Journal of Information Hiding and Multimedia Signal Processing ©2025 ISSN 2073-4212
Ubiquitous International Volume **16**, Number **2**, June **2025**

Comparative study on PCG envelope Extraction

Sidi Ahmed Taouli[1]

[1]Genie-Biomedical Laboratory (GBM), Department of Biomedical Engineering,
Faculty of Technology, University of Abou Bekr Belkaid,
Tlemcen BP 119, Tlemcen 13000, Algeria
sidiahmed.taouli@univ-tlemcen.dz

Received February 15, 2025, revised April 19, 2025, accepted April 20, 2025.

Abstract. This work presents a comparative evaluation of five envelope extraction methods for the PCG signal under added white noise conditions, namely: Shannon envelope, Hilbert envelope, Shannon entropy, the cardiac sound characteristic waveform, and Multiscale product method. The performance was assessed using the false positive rate (FPR) and the area under the ROC curve (AUC) at three SNR levels (15 dB, 5 dB, and 0 dB). At 15 dB, the methods yielded promising results, with FPRs ranging from 2% to 10% and AUCs between 89% and 97%. As the noise increased (at 5 dB and 0 dB), a general degradation in performance was observed—particularly for the Shannon, Entropy, and Hilbert methods, which exhibited significant increases in FPR and decreases in AUC (for instance, Shannon dropped from an AUC of 89% to 76%). In contrast, the cardiac sound characteristic waveform and Multi-scale methods maintained relatively robust performance, with the Multi-scale approach performing notably well, achieving AUCs of 97%, 94%, and 88.5% with corresponding FPRs of 2%, 4%, and 12% at 15 dB, 5 dB, and 0 dB, respectively. These results underscore the importance of adapting the extraction method to noise conditions and suggest that the Multi-scale approach is the most reliable for precise phonocardiogram signal extraction in noisy clinical environments.
Keywords: PCG signal; envelope extraction; Shannon; Shannon entropy; Hilbert transform; cardiac sound characteristic waveform; multi-scale product

1. **Introduction.** Cardiac auscultation, performed using a stethoscope, is a fundamental clinical examination technique. This instrument, consisting of a diaphragm or bell connected to earphones, amplifies the sounds produced by the heart. These sounds, known as heart sounds, provide valuable information about cardiac function. Normal heart sounds, S1 and S2, correspond to the closure of the heart valves. Abnormal sounds, known as murmurs, may indicate valvular stenosis (narrowing) or valvular insufficiency (regurgitation). With the advent of electronic stethoscopes, it is now possible to record and analyze these sounds in greater detail. This recording, called a phonocardiogram (PCG), helps detect subtle abnormalities that might be missed during traditional auscultation. The PCG signal plays a crucial complementary role in cardiac auscultation [1]. It not only confirms the findings obtained from a stethoscopic examination but also refines them by providing additional insights into the acoustic activity of the heart. Through PCG analysis, the precise timing of pathological signs within the cardiac cycle can be determined, facilitating the diagnosis and understanding of cardiac disorders.

A phonocardiogram (PCG) is a graphical representation of heart sounds, comprising multiple components that provide essential information about cardiac function. The primary features of the PCG include the normal S1 and S2 sounds, corresponding to the

closure of the atrioventricular and semilunar valves, respectively [1]. In addition to these fundamental sounds, the PCG may also contain additional acoustic events, such as heart murmurs, which result from turbulent blood flow caused by stenoses or valvular regurgitations. These murmurs can be classified as systolic or diastolic, depending on their occurrence within the cardiac cycle.

Furthermore, S3 sounds are often associated with ventricular dilation, whereas S4 sounds indicate ventricular stiffness. Cardiac sounds are inherently non-stationary, meaning their characteristics vary over time. They are predominantly low-frequency signals, typically ranging between 40 and 600 Hz [2, 3], mainly resulting from valvular vibrations, blood flow dynamics, and myocardial wall motion during different phases of the cardiac cycle.

The automatic analysis of cardiac signals, such as phonocardiograms, presents significant challenges due to their highly non-stationary nature. To facilitate this analysis, it is essential to extract relevant signal features. While frequency based analysis has been extensively explored, time domain techniques, particularly envelope extraction, offer promising new perspectives.

The envelope of a cardiac signal provides valuable information about its intrinsic characteristics, such as the amplitude and duration of different cardiac events. This approach enhances our understanding of the underlying physiological mechanisms involved in heart sound production and enables the correlation of these features with various cardiac pathologies. Envelope extraction is widely utilised to simplify the analysis of biomedical signals. It improves the detection of S1 and S2 heart sounds by highlighting regions of interest, thereby facilitating signal segmentation and reducing noise impact. Using time domain, frequency domain, and time frequency methods, envelope extraction enhances amplitude variation analysis, ultimately improving the robustness of heart sound classification algorithms. This technique is particularly valuable for identifying abnormalities such as heart murmurs and optimising the performance of computer aided diagnostic systems.

Recent advancements in digital signal processing have paved the way for novel clinical applications. Automated cardiac signal analysis has the potential to enable earlier detection of cardiovascular diseases and provide more personalized patient monitoring. Additionally, the development of user-friendly interfaces could make this technology accessible to a broader audience, including non-specialists.

Research on heart sound envelope extraction has explored various methods. A commonly employed approach involves segmenting the signal based on its instantaneous energy, with Shannon energy frequently used as a key measure due to its ability to capture amplitude variations in the frequency domain [4]. By normalizing this energy, different signal regions can be compared, aiding in the identification of cardiac events.

However, Shannon energy-based segmentation is not always optimal, particularly in noisy environments or when dealing with highly disturbed signals. Since most cardiovascular sounds fall below 1 kHz, the primary goal is to classify heart sounds as normal or abnormal using Shannon energy. This energy is computed from pre-processed signals, which are either low pass filtered (frequencies below 1 kHz) or band pass filtered to target specific frequency ranges [5].

Moreover, Hilbert transform based methods leverage two key features: the envelope (E), derived from the real part of an analytic signal, and the instantaneous frequency (IF), computed as the derivative of the imaginary part of the analytic signal [6, 7, 8, 9]. Previous studies have utilized entropy based metrics to characterize the complexity of heart sounds in PCG signals. By quantifying the degree of randomness and unpredictability within the signal, this approach enables precise detection and quantification of cardiac abnormalities,

which are often associated with increased local entropy [10, 11]. This method provides a valuable diagnostic tool, particularly for the detection of heart murmurs and other rhythm abnormalities.

Additionally, algorithms dedicated to extracting heart sound characteristic waveforms of heart sounds are based on one degree of freedom analytical models. These models allow for derivation of characteristic waveforms, while their response curves assist in classifying heart sounds as normal or abnormal [12, 13]. Then, the multi-scale product method, which is based on the idea that the temporal characteristics of heart sounds can be highlighted by combining information from different signal representation scales. Specifically, this method applies a continuous or discrete wavelet transform to the phonocardiogram (PCG) signal to decompose it into multiple frequency scales [14, 15].

However, to ensure effective auscultation of heart sounds, several hardware and contextual factors must be considered [16, 17]. These include stethoscope system performance, recording conditions, and individual variations. Such factors can directly influence signal quality, particularly through variations and interferences caused by ambient noise. Signal envelopes are especially sensitive to noise components, which can significantly degrade heart sound segmentation accuracy. Our study focuses on a comparative analysis of five envelope extraction algorithms for segmenting first (S1) and second (S2) heart sounds. We compare our proposed method against widely used reference techniques, such as normalized Shannon energy, Hilbert transform, Shannon entropy, cardiac sound characteristic waveform analysis, and multi-scale product techniques. These methods have been applied in various biomedical applications and evaluated in multiple studies. However, their comparative performance in noisy conditions remains an open research question, requiring further analysis to determine the most suitable method for real PCG signals. The performance of different envelope extraction techniques will be assessed based on two key criteria:

- Database diversity: The dataset used for this study includes a variety of cardiac signals, encompassing both clinically normal recordings and signals exhibiting different cardiac pathologies. This diversity will enable the evaluation of each method's ability to adapt to a broad range of clinical scenarios.
- Impact of noise: The test signals originate from a clinical environment, where they are naturally contaminated by ambient noise, respiratory sounds, and frictional noises associated with auscultation. To further examine the robustness of each method, we will artificially add three levels of Gaussian white noise to the clinical signals, corresponding to signal-to-noise ratios (SNR) of 15 dB, 5 dB, and 0 dB, respectively.

To assess the impact of noise on cardiac signals, we will compute the signal-to-noise ratio (SNR). This metric will be obtained by comparing the average signal power within a time window containing S1 and S2 sounds to the average background noise power in an equally sized window during diastole (i.e., when no heart sounds are present).

Furthermore, the second evaluation of the five algorithms for S1 and S2 localization will be based on receiver operating characteristic (ROC) curve analysis, focusing on the area under the curve (AUC) and its variation under different levels of Gaussian white noise.

The ROC curve is a widely used method for visualizing and comparing classifier performance, illustrating variations in true positive rate (sensitivity) and false positive rate (1-specificity) as a function of threshold values. This approach provides insights into each method's robustness in precisely localizing S1 and S2 sounds in noisy conditions.

This article is structured as follows: Section 2 focuses on the localization of heart sounds using envelope extraction techniques. Section 3, entitled "Algorithmic implementation of

the proposed methods", describes in detail the implementation of the algorithms developed in this study. Section 4 presents a comprehensive performance analysis, highlighting key parameters relevant to signal processing and classification. The results, discussion, and conclusions are provided in Sections 5 and 6, respectively.

2. **Envelope Extraction.** The localization of heart sound signals is a crucial step in the analysis of phonocardiographic (PCG) signals. It aims to precisely identify the occurrence times of the first (S1) and second (S2) heart sounds. This temporal localization precedes signal segmentation, which involves dividing the cardiac cycle into four distinct phases: S1, systole, S2, and diastole. Such detailed segmentation significantly facilitates automatic analysis and classification tasks, providing clinicians and diagnostic support systems with both quantitative and qualitative information about different components of the heart sound signal.

Direct heart sound localization methods are applied exclusively to the phonocardiogram (PCG) without relying on complementary signals such as the electrocardiogram (ECG). This non-invasive approach presents an attractive alternative to indirect methods, enabling a simpler and less restrictive assessment of cardiac function for the patient. Our study specifically focuses on the evaluation and development of these direct localization methods. Most heart sound localisation algorithms follow a standard process, as illustrated in Figure 1. After acquiring and filtering the PCG signal, a specific transformation is applied to extract relevant signal features.

This transformation, at the core of the localization process, highlights the primary components associated with the S1 and S2 heart sounds. An additional low pass filtering step is often employed to suppress high frequency noise and enhance the readability of the signal envelope. Finally, thresholding and local maxima detection are used to accurately identify the occurrence of the S1 and S2 heart sounds.

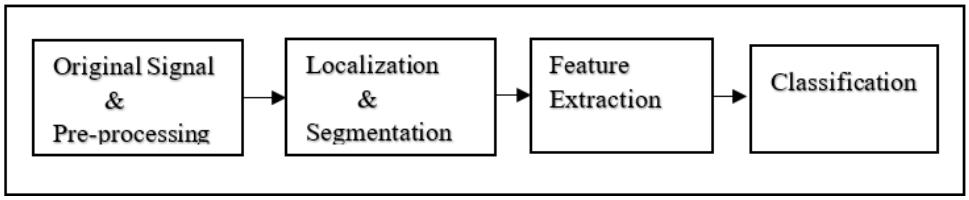

FIGURE 1. Structural diagram of PCG Signal Processing

A comparative study of heart sound localization algorithms was conducted using the PASCAL database [18]. This database includes 176 recordings intended for heart sound segmentation and 656 recordings for heart sound classification. Although the number of recordings is relatively large, their duration is limited, ranging from 1 to 30 seconds.

Furthermore, the frequency range is restricted to values below 195 Hz due to the application of a low pass filter. This filtering removes many components of heart sounds that are essential for accurate clinical diagnosis. Subsequently, normalization is performed by adjusting the signal variance to a value of 1. This step standardizes the signals, facilitating comparisons and subsequent analyses. The resulting signals can be expressed using the following equation [19, 20]:

$$x_{\text{norm}}(t) = \frac{x(t)}{|\max(x(t))|} \tag{1}$$

where $x(t)$ represents the filtered original PCG signal.
This normalization ensures that the signals have a standardized amplitude while preserving their temporal and frequency characteristics. Heart sounds are often embedded in

significant background noise originating from respiration, body movements, or medical devices. These signals must be segmented into different components during the initial stage of automatic analysis and classification. The envelope of heart sounds provides valuable information for studying the intrinsic characteristics of the signal. However, while the envelope offers a simplified representation, it contains less information than the original sound signal. For example, a normal heart sound and its corresponding envelope curve are illustrated in Figure 2.

It is evident that the envelope curve, derived from the first heart sound (S1) and the second heart sound (S2), provides a clearer and more intuitive representation than the raw acoustic signals. In this study, we analyze five of the most representative envelope extraction techniques: the Shannon envelope (based on the normalized Shannon energy), the Hilbert transform envelope, entropy-based methods, characteristic heart sound waveforms, and a nonlinear method.

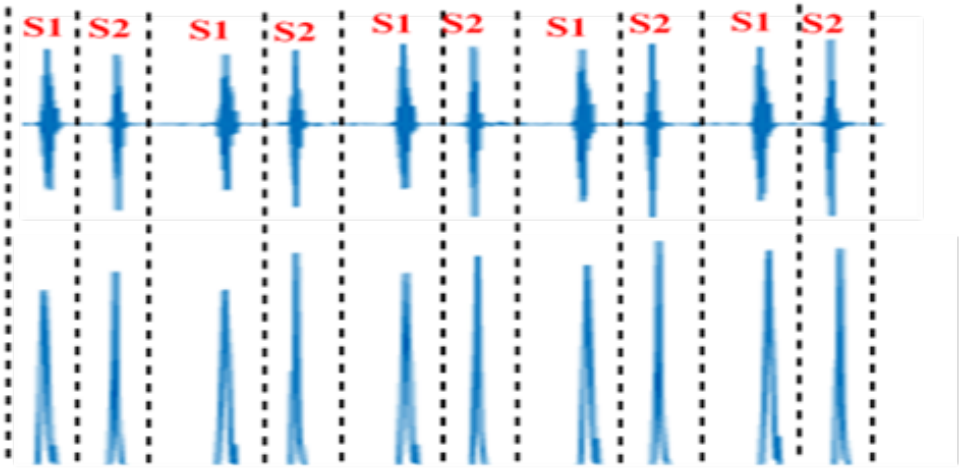

FIGURE 2. PCG signal in top row, envelope extraction in bottom row

2.1. **Shannon Envelope. Shannon Envelope.** The Shannon envelope, based on the normalized Shannon energy, is a widely used method for extracting the envelope of heart sound signals. Further details of this algorithm are explained later in the text.

The normalized Shannon energy, also known as the Shannon envelope, is a commonly used technique for heart sound envelope extraction, particularly in the analysis of phonocardiograms (PCG). It relies on nonlinear signal processing to emphasize energy variations over time, enhancing the detection of cardiac events such as heart sounds.

The main steps for obtaining the Shannon envelope of a heart sound signal are as follows:

a) Compute the Instantaneous Energy $E[n]$ of the Normalized Signal by Squaring Its Amplitude [21]:

$$E[t] = x[t]^2_{\text{norm}} \tag{2}$$

with $x[t]_{\text{norm}}$ being the amplitude of the signal at time t.

b) Transform the instantaneous energy using the Shannon formula to obtain the normalized energy:

$$E_{\text{shannon}} = -E[t] \cdot \log\left(E[t]\right) \tag{3}$$

c) Calculate the Mean Shannon Energy of the digital signal:

$$E_i = -\frac{1}{N} \sum_{i=1}^{N} E[i] \cdot \log\left(E[i]\right) \tag{4}$$

When using a sliding window with a length of 0.02 s, corresponding to $N = 40$ samples (for a sampling frequency of 2 kHz), and a 50% overlap rate, the normalized mean Shannon energy is calculated as:

$$P(t) = \frac{E_i[t] - M(E_i[t])}{S(E_i[t])} \tag{5}$$

In this context, $E_i[t]$ represents the mean Shannon energy within window i, $M(E_i[t])$ is the average value of $E_i[t]$, and $S(E_i[t])$ is its standard deviation.

These values are often used to characterize the Shannon envelope or to detect anomalies in heart sound signals.

2.2. **Hilbert Envelope.** The Hilbert envelope is a mathematical representation that allows us to track the amplitude of a signal over time while ignoring high-frequency oscillations.

The Hilbert envelope of a signal $x(t)$ is obtained from the Hilbert transform $H(x(t))$, which generates a complex analytic signal [22, 23]:

$$S_a(t) = x(t) + jH(x(t)) = A(t)e^{j\varphi(t)} \tag{6}$$

where: $x(t)$ is the normalized PCG signal,

$H(x(t))$ is its Hilbert transform, representing the imaginary part of the analytic signal defined by the following formula:

$$H(x(t)) = \frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{x(\tau)}{t-\tau}d\tau = x(t) * \frac{1}{\pi t} \tag{7}$$

The Hilbert envelope of a real signal $x(t)$, denoted as $A(t)$, is defined as the modulus of the analytic signal $S_a(t)$ associated with $x(t)$:

$$A(t) = |S_a(t)| = \sqrt{x(t)^2 + H(x(t))^2} \tag{8}$$

The main objective of the approach presented in [24] is to eliminate high-frequency vibrations present in the envelope of a signal to highlight lower-frequency trends.

2.3. **Shannon Entropy.** Shannon entropy is a fundamental measure in information theory that quantifies the degree of uncertainty or diversity in the probability distribution of a particular event x, which in this case corresponds to the PCG signal.

The entropy of a signal x is defined as:

$$H(x) = -\sum_{i=1}^{N} P(x_i)\log_2 P(x_i) \tag{9}$$

where:

$P(x)$ represents the probability density of the amplitude x_i occurring in the PCG signal, N is the total number of amplitude levels considered, and the logarithm base determines the unit of entropy, with a $\log_2$ expressing entropy in bits.

In the domain of PCG signals, this formula is used to quantify the uncertainty or variability of signal amplitudes. In other words, it measures the complexity of the PCG signal [25].

Several methods exist for estimating the probability density of a discrete signal, and these approaches are essential for calculating Shannon entropy. In probability density estimation, two major categories of methods can be distinguished:

- *Parametric Methods:* These methods assume that the probability density follows a known distribution (e.g., Gaussian) and estimate its parameters.
- *Non-Parametric Methods:* These methods do not assume a predefined model and construct the probability density directly from the data.

Parametric methods have the disadvantage of being sensitive to violations of the assumed distribution hypothesis. If the signal does not follow the presumed distribution, the estimation may be biased and inaccurate. These methods are less suitable for complex signals where the underlying distribution is unknown or difficult to model. On the other hand, non-parametric methods offer greater flexibility and are better suited for complex signals with unknown or difficult to model distributions. The choice between parametric and non-parametric methods depends on the prior knowledge of the signal's distribution and the amount of available data.

For PCG signals, non-parametric methods are generally more appropriate due to their complexity and the challenges in modeling their distribution. The histogram method is one of the simplest and fastest non-parametric approaches for estimating the probability density of a PCG signal. However, it has a major drawback: it does not provide density smoothing.

For this reason, kernel density estimators are often a better alternative to histograms, as they enable density smoothing, are less sensitive to bin selection, and offer greater accuracy.

Yaddoulahi et al. [26] employed kernel density estimation to model the distribution of heart sounds in PCG signals. They assumed that heart sounds followed a different distribution than lung sounds, using the following kernel estimator:

$$\hat{P}(x) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{h}K\left(\frac{x - X_i}{h}\right) \tag{10}$$

Where:
x is the point at which the probability density is to be estimated, X_i represents the observations (samples) of the PCG signal, N is the total number of observations, $K(x) = (\frac{1}{2\pi})e^{-\frac{x^2}{2}}$ is the kernel function, and $h = 1.06\hat{\sigma}(x)N^{-0.2}$ is the scaling factor, where $\hat{\sigma}$ is the standard deviation of the observations.

The steps for calculating Shannon entropy on a PCG signal are: The steps for calculating Shannon entropy on a PCG signal are as follows:
(i) Divide the signal into segments of 50 samples with a 50% overlap.
(ii) For each segment, compute Shannon entropy.
(iii) Determine the adaptive threshold: threshold $= (\sigma + \mu)$.
(iv) Compare the entropy of each segment to the threshold to locate heart sounds.

2.4. **Cadiac sound characteristic Waveform.** The Single Degree of Freedom (SDOF) model is a simple yet powerful tool for analyzing dynamic systems. In this context, it is used to model the cardiovascular system's response to heart sounds. The model is based on an electrical analogy, where mechanical elements (mass, spring, and damper) are represented by their equivalent electrical components (inductance, capacitance, and resistance).

Although the cardiovascular system is inherently complex, an SDOF model can provide a reasonable approximation of its response to heart sounds. It considers the system as a damped oscillator, which allows capturing the key characteristics of heart sounds. This model is defined by M as the mass, K as the spring stiffness coefficient, and C as the damping coefficient of the damper. By designating the heart sound recorded by the stethoscope as the input signal $X(t)$, the output response $Y(t)$ of the SDOF model is given by the following differential equation [27]:

$$M\ddot{Y}(t) + C\dot{Y}(t) + KY(t) = X(t) \tag{11}$$

This equation can be written as:

$$\ddot{Y}(t) + 2\omega\zeta\dot{Y}(t) + \omega^2\zeta^2 Y(t) = \overline{X}(t) \tag{12}$$

With the signal $\overline{X}(t) = \pm|X(t)/M|$, the angular resonance frequency $\omega = \sqrt{K/M}$ (rad/s), and the damping ratio parameter $\zeta = C/2\sqrt{MK} \times 100$ (%), the appropriate selection of ω and ζ is crucial for extracting heart sound features using an SDOF model. In the study [27], the authors fixed ω a value of 62.832 rad/s (10 Hz) and ζ a value of 70.7% to analyze heart sounds within a specific frequency range and with a high damping level.

2.5. **Multi-scale product envelope.** The ability of wavelets to analyze the local regularity of a function is a fundamental property that makes them a powerful and versatile tool in many scientific and engineering fields. The relationship established by Meyer et al. between the decay of wavelet coefficients and Lipschitz regularity provides a strong theoretical foundation for numerous practical applications [28].

A function f is said to be uniformly Lipschitz of order $\alpha \leq r$ over an interval $[a, b]$, and is defined as:

$$|W_{f,j}(x)| = K2^{j\alpha} \tag{13}$$

Where $W_{f,j}$ represents the wavelet coefficient of the function $f(x)$ at scale j.

Multiresolution analysis provides a powerful tool for separating the regular components of a signal from background noise by leveraging the different behaviors of these two types of signals across various scales. The increasing distance between the signal and noise with scale is a key characteristic that facilitates the detection and characterization of cardiac signals, particularly the Phonocardiogram (PCG), and aims to highlight the main components of the PCG signal (S1 and S2). The discrete wavelet transform of the pulmonary signal is computed using the Symlet5 mother wavelet [29]. This process involves decomposing the signal into different scales (or detail levels) and computing the corresponding wavelet coefficients. These coefficients represent the signal's energy at each scale and position.

The wavelet coefficient product is a powerful technique that exploits redundancy across scales to enhance certain signal properties. It is given by:

$$P_j(n) = \prod_{i=1}^{j} W_i\left(f(n)\right) \tag{14}$$

With W_i representing the wavelet coefficient at scale i, and the decomposition level j is fixed at 3 to extract the key features of the PCG signal $f(n)$. The signal obtained after the multiscale product of the wavelet coefficients is then filtered using a 5th-order Butterworth low-pass filter with a cutoff frequency of 20 Hz.

3. **Implementation of the Proposed Methods.** The localization of the S1 and S2 heart sounds in PCG signals is a fundamental step in automated cardiac signal analysis. This study investigates five envelope extraction methods applied to signals from the PASCAL database, with a focus on evaluating their robustness under noisy conditions. All signal processing was performed using MATLAB R2022a on a system equipped with an Intel Core i7 processor (8 cores), 32 GB of RAM, and running Windows 10 (64-bit). The PCG signals were sampled at 2 kHz.

All PCG signals were filtered using a 4th-order Butterworth low-pass filter with a cutoff frequency of 195 Hz to eliminate high-frequency components. Subsequently, the signal

was normalized to unit variance, as described by Equation (1), ensuring standardized amplitude while preserving temporal and spectral characteristics.

The evaluation was carried out using the PASCAL Heart Sound Database, which contains 176 recordings for segmentation and 656 recordings for classification. Signal durations range from 1 to 30 seconds and encompass normal, pathological, and noisy cases. To ensure the reproducibility of the study, the main implementation steps for each envelope extraction method are described in detail below.

a) Shannon Energy envelope algorithm

The Shannon envelope method is based on an estimation of the energy of the normalized PCG signal, which is then transformed to emphasize its most prominent components. The process begins with the computation of the instantaneous energy by squaring each sample of the signal (Equation 1). This energy is subsequently transformed using the Shannon energy formulation, where each energy value is multiplied by the negative logarithm of itself (Equation 2). This transformation enhances high-energy peaks while suppressing low-intensity regions. Finally, a moving average is applied using a 0.02-second rectangular window (corresponding to 40 samples at a 2 kHz sampling rate), with 50% overlap between frames, to smooth the resulting signal and generate a clear and informative temporal envelope (Equations 3 and 4).

Figure 3 shows a typical result obtained using Shannon envelope extraction applied to a PCG signal. The figure highlights the envelope of the normalized signal, with distinct peaks corresponding to the S1 and S2 heart sounds.

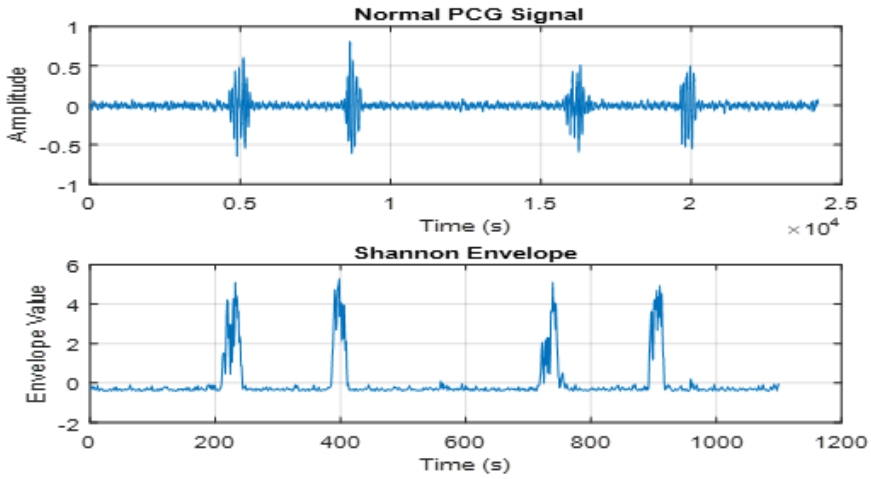

FIGURE 3. Shannon envelope extraction

b) Hilbert envelope algorithm

The Hilbert envelope method is based on the use of the analytic signal $S_a(t)$ obtained by applying the Hilbert transform (Equation 6). This operation constructs a complex signal whose real part corresponds to the normalised PCG signal and whose imaginary part is its Hilbert transform (Equation 7). The analytic signal effectively captures the amplitude variations of the cardiac signal. The envelope is then derived by calculating the modulus of this complex signal, which corresponds to the instantaneous amplitude of the original signal (Equation 8). To improve the clarity of the envelope and suppress high frequency noise, a low pass filter is applied using a 5th order Butterworth filter with a cut-off frequency set at 20 Hz. This approach results in a smooth and well localised envelope that accurately highlights cardiac acoustic events (see figure 4).

c) Shannon Entropy envelope algorithm

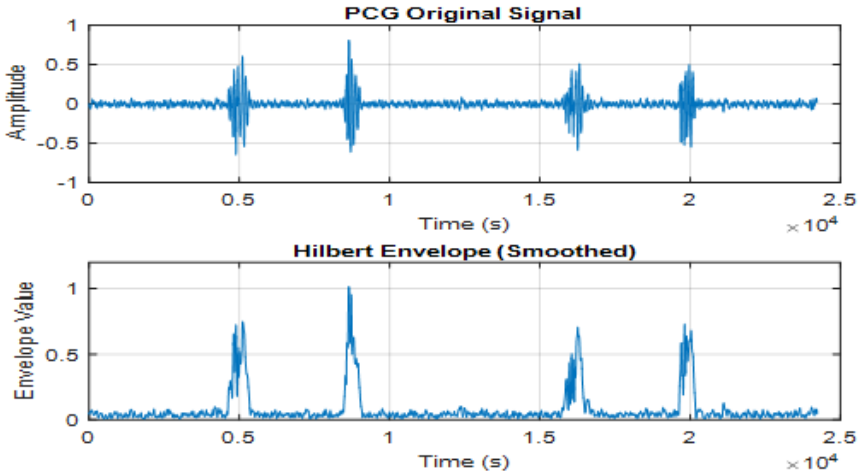

FIGURE 4. Hilbert envelope extraction

The Shannon entropy method uses entropy as a measure of local complexity within the normalised PCG signal (Equation 9) to detect cardiac events. The signal is first divided into segments of 50 samples, with 50% overlap between adjacent windows. For each segment, the probability density function of the signal amplitudes is estimated using a Gaussian kernel with bandwidth set to the local standard deviation of the PCG signal (Equation 10). Based on this estimate, the Shannon entropy is calculated to quantify the degree of disorder within each segment. An adaptive threshold, defined as the mean entropy calculated over the entire signal, is then used to identify segments containing relevant information. This approach is particularly effective in emphasizing key heart sound components. Figure 5 shows the input PCG signal along with its corresponding envelope extracted using the Shannon entropy method. This approach effectively highlights the S1 and S2 heart sounds by enhancing their temporal contrast and attenuating background noise.

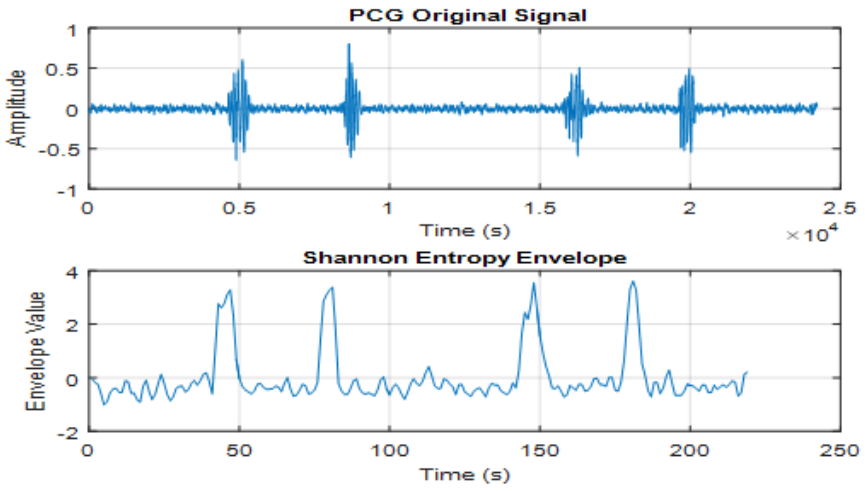

FIGURE 5. Entropy envelope extraction

d) Cadiac sound characteristic waveform envelope algorithm

The Cardiac Sound Characteristic Waveform (CSCW) method is based on a dynamic modeling approach using a Single-Degree-of-Freedom (SDOF) system applied to a pre-processed cardiac signal. Initially, the signal is normalized to its maximum absolute amplitude. This preprocessed signal, denoted as $X(t)$ is subsequently used as the input to an SDOF system modeled by a classical differential equation, where the output $Y(t)$ represents the heart's response to the excitation. This response is then corrected for any temporal delay by means of cross-correlation, allowing accurate alignment of the cardiac events in time. The corrected output signal, referred to as CSCW, forms a smoothed and noise robust envelope, clearly highlighting the first (S1) and second (S2) heart sounds. The corresponding figure 6 illustrates the entire process, showing sequentially: (1) the normalized signal, (2) the uncorrected SDOF system response, and (3) the final CSCW envelope, effectively demonstrating the method's capability to extract relevant cardiac features.

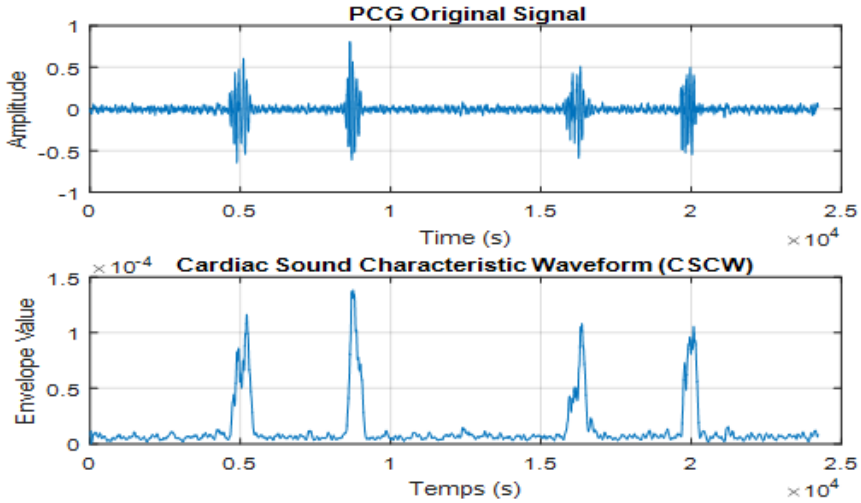

FIGURE 6. CSCW envelope extraction

e) Multi-Scale Product envelope algorithm

The multi-scale product envelope method is a wavelet based technique designed to highlight the salient components of a phonocardiogram (PCG) signal, particularly the first (S1) and second (S2) heart sounds. This method leverages the redundancy and correlation of wavelet coefficients across multiple decomposition levels to amplify relevant features and suppress background noise.

The signal is first decomposed using a Discrete Wavelet Transform (DWT) commonly with the *Symlet5* wavelet up to a predefined level (e.g., level 3). At each level j, the detail coefficients $d_j(n)$ are extracted. The multi-scale product is then computed as the pointwise multiplication of detail coefficients across adjacent levels (see Equation 14).

This operation emphasizes structures (such as S1 and S2) that persist across scales while attenuating uncorrelated noise. To further enhance the signal, the result is first smoothed using a Kalman filter [30], which provides adaptive noise reduction by optimally estimating the signal based on prior observations. This is followed by a low-pass Butterworth filter (commonly 5th-order with a cutoff frequency of 20 Hz) to ensure a clean and continuous envelope of the PCG signal. This combined approach has proven effective for denoising and enhancing cardiac events, particularly in noisy conditions, due to its multiresolution and adaptive nature.

Figure 7 illustrates the input PCG signal and its corresponding Multi-Scale Product envelope. Subfigure (a) shows the input PCG signal, while subfigure (b) shows the envelope extracted using the multi-scale wavelet product technique. This method effectively highlights the S1 and S2 heart sounds by enhancing their temporal contrast and suppressing background noise components.

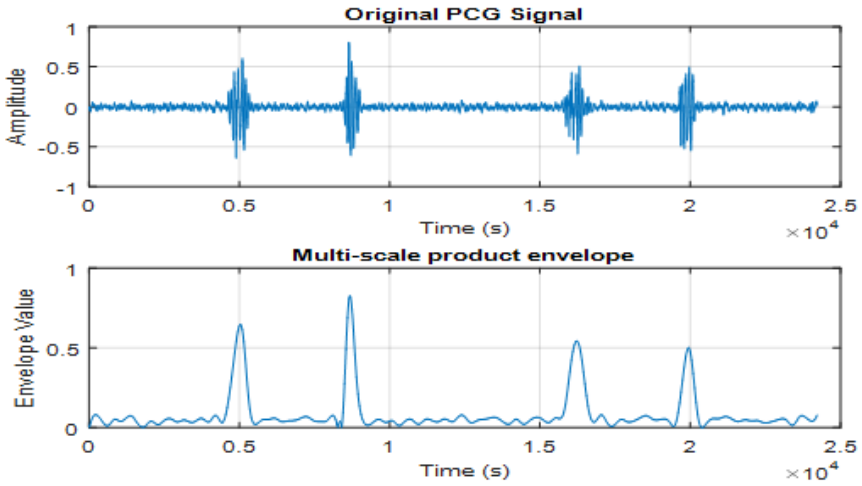

FIGURE 7. Multi-scale product envelope extraction

4. **Evaluation criteria.** The performance of PCG envelope extraction methods for locating S1 and S2 heart sounds is generally measured using the following parameters:

- Sensitivity Indicates the ability of the method to correctly detect S1 and S2 heart sounds. It is calculated as:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{15}$$

Where TP is the number of true positives (sounds detected correctly) and FN is the number of false negatives (sounds not detected).

- Positive predictive value (PPV or accuracy): Measures the proportion of correct detections among those identified as heart sounds. It is defined by:

$$PPV = \frac{TP}{TP + FP} \tag{16}$$

With FP: Number of false detections (sounds wrongly detected).

- Positive Predictive Value (PPV), used here to deduce the false positive rate (FPR) via the relation:

$$FPR = 1 - \frac{PPV}{100} \tag{17}$$

- Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC), which plots sensitivity (Se) as a function of the false positive rate (FPR). It is calculated using the following integral:

$$AUC = \int_0^1 S_e(FPR)d(FPR) \tag{18}$$

Where:
- $S_e(FPR)$ is the sensitivity function as a function of the false positive rate.
- The ROC curve plots sensitivity (S_e) against the false positive rate (FPR) for different decision thresholds.

The performance of the five envelope extraction methods is evaluated on a real dataset containing both normal and abnormal signals, with average signal-to-noise ratio (SNR) of approximately 15 dB. To simulate noisy acquisition conditions, Gaussian noise is added to achieve SNR levels of 5 dB and 0 dB. Each method is then tested at these different noise levels, and the results are compared to analyze their robustness and effectiveness in the presence of perturbations.

5. **Results and discussions.** Our comparative study is based on the Pascal database, which includes a total of 140 audio recordings. Among them, 25 correspond to normal heart sounds, 40 to normal asfounds affected by noise, and 75 to recordings of cardiac pathologies exhibiting systolic murmurs and other cardiac abnormalities. This distribution allows us to accurately assess the performance of the analysis methods applied to both normal and pathological cardiac signals.

Figures 8 and 9 show a comparative analysis of normal and pathological (late systolic murmur) heart sounds, respectively, along with their corresponding envelope representations obtained using different extraction methods. In both figures, sub-figure (b) represents the envelope extracted using Shannon energy. For the normal heart sound (Figure 8.b), the S1 and S2 components are clearly distinguishable, demonstrating the effectiveness of the method in delineating primary heart sound events, despite the relatively broad envelope. In contrast, the pathological signal (Figure 9.b) features a late systolic murmur located after S2 and before the subsequent S1, which is effectively detected by the Shannon-based envelope. Although the murmur has a more irregular and noisy structure compared to normal heart sounds, the Shannon energy method is able to highlight these abnormal components, demonstrating its potential for characterizing pathological patterns.

The Shannon entropy, shown in Figures 8.c and 9.c for the normal and pathological cases respectively, provides improved localisation of the S1 and S2 heart sounds in the normal signal compared to the standard Shannon energy method. In the pathological case (Figure 9.c), which includes a late systolic murmur, the entropy based envelope shows improved sensitivity in detecting both the primary heart sounds and the murmur components. This suggests that the Shannon entropy approach provides a more refined extraction of temporal features than the envelope obtained using Shannon energy.

The envelopes derived using the Hilbert transform are shown in Figures 8.d and 9.d for the normal and pathological cases respectively. In the normal case (Figure 8.d), the envelope shows relatively thick curves with clear and well defined peaks corresponding to the S1 and S2 heart sounds. In contrast, the pathological case with a late systolic murmur (Figure 9.d) presents a noisier envelope, especially in the segment after S2 and before the following S1, where murmur components are present. Compared to the envelopes obtained using Shannon entropy, those derived from the Hilbert transform appear slightly thicker and more affected by noise, especially in the pathological case, indicating a reduced ability to isolate finer acoustic details.

The envelope representations obtained using the Cardiac Sound Characteristic Waveform (CSCW) method are presented in Figures 8.e and 9.e for the normal and pathological cases, respectively. In the normal case (Figure 8.e), the CSCW approach enables more precise and well localized extraction of the S1 and S2 sounds compared to other methods. For the pathological signal with a late systolic murmur (Figure 9.e), the envelope appears less noisy than that produced by the Hilbert transform, particularly in the murmur region following S2.

These observations suggest that CSCW provides a cleaner and more structured envelope, highlighting its robustness in both normal and pathological conditions.

Figures 8.f and 9.f show the envelope results obtained using the multiscale product method for normal and pathological heart sounds, respectively. In both cases, the resulting envelopes show smoother and more uniform curves compared to the raw signals, improving the clarity and readability of the heart sound segmentation. In the normal case, the S1 and S2 components are clearly represented with minimal noise. Similarly, in the pathological case, the method effectively suppresses background noise while preserving the characteristics of the late systolic murmur between S2 and the following S1. Compared to other techniques such as Shannon energy, Shannon entropy, Hilbert transform and CSCW, the multiscale product method produces cleaner and more consistent envelopes, making it particularly suitable for reliable heart sound analysis.

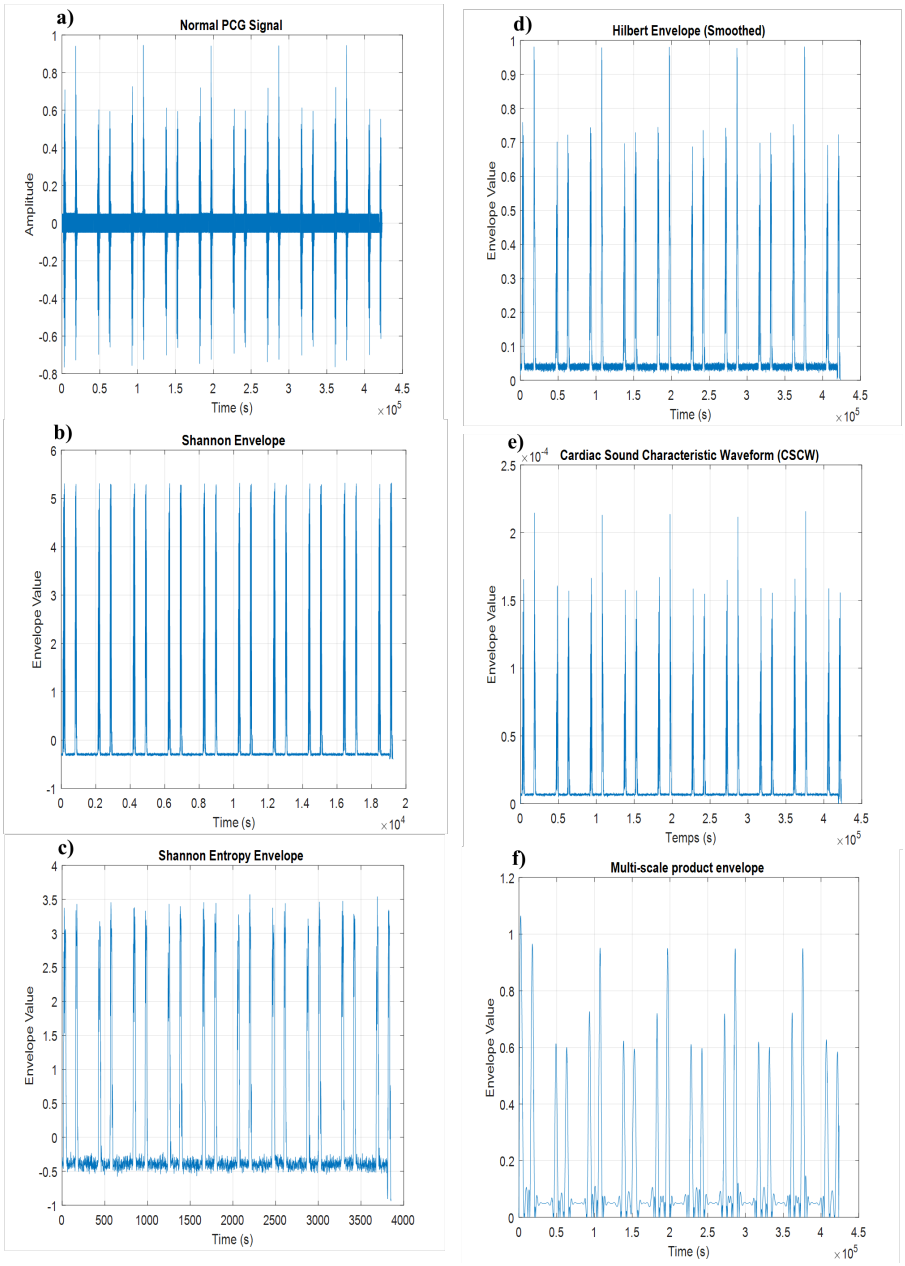

FIGURE 8. a) Normal PCG Signal, envelope extraction of: b) Shannon energy, c) Shannon entropy, d) Hilbert transform, e) CSCW, f) Multi-scale product

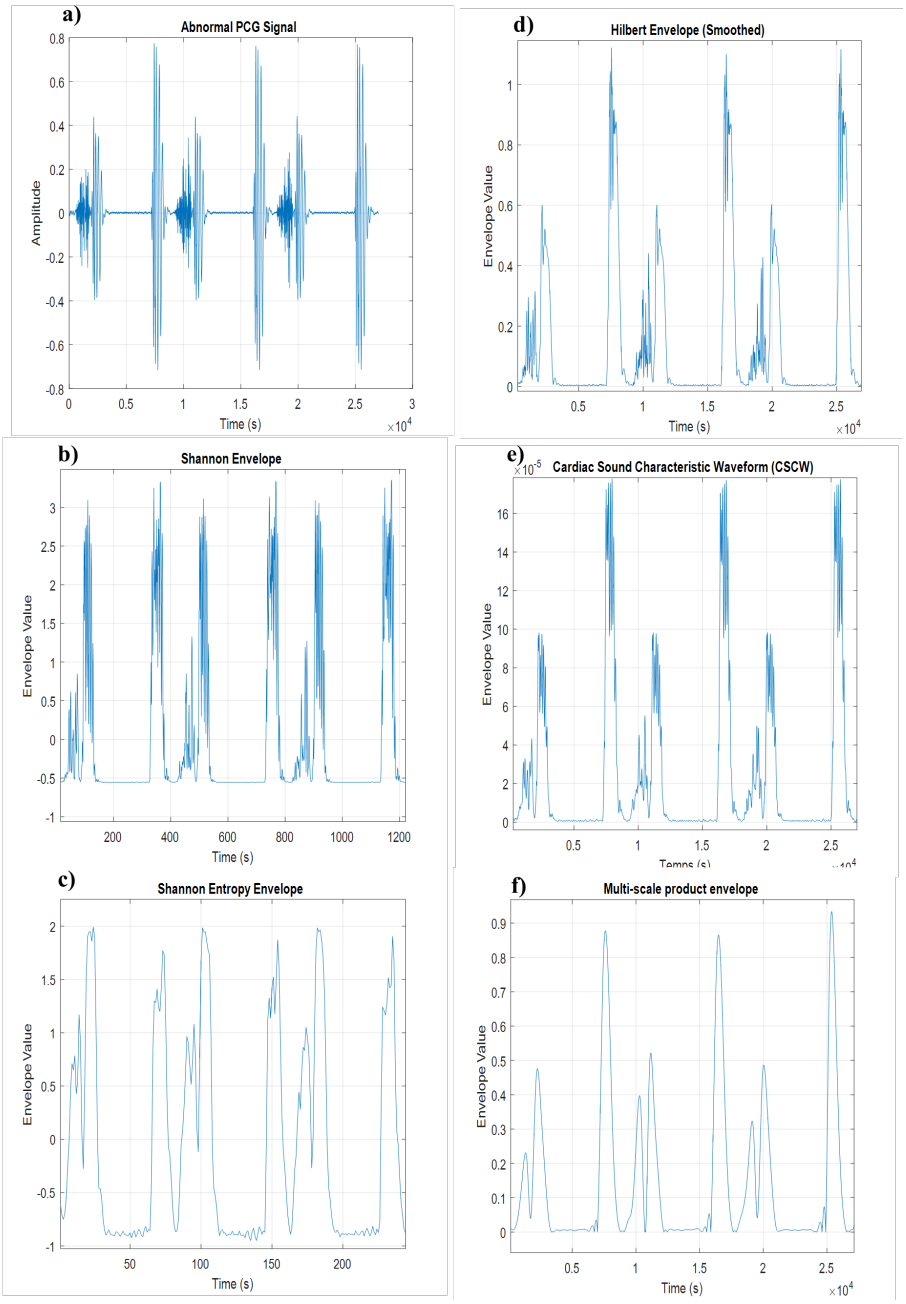

FIGURE 9. a) Heart murmur, envelope extraction of: b) Shannon energy c) Entropy d) Hilbert transform e) CSCW f) Multi-scale product

Figures 10 and 11 illustrate the effect of additive Gaussian noise on normal PCG signals at two different signal-to-noise ratio (SNR) levels: 0 dB and 5 dB, respectively. In the case of 5dB SNR (Figure 11), the analysis shows that all five envelope extraction methods, Shannon energy, Shannon entropy, Hilbert transform, CSCW and multiscale product, are able to correctly localise S1 and S2 heart sounds despite the presence of moderate noise. This is an indication of a certain robustness of the methods under conditions of relatively low noise.

Among the techniques evaluated, the multi-scale product method provides the most accurate localisation of S1 and S2 heart sounds, outperforming both the Shannon entropy and Hilbert transform approaches. The CSCW method also shows relatively high robustness and effectiveness. When the SNR is reduced to 0 dB (Figure 10), the analysis

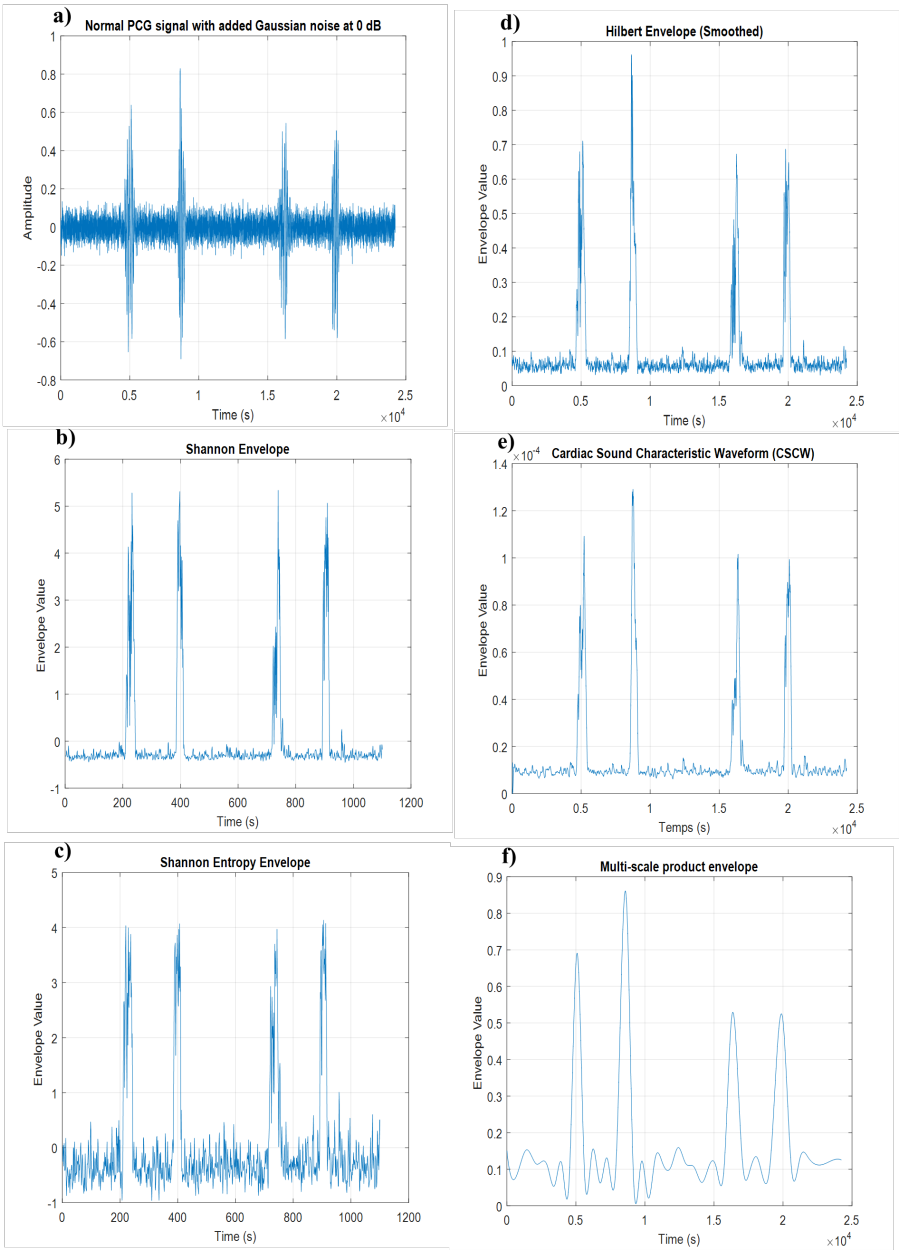

FIGURE 10. a) PCG signal with Added Gaussian noise (0 dB), Envelope extraction of: b) Shannon energy c) Entropy d) Hilbert transform e) CSCW f) Multi-scale product

shows that the Shannon energy and Hilbert transform methods are significantly affected by noise, making it difficult to accurately identify the S1 and S2 components. In contrast, the envelopes obtained using Shannon entropy (Figure 10.c) and CSCW (Figure 10.e) still allow reasonable segmentation of the heart sounds. Notably, the envelope extracted using the multi-scale product method (Figure 10.f) maintains a high level of accuracy in localising the S1 and S2 events even under severe noise conditions, highlighting its superior noise resilience.

Table 1 presents the sensitivity and positive predictive value (PPV) measurements performed on the entire dataset, comprising 140 heart sounds, for all the methods previously introduced. The row-by-row analysis of table 1 shows that the sensitivity and positive predictive value (PPV) of the tested methods are high when the SNR is 15 dB without

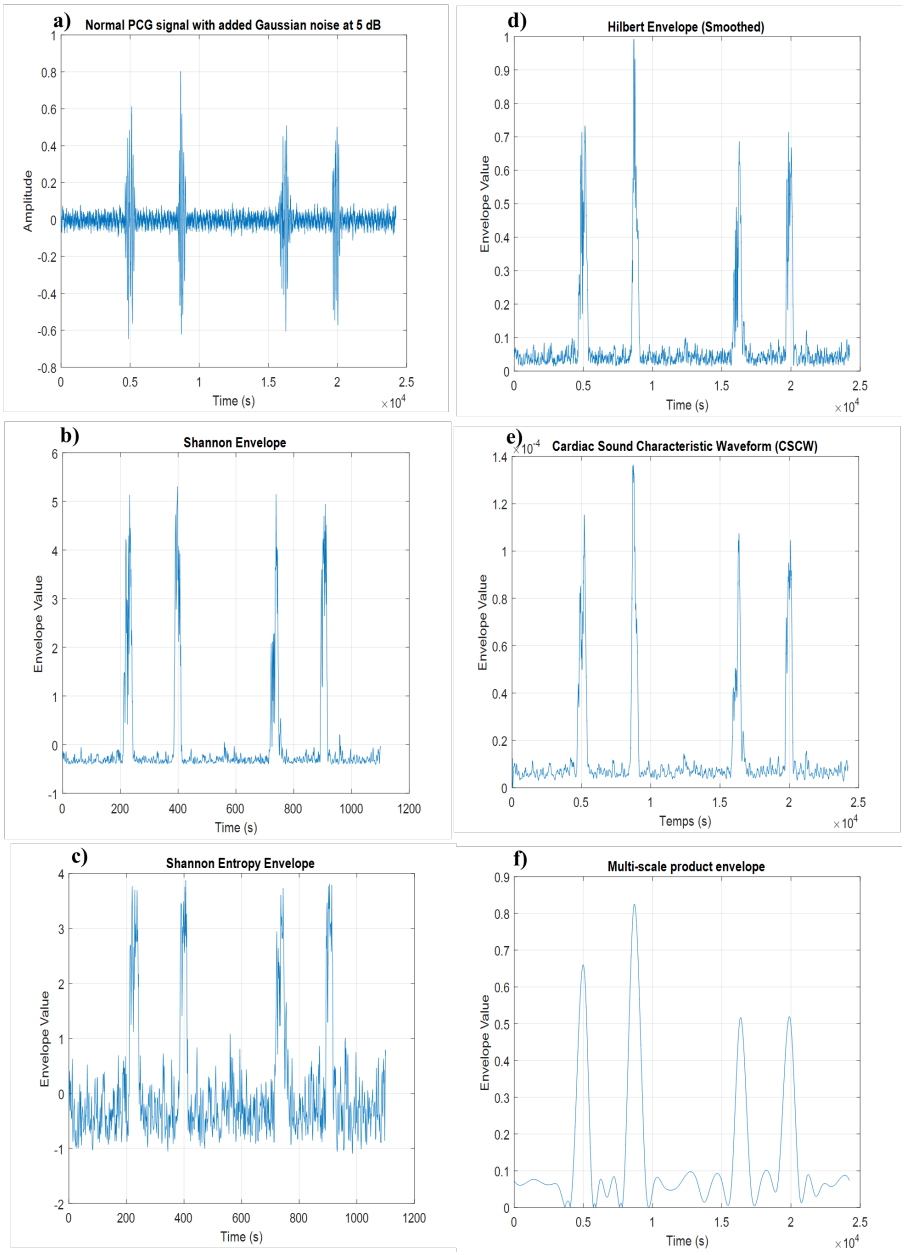

FIGURE 11. a) PCG signal with Added Gaussian noise (5 dB), envelope extraction of: b) Shannon energy c) Entropy d) Hilbert transform e) CSCW f) Multi-scale product

added noise. However, these performances significantly decrease when the SNR drops from 5 dB to 0 dB. The Shannon method offers good performance, with a sensitivity of 88% and a positive predictive value (PPV) of 90% when the signal-to-noise ratio (SNR) is 15 dB, in the absence of additional Gaussian noise. Even in the presence of higher noise levels, with an SNR of 5 dB, it maintains acceptable performance, achieving a sensitivity of 86% and a PPV of 85%. However, when the SNR decreases to 0 dB, its effectiveness significantly declines, with both sensitivity and PPV reduced to 76%. Furthermore, time-domain analysis methods, particularly the Shannon approach, prove to be highly sensitive to noise, which limits their reliability in heavily disturbed environments.

TABLE 1. Performance of five Heart Sound localization methods: Sensitivity and Positive Predictive Value across three noise levels

Methods	**SNR=15 dB**		**SNR=5 dB**		**SNR=0 dB**	
	Sens (%)	PPV(%)	Sens (%)	PPV(%)	Sens (%)	PPV(%)
Shannon	88	90	86	85	76	76
Entropy	91	90	89	87	84	81
Hilbert	89	90	85	85	78	77
CSCW	94	96	91	90	87	86
Multi-scale	96	98	92	96	89	88

Methods based on Shannon entropy and the Hilbert transform exhibit similar performance to the Shannon method when the signal-to-noise ratio (SNR) is between 15 dB and 5 dB.

Shannon entropy provides good results, with a sensitivity of 84% and a positive predictive value (PPV) of 81%. Additionally, the Hilbert-based method demonstrates comparable performance to the Shannon method at an SNR of 0 dB.

The characteristic waveform-based approach, which relies on frequency transformations, exhibits superior performance when the signal-to-noise ratio (SNR) is 0 dB, achieving a sensitivity of 94% and a PPV of 96%. This method proves to be robust against noise, particularly at an SNR of 5 dB, where it reaches a sensitivity of 91% and a PPV of 90%.

Although its performance slightly decreases at an SNR of 0 dB, it remains acceptable, with a sensitivity of 87% and a PPV of 86%.

Among the various approaches studied, the multi-scale product method stands out for its exceptional performance, surpassing those of Shannon, Shannon entropy, the Hilbert transform and the cardiac sound characteristic waveform method. This technique proves to be particularly robust against noise, maintaining high accuracy even at low SNR levels.

For an SNR of 15 dB, it achieves a sensitivity of 96% and a positive predictive value (PPV) of 98%, confirming its effectiveness under favorable conditions. Even when the SNR is reduced to 5 dB, its performance remains superior, with a sensitivity of 92% and a PPV of 96%.

At an SNR of 0 dB, although the performance slightly decreases, it remains competitive compared to other methods, with a sensitivity of 89% and a PPV of 88%.

One of the key strengths of this method lies in its time-frequency approach, which is based on the use of the Symlet 5 wavelet. This function enables a more refined signal analysis by effectively decomposing frequency components, thereby enhancing detection and robustness against noise. Thanks to this combination, the multi-scale product method emerges as a particularly reliable and high-performance solution, outperforming previously mentioned methods, especially in highly noisy environments.

Figure 12 and Table 2 comprehensively illustrate the performance of various methods for extracting the S1 and S2 envelopes from the PCG signal in the presence of noise. Based on the ROC curves shown in Figure 12, it is evident that at a clinical SNR of 15 dB, all methods yield promising results with curves that approach the upper left corner, indicating high sensitivity and a low false positive rate.

Table 2 quantifies these performances, showing, for example, AUC values of approximately 0.97 for the Multi-scale Product method and about 0.95 for the cardiac sound characteristic waveform method, compared to around 0.89 for Shannon, 0.905 for Entropy, and 0.895 for Hilbert. As the noise level increases (with SNR decreasing to 5 dB

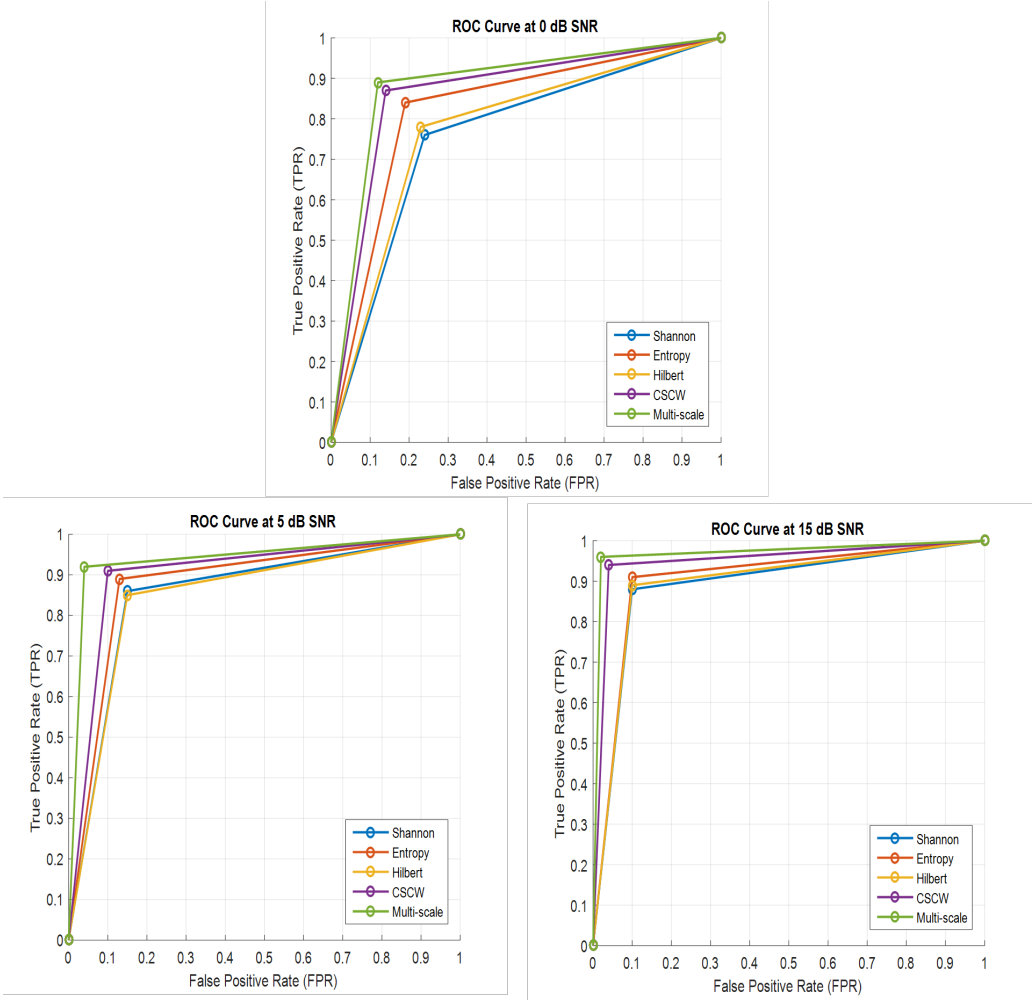

FIGURE 12. Variation of the ROC curve with added white noise

and then 0 dB), the ROC curves degrade and shift toward higher FPR values, leading to a reduction in AUC, the Multi-scale Product method declines to roughly 0.94 and then to 0.885, while the other methods experience more pronounced decreases (with Shannon dropping to about 0.855 and then 0.76). These observations suggest that although all methods perform competitively under low-noise conditions, the multi-scale Product and Characteristic Waveform methods exhibit superior robustness in noisy environments, making them the preferred choices for reliable cardiac sound extraction.

TABLE 2. Comparison of Performance Metrics (FPR and AUC) of Methods at Different SNR Levels

Methods	**SNR=15 dB**		**SNR=5 dB**		**SNR=0 dB**	
	FPR(%)	**AUC(%)**	**FPR (%)**	**AUC(%)**	**FPR (%)**	**AUC(%)**
Shannon	10	89	15	85.5	24	76
Entropy	10	90.5	13	88	19	82.5
Hilbert	10	89.5	15	85	23	77.5
CSCW	4	95	10	90.5	14	86.5
Multi-scale	2	97	4	94	12	88.5

6. **Conclusion.** In this comparative study, we analyzed several envelope extraction methods for the PCG (Phonocardiogram) signal to localize heart sounds S1 and S2, using the Pascal database. Five approaches were evaluated: the Shannon method, Shannon entropy, the Hilbert transform, the cardiac sound characteristic waveform method, and the multi-scale product method.

The results show that the multi-scale product method is the most effective, achieving high sensitivity (Se), positive predictive value (PPV), and an area under the ROC curve (AUC) of 0.973, confirming its robustness even in the presence of significant noise. The cardiac sound characteristic waveform method follows closely, with an AUC of 0.950, while the Shannon entropy and Hilbert transform methods exhibit moderate performance (AUC of 0.867 and 0.846, respectively). Finally, the Shannon method is the least effective, with an AUC of 0.832, indicating more uncertain classification and greater sensitivity to noise. This study highlights the importance of the choice of envelope extraction method to ensure reliable detection of S1 and S2 heart sounds, essential for automatic PCG analysis and screening for cardiovascular pathologies.

REFERENCES

[1] Y. He, W. Li, W. Zhang, S. Zhang, X. Pi, and H. Liu, "Research on segmentation and classification of heart sound signals based on deep learning," *Applied Sciences*, vol. 11, p. 651, 2021.

[2] Z. Dokur and T. Ölmez, "Feature determination for heart sounds based on divergence analysis," *Digital Signal Processing*, 2007.

[3] S. M. Debbal and F. Bereksi-Reguig, "Choix de l'ondelette analysante et classification des signaux phonocardiogrammes en fonction des souffles surajoutés," *Afrique SCIENCE*, vol. 1, no. 1, pp. 1–13, 2005.

[4] A. Atbi and S. Debbal, "Segmentation of pathological phonocardiogram signals using the Shannon energy envelogram," *American Journal of Cardiovascular Medicine*, vol. 2, no. 1, pp. 1–14, 2013.

[5] M. Nath, S. Srivastava, N. Kulshrestha, and D. Singh, "Detection and localization of S1 and S2 heart sounds by third-order normalized average Shannon energy envelope algorithm," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 235, no. 6, pp. 615–624, 2021.

[6] V. N. Varghees and K. I. Ramachandran, "Heart murmur detection and classification using wavelet transform and Hilbert phase envelope," in *Proc. Twenty-First National Conference on Communications*, Mumbai, India, pp. 1–6, 2015.

[7] S. Sun, Z. Jiang, H. Wang, and Y. Fang, "Automatic moment segmentation and peak detection analysis of heart sound patterns via short-time modified Hilbert transform," *Computer Methods and Programs in Biomedicine*, vol. 114, no. 3, pp. 219–230, 2014.

[8] L. H. Cherif and S. M. Debbal, "Algorithm for detection of the internal components of heart sounds and their split using a Hilbert transform," *Journal of Medical Engineering & Technology*, vol. 37, no. 3, pp. 220–230, 2013.

[9] A. Atbi, S. M. Debbal, F. Meziani, and A. Meziane, "Separation of heart sounds and murmurs by Hilbert transform envelogram," *Journal of Medical Engineering & Technology*, vol. 37, no. 6, pp. 375–387, 2013.

[10] D. L. Donoho, I. M. Johnstone, G. Kerkyacharian, and D. Picard, "Density estimation by wavelet thresholding," *Annals of Statistics*, vol. 24, pp. 508–539, 1996.

[11] D. Merritt and B. Tremblay, "Nonparametric estimation of density profiles," *Astronomical Journal*, vol. 108, no. 2, pp. 514–537, 1994.

[12] Z. W. Zhongwei and S. Samjin, "A cardiac sound characteristic waveform method for in-home heart disorder monitoring with electric stethoscope," *Expert Systems with Applications*, vol. 31, no. 2, pp. 286–298, 2006.

[13] Z. Yan, Z. Jiang, A. Miyamoto, and Y. Wei, "The moment segmentation analysis of heart sound patterns," *Computer Methods and Programs in Biomedicine*, vol. 98, no. 2, pp. 140–150, 2010.

[14] M. T. Pourazad, Z. Moussavi, and G. Thoma, "Heart sound cancellation from lung sound recordings using time–frequency filtering," *Medical and Biological Engineering and Computing*, vol. 44, no. 3, pp. 216–225, 2006.

[15] F. Li, M. Liu, Y. Zhao, L. Kong, L. Dong, X. Liu, and M. Hui, "Feature extraction and classification of heart sound using 1D convolutional neural networks," *EURASIP Journal on Advances in Signal Processing*, 2019, Art. no. 59.
[16] Q. Zhao, C. Niu, F. Liu, L. Wu, X. Ma, and G. Huang, "Accuracy of cardiac auscultation in detection of neonatal congenital heart disease by general paediatricians," *Cardiology in the Young*, vol. 29, pp. 679–683, 2019.
[17] Y. Huang *et al.*, "Large scale application of pulse oximeter and auscultation in screening of neonatal congenital heart disease," *BMC Pediatrics*, vol. 22, p. 483, 2022.
[18] P. Bentley, G. Nordehn, M. Coimbra, S. Mannor, and R. Getz, "The PASCAL classifying heart sounds challenge (CHSC2011) results," 2011.
[19] O. Deperlioglu, U. Kose, D. Gupta, A. Khanna, and A. K. Sangaiah, "Diagnosis of heart diseases by a secure Internet-of-Health-Things system based on an autoencoder deep neural network," *Computer Communications*, vol. 162, pp. 31–50, 2020.
[20] Y. Jeong, J. Kim, D. Kim, J. Kim, and K. Lee, "Methods for improving deep learning–based cardiac auscultation accuracy: Data augmentation and generalization," *Applied Sciences*, vol. 11, p. 4544, 2021.
[21] H. Liang, S. Lukkarinen, and I. Hartimo, "Heart sound segmentation algorithm based on heart sound envelogram," Helsinki University of Technology, 1997.
[22] N. N. Binta, K. N. Shoyad Ibn Sabur, and H. Taufiq, "Hilbert-envelope features for cardiac disease classification from noisy phonocardiograms," *Biomedical Signal Processing and Control*, vol. 78, 2022.
[23] A. Özkan and K. Mustafa, "Effect of Hilbert–Huang transform on classification of PCG signals using machine learning," *Journal of King Saud University — Computer and Information Sciences*, vol. 34, no. 10, pp. 9915–9925, 2022.
[24] S. Choi and J. Zhongwei, "Comparison of envelope extraction algorithms for cardiac sound signal segmentation," Micro Mechatronics Laboratory, Yamaguchi University, Japan, 2006.
[25] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. New York: Chapman & Hall, 1986.
[26] A. Yadollahi and Z. M. Moussavi, "A robust method for heart sounds localization using lung sounds entropy," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 3, pp. 497–502, 2006.
[27] M. Feldman and S. Braun, "Description of free responses of SDOF systems via the phase plane and Hilbert transform: Concepts of envelope and instantaneous frequency," in *Proc. SPIE*, Orlando, FL, vol. 3089, pp. 973–979, 1997.
[28] Y. Meyer, *Wavelets and Operators*. Cambridge, UK: Cambridge Univ. Press, 1992.
[29] J. Gnitecki and Z. Moussavi, "Variance fractal dimension trajectory as a tool for heart sound localization in lung sounds recording," in *25th Annu. Int. Conf. IEEE*, vol. 3, pp. 2420–2423, 2003.
[30] M. M. Rahman, S. Sarker, and M. M. Islam, "Underwater active noise cancellation combining Kalman filter with FxLMS," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 14, no. 1, pp. 31–40, Mar. 2023.