



OPEN

Estimation of systolic blood pressure by Random Forest using heart sounds and a ballistocardiogram

Rafael Gonzalez-Landaeta, Brenda Ramirez & Jose Mejia

Cuffless blood pressure measurement enables unobtrusive and continuous monitoring that can be integrated with wearable devices to extend healthcare to non-hospital settings. Most of the current research has focused on the estimation of blood pressure based on pulse transit time or pulse arrival time using ECG or peripheral cardiac pulse signals as proximal time references. This study proposed the use of a phonocardiogram (PCG) and ballistocardiogram (BCG), two signals detected noninvasively, to estimate systolic blood pressure (SBP). For this, the PCG and the BCG were simultaneously measured in 21 volunteers in the rest, activity, and post-activity conditions. Different features were considered based on the relationships between these signals. The intervals between S1 and S2 of the PCG and the I, J, and K waves of the BCG were statistically analyzed. The IJ and JK slopes were also estimated as additional features to train the machine-learning algorithm. The intervals S1-J, S1-K, S1-I, J-S2, and I-S2 were negatively correlated with changes in SBP (p -val < 0.01). The features were used as explanatory variables for a regressor based on the Random Forest. It was possible to estimate the systolic blood pressure with a mean error of 3.3 mmHg with a standard deviation of ± 5 mmHg. Therefore, we foresee that this proposal has potential applications for wearable devices that use low-cost embedded systems.

Arterial blood pressure (ABP) is a reliable indicator of health conditions. Current gold standard noninvasive methods rely on the use of inflatable cuff-based systems, which may be uncomfortable for some subjects. To tackle this, different cuff-less approaches have been proposed, mainly based on the estimation of pulse transit time (PTT), pulse arrival time (PAT), and pulse wave velocity (PWV). This implies multimodal measurement approaches¹, the most common being those where the ECG is used as a proximal timing reference signal versus the peripheral pulse. In these cases, the time between the R-wave and different points of the pulse signal (peak, foot, and slope) was measured to determine the correlation between ABP and PTT^{2,3}. This method is termed PAT and includes the pre-ejection period (PEP) ($PAT = PTT + PEP$)⁴, which depends on the isovolumetric contraction in the left ventricle⁵ and is influenced by sympathetic activity⁶. PEP changes independently of ABP^{2,7}, and it is not easy to estimate it.

Other biosignals have been considered proximal time references; the most common are photoplethysmography (PPG)^{8,9}, phonocardiography (PCG)¹⁰, ballistocardiography (BCG)^{7,11}, and seismocardiography (SCG)¹². With these signals, the PEP is excluded, and the PTT is estimated by measuring the time between some waves of these signals and the peak, foot, or slope of the peripheral pulse signal.

The PCG is a good indicator of the onset of mechanical systole^{13,14}, and the analysis of its main waves provides information about blood pressure (systolic, diastolic, and mean)¹⁵. Regarding the use of PCG as a proximal timing reference signal, the PTT is usually estimated by using the PPG as a distal timing signal, and it has been demonstrated that the correlation with ABP is comparable when using the ECG as reference¹⁰. However, the morphology of the PPG signal can be modified by the effects of aging¹⁶ and by the peripheral arterial stiffness¹⁷, bringing about misleading results when using PTT as an ABP indicator.

A different approach is proposed in this paper, where it is not necessary to detect the peripheral pulse. The change in systolic ABP is estimated by detecting the time delay between the two main sounds (S1, S2) of the PCG and the main waves (I, J, K) of the BCG. This idea stems from the fact that changes in blood pressure have little effect on the onset of S1 and S2. However, they do have an important influence on the genesis of the main

Department of Electrical and Computer Engineering, Autonomous University of Ciudad Juárez, 32310 Ciudad Juárez, Mexico. ✉email: jose.mejia@uacj.mx

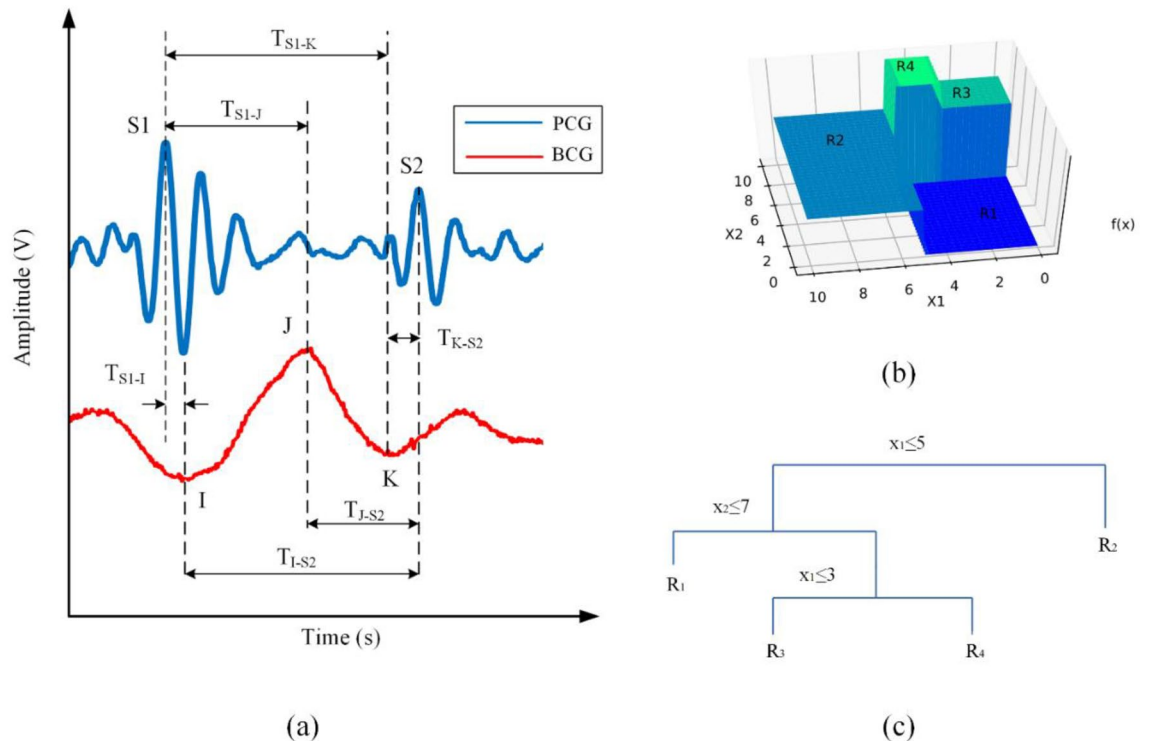


Figure 1. (a) Time intervals between the PCG and the BCG. (b) Example of a tree model, where the tree input space is a box of \mathbb{R}^2 and a partition of the input space, and (c) the associated decision tree.

waves of the BCG. In the case of PCG, S1 is due to the simultaneous closure of the atrioventricular valves, so its correlation with blood pressure is not as evident compared with that of S2, whose spectrum has been shown to have a significant correlation with systolic blood pressure (SBP)^{15,18}. In the case of BCG, pressure gradients in the ascending and descending aorta define both the amplitudes and the onsets of I, J, and K waves¹⁹. In this sense, we believe that a change in systolic pressure causes a variation in the time delay between the main waves of the PCG and the BCG. To validate this, both signals are detected simultaneously, and the time relationship between the main waves of PCG and BCG is analyzed. To conduct the measurements in a simple way, in this work, the PCG is measured at the sternum, while the BCG is measured using a weighing scale. However, the intention is that these signals can be detected using wearable systems capable of detecting BCG at the sternum, using methods such as that proposed by Wiens et al.²⁰, so both signals (PCG and BCG) could be measured at the same point. From the PCG, both S1 and S2 are used as proximal time references, and the time delay with respect to each I, J, and K wave of the BCG is estimated; see Fig. 1a. In addition, these time intervals are used as explanatory variables in a Random Forest regressor (RF) to estimate systolic blood pressure. The RF model provides more information because it evaluates the importance of each explanatory variable over the response variable (i.e., systolic blood pressure). To the best of our knowledge, these two signals have not been used together to assess the correlation with the ABP, so we consider that the results derived from this research may be relevant for future work related to the estimation of ABP using cuff-less methods.

This manuscript is organized as follows: Sect. “Random forest theory” provides a brief introduction to the RF model. Section “Methods” describes the setup used for signal acquisition, signal processing, and measurement protocol. Section “Results and discussion” shows the experimental results and discussion, and Sect. “Conclusion” draws the main conclusions.

Random forest theory

Regression trees. Classification and regression tree models work by recursively partitioning their input space into M decision regions²¹. They are named such in part because such a partition is generally represented by a tree, with its leaf denoting different regions. Let $f(x)$ be the output of the tree to an input vector x . Then the model can be written as:

$$f(x) = \sum_{m=1}^M w_m I(x \in R_m) \quad (1)$$

where $I(\cdot)$ is an indicator function, R_m is the m th region, and w_m is its associated weight. Figure 1b shows an example of an input space partitioned into four regions by the tree of Fig. 1c. Although tree models are well suited to large data sets and they handle outliers relatively well, they suffer from a lack of accuracy compared to other models. In addition, trees are known to be unstable due to small changes to the input data, and they are considered high variance estimators.

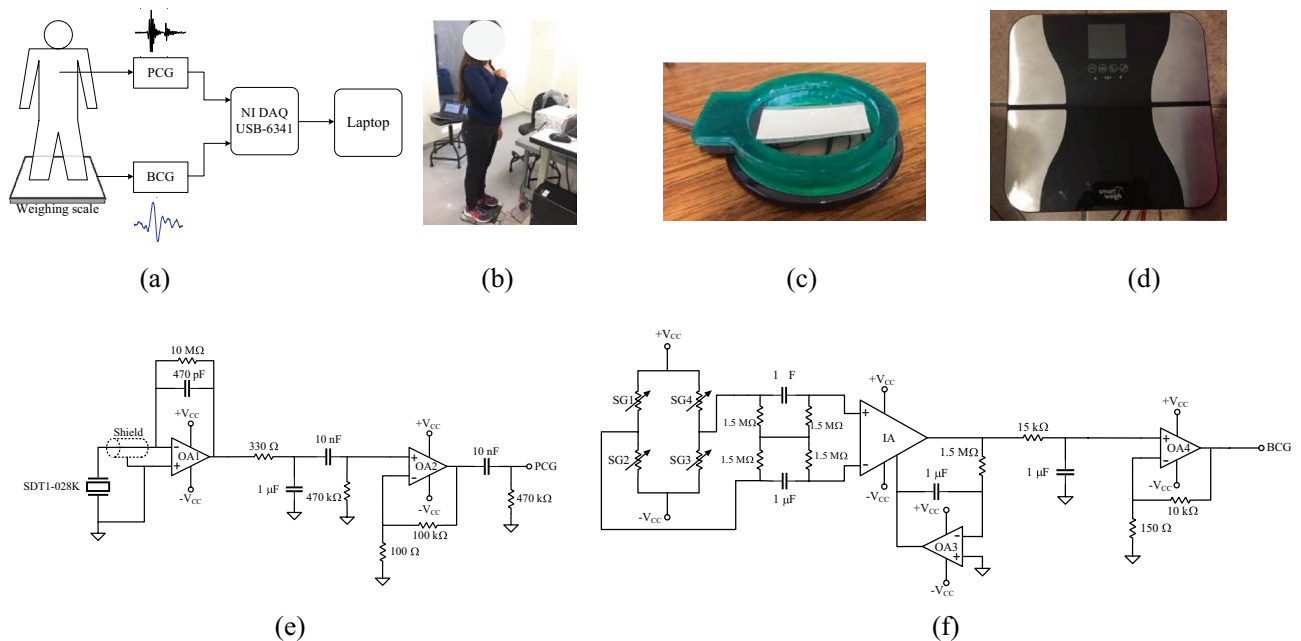


Figure 2. (a) Setup for detecting the PCG and the BCG simultaneously. (b) One of the volunteers using the measurement setup. (c) SDT1-028 K mounted on a custom-built case used for detecting the PCG. (d) Electronic weighing scale used for detecting the BCG. (e) Circuit used to detect the PCG, and (f) circuit used to detect the BCG. Figures (a, e, f) were drawn in Microsoft Visio Professional 2019²⁴ by Rafael Gonzalez-Landaeta.

Random forest. Random Forests are procedures that use an ensemble of decision trees and are recognized for their high prediction performance when dealing with real problems²². In these models, each tree was constructed using a randomly chosen subset of the training set through a random vector Θ . An RF model tries to reduce the variance of the regression tree model by averaging many tree estimates. Considering that the model uses T decision trees, $\{f_1(x, \Theta_1), f_2(x, \Theta_2), \dots, f_T(x, \Theta_T)\}$, where now each tree has an additional parameter Θ_t , the RF could be formally written as

$$f_{RF}(x) = \frac{1}{T} \sum_{t=1}^T \alpha_t f_t(x, \Theta_t) \quad (2)$$

where α_t is an associated weight. For the regression case, this weight is generally chosen as one.

Methods

Signal acquisition. In this work, PCG and BCG were detected simultaneously (Fig. 2a,b). The PCG was detected at the pulmonary point located in the second intercostal space to the left of the sternum. The reason was that in preliminary tests, the PCG was obtained with greater amplitude at this point. The BCG was obtained by detecting cranial-caudal forces using an electronic weighing scale²³.

To detect PCG, the approach proposed by Vazquez et al. was used²⁵. A shielded piezofilm was used as a microphone (Fig. 2c) to obtain a high signal-to-noise ratio (SNR) PCG. For this, the SDT1-028K (TE Connectivity) was used and placed in a custom-built case to simplify the positioning of the sensor in the auscultation point. The conditioning circuit of the piezofilm consisted of a charge amplifier followed by a first-order passive band-pass filter and a non-inverting amplifier (Fig. 2e). The sensitivity of the entire system was 2.12 V/pC in a range of frequencies between 34 and 482 Hz. OA1 and OA2 were implemented using LT1793 (Linear Technology).

The BCG was detected using a weighing scale from Smart Weigh (Fig. 2d). The strain gauges (SG) mounted on the load cells of the scale formed a full Wheatstone bridge whose output was connected to the circuit shown in Fig. 2f. The total gain of the circuit was adjusted to 61×10^3 , with a bandwidth limited to between 0.5 and 10 Hz. The instrumentation amplifier (IA) used was the INA114 (Texas Instruments), and OA3 and OA4 were implemented using the TL082 (Texas Instruments). All the circuits were powered at ± 10 V with a Power Supply E3631A (Keysight). The PCG and BCG signals were registered using the data acquisition system USB-6341 (National Instrument) connected to a laptop and configured with a sampling frequency of 1 kHz.

Signal processing. PCG is an acoustic signal that may be corrupted by noise from different sources, such as other sounds coming from the patient, skin contact with the stethoscope, and ambient noise^{26–28}. Therefore, it was necessary to filter the PCG signal. For this study, it was decided to preserve components in the band above 34 Hz, since much of the low-frequency noise has main frequency components from 0 to 25 Hz²⁹. In addition, because this study will not attempt to classify or detect sounds that indicate any abnormality, it was decided to eliminate high-frequency noises, unlike other studies in which pathologies are detected³⁰. Therefore, the upper

bound of the passband was established at 50 Hz, which facilitated the detection of sounds S1 and S2. Thus, the signal was filtered using a bandpass FIR filter designed with a Blackman window and cutoff frequencies of 34 Hz and 50 Hz. This also avoids other relatively high-frequency interferences, such as those produced by the AC line. In the case of BCG, this signal was filtered using a bandpass filter from 1 to 10 Hz, since it was observed that most of the signal power was in that band, which agrees with other studies³¹.

Measurement of the features. The features considered in this study were the time intervals between S1-J, J-S2, S1-I, I-S2, K-S1, K-S2, the slope of the I-J segment, and the slope of the J-K segment. For the extraction of the features of interest, an algorithm similar to³² was used. Initially, all signals were normalized. Next, for the signals of each subject, the detection of the I, J, and K peaks was performed by comparing neighboring points to determine the local maxima and minima in the signal. Subsequently, using point J, which was previously detected, we proceeded to search for S1 in the PCG, looking for a local maximum before the temporary location of point J. Accordingly, point S2 was sought in the temporary neighborhood of point K. Finally, the peaks of waves I, J, and K of the next cycle were searched for at a distance of 4 ms from the previous J wave peak. This process was repeated until the end of the signal was reached. Then the time intervals between S1-J, J-S2, S1-I, I-S2, K-S1, K-S2, the slope of the I-J segment, and the slope of the J-K segment were calculated for each cardiac cycle. Finally, for each subject, a global estimate of each feature was obtained using the median of the measurements made for each cardiac cycle. The median instead of the sample mean was used since this is a more robust estimator of the mean of the real distribution, especially when measures are noisy and more prone to outliers²⁷.

Data analysis. *Behavior of features regarding pressure changes induced by activities.* With the measurements made on each individual, various statistical analyses were performed. The behavior of the features in response to changes in SBP was statistically analyzed using the empirical distributions of the features when the subjects performed different activities. Sample statistics of the mean and variance were calculated, as well as approximations to a probability distribution using kernel density estimation²⁸. Moreover, to test the normality of the data, the Shapiro–Wilk test was used³³.

Relationships of the features by distinctions in time and pressure. We further explored the relationships between these features. Based on the states of rest, activity, and post-activity, variations in pressure and the timing of characteristics were evaluated for the three states. The linearity of the relationship between the characteristics and the pressure was analyzed using the least squares regression technique with a Huber regularizer²⁷, so the estimation was less influenced by outliers. To quantify the degree of association between the quantities involved, the Pearson correlation coefficient was calculated, which shows how well the data fit a linear relationship, and Spearman's correlation coefficient, which demonstrates how well the data agree with the monotone classification with relevant outliers³⁴. The p-value was calculated based on the probability of the null hypothesis: that the current result would have been found if the correlation coefficients were zero. Therefore, it is assumed that if this probability is less than 5% (p-val < 0.05), the correlation coefficient will be statistically significant.

BP estimation using machine learning algorithms. Once the features and their relationship with the SBP were analyzed, several machine learning algorithms were compared to estimate the SBP from the characteristics analyzed. It was determined which characteristics were the most predominant in the estimation process.

Dataset. A data set, $D_1 = \{(x_i, y_i)\}_{i=0}^{N_1}$, with $N_1 = 1,067$, was used to train the machine learning algorithms, which consisted of records $x_i \in \mathbb{R}^{10}$ with features of the PCG and BCG signals. In addition, $y_i \in \mathbb{R}$ contains systolic pressure. To obtain the features, the measured signals of all subjects were segmented into cardiac cycles. A record x_i then consisted of the following features obtained in one cycle: the intervals S1-J, J-S2, S1-I, I-S2, K-S1, K-S2, I-J, J-K, the slope from points I to J, and the slope from points J to K. Once the features were obtained, the corresponding systolic pressure was recorded in y_i . Notably, not all segmented cycles were used; those cycles where it was not possible to obtain all the features due to distortions or noise were eliminated. In addition, those records with S1-J intervals that were far from the mean by more than two standard deviations were discarded. Additionally, a second data set was contemplated, $D_2 = \{(x_i, y_i)\}_{i=0}^{N_2}$, with the same characteristics of D_1 , except those records with measures of S1-J, S1-K, and I-J with a deviation of the mean by more than two standard deviations. This reduced the number of records to $N_2 = 235$. The percentage of acceptance of cycles varied from person to person and with the activity; on average the percentage of acceptance for D_1 was 78% and for D_2 it was 41%.

Estimation of SBP. To estimate SBP, the following algorithms were evaluated: simple linear regression (LR); support vector machine (SVM), using a radial basis function kernel and parameters $c=45$, $\text{gamma}=0.0001$, $\text{epsilon}=0.01$ for D_1 and $c=55$, $\text{gamma}=0.0001$, and $\text{epsilon}=0.01$ for D_2 ; kernel ridge (KR), with the parameters $\text{alpha}=0.005$, $\text{gamma}=0.05$, and a Laplacian kernel function for D_1 and D_2 ; and RF, with 139 trees for D_1 and 210 trees for D_2 . The algorithms were used in regression mode instead of classifier mode because the quantity to be estimated, the BP, is a non-categorical continuous variable. The parameters were optimized using the search grid method with fivefold cross validation on the training test, and all the algorithms were implemented with the sklearn library³⁵. The algorithms were trained using data from 17 subjects, and data from four subjects were used for testing. The subjects in each set were chosen by a random permutation. It was decided not to use Deep Learning (DL) algorithms for two reasons: the tendency in DL is to obtain features automatically, using, for example, convolutional networks, which in some contexts such as medicine could make their interpretation difficult^{36,37}, and second, these types of architectures generalize better when they are trained with a large amount of data³⁸.

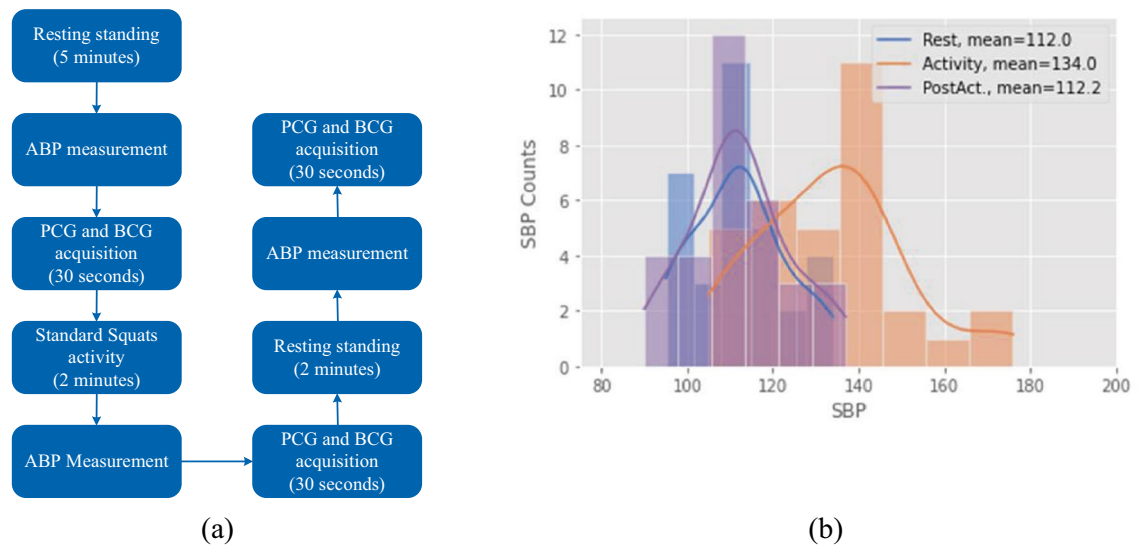


Figure 3. (a) Description of the measurement protocol. (b) Distribution of pressure measures related to each activity.

Moreover, the selected algorithms have a low computational cost, which could facilitate their implementation in an embedded system, as they do not require a high-end processor for their programming.

To evaluate the regressors, the following metrics were used: the Explained Variance Score (EVS), which calculates an estimate of the explained variance by the algorithm used, and is given by

$$EVS = 1 - \frac{Var(y - \hat{y})}{Var(y)} \quad (3)$$

where \hat{y} is the value estimated by the regressor and $Var(\cdot)$ is the variance of the argument. The coefficient of determination (R^2) estimates the proportion of the variation in the dependent variable that is predictable from the independent variable, and is given by

$$R^2 = \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (4)$$

where the variable i runs through all observations and \bar{y} is the mean of the observations. The mean absolute error (MAE) and mean square error (MSE) were also used to evaluate the regressors.

Measurement protocol. The PCG and the BCG of 21 healthy subjects (11 females and 10 males) were measured (mean \pm SD): age = (23 \pm 4) years, weight = (65 \pm 12) kg, and height (169 \pm 12) cm. Subjects with hypertension or other cardiac diseases were not considered in this study. The corresponding approval (CIEB-2019-1-106) from the institutional ethical committee of the Autonomous University of Ciudad Juarez was obtained, and a written informed consent was signed by the volunteers. All experiments were performed according to relevant guidelines and regulations. Figure 3a describes the measurement protocol for each subject. Two resting periods were considered: the first one of 5 min before the physical activity, and the second one of 2 min after the physical activity. The ABP of the volunteers and the biosignals were measured after each rest period. To induce variations in ABP, the volunteers performed standard squats for 2 min, after which the ABP and biosignals were measured. The ABP was measured with an automatic blood pressure monitor HEM-7200 (Omron), with an accuracy of ± 3 mmHg. The measurement protocol was repeated twice for each volunteer. Figure 3b reveals the distribution of the data collected with the measurement protocol. Data were summarized graphically using histograms of the samples and grouped by the subjects' states during the measurement: rest, activity, and post-activity. In addition, approximations to a probability distribution using the kernel density estimation method²⁸ are illustrated as continuous curves. As expected, the mean pressure during activity was higher than at rest and post-activity; in addition, the mean pressures at rest and post-activity were very similar.

Results and discussion

Figure 4 illustrates the histograms for each of the features obtained and as an estimate of its probability distribution (continuous line). For each characteristic, three histograms were obtained depending on the person's state, that is, at rest, physical activity, and post-activity. The first row of Fig. 4 shows that, in the intervals S1-J, S1-I, and S1-K, the means of their empirical distributions followed the same pattern; as physical activity increases, the mean of the intervals decreases. In addition, in the measurements of the intervals J-S2 and I-S2 (second row of Fig. 4), despite the distributions being noisier, their means followed the same pattern, except for the interval K-S2, where the mean of the activity data is greater than that of the post-activity, which may be due to the dispersion

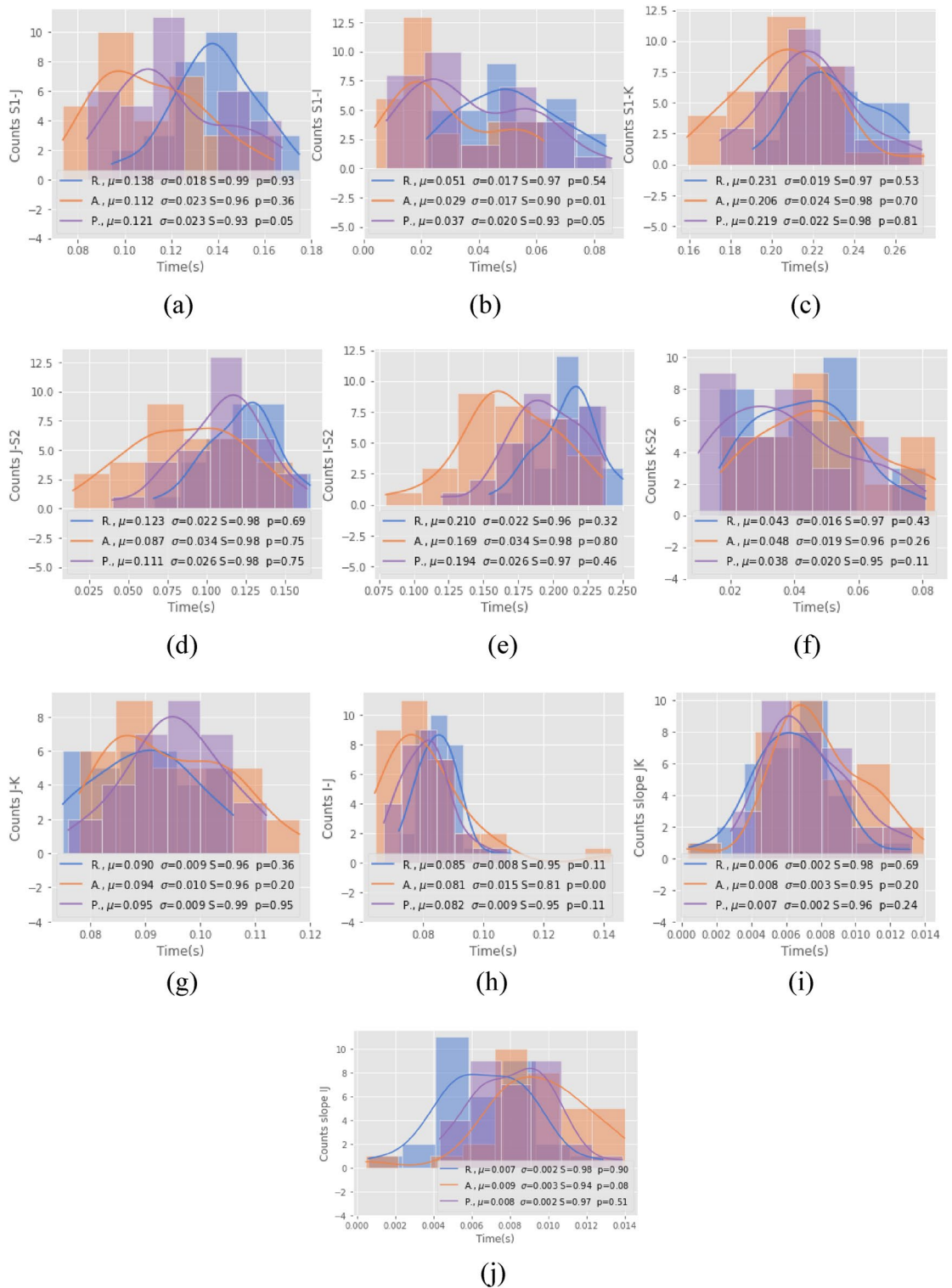


Figure 4. Histograms of the features according to the subject state: R=Rest, A=Activity, P=Post-activity, μ =Mean, σ =Std. dev., S=Shapiro test for normality, and p=p-value of the Shapiro test. Sample: 21 volunteers: (a) S1-J, (b) S1-I, (c) S1-K, (d) J-S2, (e) I-S2, (f) K-S2, (g) J-K, (h) I-J, (i) slope J-K, and (j) slope I-J.

in the data in both histograms, which revealed great variance with respect to their mean. In addition, it can be seen that the means of all the distributions are very close. Regarding the I-J and J-K intervals and the JK slope, the histograms were more defined, with less variance. In the I-J slope, their distributions presented a greater flat-tening at their peak. All the data had a normality statistic above 0.8, according to the Shapiro–Wilk test. However, for the case of the I-J interval, the physical activity data did not reach statistical significance.

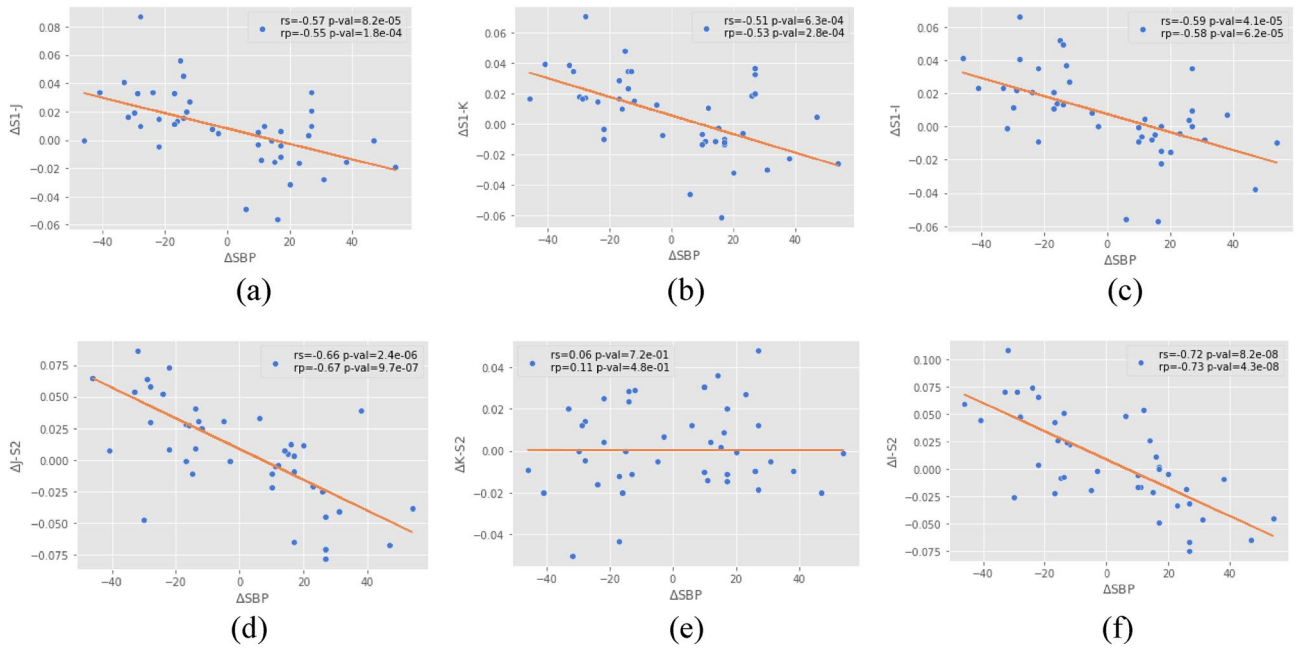


Figure 5. Scatter plots of time intervals against SBP changes in the features for different states. rs = Spearman correlation coefficient; rp = Pearson correlation coefficient.

	D_1 dataset				D_2 dataset			
	EVS	R ²	MAE	MSE	EVS	R ²	MAE	MSE
LR	0.302	0.040	8.072	125.89	0.719	0.332	5.646	46.06
KR	0.973	0.039	13.21	267.54	-0.190	-0.021	7.362	70.53
SVM	0.319	0.192	9.391	156.55	0.598	0.216	5.799	54.14
RF	0.318	0.213	7.090	103.26	0.641	0.486	4.402	35.47

Table 1. Evaluation of the regression algorithms with the metrics.

Figure 5 illustrates scatter plots of time intervals against SBP changes for all features and for different states. For each feature, correlation coefficients and the regression curve were calculated. For this, the Spearman and Pearson correlations were estimated, as well as their p-values. The results of this study showed that the intervals S1-J, S1-K, S1-I, I-S2, and J-S2 were negatively correlated with changes in SBP (p-val < 0.01). However, for the difference S2-K, a slightly positive correlation trend and a greater dispersion in the data existed. With a p-val > 0.01, its statistical significance was ruled out.

Once the proposed characteristics were analyzed, they were used as explanatory variables for training the regression algorithms. Table 1 portrays the evaluation of the regression algorithms on the test set. For the case of using all measures, D_1 , SVM was the least adapted; even though various kernels were tested, the best kernel *rbf* could not model the non-linearity of the characteristics. The regressor based on Random Forest achieved the best adaptation with the least outliers; this was observed in the lowest MSE for all methods. Regarding the use of data, D_2 , constrained by the standard deviation, however, the other regressors also obtained good performance, especially the linear regression model.

Figure 6 shows error graphs for the algorithms. For the case of using the dataset D_1 , the first column of Fig. 6, the best mean (2.9) is obtained by the KR. However, many measurements exceed the error by more than 20 units, which is reflected in a high MSE. The second-best method is the RF, which obtains a mean error of 3.7, where few measurements have more than 20 units of error compared to the other methods. The same occurs with LR, but to a lesser extent. As for SVR, it has the poorest performance of the exposed methods. Notably, for the clarity of graphs, only the first 100 samples are shown. Nonetheless, the mean and standard deviations were calculated using the entire D_1 dataset. Figure 7a demonstrates a boxplot of the statistics of each algorithm.

Regarding dataset D_2 , the RF obtains the smallest mean error of 3.3, followed by KR, though the latter presents more values greater than 20, so its MSE is large. Furthermore, in the RF method, several predicted values are zero, which does not occur in the other methods. Figure 7b depicts boxplots of the performance of the algorithms on the D_2 dataset.

Finally, using the best classifier (RF in D_2), the importance of the features was analyzed. In Fig. 7c, the bar height shows the importance in the Random Forest, while the lines show the variability between trees represented by the error bars. The importance is based on the mean decrease in impurity (MDI). The most important

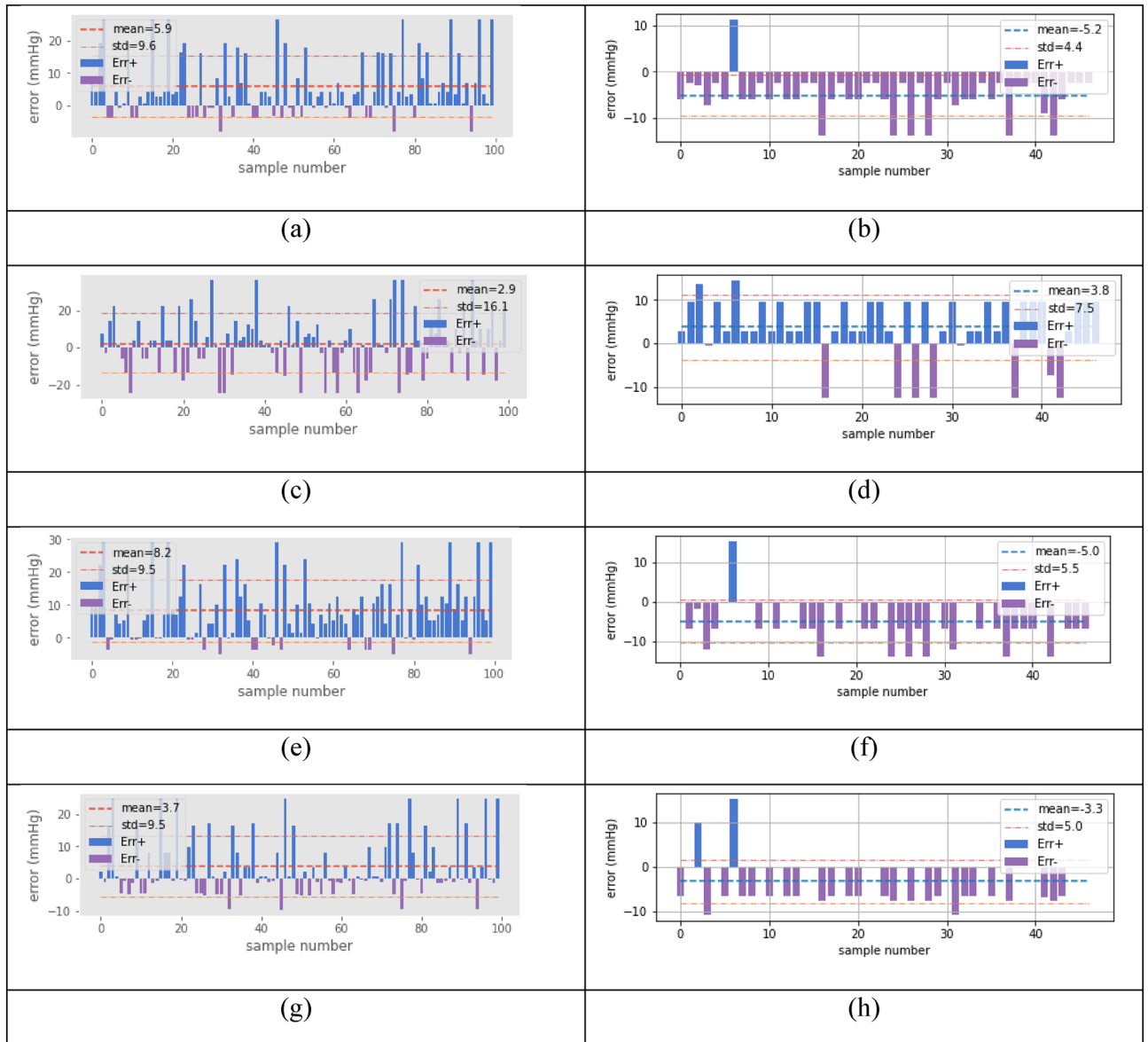


Figure 6. Error in units of pressure, left column, reveals the performance using all measures, and the right column uses restricted data. First line (a), (b) linear regression, second line (c), (d) ridge regression, third line (e), (f) SVM, and fourth line (g), (h) Random Forest.

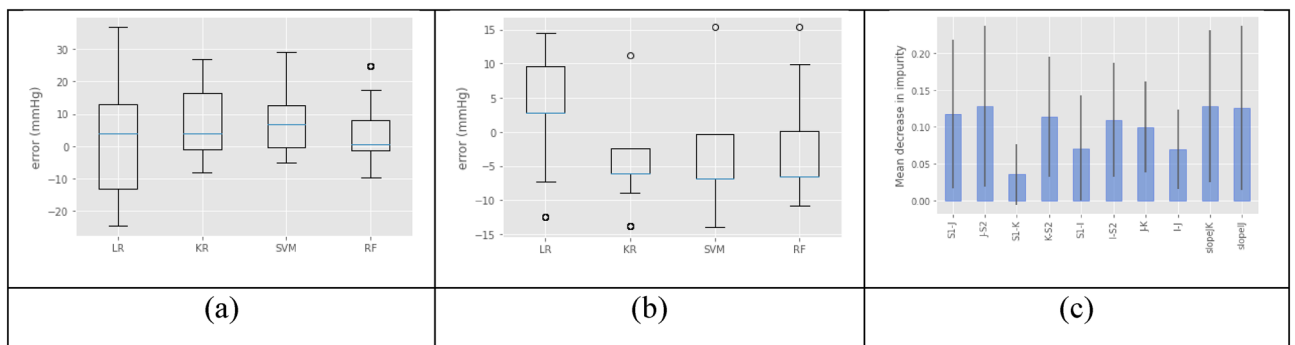


Figure 7. Boxplots: (a) Four algorithms evaluated in D_1 , (b) Four algorithms evaluated in D_2 , and (c) Feature importance based on MDI.

features are the intervals J-S2, followed by the slope J-K and slope IJ. Interestingly, point J is involved in two of the three features. Conversely, the feature with the least weight is the S1-K interval, which could be due to the distance between the two points.

Conclusion

In this work, a new cuffless method for estimating systolic blood pressure was proposed. It is not based on pulse transit time or pulse arrival time estimation, but relies on analyzing the correlation between the S1 and S2 of the phonocardiogram and the I, J, and K waves of the ballistocardiogram, which can be detected noninvasively in a simple way. For this, various intervals between these two signals were used as features to train the machine-learning algorithms, including the IJ and JK slopes of the ballistocardiogram. The means of the empirical distributions of the intervals S1-J, S1-I, and S1-K decreased as the systolic blood pressure increased. The same behavior was observed for the J-S2, I-S2, and K-S2 intervals, although with a noisy probability distribution. Except for the interval K-S2, these features were negatively correlated with the systolic pressure, with a p -val < 0.01 . Among all the regression algorithms employed, Random Forest exhibited a coefficient of determination of 0.48, and the mean error (mean difference) was 3.3 mmHg with a standard deviation of ± 5 mmHg in the estimation of systolic blood pressure. According to the mean decrease impurity, the best features to estimate systolic blood pressure were the S1-K interval and the JK slope of the BCG, which can also be used as indicators of changes in systolic blood pressure. From the results presented here, correlating the PCG and the BCG signals could be used to estimate changes in systolic blood pressure, and this could be a starting point to implement wearable systems that do not require pulse detection in peripheral arteries. Our vision is to develop a compact device capable of estimating blood pressure non-invasively by simultaneously detecting both PCG and BCG in a single area of the body (chest). This not only simplifies the blood pressure measurement procedure but may be useful for subjects with partial or total amputations of some or all of their limbs, and who cannot use current cuffless methods.

Data availability

The datasets generated during the current study are available in the MAYAS project repository, <https://github.com/MAYASproject/data>.

Received: 21 May 2022; Accepted: 11 October 2022

Published online: 13 October 2022

References

- Welykholowa, K. *et al.* Multimodal photoplethysmography-based approaches for improved detection of hypertension. *J. Clin. Med.* **9**, 1203 (2020).
- Lane, J. D., Greenstadt, L., Shapiro, D. & Rubinstein, E. Pulse transit time and blood pressure: An intensive analysis. *Psychophysiology* **20**, 45–49 (1983).
- Pereira, T., Sanches, R., Reis, P., Pego, J. & Simões, R. in *2015 IEEE 4th Portuguese meeting on bioengineering (ENBENG)*. 1–5 (IEEE).
- Escobar-Restrepo, B., Torres-Villa, R. & Kyriacou, P. A. Evaluation of the linear relationship between pulse arrival time and blood pressure in ICU patients: Potential and limitations. *Front. Physiol.* **9**, 1848 (2018).
- Wong, M. Y. M., Pickwell-MacPherson, E., Zhang, Y. T. & Cheng, J. C. The effects of pre-ejection period on post-exercise systolic blood pressure estimation using the pulse arrival time technique. *Eur. J. Appl. Physiol.* **111**, 135–144 (2011).
- Lanfranchi, A., Pépin, J. & Somers, V. Chapter 14—Cardiovascular physiology: Autonomic control in health and in sleep disorders. *Principles Practice Sleep Med.* 142–154 (2010).
- Martin, S.L.-O. *et al.* Weighing scale-based pulse transit time is a superior marker of blood pressure than conventional pulse arrival time. *Sci. Rep.* **6**, 1–8 (2016).
- Gao, M., Olivier, N. B. & Mukkamala, R. Comparison of noninvasive pulse transit time estimates as markers of blood pressure using invasive pulse transit time measurements as a reference. *Physiol. Rep.* **4**, e12768 (2016).
- Elgendi, M. *et al.* The use of photoplethysmography for assessing hypertension. *NPJ Digital Med.* **2**, 1–11 (2019).
- Dastjerdi, A. E., Kachuee, M. & Shabany, M. in *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*. 1–4 (IEEE).
- Yousefian, P. *et al.* The potential of wearable limb ballistocardiogram in blood pressure monitoring via pulse transit time. *Sci. Rep.* **9**, 1–11 (2019).
- Soerensen, K. *et al.* in *2017 Computing in Cardiology (CinC)*. 1–4 (IEEE).
- Thompson, W. B., Rappaport, M. B. & Sprague, H. B. Ballistocardiography: II. The normal ballistocardiogram. *Circulation* **7**, 321–328 (1953).
- Weissler, A. M., Harris, W. S. & Schoenfeld, C. D. Systolic time intervals in heart failure in man. *Circulation* **37**, 149–159 (1968).
- Peng, R.-C. *et al.* Cuffless and continuous blood pressure estimation from the heart sound signals. *Sensors* **15**, 23653–23666 (2015).
- Yousef, Q., Reaz, M. & Ali, M. A. M. The analysis of PPG morphology: Investigating the effects of aging on arterial compliance. *Meas. Sci. Rev.* **12**, 266 (2012).
- Pilt, K., Meigas, K., Ferenets, R., Temitski, K. & Viigimaa, M. Photoplethysmographic signal waveform index for detection of increased arterial stiffness. *Physiol. Meas.* **35**, 2027 (2014).
- Bartels, A. & Harder, D. Non-invasive determination of systolic blood pressure by heart sound pattern analysis. *Clin. Phys. Physiol. Meas.* **13**, 249 (1992).
- Kim, C.-S. *et al.* Ballistocardiogram: Mechanism and potential for unobtrusive cardiovascular health monitoring. *Sci. Rep.* **6**, 1–6 (2016).
- Wiens, A., Etemadi, M., Klein, L., Roy, S., & Inan, O. T. in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 5172–5175.
- Murphy, K. P. *Machine Learning: A Probabilistic Perspective* (MIT Press, 2012).
- Hara, S. & Hayashi, K. Making tree ensembles interpretable. *arXiv preprint arXiv:1606.05390* (2016).
- Gonzalez-Landaeta, R., Casas, O. & Pallas-Areny, R. Heart rate detection from an electronic weighing scale. *Physiol. Meas.* **29**, 979 (2008).
- Microsoft. *Visio Professional 2019* (Version 2208).
- Vazquez, K., Cota, J., Sifuentes, E. & Gonzalez, R. High Signal-to-noise ratio phonocardiogram using a shielded pvd film sensor. *IEEE Lat. Am. Trans.* **14**, 1139–1145 (2016).

26. Naseri, H., Homaeinezhad, M. R. & Pourkhajeh, H. Noise/spike detection in phonocardiogram signal as a cyclic random process with non-stationary period interval. *Comput. Biol. Med.* **43**, 1205–1213 (2013).
27. Ronchetti, E. M. & Huber, P. J. *Robust Statistics* (Wiley, 2009).
28. Zambom, A. Z. & Ronaldo, D. A review of kernel density estimation with applications to econometrics. *Int. Econometr. Rev.* **5**, 20–42 (2013).
29. Kovács, F., Torok, M. & Habermajer, I. A rule-based phonocardiographic method for long-term fetal heart rate monitoring. *IEEE Trans. Biomed. Eng.* **47**, 124–130 (2000).
30. Lee, S.-Y. *et al.* Electrocardiogram and phonocardiogram monitoring system for cardiac auscultation. *IEEE Trans. Biomed. Circ. Syst.* **13**, 1471–1482 (2019).
31. Jose, S. K., Shambharkar, C. M. & Chunkath, J. in *2015 International Conference on Computing and Communications Technologies (ICCCCT)*. 128–132 (IEEE).
32. Omari, T. & Bereksi-Reguig, F. A new approach for blood pressure estimation based on phonocardiogram. *Biomed. Eng. Lett.* **9**, 395–406 (2019).
33. Razali, N. M. & Wah, Y. B. Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *J. Stat. Model. Anal.* **2**, 21–33 (2011).
34. Schober, P., Boer, C. & Schwarte, L. A. Correlation coefficients: Appropriate use and interpretation. *Anesth. Analg.* **126**, 1763–1768 (2018).
35. Buitinck, L. *et al.* API design for machine learning software: Experiences from the scikit-learn project. *arXiv preprint arXiv:1309.0238* (2013).
36. London, A. J. Artificial intelligence and black-box medical decisions: Accuracy versus explainability. *Hastings Cent. Rep.* **49**, 15–21 (2019).
37. Quinn, T. P., Jacobs, S., Senadeera, M., Le, V. & Coghlan, S. The three ghosts of medical AI: Can the black-box present deliver?. *Artif. Intell. Med.* **124**, 102158 (2022).
38. Zhang, C., Bengio, S., Hardt, M., Recht, B. & Vinyals, O. Understanding deep learning (still) requires rethinking generalization. *Commun. ACM* **64**, 107–115 (2021).

Acknowledgements

The authors would like to thank the Autonomous University of Ciudad Juarez for its research facilities.

Author contributions

R.G.L. proposed the main idea, designed the acquisition protocol, acquired the signals, developed the theoretical framework, and wrote part of the methods; B.R. acquired the signals, and wrote part of the methods; J.M. worked on the signal processing and developed the experiments, statistical analysis, and wrote part of the methods. R.G.L., B.R., and J.M. approved the submitted version.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022