

Nonlinear Heart Rate Variability Analysis for Sleep Stage Classification Using Integration of Ballistocardiogram and Apple Watch

Dominic Jaworski ^{1,2}, Edward J Park^{1,2}

¹Mechatronic Systems Engineering, Simon Fraser University, Surrey, BC, V3T 0A3, Canada; ²WearTech Labs, Simon Fraser University, Surrey, BC, V3V 0C6, Canada

Correspondence: Edward J Park; Dominic Jaworski, Surrey Centre 2 Building, 9639 - 137A Street, Unit 206, Surrey, BC, V3V 0C6, Canada, Email ed_park@sfu.ca; djaworsk@sfu.ca

Purpose: Wearable or non-contact, non-intrusive devices present a practical alternative to traditional polysomnography (PSG) for daily assessment of sleep quality. Physiological signals have been known to be nonlinear and nonstationary as the body adapts to states of rest or activity. By integrating more sophisticated nonlinear methodologies, the accuracy of sleep stage identification using such devices can be improved. This advancement enables individuals to monitor and adjust their sleep patterns more effectively without visiting sleep clinics.

Patients and Methods: Six participants slept for three cycles of at least three hours each, wearing PSG as a reference, along with an Apple Watch, an actigraphy device, and a ballistocardiography (BCG) bed sensor. The physiological signals were processed with nonlinear methods and trained with a long short-term memory (LSTM) model to classify sleep stages. Nonlinear methods, such as return maps with advanced techniques to analyze the shape and asymmetry in physiological signals, were used to relate these signals to the autonomic nervous system (ANS). The changing dynamics of cardiac signals in restful or active states, regulated by the ANS, were associated with sleep stages and quality, which were measurable.

Results: Approximately 73% agreement was obtained by comparing the combination of the BCG and Apple Watch signals against a PSG reference system to classify rapid eye movement (REM) and non-REM sleep stages.

Conclusion: Utilizing nonlinear methods to evaluate cardiac dynamics showed an improved sleep quality detection with the non-intrusive devices in this study. A system of non-intrusive devices can provide a comprehensive outlook on health by regularly measuring sleeping patterns and quality over time, offering a relatively accessible method for participants. Additionally, a non-intrusive system can be integrated into a user's or clinic's bedroom environment to measure and evaluate sleep quality without negatively impacting sleep. Devices placed around the bedroom could measure user vitals over longer periods with minimal interaction from the user, representing their natural sleeping trends for more accurate health and sleep disorder diagnosis.

Keywords: ballistocardiogram, wearable, heart rate variability, nonstationary signals

Introduction

Maintaining a regular sleep schedule of good quality restful sleep is beneficial for a healthy life.¹ Good sleep can enhance alertness and daily performance, while lack of sleep can increase the risk of various diseases such as obesity, hypertension, and cardiovascular issues.^{1,2} Keeping a consistent sleep schedule may be difficult due to many activities that can be carried out at night, such as personal hobbies, or due to sleeping patterns related to differences between workdays and non-workdays, as well as shift work that changes working hours based on job requirements and results in re-adjusting sleeping schedules.^{3,4} This phenomenon, known as social jetlag, can disrupt the body's biological clock, resulting in daytime drowsiness.⁴ Therefore, managing daily activities to leave enough time to sleep well is very important for a healthy life.

Polysomnography (PSG) is recognized as the gold standard for assessing sleep quality, capturing a wide range of physiological signals simultaneously via electrodes and sensors attached to the body.⁵ Typically conducted within the confines of sleep clinics, PSG studies necessitate an extensive array of sophisticated equipment and the expertise of trained professionals for accurate classification of sleep stages. This requirement for specialized infrastructure and personnel not only adds to the discomfort but also renders the procedure impractical for routine application.⁶ Alternatively, a variety of consumer-grade devices, including smartphones, smartwatches, and specialized sleep monitoring tools, are accessible to the broader public. These devices, designed to be worn or placed in or around the bed, offer a non-intrusive approach to track sleep activity.⁶ They enable the long-term monitoring of sleeping cycles, facilitating the improvement of users' sleep patterns, and contributing to more restful sleep nightly.

Consumer health-monitoring wearables, often worn on the arm-like smartwatches, are equipped with sensors and algorithms that effectively track vital signs throughout the day.⁷ Wearable devices have been shown to have good agreement of sleep stages compared with PSG.^{8,9} For instance, the Apple Watch, which is equipped with optical and electrical cardiac sensors as well as sensitive accelerometers and gyroscopes, offers a native sleep quality analysis function since the release of the Apple Watch Series 6.^{10,11} Open-Source software kits are available from Apple Inc. for developers to create health solutions from the sensor data that could benefit clinicians and patients.¹² Consumer devices can connect to other user devices to transfer health information through wireless networks making them a convenient option to monitor their vitals and make appropriate adjustments to live a healthier lifestyle. However, many consumer wearables are non-clinical fitness devices that monitor one or a few physiological parameters with the goal of providing users insight into their sleeping habits.⁵ Since only a few physiological properties are measured, these devices would not be able to evaluate every kind of sleep disorder without more sensors.¹³ Nonetheless, the benefit of non-intrusive devices with sophisticated sensors and software is that they provide an accessible way for consumers to monitor their health effectively without expensive clinical equipment.^{7,9,13}

While wearables can be conveniently worn throughout the day, dedicated devices positioned near the sleep area in the bedroom can provide additional insight into a user's sleep quality and habits. One type of such device is known as ballistocardiography (BCG), which measures cardiac activity through vibrations from a mattress during sleep.¹⁴ More specifically, the recoil forces from the heartbeat and blood being pumped throughout the body can be measured by the BCG and evaluated to show heart rate activity like an ECG.¹⁵ In ECG, heart rate signals are identified with the QRS complex that reflecting systolic and diastolic activity in the heart with the R wave being a large upward deflection and generally used to evaluate cardiac signals.¹⁶ On the other hand, BCG uses analogous complexes noted as IJK, reflecting the use of a different evaluation methods.¹⁷ The I wave occurs during the systole from the blood accelerating into the aorta; the J wave shows a sharp amplitude when the direction of blood changes; and the K wave is caused by the deceleration of the blood flow, leading to the diastolic portion of the cardiac cycle.¹⁶ The IJK complexes also have more noise from movement artifacts and external vibrations, as well as lagging slightly from QRS complexes due to the time it takes for the vibrations to travel from the body to the sensor.¹⁷ Filtering and signal processing techniques are required to clean up the BCG signals into usable cardiac activity, with appropriate filtering methods used for specific sensors.^{14,16}

The time between the J–J peaks in BCG signals is referred to as heart rate variability (HRV).¹⁸ The changes in HRV can reflect the activity of the ANS, which controls the body's restful or active states, and indicate different sleep stages.¹⁹ Particularly, the high and low frequency components of HRV can signify the transition between wake and sleep stages, with a reduction of HRV showing a transition into the REM sleep stage.¹⁹ Cardiac activity and HRV have nonlinear and nonstationary characteristics that can be analyzed by nonlinear methods to show physiological autonomic changes that can be classified into different sleep stages.^{19,20} In 1996, a task force was established to standardize the HRV measurement methods, focusing on time, frequency, and nonlinear domains to assess the ANS.²¹ Although nonlinear methods showed promise for HRV analyzing, they required further development and evaluation before being standardized at the time.²¹ Since then, however, nonlinear methods have become standard for HRV analysis, as physiological signals are nonstationary and nonlinear, essential for regulating the body and maintaining stability during various activities.^{22,23}

BCG signals are inherently noisy due to the movement, creating artefacts that can affect the overall signal quality measured by the sensitive mechanical sensors.¹⁶ While an ECG utilizes electrodes attached directly to the body to measure electrical signals to analyze cardiac activity, these signals are not as affected by movement as BCG signals.¹⁶

Therefore, BCG signals need additional preprocessing to remove any movement-induced noise to perform an equivalent analysis to that of ECG, and the extra filtering could result in losing some vital information from the signals.¹⁶ However, nonlinear methods that can capture the dynamic nature of cardiac signals can be applied with minimal preprocessing and without over-filtering of BCG signals. Additionally, the body reacts and adapts to situations that are reflected by the changing cardiac activity when someone is resting or active.²⁰ Classical spectral analysis for HRV was under the assumption that cardiac signals contained linear and stationary such as requiring interpolation into evenly spaced intervals which could lose some valuable information from any biases incurred.^{19,24} Therefore, measuring these nonlinear changes in HRV with methods that can directly analyze the dynamic nature of the body could provide better insight into some of the subtle changes during the transitions among sleep stages.

Cardiac signal analysis with a BCG is usually performed similarly to an ECG, following the recommendations and guidelines set out by the task force for features and model training.^{21,22} However, using more specific techniques that can take advantage of the nonlinear and nonstationary nature of cardiac signals, as well as the unique signals from a BCG, could evaluate sleep quality equivalently to ECG but in a non-contact way.²⁰ BCG signals also produce a nonlinear and nonstationary cardiac signal due to the inconsistent J-peaks from motion artifacts.²⁰ However, the BCG signal can measure a good quality cardiac signal similar to an ECG signal with appropriate signal preprocessing and effectively be analyzed through nonlinear methods.^{25,26} Nonlinear methods offer a more accurate modeling of physiological signal dynamics, providing greater insights into cardiac state and improving sleep stage classification with non-intrusive devices.²⁷

Machine learning methods have demonstrated effective classification of sleep stages with various devices and have shown good results compared to PSG.²⁸ Applying machine learning with nonlinear methods for HRV could further improve the classification of sleep stages using the non-intrusive devices. Additionally, the Internet of Things (IoT) is a technology that can connect devices, sensors, and software together through a shared network to benefit patients and healthcare professionals.²⁹ This is particularly advantageous for sleep monitoring, as devices can collect data and transmit it to a database, allowing medical professionals to monitor patients' conditions without requiring an overnight in a laboratory or clinic. Therefore, patients can maintain their daily routines without disruption while being monitored to manage their sleeping habits.

The study in this manuscript aims to expand on the methods for detecting sleep quality with non-intrusive devices and continues from our previous publications that acted as pilot studies to formulate the experiments and methods for this study.^{30,31} Combining a non-contact BCG in the bed mattress and a wearable device like the Apple Watch to measure physiological data while they sleep. Nonlinear methods and advanced techniques are explored to measure the nonlinear and nonstationary nature of BCG cardiac signals as the participants transition from wake to sleep stages. Additionally, machine learning is used to classify the sleep stages from the expanded nonlinear features. The goal of this study was to advance methods for evaluating sleep quality with non-intrusive devices and provide additional support for future studies looking into sleep quality measurements with similar system environments.

Materials and Methods

Experiment Protocol

This study was approved by the Simon Fraser University Research Ethics Board (#20170629) and complies with WMA Declaration of Helsinki ethical principles. Participants were provided with detailed information about the protocol prior to the study commencement to ensure they were fully informed. After comprehensively understanding and agreeing to the study's requirements, participants indicated their consent by signing informed consent forms. The experimental procedure consisted of participants coming into a designated sleep lab and sleeping for three separate days and at least for three hours, while sensors collected data on and around their bodies. This timeframe was chosen based on initial testing to see if participants could enter every sleep stage in a shorter sleep cycle. Additionally, the three-hour duration was chosen to collect sleep cycles from different participants on the same day for data collection. The study consisted of three nights to alleviate some of the first night effect from PSG, which may cause poor sleep readings from participants as they are not accustomed to all the electrodes and sensors attached to their body during sleep.³² All three days were included in the experimental analysis as patients reported similar sleeping experiences for each night. The experiment consisted of six healthy participants (two female and four male) ranging in age from 25 to 40, with average BMI and no reported sleep

disorders, totaling 18 sleep sessions. A limitation of this study is the number of participants; however, the three sleep sessions per participant helped in mitigating the low number of participants. Future studies will include an expanded number of participants to strengthen the results. Upon arrival to the sleep lab, the participant put on the Natus Embletta X100 PSG by Embla Systems Inc (Broomfield, CO, USA) and the electrodes that would act as a gold standard reference to the rest of the proposed systems and devices. Additionally, the participants wore an Apple Watch Series 8 (Apple Inc., Cupertino, CA, USA) and a GT9X Link (ActiGraph LLC, Pensacola, FL, USA) on their non-dominant arm, which ended up being the left arm for all participants. The BCG device was the SCA11H bed sensor by Murata Manufacturing Co., LTD (Nagaokakyo, Japan) that was placed inside the bed. [Figure 1](#) shows the entire experimental setup and a participant wearing the devices during their sleep. An overview of the IoT system is available in a previous publication by our team and is seen on the table in [Figure 1B](#).³⁰ Each device was calibrated before the participant went to sleep, following the user manual's recommendations:

- GT9X Link performed a ten-second calibration upon removal from its USB-charging dock.
- Participants sat still for five minutes while the Apple Watch mindfulness app was activated, with the heart rate measured in BPM.
- BCG calibration involved running its calibration mode for 1 minute on an empty bed and 1 minute with the participant lying on the bed.^{33,34}

The study ended when participants woke up naturally after three hours. Data was downloaded to local computers from each device for processing and analysis. The experiment was repeated for two more days with the same procedure. All data from the experiment was stored on lab computers that only researchers had access to, and all data from sensors used anonymous user profiles without any descriptors to maintain participant privacy and security.

Polysomnography and Sleep Detection

The PSG system consisted of electroencephalogram (EEG), electrocardiogram (ECG), electrooculogram (EOG), electromyogram (EMG), photoplethysmogram (PPG), respiration bands, and nasal cannula, as shown in [Figure 1A](#) and [C](#). The PSG was used as a gold standard reference for the sleep stage classification obtained from the IoT sleep system and devices. The data were downloaded into the Remlogic software provided by Natus, which was then scored visually using

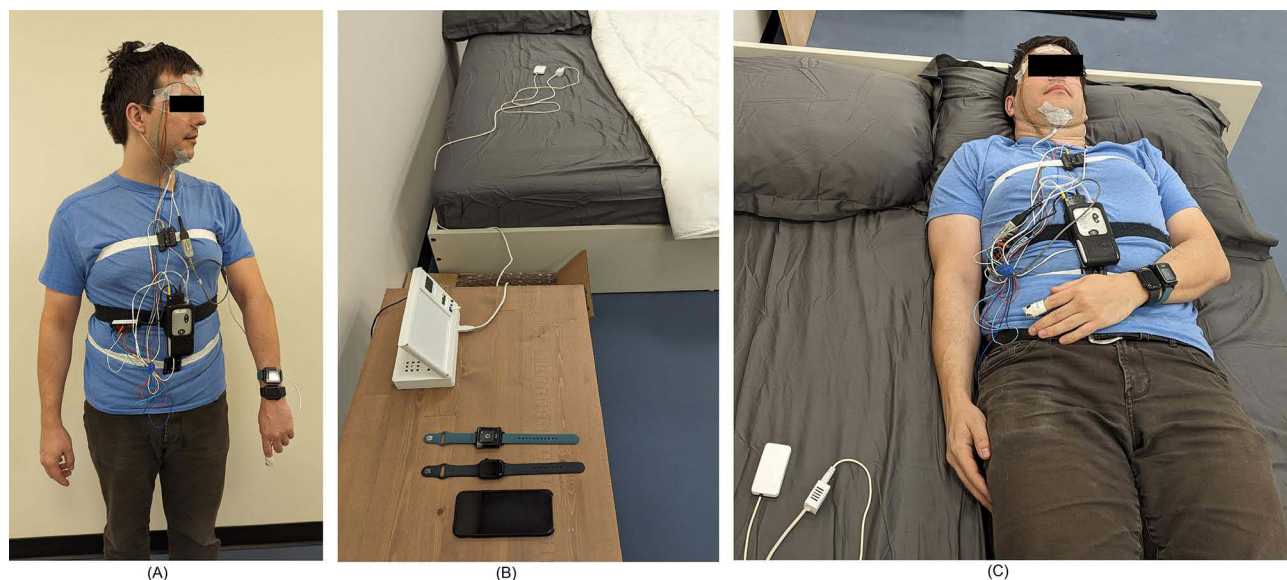


Figure 1 (A) Participant with all devices attached to their body. (B) Sleep IoT device on bedside table with the Apple Watch, iPhone and GT9X Link prior to experiment. (C) Participant lying in bed with all devices beside the BCG device that is clamped to the bed sheet. Written informed consent was obtained from the participant to use an image in this manuscript.

the rules and guidelines established by the American Academy of Sleep Medicine (AASM).³⁵ In addition to the PSG, the GT9X link is a research-grade actigraph that was used to provide an additional reference by estimating sleep/wake instances through limb movements with the Cole–Kripke method.^{7,36} The room was kept at around 21°C, and participants were provided blankets to ensure maximum comfort during sleep.

Ballistocardiography and Environmental Device

Murata's BCG device collects various cardiorespiratory signals (heart rate, respiration rate, heart stroke volume, HRV, and J–J interval time) and samples them at 1 Hz over a wireless network through highly sensitive accelerometers.³⁷ Clinical testing has validated Murata's BCG sensor, showing that its cardiac signals closely matched the readings obtained from the PSG.^{16,33}

Furthermore, low pass filtering was used to remove movement artifacts within the signals, as detailed in the above clinical trials from Murata. The non-invasive nature of the BCG sensors, placed within the bed before participant arrival, ensured undisturbed sleep by eliminating any direct interaction with the device.

Apple Watch

For the experiment, an Apple Watch Series 8 (OS version 9.6.3) was linked to an iPhone x (OS version 16.7.2) during the experiment. We had recently validated the Apple Watch against the BCG and PSG during sleep cycles, along with some of its health measurement features.³¹ The results showed that the Apple Watch physiological sensor readings were closely accurate compared to the other devices. However, an issue was identified with the Apple Watch exporting vitals signal signals at an inconsistent sampling rate when exported for analysis. In this study, this issue was alleviated by utilizing the workout app's cross-training mode, which significantly improved the sampling rate to around 3–7 seconds between samples. Consequently, linear interpolation was applied to match the 1 Hz sampling rate of the BCG. Note that only heart rate data in beats per minute (BPM) was available for extended sampling, whereas instantaneous HRV and QRS complex data were limited to 30-second intervals via the Apple Watch's ECG app. Therefore, HRV was estimated by dividing the heart rate (in BPM) by 60,000 milliseconds (1 minute) as per the following equation:³⁸

$$HRV = \frac{60000}{HR} \quad (1)$$

Heart Rate Variability Analysis

The conventional evaluation of HRV typically occurs in the intervals of at least 5 minutes, aligning with many established norms, whereas sleep stage scores are assessed in 30-second sequential intervals.^{22,35} To synchronize the data from all devices, this study processed signals within a 5-minute sliding window that advanced in 30-seconds increments. While employing the metrics and norms for HRV analysis as outlined by the task force, this study also expanded on them into the nonlinear domain.^{21,22} It is important to note that the application of nonlinear HRV analysis methods in this study represents a relatively novel approach, as most such methods have traditionally been applied to ECG signals.^{22,27,39} However, BCG signals also exhibit nonlinear and nonstationary characteristics, which are even noisier than ECG signals, making nonlinear analysis a suitable approach to model the physiological dynamics.²⁰ The BCG signals, measuring the heart's mechanical effect on the body, are distinct in nature from ECG signals, which measure the electrical activity of the heart. This distinction necessitates further evaluation to determine which nonlinear features are relevant for accurate BCG-based HRV analysis. This aspect is important due to the inherent differences between BCG and ECG signals, underscoring the need for this study's comprehensive examination of relevant nonlinear features for effective HRV analysis using BCG signals.

In this context, an additional time domain metric is introduced in this study, which leverages the relative J–J interval difference (analogous to the R–R intervals used ECG studies).⁴⁰ These intervals are weighted by the means to enhance the robustness of the analysis. The relative J–J intervals are calculated by the following equation:

$$rr_i = \frac{2(JJ_i - JJ_{i-1})}{JJ_i + JJ_{i-1}} \quad (2)$$

where rr is the relative variation of the intervals, i is the current relative interval, and JJ are the J–J interval times. The return map of the consecutive relative J–J intervals can show heartbeat dynamics through the Euclidean distances to the center and identify outliers through the median or annular intensity from the interquartile range.⁴⁰ Similarly, a nonlinear method that also uses return maps of consecutive J–J intervals for analysis is known as the Poincare plot, which is insensitive to trends in inter-beat intervals and a great tool for analysis of noisy signals from wearable and noncontact devices.^{22,41} Poincare plots are usually evaluated by the shape and area of the ellipse drawn around the scattered J–J intervals with the standard deviations representing the length and width of the ellipse represented by the following equations:^{22,40}

$$SD1 = \sqrt{\frac{1}{2} \cdot \sigma(JJ_{i+1} - JJ_i)} \quad (3)$$

$$SD2 = \sqrt{\frac{1}{2} \cdot \sigma(JJ_{i+1} + JJ_i)} \quad (4)$$

where $SD1$ and $SD2$ are the width and length, respectively, and σ is the variance of the scattered points of the Poincare plot. Both measures are related to the baroreflex sensitivity of cardiac activity and the ratio $SD1/SD2$ measures the ANS balance that could show if the body is in a more sympathetic (SNS) or parasympathetic nervous system (PNS) activation.²² Additionally, deeper analysis on Poincare plots that looks at the asymmetry of the inter-beat intervals and can show the imbalances of activity between the SNS and PNS.³⁹ Figure 2 shows an example of the Poincare plot of consecutive J–J intervals.

The deceleration and accelerations of the inter-beat intervals are what cause the heart rate asymmetry and have been modeled through various indices about the 45° line of identity, which passes through the origin of the Poincare plot shown in Figure 2 as the red line.^{39,42}

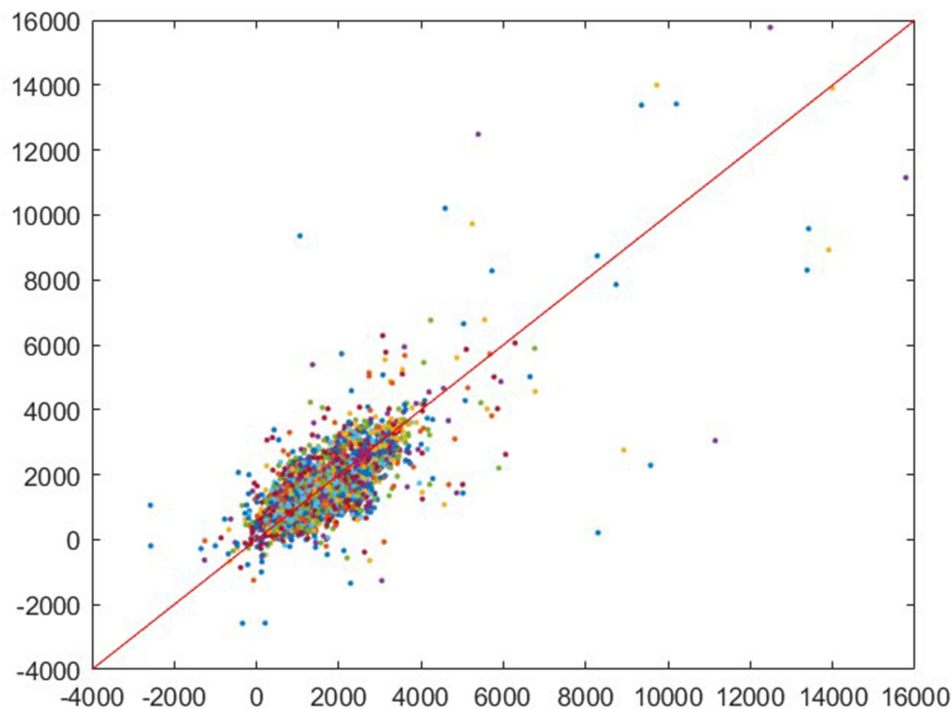


Figure 2 Poincare plot for J–J intervals from the BCG device.

The following section briefly describes each of the asymmetry indexes used in this study and their accompanying equations.^{39,43} The Porta index shows the ratio of how many points are below the line of identity:

$$PI = \frac{b}{m} \times 100 \quad (5)$$

with b being the point below and m being the total number of points. The Guzik index is the ratio of all the distances of the points above the line of identity over all the point distances:

$$GI = \frac{\sum_{i=1}^l D_i}{\sum_{i=1}^m D_i} \times 100 \quad (6)$$

$$D_i = \frac{|JJ_{i+1} - JJ_i|}{\sqrt{2}} \quad (7)$$

with l being the number of points above the line of identity, m being all the points, D_i is the distance of each point to the line of identity and is calculated by Equation (7). Next, Ehlers' index shows the skewness of the points with respect to the line of identity:

$$EI = \frac{\sum_{i=1}^{N-1} (JJ_i - JJ_{i+1})^3}{\left(\sum_{i=1}^{N-1} (JJ_i - JJ_{i+1})^2\right)^{3/2}} \quad (8)$$

Finally, the slope indexes are calculated by the ratio of the phase angle of all points above the line of identity, while the area index calculates the area from the phase angle to the line of identity:

$$\theta_i = \text{atan}\left(\frac{JJ_{i+1}}{JJ_i}\right) \quad (9)$$

$$R\theta_i = \theta_{LI} - \theta_i \quad (10)$$

$$S_i = \frac{1}{2} \times R\theta_i \times r^2 \quad (11)$$

$$SI = \frac{\sum_{i=1}^l |R\theta_i|}{\sum_{i=1}^m |R\theta_i|} \times 100 \quad (12)$$

$$AI = \frac{\sum_{i=1}^l |S_i|}{\sum_{i=1}^m |S_i|} \times 100 \quad (13)$$

where θ is the phase angle of the point, $R\theta$ is the phase angle from the line of identity to the point, S_i is the area of the sector with r being the radius of the sector, SI is the slope index ratio, and AI is the area index ratio.^{39,43}

Another scatter plot method used in this study is sequential trend analysis (STA). STA plots the differences of successive J–J intervals with the first value plotted on the x-axis and the consecutive value plotted on the y-axis.⁴⁴ The data points placed into quadrants that represent increasing/decreasing consecutively or points that are alternating in their activation. Consecutively decreasing data points correspond to the SNS while consecutively increasing points are for the PNS with the other quadrants representing the transition between the states.^{44,45} To evaluate which quadrants are the most active, the mean of all the distances for the point in each quadrant is evaluated with the following equation:⁴⁶

$$\frac{1}{n-1} \sum_{k=1}^{n-1} \sqrt{JJ_k^2 + JJ_{k+1}^2} \quad (14)$$

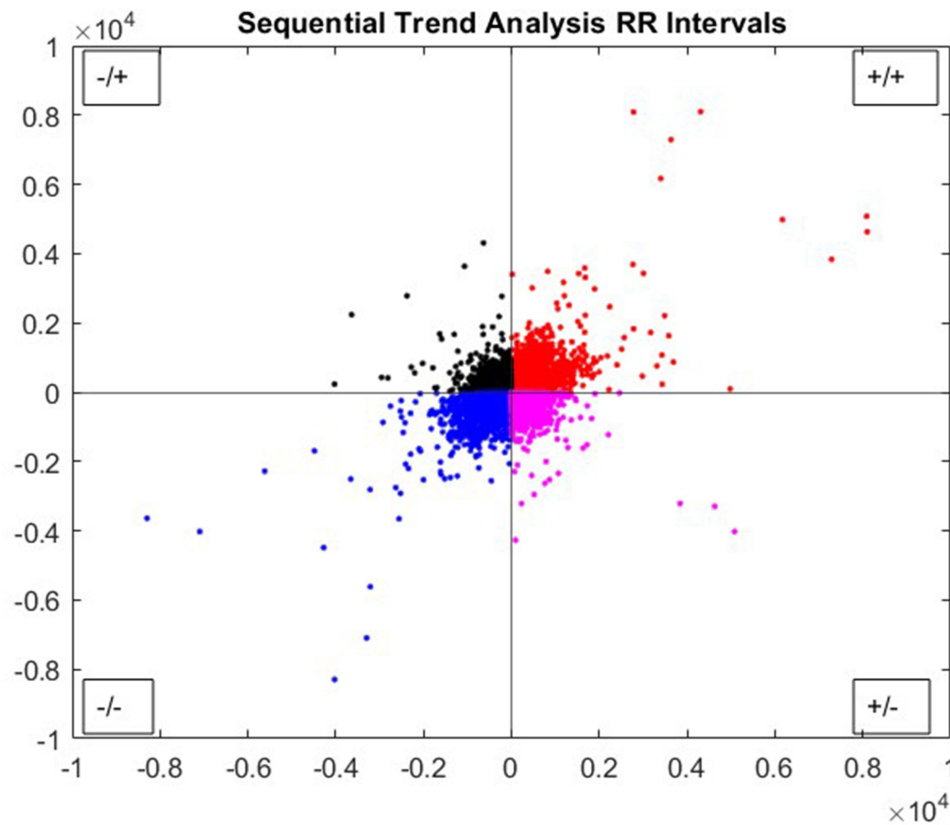


Figure 3 Sequential trend analysis plot from the BCG device.

where n is the total number of J–J intervals. Many of the Poincare plot methods in this study build off each other and can be effective for analysis as many features can be extracted from the plots. Figure 3 shows an example of the STA plot and the consecutive J–J intervals of the quadrants.

The last nonlinear method used is the Renyi entropy that is effective for short time series, nonlinear, and nonstationary signals like HRV.²⁷ This is beneficial for the BCG signals obtained in this study as the hardware limitations can cause some of the HRV readings to be distorted from noise and artifacts.²⁰ Renyi entropy is generalized from Shannon entropy and is briefly shown in the following equations:^{27,47}

$$H(\alpha) = \frac{1}{1 - \alpha} \log_2 \left(\sum_{i=1}^n \rho_i^\alpha \right) \tag{15}$$

$$\rho_i = \frac{1}{\sigma \sqrt{2\pi}} \sum_{j=1}^n e^{-\frac{dist_{ij}^2}{2\sigma^2}} \tag{16}$$

$$dist_{ij} = \sum_{k=0}^{\pi} (x_{i+k} - x_{j+k})^2 \tag{17}$$

where n is the number of values, α order of entropy measure, ρ_i is the probability from the random variable from n values and the Gaussian kernel, σ is the dispersion of the function, π is the number of J–J intervals designated from the literature and $dist$ is the sum of Euclidean distances between sequential J–J intervals designated as x in Equation (17).²⁷ Additionally, the J–J intervals were decomposed into magnitude, sign, and accelerations to be evaluated with the Renyi entropy based on the following works.^{27,48}

Additionally, the signals obtained from the Apple Watch and BCG are compared for any offset between them. The Apple Watch and BCG should provide similar physiological signals but the difference in sensing methods and movement artifacts, particularly in BCG, could distort the signal acquisition. Additional features that look at the

correlation, differences, and increasing/decreasing segments are included in the model. Dynamic time warping was also used to align and compare any differences between the signals.⁴⁹ The features are trained with a bi-directional long short-term memory (LSTM) neural network based on previous studies for sleep stage classification.^{30,50} This was used to capture the temporal order of the sleep cycle as it can store memory from the time series and generate outputs on the current time step.⁵⁰ The sleep cycle begins initially getting into bed wake, a middle section where the bulk of the sleep stages occur and ending with wake and getting out of bed. This study utilized the machine and deep learning toolbox within Matlab 2023b for the LSTM model and consisted of five layers. Beginning with the sequence input layer that added the data to the model and applied data normalization and followed by the bi-directional LSTM layer, consisting of 150 hidden units and the output mode set to the sequence. The fully connected layer processed the inputs from the biLSTM layer into the sleep stage classes into a softmax function layer. Finally, the classification layer calculated the weight sleep stages classes and output them for each epoch during the sleep cycle. This network configuration received all the processed sensor data from each device to classify sleep stages by 30 second epochs with respect to the sleep cycle temporal order.

Results

The sleep stage classifications from the LSTM model, which was trained on physiological signals collected from the BCG bed sensor and Apple Watch for each participant, are presented as accuracy and agreement compared to PSG in Table 1. Additionally, similar models were trained separately on physiological signals obtained from each device to compare their combinations. The BCG and Apple Watch provided 38 and 37 features, respectively, with 22 of the total features obtained from the nonlinear methods described in this study. Additional features looked at comparing both devices such as correlations, difference, deviation and rate of change to include instances when one device had better physiological signals. For example, when the participant was moving a lot or left the bed, the BCG would have poor or no signal while the Apple Watch would continue to function as it was always worn during the experiment. The remaining 55 features were derived from comparisons between the BCG and Apple Watch, for a total of 130 features. The BCG device has its own HRV output, as outlined in Murata's manual and clinical trials, in addition to the raw beat-to-beat times used for most of the analysis.^{16,33} Table 1 showed improved agreement for the BCG and Apple Watch combination compared to PSG when sleep stages were classified into more diverse categories. This suggests that the combined system could function together to classify different sleep stages from the cardiac signals. Figure 4 shows the predictor importance of the top features, with the nonlinear ones considered in this study underlined. Many of the top nonlinear features come from the scatterplot methods like Poincare plots and

Table 1 The Agreement of Correctly Classifying Sleep Stages for Each Device Compared to PSG. Act is the Sleep/Wake Stages Compared to the GT9X Actigraphy Stages. Sleep/Wake is the Sleep/Wake Stages. Compared to the PSG Sleep Stages. Light/Full is the Sleep Stages Combined into Wake/Light/Deep. Light/Deep Combines Stage 1/2 and 3/REM into Light and Deep, Respectively REM/Non-REM are the Sleep Stages of Wake/Non-REM/REM Stages Compared to PSG

BCG + Apple Watch			BCG		Apple Watch	
Stages	Acc	Kappa	Acc	Kappa	Acc	Kappa
Act	81.9	0.402	81.1	0.433	80.2	0.35
Sleep/Wake	90.4	0.684	86.9	0.609	81.8	0.271
Light/Full	86.4	0.615	85.5	0.519	79.5	0.382
Light/Deep	73.4	0.583	52.4	0.268	61	0.383
REM/Non-REM	72.7	0.534	64.2	0.389	55.5	0.406

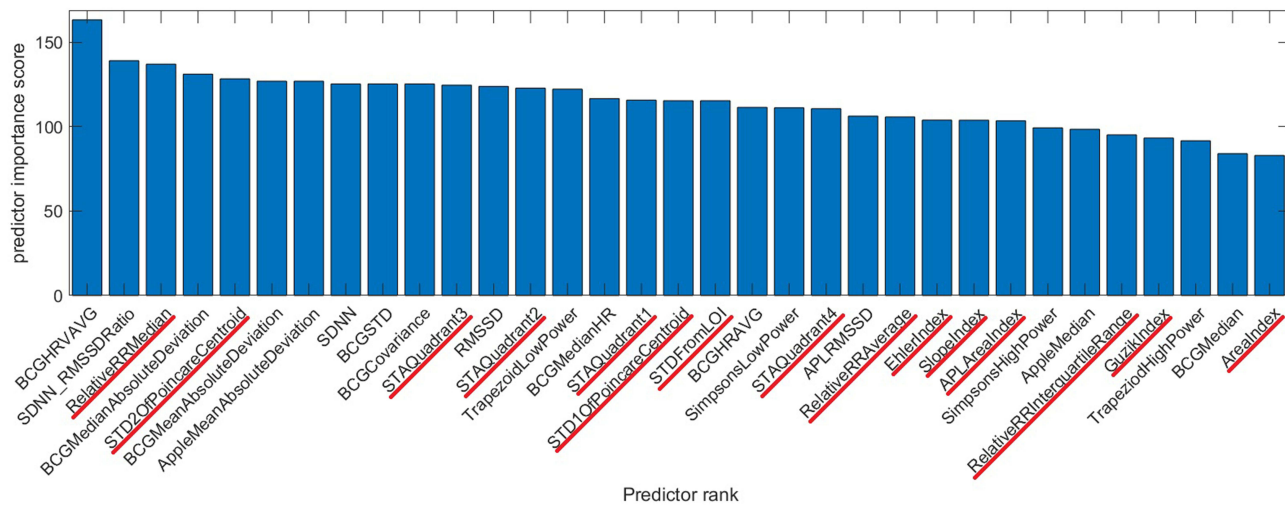


Figure 4 Zoomed in portion of top features for chi-square feature selection of most important features. Underlines are for non-linear features.

sequential trend analysis that suggests these methods can effectively distinguish ANS activity and ultimately be used for sleep stage classifications. The sleep stages were rescored to match outputs from GT9X actigraph as it only evaluates sleep or wakes and compare the Results from our previous study evaluating sleep stages with a similar setup.³⁰ The “Act” (denoting Actigraphy) and “Sleep/Wake” stages in [Table 1](#) were rescored to classify all sleep stages as “sleep”, resulting in a binary output of sleep or wake instances. “Sleep/Wake” compares instances to epochs from the PSG, while the “Act” compares them to the GT9X sleep scores. Actigraphy, which relies solely on motion activity to estimate sleep stages, is limited to evaluating sleep or wake states, while the specific sleep stages (ie, “Non-REM” and REM stages) require more sensors for accurate classification.¹³ [Table 2](#) shows statistical metrics comparing device configurations against sleep stages for PSG and the actigraphy. The sensitivity indicates correctly classified sleep stages while the specificity identifies correctly classified wake stages. The sensitivity in [Table 2](#) for all device configurations were higher than the specificity which indicates that more sleep stages were correctly classified than wake stages. However, the BCG and Apple Watch combination for “Sleep/Wake” stages showed higher sensitivity and specificity indicating that more sleep and wake stages were correctly classified compared to each device separately. Suggesting that the combination of the BCG and Apple Watch could classify sleep and wake stages correctly without overestimating sleep due to the participants lying still in a relaxed state. The classification into “Light/Full” rescoring sleep stages 2, 3, and REM as “full sleep”, while maintaining “wake” and stage 1 as “wake” and “light sleep” stages, respectively. “Light/Deep” rescoring the sleep stages 1 and 2 as “light sleep” and stages 3 and REM as “deep sleep”. “Non-REM/REM” combines all the non-REM sleep stages (stages

Table 2 Statistical Metrics for Sleep/Wake Stages from Each Device Configuration Compared Against PSG and Actigraphy Sleep Stages

Device	Method	Sensitivity	Specificity	Precision	F1
BCG + Apple Watch	ACT	0.8449	0.6604	0.9386	0.8892
	Sleep/Wake	0.8925	0.9717	0.9949	0.9409
BCG	ACT	0.8637	0.5946	0.8976	0.8803
	Sleep/Wake	0.8982	0.75	0.9366	0.917
Apple Watch	ACT	0.8364	0.5963	0.9249	0.8784
	Sleep/Wake	0.8441	0.6606	0.9366	0.888

1.2, and 3) into a single stage, with REM and wake maintaining their own designations. The sleep stages were rescored into four different target outputs to assess the model’s performance in classifying similar sleep stages together over corresponding epochs during the sleep cycle, such as grouping lighter sleep stages (stages 1 and 2) and deeper sleep stages (stages 3 and REM) and tracking the transition from wakefulness to lighter and then to deeper sleep. Figures 5–7 display the classified sleep stages in the REM/Non-REM configuration for each device configuration against the PSG stages, along with a confusion matrix for each sleep stage agreement. Figure 7 shows that the combined system was able to classify REM/non-REM sleep stages better as the BCG alone identified more non-REM stages while the Apple Watch identified more REM stages shown in Figures 6 and 5, respectively. Therefore, this suggests that the strengths of each device with an emphasis on nonlinear methods could be utilized together to identify different sleep stages.

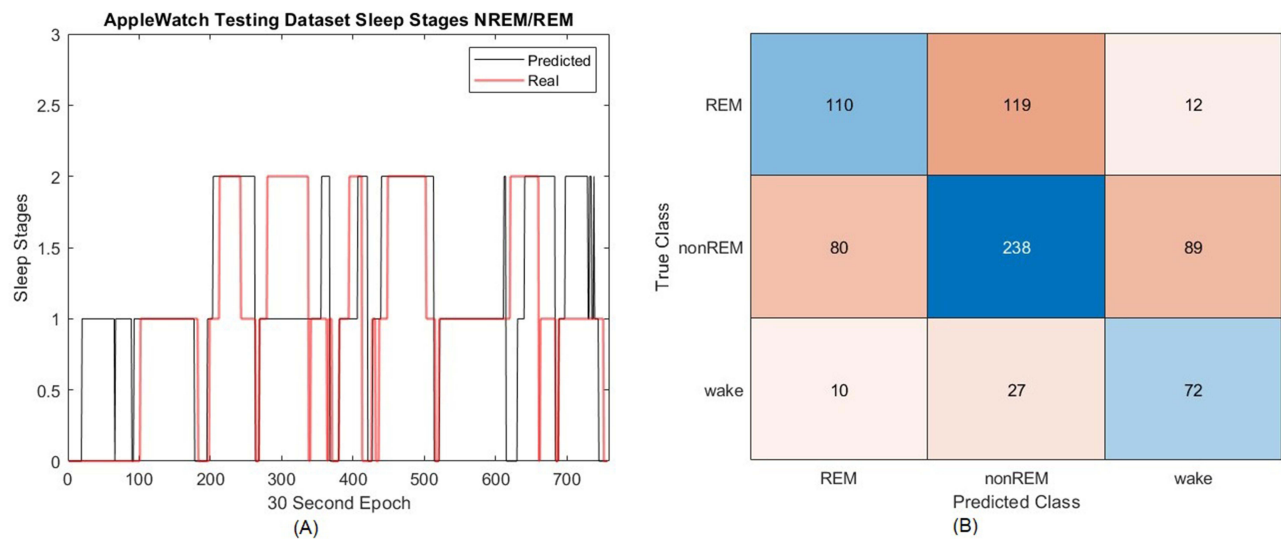


Figure 5 (A) Non-REM Sleep Stages for Apple Watch comparing to PSG. Epochs of 0, 1 and 2 are wake, non-REM, REM stages respectively. **(B)** Confusion matrix for the Apple Watch’s sleep classifications.

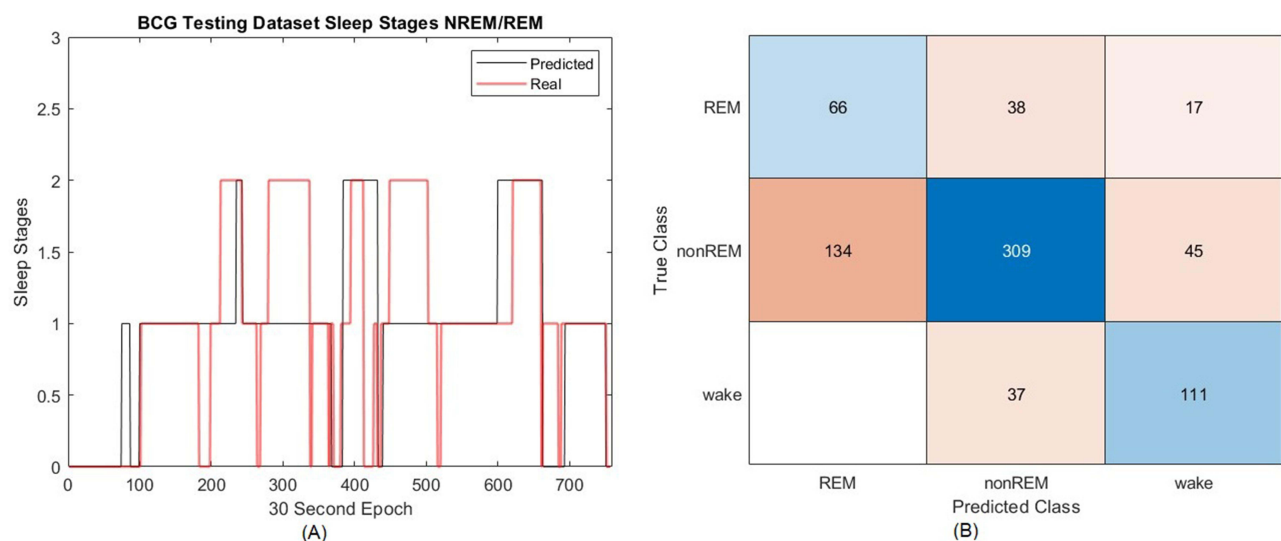


Figure 6 (A) Non-REM sleep stages for BCG comparing to PSG. Epochs of 0, 1 and 2 are wake, non-REM, REM stages respectively. **(B)** Confusion matrix for the BCG’s sleep classifications.

mean and standard deviation as a result of the inconsistent sampling rate.³¹ Table 1 shows that the Apple Watch had lower accuracy in most configurations compared to the BCG, potentially due to the interpolation used to address the inconsistent sampling rate, reducing the resolution of physiological signals and thus losing analytical detail. However, in the “Light/Deep” configuration, the Apple Watch shows higher accuracy of 61% than the BCG which had 52.4% as it was not affected by motion and distinguished between the three non-REM sleep stages more effectively. In contrast, the BCG was better at classifying the extreme stages of sleep in configurations focused on wake, light sleep, or REM. Nonetheless, the Apple Watch still maintained comparable accuracy to the BCG for most sleep stages, indicating the data’s potential utility for analysis. In the “Sleep/Wake” configuration, the BCG showed improved results in detecting sleep or wake instances compared to the Apple Watch, due to its higher sensitivity and specificity, as detailed in Table 2.

The combined use of BCG and Apple Watch showed a balanced sensitivity and specificity in Table 2, providing a comprehensive assessment of sleep and wake instances. The body’s effort to maintain stable functioning involves many organs operating in a nonlinear and nonstationary manner to adapt to external forces, with the ANS regulating cardiac activity by increasing SNS and decreasing PNS activities.^{27,39} Various indices have modeled this cardiac asymmetry to indicate the ANS dominance in heart control.³⁹ The asymmetry measured by these indices can be used to assess sleep stage levels, with plots skewed to one side indicating a deeper sleep. Similarly, the STA method measures SNS and PNS activities by placing consecutive J–J intervals in specific quadrants throughout the sleep cycle, allowing for the observation of sleep stages transitions.⁴⁴ Many influential features on the model, particularly nonlinear ones, are shown in Figure 4, enhance its performance. This study demonstrated an improvement from 50% to 86% in “Light/Full” sleep stage classification accuracy over our previous work using a BCG and an inertial measurement unit (IMU).³⁰ Furthermore, combining HRV signals from the BCG and Apple Watch yielded better results than using each device separately. Studies in the literature utilizing BCG and Apple Watch devices separately have obtained 76% precision on a 25 subject dataset⁵¹ and 72% accuracy on a 39 subject dataset,⁸ respectively, for classifying non-REM/REM sleep stages. While our study had a smaller dataset than those from literature, we were able to obtain a comparable accuracy of about 73% for non-REM/REM sleep stage classification. Many of these non-intrusive devices are relatively easy to obtain for the general population and include sophisticated sensors and algorithms to evaluate their health vitals. Incorporating more advanced non-linear techniques into these devices would allow clinics to diagnose patients in a more comfortable and natural environment and study them over longer periods of time. Additionally, the non-intrusive nature of the devices would be able to be used with almost any population group because the intrusive setup normally performed during clinical experiments would not be necessary with a small non-intrusive system. The general consumer would have a better representation of their health and activity with more algorithms that can measure the dynamic physiological activity of their body. Therefore, this study shows that nonlinear methods can be used to effectively classify sleep stages from the cardiac signals measured from BCG and Apple Watch combined.

In summary, the integration of BCG with the Apple Watch and the use of nonlinear analysis methods showed promising sleep stage classification results in this study. Future research should address this study’s limitations. A larger participant pool is necessary for more generalizable results, given the small sample size of six participants in this study. Participants were expected to complete three sleep cycles, but the average sleep duration was just under four hours, suggesting that longer sleep periods should be evaluated for a more thorough observation of sleep stage transitions. The Apple Watch’s inconsistent sampling rate necessitated interpolation to align with other devices, potentially leading to inaccurate signal readings. Apple’s APIs offer developers the opportunity to create apps that could enforce consistent data collection rates from the sensors and facilitate direct data export for analysis. Furthermore, refinement of the classification model by exploring alternative machine learning methods and diving deeper into the nonlinear methods may achieve more precise differentiation among the various sleep stages.

Conclusion

This study explored incorporating additional methods that could provide effective features to classify sleep stages using more unobtrusive devices. Specifically, methods involving scatter/return maps, such as those associated with Poincare plots, were examined for their relationship to the body’s ANS and their ability to capture the dynamics of the SNS and PNS imbalances. The combination of an Apple Watch and a bed-based BCG sensor with advanced nonlinear methods showed improved

accuracy in sleep stage scores compared to PSG. Focusing more on methods that evaluate ANS dynamics could provide more insight into modeling sleep stages with unobtrusive devices through machine learning models.

PSG, due to its inconvenience, high cost, and the necessity for medical professionals to manage setup and scoring, is not an effective method for the continuous sleep evaluation of the general population. The inter-rater reliability among manual sleep scorers, indicated by a kappa value of 0.76, demonstrates good reliability. However, discrepancies in some sleep stages among scores suggest a potential for misclassification.⁵² Therefore, the general population, interested in sleep tracking, and clinicians, aiming to observe patients' health trends over extended periods, would benefit from more accessible monitoring options. State-of-the-art consumer wearables, equipped with sophisticated sensors and algorithms, accurately measure physiological signals. Through network connectivity, these devices can be integrated with other systems for comprehensive health monitoring without relying on intrusive equipment. A home environment with this kind of system to monitor health aspects related to sleep quality and regularly provide health information to the user or medical professional. Patients with sleep health issues can be remotely monitored in their own home comfortably, diagnosing any potential sleep-related illnesses without requiring visiting a dedicated sleep clinic, which may not reflect their natural sleep. However, widespread adoption may take some time due to the initial investment in devices and networking capability required to make these devices available everywhere. As these devices connect to other devices like smart phones and online databases to provide more vitals monitoring capabilities, users would need to commit to specific system environments for adequate monitoring. Future advancements in sensor and device accessibility would alleviate barriers preventing users from incorporating a system to monitor their sleep quality and health regularly. The convenience of a non-intrusive system can regularly monitor a user's health without interrupting their sleep or activities in daily life.

Acknowledgments

This work was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) and Mitacs. Special thanks go to the team members of WearTech Labs for their guidance and feedback.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Sletten TL, Weaver MD, Foster RG, et al. The importance of sleep regularity: a consensus statement of the national sleep foundation sleep timing and variability panel. *Sleep Health*. 2023;9(6):801–820. doi:10.1016/j.sleh.2023.07.016
2. Cappuccio FP, Miller MA. Sleep and cardio – metabolic disease. In: Francesco P, editor. *Sleep, Health, and Society: From Aetiology to Public Health*. 2nd ed. United Kingdom: Oxford Academics; 2018:67–79.
3. Rocco C, Streng A, Van Kerkhof LWM, Van Der Horst GTJ, Chaves I. Social jetlag and related risks for human health: a timely review. *Nutrients*. 2021;13(12):4543. doi:10.3390/nu13124543
4. Wittman M, Dinich J, Mellow M, Roenneberg T. Social jetlag: misalignment of biological and social time. *Chronobiol. Int*. 2006;23(1–2):497–509. doi:10.1080/07420520500545979
5. Singh J, Sharma RK. Making sleep study instrumentation more unobtrusive. *IEEE Instrum Meas Mag*. 2021;21(1):50–53. doi:10.1109/MIM.2018.8278812
6. Chinoy ED, Cueller JA, Huwa KE, et al. Performance of seven consumer sleep-tracking devices compared with polysomnography. *Sleep*. 2021;44(5):1–16. doi:10.1093/sleep/zsaa291
7. Lee J, Byun W, Kiell A, Dinkel D, Seo Y. Comparison of wearable trackers' ability to estimate sleep. *Int J Environ Res Public Health*. 2018;15(6):1265. doi:10.3390/ijerph15061265
8. Welch O, Huang Y, Forger D, Goldstein C. Sleep stage prediction with raw acceleration and photoplethysmography heart rate data derived from a consumer wearable device. *Sleep*. 2019;42(12):1–19.
9. Zambotti M, Baker FC, Corlain IM. Validation of sleep-tracking technology compared with polysomnography in adolescents. *Sleep*. 2015;38(9):1461–1468. doi:10.5665/sleep.4990
10. Apple. *Apple Watch User GuideL Everything You Need to Know About Apple Watch*. WatchOS9.2, Apple Inc; 2022.
11. Lujan MR, Perez-Pozuelo I, Grandner MA. Past, present, and future of multisensory wearable technology to monitor sleep and circadian rhythms. *Front Digit Health*. 2021;3:721919. doi:10.3389/fgth.2021.721919
12. Apple. *Empowering People to Live a Healthier Day*. Apple Inc; 2022.
13. Nam Y, Kim Y, Lee J. Sleep monitoring based on a tri-axial accelerometer and a pressure sensor. *Sensors*. 2016;16(5):750. doi:10.3390/s16050750
14. Uguz DU, Tufan TB, Uzun A, Leonhardt S, Antink CH. Physiological motion artifacts in capacitive ECG: ballistocardiographic impedance distortions. *IEEE Trans Instrum Meas*. 2020;69(6):3297–3307. doi:10.1109/TIM.2020.2971336
15. Mora N, Cocconcelli F, Matrella G, Ciampolini P. Accurate heartbeat detection on ballistocardiogram accelerometric traces. *IEEE Trans Instrum Meas*. 2020;69(11):9000–9009. doi:10.1109/TIM.2020.2998644
16. Nurmi S. *Nocturnal Sleep Quality and Quantity Analysis with Ballistocardiography* [dissertation]. Acra: Aalto University; 2016.

17. Pino EJ, Chavez JAP, Aqueveque P. BCG algorithm for unobtrusive heart rate monitoring. *IEEE Healthcare Innovations and Point of Care Technologies*. Bethesda, MD, USA: IEEE; 2017.
18. Shin JH, Hwang SH, Chang MH, Park KS. Heart rate variability analysis using a ballistocardiogram during valsalva manoeuvre and post exercise. *Physiol Meas*. 2011;32(8):1239–1264. doi:10.1088/0967-3334/32/8/015
19. Tobaldini E, Nobili L, Strada S, Casali KR, Braghiroli A, Montano N. Heart rate variability in normal and pathological sleep. *Front Physiol*. 2013;4(294):1–11. doi:10.3389/fphys.2013.00294
20. Sadek I, Biswas J, Abdulrazzak B. Ballistocardiogram signal processing: a review. *Health Inf Sci Syst*. 2019;7(1):10. doi:10.1007/s13755-019-0071-7
21. Task Force of the European Society of Cardiology. The North American society of pacing electrophysiology. heart rate variability. *Circulation*. 1996;93(5):1045–1065.
22. Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. *Front Public Health*. 2017;5(258):1–17. doi:10.3389/fpubh.2017.00258
23. Goldberger AL, Amaral LAN, Hausdorff JM, Ivanov PC, Peng CK, Stanley HE. Fractal dynamics in physiology: alterations with disease and aging. *Proc Natl Acad Sci USA*. 2002;99(1):2466–2472. doi:10.1073/pnas.012579499
24. Li K, Rudiger H, Ziemssen T. Spectral analysis of heart rate variability: time window matters. *Front Neurol*. 2019;10:545. doi:10.3389/fneur.2019.00545
25. Feng J, Huang W, Jiang J, et al. Non-invasive monitoring of cardiac function through ballistocardiogram: an algorithm integrating short-time Fourier transform and ensemble empirical mode decomposition. *Front Physiol*. 2023;14:1201722. doi:10.3389/fphys.2023.1201722
26. Sadek I, Biswas J. Noninvasive heart rate measurement using ballistocardiogram signals: a comparative study. *Signal Image Video Process*. 2018;13:475–482. doi:10.1007/s11760-018-1372-z
27. Jelinek HF, Cornforth DJ, Tarvainen MP, Khalaf K. Investigation of linear and nonlinear properties of a heartbeat time series using multiscale renyi entropy. *Entropy*. 2019;21(8):727. doi:10.3390/e21080727
28. Almutairi H, Hassan GM, Datta A. Machine-learning-based-approaches for sleep stage classification utilising a combination of physiological signals: a systematic review. *Appl Sci*. 2023;13(24):13280. doi:10.3390/app132413280
29. Lin CT, Prasad M, Chung CH, et al. IoT-based wireless polysomnography intelligent system for sleep monitoring. *IEEE Access*. 2018;6:405–414. doi:10.1109/ACCESS.2017.2765702
30. Jaworski DJ, Park A, Park EJ. Internet of things for sleep monitoring. *IEEE Instrum Meas Mag*. 2021;24(2):30–36. doi:10.1109/MIM.2021.9400950
31. Jaworski DJ, Park EJ. Apple watch sleep and physiological tracking compared to clinically validated actigraphy, ballistocardiography and polysomnography. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. Sydney Australia*; 2023.
32. Herbst E, Metzler TJ, Lenoci M, et al. Adaption effects to sleep studies in participants with and without chronic posttraumatic stress disorder. *Psychophysiology*. 2010;47(6):1127–1133. doi:10.1111/j.1469-8986.2010.01030.x
33. Meriheina U. BCG Measurements in BEDS White Paper. Murata Electronics Oy: Nagaokakyo, Japan; 2019.
34. Murata Electronics LTD. Intelligent Calibration Application Note. Murata Electronics Oy: Nagaokakyo, Japan; 2017.
35. Berry RB, Quan SF, Abreu AR, et al. The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications. *Am Acad Sleep Med*. 2020;2020:1.
36. Cole RJ, Kripke DF, Gruen W, Mullaney DJ, Gillin JC. Automatic sleep/wake identification from wrist activity. *Sleep*. 1992;15(5):461–469. doi:10.1093/sleep/15.5.461
37. Murata Electronics LTD. SCA11H Operation Modes. Murata Electronics Oy: Nagaokakyo, Japan; 2015.
38. Young DW. Self-Measure of Heart Rate Variability (HRV) and arrhythmia to monitor and to manage atrial arrhythmias: personal experience with high intensity interval exercise (HIIE) for the conversion to sinus rhythm. *Front Physiol*. 2014;5(251):1–4. doi:10.3389/fphys.2014.00251
39. Khandoker AH, Karmakar C, Brennan M, Voss A, Palaniswami M. Heart Rate Asymmetry Analysis Using Poincare Plot. In: *Poincare Plot Methods for Heart Rate Variability Analysis*. Boston MA: Springer; 2013:69–91.
40. Vollmer M. A robust, simple and reliable measure of heart rate variability using relative RR intervals. *Computing in Cardiology Conference*. Nice, France; 2015.
41. Behbahani S, Dabanloo NJ, Nasrabadi AM. Ictal heart rate variability assessment with focus on secondary generalized and complex partial epileptic seizures. *Adv Biores*. 2013;4(1):50–58.
42. Pawloski R, Buszko K, Newton JL, Kujawski S, Zalewski P. Heart rate asymmetry analysis during head-up tilt test in healthy men. *Front Physiol*. 2021;2021:12.
43. Yan C, Li P, Ji L, et al. Area asymmetry of heart rate variability signal. *Biomed Eng Online*. 2017;16(112):1–14. doi:10.1186/s12938-017-0402-3
44. de Carvalho JLA, da Rocha AF, de Oliveira Nascimento FA, Neto JS, Junqueira LF. Development of a matlab software for analysis of heart rate variability. 6th Int. Conf. Signal Proc. Beijing, China; 2003.
45. Srinivas K, Reddy LRG, Srinivas R. Estimation of heart rate variability from peripheral pulse wave using PPG sensor. *Int. Conf. Biomed. Eng. Kuala Lumpur, Malaysia*; 2007.
46. Srinivas K, Reddy LRG. Detecting congestive heart failure using heart rate sequential trend analysis plot. *Int Jour Eng Science Tech*. 2010;2(12):7329–7334.
47. Renyi A. On measures of entropy and information. *Berkeley Symp on Math Statist and Prob*. 1961;4(1):547–561.
48. Askenazy Y, Ivanov PC, Havlin S, et al. Magnitude and sign correlations in heartbeat fluctuations. *Phys Rev Lett*. 2001;86(9):1900. doi:10.1103/PhysRevLett.86.1900
49. Muller M. Dynamic Time Warping. In: *Information Retrieval for Music and Motion*. Heidelberg, Berlin: Springer; 2007:69.
50. Rahda M, Fonseca R, Moreau A, et al. Sleep stage classification from heart-rate variability using long short-term memory neural networks. *Sci Rep*. 2019;9:14149. doi:10.1038/s41598-019-49703-y
51. Gasmi A, Augusto V, Beaudet P, et al. Sleep stage classification using cardio-respiratory variables. *IEEE Conf. Autom. Sci. Eng (CASE)*. Hong Kong, China; 2020.
52. Lee YJ, Lee JY, Cho JH, Choi JH. Interrater reliability of sleep stage scoring: a meta-analysis. *J Clin Sleep Med*. 2022;18(1):193–202. doi:10.5664/jcsm.9538

Nature and Science of Sleep

Dovepress

Publish your work in this journal

Nature and Science of Sleep is an international, peer-reviewed, open access journal covering all aspects of sleep science and sleep medicine, including the neurophysiology and functions of sleep, the genetics of sleep, sleep and society, biological rhythms, dreaming, sleep disorders and therapy, and strategies to optimize healthy sleep. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/nature-and-science-of-sleep-journal>