A Novel Adaptive Noise Elimination Algorithm in Long RR Interval Sequences for Heart Rate Variability Analysis

Vytautas Stankus [1,2,*], Petras Navickas [2,3], Anžela Slušnienė [2], Ieva Laucevičienė [4], Albinas Stankus [2,†] and Aleksandras Laucevičius [2,3]

[1] Department of Physics, Kaunas University of Technology, 44249 Kaunas, Lithuania
[2] State Research Institute Centre for Innovative Medicine, 08410 Vilnius, Lithuania
[3] Clinic of Cardiac and Vascular Diseases, Faculty of Medicine, Vilnius University, 03101 Vilnius, Lithuania
[4] Department of Rehabilitation, Physical and Sports Medicine, Faculty of Medicine, Vilnius University, 03101 Vilnius, Lithuania
[*] Correspondence: vytautas.stankus@ktu.lt; Tel.: +370-610-33946
[†] Deceased.

Abstract: As heart rate variability (HRV) studies become more and more prevalent in clinical practice, one of the most common and significant causes of errors is associated with distorted RR interval (RRI) data acquisition. The nature of such artifacts can be both mechanical as well as software based. Various currently used noise elimination in RRI sequences methods use filtering algorithms that eliminate artifacts without taking into account the fact that the whole RRI sequence time cannot be shortened or lengthened. Keeping that in mind, we aimed to develop an artifacts elimination algorithm suited to long-term (hours or days) sequences that does not affect the overall structure of the RRI sequence and does not alter the duration of data registration. An original adaptive smart time series step-by-step analysis and statistical verification methods were used. The adaptive algorithm was designed to maximize the reconstruction of the heart-rate structure and is suitable for use, especially in polygraphy. The authors submit the scheme and program for use.

Keywords: artifact; noise; elimination; heart rate variability; RRI; R-R interval; algorithm

1. Introduction

Heart rate variability (HRV) studies are based on the instantaneous heart rate time series analysis that uses beat-to-beat RR interval (RRI) data derived from an electrocardiographic (ECG) signal. Over the years, research on heart rate and its variability has gained a lot of traction and has spread not only to medicine but to fields of sports, ergonomics, and personal healthcare as well [1]. As it became widely available to every person who cares about his body's state and health, various time series analysis methods have been proposed. However, in all cases, the accuracy of the results is highly dependent on artifacts occurring in the time sequence [2,3]. These ECG artifacts inevitably originate due to both internal physiological processes that influence the normal sinus heart rhythm (e. g. respiration) [4], as well as external determinants caused by body movements or power line interference, external electromagnetic fields, and hardware/software failures [5], all of which are called "noise" or "artifacts". These can be classified as base-line wander [6–8], power-line interference [8–10], muscle artifacts [7,8,11], and channel noise [12]. Although hardware and software development has made great strides in proposing a wide range of solutions and tools for reducing the number of ECG artifacts, the problem of RRI sequence quality remains relevant to this day [13–15], since artifacts representing less than 0.1% of the overall record duration may cause variations of up to 50% in some HRV metrics [16].

To improve the quality of the RRI data, researchers have two options to eliminate artifacts. First, to improve the overall signal quality of an ECG signal and improve the accuracy of the ECG QRS complex recognition during the investigation [17] or the second—to identify artifacts in real time or retrospectively within the recorded RRI sequences [18]. Different methods are used for each of the aforementioned options, as the ECG and RRI signals are different in nature: the ECG signal is continuous and its amplitude is measured in volts, while the RRI signal is discrete in nature and its amplitude is measured in seconds. In addition, when the ECG is converted to digital form, the ECG time axis step is constant, while the RRI time axis step is variable. Consequently, this results in different methods being used to process these signals. Overall ECG artifact identification and removal are made possible by various algorithms and are closely dependent on the aim [19]: a data-driven mechanism of empirical mode decomposition [20,21], deep-learning-based models [22], wavelet-based models [23], and sparsity-based, Bayesian-filter-based, and Hybrid models [24]. The majority of them focus on short-duration artifacts, which are often treated in the same way as ectopic beats [25–27]. The proposed gap-filling methods are based on an integrated pulse frequency modulation model [25], a rejection filter on the basis that nonpathological artifacts are of small duration and large amplitude [28], based on the tachogram using cubic splines [29].

However, the previous methods change the overall duration of the recording, are only suitable for isolated artifacts, and have not been evaluated for longer artifact segments. There are algorithms designed for longer gap filling, some of which are using Gaussian-distribution-based methods [30]; unfortunately, they are less common and are not designed to retain the overall recording duration.

Another very common method of correction is based on the RRI signal value threshold, when RRI values appear beyond the impossible limits; for example, 350 ms and 1715 ms, respectively [31]. After removing the identified erroneous values, they are replaced by interpolated data. The resulting time sequence is usually depicted as amplitude on the abscissa axis, while the ordinary axis represents frequency or real time. This depiction of the discrete signal enforces the investigators to apply common amplitude artifacts removing algorithms to RRI sequences, despite the inadequacy of its nature, since various filters dominate between them. Therefore, applying it enables us to adequately eliminate a small number of artifacts and thus reduce the overall sequence dispersion and its distribution in the frequency axis. However, it ought to be emphasized that even though the entire duration of the RRI sequence in such a situation becomes slightly shortened or lengthened, the removal of artifacts has a low impact on the time domain and frequency domain HRV metrics [31].

To summarize, there are algorithms for ectopic beats correction [25,32–35], restoration of missing heartbeats [36], and noise elimination [37]; however, an entirely different picture is seen in which heart rate registrations occur alongside the other processes (physiological and non-physiological). In this instance, it is necessary to maintain a precise synchronization of each RRI with other measurable processes. Therefore, the use of a traditional filter becomes inadequate for the expected results of the analysis, especially when looking for associations between the heart rate and other internal (electromyography, electroencephalography, and others) as well as external factors [18]. It is well acknowledged that "accurate R–R interval artifact correction and editing methods are needed" [38], especially when RRI sequences are very long, and it is complicated to apply manual methods. Consequently, after encountering this issue when dealing with multidimensional processes, we aimed to propose an original RRI sequence artifact elimination algorithm suited for long-term (hours or days) sequences that would overcome the aforementioned limitation and would preserve the overall RRI sequence temporal structure in real-time recordings.

2. Materials and Methods

2.1. Materials

The proposed algorithm was evaluated using long RRI time-sequences collected during monitoring with a 3-channel ECG Holter monitor that has a sampling rate of 600 Hz and 12-bit sampling characteristics (Card(X)plore BPM, ECG & Actigraph from Meditech-Ltd., Budapest, Hungary). The algorithm was based on ECG data from 97 individuals. The analysis encompassed 20,000 RRIs of each patient (~5 h). Dedicated software (Lab-VIEW, National Instrument, Austin, TX, USA) was used to perform the ECG signal analysis and to employ the proposed algorithm. The study was approved by the Vilnius Regional Biomedical Research Ethics Committee and written informed consent was obtained from all participants (No. 158200-17-889-400). Data analysis was performed using the IBM SPSS Software 21.0 version (SPSS, Chicago, IL, USA). The proposed adaptive algorithm was installed on a machine with LabVIEW and Visual C++ 6.0 running Microsoft Windows 10 on a 3.5 GHz processor and 8 GB of memory.

2.2. Study Population

In total, 97 middle-aged 50–55-year-old subjects (both men and women) were included in the study. Sixty-three subjects were diagnosed with metabolic syndrome (MetS) (according to the modified National Cholesterol Education Program, Adult Treatment Panel III (NCEP-ATP III) criteria), in all of whom arterial blood pressure (BP) of more than 130/85 mmHg was present or they were treated with antihypertensive drugs. The remaining 34 subjects were the same age without MetS and had BP was less than 130/85 mmHg. All subjects were recruited from the ongoing Lithuanian High Cardiovascular Risk (LitHiR) primary prevention program as a part of an ongoing project that analyses metabolic syndrome patients [39]. In addition, daily activities according to the data from the subjects' diary are presented in Appendix A.

2.3. Proposed Novel Algorithm

We propose a novel adaptive algorithm for noise reduction of the RRI time series, which is aimed at the main limitation that is common for traditionally used algorithms: the alteration of the overall RRI sequence length. The proposed algorithm is designed to maximally restore the sinusoidal heart rate structure and is based on several principles:

1. The difference between the actual duration of the RRI recording and the sum of the identified RRI values at any point in time may not exceed the difference of one average RRI value of the measured interval.
2. The spectral characteristic of the RRI sequence obtained after artifact elimination cannot be artificially distorted.
3. The proposed algorithm is designed for long-term (hours or days) sequences, where it is difficult to precisely carry out artifact elimination without changing the timeline structure.

2.4. Procedural Steps

A detailed representation of the algorithm's processes is given in Figure 1; however, the algorithm can be conditionally divided into the following steps:
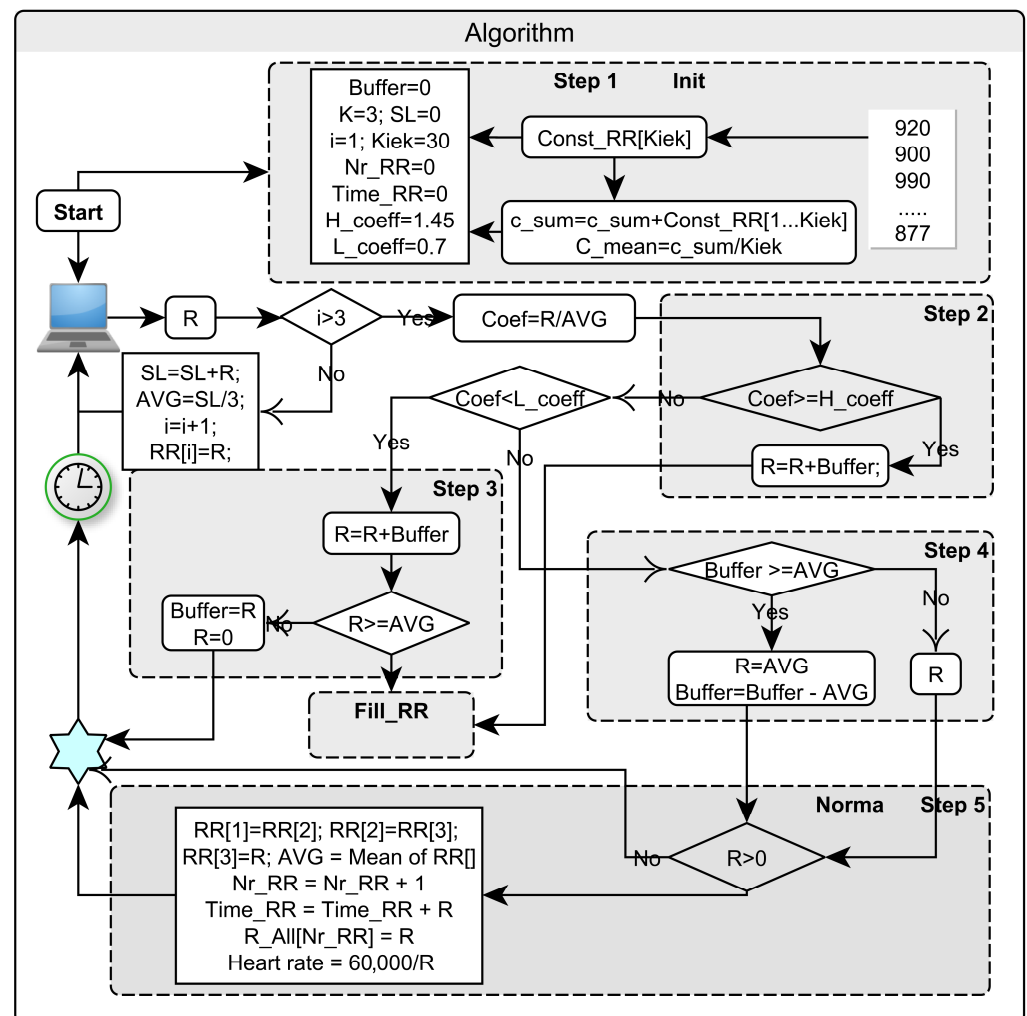
Figure 1. The stepwise procedure of adaptive algorithm (R—input of RR interval from file or device; AVG—average of 3 RR intervals; Coef—division of RRI and average; L_coeff and H_coeff—minimal and maximal criteria; RR[ ]—array of normal RRI; Time_RR—counter of time; Buffer—reserve of RR interval).

2.4.1. Step 0

The first objective is to separate artifacts from normal RRI values. To do this, the ratios of adjacent RRI values (Coef, Figure 1) were compared to the values of selected window maximum and minimum limits. If there is one artifact in the time sequence, the division of adjacent values (before and after artifact) creates two incorrect values. Therefore, each RRI value must be divided by the average of the moving window of three RRI values. The selected window indicates that all new divided values that appeared in the middle of the window remain unchanged, but values, which appeared outside of them, are named as artifacts. Statistical methods should be used to find these criteria.

2.4.2. Step 1

A very important step in initiating further actions (Init subroutine—Figure 1). A constant memory field (length = 30) filled with RRI values with normal frequency characteristics is formed and averaged. Afterwards, a temporary RRI memory field is allocated for normal and artifact-cleared RRI values with a short moving period (~3 s), from which their moving average will be continuously calculated. It is of the greatest importance that the first three initial RRI values are taken from the sequence itself and without artefacts.

2.4.3. Step 2

All RRI values whose ratio to the moving window average exceeds the upper criterion are corrected with RRI values close to the RRI at that moment (Fill_RR, look at subroutine—Figure 1). The artificial RRI field is filled with RRI values from the backup memory space (length = 30). To avoid the distortion of the frequency sequence structure, the field addresses are chosen randomly. Their average value (length = 30) must be equal to the moving window average (length = 3) at that moment. The time difference between the sum of the bad RRIs found and the sum of the time values filled with new values is stored in the buffer memory. The resulting number of RRIs in the sequence increases. To preserve the time equivalence of the last steps of the observed structure, this step can be continued indefinitely. The added RRI frequency characteristic is equivalent to white noise; it will not affect the other results and will not change their average. The RRI signal values are uncorrelated (independent), just like white noise. The mean of the white noise values is zero and the frequency response is flat and covers the entire frequency band. Thus, by filling the noise-affected sections with a sequence of values whose properties correspond to those of white noise, the characteristics of the RRI signal are preserved. If an artefact exceeds the value of only one RRI, it can be replaced by an average of the moving window values or a predictive value derived from a selected short section of the RRI sequence.

2.4.4. Step 3

RRI values, which have a ratio to the moving window average below the minimum criterion, are corrected by summing the value of the artefact first with the value of the buffer and then with the subsequent values in the time sequence up to the value of the RRI that exceeds the level of the moving field average. It may even be necessary to aggregate several short time intervals, which reduces the number of RRIs in the sequence. Consequently, the shorter RRI value(s) are replaced by the "correct" RRIs values, which are based on the moving window average (or prediction), and the difference of the resulting time sum with it is placed back in the buffer memory.

2.4.5. Step 4

The RRI values, which ratios with the average of a window are within the selected coefficients, are considered normal and are not changed. In this step, the buffer value is continuously checked. When this value exceeds the average of the moving window after the next steps, a new RRI value is added to the main stored heart rate field and the time sequence number is extended by one additional value. The buffer value decreases.

2.4.6. Step 5

After steps 2, 3, and 4, the normal or corrected RRI value(s) are stored in memory or in a file and summed. In this way, this and the following steps maintain the equality between the sum of the original (or real-time) RRI and the sum of all the corrected RRIs.

2.5. Programing Step

With the addition of the necessary inserts, the simplified program provided in Appendix B could be used. The authors deliberately retained a similar symbolism of parameters in the program block diagram and program functions as far as possible.

3. Results

3.1. Criteria Selection

In most cases, RRI values that are outside the distribution of three standard deviations (SD) from the mean (more or less ±3 ×SD) are attributed to artifacts. This criterion can be applied when the sequence distribution is normal (Gaussian). However, the distribution of the obtained values changes after the division of adjacent RRI values. It is known

that the ratio of two normally distributed random variables has a Cauchy distribution [40]. Hence this distribution has no instantaneous values: mean, variance, or higher moments defined. The mean of the sequence after the division is close to one. In order to find the required criteria, pre-whitening is often performed in such cases, which can be achieved using the autoregressive (AR) model [41]. Since the main one-step operation of artifact exclusion is the division of adjacent RRI values, we consider that in this case, the application of the first-order AR model to RRI sequences is guaranteed by the removal of the first-order link between adjacent RRI.

We use the following dependencies to solve the problem:

$$e_n = RRI_n + a_1 RRI_{n-1} \tag{1}$$

where e_n is the Gaussian white noise series with a mean of non-zero and a_1 is the first-order autoregressive coefficient. Remaining noise standard deviation:

$$\sigma_e = \sqrt{\frac{1}{N-1}\sum_n (e_n - \bar{e})^2} \tag{2}$$

where $\bar{e}$ is the mean of e_n, $e(n)$ is the Gaussian white noise series with a mean of non-zero; $a(1)$ is the first-order autoregressive coefficient; $E[x]$ is the expected x value of the mean.

The AR model was applied to the sequence that results from dividing each RRI value by the moving average of the previous three values. The noise sequence obtained after autoregressive filtering is close to a normal distribution, from which it is possible to find its mean and to find confidence limits of 0.997. These coefficients were used to distinguish normal RRIs from elongated "top" and truncated "bottom" artificial values, which were eliminated. The value three times the standard deviation in both directions of this resulting noise sequence becomes the window limits for the time sequence. The mean of adjacent RRIs ratios is equal to one; therefore, the criteria will be equal to $1 \pm 3 \times SD$. The maximum criterion is obtained by adding $3 \times SD$ value to mean and minimum—by subtracting 3*SD from it. Outside of these 99.7% limits, RRI values are classified as artifacts.

3.2. Criteria Check

The following methods were used to estimate the effectiveness of the presented algorithm. The standard time-domain variability measurements of values before (RRIb) and after (RRIa) noise elimination are as follows:

$$\mu = \frac{1}{N-1}\sum_n RRI_n \tag{3}$$

where μ is the mean of *RRI*, *N*—number of *RRI* values.

Variance of *RRI*:

$$\sigma^2 = \frac{\sum_n RRI_n^2}{N-1} - \left[\frac{\sum_n RRI_n}{N-1}\right]^2 \tag{4}$$

Difference of *RRI* variance (before–after):

$$\Delta\sigma^2 = \sigma_{RRIb}^2 - \sigma_{RRIa}^2 \tag{5}$$

Adjacent differences of *RRI* series:

$$\Delta RRI_n = RRI_n - RRI_{n-1} \tag{6}$$

for *RRIa* and *RRIb*.

Variance of the adjacent differences for $\Delta RRIa$ and $\Delta RRIb$:

$$\sigma^2_{\Delta RRI} = \frac{\sum_n \Delta RRI^2_n}{N-1} - \left[\frac{\sum_n \Delta RRI_n}{N-1} \right]^2 \tag{7}$$

Short-term standard deviation:

$$\sigma1 = \sqrt{\frac{\sigma^2_{\Delta RRI}}{2}} \tag{8}$$

Long-term standard deviation:

$$\sigma2 = \sqrt{2 \cdot \sigma^2 - \Delta\sigma^2_{\Delta RRI}} \tag{9}$$

RRI total time difference:

$$tRRI = \sum_n RRIb_n - \sum_n RRIa_n \tag{10}$$

Numbers of normal RRIs and artifacts with extended (incRRI) and shortened (decRRI) intervals identified.

3.2.1. Time Frequency Analysis

After centering both sequences (subtracting the mean from the sequence), the AR coefficients were calculated by minimizing the least square errors of the forward and backward predictions. AR coefficients return to the estimated coefficients with the value of AR order ($k \geq 16$) [42]. The following parameters were calculated in both time sequences (before and after the artifact elimination):

Residual noise before $e1_n$ and after artifact elimination $e2_n$ according to functions:

$$e_n = RRI_n + a_1 RRI_{n-1} + \ldots + a_k RRI(n-k) \tag{11}$$

Remaining noise dispersion:

$$\sigma^2_e = \frac{\sum_n e^2_n}{N} - \left[\frac{\sum_n e_n}{N} \right]^2 \tag{12}$$

where e_n is the Gaussian white noise series with a mean of zero.

Degree of noise reduction in percent:

$$\Delta\sigma^2_e = 100 \cdot \frac{\sigma^2_{e1} - \sigma^2_{e2}}{\sigma^2_{e1}} \tag{13}$$

Among the subjects, some experienced both fast and slow heart rates (their RRI averages ranged from 656 ms to 1255 ms). The mean of overall RRI was equal to 891 ± 118 ms and did not change after artifact elimination (Table 1). The data presented in the table shows that the variability indexes have significantly changed after the artifact elimination. The total size of the dispersion after the calculations has also significantly decreased ($p < 0.001$). 15–16,066 ms^2 RRI variance interval difference was obtained. This change was also confirmed by other heart rate variability indexes: the variance of adjacent RRI ($p < 0.001$) and values of the Poincare plot ($\sigma1$ and $\sigma2$) ($p < 0.001$). The residual white noise level decreased by 65.0 ± 21.0%, from 2643 ± 3374 ms^2 to 473 ± 288 ms^2 ($p < 0.001$). We did not compare the power spectral density of AR spectral density before and after the artifact elimination because, in all cases, as shown in the table, the variance of the two sequences varied greatly. It is clear that the coefficients of AR also differ for the same reasons. In addition, the algorithm was validated in RRI sequences where the total number of artifacts was known. They were not large and were repaired. Accordingly, after artifact elimination, the total number of RRIs increased by up to 230 (mean = 38 ± 51) or decreased by up

to 576 RRIs (mean = 45 ± 84) (Table 1). Compared to the entire length of the sequence (20,000) this is a small number, but it has greatly changed the sequence variance.

Table 1. Changes in RRI variability before and after artifact elimination.

Variable, Units	Variable Symbols	Before		After		Paired *t*-Test
		Range	Mean ± SD	Range	Mean ± SD	
Mean of RRI, ms	μ	656–1255	891 ± 118	656–1258	890 ± 119	1.79
Variance of RRI, ms^2	σ^2	2736–45,450	14,247 ± 8586	2721–44,277	12,298 ± 7556	6.57 *
Difference of RRI variance, ms^2	$\Delta\sigma^2$			15–16,066	1948 ± 2922	
Variance of adjacent RRI, ms^2	$\sigma^2_{\Delta RRI}$	83–29,395	3959 ± 5481	68–2561	656 ± 481	6.07 *
Variance of AR noise, ms^2	σ^2_e	77–17,174	2643 ± 3374	60–1469	473 ± 288	6.53 *
Decrease in variance AR noise, %	$\Delta\sigma^2_e$			13–98	65 ± 21	
RRI total time difference, ms	tRRI			−1596–968	215 ± 430	
Short-term, ms	$\sigma1$	6–121	37 ± 24	6–36	17 ± 6	9.29 *
Long-term, ms	$\sigma2$	74–299	156 ± 46	74–297	149 ± 44	6.78 *
Normal RRI number	Nr			19,263–19,999	19,915 ± 118	
Increased RRI number	incNr			0–230	38 ± 51	
Decreased number of RRIs	decNr			0–576	45 ± 84	
3 × SD of ratio of adjacent RRI		0.05–0.64	0.18 ± 0.11	0.02–0.04	0.051 ± 0.025	11.8 *

* denotes a statistically significant difference (*p* < 0.001). Abbreviations: RRI—R-R interval; AR—autoregressive.

3.2.2. Time Complexity Analysis

The used algorithm is based on the length of the adjacent RRI. Normal RRI values should be higher than the minimum criteria and lower than the maximum. On this basis, the normal value of RRI is separated from the artifact. For the determination of these criteria, the ratios between all adjacent RRI values were calculated. A statistical analysis of the distribution of these ratios was carried out and their histograms were monitored. Particular attention has been paid to the magnitude of the mutual limits used in the artifact-finding algorithm.

Ninety-seven case studies showed that this value (Table 1, 3 × SD of the ratio of adjacent RRI) varied around 1 from 0.05 to 0.64 with an average of 0.18 ± 0.11. Figure 2 shows that the 3 × SD value is distributed around this average. This indicates that the upper value of the criterion is very close to 1.18±0.11, and the lower value is 0.82 ± 0.11 (1.0–0.18 = 0.82). Before the beginning, the artifact elimination procedure of the first-order AR simulation must be performed, and the selection of the exact coefficients is required (Figure 2). The distribution of moments of noise generated by AR simulation is asymmetric and depends on the number of artifacts (Figure 2 and Table 1). Visually, in a separate example, the noise distribution is clearly visible (Figure 3). The values observed on both sides of the average are attributed to artifacts (±3 × SD = 0.18 ± 0.11).
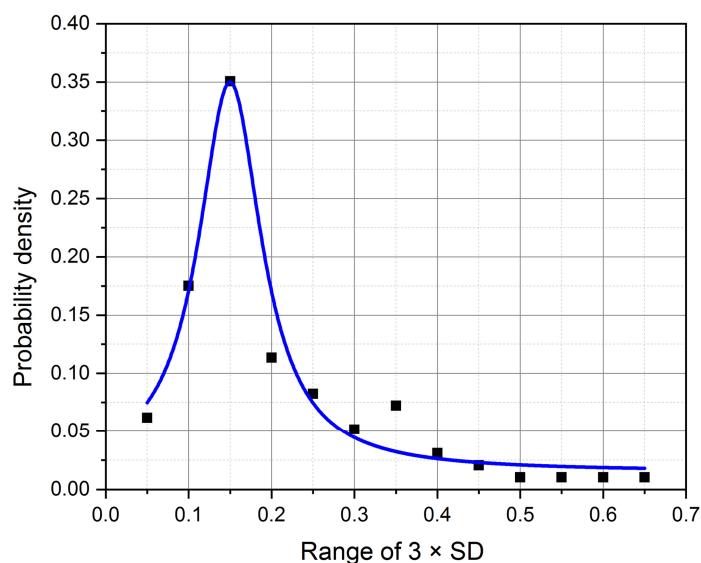
Figure 2. Distribution of averaged artifact standard deviation (3 × SD) of all 97 cases.
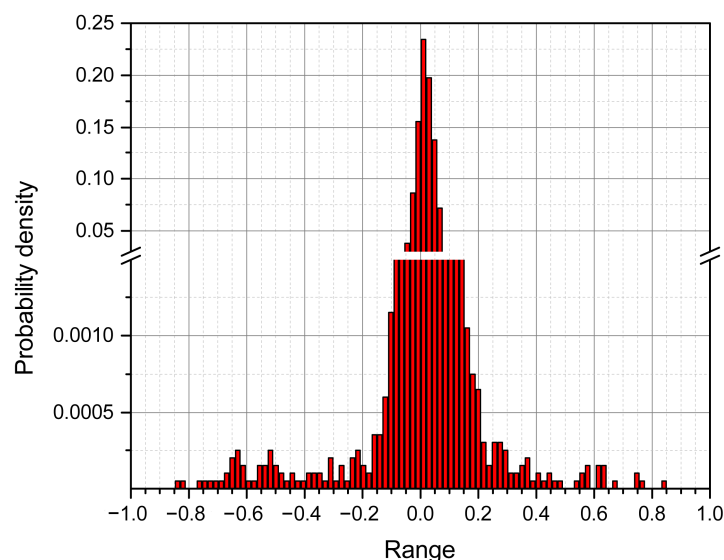

Figure 3. Artifact distribution after autoregressive RRI modeling in a separate case.

4. Discussion

There are a plethora of fundamentally different RRI processing methods being used, which sometimes makes comparing different HRV data difficult. Although many HRV software uses automatic repair systems, they usually lack information about how they behave with artifacts. Two basic principles were used to construct our proposed method. The first is that:

1. Autoregressive research and other methods show that strong relationships in the RRI sequence remain not only between adjacent RRIs but also extend to the number of RRI values from 4 to 50 (order of AR model) [43,44].

In that period, any noise will be non-stationary and very strong. Comparing artifacts with adjacent RRI values allows us to identify them within the time sequence. Therefore, changes in adjacent RRI values in our work have been used as the main indicators for separating artifacts from physiological RRI changes. A minimum of three moving RRI values (for average and for window) were selected, with which the next RRI was compared. The stability of the algorithm is conditioned by the fact that none of the artifacts found in all sequences can pass to the moving average window. The last value (third)

must include only normal or corrected RRI values in the scrolling field. The second principle is based on the perception that:

2. Each RRI sequence is made up of "fragmented" real-time intervals, the sum of which should correspond to the actual elapsed registration time.

The proposed algorithm follows the principle that the duration of the time sequence should remain identical to the actual elapsed time, independent of normal RRI values and the number of artefacts present: therefore, long-duration artefacts should be filled with similar expected duration values and very short duration artefacts should be aggregated to a value close to the former RRI (using the above-described method). Changing the length of long or short time intervals by moving the selected average window may shorten the overall sequence duration and lead to a duration difference compared to the actual elapsed time. Therefore, the time difference between the original RRI sequence and the one after artifact elimination, which is always shorter, is placed in the backup buffer and used in the next stage of the operation. Compared to other popular methods, our proposed algorithm does not change the overall length of the ECG recording. Following the use of this algorithm in the analysis, one can safely apply methods of multivariate spectral analysis, which require complete synchronization between the processes registered, since only under such conditions it is possible to accurately assess the existing relationships between the observed processes.

Nonetheless, there are some limitations that are important to note. Firstly, due to the fundamental difference of the proposed method, we were not able to directly compare it with other methods, which are being used by the majority of authors—not even if we used a simple method such as mean-square error: currently, the most commonly used classical artifact elimination algorithm, which maintains a constant number of RRI in a sequence, eliminated artifacts by increasing RRI sequence length in one place and shortening it in another. Secondly, we acknowledge that by design, the proposed algorithm is suited for long-term (hours or days) sequences and, therefore, is not applicable in very short sequences.

5. Conclusions

A novel artifact elimination algorithm that does not affect the overall structure of the RRI sequence and does not alter the overall duration of the sequence has been proposed, and its effectiveness with real-life data has been presented. The proposed algorithm is designed for long-term (hours or days) sequences, in which it is difficult to precisely carry out artifact elimination without changing the timeline structure, and it is especially suitable for use in polygraphy when multiple physiological and other processes associated with other cardiac arrhythmias are registered in parallel. Following the use of this algorithm in the analysis, one can safely apply methods of multivariate spectral analysis, which require a complete synchronization between the processes registered, since only under such conditions is it possible to accurately assess the existing relationships between the observed processes.

Abbreviations

a(1)	The first-order autoregressive coefficient.
AR	Autoregressive.
AVG	Average of 3 RR intervals.
BP	Blood pressure.
Buffer	Reserve of RR interval.
Coef	Division of RRI and average.
decRRI	RRIs with shortened intervals.
e_n	The Gaussian white noise series.
$E[x]$	The expected x value of the mean.
$e1_n$	Residual noise before artifact elimination.
$e2_n$	Residual noise after artifact elimination.
ECG	Electrocardiographic.
H_coeff	Maximal criteria.
HRV	Heart-rate variability.
incRRI	RRIs with extended intervals.
L_coeff	Minimal criteria.
LitHiR	Lithuanian High Cardiovascular Risk primary prevention program.
MetS	Metabolic syndrome.
R	Input of RR interval from file or device.
RR[ ]	Array of normal RRI.
RRI	RR interval.
RRIa	Standard time-domain variability measurements after noise elimination.
RRIb	Standard time-domain variability measurements before noise elimination.
SD	Standard deviation.
Time_RR	Counter of time.

Appendix A

A table of subjects' daily activities according to the data of the subject's diary is provided.

Table A1. Daily activities according to the data from the subjects' diary.

Daily Activities, in Hours	Subjects without MetS	Subjects with MetS
	Mean ± SD	Mean ± SD
Working	4.48 ± 4.08	3.94 ± 3.64
At home	2.12 ± 1.78	2.30 ± 2.18
Walking	1.22 ± 0.98	1.54 ± 1.76
Driving	0.71 ± 0.84	1.08 ± 1.18
Eating	1.49 ± 0.88	0.98 ± 0.45
Lying in bed	2.35 ± 2.23	1.90 ± 2.06

Abbreviations: MetS—metabolic syndrome; SD—standard deviation.

Appendix B

A simplified artifact elimination program is provided.

```
//*========================================*/
//* Title:    int albsta() */
//*========================================*/
// function declaration
int Norma();
int Init();
int Fill_RR();
int albsta();
// Fill RRIs constant
int Const_RR [30] = {920, 900, 900, 917, 883, 887, 893, 860, 890, 940,
897, 907, 903, 873, 877, 880, 863, 870, 917, 900, 857, 860, 853, 823,
837, 853, 830, 877, 907, 877};
int R;          // Normal and Corrected RRI
double AVG;             // Moving average of RRI of RR(K) array
long int Nr_RR;         // Counter of R post revision
int Buffer;         // Reserve RRI time (ms)
int K;              // Moving temporary RRI array length
const float High_coefficient = 1.45 // or other;
const float Low_coefficient = 0.75 // or other;
int Kiek;           // Constant RR array length
int RR [3];         // Moving temporary RR(K) array
int Tempor [30];        // temporary array
int R_All [1000];       // Normal and Corrected RRI array
long int Time_RR;       // Adjusted total time from R Length
int C_mean;             // Average of Const_RRI (from 1 To 30)
int Puls;           // Heart Rate
long i;             // Continuously value of a variable
long int c_sum;
//=============Initiates the initial conditions
int Init()
{
long Sum = 0;
Sum = 0;
K = 3;          // Length of moving average array
Kiek = 30;          // length of Constant RRI array
c_sum = 0;
for (i = 1; i <= Kiek; i++) {
c_sum = c_sum + Const_RR[i];
}               // Sum of Constant RR
C_mean = c_sum/Kiek;// Mean of Constant RRI
for (i = 1; i <= K; i++) {
RR[i] = R;   // Read first three good RRI
Sum = Sum + R;
}
Nr_RR = 0;          // Counter of normal RRI
AVG = Sum/K;
Buffer = 0;         // Adaptive RRI time buffer
Time_RR = 0;            // Adjusted total time from RRI
}               // Fill in the missing RRI by 30
int Fill_RR()
{
long differ;
long j = 0;
long vv = 0;
```

```
int s = 0;
c_sum=0;
differ = C_mean − AVG;
for (i = 1; i <= Kiek; i++) {
Tempor[i] = Const_RR[i] − differ;   // Equals the constant
}              // average until it is found
j = R/AVG;
vv = AVG;
Buffer = R − (vv*j);
for (i = 1; i <= j; i++) {
s = rand() % (Kiek − 2) + 1; // Random of array
R = Tempor[s];
c_sum = c_sum + R;
Norma();
}
Buffer = Buffer − c_sum;
}
//*=Fills the fields of the RRI and checks it repeatedly
int Norma()
{
long Sum;
Sum = 0;
for (i = 1; i <= K − 1; i++) {
RR[i] = RR[i + 1];
Sum = Sum + RR[i];
}
RR[K] = R;
Sum = Sum + R;
Time_RR = Time_RR + R;      // Time counter
AVG = Sum/K;          // Moving average RRI(k)
Nr_RR = Nr_RR + 1;          // All RR counter
R_All[Nr_RR] = R;          // Add a RR value
Puls = 60000/R;          // Heart rate
}
}
//===Main program============================
int albsta()
{
double Coef;
double ccc;
double Sum;
Sum =0;
Init();                // Initiates the initial values
// New RRI are expected, For example...1000
for (i = 1; i <= 1000; i++) {            // Input R from file or device
if (i <= 3) {
Sum = Sum + R;
AVG = Sum/3;
RR[i] = R;
} else {
Coef = R/AVG;          // Division RRI and average
// Long RRI
if (Coef > High_coefficient) {
R = R + Buffer;
```

```
Fill_RR();
}
// Normal RRI
if (Coef <= High_coefficient & Coef > Low_coefficient) {
Norma();
if (Buffer >= AVG) {
R = AVG;
Buffer = Buffer − AVG;
Norma();
}
}
// Short RRI
if (Coef < Low_coefficient) {
R = Buffer + R;
ccc = R/AVG;
if (ccc >= High_coefficient) {
Fill_RR();
} else {
Buffer = R;
}
}
}
}
}
```

References

1. Sammito, S.; Thielmann, B.; Seibt, R.; Klussmann, A.; Weippert, M.; Böckelmann, I. Guideline for the application of heart rate and heart rate variability in occupational medicine and occupational science. *ASU Int.* **2015**, *2015, 1–29*. https://doi.org/10.17147/asui.2015-06-09-03.
2. Malik, M. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, heart rate variability—Standards of measurement, physiological interpretation, and clinical use. *Circulation* **1996**, *93*, 1043–1065. https://doi.org/10.1161/01.CIR.93.5.1043.
3. Kher R. Signal processing techniques for removing noise from ECG signals. *J. Biomed. Eng. Res.* **2019**, *3*, 1–9. https://doi.org/10.17303/jber.2019.3.101
4. Jarchi, D.; Charlton, P.; Pimentel, M.; Casson, A.; Tarassenko, L.; Clifton, D.A. Estimation of respiratory rate from motion contaminated photoplethysmography signals incorporating accelerometry. *Heal. Technol. Lett.* **2019**, *6*, 19–26. https://doi.org/10.1049/htl.2018.5019.
5. Clifford, G.D.; Azuaje, F.; Mc Sharry, P. *Advanced Methods and Tools for Ecg Data Analysis*; Artech House Boston: Boston, MA, USA, 2006.
6. Clifford, G.D.; Azuaje, F.; Mcsharry, P. Ecg Statistics, Noise, Artifacts, and Missing Data. *Adv. Methods Tools ECG Data Anal.* **2006**, *6*, 18.
7. Friesen, G.M.; Jannett, T.C.; Jadallah, M.A.; Yates, S.L.; Quint, S.R.; Nagle, H.T. A comparison of the noise sensitivity of nine QRS detection algorithms. *IEEE Trans. Biomed. Eng.* **1990**, *37*, 85–98. https://doi.org/10.1109/10.43620.
8. Satija, U.; Ramkumar, B.; Manikandan, M.S. A Review of Signal Processing Techniques for Electrocardiogram Signal Quality Assessment. *IEEE Rev. Biomed. Eng.* **2018**, *11*, 36–52. https://doi.org/10.1109/rbme.2018.2810957.
9. Van Alste, J.A.; Schilder, T.S. Removal of Base-Line Wander and Power-Line Interference from the ECG by an Efficient FIR Filter with a Reduced Number of Taps. *IEEE Trans. Biomed. Eng.* **1985**, *BME-32*, 1052–1060. https://doi.org/10.1109/TBME.1985.325514.
10. Levkov, C.; Mihov, G.; Ivanov, R.; Daskalov, I.; Christov, I.; Dotsinsky, I. Removal of power-line interference from the ECG: a review of the subtraction procedure. *Biomed. Eng. Online* **2005**, *4*, 50. https://doi.org/10.1186/1475-925x-4-50.
11. Frølich, L.; Dowding, I. Removal of muscular artifacts in EEG signals: a comparison of linear decomposition methods. *Brain Inform.* **2018**, *5*, 13–22. https://doi.org/10.1007/s40708-017-0074-6.
12. Velayudhan, A.; Peter, S. Noise analysis and different denoising techniques of ECG signal-a survey. *IOSR J. Electron. Commun. Eng.* **2016**, *1*, 40–44.
13. Berntson, G.G.; Quigley, K.S.; Jang, J.F.; Boysen, S.T. An Approach to Artifact Identification: Application to Heart Period Data. *Psychophysiology* **1990**, *27*, 586–598. https://doi.org/10.1111/j.1469-8986.1990.tb01982.x.

14. Kaufmann, T.; Sütterlin, S.; Schulz, S.M.; Vögele, C. ARTiiFACT: a tool for heart rate artifact processing and heart rate variability analysis. *Behav. Res. Methods* **2011**, *43*, 1161–1170. https://doi.org/10.3758/s13428-011-0107-7.

15. Wu, C.-M.; Chen, S.-C.; Chen, Y.-J. A multiple bio-signal measurement analysis and warning system for the long-term health care of severe disabled. *Microsyst. Technol.* **2016**, *24*, 155–163. https://doi.org/10.1007/s00542-016-3215-6.

16. Kemper, K.J.; Hamilton, C.; Atkinson, M. Heart Rate Variability: Impact of Differences in Outlier Identification and Management Strategies on Common Measures in Three Clinical Populations. *Pediatr. Res.* **2007**, *62*, 337–342. https://doi.org/10.1203/pdr.0b013e318123fbcc.

17. Luz, E.J.D.S.; Schwartz, W.R.; Cámara-Chávez, G.; Menotti, D. ECG-based heartbeat classification for arrhythmia detection: A survey. *Comput. Methods Programs Biomed.* **2016**, *127*, 144–164. https://doi.org/10.1016/j.cmpb.2015.12.008.

18. Imtiaz, S.A.; Mardell, J.; Saremi–Yarahmadi, S.; Rodriguez–Villegas, E. ECG artefact identification and removal in mHealth systems for continuous patient monitoring. *Heal. Technol. Lett.* **2016**, *3*, 171–176. https://doi.org/10.1049/htl.2016.0020.

19. Al Mahamdy, M.; Riley, H.B. Performance Study of Different Denoising Methods for ECG Signals. *Procedia Comput. Sci.* **2014**, *37*, 325–332. https://doi.org/10.1016/j.procs.2014.08.048.

20. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.-C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **1998**, *454*, 903–995.

21. Singh, G.; Kaur, G.; Kumar, V. ECG Denoising using Adaptive Selection of IMFs through EMD and EEMD. In Proceedings of the 2014 International Conference on Data Science & Engineering (ICDSE), Kochi, India, 26–28 August 2014.

22. Cajal, D.; Hernando, D.; Lázaro, J.; Laguna, P.; Gil, E.; Bailón, R. Effects of Missing Data on Heart Rate Variability Metrics. *Sensors* **2022**, *22*, 5774. https://doi.org/10.3390/s22155774.

23. Dwivedi, A.K.; Ranjan, H.; Menon, A.; Periasamy, P. Noise Reduction in ECG Signal Using Combined Ensemble Empirical Mode Decomposition Method with Stationary Wavelet Transform. *Circuits Syst. Signal Process.* **2020**, *40*, 827–844. https://doi.org/10.1007/s00034-020-01498-4.

24. Chatterjee, S.; Thakur, R.S.; Yadav, R.N.; Gupta, L.; Raghuvanshi, D.K. Review of noise removal techniques in ECG signals. *IET Signal Process.* **2020**, *14*, 569–590. https://doi.org/10.1049/iet-spr.2020.0104.

25. Mateo, J.; Laguna, P. Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal. *IEEE Trans. Biomed. Eng.* **2003**, *50*, 334–343. https://doi.org/10.1109/tbme.2003.808831.

26. Nabil, D.; Reguig, F.B. Ectopic beats detection and correction methods: A review. *Biomed. Signal Process. Control* **2015**, *18*, 228–244. https://doi.org/10.1016/j.bspc.2015.01.008.

27. Zhao, L.; Li, J.; Xiong, J.; Liang, X.; Liu, C. Suppressing the Influence of Ectopic Beats by Applying a Physical Threshold-Based Sample Entropy. *Entropy* **2020**, *22*, 411. https://doi.org/10.3390/e22040411.

28. Mc Names, J.; Thong, T.; Aboy, M. Impulse rejection filter for artifact removal in spectral analysis of biomedical signals. In Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Francisco, CA, USA, 1–5 September 2004.

29. Lee, M.-Y.; Yu, S.-N. Improving discriminality in heart rate variability analysis using simple artifact and trend removal preprocessors. In Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2010), Buenos Aires, Argentina, 31 August–4 September 2010.

30. Benchekroun, M.; Chevallier, B.; Istrate, D.; Zalc, V.; Lenne, D. Preprocessing Methods for Ambulatory HRV Analysis Based on HRV Distribution, Variability and Characteristics (DVC). *Sensors* **2022**, *22*, 1984. https://doi.org/10.3390/s22051984.

31. Stapelberg, N.J.C.; Neumann, D.L.; Shum, D.H.K.; McConnell, H.; Hamilton-Craig, I. A preprocessing tool for removing artifact from cardiac RR interval recordings using three-dimensional spatial distribution mapping. *Psychophysiology* **2016**, *53*, 482–492. https://doi.org/10.1111/psyp.12598.

32. Lippman, N.; Stein, K.M.; Lerman, B.B. Comparison of methods for removal of ectopy in measurement of heart rate variability. *Am. J. Physiol. Circ. Physiol.* **1994**, *267*, H411–H418. https://doi.org/10.1152/ajpheart.1994.267.1.h411.

33. Acar, B.; Savelieva, I.; Hemingway, H.; Malik, M. Automatic ectopic beat elimination in short-term heart rate variability measurement. *Comput. Methods Programs Biomed.* **2000**, *63*, 123–131. https://doi.org/10.1016/s0169-2607(00)00081-x.

34. Solem, K.; Laguna, P.; Sörnmo, L. (Eds.) Handling of ectopic beats in heart rate variability analysis using the heart timing signal. *Proc. Med. Conf. Med. Biol. Eng.* **2004**, *2004*.

35. Barbieri, R.; Brown, E. Analysis of Heartbeat Dynamics by Point Process Adaptive Filtering. *IEEE Trans. Biomed. Eng.* **2005**, *53*, 4–12. https://doi.org/10.1109/tbme.2005.859779.

36. Zhang, Q.; Zhou, D.; Zeng, X. A novel machine learning-enabled framework for instantaneous heart rate monitoring from motion-artifact-corrupted electrocardiogram signals. *Physiol. Meas.* **2016**, *37*, 1945–1967. https://doi.org/10.1088/0967-3334/37/11/1945.

37. Mishra, P.; Singla, S.K. Artifact Removal from Biosignal using Fixed Point ICA Algorithm for Pre-processing in Biometric Recognition. *Meas. Sci. Rev.* **2013**, *13*, 7–11. https://doi.org/10.2478/msr-2013-0001.

38. Peltola, M.A. Role of editing of R–R intervals in the analysis of heart rate variability. *Front. Physiol.* **2012**, *3*, 148. https://doi.org/10.3389/fphys.2012.00148.

39. Slušnienė, A.; Laucevičius, A.; Navickas, P.; Ryliškytė, L.; Stankus, V.; Stankus, A.; Navickas, R.; Laucevičienė, I.; Kasiulevičius, V. Daily Heart Rate Variability Indices in Subjects with and Without Metabolic Syndrome Before and After the Elimination of the Influence of Day-time Physical Activity. *Medicina* **2019**, *55*, 700. https://doi.org/10.3390/medicina55100700.

40. Weisstein, E.W. Normal Ratio Distribution. From MathWorld—A Wolfram Web Resource. Available online: https://math-world.wolfram.com/NormalRatioDistribution.html (accessed on 1 March 2022).

41. Marple, S. *Digital Spectral Analysis with Applications;* Prentice-Hall: Englewood Cliffs, NJ, USA, 1987; pp. 130–273.

42. Boardman, A.; Schlindwein, F.S.; Rocha, A.P.; Leite, A. A study on the optimum order of autoregressive models for heart rate variability. *Physiol. Meas.* **2002**, *23*, 325–336. https://doi.org/10.1088/0967-3334/23/2/308.

43. Arsenos, P.; Manis, G. Deceleration Capacity of heart rate: Two new methods of computation. *Biomed. Signal Process. Control* **2014**, *14*, 158–163. https://doi.org/10.1016/j.bspc.2014.07.013.

44. Pereira, T.; Almeida, P.R.; Cunha, J.P.; Aguiar, A. Heart rate variability metrics for fine-grained stress level assessment. *Comput. Methods Programs Biomed.* **2017**, *148*, 71–80. https://doi.org/10.1016/j.cmpb.2017.06.018.