



This is a postprint version of the following published document:

Muñoz-Organero, M.; Corcoba-Magaña, V. (2017). Predicting upcoming values of stress while driving. *IEEE Transactions on Intelligent Transportation Systems*, 18(7), 1802-1811.
Doi: <https://doi.org/10.1109/TITS.2016.2618424>

© 2016. IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Predicting Upcoming Values of Stress While Driving

Mario Muñoz-Organero and Victor Corcoba-Magaña

Abstract—The levels of stress while driving affect the way we drive and have an impact on the likelihood of having an accident. Different types of sensors, such as heart rate or skin conductivity sensors, have been previously used to measure stress related features. Estimated stress levels could be used to adapt the driver’s environment to minimize distractions in high cognitive demanding situations and to promote stress-friendly driving behaviors. The way we drive has an impact on how stressors affect the perceived cognitive demands by drivers, and at the same time, the perceived stress has an impact on the actions taken by the driver. In this paper, we evaluate how effectively upcoming stress levels can be predicted considering current stress levels, current driving behavior, and the shape of the road. We use features, such as the positive kinetic energy and severity of curves on the road to estimate how stress levels will evolve in the next minute. Different machine learning techniques are evaluated and the results for both intra and inter-city driving and for both intra and inter driver data are presented. We have used data from four different drivers with three different car models and a motorbike and more than 220 test drives. Results show that upcoming stress levels can be accurately predicted for a single user (correlation $r = 0.99$ and classification accuracy 97.5%) but prediction for different users is more limited (correlation $r = 0.92$ and classification accuracy 46.9%).

Index Terms—Stress level prediction, stress-friendly driving behavior, stress level classification, machine learning.

I. INTRODUCTION

HUMAN factors are responsible for many of the traffic accidents on the road. The data presented in [1] categorizes the major risk factors responsible for traffic accidents according to their impact as: human factors (92%), vehicle factors (2.6%), road/environmental factors (2.6%), and others (2.8%). Among these, drivers’ human factors consist of cognitive errors (40.6%), judgment errors (34.1%), execution errors (10.3%), and others (15%). Cognitive errors appear in highly cognitive demanding situations in which the cognitive load as perceived by the driver is high and the actions taken by the driver to handle those situations are in many occasions not appropriate. Being able to detect the evolution of the

driver’s cognitive load and stress levels and to predict highly demanding situations is crucial in order to provide help to the driver to better handle these situations. In this paper, the impact that current stress levels, driving behavior and environmental road conditions have on the prediction of upcoming stress levels is analyzed and major results presented.

There are many proposals on measuring and quantifying the current driver’s cognitive load and stress levels in previous research studies in literature. The paper in [2] described a queuing network modeling approach to model the subjective mental workload and the multitask performance. Using this model, the interface of driving assistants could be automatically adapted according to the workload. In [3], Itoh et al. measured electrocardiogram (ECG) signals as well as head rotational angles, pupil diameters, and eye blinking with a faceLAB device installed in a driving simulator to calculate the driving workload. In the study captured in [4], the driver’s workload was estimated from lane changing. In [5], the authors proposed a multiple linear regression equation to estimate the driving workload. The model employs variables such as: speed, steering angle, turn signal, and acceleration.

On the other hand, the impact of the cognitive load on the driver behavior has been studied in many research papers. In [6], Kim et al. analyzed the relationship between drivers’ distraction and the cognitive load. It was discovered that heart rate, skin conductance, and left-pupil size were effective measurement variables for observing drivers’ distractions. The work described in [7] showed that the visual demand caused a reduction in the speed and increased the variation in lane maintenance. The authors highlighted that the detection of events is very important in order to capture the main safety related effects of cognitive load and visual tasks. In [8], the authors propose to use a set of variables (vehicle speed, steering angle, acceleration, and gaze information) to estimate the current workload of the driver. The authors achieved an accuracy of 81% with this method. The authors concluded that the driver’s workload required to perform primary actions (driving) is influenced with the addition of secondary verbal and special mental activities that the drivers have to perform while driving using a simulator. Other studies such as [9] also propose to use the movement of the steering wheel as an indicator of the perceived workload by the driver.

In conclusion, although there exist significant research works in previous literature on understanding the impact of different cognitive load and stress levels both in physiological signals (such as heart rate of skin conductance) and in driving actions (such as moving the steering wheel or changing speed) and how to use data from physiological and vehicle telemetry

sensors to estimate cognitive load and stress levels, a better understanding about how current sensed information could be used to predict upcoming stressful and highly cognitive demanding situations is needed. Although simulators have some positive features such as providing a controlled and replicable environment for tests, their differences with real world driving are important in how the driver perceives danger and the impact that different stressors have on the drivers' actions. Therefore, this paper will study correlations between current levels of stress, cognitive load and current driver's driving behavior with upcoming levels of cognitive load and stress in order to estimate and predict them in real driving scenarios. The stress levels will be assessed by computing time and frequency features from the Heart Rate Variability (HRV) signal measured as a proxy physiological signal.

II. PHYSIOLOGICAL SENSORS AND STRESS LEVELS

Our goal is to predict upcoming levels of stress based on current levels of stress, driving actions and road conditions. A method to measure stress levels is therefore required. Among the different alternatives to measure current stress levels based on physiological sensors, we will use both time and frequency computations based on the Heart Rate Variability signal from the driver while driving. This section describes an introduction to the major physiological sensors that have been previously described in related research studies and their merits, paving the way to justify the method used in this paper.

Several physiological sensors have been proposed in literature able to provide proxy variables for quantifying the driver's cognitive load and stress levels. These sensors include heart-rate, heart-rate variability, pupil-dilation, blood-pressure, respiration rate and GSR (Galvanic Skin Response) [10]. Some references for estimating the cognitive load from sensed data include Heart Period [11], Pupillary Response [12] and Heart Rate Variability [13]. Studies about correlations between heart rate, respiration rate and skin conductance with the amount of encountered stressors while driving in real scenarios can be found in [14]–[16]. Authors in [17] used a wrist device including an accelerometer and a skin conductance sensor to find correlations with stress levels. Using techniques such as Support Vector Machines (SVM), k-nearest neighbors (KNN) and Principal Component Analysis (PCA) and adding additional information coming from the user's mobile device, a 75% accuracy is obtained. The authors in [18] also estimate the level of stress based on physiological sensors based on the measured heart rate (HR) and galvanic skin response (GSR) combined with other sensors such as electroencephalography (EEG) and electromyography (EMG). In order to deal with inter-person differences the authors use a clustering algorithm to first divide the users into similar clusters using the k-means algorithm and then to perform cluster-wise stress evaluation using the general regression neural network taking into account the variations in individual stress response. The authors in [19] also combine physiological wearable sensors (heart rate and skin conductance) with data coming from mobile phone usage and surveys to find correlations with stress levels.

In the particular area that focuses on the impact that stress and perceived cognitive load have on driving, among the

different sensors and physiological signals, the authors in [14] found out that skin conductivity and heart rate based metrics were the most closely correlated with stress levels. We will use heart rate variability (HRV) based metrics both in the frequency and time domains as the proxy features to estimate the driver's stress. Generally, both frequency and time based parameters extracted from the heart rate variability (HRV) analysis have been used as stress related factors. High frequency power (HF; 0.15-0.4 Hz) relates to parasympathetic nerve activity and low frequency power (LF; 0.04-0.15 Hz) is modulated by both the sympathetic and parasympathetic nervous systems. Thus, the LF/HF ratio reflects the global sympatho-vagal balance and could be used as a measure of this balance [20]. Time domain parameters include the standard deviation of NN intervals (SDNN) or time intervals between consecutive heartbeats normally measured as R to R times, the square root of the mean of the squares of the successive differences between adjacent heartbeats (RMSSD) and the proportion of consecutive heartbeats differing more than 50ms divided by total number of measured heartbeats (pNN50) [20]. To be able to select among the different features derived from the HRV signal, the authors in [26] assessed the validity and reliability of telemetry-derived HRV responses to an orthostatic challenge. They found the RMSSD and HF can be used to provide a sensitive, valid and reliable assessment of autonomic control of heart rate while SDNN did not significantly respond to the orthostatic challenge.

Many laboratory studies have demonstrated changes in cardiac autonomic control during psychological stress. The authors in [21] attempted to demonstrate this effect in ambulatory subjects finding that psychological stress was significantly associated with an increase in the LF/HF ratio, suggesting increases in the relative predominance of sympathetic nervous system activity during stressful periods of the day. Vagal modulation of heart period appears to be sensitive to the recent experience of persistent emotional stress, regardless of a person's level of physical fitness and disposition toward experiencing anxiety [22]. A 5 minute interval has been normally used as the shortest time period over which HRV metrics should be assessed [23]. However, there are studies that try to assess how 5 minute windows correlate with shorter period timeframes. The authors in [14] proposed that 1 minute windows could be used. If not considering the very low frequency (VLF) part of the spectrum (0.014-0.05) the lowest frequency to be measured in the LF range is 0.05 Hz which provides a 20 second period bound to the time window to be used.

Several previous research papers have studied the correlations between HRV related features and perceived stress while driving. The authors in [14] correlated several heart rate based metrics with stress levels (the stress levels are estimated by counting the number of stressors that the driver has to face and the estimated time response of the human being to these stressors). Each stressor contributes to increase the cognitive load in the driver in order to be able to deal with it. The correlations were found to be valid both for highway as well as for city driving. The research work presented in [15] uses the heart rate variability (HRV) as a proxy

variable to get stress level measurements as the response by the autonomic nervous system. The authors extract different features from the HRV signal and use them as the input of several classification algorithms in order to assess whether it is possible to classify stressful situations from HRV features. They use the same database as [14] comprising 16 drivers and are able to classify stressful situations with an 83% accuracy using a SVM-RBF classifier. The authors in [16] reported on the autonomic nervous system changes and driving style modifications as a response to incrementally stressing situations during 3 simulated driving tests with increasing stress loads. They measured the impact on physiological signals and some vehicle's mechanical parameters (steering wheel angle corrections, velocity changes, and time to reposes). They also used heart rate measurements as the base physiological signal. Only driving simulations were used (lacking the richness of features found in real driving environments). The research conducted in [24] presents a heart rate variability (HRV) parameter-based feature transformation algorithm for driving stress recognition. The proposed parameter-based transformation algorithm consists of feature generation, feature selection, and feature dimension reduction. A parameter-based feature generation method from five-minute HRV analysis is proposed in this study. The kernel-based class separability (KBCS) is employed as the selection criterion for feature selection. Principal component analysis (PCA) and linear discriminant analysis (LDA) are adopted for feature dimension reduction. The combination of KBCS, LDA, and PCA achieved recognition rates around 81%. The authors in [25] combined the HRV signal with other physiological signals (EDA and respiration) for detecting drivers' stress and fatigue. Dividing the data samples into 2 degrees of stress they were able to classify 86% of the samples correctly.

Each physiological sensor provides only a proxy variable to estimate the cognitive load and stress levels and could also be affected by other external and environmental variables. Physical exercise and tiredness are just two examples of external factors that also have an influence on the HRV proxy variables. At the same time, different types of stressors may have a concurrent impact on the user's stress levels. Our design for the data gathering process has been conceived in order to minimize the impact of factors which are not under study. We have only taken data from drivers driving alone to work and back home in similar situations each day (same hour, same traffic conditions, with moderated previous walking to get into the car and a relaxation period before driving, with the mobile phone muted, with the radio switched off and without using any navigation system). We have used the heart rate variability as the proxy variable to estimate the driver's cognitive load and stress levels. We have analyzed the impact that current levels of stress and recent driving actions such as accelerations and deceleration and environmental factors such as the curvilinear shape of the road have on upcoming levels of stress.

To the best of our knowledge, previous studies have mainly focused on current stress estimation and monitoring. Our major contribution goes a step forward and analyzes how much information is there in the current stress level, driving behavior and environmental road conditions in order to predict

how stress levels will evolve in the near future both for intra and intercity environments and for intra and inter user generalization. Predicting how stress levels associated to the driving task will evolve over the next minute of time based on the current observed information could be the basis for stress-aware driving support and recommender tools.

III. THE IMPLEMENTED SCENARIO

In this section we present the objectives of the research conducted and the details of the method used to implement the experiment to achieve our objectives.

A. Objective

Our objective is to find the accuracy in predicting upcoming stress levels while driving based on the current measurements of stress levels, driving behavior and road shape and type.

B. Input Variables

In order to predict upcoming stress levels based on current conditions, 3 different types of variables have been used: stress related, driving behavior related and road shape and type related.

1) *Stress Related*: We have computed both frequency and time domain features from the Heart Rate Variability (HRV) that have been assessed as having a higher correlation with stress levels [14]. In particular, we have computed the LF/HF ratio, considering low-frequency (LF; 0.05–0.15 Hz) and high-frequency (HF; 0.15–0.5 Hz), and the RMSSD values as proposed in [26] to have the higher reliability.

A 5 minute window is used to estimate the current levels of stress based on the LF/HF and RMSSD features of the HRV signal over the last 5 minutes of driving. The 5 minute window continuously moves with the driving each second (as has been proposed in [14]). The average of the 5 minute estimated stress calculated over the next minute from the actual driving point is used as the target for estimating how stress levels will evolve in the near future. We want to assess how much information is already contained in the current observable features about this upcoming stress.

2) *Driving Behavior Related*: The driver's behavior has been captured from the vehicle's telemetry and GPS data. Driving actions imply direct control over the pedals and the steering wheel, being the steering wheel movements mainly related to the shape of the road as captured in the environment related set of features. In order to predict upcoming stress levels we have concurrently used 2 time windows. First, we calculate driving behavior related features on the same 5 minute window used to compute the stress related features in order to take into account the driving information that caused an influence on the assessed current stress. Second, in order to estimate how stress will evolve in the next minute, the more recent driving data from the past minute has also been used since more recent driving related data will correlate with recently encountered stressors which will influence on how stress will evolve in the near future.

Taking both 1 and 5 minute windows into account, the following variables have been selected in our experiment to

measure the behavior of the driver: average driving speed over the last 5 minutes, average driving speed over the last minute, average acceleration over the last 5 and 1 minutes and positive Kinetic Energy (PKE) on a 5 and 1 minute window.

3) *Road Shape and Type Related*: Two features have been included in the study to take into account the shape and type of the road. The curvilinear shape of the road has been introduced by averaging the angle between the current speed and next speed vectors over a 5 minute and a 1 minute windows. Since speed vectors are computed from GPS data, in order to minimize errors when moving at low speeds, the speed vectors are taken considering 5 second averages when travelling over a threshold speed of 1m/s. The type of the road has been considered by dividing intra-city and inter-city (rural) driving samples.

C. Additional Measuring Considerations

One major limitation of the HRV signal in order to estimate the level of stress and cognitive load is that there are other factors such as the physical exercise that also impact the measured values. As described in the previous section, the experiment has been designed to minimize the impact that factors outside the study have on the measurements. In this way, only data from drivers driving alone to work and back home in similar situations each day (same hour, same traffic conditions, with moderated previous walking to get into the car and a relaxation period of at least 30 seconds before driving, with the mobile phone muted, with the radio switched off and without using any navigation system) have been taken.

D. Method

In order to validate results taking into account different drivers in different driving environments, 3 different users using 3 different cars in 3 different regions have been selected. The regions were Sheffield in the UK and Madrid and Seville in Spain. The vehicle models were an Opel Zafir Tourer, a Citroen Xsara Picasso and a Citroen C5. In total, we have obtained 200 test drives with around 5000 minutes of driving. Each test drive comprised both urban and inter-urban (rural and highway) sections. In order to validate the generalization of results, a fourth driver using a motorbike in 22 drives in an urban route (Madrid) has been added.

A Polar H7 band was used to record the HRV signal. The band was paired with a Nexus 6 Android Mobile device running an application implemented for the experiment which recorded the HRV together with GPS data and telemetry data such as the driving speed.

We took samples for the input variables each second. A pre-processing technique is used in order to detect outliers due to errors in the sensor device. The body movements may affect the measurements. Each measurement error results in very low followed by a very high (or vice-versa) pairs of values which are replaced by the average value.

The current acceleration of the vehicle is calculated based on the measured speed as follows:

$$a_i = \frac{v_i - v_{i-1}}{t_i - t_{i-1}} \quad (1)$$

In which v_i represents the speed at the sample number i , a_i the estimated acceleration at that sample and the derivative of the speed is estimated by dividing the increment in speed by the time elapsed between the consecutive samples $i-1$ and i .

The PKE is estimated over a period of time as follows:

$$PKE = \frac{\sum (v_i - v_{i-1})^2}{d}; v_i > v_{i-1} \quad (2)$$

Where the sum is performed for the period considered and d is the cumulated distance traveled during this time.

The intensity of turning is estimated using the formula:

$$TII_i = \cos^{-1} \frac{\bar{l}_i \bullet \bar{l}_{i-1}}{\|\bar{l}_i\| \|\bar{l}_{i-1}\|}; v_i > th \quad (3)$$

Where the numerator represents the dot product between the average direction vectors in the last 5 seconds and the average direction vectors in the next 5 seconds and the denominator captures the norm of such averaged vectors. The direction vectors are calculated from the GPS coordinates. The average over a period of 5 seconds is used to minimize the impact of random errors in the GPS signal. In order to eliminate the errors introduced at low speeds, a threshold in the speed is used. This threshold has been empirically evaluated and a value of 1 m/s has been found to perform well and therefore selected for the experiment.

The target of the study is to assess if upcoming stress levels can be predicted from past and current measured and computed data. We have used the time average over the following driving minute for two features calculated from the last 5 minute sliding window from the HRV signal: the ratio LF/HF and the RMSSD. We have evaluated several machine learning tools in 2 scenarios:

- 1) *Use the Average Stress Values for the Next Minute and Regression Techniques*
- 2) *Cluster the average stress values for the next minute in 3 groups (low, medium and high stress) and use classification techniques*

Correlation indexes for the estimated upcoming stress signal and the real one are calculated for case 1). Confusion matrixes are calculated for case 2). In all cases, 4 different training and validation scenarios have been considered to assess the validity of results and their generalization validity:

- 1) *Training the system with part of the data for a particular user and validating it with the rest of the data for the same user (10-fold cross-validation and leave on journey out)*
- 2) *The same as 1) but differentiating the case of inter and intra city segments.*
- 3) *Training the system with all the data for a particular user and validating it with all the data for a different user*
- 4) *The same as 3) but differentiating the case of inter and intra city segments.*

IV. RESULTS

This section captures some of the results from the conducted experiments. The results are presented in 4 sub-sections.

TABLE I
CORRELATION COEFFICIENTS WITH UPCOMING LH/HF VALUES

Variable	correlation coefficient
acc (5 min)	0.058610547
curvilinear (5 min)	-0.057822176
PKE (5 min)	0.109690541
v (5 min)	0.124882195
acc (1 min)	-0.135669081
curvilinear (1 min)	-0.154139774
PKE (1 min)	-0.200308376
v (1 min)	0.18470252
LF/HF (current)	0.838362352
LF/HF next	1

The first 3 sub-sections use data from the first 3 drivers using 3 different models of cars. The last sub-section is based on data collected from the fourth driver using a motorbike. The first sub-section is dedicated to show how accurately upcoming levels of stress can be assessed based on recent levels of stress, driving behavior and road shape for a single user. The second sub-section captures the same results when training the algorithms for one particular user and validating them for a different one. The third subsection is dedicated to show how well current levels of stress can be assessed only by measuring the driving behavior and road shape (without taking the HRV signal into account). Finally, the fourth sub-section is dedicated to validate the generalization of results to a different vehicle type. Conclusions and discussions by comparing our approach with previous related results found in literature are presented in section V.

A. Single User Scenario

In this first case, part of the data of a single driver is used to train the system and the rest of the data to validate it. A leave one journey out and a 10-fold cross validation techniques have been used and compared. Both the LF/HF ratio and the RMSSD values have been studied.

Let's present first the LF/HF ratio related results. Before trying to predict upcoming levels of stress based on the LF/HF ratio, Table I captures the correlation indexes between the upcoming stress levels and the independent variables that we will use for its prediction. Table I shows that upcoming levels of stress are more correlated with current levels of stress (as measured over the last 5 minutes window). Temporal averages over the last minute for the driving behavior and road shape have a higher correlation with upcoming stress than averages over the 5 minute window used for stress estimation. This shows that recent driver behavior has a higher impact in determining how the user's perceived stress will evolve in the near future. The PKE averaged over the last minute is the variable most correlated with upcoming levels of stress.

Table II captures the correlation coefficients between the real signal for the upcoming stress based on the average of assessed values over a 5 minute sliding window of the LF/HF ratio and the predicted signal based on 3 different regression models: multilayer perceptron (MLP), linear regression and Gaussian Processes (GP). Table II captures the results for a single user over all the entire journeys (considering both

TABLE II
LF/HF CORRELATION COEFFICIENTS
(CITY + RURAL ROADS)

	Leave one out	10-fold crossval.
MLP	0.915	0.8686
Linear regression	0.9147	0.8906
GP	0.9113	0.9632

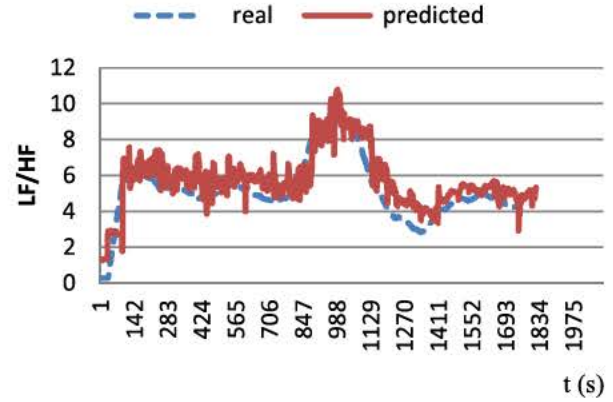


Fig. 1. Predicted vs real values for a particular test drive using linear regression.

TABLE III
LF/HF CORRELATION COEFFICIENTS (RURAL ROADS)

	Leave one out	10-fold crossval.
MLP	0.8843	0.8804
Linear regression	0.8862	0.8882
GP	0.7102	0.9854

TABLE IV
LF/HF CORRELATION COEFFICIENTS (CITY)

	Leave one out	10-fold crossval.
MLP	0.7195	0.9499
Linear regression	0.8335	0.915
GP	0.8842	0.9777

rural and city roads) for 2 different validation strategies: leave one journey out and 10-fold cross-validation. MLP and Linear regression provide similar results in both cases. GP are able to achieve a correlation of 0.96 if a 10-fold cross-validation is used. Figure 1 shows in a graphical way the real and the predicted signals for one particular journey left out in the training phase using linear regression.

Tables III and IV present the same type of results but restricting the data used both for training and validation to rural and city roads. There is a slight improvement in both cases for the case of using the 10-fold cross-validation technique but the results degrade if using a leave one journey out validation approach. Based on these results, we can say that it is better to use all the data from previous journeys for training instead of filtering it by road type in order to have better predictions for new journeys.

Table V captures the correlation coefficients for the case of the RMSSD signal. Upcoming averages for 1 minute segments

TABLE V
CORRELATION COEFFICIENTS WITH UPCOMING RMSSD VALUES

Variable	correlation coefficient
acc (5 min)	-0.26605309
curvilinear (5 min)	-0.39018718
PKE (5 min)	-0.23994482
v (5 min)	-0.44591002
acc (1 min)	-0.18700188
curvilinear (1 min)	-0.41883065
PKE (1 min)	-0.14808566
v (1 min)	-0.4356931
RMSSD (current)	0.64070573
RMSSD next	1

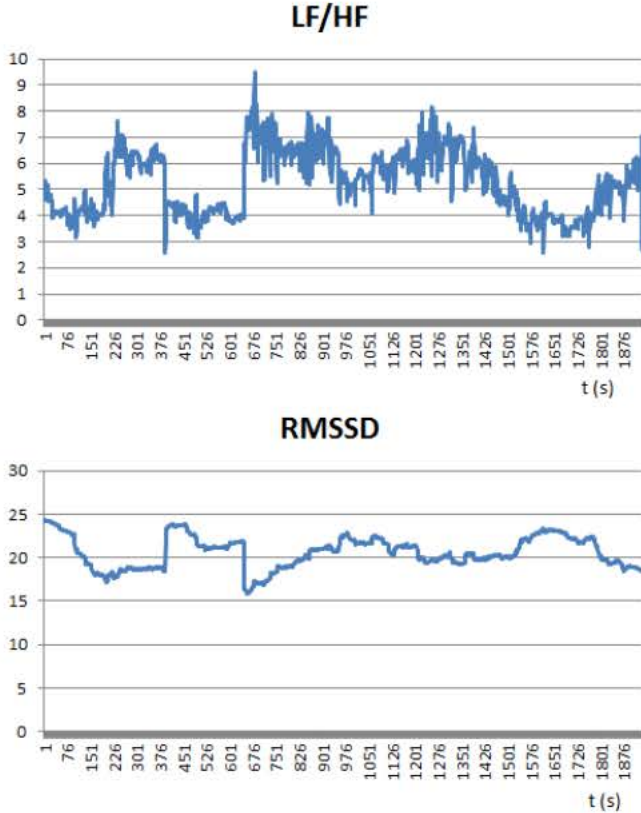


Fig. 2. RMSSD vs LF/HF graphical comparison.

of the RMSSD signal are less correlated with current values for the last 5 minutes window than the LF/HF ratio with a correlation coefficient of 0.64. Average acceleration values and road shape also show a correlation coefficient higher than 0.4 (in this case negative since more accelerations and curvilinear shapes negatively impact the RMSSD signal). In this case, the 5 minute averages and the last 1 minute averages for the driving behavior and road shape based features show similar correlation coefficients with upcoming RMSSD values.

Figure 2 shows in a graphical way the LF/HF and RMSSD signals for a particular journey. The figure shows an inverted relationship (when the LF/HF ratio grows the RMSSD values decrease and vice versa). The LF/HF ratio provides a direct estimation of perceived stress levels while the RMSSD shows a negative correlation with perceived stress (the bigger the stress level the smaller the variability of inter-beat times).

TABLE VI
RMSSD CORRELATION COEFFICIENTS (CITY + RURAL)

	Leave one out	10-fold crossval.
MLP	0.5794	0.9255
Linear regression	0.8886	0.881
GP	0.9944	0.9931

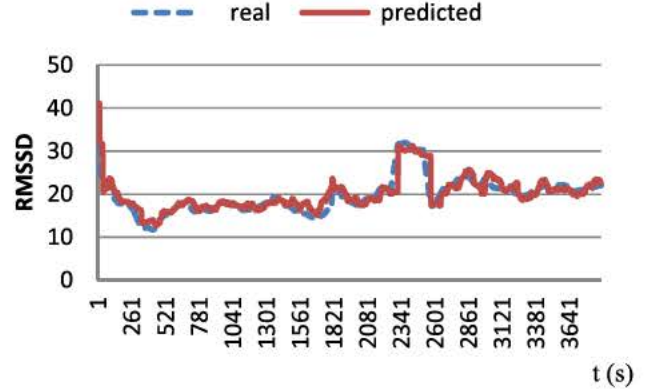


Fig. 3. Predicted vs real values for a particular test drive using linear regression.

TABLE VII
RMSSD CORRELATION COEFFICIENTS (RURAL ROADS)

	Leave one out	10-fold crossval.
MLP	-0.3471	0.9423
Linear regression	0.8109	0.8143
GP	0.2996	0.9683

TABLE VIII
RMSSD CORRELATION COEFFICIENTS (CITY)

	Leave one out	10-fold crossval.
MLP	0.1406	0.8072
Linear regression	0.4789	0.8847
GP	-0.5925	0.9683

Table VI captures the correlations coefficients between the real RMSSD upcoming 1 minute average and the estimated one for the same regression techniques and using both a leave one out and a 10-fold cross-validation validation approaches. The GP outperforms in this case both MLP and linear regression improving the value for the correlation coefficient. Figure 3 presents the real and the predicted values for a particular journey following the leave one out validation approach and the linear regression algorithm (to be consistent with the results in figure 1).

Tables VII and VIII capture the results for the correlation coefficients for the RMSSD predicted signal when restricting both training and validation data to inter and intra city road segments. In this case, we see a similar effect as the one we presented for the LF/HF case in terms of degradation of performance for the leave one journey out validation approach. Although the 10-fold cross-validation approach indicates that good prediction results can be achieved for new segments in the same journey, the leave one journey out

TABLE IX
PREDICTED VALUE FOR UPCOMING LF/HF USING
SVM AND 10-FOLD VALIDATION

Actual/Predicted	Low	Medium	High
Low	0.94520548	0.05479452	0
Medium	0.00801806	0.98816752	0.00381442
High	0	0.06252299	0.93747701

TABLE X
PREDICTED VALUE FOR UPCOMING LF/HF USING
NAÏVE BAYES AND 10-FOLD VALIDATION

Actual/Predicted	Low	Medium	High
Low	0.63387298	0.34806974	0.01805729
Medium	0.07877939	0.809357	0.11186362
High	0.00073556	0.22912836	0.77013608

TABLE XI
PREDICTED VALUE FOR UPCOMING LF/HF USING J.48 AND 10-FOLD
VALIDATION

Actual/Predicted	Low	Medium	High
Low	0.95765878	0.04171856	0.00062267
Medium	0.00513779	0.98583217	0.00903005
High	0.00073556	0.06803972	0.93122471

approach indicates that estimated values for new journeys are not so good.

In order to compare our proposed approach with previous research works in literature, we have performed a second prediction test, in this case dividing the stress levels into 3 categories: low, medium and high stress levels. This classification of stress levels has been used in previous research studies such as [14] and [15] in which low, medium and high stress levels are characterized by the properties found during rest, highway driving and city driving periods and validated using questionnaires. In order to further divide the perceived stress while driving in similar road types, we have used the standard deviation of the computed signal (both for the LF/HF and the RMSSD) to classify as medium/average samples those comprised between the mean value plus/minus the standard deviation. This is consistent with the data presented in [14] for the stressor based computed signal. For the LF/HF ratio, values above the mean plus the standard deviation are labeled as high stress. For the RMSSD, values behind the mean minus the standard deviation are labeled as high stress levels. We have used different classification techniques. The confusion matrixes for the Support Vector Machine (SVM), the Naïve Bayes and the J.48 algorithms for a 10-fold cross-validation are captured in tables IX, X and XI. J.48 and SVM show better results than Naïve Bayes. For the case of SVM, 97.5% of the predicted samples are located in the correct group. With the SVM none of the high stress segments is predicted as low stress and vice-versa.

TABLE XII
LF/HF CORRELATION COEFFICIENTS (CITY + RURAL)

	Train one user-test with a different one
MLP	0.6565
Linear regression	0.8676
GP	-0.2029

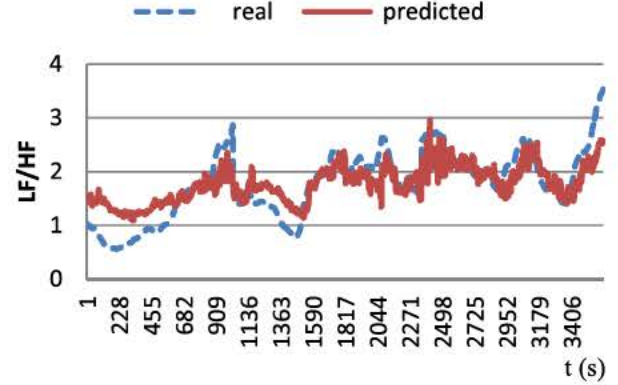


Fig. 4. Predicted vs real values for a particular test drive using linear regression.

TABLE XIII
RMSSD CORRELATION COEFFICIENTS (CITY + RURAL)

	Train one user-test with a different one
MLP	0.6922
Linear regression	0.9247
GP	-0.4148

B. Cross-User Scenario

This sub-section captures the prediction results when training the algorithms with all the data available for one particular user and applying the trained algorithms to validate the data from a second driver. This will show if results from one user are generalizable to the rest of the users.

Table XII captures the correlation coefficients for the LF/HF ratio when training the algorithm for the user in Sheffield and applying it to predict the next levels of stress for the user in Seville. In this case, using a linear regression technique we are able to achieve a correlation coefficient of 0.87. The graphical results are captured in figure 4.

Table XIII captures the same results for the case of the RMSSD signal. As in the previous case, the linear regression technique outperforms both MLP and GP and in this cases achieve a correlation coefficient of 0.92. The graphical results are captured in figure 5.

The results for the inter and intra city are not captured for this scenario since they proved to be worse than those for the entire journey for the intra-user case.

Dividing the samples in the 3 level scale as we did for the single-user case, and training the algorithm for the user in Sheffield and predicting the values for the driver in Seville, we get the results in tables XIV, XV and XVI. In this case, we can see that the results show very poor performance for inter-user predictions, especially for high values of stress.

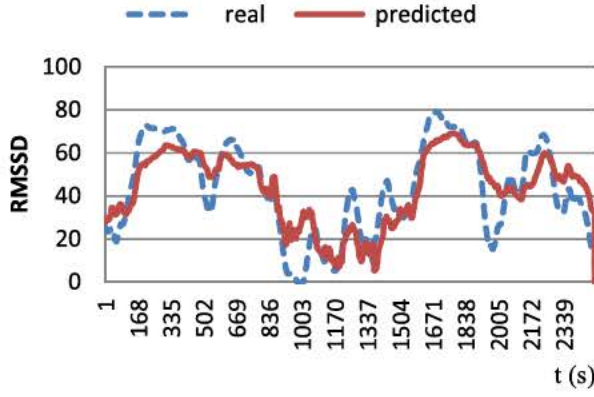


Fig. 5. Predicted vs real values for a particular test drive using linear regression.

TABLE XIV
PREDICTED VALUE FOR UPCOMING LF/HF USING
SVM AND 10-FOLD VALIDATION

Actual/Predicted	Low	Medium	High
Low	0.02105263	0.97894737	0
Medium	0.25121628	0.74568775	0.00309598
High	0.28336079	0.71663921	0

TABLE XV
PREDICTED VALUE FOR UPCOMING LF/HF USING
NAÏVE BAYES AND 10-FOLD VALIDATION

Actual/Predicted	Low	Medium	High
Low	0.96052632	0.03947368	0
Medium	0.82308713	0.17691287	0
High	0.91103789	0.08896211	0

TABLE XVI
PREDICTED VALUE FOR UPCOMING LF/HF USING
J.48 AND 10-FOLD VALIDATION

Actual/Predicted	Low	Medium	High
Low	0.42763158	0.57236842	0
Medium	0.48827952	0.51172048	0
High	0.42833608	0.57166392	0

This result is aligned with those found in [18] in which, in order to deal with inter-person differences, the authors use a clustering algorithm to first divide the users into similar clusters.

C. Predicting Stress Without the HRV Signal

Results in tables I and V show that there is certain correlation between the driving behavior and road shape and the stress signal. In this section we present the results about up to what level the current LH/HF and RMSSD signals (based on the last 5 minute window data) can be estimated based only on the information of the driving related and road related variables.

Table XVII captures the correlation coefficients between the LF/HF measured signal and the estimated one using the same

TABLE XVII
LF/HF PREDICTIONS

	Same user	Different User
MLP	0.3314	-0.3252
Linear regression	0.4435	-0.5601
GP	0.1463	-0.3109

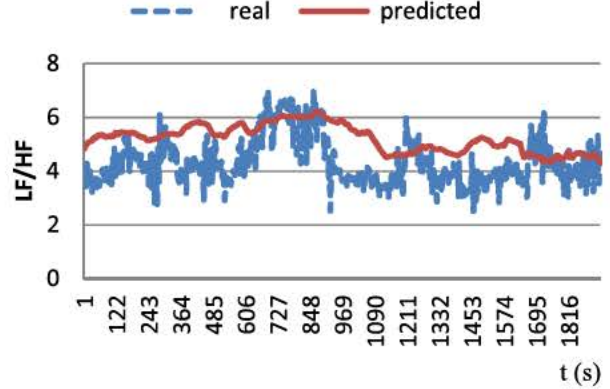


Fig. 6. Predicted vs real values for a particular test drive using linear regression.

TABLE XVIII
RMSSD PREDICTIONS

	Same user	Different User
MLP	0.1528	-0.0502
Linear regression	0.0858	-0.0117
GP	-0.0578	0.1004

regression techniques as in the previous sub-sections and using the leave one journey out validation for the single-user case and using all data from one user for training and all the data for a different user for validation for the cross-user case. In this case, the linear regression method shows the best results for the single-user scenario. The graphical representation of the real signal and the estimated one are captured in figure 6. The results for the cross-user case show that it is not possible to use the data from one user to predict the LF/HF signal for a different one (showing that different users react in not the same way to stressful situations). Again, a clustering algorithm for grouping the users into similar sets is needed and generalization results are bad outside each cluster of users as presented in [18].

The same results for the RMSSD variable are captured in table XVIII. In this case, the correlation is very limited even for the single user scenario and estimating stress based on the estimation of this variable from driving behavior and road shape will provide very poor results.

D. Results for a Different Type of Vehicle

In order to assess the generalization of results, a different driver driving a different type of vehicle has been included in the study. The driver used a motorbike to travel an urban route in Madrid (Spain). 22 different drives in similar traffic conditions on the same route have been used. Figure 7 captures

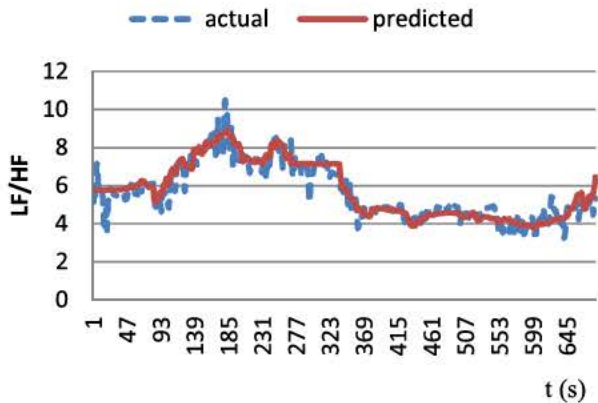


Fig. 7. Predicted vs real values for a particular test drive using MLP.

TABLE XIX
LF/HF CORRELATION COEFFICIENTS (CITY-MOTORBIKE)

	Leave one out	10-fold crossval.
MLP	0.9447	0.9271
Linear regression	0.8279	0.8225
GP	0.7809	0.7763

the actual values for the next minute stress compared to predicted values for a particular drive using the leave-one-out approach and the MLP algorithm. Table XIX captures the results for the same algorithms and validations techniques previously used in table IV for the case of the drivers using a car as the vehicle. In this case, the MLP algorithm slightly outperforms linear regression and Gaussian Processes (GP) but results are similar to those presented in table IV.

V. DISCUSSION BASED ON COMPARING RESULTS WITH PREVIOUS STUDIES

This section is dedicated to compare the results presented in this paper with related results in previous research studies. Table XX captures the results obtained in some of the references in this paper. For each reference, the input variables used are captured together with the objective of the study and the accuracy reported in each paper with the implemented methods. To the best of our knowledge, the major previous research initiatives have tried to assess current values of stress based on the information from several physiological sensors or driving behavior data. Although the objective of our research is trying to go a step forward and predict the upcoming levels of stress, the accuracy obtained in this prediction is captured as compared with the stress measurement existing techniques.

The first reference in table XIX achieves a 81% accuracy for estimating the workload while driving (which is assumed to correlated with the perceived stress). Several driving tasks are monitored in order to compose a workload index. Some of the variables used in this study have been also captured in our system. The references [14], [15], and [24] focus on validating the accuracy of using certain physiological sensors for measuring the driver's perceived stress. Reference [25] augments the previous studies to include fatigue. References [15] and [16] show a similar accuracy (around 82%) in detecting the driver's

TABLE XX
COMPARISON OF RESULTS WITH PREVIOUS WORKS

Ref.	Inputs	Objective	Accuracy
[8]	vehicle velocity, lane position, steering angle, acceleration, pupil diameter and gaze movements	Correct estimation of work load	81%
[14]	22 features from EMG, respiration, ECG and skin conductivity	Detect stress per 5 minute segment	97.4%
[15]	ECG	Detect stress per 5 minute segment	83%
[24]	ECG	Detect stress per 5 minute segment	81.6%
[25]	respiration, ECG and skin conductivity, gaze movement, environmental condions	Detect stress and fatigue	86% for stress
Our approach	HRV, speed, acceleration, PKE, shape of the road	Predict upcoming stress levels	97.5% same user 46.9% different user

perceived stress levels from the single ECG signal. The HRV signal is obtained from the ECG signal and used in the computations for correlations with a 5 minute window to perceived stress (reported by users using questionnaires). Reference [14] is able to achieve better accuracy results (97.4%) by adding other physiological sensors such as the skin conductivity or respiration rate. Our approach is able to predict with a similar accuracy (97.5%) upcoming levels in the LH/HF signal based on past values for the HRV, driving related and road shape related variables (in the case of a single user). The major novelty of our approach is in predicting upcoming levels and not simply assessing current levels of stress based on proxy physiological signals.

VI. CONCLUSIONS

The results in this paper show that the current stress levels while driving, the driving actions taken by the user to respond to previous and recent stressors and road shape show positive correlations in order to estimate upcoming levels of stress (taken as the moving 5 minute average for both the LH/HF and RMSSD over the upcoming 1 minute of driving) for future journeys by the same driver. The current driving related actions taken by the user in response to different stressors will have an impact in how the stress levels evolve in the next driving minute. A correlation index 0.99 for both the LF/HF and RMSSD predicted and real signals are achieved for the single user scenario while 97.5% of the samples are classified in the right clusters when dividing the driving samples in a 3 level stress scale.

The results for the stress prediction in the cross-user scenario show that the way each user handles stress (or executes driving actions as a response to different stressors) is different and worse results are achieved both for correlation coefficients and classification accuracy. As a further work, we plan to gather data from more drivers in order to assess the improvement in results when performing pre-clustering of users.

The results have shown that although it is possible to get an estimation of stress based on current driving behavior and road

shape for a single-user (independently of the vehicle being a car or a motorbike) these results do not generalize a cross-user scenario.

Finally, a comparison table has been included capturing the accuracy obtained in previous related studies and the current paper. We are able to obtain similar results as the best previous studies using less physiological sensors and applying the algorithm to predict upcoming stress rather than assess current levels of stress.

REFERENCES

- [1] "National motor vehicle crash causation survey," U.S. Dept. Transp., Washington, DC, USA, Tech. Rep. DOT HS 811 059, Jul. 2008.
- [2] C. Wu and Y. Liu, "Queuing network modeling of driver workload and performance," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 3, pp. 528–537, Sep. 2007.
- [3] M. Itoh, E. Kawakita, and K. Oguri, "Real-time estimation of driver's mental workload using physiological indexes," in *Proc. ITS World Congr.*, 2010, pp. 1–11.
- [4] E. T. T. Teh, S. Jamson, and O. Carsten, "How does a lane change performed by a neighbouring vehicle affect driver workload?" in *Proc. ITS World Congr.*, 2012, pp. 1–8.
- [5] S. Sega, H. Iwasaki, H. Hiraishi, and F. Mizoguchi, "Verification of driving workload using vehicle signal data for distraction-minimized systems on ITS," in *Proc. ITS World Congr.*, 2011, pp. 1–12.
- [6] J. H. Kim, Y. S. Kim, and W. S. Lee, "Real-time monitoring of driver's cognitive distraction," in *Proc. Spring Conf. Korean Soc. Autom. Eng.*, May 2011, pp. 1197–1202.
- [7] J. Engström, E. Johansson, and J. Östlund, "Effects of visual and cognitive load in real and simulated motorway driving," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 8, no. 2, pp. 97–120, Mar. 2005.
- [8] Y. Zhang, Y. Owechko, and J. Zhang, "Learning-based driver workload estimation," in *Computational Intelligence in Automotive Applications*, vol. 132, D. Prokhorov, Ed. Berlin, Germany: Springer-Verlag, 2008, pp. 1–17.
- [9] E. Adell, A. Várhelyi, and M. dalla Fontana, "The effects of a driver assistance system for safe speed and safe distance—A real-life field study," *Transp. Res. C, Emerg. Technol.*, vol. 19, no. 1, pp. 145–155, Feb. 2011.
- [10] D. Conway, I. Dick, Z. Li, Y. Wang, and F. Chen, "The effect of stress on cognitive load measurement," in *Human-Computer Interaction—INTERACT*. Berlin, Germany: Springer, 2013, pp. 659–666.
- [11] J. A. Veltman and A. W. K. Gaillard, "Physiological workload reactions to increasing levels of task difficulty," *Ergonomics*, vol. 41, no. 5, pp. 656–669, 1998.
- [12] J. Xu *et al.*, "Pupillary response based cognitive workload index under luminance and emotional changes," presented at the CHI, Vancouver, BC, Canada, 2011.
- [13] J. Aasman, G. Mulder, and L. J. M. Mulder, "Operator effort and the measurement of heart-rate variability," *Human Factors, J. Human Factors Ergonomics Soc.*, vol. 29, no. 2, pp. 161–170, 1987.
- [14] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 2, pp. 156–166, Jun. 2005.
- [15] N. Munla, M. Khalil, A. Shahin, and A. Mourad, "Driver stress level detection using HRV analysis," in *Proc. Int. Conf. Adv. Biomed. Eng. (ICABME)*, Sep. 2015, pp. 61–64.
- [16] A. Lanatà *et al.*, "How the autonomic nervous system and driving style change with incremental stressing conditions during simulated driving," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 3, pp. 1505–1517, Jun. 2015.
- [17] A. Sano and R. W. Picard, "Stress recognition using wearable sensors and mobile phones," in *Proc. IEEE Humaine Assoc. Conf. Affective Comput. Intell. Interact. (ACII)*, Sep. 2013, pp. 671–676.
- [18] Q. Xu, T. L. Nwe, and C. Guan, "Cluster-based analysis for personalized stress evaluation using physiological signals," *IEEE J. Biomed. Health Informat.*, vol. 19, no. 1, pp. 275–281, Jan. 2015.
- [19] A. Sano *et al.*, "Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones," in *Proc. IEEE 12th Int. Conf. Wearable Implant. Body Sensor Netw. (BSN)*, Jun. 2015, pp. 1–6.
- [20] T. H. Tarkiainen *et al.*, "Stability over time of short-term heart rate variability," *Clin. Auton. Res.*, vol. 15, no. 6, pp. 394–399, 2005.
- [21] R. P. Sloan *et al.*, "Effect of mental stress throughout the day on cardiac autonomic control," *Biol. Psychol.*, vol. 37, no. 2, pp. 89–99, 1994.
- [22] R. K. Dishman, Y. Nakamura, M. E. García, R. W. Thompson, A. L. Dunn, and S. N. Blair, "Heart rate variability, trait anxiety, and perceived stress among physically fit men and women," *Int. J. Psychophysiol.*, vol. 37, no. 2, pp. 121–133, 2000.
- [23] V. N. Hegde, R. Deekshit, and P. S. Satyanarayana, "Heart rate variability analysis for abnormality detection using time frequency distribution-smoothed pseudo Winger Ville method," *Power (dB)*, vol. 30, no. 20, p. 10, 2013. [Online]. Available: <http://airccj.org/CSCP/vol3/csit3103.pdf>
- [24] J.-S. Wang, C.-W. Lin, and Y.-T. C. Yang, "Using heart rate variability parameter-based feature transformation algorithm for driving stress recognition," in *Advanced Intelligent Computing*. Heidelberg, Germany: Springer, Aug. 2011, pp. 532–537.
- [25] G. Rigas, Y. Goletsis, P. Bougia, and D. I. Fotiadis, "Towards driver's state recognition on real driving conditions," *Int. J. Veh. Technol.*, vol. 2011, p. 617210, Apr. 2011.
- [26] A. Montaña, F. Brown, D. P. Credeur, M. A. Williams, and L. Stoner, "Telemetry-derived heart rate variability responses to a physical stressor," *Clin. Physiol. Funct. Imag.*, to be published.