Computers in Biology and Medicine 196 (2025) 110803

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Effects of electrocardiographic noise on ultra-short term Heart Rate Variability indices

Anna Raimondi [a], Alessandro Busacca [a], Giuseppe Costantino Giaconia [a], Yuri Antonacci [a], Salvatore Stivala [a], Luca Faes [a,b], Riccardo Pernice [a,*]

[a] Department of Engineering, University of Palermo, Viale delle Scienze, 90128, Palermo, Italy
[b] Faculty of Technical Sciences, University of Novi Sad, Trg Dositeja Obradovića 6, 21102, Novi Sad, Serbia

ARTICLE INFO

Keywords:
Heart Rate Variability (HRV)
Ultra-short term HRV analysis
Electrocardiography (ECG)
Noise
Signal-to-noise ratio (SNR)
Entropy
Time-series analysis

ABSTRACT

Heart Rate Variability (HRV) is a key metric for assessing cardiovascular health and autonomic nervous system function. The increasing use of wearable devices for continuous health monitoring during daily-life activities presents significant challenges, since the acquired signals are often noisy or affected by artifacts, resulting in a low signal-to-noise ratio (SNR). This study aims to investigate how electrocardiographic (ECG) noise affects the accuracy of ultra-short term (~ 2 min) HRV analysis. Time-, frequency- and information-domain HRV indices, computed on interbeat interval time series extracted from ECG signals contaminated by different types of simulated noise (white and frequency-specific) at various SNR levels (−3, 1, 5, 10 and 20 dB) were compared to those obtained on reference noise-free waveforms. The results show that low-frequency noise (i.e., at 0.01, 0.1, 0.3, and 3 Hz) at an SNR lower or equal than 5 dB has a significant impact on the reliability of HRV measures, leading to remarkably diminished correlation with reference values. On the other hand, white and higher-frequency noise (i.e., 50 Hz and 300 Hz) had a reduced impact on the computed indices even for very low SNR values. Overall, a SNR level of at least 10 dB seems enough for ensuring reliable HRV analysis across all domains. These findings are valuable for improving the reliability of HRV analysis especially in the case of short-duration signals acquired in noisy or extreme environments, to ensure that wearable devices can provide reliable physiological information even in challenging conditions.

1. Introduction

The rapid advancement in microelectronics and signal processing technologies has driven the widespread adoption of sensors and wearable devices for continuous health monitoring during daily-life activities [1–3]. These devices have allowed the continuous recording of biosignals on patients directly at home, which are then remotely sent to healthcare professionals and clinicians for diagnosis of diseases or follow-up through telemedicine services, with significant benefits for both the patients and the healthcare system [2,4–6].

Despite its potential, the use of wearable devices for physiological monitoring presents significant challenges, since the acquired signals are often noisy or severely affected by artifacts, with a low signal-to-noise ratio (SNR) which requires heavy preprocessing or filtering procedures demanding high computational costs; in worst cases such procedures may be not enough to avoid unreliable information, forcing to discard severely corrupted data [7–9]. Additionally, wearable devices are constrained by limited processing power and memory, posing challenges for collecting the minimum number of samples required for an accurate extraction or analysis of widely used cardiovascular physiological indices [10–13]. From these premises, it is evident that the design of fast and efficient signal processing algorithms for analysing data acquired from wearable devices is still an ongoing challenge.

One important application of wearable devices is the study of physiological regulation and autonomic tone from time series extracted from single-lead electrocardiographic (ECG) or photoplethysmographic (PPG) recordings under different physiological conditions, in order to provide valuable insights into the dynamics of the autonomic nervous system (ANS) [3,14,15]. Heart Rate Variability (HRV) is a key metric for assessing cardiovascular status and autonomic tone [16,17] and can be evaluated starting from the sequence of consecutive interbeat intervals. It provides valuable information on the ANS functional state and on the balance between the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS) activities [16,17], having as well a significant prognostic value in ageing and cardiac diseases [18,19]. A healthy cardiovascular system usually exhibits a higher HRV, reflecting its ability to respond to physiological changes, while a reduced

* Corresponding author.
E-mail address: riccardo.pernice@unipa.it (R. Pernice).

https://doi.org/10.1016/j.compbiomed.2025.110803
Received 8 March 2025; Received in revised form 29 June 2025; Accepted 17 July 2025
Available online 6 August 2025

HRV has been often associated with poor cardiovascular outcomes or pathological conditions, such as arrhythmias and an increased risk of sudden cardiac death [16,18,20–23]. However, abnormally high HRV can be sometimes linked to pathologies altering the heart rate pattern, and, for example, has been reported in patients with atrial fibrillation, given the overall higher degree of irregularity of ventricular rhythm [24,25].

HRV can be assessed from cardiovascular time series using metrics in both the time and frequency domains, the former quantifying the average heart rhythm and the extent of beat-to-beat variability, the latter instead providing insight into the distribution of absolute or relative power in specific frequency bands to assess autonomic function [17,26]. Moreover, since conventional time-domain and frequency-domain HRV metrics fail to capture the full complexity and dynamic behaviour of heart rate fluctuations [18], information-domain analysis has emerged as a promising approach to quantify the complexity and information content of time series data through the use of entropy measures [27]. Such content is related to the regularity of temporal patterns found in the time series and has been also put in relation to the balance between sympathetic and parasympathetic ANS activity [28,29].

HRV analysis has been standardised for long-term recordings (i.e., 24 h) obtained from ambulatory monitoring, while short-term registrations (~5 min) are widely employed for practical purposes being able to evidence the dynamic relationship between the sympathetic and parasympathetic branches of ANS [17]. However, the growing interest in ultra-short term (UST) recordings (i.e., less than 5 min) acquired from wearable devices has highlighted the need to better understand their reliability compared to short-term analysis and the effect of noise and artifacts on HRV indices [10,11,13,16]. This is especially important for signals acquired in noisy or even extreme environments [30,31].

While previous research has primarily focused on the effects of white noise on HRV indices computed from ECG signals [32] or of white/pink noise on heart rate complexity metrics [33], the literature lacks a comprehensive comparison of how different types of noise, including those caused by physiological mechanisms, affect the most used HRV indices. Indeed, ECG recordings not only are strongly affected by external (e.g. environmental or instrumental) noise, but are also often influenced by other physiological mechanisms or biosignals that can represent a sort of "physiological noise" [34], such as respiration or muscular activity.

Within this context, our study aims to systematically investigate the impact of simulated noise added to ECG signals on the accuracy of time-, frequency- and information-domain HRV indices of ultra-short term HRV assessed from ECG recordings of ~2 min duration. The indices computed from reference noise-free ECG signals will be compared with those obtained from recordings contaminated by different types of noise at various SNR levels, including low-frequency, white and physiological noise. In our simulations, the latter will be modelled with spectral properties similar to the underlying physiological phenomena, to more realistically simulate the conditions under which ECG data acquired by wearable or portable devices can be affected. The main aim is to provide valuable insights into the reliability of R-peak detection and thus HRV indices from ultra-short recordings, enabling in perspective a more careful use of wearable devices for health monitoring in real-life scenarios.

2. Materials and methods

2.1. Data acquisition and pre-processing

In this study, the noise-free ECG signals belonged to a database collected at the Department of Engineering of the University of Palermo from 17 young healthy volunteers (10 males, 7 females) aged 24.7 ± 4.1 years using a in-house multisensor portable system. Further details on a previous prototype of the device, supporting lower sampling rates and integrating a different type of photodetector [35] can be found in [36]. The analysed ECG tracks were acquired at a sampling frequency f_s of 1 kHz during a sitting resting condition lasting for 10 min to ensure stabilisation of physiological signals to baseline levels. The study was conducted in accordance with the declaration of Helsinki and was approved by the Bioethics Committee of the University of Palermo (ref. no. 125/2023).

Starting from the ECG signals acquired for each subject, a filtering process was carried out to improve signal quality and enable accurate detection of R-peaks. The pre-processing and all the subsequent analyses were conducted on MATLAB® R2021b (The MathWorks, Inc., Natick, MA, USA) software. The ECG signals were first filtered using a zero-phase fourth-order digital Butterworth bandpass filter (passband from 0.1 to 20 Hz) to remove the baseline drift and noise from the powerline (50 Hz). This filter was chosen for its maximally flat frequency response in the passband. An additional fourth-order Chebyshev type II high-pass filter with a cut-off frequency of 0.40 Hz and a stopband attenuation of 30 dB was subsequently applied, to remove residual signal fluctuations caused by respiratory modulation still present after Butterworth filtering due its relatively wide transition band [37]. The narrower transition band of the Chebyshev filter allowed to effectively remove all the undesired slower fluctuations, thus improving the overall effectiveness of the pre-processing. Finally, an initial transient of 2 s was removed from the filtered signal to keep only the stationary parts of the ECG tracks.

The above-described processing procedure allowed to obtain clean ECG signals (one per each subject) which were then treated as reference waveforms for the subsequent analyses.

2.2. Noise simulation

To evaluate the effects of noise on cardiovascular variability indices, in this work two types of noise were simulated, i.e., a Gaussian white noise and 'coloured' noises (i.e., at specific different frequencies) at various SNR levels. As known, SNR is a commonly used metric to quantify signal quality and is defined by:

$$\mathrm{SNR} = 10\log_{10}\left(\frac{P_{signal}}{P_{noise}}\right)\mathrm{dB}, \tag{1}$$

where P_{signal} and P_{noise} represent the power of the signal and the power of the noise, respectively [38].

Let us consider X a stochastic process, with X_n denoting the random variable representing the state at the current time n. White noise is characterised by its completely random nature, consisting of independent and uncorrelated samples, and has a constant power spectrum at all frequencies. The autocorrelation function $r_X(m)$ of a stationary white noise process X with zero mean and variance σ^2_{noise} is an impulse function defined as in Eq. (2), which takes the value of zero for all lags m except when $m = 0$, in which case it is equal to its variance [38]:

$$r_X(m) = E[X_n X_{n+m}] = \delta(m)\sigma^2_{noise}. \tag{2}$$

The equivalent description in the frequency-domain is obtained by applying the Fourier transform of the autocorrelation function $r_X(m)$, highlighting that the power spectrum density (PSD) is uniformly distributed over all frequency bands with a value equal to its variance σ^2_{noise} [38]:

$$P_X(f) = \sum_{m=-\infty}^{\infty} r_X(m)e^{-j2\pi\frac{f}{f_s}m} = r_X(0) = \sigma^2_{noise}. \tag{3}$$

In the present study, a Gaussian white noise signal was generated by using MATLAB built-in *mvnrnd* function, which allows the generation of random samples extracted from a normal distribution, using as input parameters a null mean, the variance of the noise and a number of samples in this case equal to the length of the signal to be corrupted [39]. Since variance is mathematically equal to the total power contained in the whole frequency range [16], control of noise variance σ^2_{noise} plays a

crucial role in modulating the SNR, because if the signal variance σ^2_{signal} remains constant, an increase of the variance consequently causes an SNR reduction. Therefore, the noise variance was determined by setting the desired SNR value as follows:

$$\sigma^2_{noise} = \frac{\sigma^2_{signal}}{10^{\frac{SNR}{10}}}. \quad (4)$$

The noise at a given frequency f was simulated as the output process of a linear univariate autoregressive (AR) model driven by uncorrelated stochastic noise. The present and past variables of the process are related by the following AR model of order p, considering m past samples [40]:

$$X_n = \sum_{m=1}^{p} a_m X_{n-m} + U_n, \quad (5)$$

where a_m, $m = 1, \ldots, p$ are the linear regression coefficients and U is a zero-mean white Gaussian noise process with variance σ^2_U. In this study, we have simulated an AR process of order $p = 2$ by placing two complex-conjugate poles (roots of the characteristic polynomial $A(z) = 1 - \sum_{m=1}^{p} a_m z^{-m}$, which is the representation of the coefficients in the z-domain, with $z = e^{j2\pi \frac{f}{f_s}}$), these poles have a modulus ρ and a phase of $\pm 2\pi f$ [26]. This configuration is arranged to determine the coefficients of the AR polynomial as follows [40]:

$$a_1 = 2\rho \cos(2\pi f), \quad a_2 = -\rho^2. \quad (6)$$

The procedure to derive the autoregressive (AR) model coefficients (Eq. (6)) from the parametric representation of the process (Eq. (5)) is detailed in Appendix A.1.

The choice of poles allows us to simulate a second-order AR process with specific spectral properties given by the selected frequency f. In this configuration, the parameters ρ and f define the amplitude and frequency of the imposed stochastic oscillation. The process is stationary for $\rho \in [0, 1)$, with ρ influencing the regularity of the oscillation. Specifically, when $\rho = 0$, the process represents completely unpredictable white noise, while as ρ approaches 1, it becomes a highly predictable stochastic model characterised by significant oscillatory behaviour around the frequency f [40].

For this study, we have considered a two-pole AR polynomial with a modulus fixed at $\rho = 0.95$ varying the frequency f to simulate different types of noise, while the variance σ^2_U is determined by the SNR level, according to the procedure herein detailed. First, the power of the signal is calculated and used as a reference for scaling the power of the noise at frequency f as follows:

$$P_{noise} = \frac{P_{signal}}{10^{\frac{SNR}{10}}}. \quad (7)$$

Subsequently, given that the AR coefficients and the transfer function are related by $H(z) = \frac{1}{A(z)}$ and introducing the normalised angular frequency $\Omega = \frac{2\pi f}{f_s}$, the Spectral Factorisation Theorem [41] directly relates the PSD of the AR process to the variance of the Gaussian white noise:

$$P_{noise}(\Omega) = \left|H(\Omega)\right|^2 \sigma^2_U. \quad (8)$$

According to the spectral integration property, integrating a spectral measure over all frequencies provides the equivalent measure in the time domain. Therefore, the total power of the generated noise as a function of the variance of the Gaussian white noise is obtained, and the variance σ^2_U can be computed as follows:

$$\sigma^2_U = \frac{P_{noise}}{\frac{1}{\pi}\int_0^{\pi} \left|H(\Omega)\right|^2 d\Omega}. \quad (9)$$

The mathematical steps to derive Eq. (9) starting from the Eq. (8) are reported in Appendix A.2.

In this work, we have taken into account six frequencies (i.e. 0.01, 0.1, 0.3, 3, 50 and 300 Hz) to simulate different noise types. In detail, the frequency of $f = 0.01$ Hz simulates a slowly varying baseline at the lower end of the frequency spectrum typically present in ECG signals [42]. The frequencies $f = 0.1$ Hz and $f = 0.3$ Hz have been instead chosen to model the noise introduced by breathing [43], while the frequency $f = 3$ Hz simulates the effects on the ECG signal of atrial fibrillation, a cardiac arrhythmia characterised by a rapid and irregular electrical activity; this frequency has been selected in accordance to previous works using the 3 Hz dominant frequency as a threshold for distinguishing between normal sinus rhythm and arrhythmic activation rates [44,45]. The frequency $f = 50$ Hz corresponds to the interference of the powerline [9] and, finally, the frequency $f = 300$ Hz has been selected to simulate muscular noise, caused by electrical activity of the muscle [9]. Each type of noise was added to the noise-free ECG signals at five different SNR levels, i.e., −3 dB, 1 dB, 5 dB, 10 dB and 20 dB.

2.3. Time series extraction

Herein, the study has been carried out using 2-min ECG windows in accordance with the ultra-short term HRV analysis [12,17,46]. For each subject, the 2-min window of the reference signal was selected during stable physiological conditions, starting at least 3 min after the beginning of the recording, to favour stationarity of the time series. Starting from both the clean reference and the noise-corrupted ECG signals, a simplified version of the Pan-Tompkins algorithm [47] was used to identify R-peaks and to consequently extract the R-R interval time series (RRI). The working principle of this algorithm relies on the fact that the QRS complex is concentrated in the ECG frequency band between 5 and 15 Hz [47]. The algorithm consists of five main steps: bandpass filtering (5–15 Hz), 5-point derivative filter, squaring, moving window integration and peak detection. In the simplified version of the algorithm, a span value of $\frac{f_s}{7}$ samples was used for the moving window integration, while the peak detection was carried out using the MATLAB built-in *findpeaks* function with custom thresholds, i.e. a minimum temporal distance between the peaks kept fixed at 300 ms, and a minimum peak prominence defined as three times the median of the signal obtained after the previous processing steps. Following the application of the simplified Pan-Tompkins algorithm, the R-peaks were extracted from the clean and noisy ECG signals. Afterwards, the RRI time series were obtained computing the time interval between the nth and (n+1)th QRS apexes. A visual inspection was carried out by an expert researcher to assess the quality of the clean ECG waveforms in the selected window, to verify that filtering did not alter the morphology of the signals, and finally to ensure the stationarity of the obtained time series. The average length of the RRI time series extracted from noise-free signal was of 145 ± 18 beats.

Fig. 1(a) shows a 5-s window of the reference ECG signal for a sample subject alongside the RRI time series extracted from the entire 2-min recording. Similarly, Fig. 1(b) depicts 5-s windows of exemplary ECG signals corrupted by different types of noise at an SNR level of −3 dB for a sample subject alongside the RRI time series extracted from the entire 2-min recordings.

2.4. Time-domain analysis

In this study, we have first carried out a time-domain HRV analysis on noise-free and noise-corrupted ECG signals, computing the average value (MEAN), the standard deviation (SD) and the root mean square of successive differences (RMSSD) of the RRI time series [48]. As known, time-domain HRV metrics assess the variability of interbeat intervals using statistical measures, and in particular the MEAN is the basic cardiovascular measure reflecting the cardiac rhythm [17].

The SD of the RRI time series, expressed in milliseconds, serves as a comprehensive indicator of the multiple factors influencing heart rate variability [16,49]. This measure is particularly relevant because both

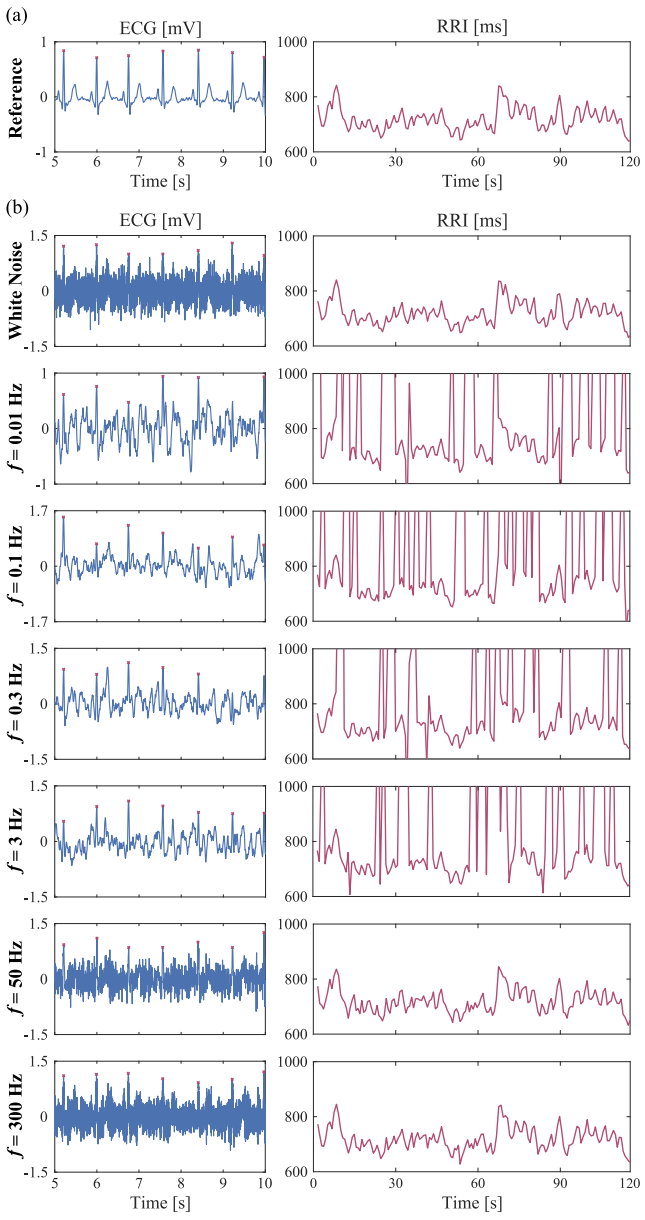

Fig. 1. (a) The reference ECG signal for a sample subject and (b) the ECG signals corrupted by different types of noise at an SNR level of −3 dB, alongside the RRI time series extracted from the entire 2-min recordings.

the SNS and PNS contribute to its variability, and it is also considered a gold standard measure in the medical stratification of cardiac risk, particularly in the context of 24-h recordings [29]. The formula for SD is given by [48]:

$$SD = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(x_n - \overline{x})^2}, \tag{10}$$

being x_n the nth RRI sample, $\overline{x}$ the average interval and N the time series length.

The RMSSD can be obtained by first squaring each successive time difference in milliseconds between heartbeats, then the result is averaged before computing the square root of the total [29,48].

$$RMSSD = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N-1}(x_{n+1} - x_n)^2}. \tag{11}$$

RMSSD is the primary time-domain measure used to estimate vagally-mediated changes in HRV and reflects the beat-to-beat variance in heart rate [49].

2.5. Frequency-domain analysis

Frequency-domain analysis relies on computation of the PSD, which provides essential information on how power (or variance) is distributed as a function of frequency [16]. Prior to performing frequency-domain analyses, the RRI time series were pre-processed by subtracting the mean value. Then, spectral analysis was performed employing the weighted covariance method to obtain the spectrum of each time series. In detail, the non-parametric Blackman-Tukey approach was applied by FFT-transforming (512 points) the truncated and windowed sample autocovariance of RRI data, using the Parzen window with bandwidth of $BW = 0.04$ Hz and setting the truncation lag M of the autocovariance sequence as $M = 1.273 f_{s_{RR}}/BW$ [50,51], where $f_{s_{RR}} = 1/\overline{RR}$ is the reciprocal of the mean RRI of the time series. $f_{s_{RR}}$ was computed for each subject assuming the series as uniformly sampled with a sampling period equal to the mean heart period. The Parzen window was used given its significantly lower side-lobe level compared to other common windows (e.g. Hanning and Hamming), while the values of BW and M were chosen in accordance to previous works on HRV spectral analysis [51].

The power values in the low-frequency band (LF, range 0.04–0.15 Hz), which can reflect both PNS and SNS activity, and in the high-frequency band (HF, range 0.15–0.40 Hz), which reflect parasympathetic activity and respiratory influence, were computed, both in absolute power (milliseconds squared) [16] and normalised (in percentage) dividing the power value in each band by the total power excluding the very-low-frequency range (VLF, 0.003–0.04 Hz) [16,17]. The normalised values of the LF and HF components allow to better highlight the balance of the ANS reducing the inter-subject variability in total power [16].

In this study, we have taken into account the LF absolute power (LF) and the normalised value (LF_{norm}), together with the low-to-high frequency power ratio (LF/HF), commonly used as an index of the balance between SNS and PNS activity under controlled conditions [17].

2.6. Information-domain analysis

Information-theoretic analysis was performed to compute indices quantifying the information content and complexity of the analysed time series. Before evaluating the information-domain indices, all the time series were pre-processed using an IIR zero-phase high-pass AR filter [52], with cut-off frequency of 0.0156 times the equivalent sampling rate of the series ($f_{s_{RR}}$), removing the slow trends to improve stationarity, and then normalised to zero mean and unit variance.

The entropy (H) and the conditional entropy (CE) measures were computed on RRI time series obtained from both noise-free and noise-corrupted ECG signals. These measures reflect, respectively, the average amount of information contained in the current state of the process X and the average uncertainty that remains about the present state of the process when its past states are known [28,53]. To define the measures, let us denote X_n as the random variable obtained by sampling the process X at the present time n, and $\mathbf{X}_n^m = [X_{n-1}, X_{n-2}, \ldots, X_{n-m}]$ as the vector variable describing the past m states of X. Formally, H and CE are defined under the hypothesis of stationarity as follows [28]:

$$H = H(X_n) = -E[\log p(x_n)], \tag{12}$$

$$\begin{aligned} CE &= H(X_n|X_n^m) = H(X_n, X_n^m) - H(X_n^m) \\ &= -E[\log p(x_n|x_{n-1}, x_{n-2}, \ldots, x_{n-m})], \end{aligned} \tag{13}$$

where $E[\cdot]$ represents the expectation operator, and $p(\cdot)$ is the probability density function.

In this study, the H and CE measures were estimated using a model-free approach based on the k-nearest neighbour (*knn*) method. This approach exploits the intuitive concept that the local probability density around a given data point is inversely related to the distance between the point and its neighbours. Estimates of H and CE for the process X can be computed through the following formulations [40]:

$$H_{knn} = \psi(N) - \psi(k) + \langle \ln \epsilon_n \rangle, \tag{14}$$

$$CE_{knn} = -\psi(k) + \langle \psi(N_{X_n^m}) \rangle + \langle \ln \epsilon_n \rangle, \tag{15}$$

where $\psi(\cdot)$ denotes the digamma function, k is the number of neighbours selected for the analysis, and ϵ_n represents twice the distance between the nth realisation of (X_n, X_n^m) and its kth nearest neighbour, computed using the Chebyshev distance (maximum norm, i.e., the maximum distance between the scalar components). $N_{X_n^m}$ represents the number of points within a distance from X_n^m smaller than $\frac{\epsilon_n}{2}$ and $\langle \cdot \rangle$ is the average operator. The average is taken over all the $N - m$ realisations of the patterns (X_n, X_n^m) extracted from a time series of length N. In our analysis, the number of neighbours was set at $k = 10$, while the number of past components considered for the time series was set equal to $m = 3$, according to previous works [54].

In addition, the CE was estimated using a linear parametric approach (*lin*). This method assumes that the observed process X follows a Gaussian distribution, which is a reasonable assumption for many physiological processes. Under this assumption, the CE can be estimated by modelling the dynamics of the process X using an AR model (Eq. (5)), which expresses the current state of the process in terms of its past states. The AR model was identified using ordinary least squares method to estimate the regression parameters and the variance of the residuals $\hat{\sigma}_U^2$, from which the CE measure can be computed as [40]:

$$CE_{lin} = \frac{1}{2} \ln\left(2\pi e \hat{\sigma}_U^2\right), \tag{16}$$

being e the Euler's number. The optimal order of the AR model for the *lin* approach was determined using the Bayesian Information Criterion [55].

2.7. Statistical analysis

The main aim of the analyses carried out in this work was to assess to what extent the various types and levels of noise affect HRV measures. Statistical analyses were performed on the distributions of time-domain, frequency-domain, and information-domain HRV indices computed on the RRI time series. In detail, the distribution of values across the 17 subjects obtained from the noise-free (reference) signal was compared to that computed on the noisy signal, for each type of noise and SNR level. The pairwise comparisons were conducted using the non-parametric Wilcoxon signed-rank test [56], with a significance threshold of $p\text{-}value < 0.05$.

Moreover, since reliance on statistical tests alone is often considered not enough for assessing the reliability of HRV indices derived from different techniques or signals [13,29,57], a correlation analysis was also performed to achieve a further indication of the extent of the agreement between the two distributions taken into account [58]. To this aim, the non-parametric Spearman's rank correlation coefficient (R_s) was computed to assess the strength of the monotonic relationship between the distributions of HRV measures computed from the noise-free and the noisy signals, with a statistical significance level of $p\text{-}value < 0.05$ [59]. In addition, the effect size between the two distributions has been computed, in order to provide an indication of the significance of observed effects that may be independent of the sample size [13,60]. To this purpose, the Cohen's d index was calculated as the difference between the means of the two distributions divided by the pooled standard deviation [61]:

$$d = \frac{\mu_1 - \mu_2}{\sqrt{\frac{(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2}{n_1 + n_2 - 2}}} \tag{17}$$

where μ, σ^2 and n are the mean, the variance and the number of samples of the compared distributions (1 and 2 in the subscript, i.e. HRV indices from noisy and noise-free signals). Typically, the effect size is considered small for d lower than 0.2, medium when around 0.5, and large if d is greater than 0.8 [61].

Finally, Bland-Altman analysis was used [62] to assess the accordance in terms of the differences between HRV metrics computed from noisy and noisy-free signals (in the y axis) plotted versus their average (in the x axis), estimating the bias and quantifying a measure of agreement as the ratio between half of the 95% confidence interval for the difference (i.e. $1.96\,\sigma_{\text{diff}}$) and the mean of the averaged values (μ_{mean}) [63]:

$$agreement = \frac{1.96\,\sigma_{\text{diff}}}{\mu_{\text{mean}}} \tag{18}$$

Higher values of this measure reflect a poorer agreement, thus evidencing a stronger influence of the noise on the HRV indices.

3. Results

Results are presented separately for time-domain (Fig. 2), frequency-domain (Fig. 3), and information-domain indices (Fig. 4). The upper part of each panel shows boxplots distributions and individual values of the given index, with an asterisk marking statistically significant differences ($p\text{-}value < 0.05$) compared to the reference RRI time series, according to the Wilcoxon test. The lower part of each panel reports the Spearman's rank correlation coefficient, highlighted in red if significant ($p\text{-}value < 0.05$) and in black if not significant, alongside the Cohen's d value. The results for SNR equal to 20 dB are not reported because no significant differences were observed under any noise conditions, the Pearson correlation coefficient was always equal to 1 and the Cohen's d index was 0.

Fig. 2 compares the distributions of the time-domain indices (MEAN, SD and RMSSD) calculated over the RRI time series of the 17 subjects obtained from both noise-free (reference) and noise-corrupted ECG signals with different types of noise at various SNR levels. Overall, the analysis reveals that a decrease of SNR ($\leq$1 dB) causes a reduction of the Spearman's rank correlation coefficient and statistically significant differences of the obtained measures compared to the reference, especially for noise at lower frequencies ($f \leq 3$ Hz). In detail, a statistically significant increase of MEAN (Fig. 2(a)), SD (Fig. 2(b)) and RMSSD (Fig. 2(c)) indices has been detected for a SNR of −3 dB for noise at frequencies of 0.01, 0.1, 0.3 and 3 Hz, with low ($\leq$0.62) values of the correlation coefficient (not significant for SD and RMSSD) and large effect size. With regard to SD and RMSSD, this trend is also observed for noise at powerline frequency (i.e., 50 Hz) with medium effect size. Interestingly, the MEAN index remains unaffected by white noise, showing no statistically significant difference and a perfect correlation ($R_s = 1.00$). With regard to SD and RMSSD, statistically significant differences are observed also for white noise and noise at 300 Hz, but still with a very high correlation coefficient ($R_s \geq 0.88$) and a medium-small effect size ($|d| \leq 0.31$). When decreasing the noise intensity to an SNR = 1 dB, similar trends can be found, with however high and statistically significant values ($R_s > 0.80$) of correlation coefficient for MEAN index. Moreover, almost perfect (and statistically significant) correlation and very low Cohen's d values are obtained also for SD and RMSSD in case of 50 Hz noise. A perfect agreement (i.e., no statistically significant differences and $R_s = 1.00$) between distributions for the MEAN index is observed at a SNR of 5 dB, while statistically significant differences can still be found for SD and RMSSD indices. When the signal quality increases (i.e., SNR = 10 dB), the distributions for MEAN index show an almost perfect agreement with the reference ($R_s = 1.00$ and no statistically significant differences). However, for SD and RMSSD, sporadic statistically significant differences can still be observed (mainly due to some wrongly detected consecutive R-peaks) but the correlation coefficient is 1.00 and the effect size negligible ($|d| \leq 0.04$).

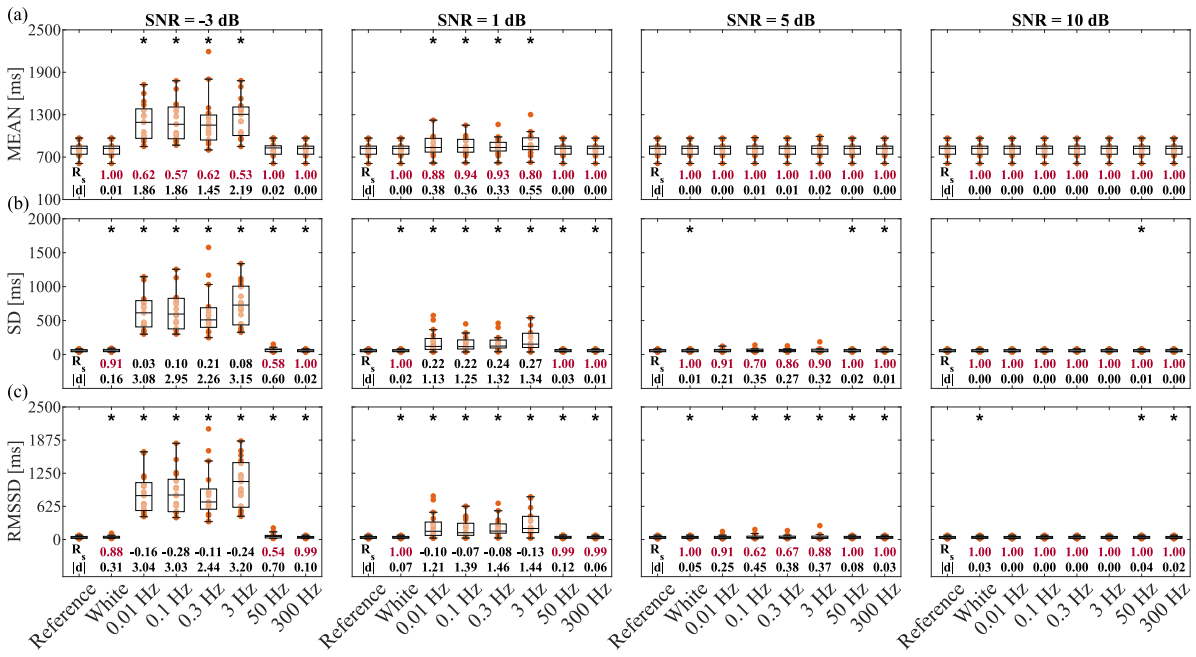

Fig. 2. Results of time-domain HRV analysis: (a) MEAN, (b) SD and (c) RMSSD. The upper part of each panel shows boxplot distributions and individual values of the given index, with an asterisk marking a statistically significant difference ($p\text{-}value < 0.05$) compared to the reference RRI time series, according to Wilcoxon test. The lower part of each panel reports the Spearman's rank correlation coefficient – highlighted in red if significant ($p\text{-}value < 0.05$) and in black if not significant – and the Cohen's d index.

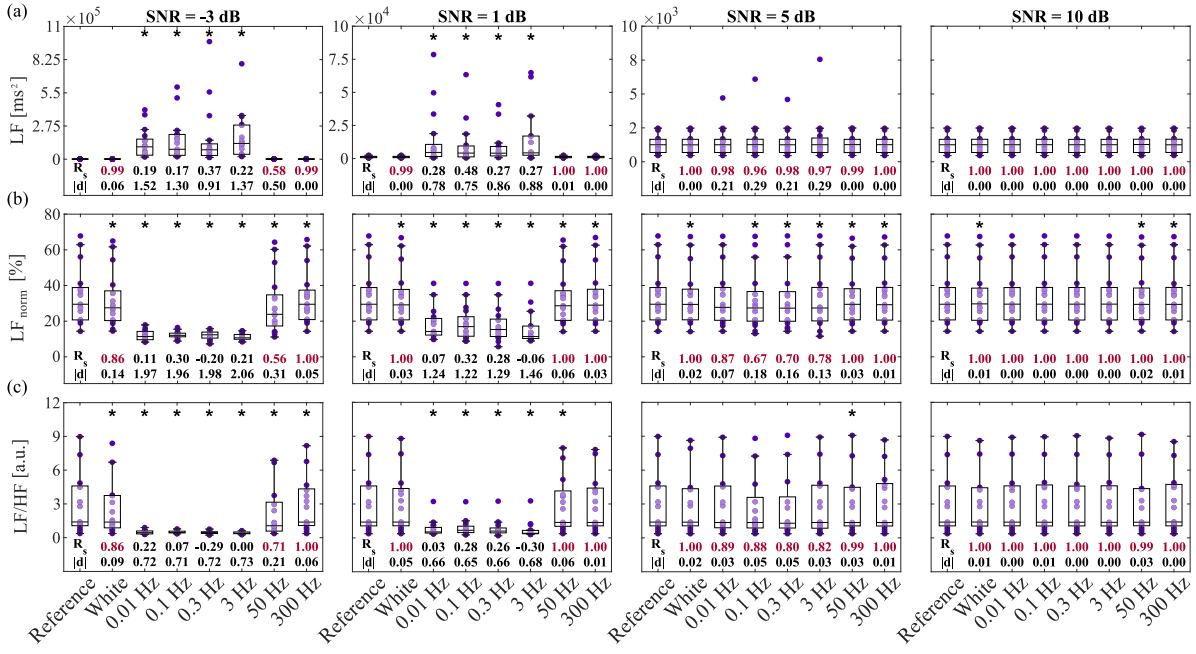

Fig. 3. Results of frequency-domain HRV analysis: (a) LF, (b) LF_{norm} and (c) LF/HF. The upper part of each panel shows boxplot distributions and individual values of the given index, with an asterisk marking a statistically significant difference ($p\text{-}value < 0.05$) compared to the reference RRI time series, according to Wilcoxon test. The lower part of each panel reports the Spearman's rank correlation coefficient – highlighted in red if significant ($p\text{-}value < 0.05$) and in black if not significant – and the Cohen's d index.

Fig. 3 illustrates the results of the frequency-domain analysis, i.e., the LF, LF_{norm} and LF/HF indices derived from the RRI time series of the 17 subjects, using both noise-free (reference) and noise-corrupted ECG signals across various SNR levels and noise types. Similar to the time-domain analysis, a decrease in SNR ($\leq$1 dB) leads to a decline in the Spearman's rank correlation coefficient and statistically significant differences in the measures compared to the reference, particularly for low-frequency noise ($f \leq 3$ Hz). Specifically, at an SNR of −3 dB, a statistically significant increase in the LF index (Fig. 3(a)), and a decrease in both LF_{norm} (Fig. 3(b)) and LF/HF (Fig. 3(c)) indices have been observed for noise at frequencies of 0.01, 0.1, 0.3 and 3 Hz, with a low ($\leq$0.37) and not significant correlation coefficient. For noise at 50 Hz frequency, no significant differences are found in LF, while LF_{norm} and LF/HF show significant variations. In these cases, the correlation coefficient is in the range of 0.56–0.71 and the Cohen's effect size is medium. For both white noise and for noise at 300 Hz, no significant differences are observed in the LF index, with perfect agreement ($R_s = 0.99$). Although LF_{norm} and LF/HF show statistically significant differences, the correlation coefficient remains very high ($R_s \geq 0.86$) and the effect size is small. When the noise level is

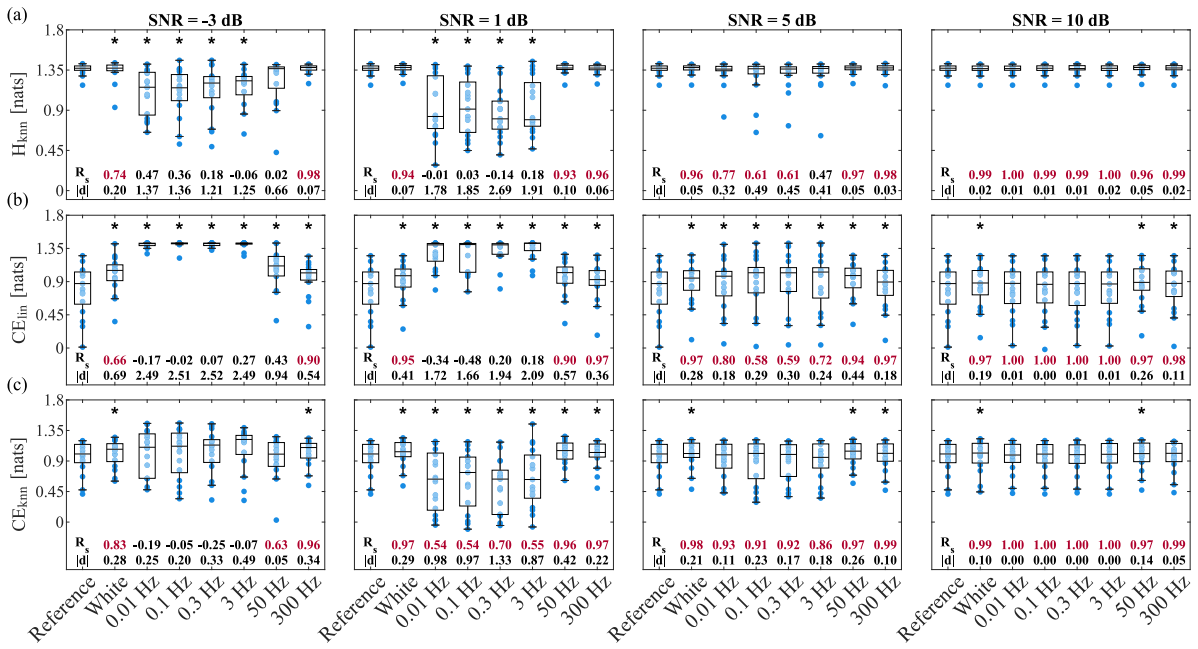

Fig. 4. Results of information-domain HRV analysis: (a) H_{knn}, (b) CE_{lin} and (c) CE_{knn}. The upper part of each panel shows boxplot distributions and individual values of the given index, with an asterisk marking a statistically significant difference (p-$value < 0.05$) compared to the reference RRI time series, according to Wilcoxon test. The lower part of each panel reports the Spearman's rank correlation coefficient – highlighted in red if significant (p-$value < 0.05$) and in black if not significant – and the Cohen's d index.

reduced to SNR = 1 dB similar trends are noted, while as the signal quality improves (i.e., SNR = 5 dB and 10 dB), perfect agreement is achieved for all three indices, with no significant differences and a high Spearman's rank correlation coefficient ($R_s \geq 0.80$), except for LF_{norm}, which exhibits some statistically significant differences, although with a high Spearman's rank correlation coefficient and a small effect size.

Fig. 4 shows the results of the information-domain analysis, where the indices H (using the *knn* approach) and CE (using both *lin* and *knn* estimators) were computed on the RRI time series from both reference and noise-corrupted ECG signals with different types of noise at various SNR levels. Similarly to previous results, a decrease in SNR (≤1 dB) produces a reduction in the Spearman's rank correlation coefficient and statistically significant differences in the obtained measures compared to the reference, especially for lower frequencies noise (0.01 Hz to 3 Hz). In particular, a statistically significant decrease in the H_{knn} index (Fig. 4(a)) and an increase in the CE_{lin} index (Fig. 4(c)) have been detected for SNR of −3 dB for noise at frequencies of 0.01, 0.1, 0.3 and 3 Hz, with low (<0.5) and not significant Spearman's rank correlation coefficient. Similar trends are also reported for CE_{knn} index (Fig. 4(b)) though the differences are not significant. For noise at 50 Hz, only CE_{knn} shows together no statistically significant variations, a moderate correlation ($R_s = 0.63$) and a negligible effect size ($|d| = 0.05$). In the presence of noise at 300 Hz, only H_{knn} shows no statistically significant variations, together with a very high correlation coefficient ($R_s = 0.98$) and a small effect size. Similar trends can be found for SNR = 1 dB with regard to H_{knn} and CE_{lin}, but in this case with also statistically significant differences for CE_{knn}. At an SNR of 5 dB, H_{knn} shows no statistically significant differences across all noise conditions, but only in case of white noise and noise at 50 and 300 Hz the correlation coefficient is high ($R_s \geq 0.96$) and the effect size small. With regard to CE_{knn} and CE_{lin}, even if statistically significant variations are reported for white noise and for noise at 50 and 300 Hz, similar remarks in terms of high correlation coefficients and small effect size can be made. As the signal quality improves to SNR = 10 dB, perfect agreement is observed for all indices (i.e., no statistically significant differences and $R_s \geq 0.96$), except for sporadic significant differences in CE_{lin} and CE_{knn}. However, in all cases, correlation remains high ($R_s \geq 0.97$), and effect sizes are small ($|d| \leq 0.26$).

Bland-Altman analysis confirms the previously described findings. Fig. 5 presents exemplary Bland-Altman plots for the MEAN, LF_{norm} and CE_{knn} indices in presence of 0.1 Hz noise at an SNR level of −3 dB, 1 dB and 5 dB, depicting the differences between metrics computed for each subject from noisy and noise-free signals versus their average. The dashed lines represent the bias, i.e., the mean of the differences between the HRV values, while the agreement values, computed according to Eq. (18), are indicated on the bottom-right of each panel. As reported in the figure, increasing the SNR results in a decrease of the bias (in absolute value) and of the agreement measure, hence reflecting a better agreement between metrics computed from noisy and noise-free signals, and thus a weaker influence of the noise.

The complete results obtained from Bland-Altman analysis are summarised in Figs. 6 and 7, respectively reporting the bias and agreement values for all the time-, frequency- and information-domain HRV indices, as a function of the SNR level. Results are graphically reported up to a SNR level of 10 dB, given that for SNR = 20 dB zero bias and agreement have been achieved. With regard to time-domain indices, results show that noise at an SNR of −3 dB has a strong effect on all HRV measures, with the exception of MEAN index in case of white noise or noise at 50 Hz or 300 Hz (bias lower than 2 ms and agreement values ≤0.02). Also SD in case of 300 Hz noise presents a very low bias (around 0.4 ms) and agreement of around 0.02. On the contrary, low-frequency noise (≤3 Hz) causes the worst corruption of the indices, with unacceptable values of bias and agreement (bias > 370 ms for MEAN, >560 ms for SD, and >820 ms for RMSSD; agreement >2 in the worst cases). As SNR increases to 1 dB, noise effects are reduced with lower bias, although low-frequency noise continues to affect SD and RMSSD (agreement >2). The effects of white, 50 Hz, and 300 Hz noise become negligible also for SD index (bias < 1 ms and agreement around 0.01) and almost for RMSSD (bias of < 3 ms and agreement of around 0.06). Finally, it is important to highlight that the high bias values reported for very low SNR values (<1 dB) can be related to undetected R-peaks in noisy ECG tracks. In detail, the strong noise generates errors in peak detection and, in particular, missing beats that cause longer R-R intervals, in turn producing higher MEAN, SD and RMSSD values; the more peaks are undetected, the more this behaviour is pronounced, thus leading to a systematic increase in the difference between the

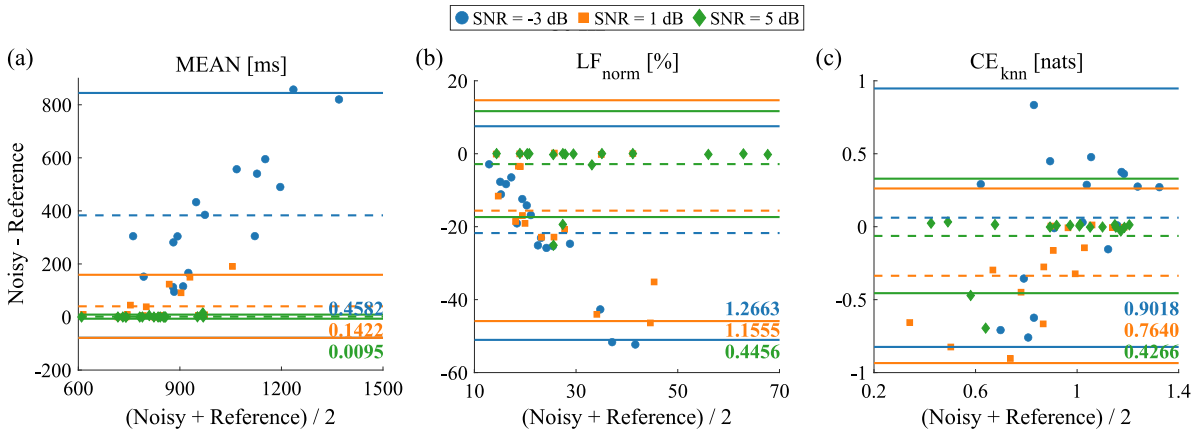

Fig. 5. Bland-Altman plots of (a) MEAN, (b) LF_{norm} and (c) CE_{knn} indices in case of 0.1 Hz noise at an SNR level of −3 dB, 1 dB and 5 dB. Each panel shows the difference between noisy and noise-free index values against their averages. The horizontal dashed lines indicate the bias, while the solid lines denote the 95% limits of agreement. The computed agreement values are reported on the bottom-right of each panel.

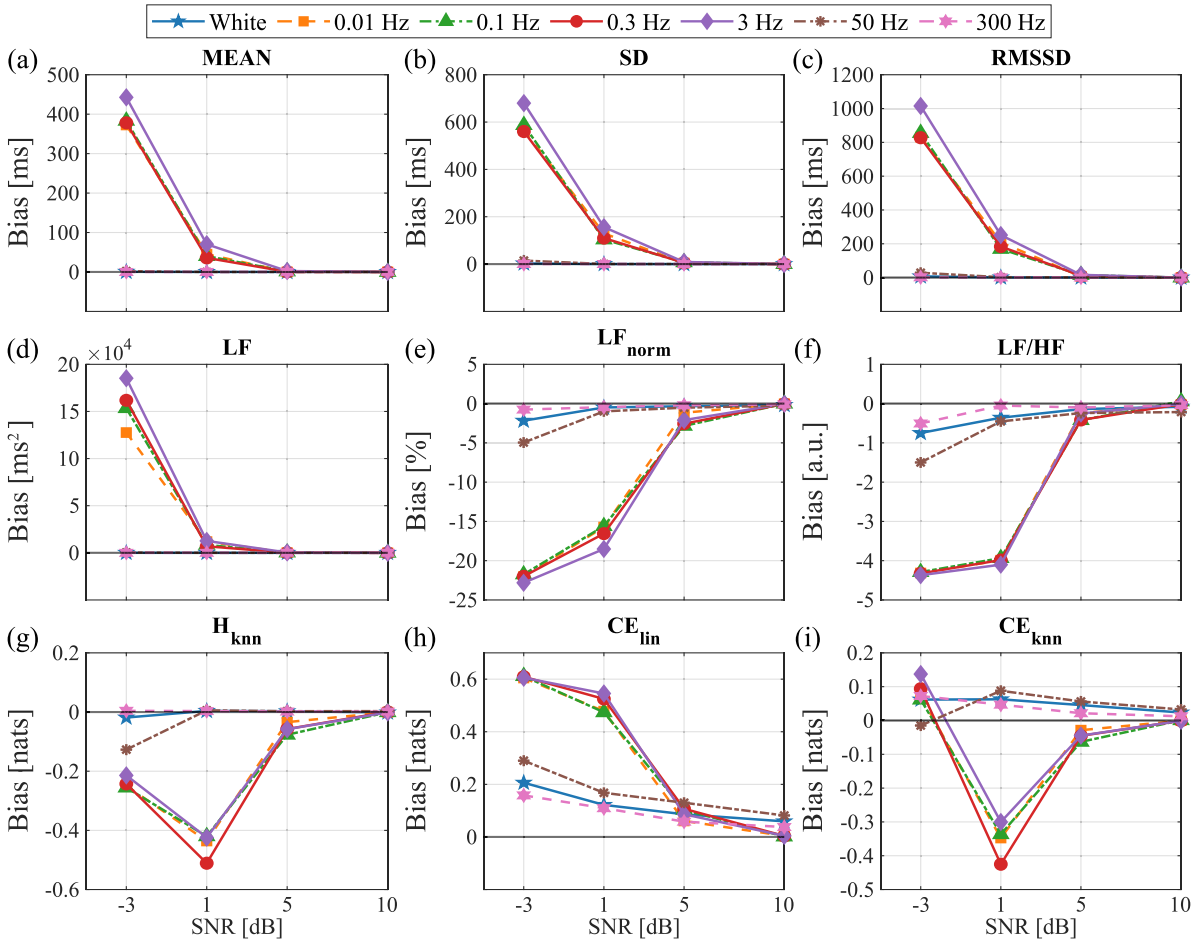

Fig. 6. Bias values computed from Bland-Altman analysis as a function of the SNR level for the (a), (b), (c) time-domain, (d), (e), (f) frequency-domain, and (g), (h), (i) information-domain HRV indices. Each colour indicates a different type of noise.

noisy and noise-free indices. At 5 dB, MEAN shows a perfect agreement (no bias and agreement index almost null), while SD and RMSSD are still affected by low-frequency noise. At 10 dB, all time-domain indices show excellent agreement with the reference, with near-zero bias and agreement values, regardless of noise type.

Similar remarks apply to the Bland-Altman analysis performed on frequency-domain indices, which confirm also in this case the previous findings. Overall, for low SNR values the frequency-domain indices present a poorer agreement if compared to time-domain metrics. In detail, at an SNR of −3 dB noise produces an unrecoverable degradation of all the indices, except for LF and LF_{norm} for noise at 300 Hz (with bias respectively of around 1 ms^2 and −1% with agreement ≤0.05). At an SNR of 1 dB, frequency-domain indices are still completely unreliable in case of low-frequency noise. On the contrary, white and ≥50 Hz noise cause an almost negligible decrease of reliability of LF and LF_{norm} (with bias respectively of around ±5 ms^2 and −1% with agreement

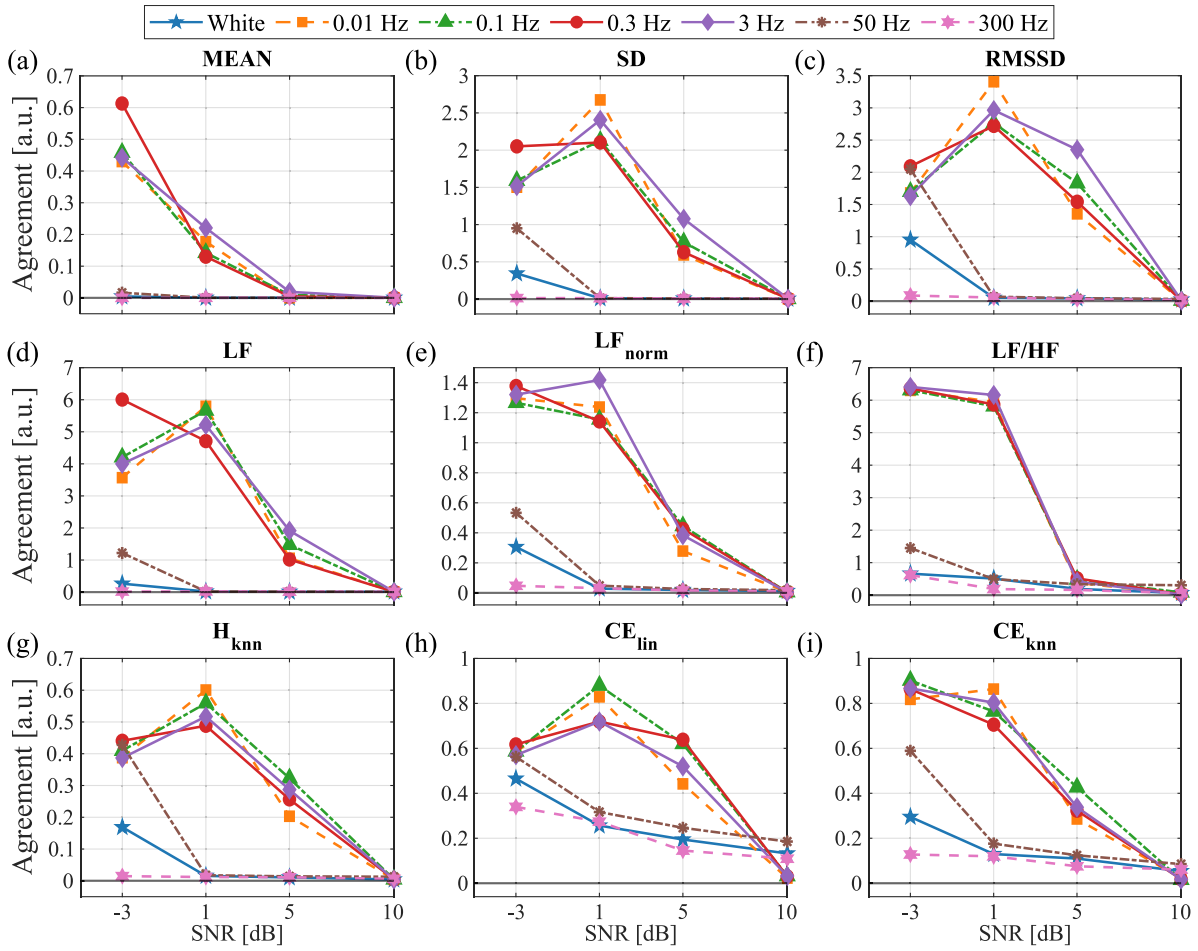

Fig. 7. Agreement values computed from Bland-Altman analysis as a function of the SNR level for the (a), (b), (c) time-domain, (d), (e), (f) frequency-domain, and (g), (h), (i) information-domain HRV indices. Each colour indicates a different type of noise.

≤0.05). Similar remarks can be made for SNR = 5 dB, with an overall decrease of bias and an overall better agreement. Increasing the SNR to 10 dB, all frequency-domain indices show highly reliable behaviour with negligible bias for all noise types.

The Bland-Altman analysis performed on information-domain HRV indices again confirms the previous results. A noise at a SNR of −3 dB produces an unrecoverable degradation of all the indices, apart from H_{knn} only in case of 300 Hz noise (bias of around 0.004 nats and agreement of 0.01). At 1 dB SNR, noise effects remain noteworthy in terms of bias and agreement, with only H_{knn} exhibiting robust performance under white, 50 Hz and 300 Hz noise (bias < 0.006 nats, agreement < 0.02). At 5 dB, the influence of noise decreases further, although the values of bias and agreement remain quite high for CE indices. Finally, at 10 dB SNR, noise has a negligible influence across all information-domain indices, with bias values lower than 0.03 nats and agreement ≤0.08 for H_{knn} and CE_{knn}; CE values estimated with linear approach exhibit a worse performance with higher bias if compared to model-free method.

4. Discussion

The aim of this study was to comprehensively assess to what extent different types of electrocardiographic noise at various SNR levels influence the reliability of ultra-short term indices of HRV measured in the time, frequency and information domains, using statistical tests, correlation analysis, Cohen's *d* effect size and Bland-Altman method. In general, it is important to emphasise that the presence of statistically significant differences between noise-free and noisy distributions, complemented by high correlation coefficient and small effect size hints that the influence of noise may not have a critical impact on the validity of the metrics. Bland-Altman analysis allows to add further information, evidencing as well the bias and the level of agreement between the distributions.

Our results evidence that a minimum SNR level of 10 dB is needed to obtain reliable HRV estimates, as it ensures near-zero bias and agreement values for all indices, whatever is the frequency of the noise. This finding is supported by statistical tests, correlation analysis (Spearman's coefficients always higher than 0.96) and small effect sizes according to Cohen's *d*. Reducing the SNR level to 5 dB already causes a worse agreement for most indices, even if a reliable estimation of MEAN, LF/HF and CE_{knn} is still feasible, given the correlation values higher than 0.80 and the small effect size for most noise types. For the other indices, instead, lower correlation values, medium effect sizes, and higher bias have been observed for noise overlapping to cardiovascular frequencies (≤3 Hz). These finding suggest that caution should be exercised when computing HRV measures on cardiovascular time series affected by noise around 0.1–10 Hz and in case of pathological conditions such as atrial fibrillation that alter the normal heart rhythm [44].

Reducing the SNR level to ≤1 dB causes not only an overall worsening of the reliability of the computed indices, but also a strongly frequency-dependent behaviour. In particular, low-frequency noise (i.e., 0.01, 0.1, 0.3, and 3 Hz) resulted in a strong decrease in correlation (R_s ≤ 0.8) and increase in agreement value (often even ≥ 1) for all the indices taken into account, except for the MEAN index that still maintains an acceptable reliability for an SNR equal to 1 dB. This behaviour at low frequencies is somewhat expected, given the

overlap of the noise with physiological mechanisms operating within these frequency ranges, e.g., baroreflex within LF band [64,65] or respiration within HF band, heartbeat around 1 Hz (or few hertz in case of arrhythmias like atrial fibrillation [44]).

Powerline noise (i.e., 50 Hz), representing the most common type of disturbance in ECG signals, caused a strong reduction of reliability for a SNR of −3 dB, causing correlation indices lower than 0.7 and poor agreement values according to Bland-Altman (often even $\geq$ 1) for all HRV indices except for the MEAN. These remarks evidence that even in case of a very strong noise level, the powerline noise does not irremediably alter the validity of the MEAN index (commonly considered the basic cardiovascular measure and widely used to compute the heart rhythm). At higher SNR levels (1 and 5 dB) for 50 Hz, the correlation with the reference signals remained above 0.90, suggesting that the overall information contained in the ECG signal remains largely unchanged in case of powerline noise, even if LF/HF and CE measures seem to be more strongly affected according to Bland-Altman analysis. It is worth underlying that this is valid when using robust peak detection methods, as the Pan-Tompkins algorithm [47] in this work. Moreover, this type of noise can be filtered more easily than low-frequency noise, whose band overlaps with that of the signal of interest.

Our results evidence that muscle-like noise (at higher frequencies, $\geq$300 Hz) did not alter the reliability of most HRV indices even for the worst cases of SNR $\leq$1 dB, with a very high correlation compared to the noise-free measures, as also confirmed by the Bland-Altman analysis. These finding suggest that high-frequency noise has a relatively minor effect on HRV measures computed from ECG signals. This is important for practical applications, since muscular noise can be acquired together with useful ECG signal, but being its frequency band much higher and farther away from that of the QRS complex, the effect can be neglected and its filtering is easier as well. Similarly, also white noise has a relatively minor effect on HRV features even for very low SNR levels ($\leq$1 dB), even if the impact seems to be overall a bit stronger than high-frequency noise, evidenced by slightly lower correlation values and higher agreement index according to Bland-Altman analysis. Nonetheless, although white noise is more difficult to be filtered because its spectrum also overlaps with that of the QRS complex, its effect does not appear to irremediably affect the HRV indices, again when a robust peak detection algorithm is exploited.

It is worth noting that noise has a noticeable effect for entropy measures also at 5 dB SNR, with lower correlation coefficients and higher effect sizes, especially in case of linear estimator, suggesting that information-theoretic measures are more affected than time- and frequency-domain indices. Nonetheless, the model-free *knn* estimator appears to offer greater robustness to noise. Moreover, our results confirm previous findings showing that CE_{lin} tends to overestimate the results if compared to CE_{knn}. This behaviour has been observed in previous studies using cardiovascular data similar to those considered in this work [13], as well as in theoretical simulations [54]. The reasons for such a difference are difficult to explain and may be related to several factors. The model-free *knn* method has the advantage of capturing non-linear mechanisms commonly occurring in cardiovascular dynamics, which are instead neglected by linear model-based approaches. However, the underestimation observed with *knn* can be attributed to the presence of local non-linearities and/or non-stationarities, whereas it is not possible to exclude as well the effect of a bias in the non-linear measures obtained using the *knn* estimator, due to the difficulty of working on high-dimensional spaces [66].

A limitation of this study consists in the use of fixed custom thresholds for the R-peak detection algorithm, as described in Section 2.3. This has been done on purpose to allow a direct comparison of the results among different noise types and levels. Nonetheless, better results could be obtained using adaptive thresholds for the span of the moving window, for the minimum peak prominence and for the minimum temporal distance between consecutive peaks (required also for correctly detecting peaks in case of ectopic/missing beats or very irregular cardiac rhythm). Although the Pan-Tompkins algorithm has been applied in this study thanks to its widespread use and simplicity (especially in the implemented simplified version), it is worth underlining that alternative peak detection algorithms (such as the Ensemble Empirical Mode Decomposition [67]) may provide even greater robustness in noisy environments and should be taken into account in future studies. Moreover, it is important to point out that, although the present study used ECG signals artificially contaminated by continuous noise with SNR levels even lower than 5 dB, this methodological choice may not fully reflect the typical conditions of wearable devices, where noise is usually transient and movement-related. However, this simulation allowed us to explore the robustness of HRV indices by analysing possible worst-case scenarios.

Finally, it should be noted that the comparison of the results of this study with those reported in the literature is not feasible given the lack of works evaluating the effect of different types of noise on HRV indices. Nonetheless, our results are in agreement with previous works reporting the negative effect of white and/or pink noise on the quality of ECG signal or on HRV features [9,32,33], even if a direct comparison cannot be carried out given the use of different time series length and study protocols. The analysis conducted in our work comparing noise at different frequencies is able to provide further insights on the magnitude of the distorting effects on cardiovascular variability metrics.

5. Conclusion

This study investigated on the effect of different types and levels of ECG noise on conventional and entropy-based ultra-short term HRV measures. Our results evidenced that low-frequency noise (below 3 Hz) with a SNR lower or equal to 5 dB has a significant negative impact on the reliability of HRV indices, while the effect of white and higher-frequency noise (powerline and 300 Hz) is minor when using robust peak detection algorithms. Moreover, complexity measures (especially when computed through linear-based approaches) appear to be more strongly affected than time- and frequency-domain indices at higher SNR levels.

Overall, a SNR of 10 dB has been demonstrated to be enough to guarantee a very good reliability with respect to noise-free reference as confirmed by statistical tests, correlation analysis, Cohen's *d* effect size, and Bland-Altman method. These findings are important for improving the reliability of HRV analysis in real-world applications when using wearable devices for continuous health monitoring during daily-life activities or especially in noisy or extreme environments.

Although the obtained results are robust, it is important to note that the analysis took into account a relatively small number of subjects, which may limit the generalisability of the conclusions. Moreover, since multiple types of noise are often simultaneously present during ECG signal acquisition, future research should explore the effects of more complex noise patterns (e.g., within different frequency ranges) in order to provide a more thorough understanding of the challenges faced in daily-life monitoring. Finally, it would be important to carry out the same analyses on pulse rate variability indices computed from PPG signals [15,17,29], given that wearable devices usually rely on photoplethysmographic technique being easier and less expensive than ECG, although being more prone to movement artifacts.

CRediT authorship contribution statement

Anna Raimondi: Writing – original draft, Visualization, Software, Investigation, Formal analysis. **Alessandro Busacca:** Writing – review & editing, Resources, Investigation, Funding acquisition, Data curation. **Giuseppe Costantino Giaconia:** Writing – review & editing, Investigation, Funding acquisition. **Yuri Antonacci:** Writing – review &

editing, Investigation, Data curation. **Salvatore Stivala:** Writing – review & editing, Validation, Investigation. **Luca Faes:** Writing – review & editing, Supervision, Software, Methodology, Funding acquisition, Conceptualization. **Riccardo Pernice:** Writing – review & editing, Supervision, Software, Methodology, Investigation, Funding acquisition, Conceptualization.

Ethics statement

The study was conducted in accordance with the declaration of Helsinki and was approved by the Bioethics Committee of the University of Palermo (ref. no. 125/2023). Informed consent was obtained from all subjects involved in the study.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU: partnership on “Telecommunications of the Future” (PE00000001-program “RESTART”), projects S10 - SEXTET, CUP E63C22002070006, and by SiciliAn MicronanOTecH Research And Innovation CEnter “SAMOTHRACE”, Italy (MUR, PNRR-M4C2, ECS 00000022), spoke 3 - Università degli Studi di Palermo “S2-COMMs - Micro and Nanotechnologies for Smart & Sustainable Communities”.

Appendix

A.1. Autoregressive model

In this section, we describe the procedure to derive the autoregressive (AR) model coefficients in terms of its pole locations starting from the parametric representation of the process. Let us consider a generic zero-mean scalar random process X with variance σ_X^2 and identify the following AR model [68]:

$$X_n = \sum_{m=1}^{p} a_m X_{n-m} + U_n, \tag{A.1.1}$$

where p denotes the model order, defining the maximum lag used to quantify interactions, X_n represents the present state of the process X, a_m are the AR model coefficients relating the present with the past of the process at lag m, and U_n is a zero-mean white process, i.e., a process composed by independent and identically distributed variables with variance σ_U^2.

The AR model can be suitably represented in the z-domain. To this end, the z-transform of Eq. (A.1.1) is computed to obtain:

$$X(z) = \left(1 - \sum_{m=1}^{p} a_m z^{-m}\right)^{-1} U(z) = A(z)^{-1} U(z), \tag{A.1.2}$$

where $X(z)$ and $U(z)$ are the z-transform of X_n and U_n, respectively, and $A(z) = 1 - \sum_{m=1}^{p} a_m z^{-m}$ is the characteristic polynomial, with $z = e^{j2\pi \frac{f}{f_s}}$ and $j = \sqrt{-1}$, being f_s the sampling frequency of the process.

Considering an AR process of order $p = 2$, the characteristic polynomial is:

$$A(z) = 1 - a_1 z^{-1} - a_2 z^{-2}. \tag{A.1.3}$$

An AR process of order $p = 2$ is simulated by placing two complex-conjugate poles (i.e., the roots of the characteristic polynomial) at $z_1, z_2 = \rho e^{\pm j2\pi f}$, where ρ is the modulus and $\pm 2\pi f$ is the phase. Factoring the polynomial in terms of its roots gives:

$$A(z) = (1 - z_1 z^{-1})(1 - z_2 z^{-1}) = 1 - (z_1 + z_2) z^{-1} + z_1 z_2 z^{-2}. \tag{A.1.4}$$

Comparing Eq. (A.1.4) with the original form in Eq. (A.1.3), the AR coefficients can be directly identified:

$$a_1 = z_1 + z_2 = \rho e^{j2\pi f} + \rho e^{-j2\pi f} = 2\rho \cos(2\pi f), \tag{A.1.5}$$

$$a_2 = -z_1 z_2 = -\rho^2 e^{j2\pi f} e^{-j2\pi f} = -\rho^2. \tag{A.1.6}$$

A.2. Spectral factorisation

In this section, we explain how spectral factorisation can be used to obtain the PSD of the resulting stochastic process, and how this frequency-domain representation can be integrated to recover the total power of the process.

Starting from the z-domain representation of the AR model in Eq. (A.1.2), and recalling the definition of the transfer function for a linear time-invariant system, the transfer function is given by:

$$H(z) = \frac{1}{A(z)}. \tag{A.2.1}$$

Computing $H(z)$ on the unit circle in the complex plane - i.e., by setting $z = e^{j\Omega}$, where $\Omega = \frac{2\pi f}{f_s}$ is the normalised angular frequency ranging in $\Omega \in [-\pi, \pi]$ - and exploiting spectral factorisation, it is possible to derive the power spectral density of the stationary random process X:

$$P_X(\Omega) = \left|H(\Omega)\right|^2 \sigma_U^2. \tag{A.2.2}$$

Finally, exploiting the spectral integration property, which guarantees that the observation of a given spectral measure extended over the entire frequency range leads it back to the corresponding time-domain measure, it is possible to obtain:

$$P_X = \frac{1}{\pi} \int_0^{\pi} P_X(\Omega)\, d\Omega = \frac{\sigma_U^2}{\pi} \int_0^{\pi} \left|H(\Omega)\right|^2 d\Omega \tag{A.2.3}$$

References

[1] L. Jóźwiak, Advanced mobile and wearable systems, Microprocess. Microsyst. 50 (2017) 202–221.
[2] S. Majumder, T. Mondal, M.J. Deen, Wearable sensors for remote health monitoring, Sensors 17 (1) (2017) 130.
[3] G. Volpes, S. Valenti, G. Genova, C. Barà, A. Parisi, L. Faes, A. Busacca, R. Pernice, Wearable ring-shaped biomedical device for physiological monitoring through finger-based acquisition of electrocardiographic, photoplethysmographic, and galvanic skin response signals: Design and preliminary measurements, Biosensors 14 (4) (2024) 205.
[4] M.J. Deen, Information and communications technologies for elderly ubiquitous healthcare in a smart home, Pers. Ubiquitous Comput. 19 (2015) 573–599.
[5] B. Raghavendra, D. Bera, A.S. Bopardikar, R. Narayanan, Cardiac arrhythmia detection using dynamic time warping of ECG beats in e-healthcare systems, in: 2011 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks, IEEE, 2011, pp. 1–6.
[6] N. Agoulmine, M.J. Deen, J.-S. Lee, M. Meyyappan, U-health smart home, IEEE Nanotechnol. Mag. 5 (3) (2011) 6–11.
[7] K.T. Sweeney, T.E. Ward, S.F. McLoone, Artifact removal in physiological signals—Practices and possibilities, IEEE Trans. Inf. Technol. Biomed. 16 (3) (2012) 488–500.
[8] C. Zou, Y. Qin, C. Sun, W. Li, W. Chen, Motion artifact removal based on periodical property for ECG monitoring with wearable systems, Pervasive Mob. Comput. 40 (2017) 267–278.
[9] G.M. Friesen, T.C. Jannett, M.A. Jadallah, S.L. Yates, S.R. Quint, H.T. Nagle, A comparison of the noise sensitivity of nine QRS detection algorithms, IEEE Trans. Biomed. Eng. 37 (1) (1990) 85–98.
[10] F. Shaffer, Z.M. Meehan, C.L. Zerr, A critical review of ultra-short-term heart rate variability norms research, Front. Neurosci. 14 (2020) 594880.
[11] R. Castaldo, L. Montesinos, P. Melillo, C. James, L. Pecchia, Ultra-short term HRV features as surrogates of short term HRV: A case study on mental stress detection in real life, BMC Med. Inform. Decis. Mak. 19 (2019) 1–13.

[12] H.J. Baek, C.-H. Cho, J. Cho, J.-M. Woo, Reliability of ultra-short-term analysis as a surrogate of standard 5-min analysis of heart rate variability, Telemed. E-Heal. 21 (5) (2015) 404–414.

[13] G. Volpes, C. Barà, A. Busacca, S. Stivala, M. Javorka, L. Faes, R. Pernice, Feasibility of ultra-short-term analysis of heart rate and systolic arterial pressure variability at rest and during stress via time-domain and entropy-based measures, Sensors 22 (23) (2022) 9149.

[14] N. Alugubelli, H. Abuissa, A. Roka, Wearable devices for remote monitoring of heart rate and heart rate variability—what we know and what is coming, Sensors 22 (22) (2022) 8903.

[15] F. Scardulla, G. Cosoli, S. Spinsante, A. Poli, G. Iadarola, R. Pernice, A. Busacca, S. Pasta, L. Scalise, L. D'Acquisto, Photoplethysmograhic sensors, potential and limitations: Is it time for regulation? A comprehensive review, Measurement 218 (2023) 113150.

[16] M. Malik, J.T. Bigger, A.J. Camm, R.E. Kleiger, A. Malliani, A.J. Moss, P.J. Schwartz, Heart rate variability: Standards of measurement, physiological interpretation, and clinical use, Eur. Heart J. 17 (3) (1996) 354–381.

[17] F. Shaffer, J.P. Ginsberg, An overview of heart rate variability metrics and norms, Front. Public Heal. 5 (2017) 258.

[18] H.F. Jelinek, D.J. Cornforth, A.H. Khandoker, ECG Time Series Variability Analysis: Engineering and Medicine, CRC Press, 2017.

[19] R. Romero-Ortuño, N. Martínez-Velilla, R. Sutton, A. Ungar, A. Fedorowski, R. Galvin, O. Theou, A. Davies, R.B. Reilly, J. Claassen, et al., Network physiology in aging and frailty: The grand challenge of physiological reserve in older adults, 2021.

[20] B. Xhyheri, O. Manfrini, M. Mazzolini, C. Pizzi, R. Bugiardini, Heart rate variability today, Prog. Cardiovasc. Dis. 55 (3) (2012) 321–331.

[21] H.V. Huikuri, T. Mäkikallio, K.J. Airaksinen, R. Mitrani, A. Castellanos, R.J. Myerburg, Measurement of heart rate variability: A clinical tool or a research toy? J. Am. Coll. Cardiol. 34 (7) (1999) 1878–1883.

[22] E. Mejía-Mejía, K. Budidha, T.Y. Abay, J.M. May, P.A. Kyriacou, Heart rate variability (HRV) and pulse rate variability (PRV) for the assessment of autonomic responses, Front. Physiol. 11 (2020) 779.

[23] R.R. Dos Santos, T.M. da Silva, L.E.V. Silva, A.L. Eckeli, H.C. Salgado, R. Fazan Jr., Correlation between heart rate variability and polysomnography-derived scores of obstructive sleep apnea, Front. Netw. Physiol. 2 (2022) 958550.

[24] P.K. Stein, P.P. Domitrovich, N. Hui, P. Rautaharju, J. Gottdiener, Sometimes higher heart rate variability is not better heart rate variability: Results of graphical and nonlinear analyses, J. Cardiovasc. Electrophysiol. 16 (9) (2005) 954–959.

[25] M.P. van den Berg, J. Haaksma, J. Brouwer, R.G. Tieleman, G. Mulder, H.J. Crijns, Heart rate variability in patients with atrial fibrillation is related to vagal tone, Circulation 96 (4) (1997) 1209–1216.

[26] S.L. Marple Jr., Digital Spectral Analysis, Courier Dover Publications, 2019.

[27] A. Porta, S. Guzzetti, N. Montano, M. Pagani, V. Somers, A. Malliani, G. Baselli, S. Cerutti, Information domain analysis of cardiovascular variability signals: Evaluation of regularity, synchronisation and co-ordination, Med. Biol. Eng. Comput. 38 (2000) 180–188.

[28] H. Azami, L. Faes, J. Escudero, A. Humeau-Heurtier, L.E. Silva, Entropy analysis of univariate biomedical signals: Review and comparison of methods, Front. Entropy Across Discipl.: Panor. Entropy: Theory, Comput. Appl. (2023) 233–286.

[29] R. Pernice, M. Javorka, J. Krohova, B. Czippelova, Z. Turianikova, A. Busacca, L. Faes, I. Member, Comparison of short-term heart rate variability indexes evaluated through electrocardiographic and continuous blood pressure monitoring, Med. Biol. Eng. Comput. 57 (2019) 1247–1263.

[30] B. Le Roy, C. Aufauvre-Poupon, A. Ferragu, A. Vannier, C. Martin-Krumm, M. Trousselard, Cardiac biosignal in confined nuclear submarine patrol: Heart rate variability a marker of adaptation, Acta Astronaut. 203 (2023) 469–482.

[31] M.M. Moraes, T.T. Mendes, R.M. Arantes, Smart wearables for cardiac autonomic monitoring in isolated, confined and extreme environments: A perspective from field research in antarctica, Sensors 21 (4) (2021) 1303.

[32] G. Manis, A. Alexandridi, S. Nikolopoulos, K. Davos, The effect of white noise and false peak detection on HRV analysis, in: The First International Workshop on Biosignal Processing and Classification, vol. 2, SciTePress, 2005, pp. 161–166.

[33] N.J. Napoli, M.W. Demas, S. Mendu, C.L. Stephens, K.D. Kennedy, A.R. Harrivel, R.E. Bailey, L.E. Barnes, Uncertainty in heart rate complexity metrics caused by R-peak perturbations, Comput. Biol. Med. 103 (2018) 198–207.

[34] A. Scarciglia, V. Catrambone, C. Bonanno, G. Valenza, Physiological noise: Definition, estimation, and characterization in complex biomedical signals, IEEE Trans. Biomed. Eng. 71 (1) (2023) 45–55.

[35] G. Adamo, D. Agrò, S. Stivala, A. Parisi, A. Tomasino, L. Curcio, R. Pernice, C. Giaconia, A. Busacca, G. Fallica, Signal to noise ratio of silicon photomultipliers measured in the continuous wave regime, in: 2014 Third Mediterranean Photonics Conference, IEEE, 2014, pp. 1–3.

[36] G. Volpes, L. Sparacino, S. Valenti, A. Parisi, A. Busacca, L. Faes, R. Pernice, A portable multisensor system to assess cardiorespiratory interactions through photoplethysmography, in: 2022 IEEE International Symposium on Medical Measurements and Applications (MeMeA), IEEE, 2022, pp. 1–6.

[37] W.M. Laghari, M.U. Baloch, M.A. Mengal, S.J. Shah, Performance analysis of analog butterworth low pass filter as compared to Chebyshev type-I filter, Chebyshev type-II filter and elliptical filter, Circuits Syst. 5 (9) (2014) 209–216.

[38] S.V. Vaseghi, Advanced Digital Signal Processing and Noise Reduction, John Wiley & Sons, 2008.

[39] F. Arteaga, A. Ferrer, How to simulate normal data sets with the desired correlation structure, Chemometr. Intell. Lab. Syst. 101 (1) (2010) 38–42.

[40] W. Xiong, L. Faes, P.C. Ivanov, Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: Effects of artifacts, nonstationarity, and long-range correlations, Phys. Rev. E 95 (6) (2017) 062114.

[41] M. Priestley, Spectral analysis and time series, 1981.

[42] H.K. Jayant, K. Rana, V. Kumar, S.S. Nair, P. Mishra, Efficient IIR notch filter design using minimax optimization for 50Hz noise suppression in ECG, in: 2015 International Conference on Signal Processing, Computing and Control, ISPCC, IEEE, 2015, pp. 290–295.

[43] H. Limaye, V. Deshmukh, ECG noise sources and various noise removal techniques: A survey, Int. J. Appl. Or Innov. Eng. Manag. 5 (2) (2016) 86–92.

[44] M. Tourni, S.J. Han, R. Weber, M. Kucinski, E.Y. Wan, A.B. Biviano, E.E. Konofagou, Electromechanical cycle length mapping for atrial arrhythmia detection and cardioversion success assessment, Comput. Biol. Med. 163 (2023) 107084.

[45] L. Uldry, J. Van Zaen, Y. Prudat, L. Kappenberger, J.-M. Vesin, Measures of spatiotemporal organization differentiate persistent from long-standing atrial fibrillation, Europace 14 (8) (2012) 1125–1131.

[46] L. Salahuddin, J. Cho, M.G. Jeong, D. Kim, Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings, in: 2007 29th Annual International Conference of the Ieee Engineering in Medicine and Biology Society, IEEE, 2007, pp. 4656–4659.

[47] J. Pan, W.J. Tompkins, A real-time QRS detection algorithm, IEEE Trans. Biomed. Eng. (3) (1985) 230–236.

[48] M. Vollmer, A robust, simple and reliable measure of heart rate variability using relative RR intervals, in: 2015 Computing in Cardiology Conference (CinC), IEEE, 2015, pp. 609–612.

[49] F. Shaffer, R. McCraty, C.L. Zerr, A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability, Front. Psychol. 5 (2014) 1040.

[50] R.B. Blackman, J.W. Tukey, The measurement of power spectra from the point of view of communications engineering—Part I, Bell Syst. Tech. J. 37 (1) (1958) 185–282.

[51] G. Pinna, R. Maestri, A. Di Cesare, Application of time series spectral analysis theory: Analysis of cardiovascular variability signals, Med. Biol. Eng. Comput. 34 (1996) 142–148.

[52] G. Nollo, L. Faes, B. Pellegrini, A. Porta, R. Antolini, Synchronization index for quantifying nonlinear causal coupling between RR interval and systolic arterial pressure after myocardial infarction, in: Computers in Cardiology 2000. Vol. 27 (Cat. 00CH37163), IEEE, 2000, pp. 143–146.

[53] H. Pinto, I. Lazic, Y. Antonacci, R. Pernice, D. Gu, C. Barà, L. Faes, A.P. Rocha, Testing dynamic correlations and nonlinearity in bivariate time series through information measures and surrogate data analysis, Front. Netw. Physiol. 4 (2024) 1385421.

[54] C. Barà, R. Pernice, C.A. Catania, M. Hilal, A. Porta, A. Humeau-Heurtier, L. Faes, Comparison of entropy rate measures for the evaluation of time series complexity: Simulations and application to heart rate and respiratory variability, Biocybern. Biomed. Eng. 44 (2) (2024) 380–392.

[55] G. Schwarz, Estimating the dimension of a model, Ann. Statist. (1978) 461–464.

[56] F. Wilcoxon, Individual comparisons by ranking methods, in: Breakthroughs in Statistics: Methodology and Distribution, Springer, 1992, pp. 196–202.

[57] A. Bánhalmi, J. Borbás, M. Fidrich, V. Bilicki, Z. Gingl, L. Rudas, Analysis of a pulse rate variability measurement using a smartphone camera, J. Heal. Eng. 2018 (1) (2018) 4038034.

[58] J. Benesty, J. Chen, Y. Huang, I. Cohen, Noise reduction in speech processing, vol. 2, Springer Science & Business Media, 2009.

[59] E. van den Heuvel, Z. Zhan, Myths about linear and monotonic associations: Pearson'sr, Spearman's ρ, and Kendall's τ, Amer. Statist. 76 (1) (2022) 44–52.

[60] C.O. Fritz, P.E. Morris, J.J. Richler, Effect size estimates: Current use, calculations, and interpretation, J. Exp. Psychol. [Gen.] 141 (1) (2012) 2.

[61] J. Cohen, Statistical Power Analysis for the Behavioral Sciences, routledge, 2013.

[62] J.M. Bland, D. Altman, Statistical methods for assessing agreement between two methods of clinical measurement, Lancet 327 (8476) (1986) 307–310.

[63] D. Giavarina, Understanding bland altman analysis, Biochem. Medica 25 (2) (2015) 141–151.

[64] J. Krohova, L. Faes, B. Czippelova, R. Pernice, Z. Turianikova, R. Wiszt, N. Mazgutova, A. Busacca, M. Javorka, Vascular resistance arm of the baroreflex: methodology and comparison with the cardiac chronotropic arm, J. Appl. Physiol. 128 (5) (2020) 1310–1320.

[65] A. Cevese, G. Gulli, E. Polati, L. Gottin, R. Grasso, Baroreflex and oscillation of heart period at 0.1 Hz studied by α-blockade and cross-spectral analysis in healthy humans, J. Physiol. 531 (1) (2001) 235–244.

[66] C. Barà, A. Zaccaro, Y. Antonacci, M. Dalla Riva, A. Busacca, F. Ferri, L. Faes, R. Pernice, Local and global measures of information storage for the assessment of heartbeat-evoked cortical responses, Biomed. Signal Process. Control. 86 (2023) 105315.

[67] S. Romagnoli, I. Marcantoni, K. Campanella, A. Sbrollini, M. Morettini, L. Burattini, Ensemble empirical mode decomposition for efficient R-peak detection in electrocardiograms acquired by portable sensors during sport activity, in: 2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA), IEEE, 2021, pp. 1–6.

[68] H. Lütkepohl, New Introduction to Multiple Time Series Analysis, Springer Science & Business Media, 2005.