Drowsiness Detection Using Multivariate Statistical Process Control

Ana Rita Antunes[1,2(✉)], Ana Cristina Braga[2], and Joaquim Gonçalves[1]

[1] 2Ai, Polytechnic Institute of Cávado and Ave, 4750-810 Barcelos, Portugal
{arantunes,jgoncalves}@ipca.pt
[2] ALGORITMI Center, University of Minho, 4710-057 Braga, Portugal
acb@dps.uminho.pt

Abstract. Drowsiness at the wheel has been studied for different countries since it is important for road safety and its prevention. Since it is considered a public health problem, solutions must be found to avoid worse scenarios and to identify a low-cost system.

Therefore, this work aims to detect the drowsy state, without labeling it manually, considering the heart rate variability. To make this possible, driving simulations were performed, using a wearable device. In terms of methodology, multivariate statistical process control, considering principal component analysis, was implemented, and compared with a similar study. Three principal components were computed taking into consideration time, frequency, and non-linear domain, every two minutes. Thereafter, Hotelling T^2 and squared prediction error statistics were estimated. These statistics were estimated considering each principal component, individually. Thereby, the results achieved seemed to be promising to identify drowsiness peaks. However, the study developed has limitations, like the identification of points out-of-control occurred due to signal noise and it does not identify all the drowsiness peaks. Conversely, it was not used information from the participants' awake states as a reference. Therewith, new simulations must be done, and new information must be added to avoid noise and to detect more drowsiness peaks.

Keywords: Drowsy · Driving · Simulation · Heart rate variability · Multivariate statistical process control · Principal component analysis

1 Introduction

Sleep is essential for the physical and psychological well-being and longevity of a human being, where it is crucial for cognitive processes, emotional regulation,

This paper was funded by the project "NORTE-01-0247-FEDER-0039720", supported by Northern Portugal Regional Operational Programme (Norte2020), under the Portugal 2020 Partnership Agreement, through the European Regional Development Fund (ERDF)".

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022
O. Gervasi et al. (Eds.): ICCSA 2022 Workshops, LNCS 13377, pp. 571–585, 2022.
https://doi.org/10.1007/978-3-031-10536-4_38

executives, and attention. These processes are very important for memory, learning, productivity, and concentration [1,2]. Consequently, sleep deprivation can have serious problems in security, health, and quality of life that can lead to a decrease in the individual's productivity, as well as an increase in work accidents, such as driving [1,3,4]. Sleep disorders are a public health problem and have been associated with increased risk of hypertension, depression, obesity, heart attack, and stroke [5]. In Portugal, 22.3% of drivers confirmed that fell asleep while they were driving and 1.7% had an accident for falling asleep [6].

The devices available use moves sensors at the wheel and/or cameras located in the rearview mirror and the system recognize the driver fatigue and send an alert. These technologies are expensive and are only available in major brands. Therefore, these technologies are available for those who have purchase power [7]. It is necessary to replace the detection systems with a forecasting one and detection of disturbance problems or sleep deprivation.

Thus, drowsiness at the wheel has been studied over the years and there are still aspects to be improved. Most of the studies found intended to classify the sleepy state, using different machine learning algorithms. Even thought Multivariate Statistical Process Control (MSPC) is widely used in industry area, this methodology has been applied in the healthcare area, in recent years, considering the Heart Rate Variability (HRV). For example, this approach was developed for epileptic seizure prediction and to detect driver drowsiness [8–10].

In this work, MSPC based on Principal Component Analysis (MSPC-PCA) is performed to identify a drowsy state using biometric data, from driving simulator experiments. The work is carried out by the comparison of the work [10], where the methodology was implemented differently in an attempt to obtain more out-of-control points that are related to the detection of sleepiness.

This article is structured as follows. Section 2 presents a literature review, where some concepts about drowsy detection methods, questionnaires to evaluate sleep disorders and MSPC-PCA concepts are briefly explained. Materials and Methods are presented in Sect. 3, where it is explained the proposed drowsiness detection procedure, variables description and the implementation details. In Sect. 4 there are the information of the results achieved. Firstly, it is introduced the participants description, then the HRV analysis is performed and the discussion of results. Finally, the main conclusions and the proposed of future work are described in Sect. 5

2 Literature Review

In this section, it will be firstly introduced the existing methods for detecting sleepiness, than the questionnaires to identify sleep disorders. Therefore, the MSPC-PCA concepts are presented.

2.1 Methods for Detecting Sleepiness

For the detection of sleepiness, there are different types of methods, that can be divided into subjective, vehicle, behavioral and physiological measures.

Subjective Measures. Subjective measures are used to assess the level of sleepiness considering the drivers' opinion. Where the Karolinska sleepiness scale [11,12] is widely implemented and has nine situations: (1) extremely alert, (2) very alert, (3) alert, (4) fairly alert, (5) neither alert nor sleepy, (6) some signs of sleepiness, (7) sleepy, but no effort to keep alert, (8) sleepy, some effort to keep alert, (9) very sleepy, great effort to keep alert, fighting sleep. This measure is not practical in real driving situations and is based on opinions. Note that the driver is not always aware that he/she is showing signs of sleepiness [13].

Vehicle-Based Measures. The vehicle-based measures use the information available in the cars' intelligent systems. Thus, it can extract information about driving speed, and standard deviation of lane position, this is the variability of lane position, steering wheel movement, and driving duration [14]. This type of measure can be easily collected but, in some situations, is very limited since it depends on external factors like the road's geometric characteristics. For example, road marking, climatic, and lighting conditions. On the other hand, when the alert of sleepiness is sent to the driver it means that the driver presents already signs of sleepiness. This can be a problem since it may not prevent road accidents due to late detection [13,15].

Behavioral Measures. Behavioral measures aim to evaluate the driver behavior in order to detect the drive fatigue. Thus, it can be acquired the number and duration of blinks, frequent yawning, head position, and, a well-known metric, the percentage of eyelid closure over the pupil over time [16,17]. These metrics are non-intrusive for the driver activity, but the illumination and the use of glasses can be a problem, even though there are cameras with better performance during the day and others during the night [13,18].

Physiological Measures. The physiological measures are the most reliable metrics to detect sleepiness since it can alert the driver beforehand. Thus, electroencephalogram (EEG), electrocardiogram (ECG), electrooculogram (EOG), electromyogram (EMG), and photoplethysmograph (PPG) are the most known signs. Where these signs take into consideration brain activity, heart rate, eye movement, muscles activity, and blood volume changes, respectively [19–21]. Note that the EEG, EOG, and EMG are the signs used to identify the respective sleep phases [22]. Even though these metrics are the most reliable, they can be extracted wrong information and it is very intrusive to drivers' activity. Moreover, it is important to rectify this disadvantage, where non-intrusive solutions have been developed using wearable devices to measure these signs [13].

2.2 Questionnaires

The Epworth sleepiness scale [23] is widely used to measure the level of daytime sleepiness. There are eight questions based on different situations, where

the subject has to rate on a scale of 0 to 3 how likely they would fall asleep, considering their recently routine. So, the final score can be between 0 and 24. If the final score is greater than 10, the subject has hypersomnia.

In contrast, the STOP-Bang questionnaire [24] is applied to identify the risk of developing obstructive sleep apnea. This one takes into consideration the snoring, tiredness, observed apnea, high blood pressure, body mass index, age, neck circumference, and male gender. Thus, there are eight dichotomous questions and the final score is between 0 and 8. Those with scores equal to or greater than 5 have a high risk of developing moderate to severe obstructive sleep apnea.

Another questionnaire available and commonly used is the Pittsburgh sleep quality [25] that assesses the sleep quality in the previous month. This one has 19 questions that are divided into seven components: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medications, and daytime dysfunction. Each component has a score of 0 to 3 and, the final score (0 to 21) is the sum of the scores of each component. The subjects with a final score greater than 5 have a bad sleep quality.

The circadian rhythm, our twenty-four-hour rhythm, is another aspect to analyze, when sleepiness is being studied since it is possible to identify when the subject wants to be awake or asleep. For example, a diurnal person has a preference to stay awake during the day and sleep at night [26]. Thus, the Morning-Eveningness questionnaire [27] consists of 19 questions and the final score is between 16 to 86. The subjects with classification between 16 to 30 is a definitely evening type, 31 to 41 moderately evening type, 42 to 58 neither type, 59 to 69 moderately morning type and 70 to 86 definitely morning type.

2.3 Multivariate Statistical Process Control Based on Principal Component Analysis

Statistical process control is used to monitor the process performance to verify if it is working like should be and to detect when an anomaly occurs. Control charts are a very practical visual method for understanding the operation of a given process, where it is possible to do a univariate or multivariate analysis. This is, evaluate one variable individually or all variables in simultaneously [28].

However, some analyses can be misleading and difficult to interpret due to the highly correlated between variables. Thus, there are some difficulties when multivariate process data is being analyzed, like the dimensionality, colinearity, the presence of noise, and missing data [29].

Principal Component Analysis (PCA) is an alternative method capable to simplify the analysis and the process of monitoring faults since it projects the data into low dimensional spaces. Thus, in order to eliminate colinearity and dimensionality, PCA computes a set of orthogonal and linearly uncorrelated variables capable to explain the original information, call as principal components [28].

Considering the matrix $\mathbf{X} \in \mathbb{R}^{N \times M}$, where N and M are the number of observations and the number of variables, respectively. Where the firsts compo-

nents explain most of the variability, in the original data, and R ($\leq M$) is the number of components retained. Note that the values in $\mathbf{X}$ are scaled to zero mean and unit variance. The principal components can be expressed as Eq. 1,

$$\mathbf{T} = \mathbf{XP} + \mathbf{E} \tag{1}$$

where $\mathbf{T} \in \mathbb{R}^{N \times R}$ is called the scores matrix, $\mathbf{P} \in \mathbb{R}^{M \times R}$ is the loading matrix and $\mathbf{E} \in \mathbb{R}^{N \times M}$ is the residuals, in other words, is the information that is not explained by the PCA model. Furthermore, the estimation of $\mathbf{X}$ can be made considering Eq. 2 and, consequently, the errors can be written as Eq. 3 [28,30,31].

$$\hat{\mathbf{X}} = \mathbf{TP^T} \tag{2}$$

$$\mathbf{E} = \mathbf{X} - \hat{\mathbf{X}} \tag{3}$$

Considering this information, the Squared Prediction Error (SPE), also known as the Q statistic, can be expressed like in Eq. 4. This statistic is used to measure the difference between the original information and the R dimensional subspace.

$$SPE = \sum_{i=1}^{M} (\mathbf{x}_i - \hat{\mathbf{x}_i})^2 \tag{4}$$

Another statistic widely used to monitoring the process stability is the Hotelling's T^2. The expression is given by Eq. 5,

$$T_i^2 = \mathbf{t_i^T \Lambda^{-1} t_i} = \sum_{\mathbf{r=1}}^{\mathbf{R}} \frac{\mathbf{t_r^2}}{\lambda_\mathbf{r}} \tag{5}$$

where $\Lambda \in \mathbb{R}^{R \times R}$ is the covariance matrix of $\mathbf{T}$, $\mathbf{t_i^T} = \{t_{i1}, t_{i2}, \ldots, t_{iR}\}$ is the score vector for the ith observation and λ_r is the eigenvalues of the R component.

Hence, the upper limit control (ULC) for Hotelling's T^2 (Eq. 6) can be compute considering the F-distribution, considering the $100(1-\alpha)\%$ confidence.

$$ULC(T^2) = \frac{R(N^2-1)}{N(N-R)} \times F_{R,(N-R),\alpha} \tag{6}$$

However, the SPE control limit, considering the significant level α, is express like in Eq. 7,

$$UCL(SPE) = \frac{v}{2b} \chi^2 \left(\frac{2b^2}{v}, \alpha \right) \tag{7}$$

where b and v are the sample mean and variance values.

3 Materials and Methods

This section addresses the drowsiness detection procedure, where all the steps are explained. Besides that, a brief explanation of the variables, in the study, and the implementation details is also presented.

3.1 Drowsiness Detection Procedure

The drowsiness detection procedure developed requires several steps. Figure 1 shows a summary of the required steps to implement the proposed methodology.

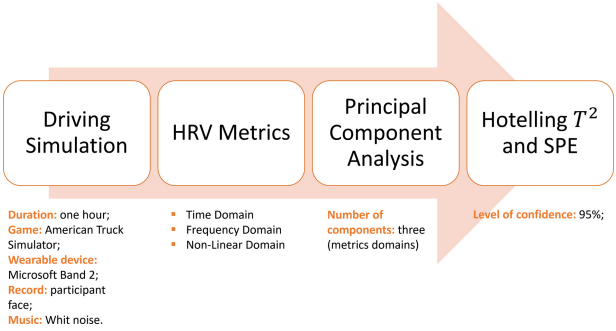


Fig. 1. Work flowchart.

Thus, in order to identify the participants' awake and drowsy state, a strategy was designed to distinguish these different states. For the achievement of this a driving simulation was conducted, between July 5 and August 3, 2021. All the participants agreed to be part of the study and the anonymity was guaranteed.

Thus, the participants had to fill out four questionnaires about daytime sleepiness, the risk of obstructive sleep apnea, sleep quality, and the circadian rhythm [23–25,27]. Therefore, during the simulation, the participant used a wearable device, Microsoft Band 2, to collect physiological information. With this device, there is information about the heart rate and R-R intervals. Note that the simulation room was completely empty to not compromise the participant's condition with noise and the lights were off.

In terms of the game, it was used for the `American Truck Simulator` since there are highways with long kilometers and it is possible to define a single route for all participants. The simulation duration was about one hour [32,33]. When the participant is driving within a city, the maximum speed is 30 km/h and on a highway is 55 km/h. It was asked to each participant to not exceed the speed limits to make it more real and to induce sleepiness. Another way to induce more sleepiness was to play a monotonous music [1] along the route and with the sound of the truck working.

The next step was to extract the HRV metrics, known as time, frequency and non-linear domains. This information it will be used to detect the drowsy state. Where the implementation of PCA is the next step to take into consideration, which aims to divide the variables by their respective domains as they are properly validated.

[1] Available in: https://www.youtube.com/watch?v=wzjWIxXBs_s. Acceded between July 5 and August 3, 2021.

Lastly, the identification of the Hotelling T^2 and SPE statistics and control limits are the final step. This information was estimated for each principal component, individually, to identify the maximum number of observation out-of-control and verify if there are signs of drowsiness using the video recording.

3.2 Variables Description

Heart Rate Variability (HRV) is the time between consecutive heartbeats, which gives rise to the R-R intervals or Inter-beat interval (IBI) [34,35]. It is an essential tool for characterizing and understanding the regulation of the cardiovascular system through the autonomic nervous system (ANS) [36]. In Fig. 2 it is presented the ECG signal, where the time between two adjacent highest peaks, known as R-peak, is the R-R interval [34].

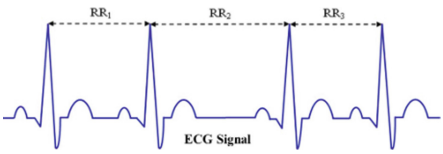


Fig. 2. RR-Intervals for ECG signals (from [34]).

Furthermore, there are time domain, frequency and non-linear metrics. Time domain measures quantify the amount of HRV over a given period of time. However, frequency metrics compute absolute or relative signal power and, conversely, the non-linear metrics measure the unpredictability and complexity [35]. The time [T], frequency [F] and non-linear [NL] variables were extracted using the python open-source package *HRV-analysis*, every two minutes, and the follow descriptions [37].

T mean_{nni}: R-R intervals mean;
T sdnn: R-R standard deviation;
T sdsd: The standard deviation of differences between adjacent R-R intervals;
T rmssd: The square root of the mean of the sum of the squares of differences between adjacent N-N intervals;
T median_{nni}: Median Absolute values of the successive differences between the R-R intervals;
T nni_{50}: Number of interval differences of successive R-R intervals greater than 50 ms;
T pnni_{50}: The proportion derived by dividing nni_50 by the total number of R-R intervals;
T nni_{20}: Number of interval differences of successive R-R intervals greater than 20 ms;
T pnni_{20}: The proportion derived by dividing nni_20 by the total number of RR-intervals;

T range$_{nni}$: Difference between the maximum and minimum nn_interval;
T cvsd: Coefficient of variation of successive differences equal to the rmssd divided by mean$_{nni}$;
T cvnni: Coefficient of variation equal to the ratio of sdnn divided by mean_nni;
T mean$_{hr}$: The mean Heart Rate;
T max$_{hr}$: Maximum heart rate;
T min$_{hr}$: Minimum heart rate;
T std$_{hr}$: Standard deviation of heart rate;
F total power: Total power density spectral;
F vlf: variance in HRV in the very low frequency;
F lf : variance in HRV in the low frequency;
F hf: variance in HRV in the high frequency;
F lf_hf_ratio: lf/hf ratio;
NL csi: Cardiac Sympathetic Index.
NL cvi: Cadiac Vagal Index.
NL Modified csi: Modified CSI;
NL sampen: sample entropy.

Before the variables extraction, the R-R intervals must be resample at equal intervals and afterward must be clean, this is, the outliers and ectopic beats must be treated. In terms of outliers, the values below 300 and above 2000 were considered. The package *HRV-analysis* has two functions to remove the outliers and ectopic beats, *remove_outliers* and *remove_ectopic_beats*. The next step was to interpolate the missing values, considering the cubic method and the function *interpolate_nan_values* [37]. Figure 3 presents a summary of what was done to extract the variables through a flowchart.

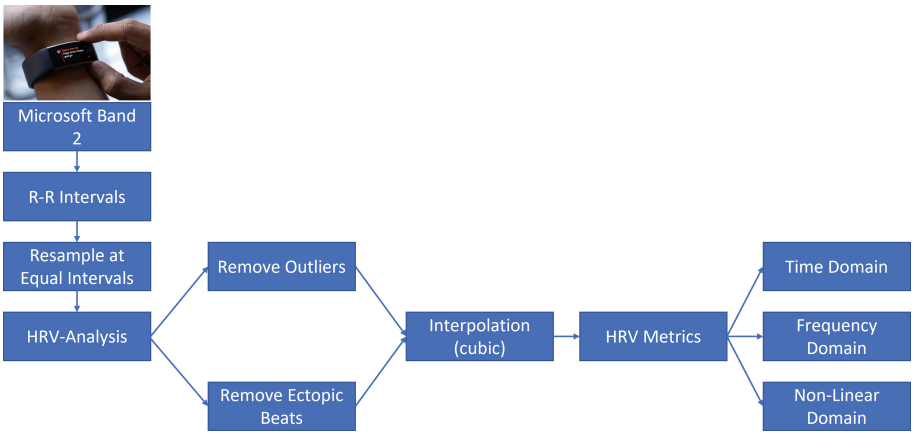


Fig. 3. Flowchart for the variables extraction.

3.3 Implementation Details

In order to produce the HRV analysis, the software Python (version 3.8.5) was used [38]. Firstly, the csv file was imported using the library `pandas` considering the function `read_csv` [39]. In a second instance, the library `matplotlib` was employed to graphically visualize the biometric data collected and the questionnaire results using the `pyplot` function [40].

Then, after the variables extraction, the standardization of data was performed using the `mean()` and `std()` `pandas` functions [39]. Moreover, PCA was conducted considering the library `pca`, the functions `pca`, and `fit_transform` to identify the scores of the three principal components and the loading's values [41].

Thereafter, the identification of the Hotelling T^2 and SPE statistics and control limits were assessed with the `numpy` and `SciPy` libraries [42,43]. The covariance matrix was computed with the fuction `cov` and its inverse value by the function `linalg.inv`. Then, matrix multiplication was performed with the `dot` function. In the case of the SPE statistic, the sum of the errors was done with the `sum` function. Finally, the distributions were calculated by the functions `f.ppf` and `chi2.ppf`.

4 Results

In this section, the participants information is provided and the questionnaires results are shown. Moreover, the HRV analysis is conducted by applying MSPC-PCA, and finally, it is made a discussion of the results.

4.1 Participants Description

The simulation was conducted during one month, where the number of participants was twelve. Each volunteer performed at least two simulations, one in the morning and one in the afternoon. Thus, thirty-one simulations were carried out with voluntary participation, at the Polytechnic Institute of Cávado and Ave.

The age range of the participants varies from 22 to 53 years old. Although, 50% are under the age of 28. According to Fig. 4, most of the participants are male (83.3%), with normal weight (50%), 41.7% have overweight and 8.3% have class I obesity.

In terms of sleep disorders, the questionnaires results are present in Fig. 5. There are 33.3% of participants that are a morning type, 41.7% neither type, 8.3% as definitively morning and 16.7% evening type. Nonetheless, 25% has excessive daytime sleep and risk of developed obstructive sleep apnea, and 8.3% has bad sleep quality.

4.2 Heart Rate Variability

In this subsection, only the results of two participant will be presented. The chosen participants had different experiences, where Participant 1 did not experience drowsiness, unlike Participant 2, according to their opinions.

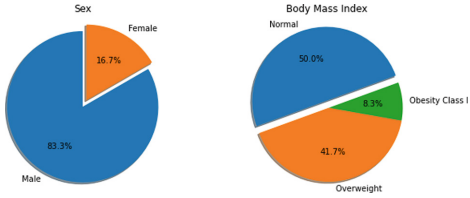


Fig. 4. Sex and body mass index for the participants.

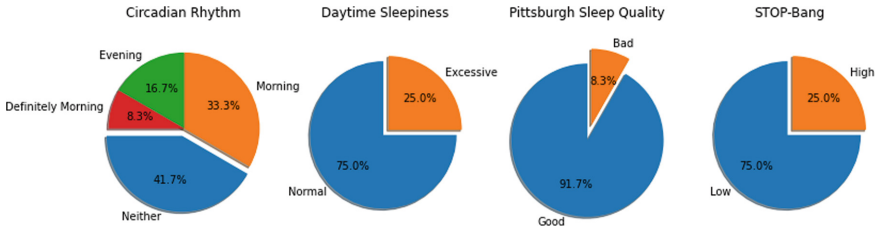


Fig. 5. Questionnaires results.

Moreover, the heart rate descriptive statistics are shown in Table 1. According to the results presented, Participant 1 has a lower mean heart rate than Participant 2. However, Participant 2 has a greater variability when comparing the minimum and maximum heart rate values. Although, the standard deviation and the coefficient of variation (CV) are smaller for Participant 2.

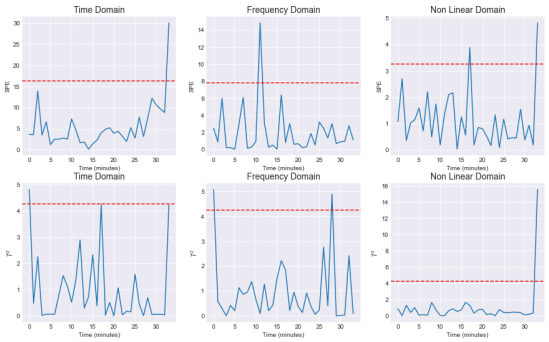
Table 1. Heart rate descriptive statistics for Participant 1 and 2.

Participant	Mean	Std	CV	Minimum	Q1	Median	Q3	Maximum	Range
1	63.16	3.53	5.59	5 2.64	60.87	62.80	65.20	77.61	24.97
2	77.86	3.46	4.44	63.96	75.83	78.17	79.94	91.41	27.45

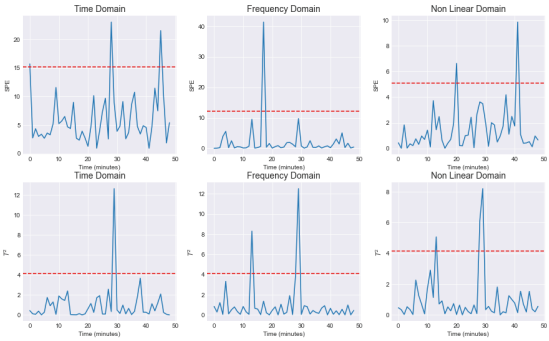
Therefore, in order to detect drowsy state, considering the HRV, three principal components were firstly estimated. Since the HRV metrics are divided into time, frequency, and non-linear domains, the principal components were divided considering this separation.

After this estimation, the Hotelling T^2 and SPE statistic were assessed considering each principal component, using 95% confidence level. Figure 6 presents the results achieved for the Participant 1 (Fig. 6a) and Participant 2 (Fig. 6b). With this visualization, it is possible to identify that there are five and eight points out-of-control, for Participant 1 and Participant 2, respectively.

The next step was to identify when the out-of-control points occurred, in terms of time, and visualize the video to identify if there are clear signs of drowsiness.



(a) Participant 1



(b) Participant 2

Fig. 6. Hotelling T^2 and SPE statistic.

In the case of Participant 1, the first two points occurred at the beginning of the wearable device placement, where there was more movement. On the remaining points, there is more frequent blinking and, in some cases, slightly slower blinking and deeper breathing.

In terms of Participant 2, the first point is noise since there was more movements. Moreover, it was identified as slower blinking, arms movement, difficulty keeping eyes open, more frequent blinking, yawning, head landing and closes eyes for a few seconds. The fifth and sixth out-of-control points are followed, where in the first one there are clear signs of drowsiness and since the participant closed his eyes for a few seconds, when he opened his eyes he had to control the truck. Therefore, the participant needed to be more attentive and concentrated. In the last point out-of-control, the participant seemed to be more concentrated.

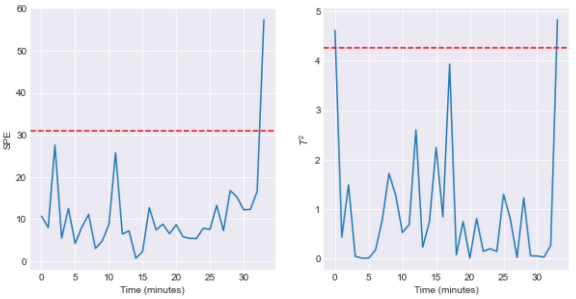
Furthermore, the same analysis was replicated for the remaining participants and it was possible to identify signs of drowsiness, in some participants. Although, in some cases, it was difficult to understand or identify the drowsy state.

4.3 Discussion of Results

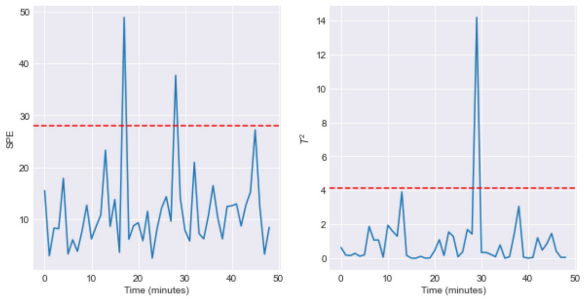
Drowsiness at the wheel has been study over the years, since it is one of the causes of road accidents. Different approaches have been developed to detect signs of drowsiness and implement machine learning algorithms.

In this study, MSPC-PCA is implemented based on the work developed by [10]. Although, in the cited work, the authors considered only one principal component with time and frequency domain variables. Hotelling T^2 and SPE statistics were assessed considering the principal component.

If the methodology implemented at [10] is considered, the results achieved are presented in Fig. 7. Only two and three points are out-of-control for Participant 1 and 2, respectively. These points were identified in our approach.



(a) Participant 1



(b) Participant 2

Fig. 7. SPE statistic considering one principal component.

In our proposed methodology, three principal components were implemented take into consideration the different HRV domains, this is, time, frequency, and non-linear. SPE and Hotelling T^2 statistics were computed for the three principal components, individually. Using this approach it was possible to identify five and eight points out-of-control, for Participant 1 and Participant 2, respectively.

The Participant 2 was the one that showed the most signs of drowsiness, to the point of falling asleep briefly. However, there were more peaks of drowsiness

that were not identified using the proposed methodology. The same analysis was made for the others participants, where our approach detects more points out-of-control. Thus, the proposed methodology achieved promising results to detect more peaks of drowsiness, although it does not yet detect all the drowsiness peaks. Besides that, some points out-of-control were due to greater movement of the arms, for example, when it is necessary to turn left or right, in the city, or at freeway exits.

5 Conclusions and Future Works

The present work propose a new approach to detect signs of drowsiness, without labelling it manually, taking into consideration the HRV. A driving simulation was conducted, for one hour, where the participants used a wearable device, and their face was recorded to help identify signs of sleepiness. Four questionnaires were answered to identify possible sleep disorders, where most of the participants are healthy.

Using the R-R intervals, time, frequency, and non-linear domain variables were extracted, every two minutes. After this, three principal components were evaluated taking into account the division of the variables. Moreover, the Hotelling T^2 and SPE statistic for each principal component, separately, were computed, considering 95% confidence level. The proposed methodology presented promising results when compared with a similar study since it identified more signs of drowsiness.

Although, this study has some limitations. Some points out-of-control occurred due to major movements and caused noise in the signal. Moreover, there are signs of sleepiness that were not identified considering the proposed approach. Another limitation is that information from the participants' awake state was not used as a reference.

As future work, a new principal component can be added with the accelerometer information to identify the points caused by noise. Besides that, it is important to consider awake information and then apply the proposed methodology. Wherefore, it is essential to perform more simulations and add more measures to detect drowsiness. Moreover, the contribution plots must be performed and assessed to identify which features contribute to the out-of-control points.

References

1. Samouco, A.I., et al.: A Ciência do Sono - Da Fisiologia à (Psico)patologia. 1 ed. (2020)
2. Kortelainen, J.M., Mendez, M.O., Bianchi, A.M., Matteucci, M., Cerutti, S.: Sleep staging based on signals acquired through bed sensor. IEEE Trans. Inf Technol. Biomed. **14**(3), 776–785 (2010)
3. Higgins, J.S., et al.: Asleep at the wheel-the road to addressing drowsy driving. Sleep **40**(2) (2017)
4. Strine, T.W., Chapman, D.P.: Associations of frequent sleep insufficiency with health-related quality of life and health behaviors. Sleep Med. **6**(1), 23–27 (2005)

5. Altevogt, B.M., Colten, H.R., et al.: Sleep disorders and sleep deprivation: an unmet public health problem (2006)
6. Gonçalves, M., et al.: Sleepiness at the wheel across Europe: a survey of 19 countries. J. Sleep Res. **24**(3), 242–253 (2015)
7. Spurnỳ, P., Andrš, J., Bouchner, P., Pučelík, J., Rokyta, R.: Testing a system for predicting microsleep. Lékař a technika-Clinician and Technology **46**(2), 51–54 (2016)
8. Braga, A.C., et al.: Multivariate statistical process control based on principal component analysis: implementation of framework in R. In: Gervasi, O., et al. (eds.) ICCSA 2018. LNCS, vol. 10961, pp. 366–381. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-95165-2_26
9. Fujiwara, K., et al.: Epileptic seizure prediction based on multivariate statistical process control of heart rate variability features. IEEE Trans. Biomed. Eng. **63**(6), 1321–1332 (2015)
10. Abe, E., Fujiwara, K., Hiraoka, T., Yamakawa, T., Kano, M.: Development of drowsiness detection method by integrating heart rate variability analysis and multivariate statistical process control. SICE J. Control, Measur. Syst. Integr. **9**(1), 10–17 (2016)
11. Åkerstedt, T., Gillberg, M.: Subjective and objective sleepiness in the active individual. Int. J. Neurosci. **52**(1–2), 29–37 (1990)
12. Åkerstedt, T., Anund, A., Axelsson, J., Kecklund, G.: Subjective sleepiness is a sensitive indicator of insufficient sleep and impaired waking function. J. Sleep Res. **23**(3), 242–254 (2014)
13. Sahayadhas, A., Sundaraj, K., Murugappan, M.: Detecting driver drowsiness based on sensors: a review (2012)
14. Liu, C.C., Hosking, S.G., Lenné, M.G.: Predicting driver drowsiness using vehicle measures: recent insights and future challenges. J. Saf. Res. **4**, 239–245 (2009)
15. Choudhary, P., Sharma, R., Singh, G., Das, S., et al.: A survey paper on drowsiness detection & alarm system for drivers. Int. Res. J. Eng. Technol. (IRJET) **3**(12), 1433–1437 (2016)
16. Fan, X., Yin, B.C., Sun, Y.F.: Yawning detection based on Gabor wavelets and LDA. Beijing Gongye Daxue Xuebao / J. Beijing Univ. Technol. (2009)
17. Yin, B.C., Fan, X., Sun, Y.F.: Multiscale dynamic features based driver fatigue detection. Int. J. Pattern Recogn. Artifi. Intell. **23**, 575–589 (2009)
18. Salles, A.F., et al.: Detecção automática de sonolência em condutores de veículos utilizando redes neurais artificiais (2018)
19. Khushaba, R.N., Kodagoda, S., Lal, S., Dissanayake, G.: Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm. IEEE Trans. Biomed. Eng. **58**, 121–131 (2011)
20. Kurt, M.B., Sezgin, N., Akin, M., Kirbas, G., Bayram, M.: The ANN-based computing of drowsy level. Expert Syst. Appl. **36**, 2534–2542 (2009)
21. Allen, J.: Photoplethysmography and its application in clinical physiological measurement. Physiol. Meas. **28**, 1–39 (2007)
22. Rundo, J.V., Downey, R.: Polysomnography. In: Handbook of Clinical Neurology, vol. 160 (2019)
23. Johns, M.W.: A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep **14**(6), 540–545 (1991)
24. Chung, F., Abdullah, H.R., Liao, P.: Stop-bang questionnaire: a practical approach to screen for obstructive sleep apnea. Chest **149**(3), 631–638 (2016)

25. Buysse, D.J., Reynolds, C.F., III., Monk, T.H., Berman, S.R., Kupfer, D.J.: The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. Psychiatry Res. **28**(2), 193–213 (1989)
26. Walker, M.: Why we sleep: unlocking the power of sleep and dreams (2017)
27. Horne, J.A., Östberg, O.: Individual differences in human circadian rhythms. Biol. Psychol. **5**(3), 179–190 (1977)
28. Ferrer, A.: Multivariate statistical process control based on principal component analysis (MSPC-PCA): some reflections and a case study in an autobody assembly process. Qual. Eng. **19**(4), 311–325 (2007)
29. MacGregor, J.F.: Using on-line process data to improve quality: challenges for statisticians. Int. Stat. Rev. **65**(3), 309–323 (1997)
30. Jackson, J.E., Mudholkar, G.S.: Control procedures for residuals associated with principal component analysis. Technometrics **21**(3), 341–349 (1979)
31. Alcala, C.F., Qin, S.J.: Analysis and generalization of fault diagnosis methods for process monitoring. J. Process Control **21**(3), 322–330 (2011)
32. Li, G., Chung, W.-Y.: A context-aware EEG headset system for early detection of driver drowsiness. Sensors **15**(8), 20873–20893 (2015)
33. Leng, L.B., Giin, L.B., Chung, W.-Y.: Wearable driver drowsiness detection system based on biomedical and motion sensors. In: 2015 IEEE SENSORS, pp. 1–4. IEEE (2015)
34. Akhter, N., Tharewal, S., Gite, H., Kale, K.: Microcontroller based RR-interval measurement using PPG signals for heart rate variability based biometric application. In: 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 588–593. IEEE (2015)
35. Shaffer, F., Ginsberg, J.P.: An overview of heart rate variability metrics and norms. Frontiers in public health **5**, 258 (2017)
36. Pinheiro, N., et al.: Can PPG be used for HRV analysis? In: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 2945–2949. IEEE (2016)
37. Champseix, R., Ribiere, L., Le Couedic, C.: A python package for heart rate variability analysis and signal preprocessing. J. Open Res. Software **9**(1) (2021)
38. Van Rossum, G., Drake, F.L.: Python 3 Reference Manual. CreateSpace, Scotts Valley (2009)
39. McKinney, W., et al.: pandas: a foundational python library for data analysis and statistics. Python High Perform. Sci. Comput. **14**(9), 1–9 (2011)
40. Ari, N., Ustazhanov, M.: Matplotlib in python. In: 2014 11th International Conference on Electronics, Computer and Computation (ICECCO), pp. 1–6. IEEE (2014)
41. pca's documentation!. https://erdogant.github.io/pca/pages/html/index.html. Accessed 28 Mar 2022
42. Harris, C.R., et al.: Array programming with Numpy. Nature **585**(7825), 357–362 (2020)
43. Virtanen, P., et al.: Scipy 1.0: fundamental algorithms for scientific computing in python. Nat. Methods **17**(3), 261–272 (2020)