Fit for Duty Assessment of Driver Fatigue based on Statistical Modelling of Cardiovascular Parameters

Ciara Pircher[1], Martin Bachler[1,2*], Christer Ahlström[3], Christopher C. Mayer[2],

Bernhard Hametner[2]

[1]Institute of Analysis and Scientific Computing, TU Wien, Wiedner Hauptstraße 8–10, 1040 Vienna, Austria
[2]Center for Health & Bioresources, AIT Austrian Institue of Technology, Giefinggasse 4, 1210 Vienna, Austria
*martin.bachler@ait.ac.at
[3]Swedisch National Road and Transport Institute (VTI), Linköping, Sweden

SNE 33(4), 2023, 157-166, DOI: 10.11128/sne.33.tn.10663
Received: 2023-10-15; Revised: 2023-11-22
Accepted: 2023-11-24
SNE - Simulation Notes Europe, ARGESIM Publisher Vienna
ISSN Print 2305-9974, Online 2306-0271, www.sne-journal.org

Abstract. Driver fatigue is a risk factor for road crashes. Fit for duty technologies could play a pivotal role in countering these crashes. Heart rate variability (HRV) and the pulse wave shape are influenced by the autonomic nervous system and are therefore affected by fatigue. This work focusses on modelling their relationship with fatigue and is based on data recorded in a simulated driving study. Six different multivariate linear regression models, using either stepwise variable selection or principal component analysis, are presented in this study. To account for differences in physiology, individual participant baselines for HRV and pulse wave parameters are introduced. Stepwise regression using any kind of baseline yields the most promising results. The most promising predictors are the ratio $\frac{\text{LF}}{\text{HF}}$ between low and high frequency components of HRV and heart rate. Finally, a stepwise regression model with a baseline, which has an adjusted R^2 statistic of 0.17, is proposed for further use. Nevertheless, further research with an extended dataset is necessary, incorporating a more diverse participant group and a higher number of recordings from severely sleepy drivers.

Introduction

Around 7% of European road crashes and around 13% crash-related injuries can be linked to driver fatigue [1]. A fit-for-duty assessment system, which can alert a driver of possible fatigue, has the potential to reduce the number of crashes, injuries and deaths on Europe's roads. Fit for duty assessments are typically based on ocular parameters [15, 16] or cognitive performance [17], and often evaluate the driver response to some sort of stimulus [18].

This work aims to develop a predictive model to estimate fatigue from cardiovascular parameters, derived from heart rate variability (HRV) and pulse wave shape. The model is meant for use in the field of commercial driving within the EU-funded PANACEA project, which stands for "practical and effective tools to monitor and assess commercial drivers' fitness to drive", and aims to take various driving impairments, such as alcohol or stress, into account.

1 Physiologial Background

1.1 Heart Rate Variability

The autonomous nervous system (ANS), which keeps the body in homoeostasis, a state of stable physical conditions, is constantly monitoring and correcting the heart rate (HR) to the needed pace via the sinus node. This means certain fluctuations in the time between two successive heartbeats are in fact healthy.

This variance indicates that our bodies can quickly adapt to environmental change or stressors and shows a degree of resilience. This fluctuation in time between successive heartbeats is termed the heart rate variability (HRV). [3]

Parameters derived from HRV can be extracted from Electrocardiography (ECG) and are usually separated in time and frequency domain parameters. An important measure for most time-domain HRV parameters is the normal-to-normal interval (NNI), which is a time series describing the time differences between successive normal heartbeats. After transforming NNI to the frequency domain, the power in certain frequency bands are widely used parameters. A description of used time and frequency domain parameters is given in table 1.

Table 1: Overview of time (1-4) and frequency (5-10) domain HRV parameters derived from ECG data. Units are given in parenthesis. [14]

	HRV Parameter	Description
1	***mean HR*** (bpm)	Mean haert rate (HR) throughout a recording
2	***SDNN*** (ms)	Standard deviation of NNIs (i.e. the square root of variance of NNIs)
3	***RMSSD*** (ms)	Root mean square of successive differences of NNIs
4	***pNN50*** (%)	Percentage of successive NNIs, that differ by more than 50 ms
5	***TP*** (ms^2)	Total power in all frequency bands
6	***LF*** (ms^2)	Power in the low frequency band (0.04 – 0.15 Hz)
7	***LFnorm*** (-)	LF power divided by absolute power of LF+HF
8	***HF*** (ms^2)	Power in the high frequency band (0.15 – 0.4 Hz)
9	***HFnorm*** (-)	HF power divided by absolute power of LF+HF
10	$\frac{LF}{HF}$ ***ratio*** (-)	Ratio of low frequency and high frequency power

As the body prepares for sleep, the heart rate decreases, allowing for more variability between beats, and the parasympathetic branch of the ANS becomes dominant while activity in the sympathetic branch of the ANS decreases [4]. Thus, we hypothesize that as fatigue arises, SDNN and HF should increase due to higher parasympathetic activity, whereas LF, heart rate and the $\frac{\mathrm{LF}}{\mathrm{HF}}$ ratio, which is said to describe the balance between the branches of the ANS, should decrease due to lower sympathetic activity [5].

Even though HRV parameters, especially $\frac{\mathrm{LF}}{\mathrm{HF}}$ ratio, are appealing parameters for fatigue assessment due to their physiological interpretation, they tend to show some controversy. There are inconsistencies in findings for all HRV parameters in connection to fatigue [19]. Concerning the frequency-domain HRV parameters, it should be noted that multiple different procedures are used to estimate the power spectrum and the applied method is often not clarified. Due to various anatomical factors, such as age or sex, there can also be large differences between individuals in HRV parameters [9, 10].

1.2 Pulse Wave

The ejection of blood from the heart causes a pressure wave that is partly reflected as it propagates through the arterial system. The pulse wave is the superposition of this pressure pulse and its reflections. Pulse arrival time (PAT) is the time from a point in the ECG, usually the prominent R-peak, to the detection of the pulse wave in a certain location of the body, in this case the finger. The pulse wave can be measured using photoplethysmography (PPG). [11]

Even though pulse waves are dependent on measurement location and individual factors, they mostly have similar main features that can be extracted as parameters. A general depiction of the pulse wave is shown in figure 1. Characteristic points of the pulse wave include the onset (P_O, the point before blood pressure begins to rise), diastolic blood pressure (P_{dia}, the minimal blood pressure), systolic blood pressure (P_{sys}, the first peak of blood pressure), the dicrotic wave amplitude (P_{dwa}, the second peak of the wave) and the dicrotic notch (P_{notch}, the through between first and second peak of the wave) [26]. The total pulse duration (TPD) t_{T} is measured as the time from wave onset to the onset of the next wave. Descriptions of the parameters derived from the pulse wave are given in table 2.

It has been shown, that sleep deprivation affects blood pressure and therefore also the pulse wave [20]. Nevertheless, only one study was found that links changes in the pulse wave shape parameters to fatigue. An evaluation of sleepiness while flying, rather than driving, showed a significant increase in PAT, systolic time and diastolic time, which is the time from wave onset P_O to the diastolic peak P_{dwa} [12].

Table 2: Overview of pulse wave parameters derived from PPG and ECG data. [25, 27]

Pulse Wave Parameter	Description
t_T	Total pulse duration (TPD): time from wave onset P_O to the onset of the next wave
t_{sys}	Time from wave onset to systolic pressure P_{sys}
t_{sys_rel}	Time from wave onset to systolic pressure P_{sys}, relative to the TPD t_T
t_{notch}	Time from wave onset to dicrotic notch P_{notch}
t_{notch_rel}	Time from wave onset to dicrotic notch P_{notch}, relative to the TPD t_T
t_{dwa_rel}	Time from wave onset to dicrotic wave amplitude P_{dwa}, relative to the TPD t_T
P_{dwa_sys}	Dicrotic wave amplitude relative to systolic blood pressure: $\frac{P_{dwa}}{P_{sys}}$
P_{notch_sys}	Amplitude of the dicrotic notch relative to systolic blood pressure: $\frac{P_{notch}}{P_{sys}}$
P_{notch_dwa}	Amplitude of the dicrotic notch relative to dicrotic wave amplitude: $\frac{P_{notch}}{P_{dwa}}$
PAT	Pulse arrival time

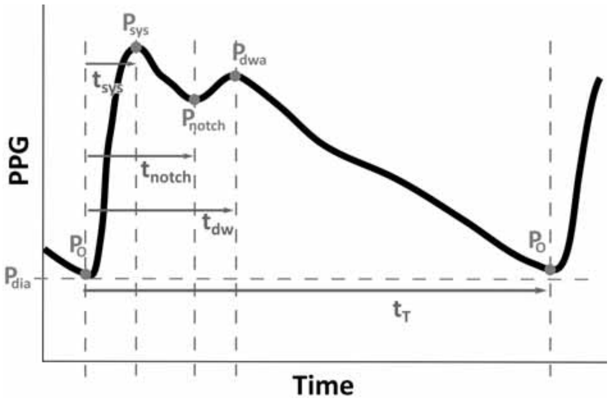

Figure 1: The image shows all characteristic points of the pulse wave (labelled in green) and all significant time durations (labelled in blue) used in this study.

2 Methods

2.1 Data Collection and Processing

The data collection was conducted by the Swedish National Road and Transport Research Institute (VTI) in Linköping, Sweden. In total, 30 male professional drivers, who did not work nights and who are free of motion sickness and sleep disorders, completed six driving simulation tasks each. While the primary focus of the pilot trial was to examine effects of social drinking in the evening (target blood alcohol content (BAC) of 0.5‰) on next-day driving performance, the secondary focus was on fatigue data and modelling.

Driving tasks took approximately 35 minutes and were completed in a driving simulator in three different conditions: a control condition (C), where participants were under no known influence, a condition for the effect of alcohol (condition A), where drivers were intoxicated for half the measurements, and measurements conducted the day after drinking (condition B). Details concerning the measurement conditions are shown in table 3. Figure 2 shows the driving simulator and examples of scenery shown during the driving tasks.

Fatigue is measured using the subjective nine-point Karolinska Sleepiness Scale (KSS), depicted in table 4.

Table 3: Summary of basic parameters for the different conditions in which each participant completed the simulated driving exercises. The blood alcohol content is abrreviated by BAC.

Condition	Purpose	Influence	Time of Day
A	Alcohol	BAC 0.3‰- 0.7‰ for half the drives	3 p.m. - 9.30 p.m
B	Day After	Residual Alcohol	7 a.m. - 1 p.m
C	Control	None	7 a.m. - 1 p.m

Figure 2: **Left:** The driving simulator used in the data collection. **Right:** Two examples of the simulated driving environment in rural (top) and urban (bottom) surroundings.

Drivers were asked to rate their level of fatigue on the KSS before and after driving. Using an objective ground truth of sleepiness would have been favourable, but available objective technologies suffer from intra- and interindividual differences and small effect sizes, whereas KSS has been found to be the measure of driver sleepiness least affected by inter-individual variations [2].

AIT Austrian Institute of Technology's proprietary device, the SmartPWA, was used to record ECG and PPG signals both before and after driving. Measurements while driving are not possible, since the device must be held with both hands for recording signals. A detailed description of the device can be found in Mengden et al. [13]. In accordance with the standards of measurement for heart rate variability [14], at least two minutes of ECG and PPG data were recorded for each measurement in this trial.

For HRV measures in the frequency domain, the series of NNIs is transformed using the Lomb-Scargle-periodogram without interpolation [22, 23]. From this periodogram, the power in the low frequency band (0.04 - 0.15 Hz), the power in the high frequency band (0.15 - 0.04 Hz) as well as the total power are derived using the integral in the respective intervals. Data processing and modelling were conducted in MATLAB R2022b (The MathWorks Inc., Natick, USA).

Table 4: Levels of the Karolinska Sleepiness Scale (KSS) [2]

Level	Description
1	Extremely alert
2	Very alert
3	Alert
4	Rather alert
5	Neither alert nor sleepy
6	Some signs of sleepiness
7	Sleepy, but no effort to keep awake
8	Sleepy, some effort to keep awake
9	Very sleepy, great effort to keep awake, fighting sleep

2.2 Modelling

In addition to the ECG and PPG parameters already described, age, height, and weight were also included as predictors in the regression models. The number of parameters is too large to sensibly include all in one predictive model, which raises the question, which parameters attribute most to accurate prediction of fatigue. MATLAB's predefined functions for dimension reduction using principal component analysis (PCA) and stepwise variable selection, respectively, were used to determine the most valuable predictors and generate multivariate linear regression models.

Since alcohol is a known confounder of HRV [21], all models using no baseline were trained on data from condition C and, due to lack of more uninfluenced data, tested on data from condition A and B.

The individual differences in HRV and pulse wave parameters used as predictors in the generated regression models can have a huge effect on the generality of these models. Therefore, two different versions of an individual baseline were pursued: a fixed baseline (F) and a dynamic baseline (D).

Fixed Baseline (F). For each individual, the first measurement taken in control condition C, before driving, is used as the baseline. The training data then consists of the differences between any other measurement and the allocated participant baseline.

However, using recordings from condition C as a baseline, does not leave enough condition C recordings to train a model. Instead, for each participant, the two measurements for a given recording time (before or after driving) and a given condition (A,B or C) are randomly divided between the test and training data set.

Dynamic Baseline (D). The measurement before driving serves as a baseline in values for each participant for this particular drive. This baseline is dynamic in the sense that for each driving simulation a new participant baseline is set. The model is trained on the differences between before and after driving for data measured in condition C. Hence, the focus lies on the change in parameters throughout a simulated drive. The model is tested on the differences of parameters between before and after driving for data measured in conditions A and B.

Modelling Approaches. All generated models use combinations of the HRV and pulse wave parameters, presented in tables 1 and 2, as well as metadata (height, weight or age) to predict the level of fatigue on the KSS.

Six different approaches, depicted in figure 3, were pursued when generating multivariate linear regression models.

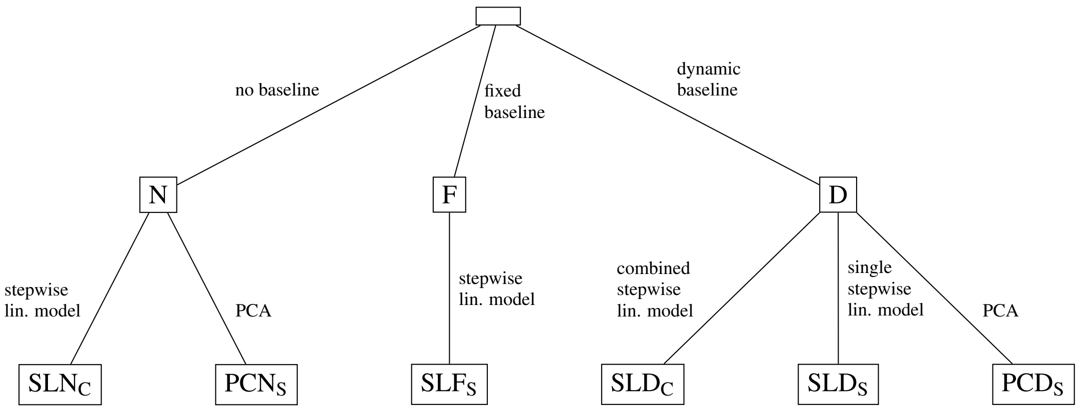

Figure 3: The tree represents an overview of different choices of baselines and modelling approaches for generating a model. A model can use no baseline (N), a fixed baseline (F) or a dynamic baseline (D). Applied methods can be principal component analysis (PC-) or stepwise linear regression (SL-). The subscript C indicates that the model is a combination of two models, where HRV and pulse wave data were modelled separately, while S indicates that one single model was generated from all data.

Each approach consists of a choice of baseline (fixed (-F), dynamic (-D) or none (-N)) and a choice of modelling method (PCA (PC-) or stepwise variable selection (SL-)). The models are given a three-letter name, where the first two indicate the chosen method while the last indicates the choice of baseline.

In some cases, HRV and pulse wave data are modelled separately, since, due to lack of signal quality, there are many missing values in pulse wave parameters. Generating separate models for the parameter groups allows the use of a larger training set for the HRV component of the model.

The final prediction for such models, which are in fact a combination of two models, is set as the average of both contributing predictions.

Models that are actually a combination of two separate models for HRV and pulse wave data are marked with a subscript C, while those generated as a single model from all data simultaneously are marked with a subscript S.

Evaluation. The models are evaluated using the F-test, which determines the statistical significance of the relationship between a group of predictors and the response. The relationship given by a model is significant, if the p-value determined by the F-test is below the level of significance $\alpha = 0.05$. Residual plots are used to detect systematic error or non-normality of errors and residuals are tested for normality using the Anderson-Darling test.

Models are also compared to each other with respect to quality of fit, using the adjusted R^2 statistic as well as root mean square error (RMSE), based on the difference between measured and predicted KSS, on test and training data as main indicators of goodness of fit.

3 Results

The median age of the drivers was 40 years with an interquartile range (IQR) of 12 years. The participants had a median height of 183 cm with an IQR of 9 cm. The median weight of drivers was 91 kg with an IQR of 21 kg.

Table 5 shows a summary of model results. The results of the F-test indicate that all but one model are statistically significant. The model PCN_S is not statistically significant. While the model SLD_S has an adjusted R^2 statistic of 0.6, all other models are below 0.25. The root mean square error of the models ranges from 0.76 to 1.26 for training data. For test data RMSE is between 1.3 and 1.75.

The tables 6 and 7 show the RMSE for data from each of the conditions separately. In general, the prediction error seems to be highest in condition B, the day after drinking, while it is lowest on data from condition C, on which most models were trained.

Graphical residual analysis using residual plots did not reveal any inappropriate model choices or correlated errors.

Table 5: The table shows key characteristics of quality of fit, the adjusted R^2 statistic and the root mean square error (RMSE) on both training and test data on the KSS scale (1-9), for each of the models discussed. Additionally, the p-value of the F-test versus a constant model is given for each of the models.

Model	Adjusted R^2	Training RMSE	Test RMSE	F-Test
SLN_C	0.21	1.11	1.55	$p_{HRV} = 0.0053$ $p_{PW} = 0.0274$
SLD_C	0.17	1.19	1.41	$p_{HRV} = 0.0035$ $p_{PW} = 0.0143$
SLD_S	0.60	0.76	1.75	$p = 0.0002$
SLF_S	0.14	1.14	1.3	$p = 0.0037$
PCN_S	0.07	1.26	1.55	$p = 0.4924$
PCD_S	0.25	1.19	1.52	$p = 0.0235$

Table 6: RMSE for KSS estimates on the scale of 1 to 9 of SLN_C, SLF_S and PCN_S models. The RMSE is given for each data category separately. The training data set of SLN_C and PCN_S is marked by an asterisk.

Data	SLN_C	SLF_S	PCN_S
C (before)*	1.04	-	1.26
C (after)	1.21	1.08	1.23
B (before)	1.68	1.35	1.49
B (after)	1.78	1.52	1.65
A (before)	1.35	1.03	1.35
A (after)	1.61	1.08	1.69

Table 7: RMSE for KSS estimates on the scale of 1 to 9 of models using dynamic baselines, i.e. SLD and PCD_S models. The RMSE is given for each data category separately. The training data set is marked by an asterisk.

Data	SLD_C	SLD_S	PCD_S
C*	1.19	0.76	1.19
B	1.52	1.48	1.68
A	1.30	1.97	1.36

While the residuals of the models SLN_C, SLD_C, SLD_S and the principal component models PCN_S and PCD_S passed the Anderson-Darling test of normality, those of the model SLF_S did not.

Table 8 gives an overview of the variables included in each of the generated stepwise linear models as well as their estimated coefficient values.

The most important variables in the generated regression models seem to be age and heart rate, which are both selected in three models (with a statistical significant relationship in two of them), as well as the $\frac{LF}{HF}$-ratio for HRV data, which is selected in all models and is significant in two of these. Concerning pulse wave parameters, systolic time and total pulse duration are chosen with a statistically significant coefficient comparatively often: in two and three models, respectively.

Table 8: This table shows an overview of the variables chosen by each stepwise regression model as well as their computed coefficients. An asterisk indicates statistical significance at the level $\alpha = 0.05$ in the corresponding model. Fields of coefficients, that were not selected, are shaded in dark grey, while those that were not statistically significant are shaded in light grey.

Variable	SLN_C	SLD_C	SLD_S	SLF_S
intercept	-1.5800	0.3155	-7.2307	4.8773
age	-0.0915^*		0.0379	-0.0378^*
height	-0.0413		0.0453	0.318
weight			-0.0298	
mean HR	-0.0688^*		-0.0966^*	-0.0736
LF	418.15		-1175.50^*	-394.46
HF			5567.10^*	
$\frac{LF}{HF}$ **ratio**	0.1858	0.2751^*	1.0477^*	-0.1314
RMSSD				
SDNN	-0.0967^*			
pNN50	-2.282		-13.4670^*	
TP	358.86			325.96^*
LFnorm				4.3528^*
HFnorm				
t_T	0.0206^*			-0.0093^*
t_{notch}				
t_{sys_rel}	65.2190^*			
t_{notch_rel}			22.762^*	
t_{dwa_rel}		8.4928		
t_{sys}	-0.0690^*	-0.0160^*	-0.0237^*	
P_{dwa_sys}	29.6110	5.3830	6.3948^*	
P_{notch_sys}	-32.1640			
P_{notch_dwa}	18.7490			
PAT	-0.0142	0.0598^*		

The principal component model with no baseline (PCN_S) uses eight of the 23 computed principal components, none of which are statistically significant.

The model PCD_S uses five principal components, two of which have a significant relationship to the response.

The model SLD_C is described in greater detail, since this is later chosen as the most favourable model. Predictions are calculated using the formula

$$\mathrm{KSS_{est}} = \frac{(\mathrm{KSS_{HRV}} + \mathrm{KSS_{PW}})}{2}$$

where the KSS values from HRV and PW data are estimated by

$$\mathrm{KSS_{HRV}} = K_{\mathrm{HRV}} + c_1 \frac{\mathrm{LF}}{\mathrm{HF}}$$
$$\mathrm{KSS_{PW}} = K_{\mathrm{PW}} + c_2 t_{\mathrm{dwp\_rel}} + c_3 t_{\mathrm{sys}} + c_4 P_{\mathrm{dwp\_sys}} + c_5 \mathrm{PAT}.$$

K_{HRV} and K_{PW} denote the estimated intercept included in the corresponding models while the coefficients are referred to as $c_j,\ j = 1,\ldots,5$. Figure 4 shows the estimated KSS values, fitted by the model SLD_C, plotted against the corresponding measured KSS values.

4 Discussion

This work aims to predict fatigue on the KSS from HRV and pulse wave parameters. Four different stepwise linear regression models, the coefficients of which can be found in table 8, and two models using principal component analysis were generated. Using linear regression resulted in fatigue predictions that were on average about 1.5 KSS units wrong. Principal component analysis and regression did not lead to improvement compared to stepwise linear regression models, especially with respect to statistical significance. This could be due to the complexity of the cardiovascular system and the multitude of confounding factors, many of which could not be included in the regression data. Due to the large individual differences in HRV and pulse wave data, using a baseline improved model results.

While many studies investigate the connection between single HRV measures and fatigue [4, 5, 8, 19], only one study was found, that evaluates this relationship for pulse wave shape parameters [12]. In these previous studies HRV and pulse wave shape parameters are used for continuous fatigue monitoring during driving. As this study attempts to assess driver sleepiness prior to driving and no previous research on predictive regression models for fatigue based on cardiovascular parameters was found, the results of this study are difficult to put into context.

The variables age, heart rate and $\frac{\mathrm{LF}}{\mathrm{HF}}$-ratio are often included as variables, while also being statistically significant in the stepwise regression models of this work. They are therefore considered to be the most important HRV parameters connected to fatigue. Similarly, when looking at the results of this study, systolic time and total pulse duration are very important pulse wave parameters for fatigue assessment.

These parameters are also considered to be important variables in previous studies, however, for the systolic time and frequency-domain HRV measures, the coefficient signs are mostly the opposite of what could have been expected from previous research [6, 8, 12]. In comparison to literature, the influence of height and weight is smaller than expected [10]. They are not chosen with a statistically significant coefficient in any model. RMSSD is not included at all and therefore seems to be of low importance. PAT is chosen with a statistically significant coefficient in one model, but seems to be less sensitive to changes in fatigue than expected.

Both, previous studies and the stepwise regression models of this study suggest that the $\frac{\mathrm{LF}}{\mathrm{HF}}$ ratio is of utmost importance for fatigue prediction, but contradictory results pose challenges in its use as a predictor. While most studies, such as [6], show a negative trend for rising sleepiness, some research, such as Rodriguez-Ibañez et al. [7], shows the opposite or, as in Abtahi et al. [8], finds no significant change. In this work the $\frac{\mathrm{LF}}{\mathrm{HF}}$ ratio is included significantly with a positive trend in two models, SLD_S and SLD_C. Contradictory results could be caused by confounding factors or, as suggested in the standards of measurement for HRV [14], could be the result of varying methods in use to obtain the frequency-domain measures. Since most studies do not clarify the applied method, the exact influence can not be determined. Alternatively, increased stress due to fighting fatigue, when driving on real roads compared to a simulated environment, could increase sympathetic activity and therefore affect HRV and limit the comparability of studies [8].

For the model SLD_S the heart rate and HF power are included as expected [8]. The models SLF_S, SLN_C and SLD_C each include one variable as the literature review would suggest (TP, HR and PAT, respectively) [8, 12].

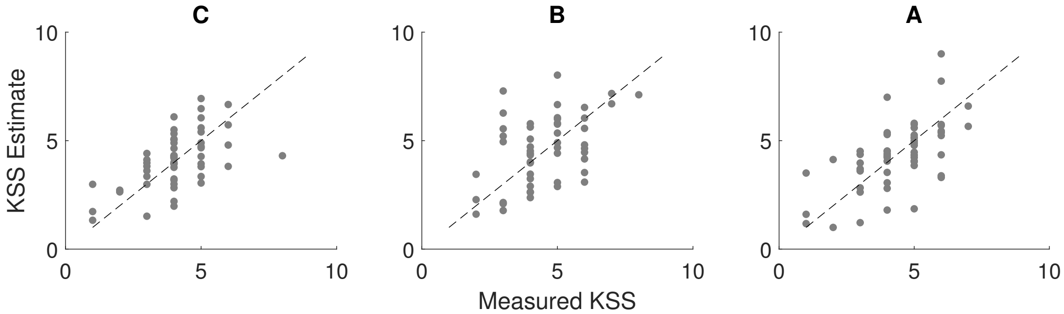

Figure 4: KSS values fitted by the model SLD_C are plotted against the corresponding measured KSS values in each group. The black dotted line indicates the line of equality between measures and estimates.

Even though all residual plots look acceptable, the residuals of the model SLF_S failed the test of normality, indicating possible systematic error. Therefore, its results should be used with caution. The model SLD_S on the other hand, shows signs of overfitting. The adjusted R^2 statistic indicates a higher portion of explained variance than could be expected in such a complex system and especially the high difference between test and training RMSE causes doubts, whether this is a suitable choice.

The model SLF_S has an acceptable R^2 value and retains a higher generality than other models presented in this study. However, the better quality of fit of SLF_S must be seen in the context of test and training data. While all other models were trained on condition C and tested on conditions A and B, both training and test data sets of the SLF_S model contained measurements from all conditions.

In the context of predictions, low p-values and RMSE are essential. Therefore, even though SLN_C has low training error and high adjusted R^2, it may not be suited for the intended use, since only 6 out of 15 variable coefficients are statistically significant. The combination of all variables is considered to be significant at the level $\alpha = 0.05$, but the p-value is higher than that achieved by other models.

The model SLD_C seems to strike a balance, where a good amount of variance is explained through a small number of variables, while statistical significance, prediction error and the normality of residuals are all acceptable. Nevertheless, the results should be interpreted with caution, since this model includes some variables, most notably the $\frac{\mathrm{LF}}{\mathrm{HF}}$ ratio and systolic time, in a different manner, i.e. opposite sign of the coefficient, than the majority of previous research.

One limitation of this work is that the data used for the purpose of generating the regression models is not perfectly suited to the task. Considering the fact that predicting high KSS values is of most interest in the context of driving, it is unfortunate, that over 90% of all recorded KSS values are below 7.

During the entire trial, no participant was tired enough to evaluate themselves at the highest KSS value of 9. This fact does not allow to generate or even test a model, that predicts fatigue accurately at the top end of the scale, thus a generally valid model.

Additionally, the training data set is rather small, after removing data influenced by alcohol. Therefore, no data without known influences can be reserved for model testing, which makes the interpretation of results difficult. Multiple factors, such as age, sex, shift work or certain medical conditions can affect HRV and pulse wave parameters and should also be accounted for [14, 24].

5 Conclusion

In the context of quantifying the relationship between fatigue and physiological parameters of the cardiovascular system, which are sensitive to changes in fatigue due to their connection to the autonomic nervous system, the presented linear regression models using stepwise variable selection produce promising results.

Even with the restriction of a small data set with avoidable confounding factors, this work shows that the prediction of fatigue on the KSS scale through a regression model using cardiovascular parameters is not only feasible in theory, but also in practice.

Of course, the methods and models presented and discussed in this work need to be refined, especially by incorporating larger and more heterogeneous data sets, before fatigue assessment for commercial drivers can be used at a large scale. The dataset included only healthy, male participants, mainly between the ages of 30 and 50. Even though the homogeneous participant group has the advantage that much variability can be avoided at such an early stage, it also means the models must be generalised and re-evaluated to be applicable to the entire adult population.

Acknowledgement

This work was carried out in the PANACEA project which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 953426. We are also very grateful to the data collection team at VTI.

Ciara Pircher was supported through a scholarship within the Talents programme of the Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology.

References

[1] Gonçalves M, Amici R, Lucas R, et al. Sleepiness at the wheel across Europe: a survey of 19 countries. *Journal of Sleep Research*. 2015; 24(3): 242–253. doi: 10.1111/jsr.12267

[2] Åkerstedt T, Anund A, Axelsson J, Kecklund G. Subjective sleepiness is a sensitive indicator of insufficient sleep and impaired waking function. *Journal of Sleep Research*. 2014; 23(3): 240-252. doi: 10.1111/jsr.12158

[3] Shaffer F, Ginsberg J P. An overview of heart rate variability metrics and norms. *Frontiers in public health*. 2017; 5(258). doi: 10.3389/fpubh.2017.00258

[4] Persson A, Jonasson H, Fredriksson I, Wiklund U, Ahlström C. Heart Rate Variability for Classification of Alert Versus Sleep Deprived Drivers in Real Road Driving Conditions. *IEEE Transactions on Intelligent Transportation Systems*. 2021; 22(6): 3316–3325. doi: 10.1109/TITS.2020.2981941

[5] Burlacu A, Brinza C, Brezulianu A, Covic A. Accurate and early detection of sleepiness, fatigue and stress levels in drivers through Heart Rate Variability parameters: a systematic review. *Reviews in cardiovascular medicine*. 2021; 22(3): 845–852. doi: 10.31083/j.rcm2203090

[6] Awais M, Badruddin N, Drieberg M. A non-invasive approach to detect drowsiness in a monotonous driving environment. *TENCON 2014-2014 IEEE Region 10 Conference*. 2014. doi: 10.1109/TENCONSpring.2014.6863035

[7] Rodriguez-Ibañez N, García-Gonzalez M A, de la Cruz M A F, Fernández-Chimeno M, Ramos-Castro J. Changes in heart rate variability indexes due to drowsiness in professional drivers measured in a real environment. *2012 Computing in Cardiology*. 2012; 913–916.

[8] Abtahi F, Anund A, Fors C, Seoane F, Lindecrantz K. Association of drivers' sleepiness with heart rate variability: A pilot study with drivers on real roads. *EMBEC & NBC 2017*. 2017; IFMBE Proceedings, vol 65. doi: 10.1007/978-981-10-5122-7_38

[9] Lutfi M F, Sukkar M Y. The effect of gender on heart rate variability in asthmatic and normal healthy adults. *International journal of health sciences*. 2011; 5(2): 146–154.

[10] Lutfi M F, Sukkar M Y. Relationship of height, weight and body mass index to heart rate variability. *Sudan Med J*. 2011; 47(1): 14–19.

[11] Tamura T, Chen W. Blood Pressure. *Seamless Healthcare Monitoring*. Springer; 2018. p 103–127.

[12] Majumder S, Verma A K, Wang C et al. Using photoplethysmography based features as indicators of drowsiness: Preliminary results. *2019 Design of Medical Devices Conference*. 2019. American Society of Mechanical Engineers. doi: 10.1115/DMD2019-3236

[13] Mengden T, Bachler M, Sehnert W, Marschall P, Wassertheurer S. Device-guided slow breathing with direct biofeedback of pulse wave velocity–acute effects on pulse arrival time and self-measured blood pressure. *Blood Pressure Monitoring*. 2023; 28(1): 52–58. doi: 10.1097/MBP.0000000000000628

[14] Malik M. Heart rate variability: Standards of measurement, physiological interpretation, and clinical use: Task force of the European Society of Cardiology and the North American Society for Pacing and Electrophysiology. *Annals of Noninvasive Electrocardiology*. 1996; 1(2): 151-181. doi: 10.1111/j.1542-474X.1996.tb00275.x

[15] Ahlström C, Nyström M, Holmqvist K, Fors C, Anund A, Kecklund G, Åkerstedt T. Fit-for-duty test for estimation of drivers' sleepiness level: Eye movements improve the sleep/wake predictor. *Transportation Research Part C: Emerging Technologies*. 2013; 26: 20–32. doi: 10.1016/j.trc.2012.07.008

[16] Di Stasi L, Renner R, Catena A, Cañas J, Velichkovsy M, Pannasch S. Towards a driver fatigue test based on the saccadic main sequence: A partial validation by subjective report data. *Transportation Research Part C: Emerging Technologies*. 2012; 21(1): 122–133. doi: 10.1016/j.trc.2011.07.002

[17] Perelli L. CogSpeed: A Proposed Technology for Instantly Measuring Cognitive Performance in a Laboratory or a Real World Setting. 2023.

[18] Cori J, Manousakis J, Koppel S, Ferguson S, Sargent C, Howard M, Anderson C. An evaluation and comparison of commercial driver sleepiness detection technology: a rapid review. *Physiological Measurement*. 2021; 42(7). doi: 10.1088/1361-6579/abfbb8

[19] Lu K, Dahlman A, Karlsson J, Candefjord S. Detecting driver fatigue using heart rate variability: A systematic review. *Accident Analysis & Prevention*. 2022; 178. doi: 10.1016/j.aap.2022.106830

[20] Kato M, Phillips B, Sigurdsson G, Narkiewicz K, Pesek C, Somers V. Effects of Sleep Deprivation on Neural Circulatory Control. *Hypertension*. 2000; 35(5): 1173-1175. doi: 10.1161/01.HYP.35.5.1173

[21] Romanowicz M, Schmidt J, Bostwick J, Mrazek D, Karpyak V. Changes in heart rate variability associated with acute alcohol consumption: current knowledge and implications for practice and research. *Alcoholism: Clinical and Experimental Research*. 2011; 35(6): 1092-1105. doi: 10.1111/j.1530-0277.2011.01442.x

[22] Bachler M, Sehnert W, Mikisek I, Wassertheurer S, Mengden T. Non-invasive quantification of the effect of device-guided slow breathing with direct feedback to the patient to reduce blood pressure. *Physiological Measurement*. 2020; 41(10): 1092-1105. doi: 10.1088/1361-6579/abb320

[23] Fonseca D, Netto A, Ferreira R, De Sa A. Lomb-scargle periodogram applied to heart rate variability approach. *ISSNP Biosignals and Biorobotics Conference: Biosignals and Robotics for Better and Safer Living (BRC)*. 2013.

[24] Skornyakov E, Gaddameedhi S, Paech G, Sparrow A, Satterfield B, Shattuck N, Van Dongen H. Cardiac autonomic activity during simulated shift work. *Industrial health*, 57(1): 118-132. 2019.

[25] Korpas D, Halek J, Doležal L. Parameters describing the pulse wave. *Physiological Research*, 58(4): 473-479. 2009.

[26] Solà J, Delgado-Gonzalo R. *The Handbook of Cuffless Blood Pressure Monitoring*. Springer; 2019.

[27] Tamura T, Chen W. *Seamless Healthcare Monitoring*. Springer; 2018.