Outlier Management for Pulse Rate Variability Analysis from Photoplethysmographic Signals

Elisa Mejía-Mejía[1] and Panayiotis A. Kyriacou[2], *Member, IEEE*

***Abstract*— Pulse rate variability (PRV) has been proposed as a surrogate for the estimation of Heart Rate Variability (HRV), which is a non-invasive technique used to assess the cardiac autonomic activity. However, both physiological and technical factors may affect the relationship between HRV and PRV, and there are no standards for the analysis of PRV from photoplethysmographic (PPG) signals. The aim of this study was to determine the best outlier management strategies for PRV analysis. 117 PPG signals with randomly generated PRV information were simulated using Gaussian signals. From these, interbeat intervals were detected and different outlier detection and correction techniques were applied. Time and frequency-domain and non-linear PRV indices were extracted and compared with respect to the gold standard values obtained from the simulated PRV information. The results show that, in good quality PPG signals, there is no need to apply any outlier management technique for the extraction of PRV information.**

***Clinical relevance*— Establishing guidelines for PRV measurement can lead to more reliable and comparable results, as well as to the increase in the use of this variable for the diagnosis and monitoring of cardiovascular and autonomic conditions.**

I. INTRODUCTION

Pulse Rate Variability (PRV) refers to the changes in the duration of cardiac cycles through time, measured from pulsatile signals such as the photoplethysmogram (PPG) [1]. PRV has been used in recent years as a surrogate for Heart Rate Variability (HRV), which is measured from the electrocardiogram (ECG) and has been used for the assessment of the cardiac autonomic nervous activity [2], [3]. However, the relationship between HRV and PRV is not straightforward and there is still a debate regarding the use of PRV as an estimate of HRV [1], [4].

Physiological factors may explain the differences between HRV and PRV. PRV is measured from the PPG, a non-invasive optical technique widely used in the study and monitoring of the pulsations associated with changes in blood volume in a peripheral vascular bed [5], while ECG is an electrical signal, and both contain distinct information and are affected by different processes. Moreover, the ECG directly reflects the electrical activity of the heart, while PPG is measured in peripheral tissue and may be affected by vascular behaviour and other physiological variables, such as blood pressure [6], [7], [8].

*This work was not supported by any organization

[1]Elisa Mejía-Mejía is with the Research Centre for Biomedical Engineering, City, University of London, London, United Kingdom. elisa.mejia-mejia@city.ac.uk

[2]Panayiotis A. Kyriacou is with the Research Centre for Biomedical Engineering, City, University of London, London, United Kingdom. p.kyriacou@city.ac.uk

However, technical aspects may also be a source of disagreement between HRV and PRV. The measurement of HRV, which has been standardised before [9], relies on the detection of the R peaks from the ECG trace and the measurement of the so called R-to-R intervals (RRI). The ECG QRS complex possess a distinct shape whose identification is relatively simple [10]. Conversely, the PPG has a smoother shape without a remarkably distinct fiducial point [11] and there has not been any standardisation of the processes to obtain PRV information from the PPG signal. Factors such as the selection of the fiducial point to use for interbeat interval (IBI) extraction, sampling rate used to acquire the PPG signal, the methods used to pre-process the PPG signal, and the management of outliers in IBI time series, among others, can affect the results obtain from PRV and its relationship with HRV [1].

The aim of this study was to establish guidelines for outlier management in PRV analysis, using simulated PPG signals and simulated PRV information as gold standard. IBIs were extracted from simulated PPG signals and different outlier detection and management techniques were used to determine the effects of these methods in PRV indices. The use of simulated PPG signals and PRV information allows for the direct comparison of the expected results to the extracted PRV information, instead of using HRV information as gold standard. Thus, this is a direct analysis of how the different techniques affect PRV analysis, and controls for the physiological differences that may be included in the comparison against ECG-derived HRV indices.

II. MATERIALS AND METHODS

A. Signal simulation

PPG signals were simulated using a modified version of the model proposed by Tang et al [12], [13]. In this model, each single PPG pulse is simulated as the summation of two Gaussian functions with different parameters according to the quality of the PPG signal to be simulated. Tang et al. [12] determined these parameters based on PPG signals available in the MIMIC database. In the modified version of this model, the parameters for both Gaussian functions are not fixed, and the user can determine a ratio of amplitudes (r) of the Gaussian functions, being able to simulate almost any morphology of the PPG pulse.

Then, PRV information was simulated by generating randomly features for the expected PRV trace, considered as the gold standard, and the duration of each individual pulse was set according to the generated PRV trace. PRV information was simulated as a summation of four sinusoidal waves with

Manuscript 124 submitted to 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society . Received January 18, 2022.

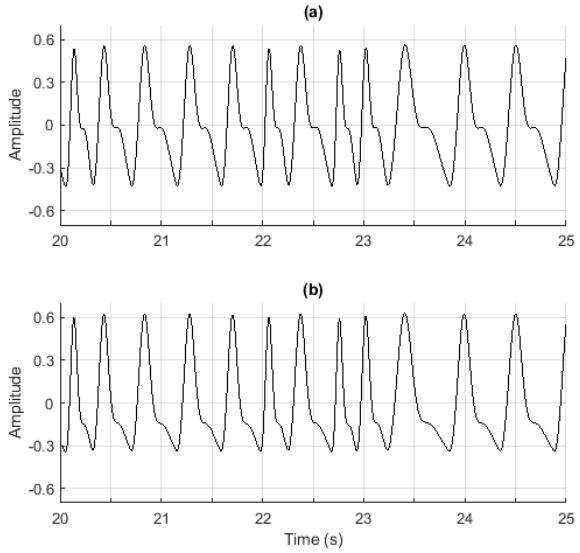

Fig. 1. Simulated photoplethysmographic signals with varying pulse rate variability information. (a) Excellent quality signal. (b) Acceptable quality signal.

randomly selected frequency, amplitude and offset, as shown in (1). The offset μ, which represents the mean duration of the pulses was set to be a random selected number between 300 and 1500 ms (for 200 and 40 beats per minute, respectively), while the amplitude, A, was obtained also as a random number between 50 and 80 ms. For the frequency information, two randomly generated values between 0.04 and 0.15 Hz, f_{LF1} and f_{LF2}, and two randomly generated values between 0.15 and 0.40 Hz, f_{HF1} and f_{HF2}, were used to simulate the low frequency (LF) and high frequency (HF) components from the PRV. These sinusoidal waves were considered as gold standard for PRV analysis.

$$PRV = \mu + A(sin(2\pi t f_{LF1}) + sin(2\pi t f_{LF2}) + sin(2\pi t f_{HF1}) + sin(2\pi t f_{HF2})) \quad (1)$$

Finally, all individual pulses were concatenated using a cubic spline interpolation, and filtered using a low-pass, second order Butterworth filter with cutoff frequency of 15 Hz. A total of 117 PRV traces were randomly generated, and used for the simulation of excellent ($r = 2$) and acceptable ($r = 4$) quality PPG signals. PPG signals were simulated using a sampling rate of 256 Hz and with a duration of 5 minutes. Fig. 1 exemplifies the simulated PPG signals. Signal simulation and analyses were performed in MATLAB R2020b.

B. Interbeat intervals extraction, and outlier detection and management

Cardiac cycles were identified from the simulated PPG signals using the algorithm described in [14]. Then, the *a* points from the second derivative of each cardiac cycle were used to extract the IBIs as the time difference between the location of consecutive *a* points. Outliers in these IBIs traces were detected and corrected applying the methods described in Tables I and II, respectively.

C. Pulse Rate Variability analysis

From the extracted IBIs, the corrected IBIs, and gold standard PRV traces, time-domain, frequency-domain and Poincaré plot indices were extracted.

From the time domain, the average duration of the IBIs (AVNN), their standard deviation (SDNN), the root mean squared value of the consecutive differences (RMSSD) and the proportion of consecutive differences larger than 50 ms (pNN50) were obtained. For the Poincaré plot analysis, the ellipse fitting technique was applied [15] and the area of the ellipse (S), its minor and major diameters (SD1 and SD2, respectively) and the ratio between diameters (SD1/SD2) were estimated.

For the frequency domain analysis, the IBIs were interpolated using cubic spline interpolation with 4 Hz sampling rate. The fast Fourier transform (FFT) with 512 data points was used to obtain the frequency spectra. From these, the low (LF, $0.04 \leq f < 0.15Hz$) and high frequency bands (HF, $0.15 \leq f < 0.40Hz$) were calculated, as well as the total power of the spectrum (TP, $0.0033 \leq f < 0.40Hz$). The ratio between LF and HF (LF/HF) and the normalised power of LF and HF bands (nLF and nHF, respectively) were obtained, as were the x and y coordinates of the centroid of LF, HF and TP (cLF_x, cLF_y, cHF_x, cHF_y, cTP_x and cTP_y).

D. Statistical analysis

The differences between indices extracted from the gold standard and the original and corrected IBIs were calculated and used for the statistical analyses.

Factorial analyses were performed to evaluate the effects of the detection and correction methods, as well as their interaction on the differences of each of the PRV indices.

TABLE I
OUTLIER DETECTION METHODS

Name	Outlier definition
Median	IBIs with values more than three scales median absolute deviations from the median
Mean	IBIs with values more than three standard deviations from the mean
Quartiles	IBIs with more than 1.5 interquartile ranges above the upper quartile or below the lower quartile
Grubb's test	IBIs are detected in an iterative manner, assuming the sample as normally distributed. IBIs are classified as outliers using the largest absolute deviation from the sample mean in units of the sample standard deviation as the statistic
Generalized extreme Studentized deviate (GESD) test	Similar to the Grubb's test but optimised for multiple outliers
Moving mean	IBIs with values more than three standard deviations from the mean over a window of 5 consecutive samples
Moving median	IBIs with values more than three scales median absolute deviations from the median over a window of 5 consecutive samples

Manuscript 124 submitted to 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society . Received January 18, 2022.

TABLE II

OUTLIER MANAGEMENT METHODS

Name	Description
Mean $k=5$	Replaces each outlier for the mean value of the 5 previous IBIs
Median $k=5$	Replaces each outlier for the median value of the 5 previous IBIs
Mean	Replaces each outlier for the mean value of the IBIs
Median	Replaces each outlier for the median value of the IBIs
Clip	Replaces each outlier with the lower or upper threshold value for elements smaller than or higher than three scaled median absolute deviations from the median
Previous	Replaces each outlier with the previous non-outlier value
Next	Replaces each outlier with the next non-outlier value
Nearest	Replaces each outlier with the nearest non-outlier value
Linear	Replaces each outlier after linearly interpolating neighbouring, non-outlier values
Spline	Replaces each outlier after applying a cubic spline interpolation with neighbouring, non-outlier values
Piecewise Spline	Replaces each outlier after applying a shape-preserving piecewise cubic spline interpolation with neighbouring, non-outlier values
Makima	Replaces each outlier after applying a modified Akima cubic Hermite interpolation with neighbouring, non-outlier values

Since the data did not follow a normal distribution, as checked using the Lilliefors test, Box-Cox transformations were applied to the differences, after finding the optimal lambda for each case. Then, the combination of factors that gave the lowest difference for each index was compared to the differences obtained if no outlier detection and management strategy was applied, using Wilcoxon rank sum tests. A 95% significance value was used for all the analyses.

Statistical analyses were performed in RStudio (version 1.4.1717).

III. RESULTS

From the factorial analyses, it was found that there were significant differences due to the interaction of factors only for measuring pNN50, both with excellent and acceptable quality signals. The detection method was a significant factor for SDNN, RMSSD, SD1 and SD2 in both types of signals, while AVNN and SD1/SD2 showed significant differences due to the detection method only when measured in acceptable quality signals. The correction method did not show statistically significant differences for any index or any type of signal. Table III summarises the combination of factors that gave the lowest difference to the gold standard for the estimated indices.

After comparing these best combinations to the indices extracted from the original IBIs, i.e. without managing outliers, it was found that only RMSSD ($p < 0.001$) and SD1 ($p < 0.001$) showed significant differences when measured from excellent PPG signals, while S showed significant differences ($p < 0.001$) when measured from both types of signals. In all these cases, the mean difference between extracted and gold standard indices were lower when no outlier management strategy was applied, as shown in Table IV.

TABLE III

COMBINATION OF FACTORS WITH THE LOWEST DIFFERENCE TO INDICES EXTRACTED FROM THE GOLD STANDARD.

Index	Excellent PPG		Acceptable PPG	
	A	B	A	B
AVNN †	Quartiles	Median	Quartiles	Median
SDNN ‡	GESD test	Next	Mean	Next
RMSSD ‡	Moving median	Next	Grubb's test	Next
pNN50 ⋆	Mean	Median $k=5$	Mean	Median
LF	Moving median	Previous	Moving median	Previous
HF	Moving median	Median $k=5$	Moving median	Piecewise Spline
TP	Moving median	Next	Moving median	Nearest
nLF	Moving median	Linear	Moving median	Linear
nHF	Moving median	Median $k=5$	Mean	Linear
LF/HF	Moving median	Median $k=5$	Moving median	Median $k=5$
cLF_x	Mean	Clip	Moving median	Previous
cLF_y	Moving median	Linear	Moving median	Median $k=5$
cHF_x	Moving median	Previous	Moving median	Previous
cHF_y	Moving median	Spline	Moving median	Linear
cTP_x	Moving median	Mean	Mean	Clip
cTP_y	Moving median	Nearest	Moving median	Clip
S	Moving median	Next	Moving median	Next
SD1 ‡	Moving median	Next	Grubb's test	Median
SD2 ‡	Quartiles	Linear	Quartiles	Median
SD1/SD2 †	Mean	Next	Grubb's test	Makima

A: Detection method
B: Correction method
‡ Significant difference due to A on both types of signals
† Significant difference due to A on acceptable signals
⋆ Significant difference due to the A*B on both types of signals

IV. DISCUSSION

PRV describes the changes in pulse rate over time [1]. It is usually measured from PPG signals, which are obtained using an optical, non-intrusive and low-cost device widely used in clinical and wearable devices [16]. PRV has

TABLE IV

AVERAGE DIFFERENCE TO GOLD STANDARD FOR INDICES THAT SHOWED SIGNIFICANT DIFFERENCES BETWEEN APPLYING OR NOT APPLYING OUTLIER MANAGEMENT STRATEGIES

Index	Excellent PPG		Acceptable PPG	
	No management	Best combination	No management	Best combination
RMSSD	-0.0014 ± 0.0062	0.0045 ± 0.0064	-	-
S	-0.0100 ± 0.0408	0.0115 ± 0.0448	-0.0076 ± 0.0374	0.0135 ± 0.0410
SD1	$-9.78x10^{-4} \pm 0.0043$	0.0032 ± 0.0045	-	-

Manuscript 124 submitted to 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society . Received January 18, 2022.

been suggested as an alternative to HRV, although their relationship is not straightforward and both physiological and technical aspects could affect the measurement of PRV [1], [4]. Moreover, the measurement of PRV has not been standardised and different studies usually do not align their methodologies, precluding the comparison of the results and the validation of the technique. Some aspects that could affect PRV analysis are the sampling rate used to acquire the PPG, the identification and selection of fiducial points to segment the cardiac cycles, the length of the window used for PRV analysis, and the management of outliers and ectopic beats from the PRV traces. In this study, the aim was to determine the best way to manage these outliers, by investigating the effects of different detection and correction strategies.

The obtained results showed that most of the extracted indices were not affected by the selection of outlier management strategies, and even not controlling for outliers showed good results when compared against the gold standard. This means that with good quality signals there is no need to manage outliers before extracting PRV indices. This is especially true for the assessment of frequency domain indices, which did not show any difference among outlier management strategies. These results are both for excellent and acceptable PPG signals, suggesting that PRV analysis might be performed without managing outliers when the signal has a good signal-to-noise ratio and when the signal is measured from different sites of the body, which have been shown to alter the morphology of the signal [17].

This study has some limitation. Firstly, simulated PPG signals with simulated PRV information were used in this study. This was done with two main purposes. It is simpler to obtain larger number of samples using simulated data, which gives statistical validity to the experiment. The sample size for this study was estimated to be the optimal value in order to observe differences of 2% in the measurement of the indices, compared to the gold standard. Also, by simulating PRV information it was possible to obtain a gold standard that was not HRV information obtained from the ECG. As mentioned, physiological aspects may explain part of the differences between HRV and PRV, hence comparing them in order to establish methodologies and strategies for obtaining PRV information its not ideal. However, the results from this study need to be validated using real PPG data. Secondly, there was no noise in the simulated signals, which could enhance the performance of the algorithm used to segment the cardiac cycles and diminish the presence of outliers. A similar analysis should be performed with PPG signals that contain different types and levels of noise in order to evaluate the performance of outlier management strategies in these cases. Finally, the agreement between indices was not assessed. Future studies should investigate not only the significance of the difference but also determine how the indices agree using techniques such as Bland-Altman plots [18].

REFERENCES

[1] E. Mejía-Mejía, J.M. May, R. Torres, and P.A. Kyriacou, "Pulse rate variability in cardiovascular health: A review on its applications and relationship with heart rate variability," *Physiol. Meas.*, vol. 41, 2020, pp. 07TR01. DOI: 10.1088/1361-6579/ab998c.

[2] D. Quintana, "Statistical considerations for reporting and planning heart rate variability case-control studies," *Psychophysiology*, vol. 54 (3), 2017, pp. 344–349. DOI:10.1111/psyp.12798.

[3] M. Malik, H. Huikuri, F. Lombardi, and G. Schmidt, "The purpose of heart rate variability measurements," *Clin. Aut. Res.*, vol. 27, 2017, pp. 139–140. DOI: 10.1007/s10286-017-0416-8.

[4] A. Schäfer, and J. Vagedes, "How accurate is pulse rate variability as an estimate of heart rate variability? A review on studies comparing photoplethysmographic technology with an electrocardiogram," *Int. J. Cardiol.*, vol. 166, 2013, pp. 15–29. DOI: 10.1016/j.ijcard.2012.03.119.

[5] P.A. Kyriacou, "Introduction to photoplethysmography," in Photoplethysmography: Technology, Signal Analysis, and Applications, P.A. Kyriacou and J. Allen, Eds. London, UK: Elsevier, 2021, pp. 69-145.

[6] E. Mejía-Mejía, K. Budidha, T.Y. Abay, J.M. May and P.A. Kyriacou, "Heart Rate Variability (HRV) and Pulse Rate Variability (PRV) for the Assessment of Autonomic Responses," *Front. Physiol.*, vol. 11, 2020, p. 779. DOI: 10.3389/fphys.2020.00779.

[7] E. Gil, M. Orini, R. Bailón, J. M. Vergara, L. Mainardi and P. Laguna, "Photoplethysmography pulse rate variability as a surrogate measurement of heart rate variability during non-stationary conditions," *Physiol. Meas.*, vol. 31 (9), 2010, pp. 1271–1290. DOI: 10.1088/0967-3334/31/9/015.

[8] E. Mejía-Mejía, J. M. May, M. Elgendi and P. A. Kyriacou, "Differential effects of the blood pressure state on pulse rate variability and heart rate variability in critically ill patients," *npj Digit. Med.*, vol. 24, 2021, p. 82. DOI: 10.1038/s41746-021-00447-y.

[9] Task Force of the European Society of Cardiology and The North American Society of Pacing and Electrophysiology, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use," *Eur. Heart J.*, vol. 17, 1996, pp. 354–381. DOI: 10.1161/01.CIR.93.5.1043.

[10] G.D. Clifford, F. Azuaje, and P.E. McSharry, Advanced Methods and Tools for ECG Data Analysis. Norwood, MA: Artech House, 2006.

[11] E. Mejía-Mejía, J. Allen, K. Budidha, C. El-Hajja, P.A. Kyriacou and P.H. Charlton, "Photoplethysmography signal processing and synthesis," in Photoplethysmography: Technology, Signal Analysis, and Applications, P.A. Kyriacou and J. Allen, Eds. London, UK: Elsevier, 2021, pp. 69-145.

[12] Q. Tang, Z. Chen, R. Ward and M. Elgendi, "Synthetic photoplethysmogram generation using two Gaussian functions," *Sci. Rep.*, vol. 10, 2020, p. 13883. DOI: 10.1038/s41598-020-69076-x.

[13] Q. Tang, Z. Chen, J. Allen, A. Alian, C. Menon, R. Ward and M. Elgendi, "PPGSynth: An Innovative Toolbox for Synthesizing Regular and Irregular Photoplethysmography Waveforms," *Front Med (Lausanne)*, vol. 7, 2020, p. 597774. DOI: 10.3389/fmed.2020.597774.

[14] M. Elgendi, I. Norton, M. Brearley, D. Abbott and D. Schuurmans, "Systolic peak detection in acceleration photoplethysmograms measured from emergency responders in tropical conditions," *PLoS One*, vol. 8 (10), 2013, p. e76585. DOI: 10.1371/journal.pone.0076585.

[15] A.H. Khandoker, C. Karmakar, M. Brennan, A. Voss and M. Palaniswami, Poincaré Plot Methods for Heart Rate Variability Analysis. New York, NY: Springer, 2013. DOI: 10.1007/978-1-4614-7375-6.

[16] P.H. Charlton and V. Marozas, "Wearable photoplethysmography devices," in Photoplethysmography: Technology, Signal Analysis, and Applications, P.A. Kyriacou and J. Allen, Eds. London, UK: Elsevier, 2021, pp. 69-145.

[17] E. Peralta, J. Lazaro, R. Bailon, V. Marozas and E. Gil, "Optimal fiducial points for pulse rate variability analysis from forehead and finger photoplethysmographic signals," *Physiol Meas*, vol. 40, 2019, p. 025007. DOI: 10.1088/1361-6579/ab009b.

[18] J.M. Bland and D.G. Altman, "Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement," *Lancet*, vol. 1 (8476), 1986, pp. 307-310.

Manuscript 124 submitted to 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society . Received January 18, 2022.