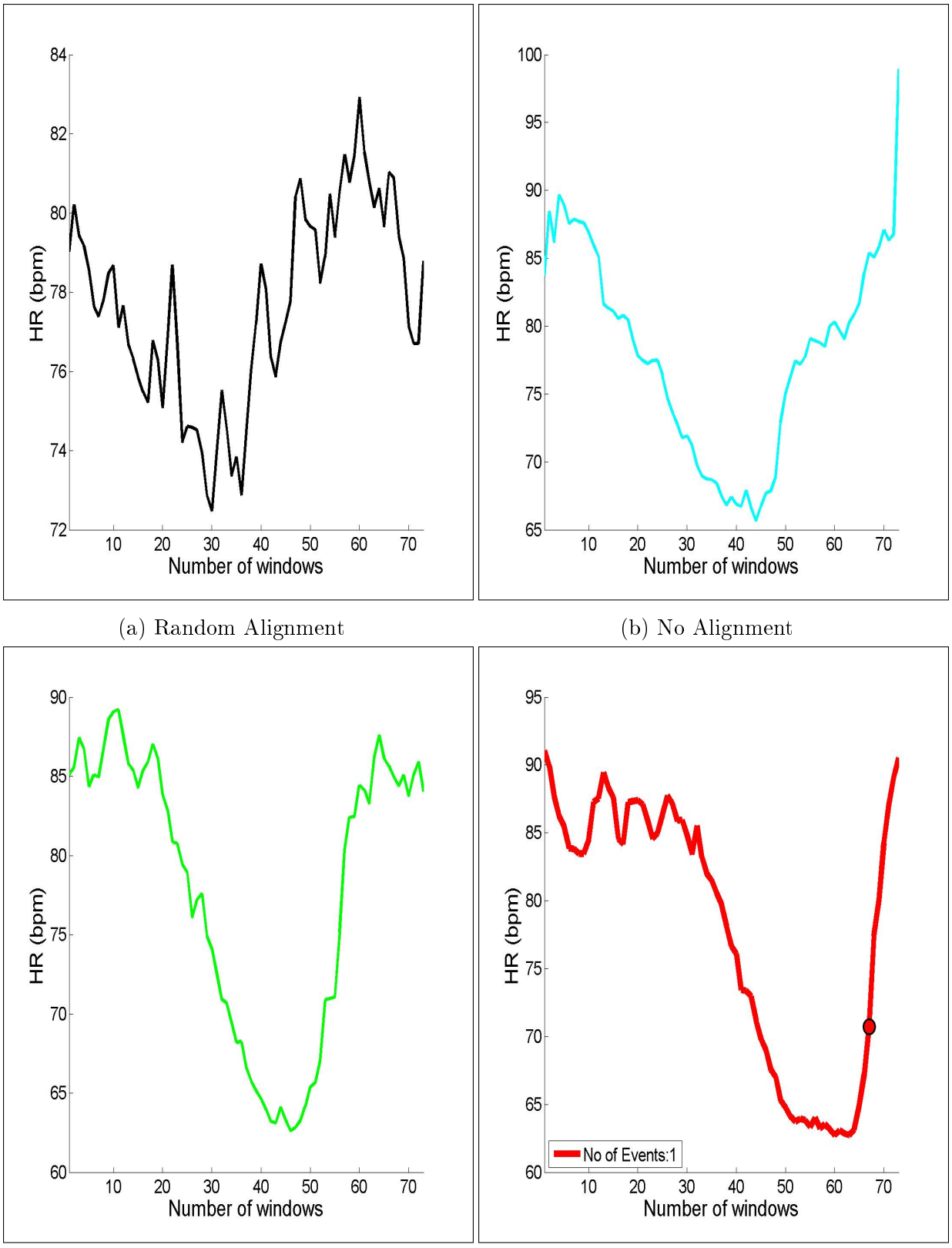
(a) Random Alignment

(b) No Alignment

(c) PPA result

(d) EBA result

Figure 5.1: **Parameters:** Database = NSR, $T = 10$ minutes



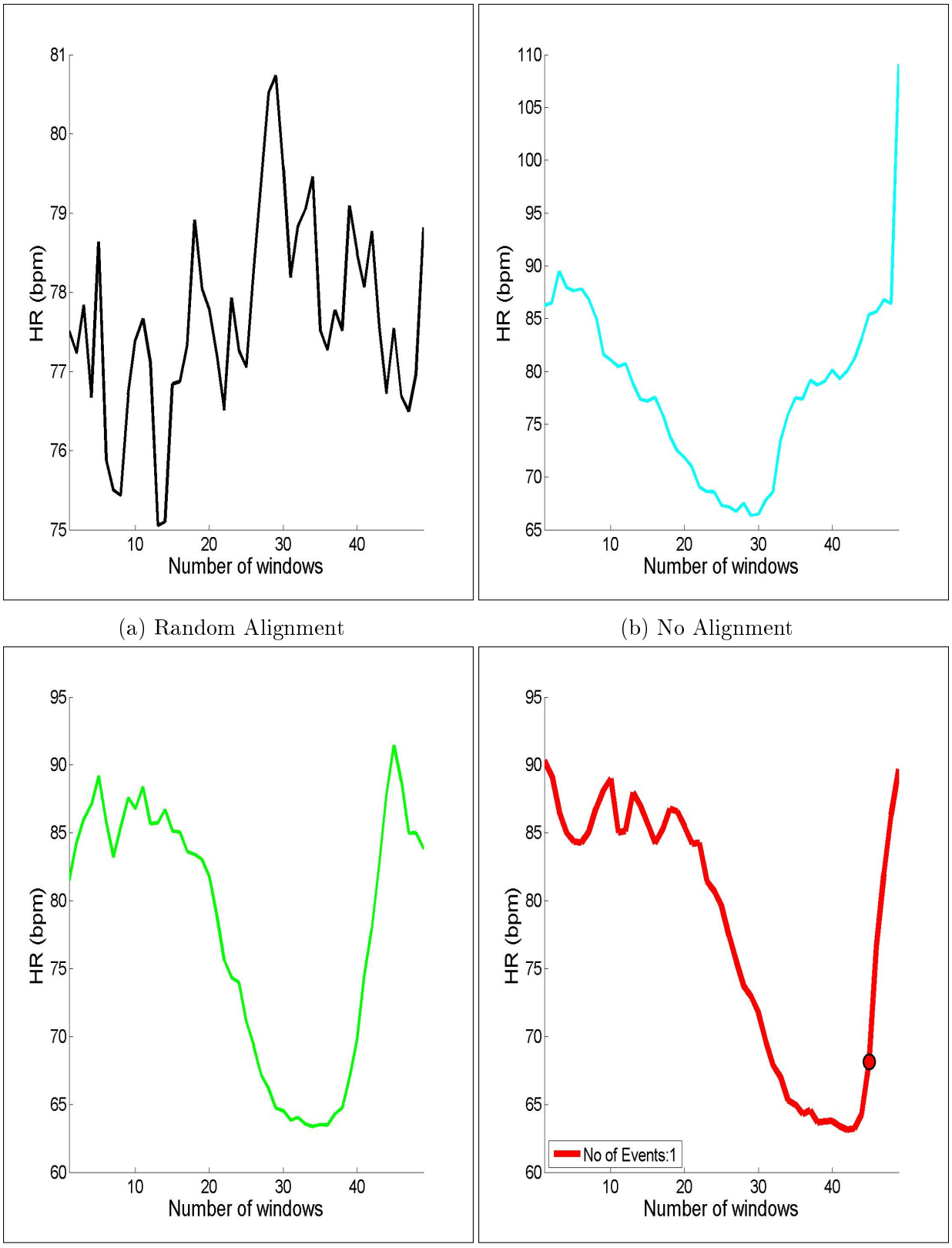
(a) Random Alignment

(b) No Alignment

(c) PPA result

(d) EBA result

Figure 5.2: **Parameters:** Database = NSR, $T = 20$ minutes



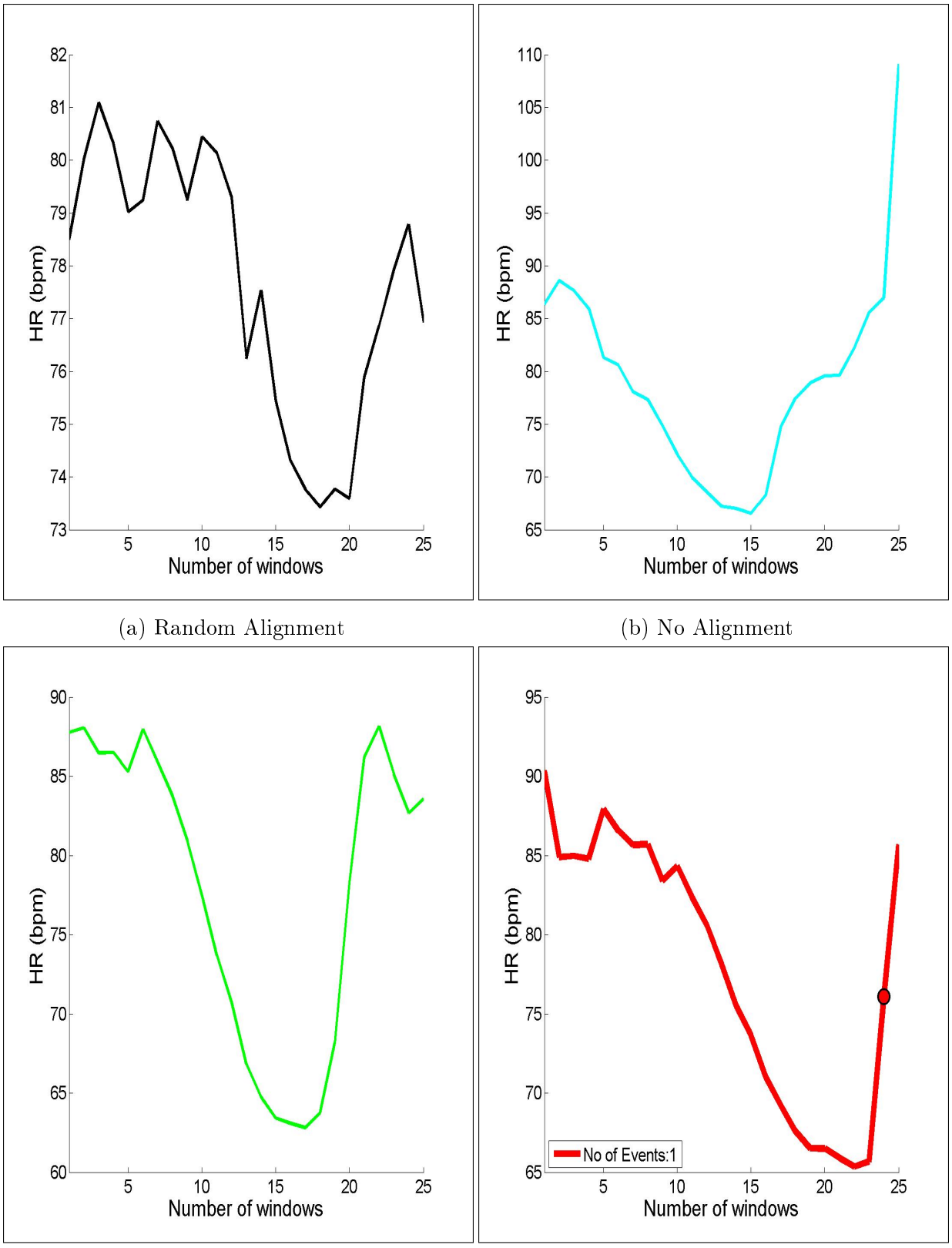
(a) Random Alignment

(b) No Alignment

(c) PPA result

(d) EBA result

Figure 5.3: **Parameters:** Database = NSR, $T = 30$ minutes



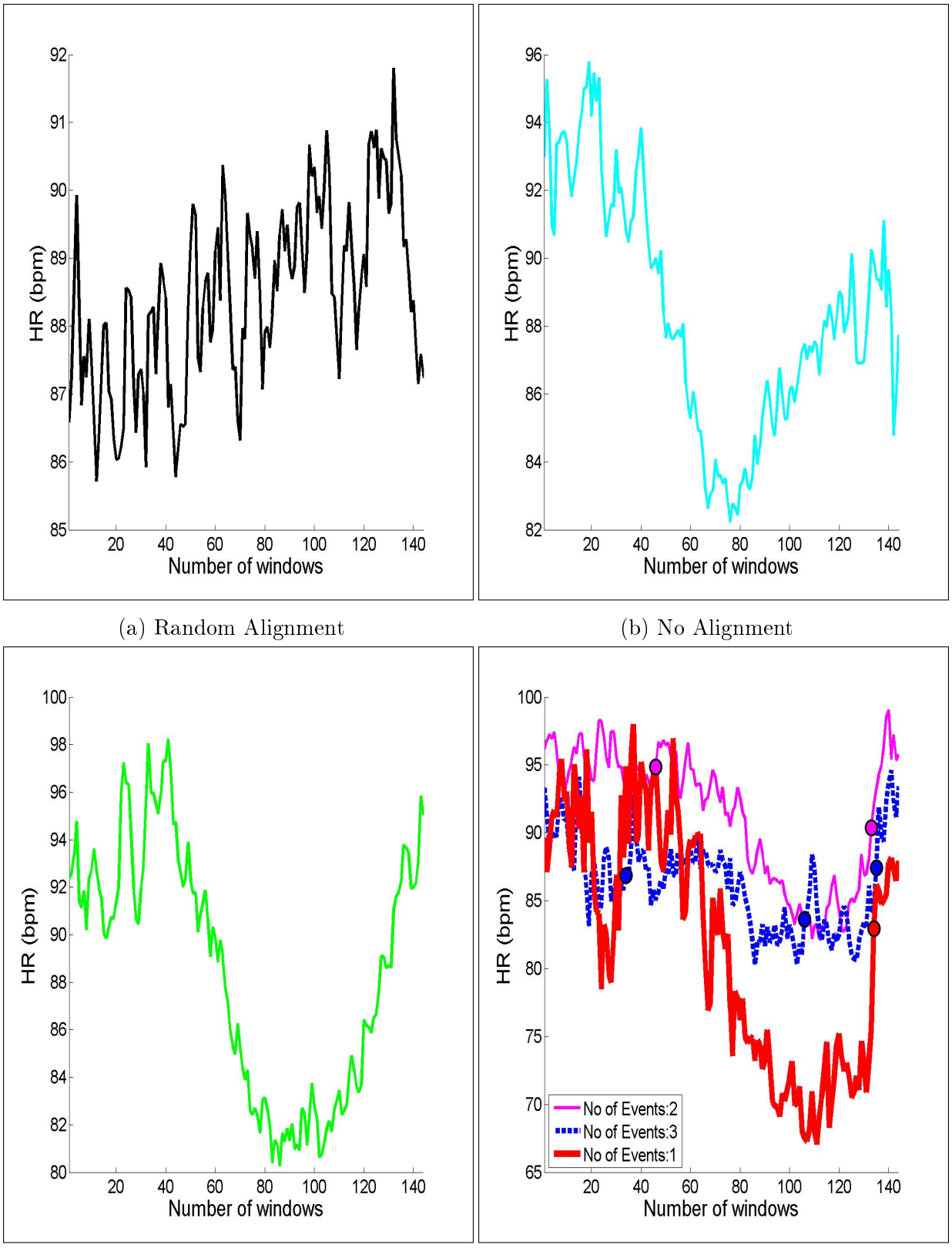
(a) Random Alignment

(b) No Alignment

(c) PPA result

(d) EBA result

Figure 5.4: **Parameters:** Database = NSR, $T = 60$ minutes



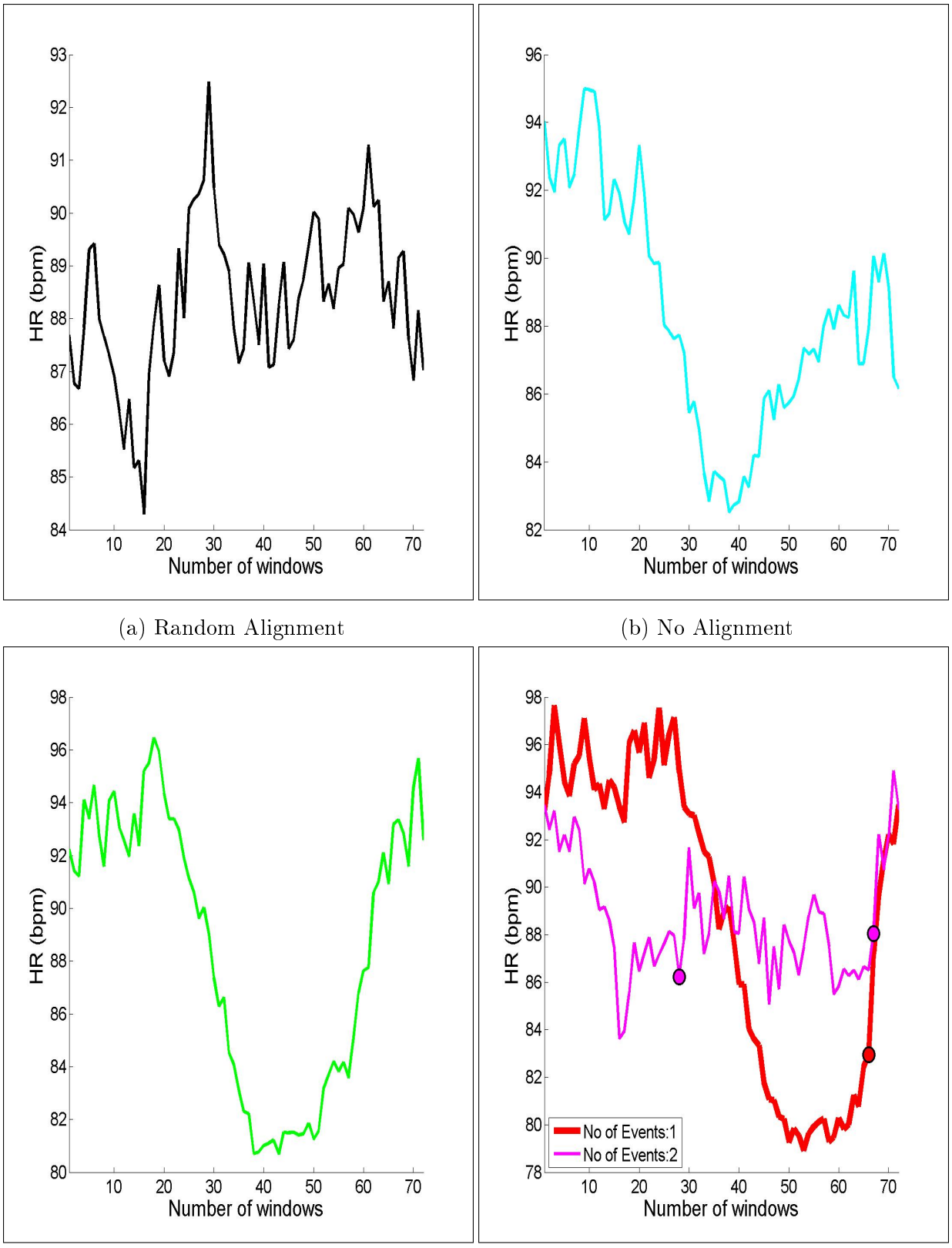
(a) Random Alignment

(b) No Alignment

(c) PPA result

(d) EBA result

Figure 5.5: **Parameters:** Database = CHF, $T = 10$ minutes



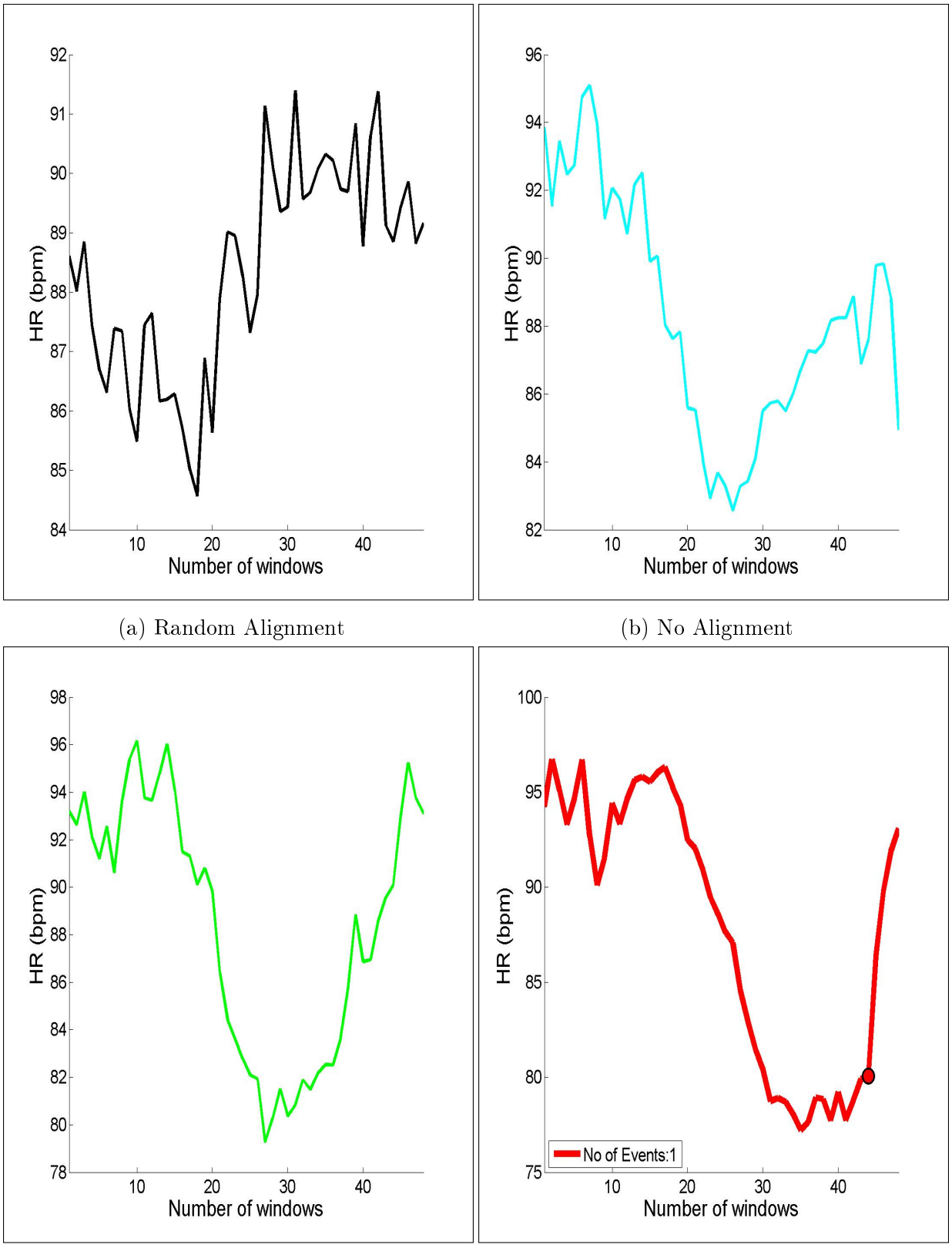
(a) Random Alignment

(b) No Alignment

(c) PPA result

(d) EBA result

Figure 5.6: **Parameters:** Database = CHF, $T = 10$ minutes



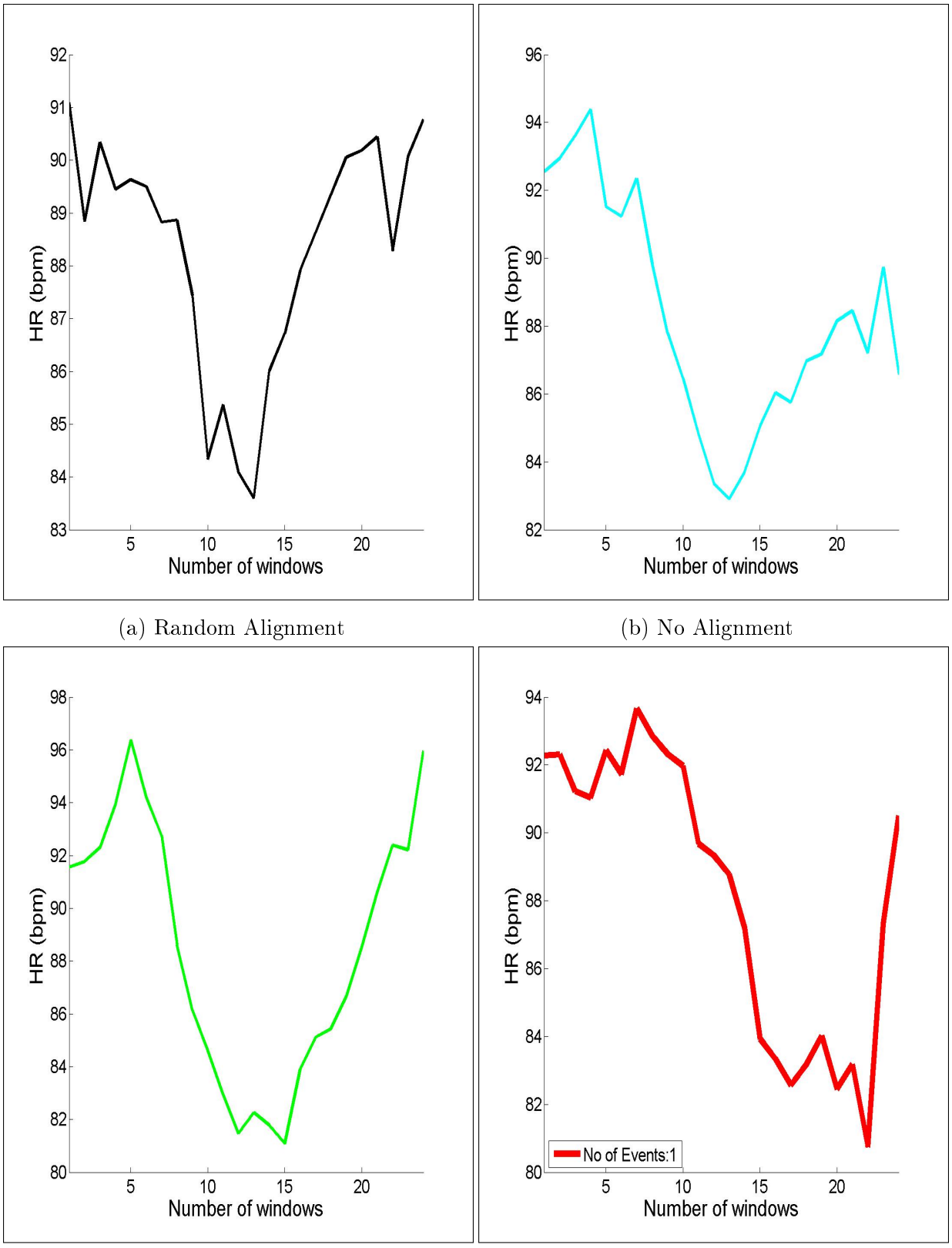
(a) Random Alignment

(b) No Alignment

(c) PPA result

(d) EBA result

Figure 5.7: **Parameters:** Database = CHF, $T = 10$ minutes



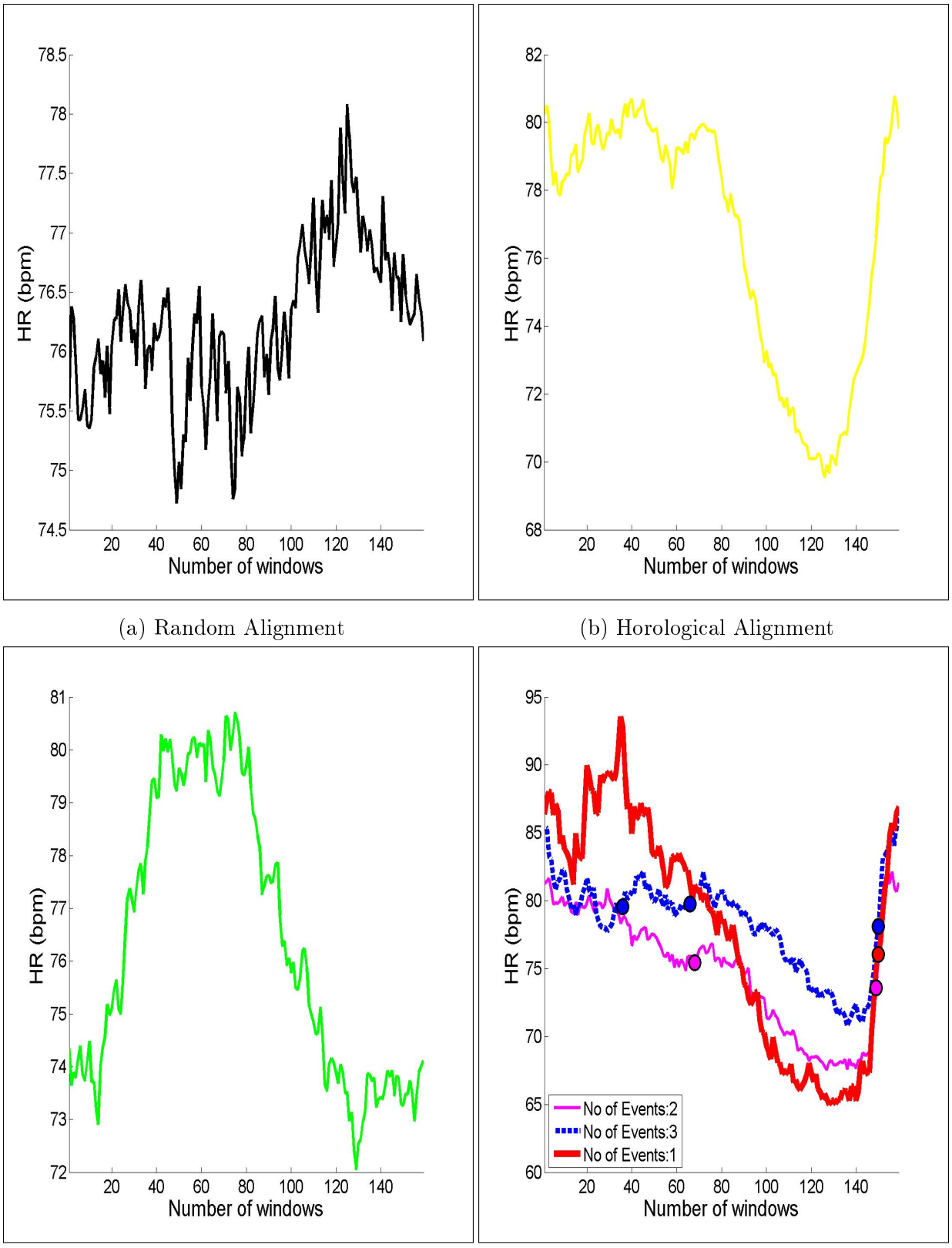
(a) Random Alignment

(b) No Alignment

(c) PPA result

(d) EBA result

Figure 5.8: **Parameters:** Database = CHF, $T = 60$ minutes



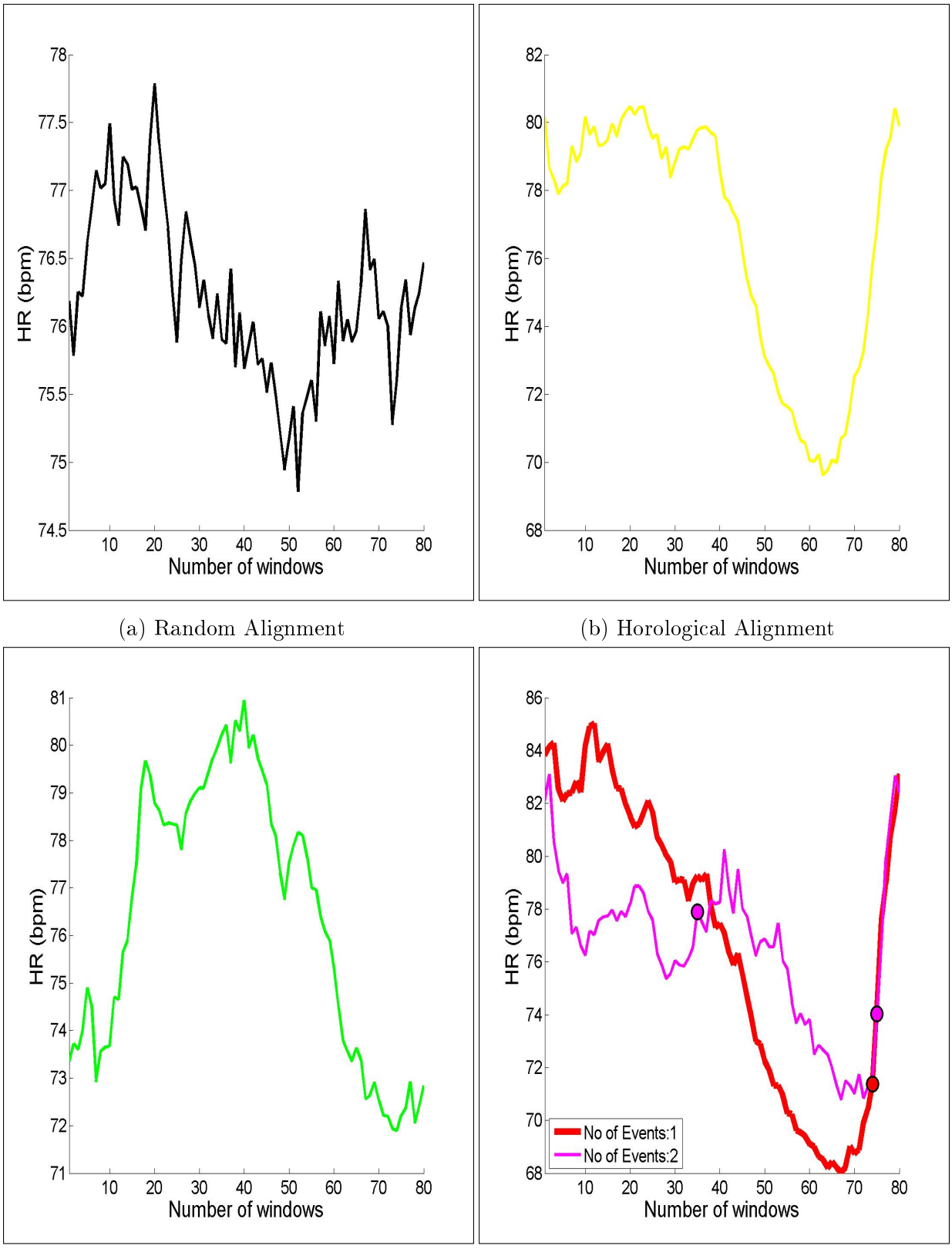
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.9: **Parameters:** Database = Ea, $T = 10$ minutes



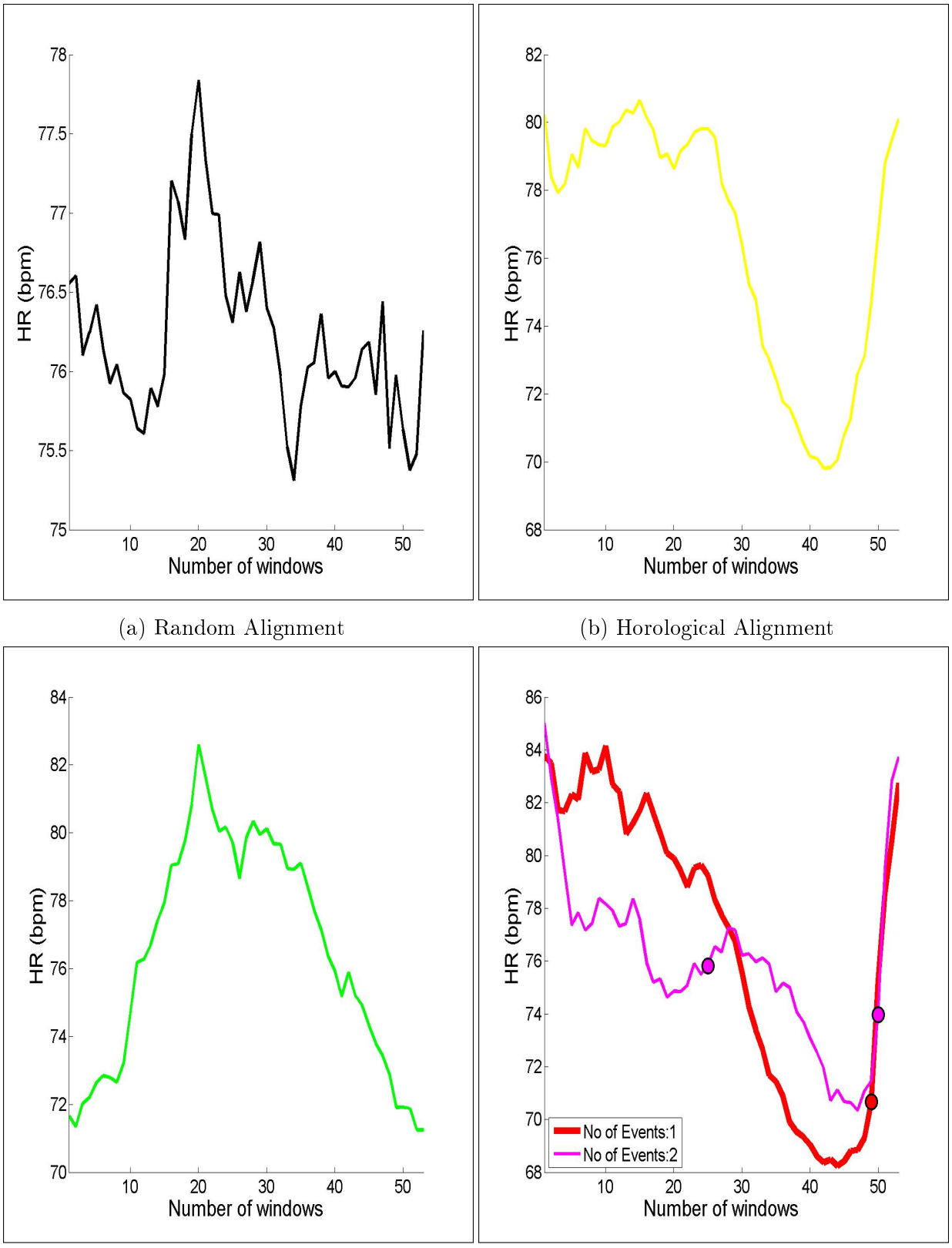
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.10: **Parameters:** Database = Ea, $T = 20$ minutes



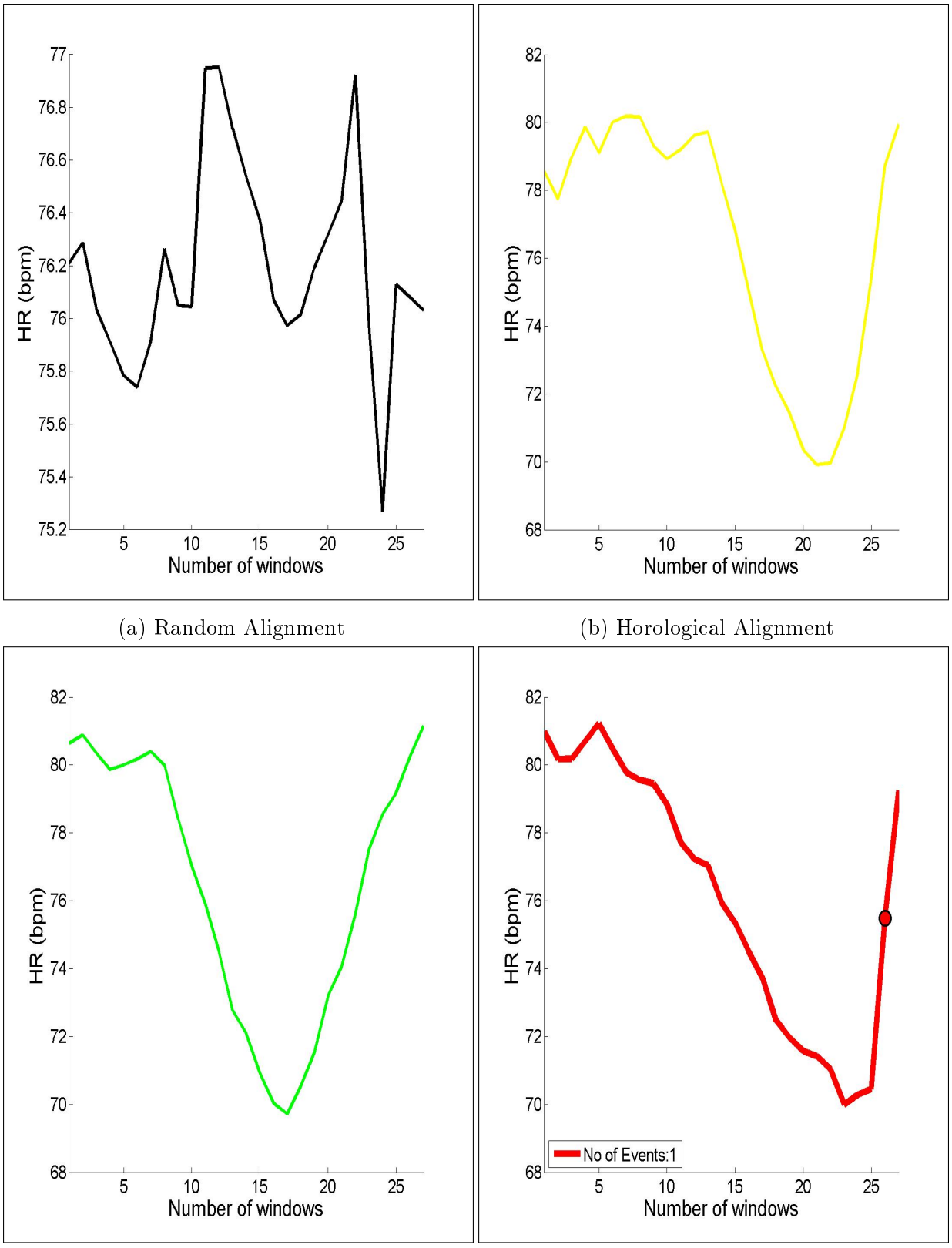
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.11: **Parameters:** Database = Ea, $T = 30$ minutes



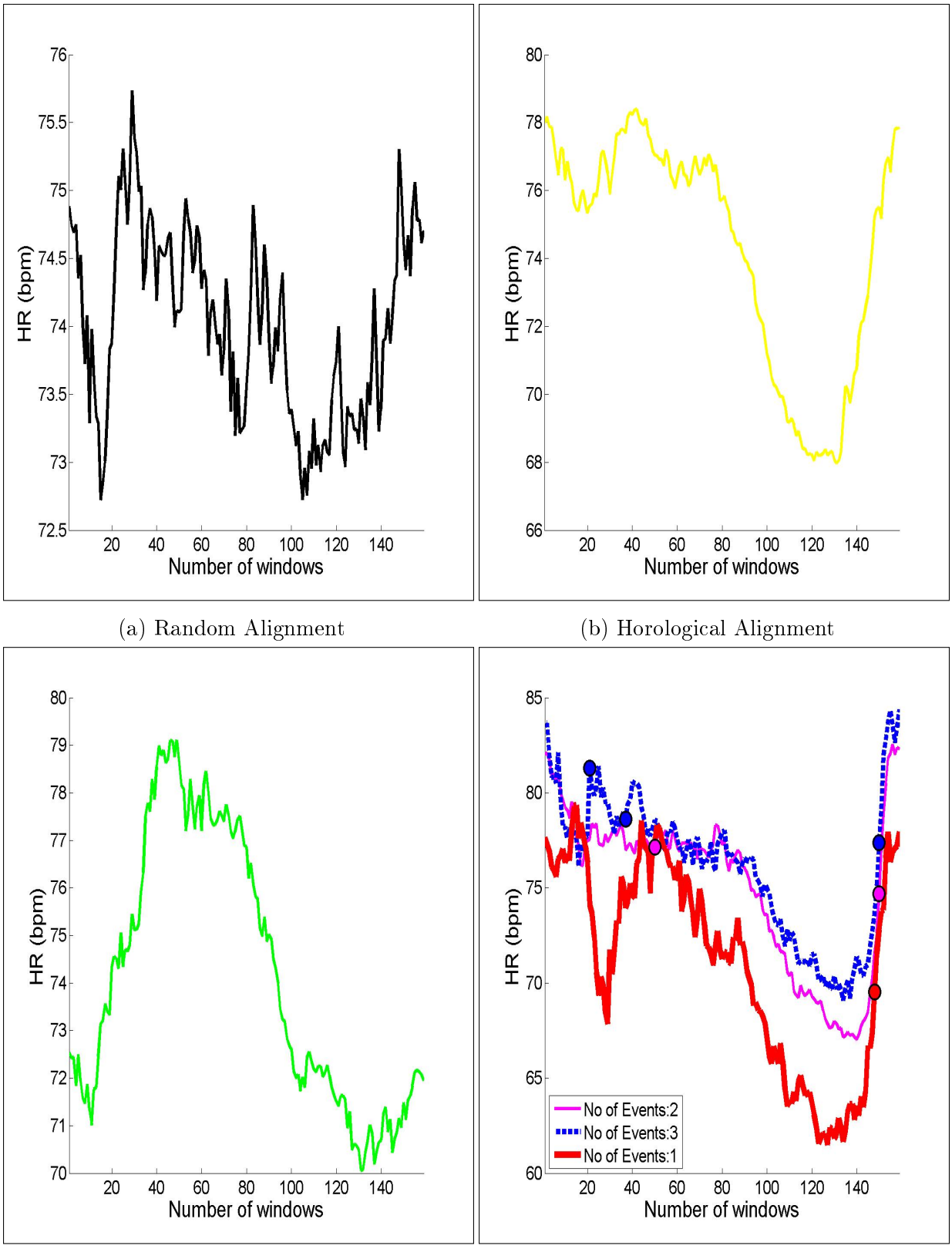
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.12: **Parameters:** Database = Ea, $T = 60$ minutes



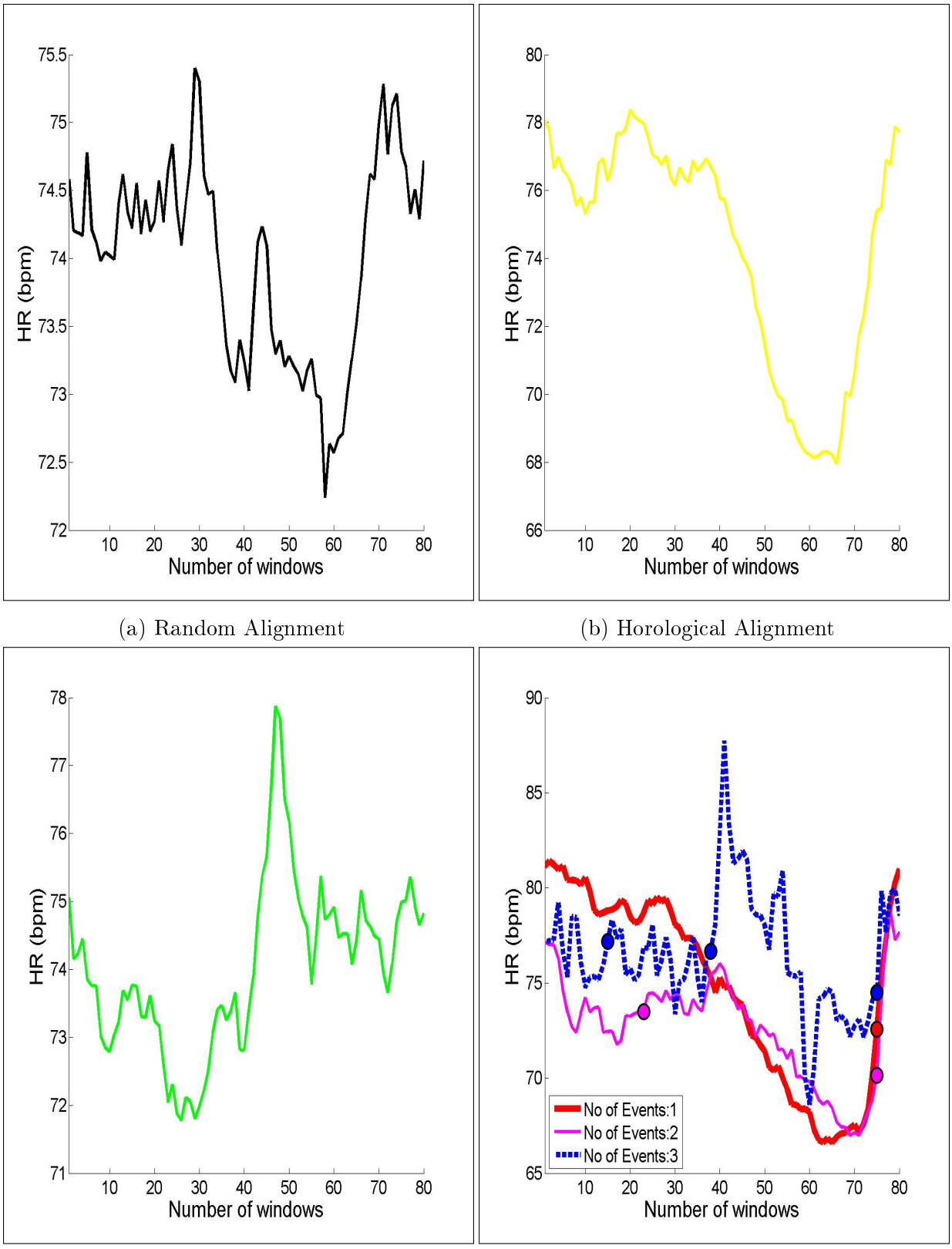
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.13: **Parameters:** Database = Eb, $T = 10$ minutes



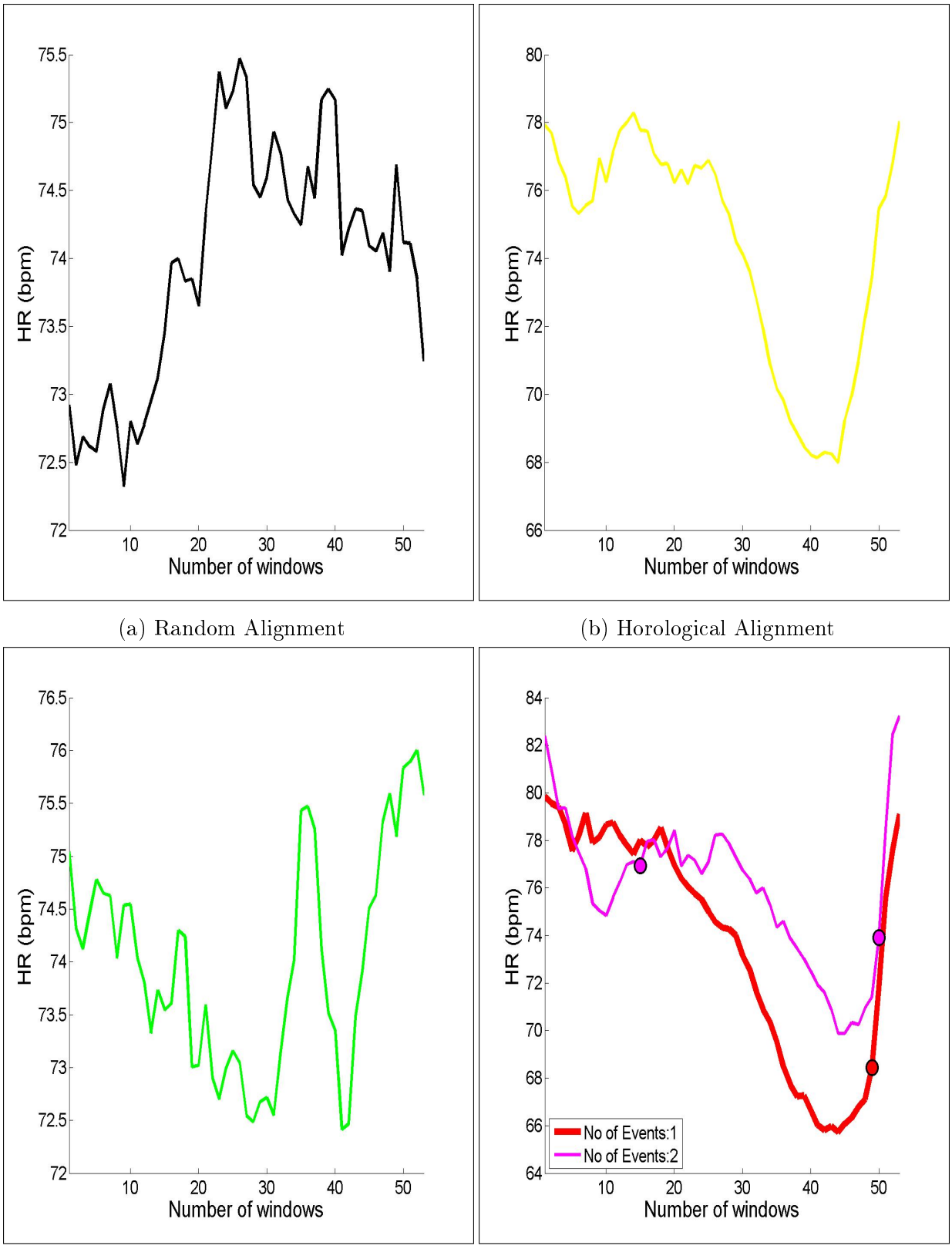
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.14: **Parameters:** Database = Eb, $T = 20$ minutes



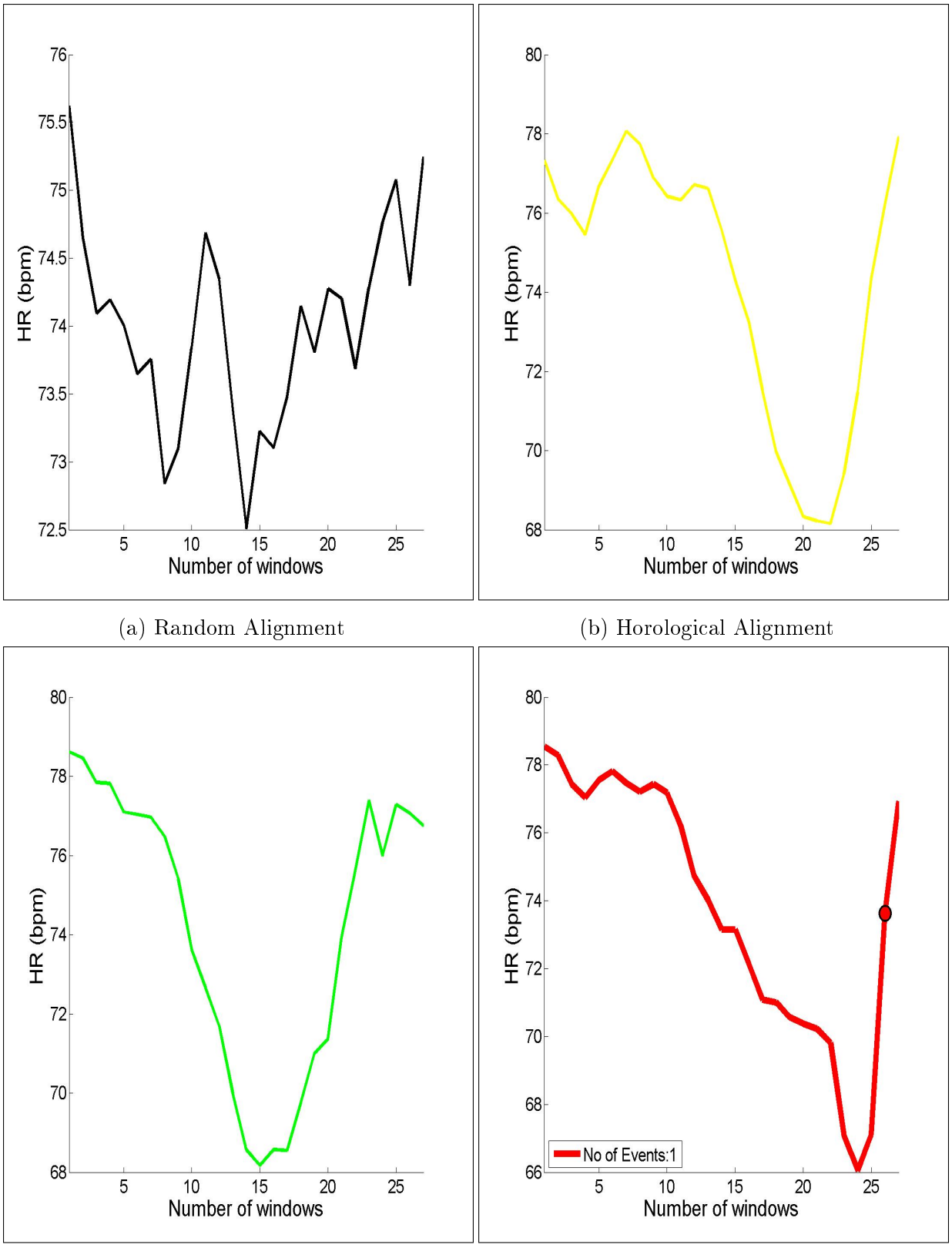
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.15: **Parameters:** Database = Eb, $T = 30$ minutes



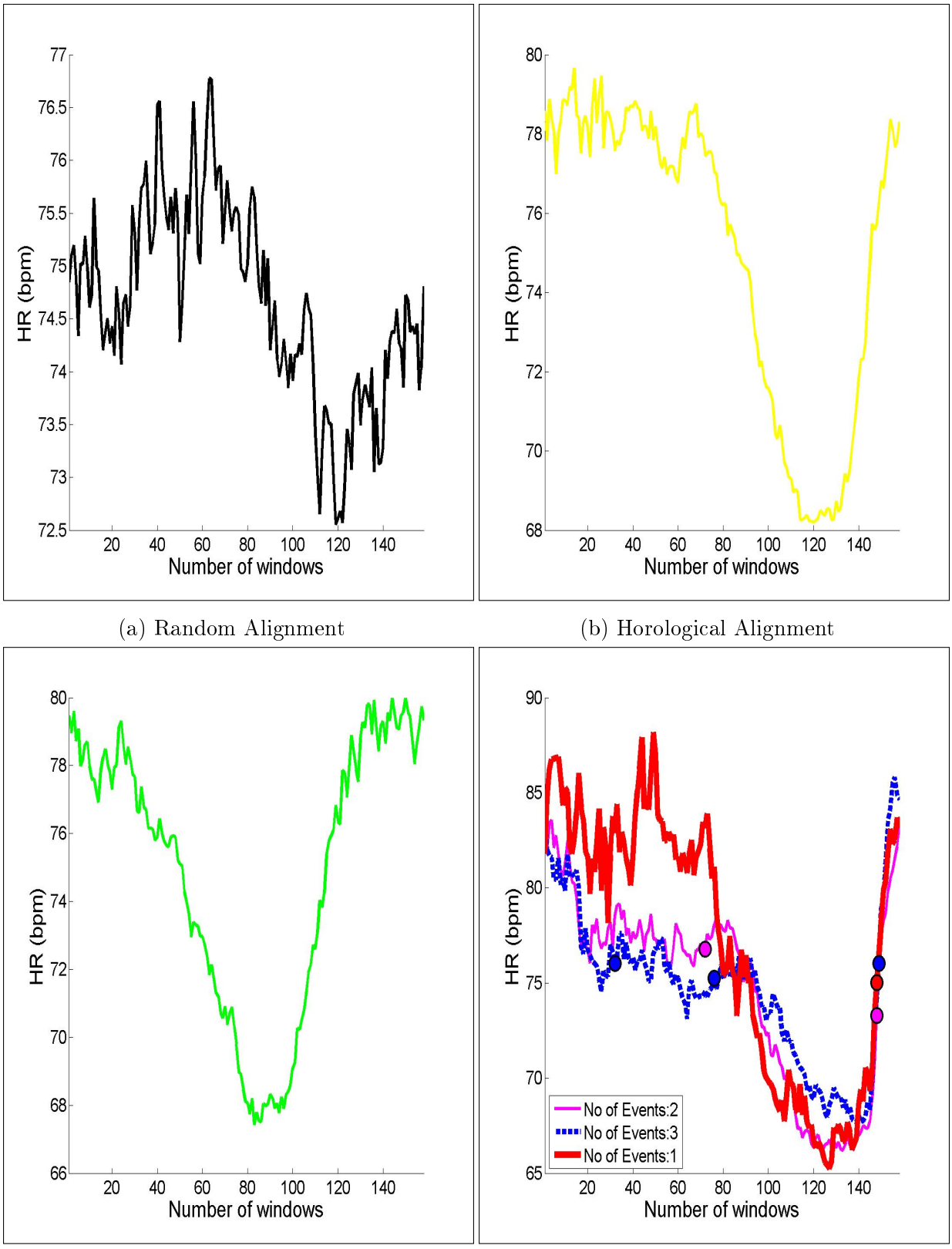
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.16: **Parameters:** Database = Eb, $T = 60$ minutes



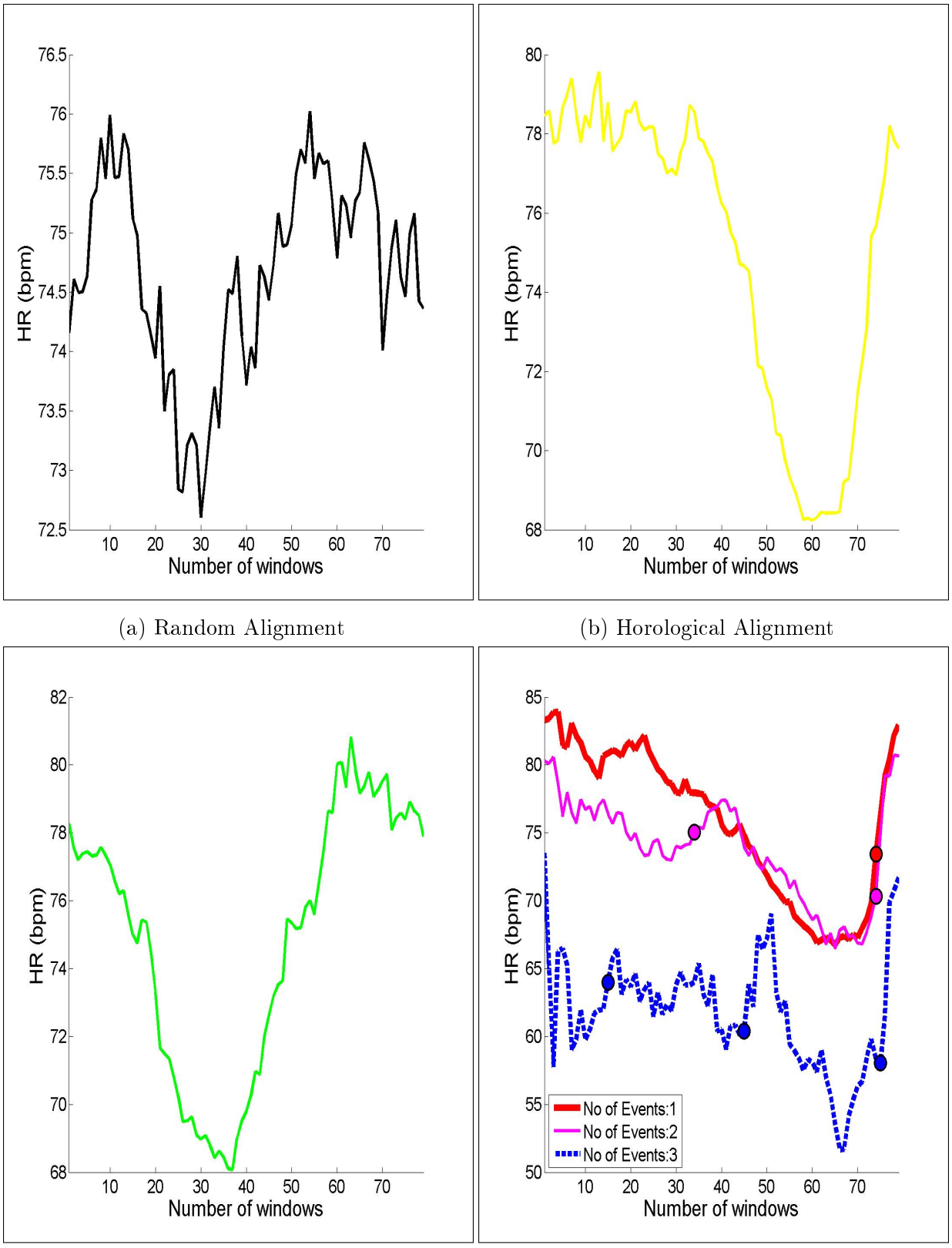
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.17: **Parameters:** Database = Fa, $T = 10$ minutes



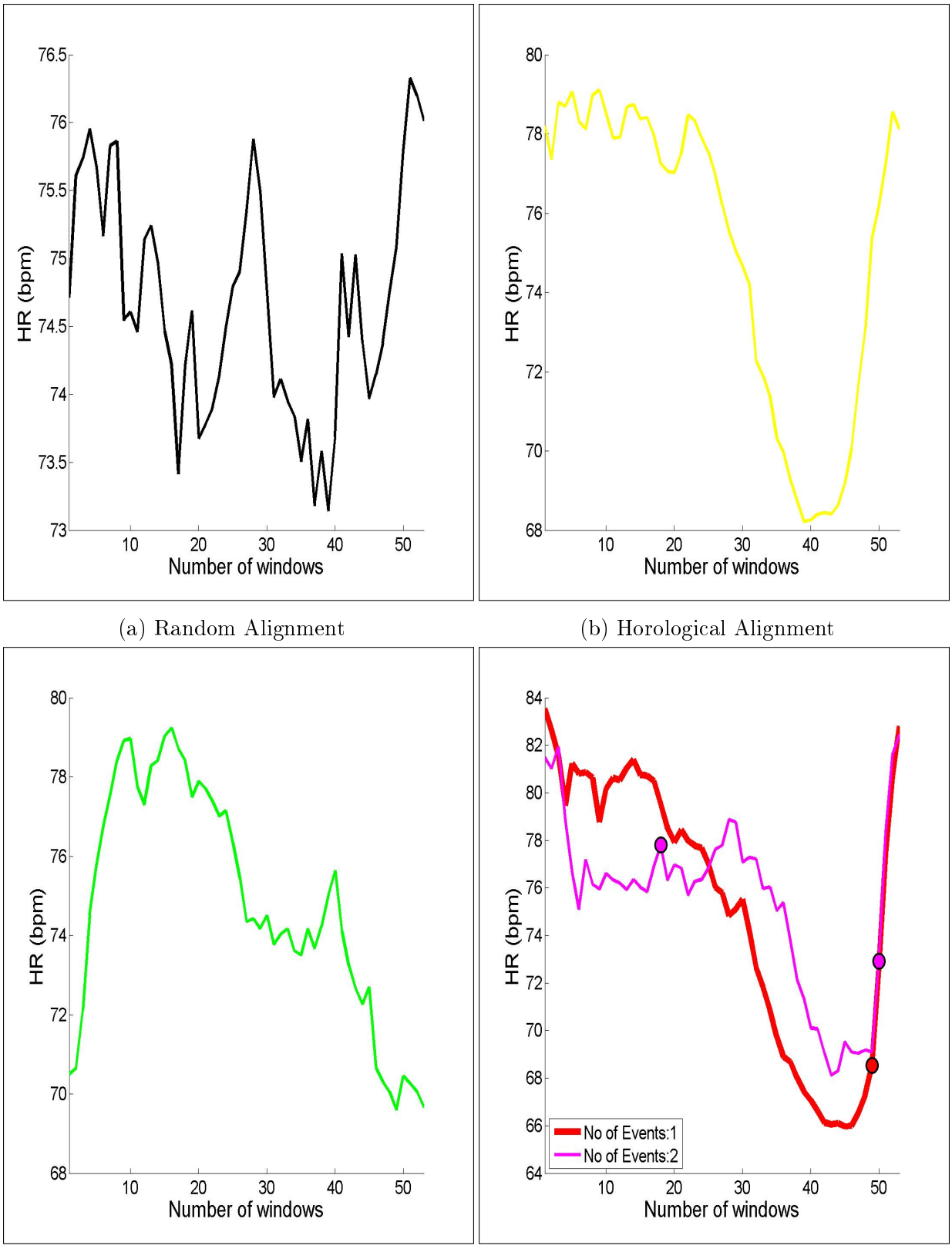
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.18: **Parameters:** Database = Fa, $T = 20$ minutes



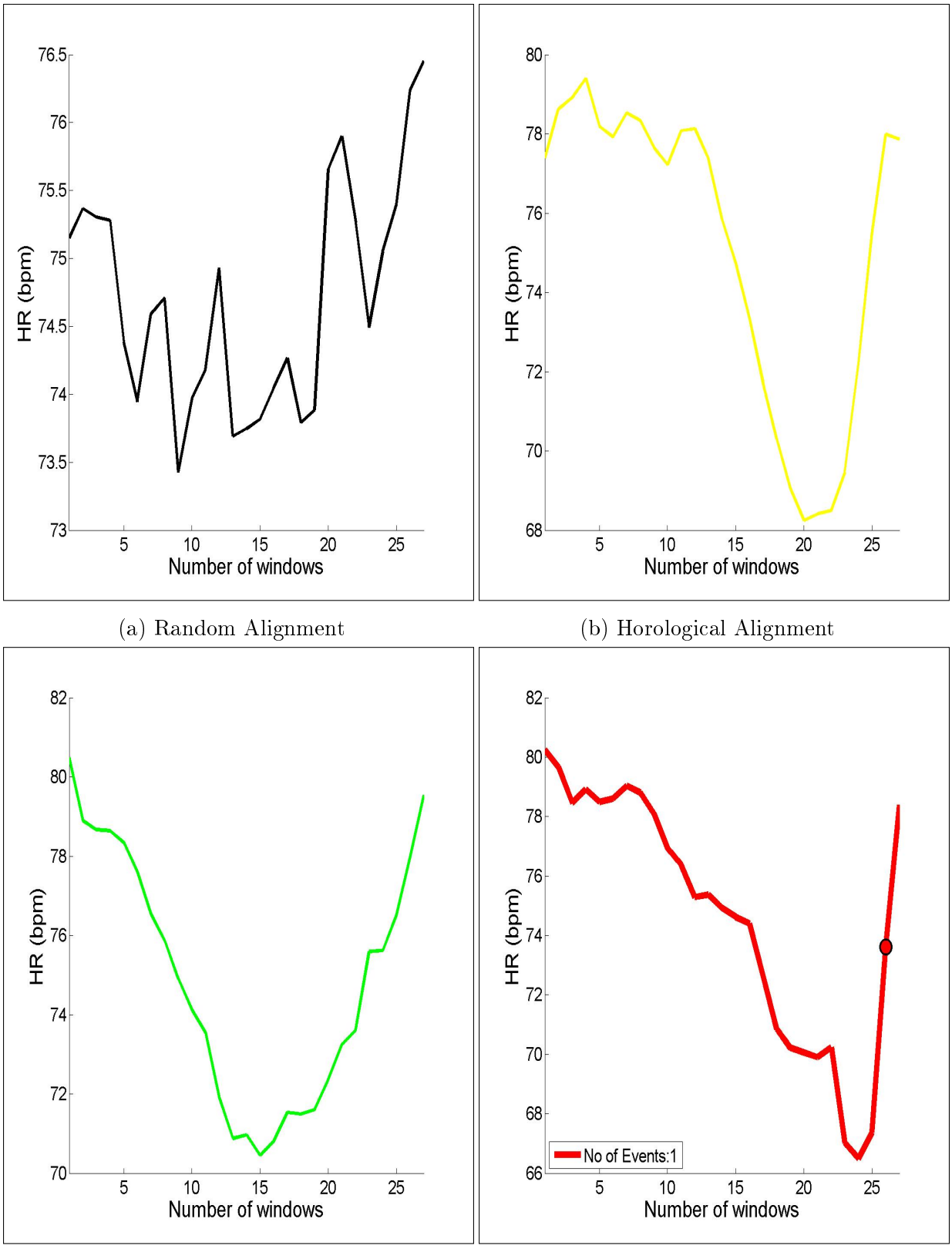
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.19: **Parameters:** Database = Fa, $T = 30$ minutes



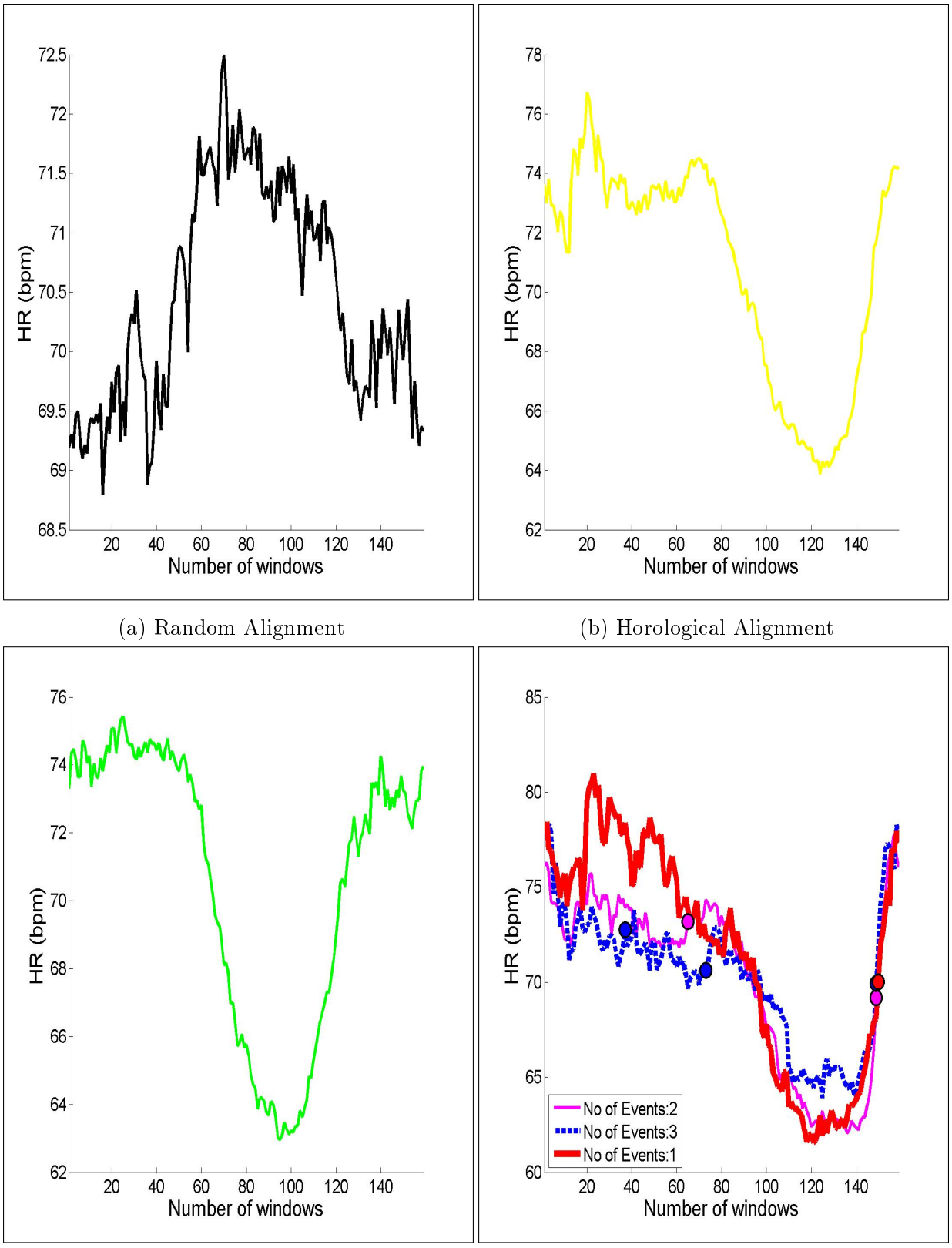
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.20: **Parameters:** Database = Fa, $T = 60$ minutes



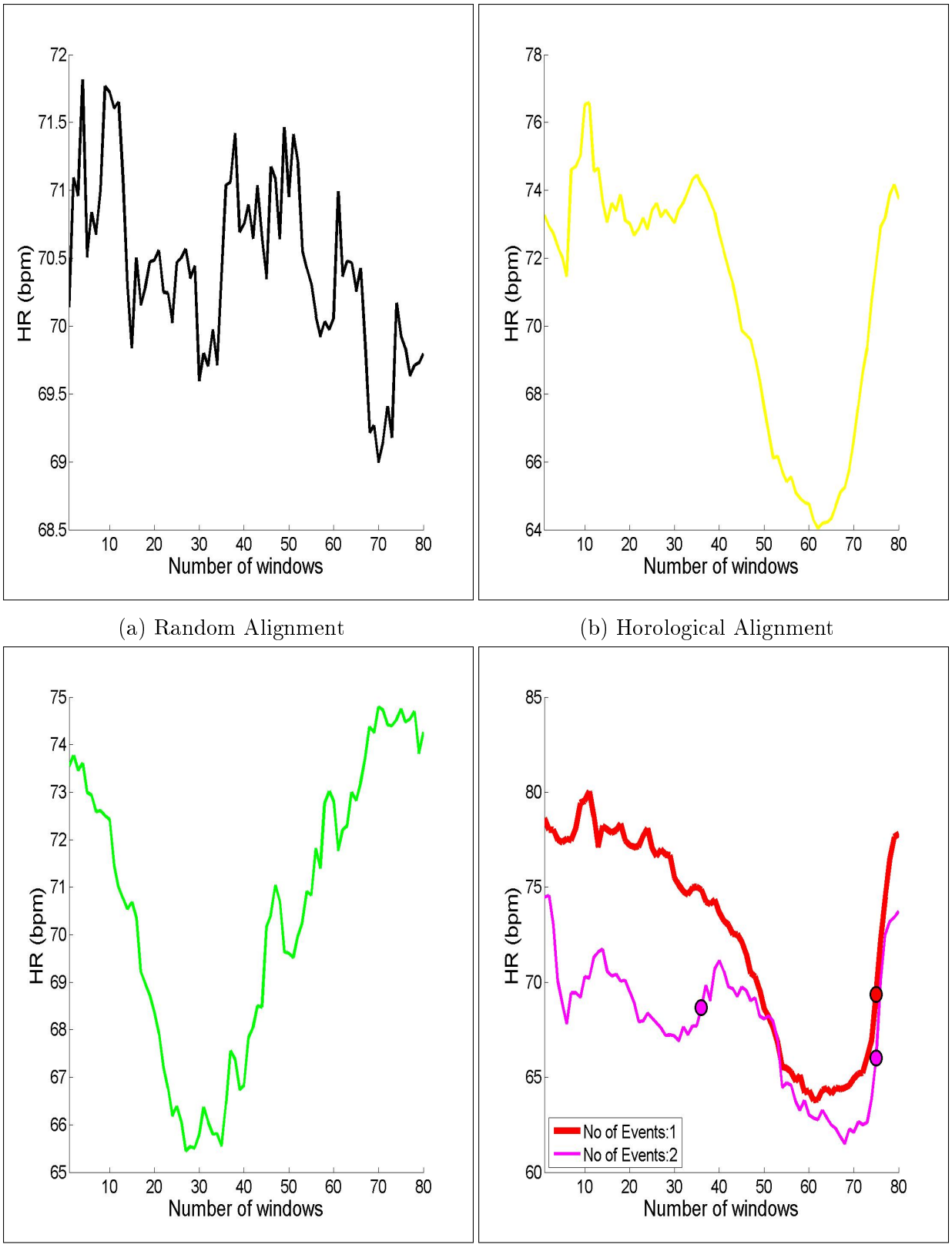
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.21: **Parameters:** Database = Fb, $T = 10$ minutes



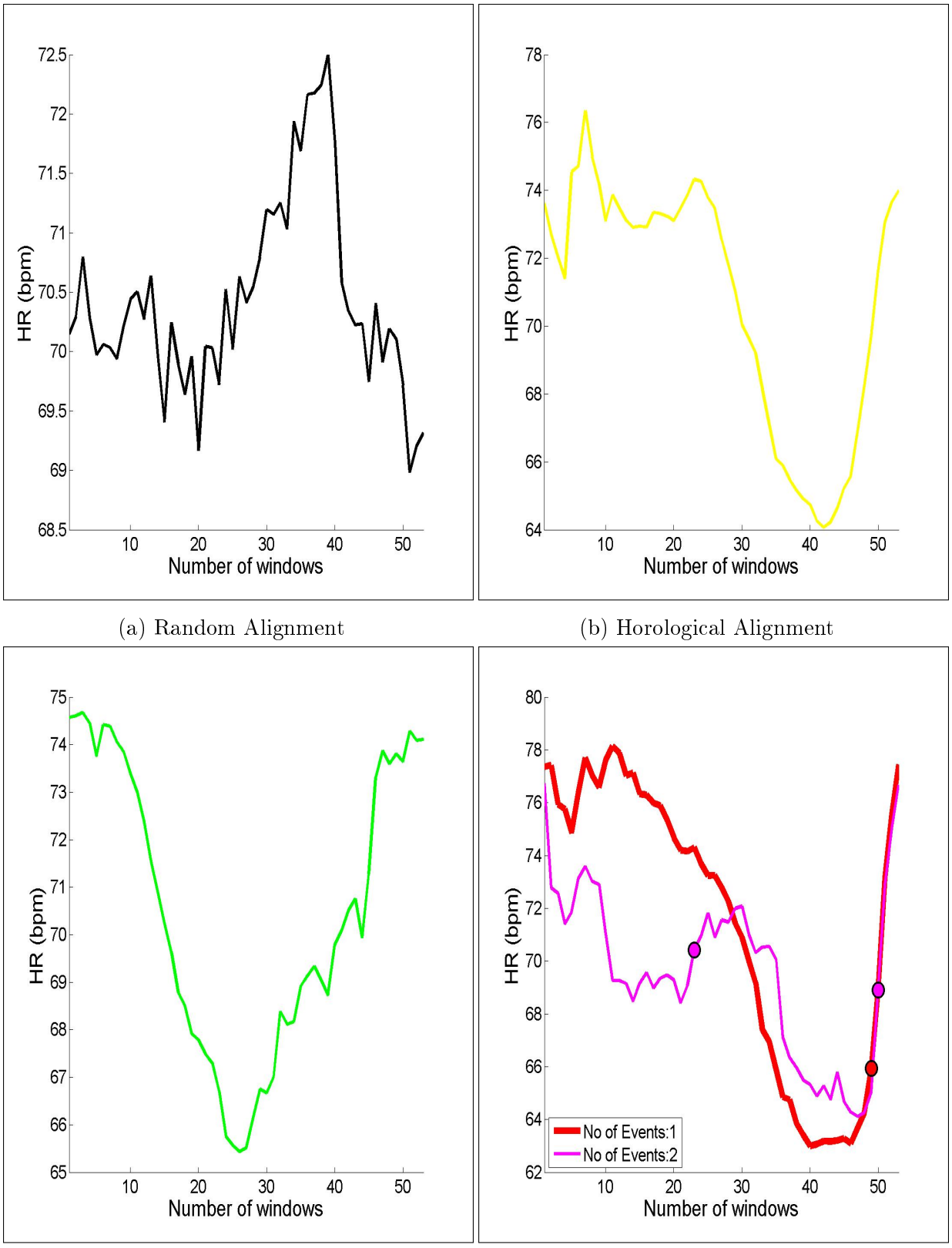
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.22: **Parameters:** Database = Fb, $T = 20$ minutes



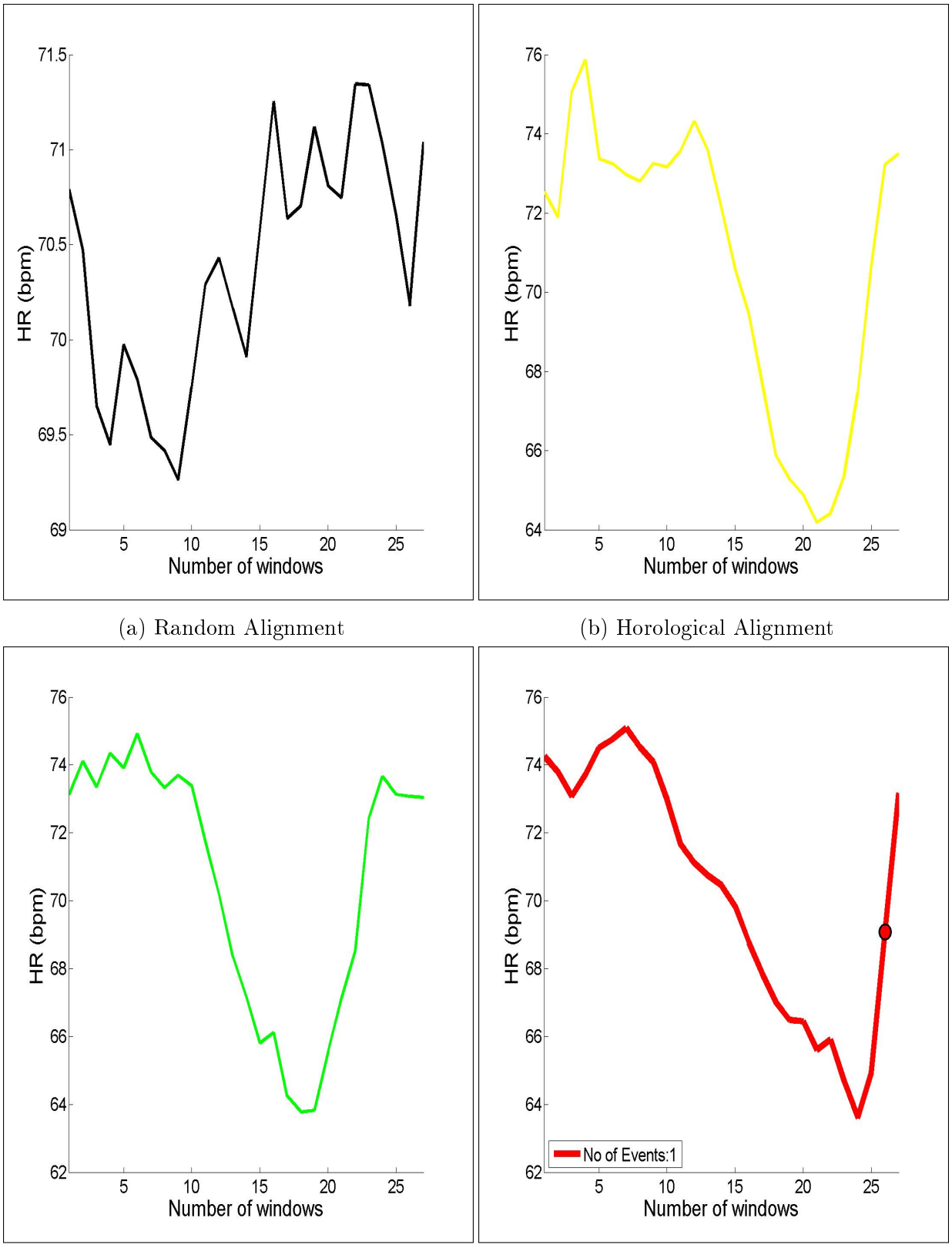
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.23: **Parameters:** Database = Fb, $T = 30$ minutes



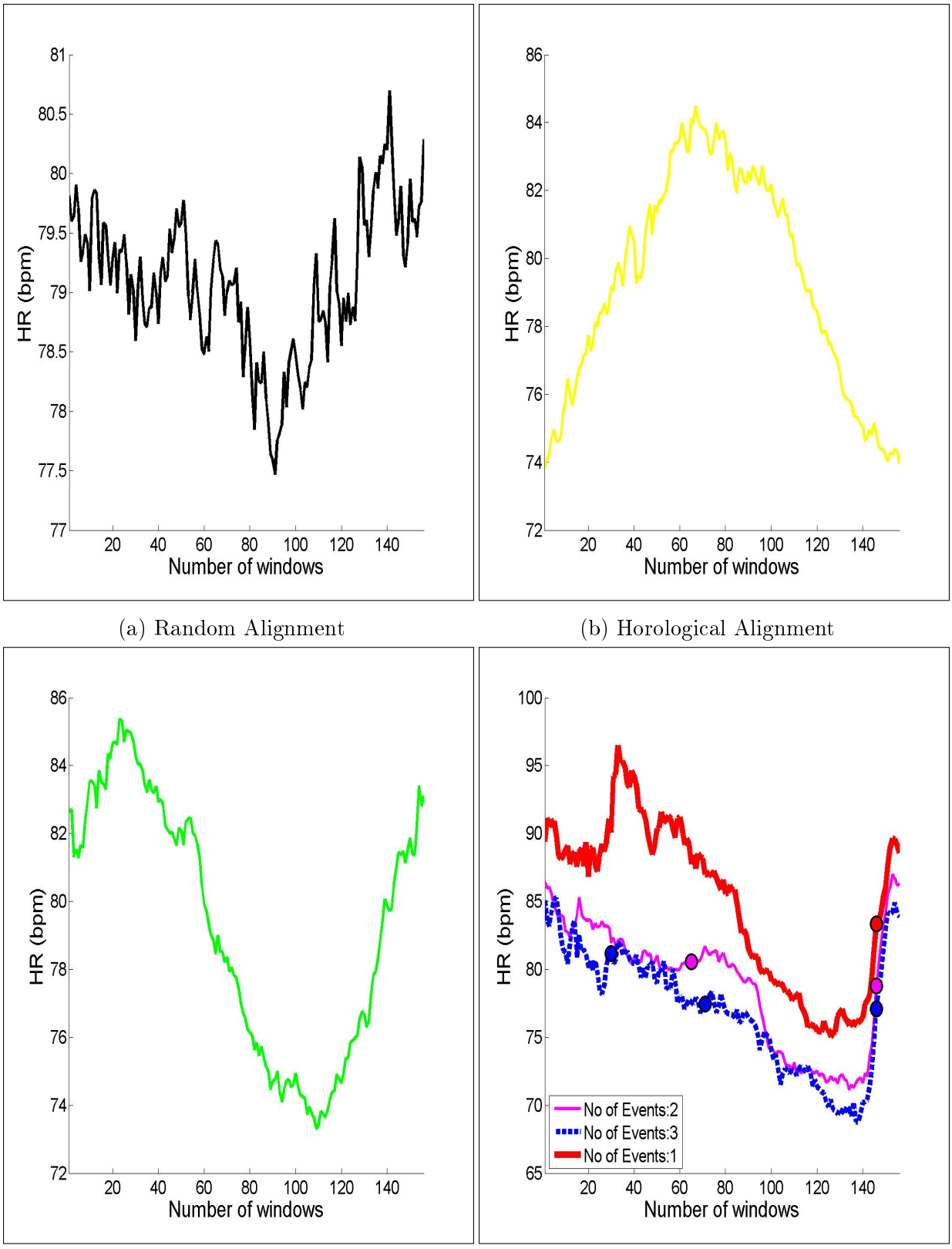
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.24: **Parameters:** Database = Fb, $T = 60$ minutes



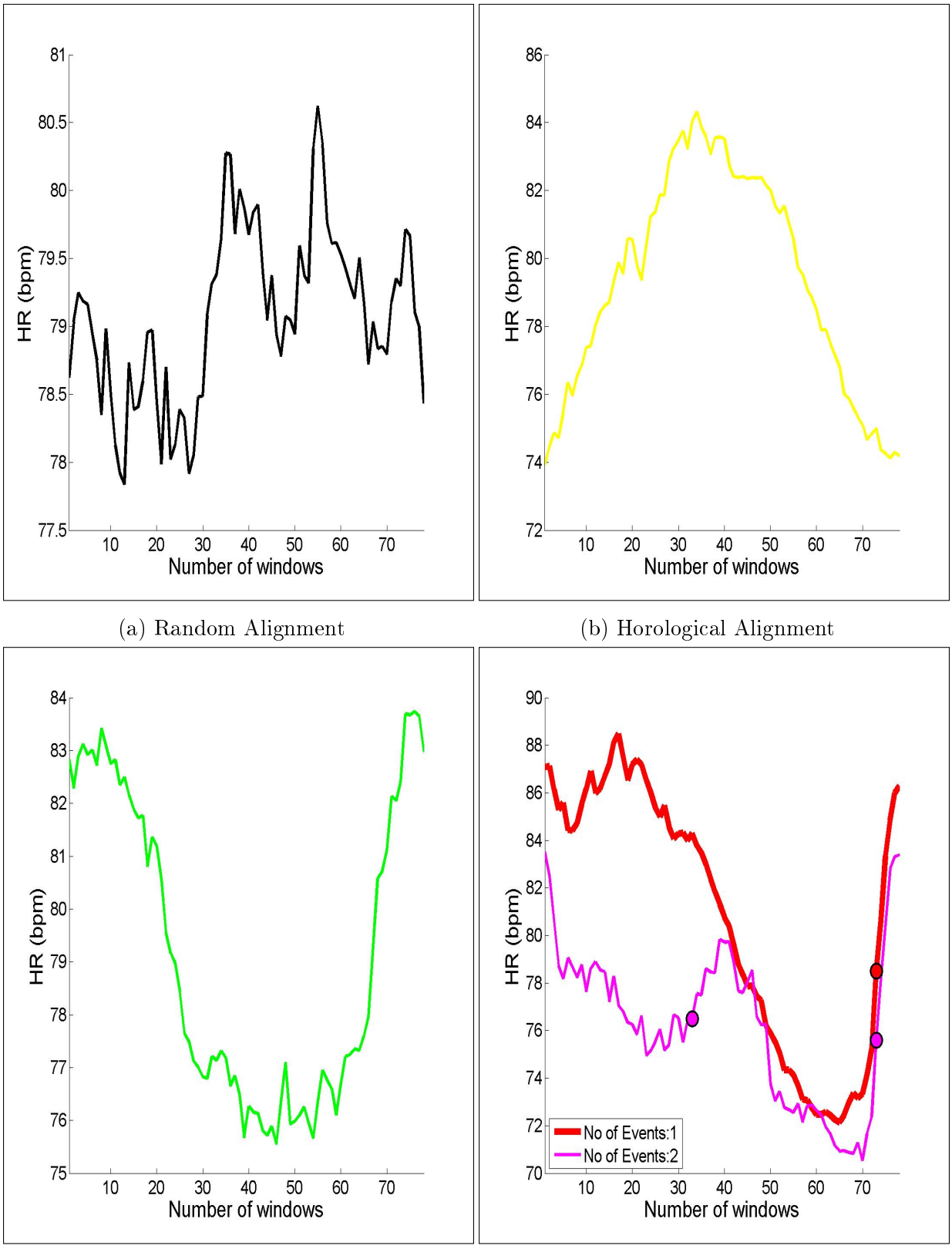
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.25: **Parameters:** Database = Ma, $T = 10$ minutes



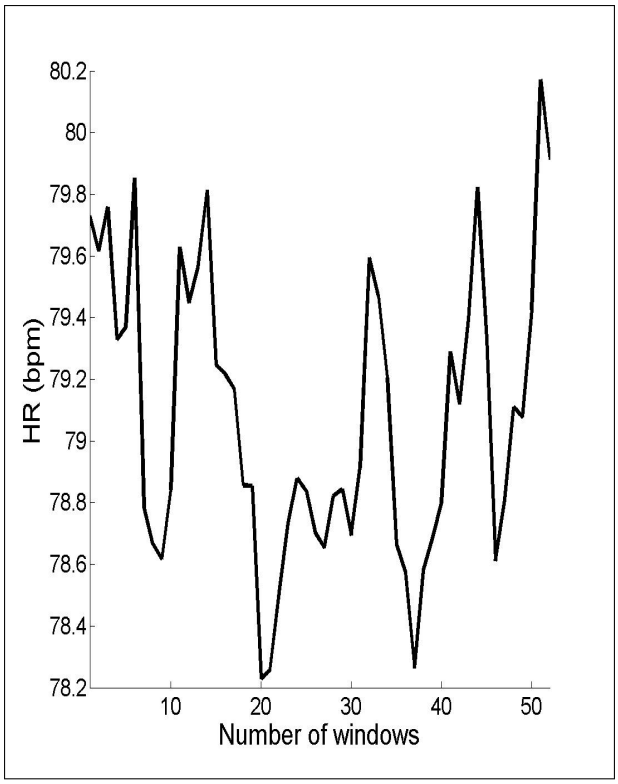
(a) Random Alignment

(b) Horological Alignment

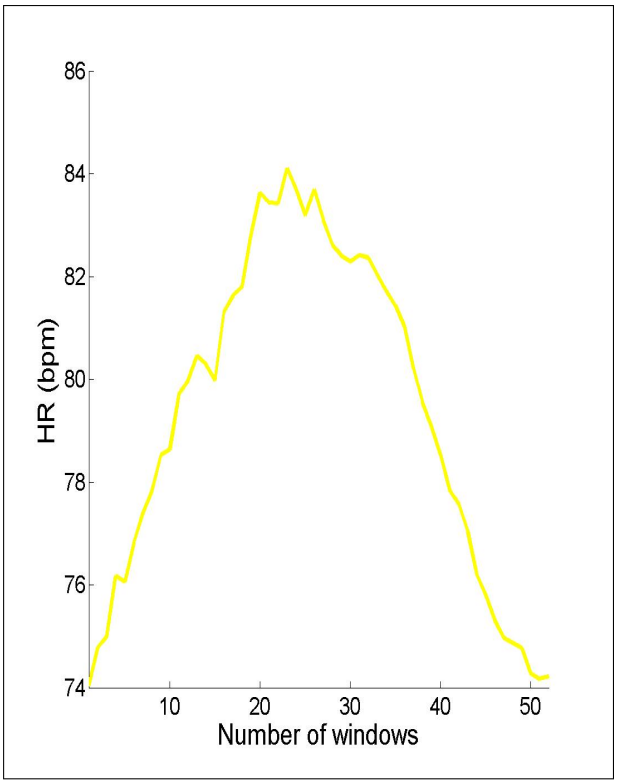
(c) PPA Alignment

(d) EBA Alignment

Figure 5.26: **Parameters:** Database = Ma, $T = 20$ minutes

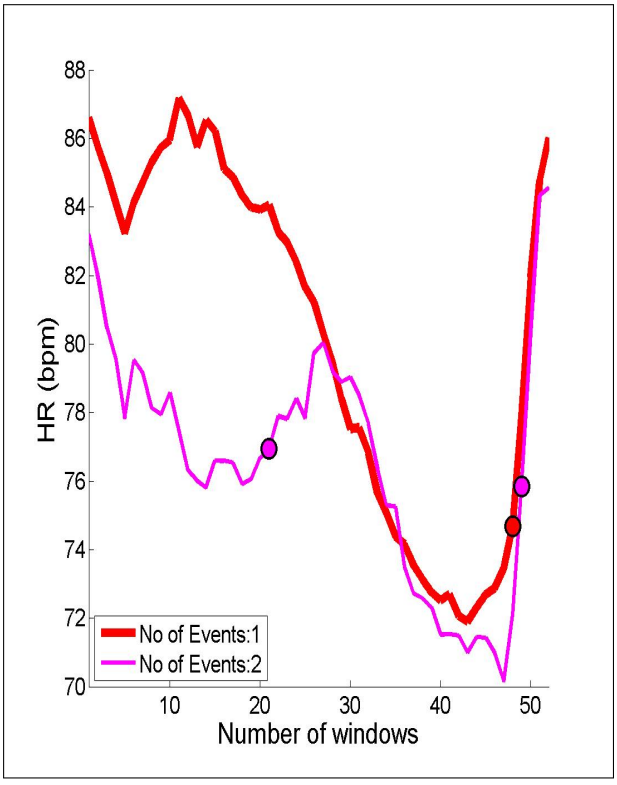


(a) Random Alignment



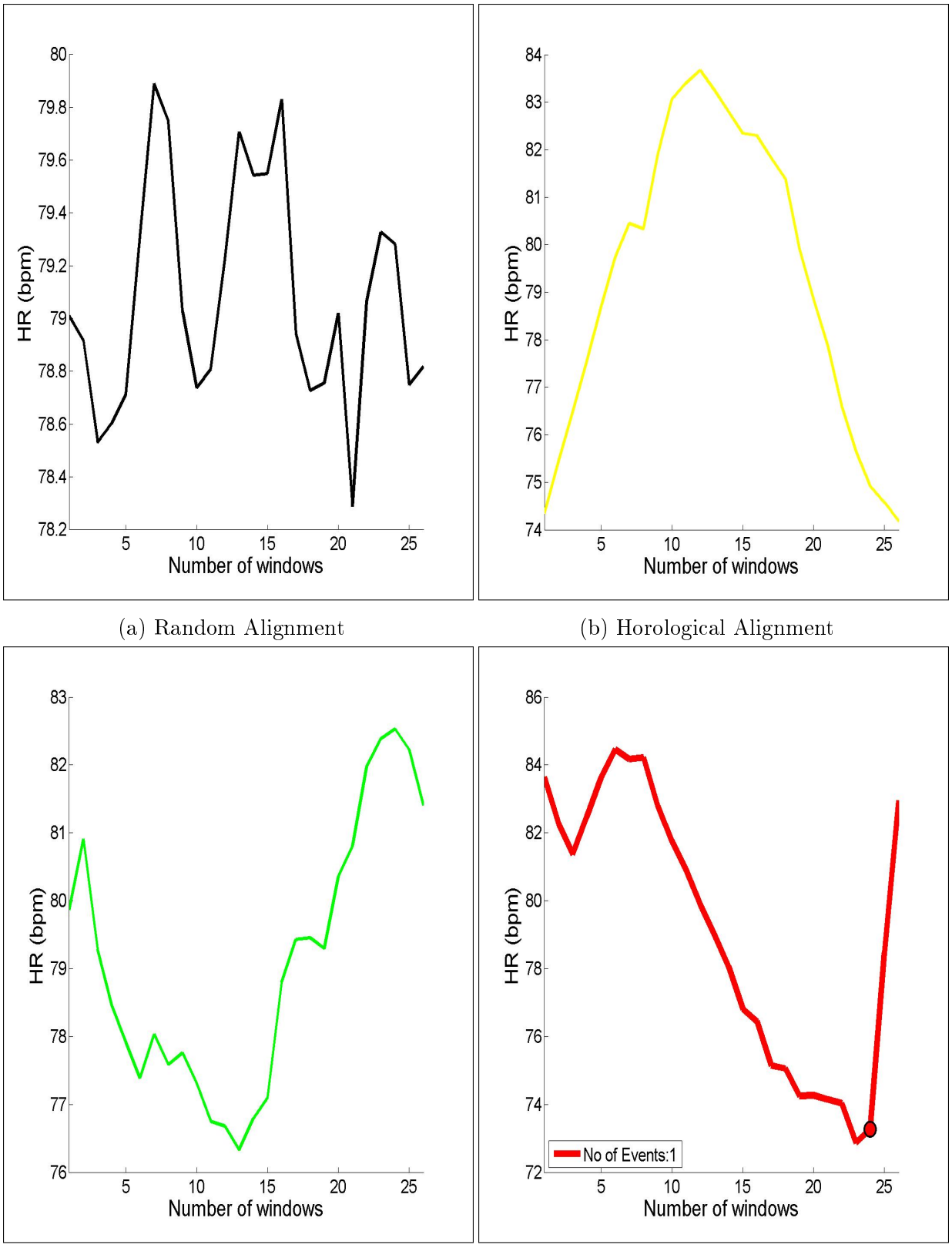
(b) Horological Alignment

(c) PPA Alignment



(d) EBA Alignment

Figure 5.27: **Parameters:** Database = Ma, $T = 30$ minutes



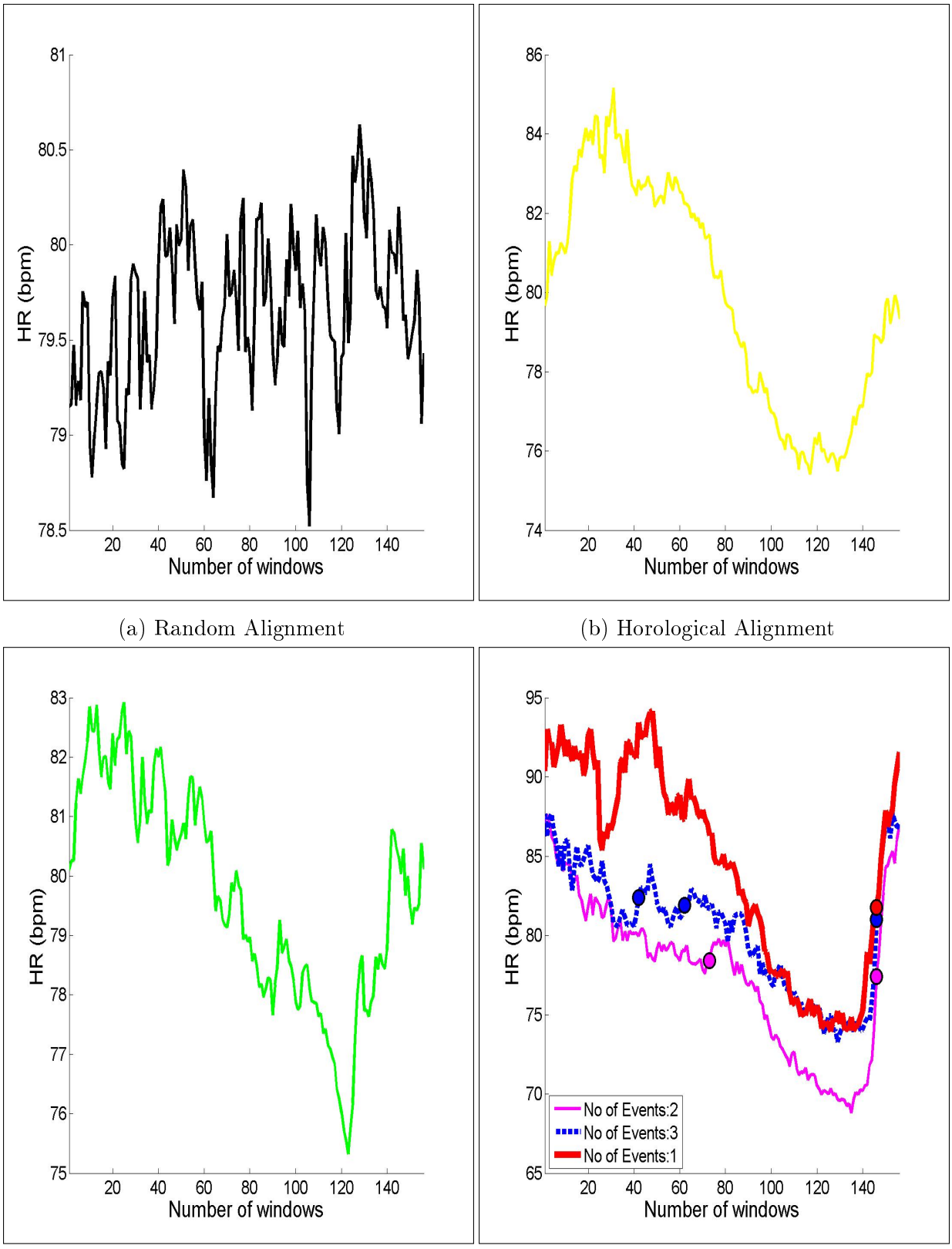
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.28: **Parameters:** Database = Ma, $T = 60$ minutes



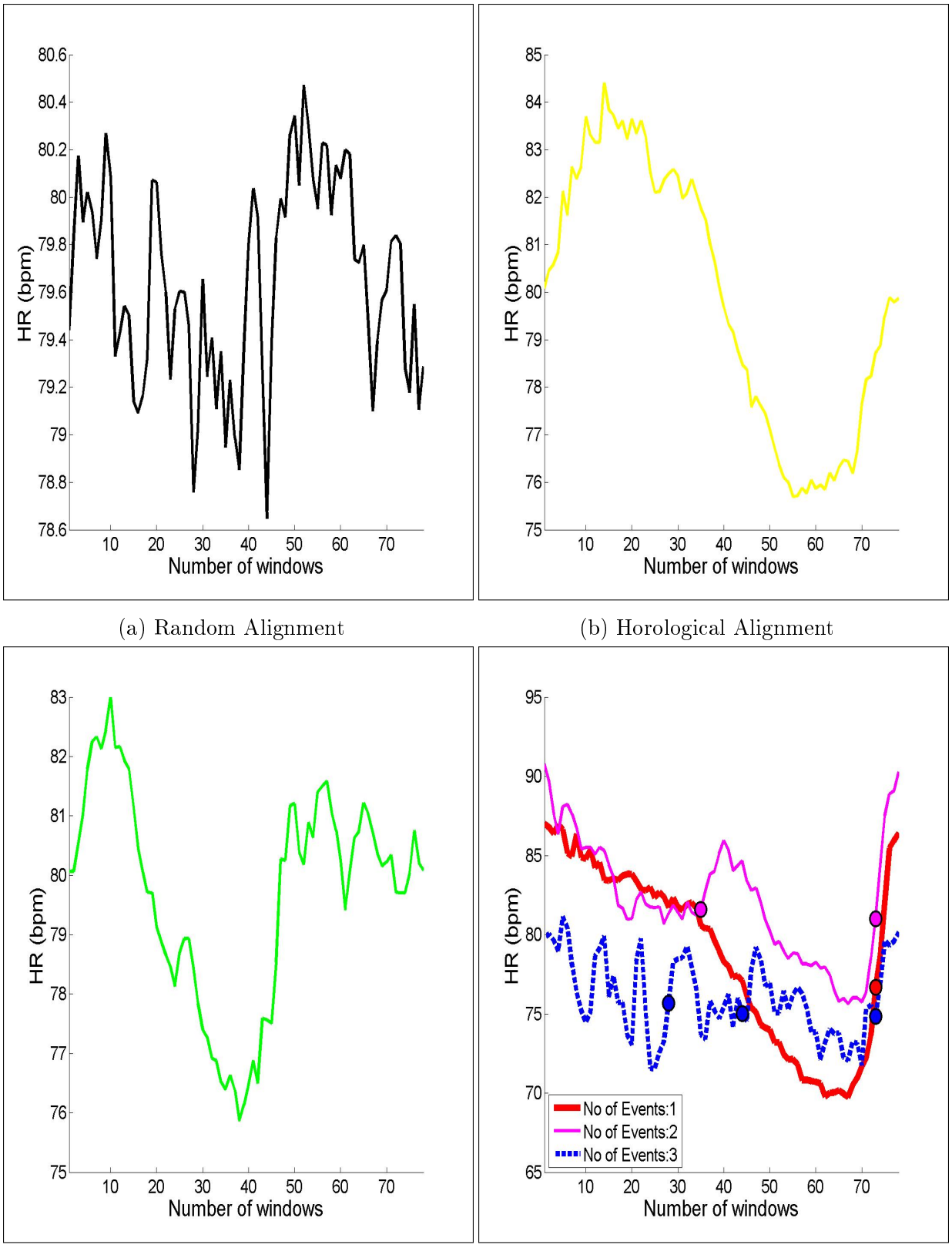
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.29: **Parameters:** Database = Mb, $T = 10$ minutes



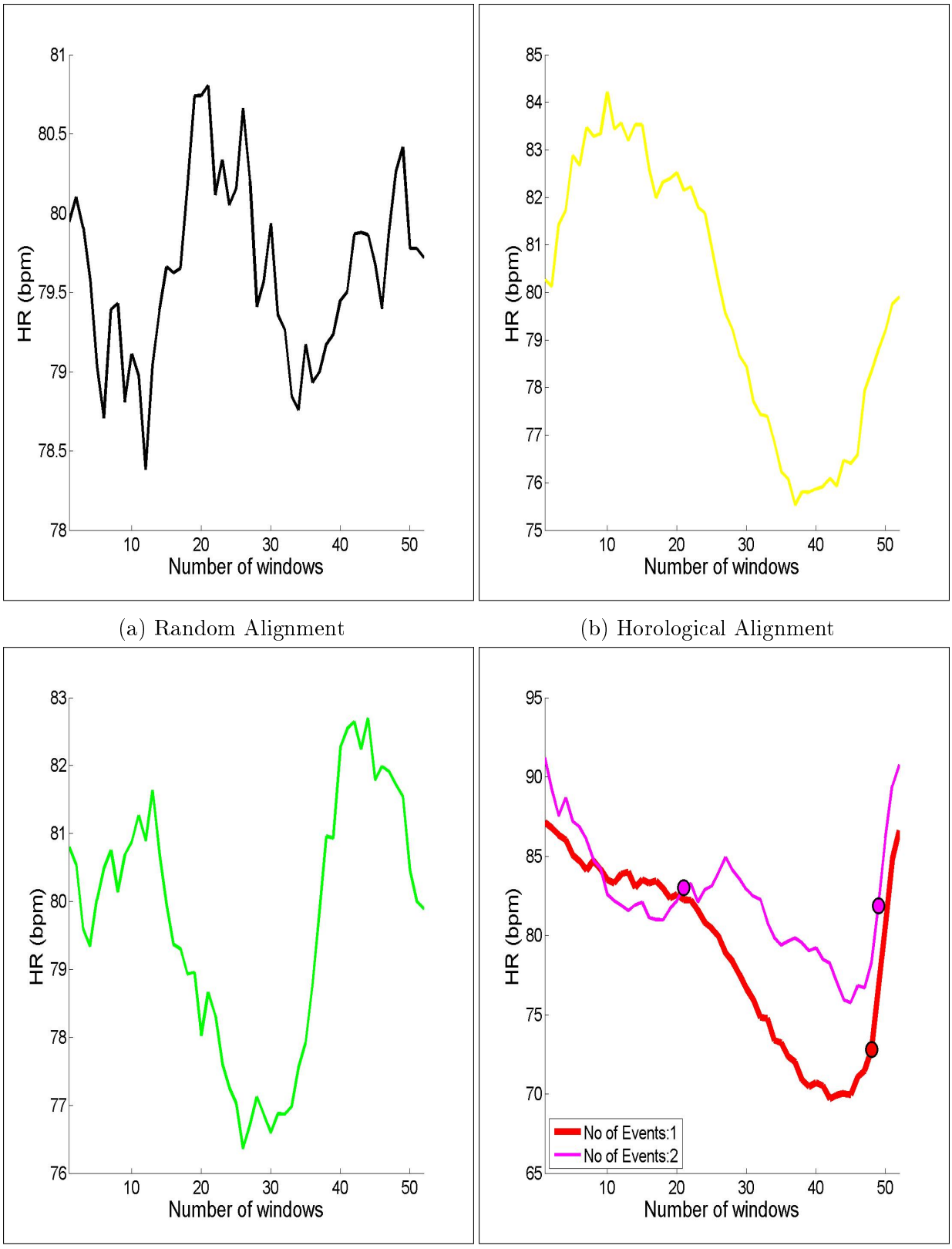
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.30: **Parameters:** Database = Mb, $T = 20$ minutes



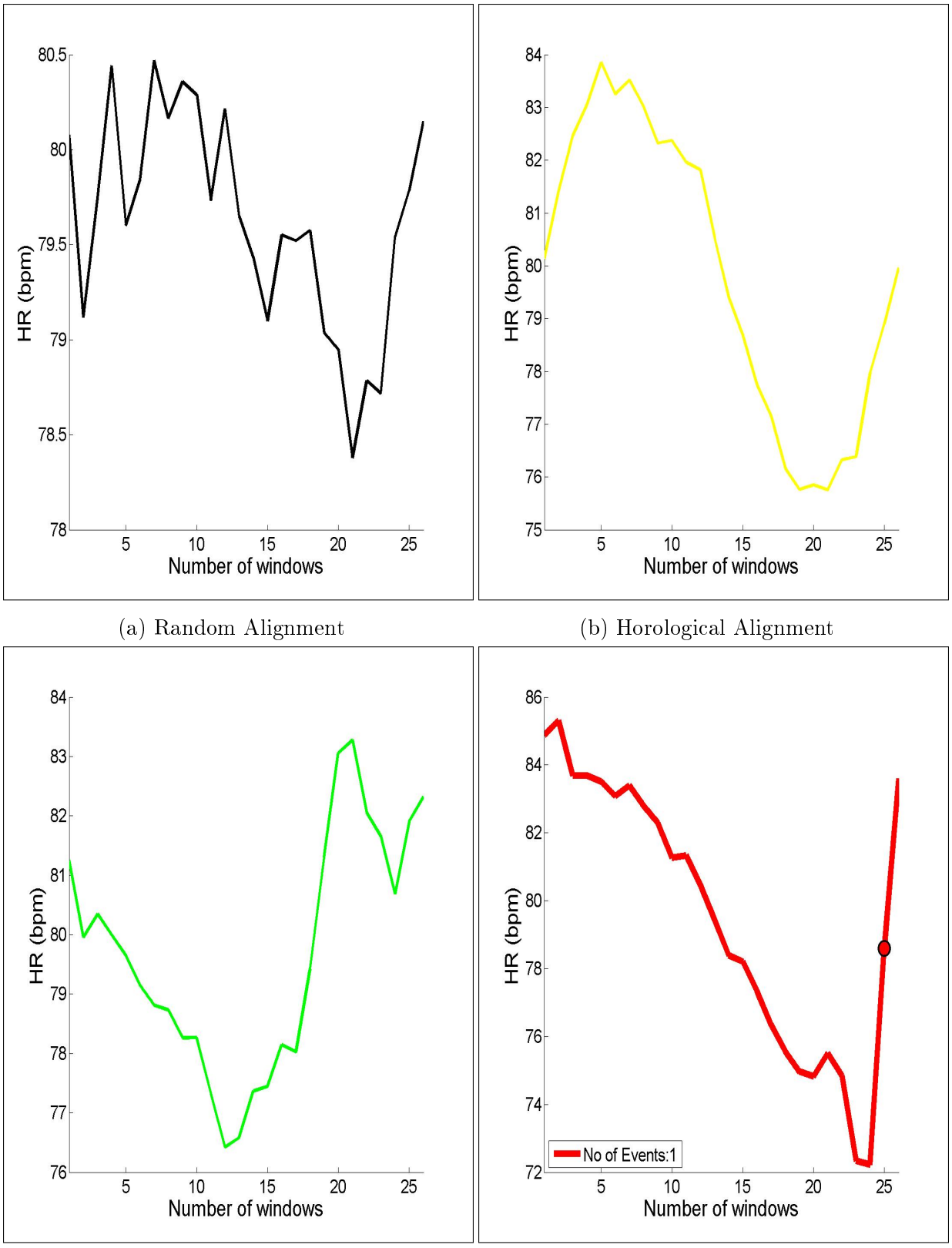
(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.31: **Parameters:** Database = Mb, $T = 30$ minutes



(a) Random Alignment

(b) Horological Alignment

(c) PPA Alignment

(d) EBA Alignment

Figure 5.32: **Parameters:** Database = Mb, $T = 60$ minutes

Author's Publications

Nathalie T.H. Gayraud and George Manis, "Alignment of R-R Interval Signals using the Circadian Heart Rate Rhythm," *37 th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2015

Marina Argyri, Nathalie T.H. Gayraud and George Manis, "Fetal Heart Rate in the Course of Delivery," *37 th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2015

Short Vita

Gayraud Nathalie Therese Helene

PERSONAL INFORMATION

- *Contact Details*
  Address: Kouremenou V. 22, Ioannina, 45333, Greece
  Telephone: +30 6932 307705
  Email: nat.gayraud@gmail.com
- *Nationalities*
  French
  Greek
- *Date of Birth*
  04/12/1986
- *Marital Status*
  Single

EDUCATION

- University of Ioannina - Computer Science Department, 2015
  M.Sc. in Computer Science, Technologies - Applications
- University of Ioannina - Computer Science Department, 2009
  Computer Science Degree with a final grade of 7,89 / 10
- Musical High School of Ioannina, 2004
  Graduated from high school with a final grade of 18,9 / 20

RESEARCH EXPERIENCE

- Heart Rate Variability, 2014
  Supervisor: George Manis
  We perform HRV analysis and try to extract various features from the signals.

- Diffusion in Evolving Social Networks, 2014
  Supervisors: Evaggelia Pitoura, Panayotis Tsaparas
  We study diffusion on time-evolving networks, extending known diffusion models to account for network evolution, and consider the problem of influence maximization on such networs.

- Implementing Routing and Wavelength Assignment Algorithms in Multifiber WDM Networks, 2009
  Supervisor: Christos Nomikos
  We implement algorithms to solve the maximum routing and wavelength assignment problem on ring networks, by reducing it to the maximum path coloring problem.

SKILLS

- Strong programming skills
  - C/C++
  - Matlab
  - Java
- Good programming skills
  - Python
  - Parallel programming
  - Open MP
  - MPI
- Experience in the following areas
  - Database Management
  - Algorithms
  - Compilers
  - Web design
  - User interface design

LANGUAGES

- French - Fluent (Native language)
- Greek - Fluent (Native language)
- English - Fluent (Certified by the university of Cambridge)
- Italian - Fluent (Certified by the university of Perugia)
- Spanish - Good
- Portuguese - Basic

TEACHING EXPERIENCE

- IT Teacher, Pronomio Pliroforiki, 2010 – 2012
  Owner / Teacher at the IT school βἐώPronomioβἐώ.
  Taught the following courses to candidates of all ages:
  Application development
  ECDL preparation courses
  Programming languages
- English Teacher, 2005 – 2012
  Teaching English to children to prepare them for the First Certificate (Lower) of Cambridge

PROFESSIONAL EXPERIENCE

- Website design, Freelance, 2012 – 2014
  Dynamic and static website design using CSS, HTML, Javascript and various other tools and languages.
- Senior Cashier, Kotsovolos Dixonβἐΰs Retail, 2012 – 2013
  Loan approvals for purchases with installments
- AVS – Special needs assistant – School aid, 2010 – 2012
  French term : βἐώAuxiliaire de vie scolaireβἐώ. I catered to a young boy diagnosed with trisomy 21, who was going to elementary school, during a period of three years, to improve his inclusion in a βἐώnormalβἐώ scholar environment and society in general.
- Sales Manager, Kotsovolos Dixonβἐΰs Retail, 2005 – 2010
  Top salesman in the computing/telecoms department
  Computer software & expert technician

CERTIFICATES

- The Hellenic Association of Continuous Education, 2010
  (Teaching and catering to children with) Specific learning disabilities - Dyslexia
- Kotsovolos Academy, 2010
  Stepping into Management

DISTINCTIONS

- University of Ioannina – Computer Science Department, 2012
  Received an Honorary Scholarship from the Academy of Athens (Year 2012)
- University of Ioannina – Computer Science Department, 2004
  Received a Best Performance Honorary Scholarship from the State Scholarships Foundation
- Musical High School of Ioannina, 1998 – 2004
  Received six High School Best Performance Honorary distinctions – Excellency Awards for remarkable grades.

OTHER INTERESTS

- Music
  Piano (20 years)
- Sports
  Capoeira