Accepted Manuscript

Driver Workload Estimation using a Novel Hybrid Method of Error Reduction Ratio Causality and Support Vector Machine

Yang Xing, Chen Lv, Dongpu Cao, Huaji Wang, Yifan Zhao

PII: S0263-2241(17)30626-7
DOI: https://doi.org/10.1016/j.measurement.2017.10.002
Reference: MEASUR 5004

To appear in: *Measurement*

Received Date: 5 May 2017
Revised Date: 29 September 2017
Accepted Date: 2 October 2017

Measurement, Vol. 114, January 2018, Pages 390-397
DOI:10.1016/j.measurement.2017.10.002

Driver Workload Estimation using a Novel Hybrid Method of Error Reduction Ratio Causality and Support Vector Machine

Yang Xing[a], Chen Lv[a], Dongpu Cao[a], Huaji Wang[a] and Yifan Zhao[b*]

[a]*Advanced Vehicle Engineering Centre, School of Aerospace, Transport & Manufacturing, Cranfield University UK*
[b]*Through-life Engineering Services Centre, School of Aerospace, Transport & Manufacturing, Cranfield University, UK*

* Corresponding author: yifan.zhao@cranfield.ac.uk

Abstract

Measuring driver workload is of great significance for improving the understanding of driver behaviours and supporting the improvement of advanced driver assistance systems technologies. In this paper, a novel hybrid method for measuring driver workload estimation for real-world driving data is proposed. Error reduction ratio causality, a new nonlinear causality detection approach, is being proposed in order to assess the correlation of each measured variable to the variation of workload. A full model describing the relationship between the workload and the selected important measurements is then trained via a support vector regression model. Real driving data of 10 participants, comprising 15 measured physiological and vehicle-state variables are used for the purpose of validation. Test results show that the developed error reduction ratio causality method can effectively identify the important variables that relate to the variation of driver workload, and the support vector regression based model can successfully and robustly estimate workload.

Keywords

Driver workload estimation; Driver behaviour; Causality detection; Machine learning; Nonlinear system identification; Correlation analysis

1. Introduction

Intelligent vehicles have been gaining increasing attention from both academia and industrial sectors [1]. The field of intelligent vehicles exhibits a multidisciplinary nature, involving transportation system, automotive engineering, information technology, energy, and security [2]–[9]. Intelligent vehicles have increased their capabilities in highly, and even fully, automated driving. However, unresolved problems do arise due to strong uncertainties surrounding driving experience and complex driver-vehicle interactions. Before transitioning to fully autonomous driving, driver behaviour should be better understood and integrated to enhance vehicle performance and traffic efficiency [10], [11].

Measuring driver workload is of great significance for improving the understanding of driver behaviours, and worthwhile investigating for the purposes of enhancing driver-vehicle interactions [12], [13]. The workload indicates the proportion of an operator's limited capacity that is needed to conduct a specific task [14]. Driving tasks also require drivers to allocate certain amounts of physical and cognitive workload. A driver's workload is dynamically varied with their different driving behaviours, including straight-line driving, cornering, U-turns, rapid acceleration and deceleration, shifting gears, and changing lanes. Furthermore, the level and the variation of drivers' workload that are affected by the above behaviours could be also influenced by many subjective and objective factors, including driving skills, driving styles, trip objectives, personal tendencies, gender, road conditions, traffic conditions, and so on.

A lot of studies focusing on the measurement and estimation of drivers' workloads have been conducted via different methodologies in recent years. In [15], a method of quantifying driver's workload with five discrete levels was proposed by subjective measurement of vehicle data. In [16], the correlation between distraction condition and drivers' mental load was investigated. The results showed that three variables, namely driver's left-pupil size, skin conductance and pulse-to-pulse interval, could be used for efficiently identifying a driver's distraction. In [17], the driver's subjective mental workload and the multiple task performance were modelled through a proposed queuing network method. In [18], drivers' workloads under lane change manoeuvres were investigated through driving simulations. During simulations, the drivers were required to verbally rate the level of their workloads. In [18], the drivers' physiological information, including the electrocardiogram (ECG) signals, eye blinking, pupil diameters and head rotational angles, was measured and used to estimate the workload in a driving simulation scenario. Although the above studies provided multiple potential means to estimate drivers' workload quantitatively by using various physiological signals of drivers, all these studies were performed in the simulation environment, which could not reflect the real driving situations and therefore replicate and measure the impact of the potential uncertainties.

Outside of the simulation environment, driver workload estimation with data from real driving environment has also been studied. Analysis of drivers' workload was conducted using the driving data of a real vehicle [19]. Electroencephalogram (EEG) data was collected by a sensor mounted on the driver's head during actual driving conditions. The experimental data showed that the EEG signals increased when the vehicle speed went over a threshold limit. Moreover, with respect to the driving scenarios, the EEG signals tended to rise with left cornering and downhill. However, the EEG measurement is very sensitive to external disturbances [20], [21]. In this analysis, the original EEG data was used without filtering noise signals caused by vehicle vibrations and other factors, which may affect the reliability of results. Nevertheless, the existing research in driver workload estimation is mainly in the stage of driving simulations, and methodology of measuring workload with real vehicle data is still very challenging, it is still worthwhile improving this.

To further enhance the algorithm of drivers' workload measurement in real world driving situations, in this study, a novel hybrid method of Error Reduction Ratio Causality and Support Vector Machine (ERRC-SVM) is being proposed. To evaluate the performance of the proposed method, this paper uses a real-world driving data set, including driver's physiological states and vehicle states. The rest of this paper is composed as follows. In Section 2, the hybrid method to effectively estimate the driver's workload is proposed. Sections 3 presents the real world driving data set with a pre-processing method. Section 4 shows the analysis results with discussions, and conclusions are given in Section 5.

2. The Hybrid Methods

2.1. Methodology

This paper proposes a novel method, called ERRC-SVM, that combines the knowledge in nonlinear system identification and machine learning to effectively estimate the driver's workload.

There are a variety of sensors available to provide signals that may help to estimate workload, where the sampling rate of each measure could be different. Most data-driven modelling techniques require that the observed measurements have the same sampling rate. A data pre-process procedure is usually required, such as noise filtering, removing mean or normalisation, to prepare the data as required by the methods being used.

There are different methods for developing predictive models and it is very challenging to maintain balance between including too many variables (and therefore loss of precision) and omitting important variables (and therefore risk biased prediction) [22]. Including too many types of measurements will a) increase the cost of system; b) cause the overfitting problem if the number of observed data is limited. It is, therefore, important to remove the irrelevant or less relevant variables before modelling. This paper proposes to use a novel causality detection method, named Error Reduction Ratio Causality (ERRC), to assess the importance of each available measure to the variation of workload using a Nonlinear Finite Impulse Response (NFIR) model.

A full model to describe the relationship between the workload and the selected important input variables can then be trained using a support vector regression (SVR) model, where the training data can be either randomly selected if no temporal lag is considered, or a sequence of data. The model can then be applied to the testing data to evaluate the performance.

The routine of the introduced method can be summarised by Fig. 1.

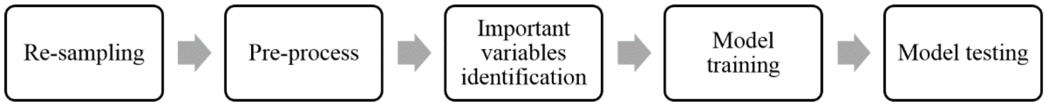

Fig. 1: Process of the proposed ERRC-SVM method.

2.2. Identification of important variables using ERRC

The NFIR model, also known as the Volterra Non-linear Regressive with eXogenous (VNRX) Inputs model, can be used to represent a multi-input and single-output system. The model can be expressed as:

$$y(k) = f\left({u_1}^{[k-1]}, {u_2}^{[k-1]}, \dots, {u_R}^{[k-1]}\right) + \varepsilon(k) \quad (1)$$

where $k(k = 1,2,..)$ is a time index, R is the number of the system inputs, f is some unknown linear or non-linear mapping which links the system output y to the system inputs $u_1, u_2, \dots, u_R$; $\varepsilon(k)$ denotes the model residual. The symbol ${u_i}^{[k-1]}(i = 1,2, \dots, R)$ denotes the past information of the input u_i, which can be expanded as

$${u_i}^{[k-1]} = \cup_{j=0}^{n_i} u_i(k-j) \quad (2)$$

where n_i is the maximum temporal lag to be considered for the input u_i.

A commonly employed model type to specify the function f in Eq. (1) is a polynomial function [23], [24], which can be expressed as

$$y = \theta_0 + \sum_{m=1}^{N} \theta_m \emptyset_m + \varepsilon \quad (3)$$

where $\emptyset_m$ is the m^{th} model term generated from all input vectors; θ_m is the corresponding unknown parameters; N is the total number of potential model terms. Note that $\emptyset_m$ is, in general, non-linear.

If the inputs and output of a system are observable, the model (3) can then be identified. In this paper, the orthogonal least squares (OLS) algorithm [25], is used to determine the model structure from the observations but without estimating the unknown parameters. The OLS algorithm is a popular approach that has been widely used in non-linear system identification where it searches through all the possible candidate model terms to select the most significant model terms which are then included to build model term by term. The significance of each selected model term is measured by an index, called the Error Reduction Ratio (ERR), which indicates how much of the variance change in the system response, in percentage terms, can be accounted for by including the relevant model terms.

Consider a function in linear-in-the-parameters form:

$$y(k) = \sum_{i=0}^{N} \theta_i p_i(k), k = 1,2, \dots, M \quad (4)$$

where $y(k)$ is the dependent variable or the term to regress upon, $p_i(k)$ are regressors, θ_i are unknown parameters to be estimated, M denotes the number of data points in the data set, and N denotes the number of terms in the model that is yet to be determined. Equation (4) can be written as

$$Y = P\Theta \tag{5}$$

where

$$Y = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(M) \end{bmatrix}, P = \begin{bmatrix} P^T(1) \\ P^T(2) \\ \vdots \\ P^T(M) \end{bmatrix}, \Theta = \begin{bmatrix} \theta(1) \\ \theta(2) \\ \vdots \\ \theta(M) \end{bmatrix} \tag{6}$$

where

$$P^T(k) = \left(p_1(k), p_2(k), \dots, p_N(k)\right) \tag{7}$$

Matrix P can be decomposed as $P = W \times A$ where

$$W = \begin{bmatrix} w_1(1) & w_2(1) & \dots & w_N(1) \\ w_1(2) & w_2(2) & \dots & w_N(2) \\ \vdots & \ddots & \ddots & \vdots \\ w_1(M) & w_2(M) & \dots & w_N(M) \end{bmatrix} \tag{8}$$

and A is an upper triangular matrix with unity diagonal elements

$$A = \begin{bmatrix} 1 & a_{12} & a_{13} & \dots & a_{1N} \\ & 1 & a_{23} & \dots & a_{2N} \\ & & \ddots & \ddots & \vdots \\ & & & 1 & a_{N-1N} \\ & & & & 1 \end{bmatrix} \tag{9}$$

Therefore, Equation (5) can be rewritten as

$$Y = WG \tag{10}$$

where $G = A\Theta = [g_1 \quad g_2 \quad \cdots \quad g_N]^T$.

Now we introduce how to estimate the importance of each input variable to the variation of the system output. Initially, set values $a_{11} = 1, w_1(k) = p_1(k)$, and calculate

$$\hat{g}_1 = \frac{\sum_{k=1}^{M} w_1(k) y(k)}{\sum_{k=1}^{M} w_1{}^2(k)} \tag{11}$$

For $j = 2,3, \dots, M$, set $a_{jj} = 1$ and then calculate

$$a_{ij} = \frac{\sum_{k=1}^{M} w_i(k) p_j(k)}{\sum_{k=1}^{M} w_i{}^2(k)} \tag{12}$$

where $i = 1,2, \dots, j-1$. Next calculate

$$w_j(k) = p_j(k) - \sum_{i=1}^{j-1} a_{ij} w_i(k) \tag{13}$$

and

$$\hat{g}_j = \frac{\sum_{k=1}^{M} w_j(k) y(k)}{\sum_{k=1}^{M} w_j{}^2(k)} \tag{14}$$

The ERR value for each team p_i, as a criterion to select the model structure, is defined as

$$ERR_i = \frac{\hat{g}_i{}^2 \sum_{k=1}^{M} w_i{}^2(k)}{\sum_{k=1}^{M} y^2(k)} \tag{15}$$

Values of ERR range always from 0% to 100%. The larger ERR is, the higher dependence is between this term and the output. It is, therefore, a very important index to indicate the importance of each term to the output. To calculate the contribution of each input variable to the output, the sum of ERR values of all selected terms, denoted by $SERR$, is calculated by

$$SERR = \sum_{i=1}^{N} ERR_i \tag{16}$$

Note N is the number of the selected terms, not the number of total candidate terms. The value of $SERR$ describes the percentage explained by the identified model to the system output. If the considered inputs can fully explain the variation of the system output, the value of $SERR$ is equal to 100%. It is an indicator of model performance and uncertainty. The contribution of the i^{th} input variable, u_i, to the variation of the system output, denoted as $ERRC_i$, is defined as the sum of ERR values of the selected terms that include this input variable. Because some selected terms may involve more than one input variable due to nonlinearity, the sum of $ERRC_i$ for all input variables can be greater than $SERR$. To overcome this problem, the value of $ERRC_i$ is normalised and is written as:

$$ERRC_i = \frac{\sum_{j=1}^{N} \left(ERR_j | u_i \in \emptyset_j\right)}{\sum_{p=1}^{R} \sum_{j=1}^{N} \left(ERR_j | u_p \in \emptyset_j\right)} \times SERR \tag{17}$$

The value of $ERRC_i$ should be always between 0% and 100%. This paper proposes to use this value to represent the importance of each measures to estimate the driver workload.

2.3. Model estimation using SVM

After the identification of important measures, this paper proposes to use a support vector machine (SVM) to identify the driver workload model. Being a popular machine learning method, SVM was firstly designed to solve binary classification problems by maximising the margin between different data classes [26]. After that, it was developed to be suitable for both multiple classes classification and regression problems [27]. By utilising the kernel technique, SVM can map linearly inseparable low dimensional data into separable high dimensional data and construct classification hyperplanes efficiently.

In this paper, a support vector regression (SVR) model is used. Considering a system with a single input and single output, given data set $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_M, y_M)\}$, where $x_1, x_{2,\ldots,} x_M$ and $y_1, y_{2,\ldots,} y_M$ are observed input and output respectively. The goal of the SVR model is to find a function $f(x)$ that has at most ε deviation from actual target y_i. Meanwhile, the linear decision boundary $f(x) = \langle w, x \rangle + b$ should be as flat as possible. Therefore, by applying a soft margin, the loss function for SVR can be represented as:

$$\min_{w,\xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{l} (\xi_i + \xi_i^*) \tag{18}$$

subject to

$$\begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \tag{19}$$

where ξ_i and ξ_i^* are the slack variables in the soft margin loss function that allows error cases larger than ε. $C > 0$ controls the trade-off between the flatness of function $f(x)$ and the tolerate ratio for the cases that has larger deviations than ε. To efficiently solve the optimization problem, the linear decision function can be further nonlinearized with kernel function, which has the following form.

$$K(x_i, x_j) = \emptyset(x_i)^T \emptyset(x_j) \tag{20}$$

where $\emptyset(x)$ is the mapping function that maps the raw data points into higher dimensions. In this study, a specific kernel function named radial basis function (RBF) is adopted, which has the following form:

$$K(x_i, x_j) = \left(-\gamma \left\|x_i - x_j\right\|^2\right), \ \gamma > 0 \tag{21}$$

According to the aforementioned concept, the penalty term C and the kernel function parameter γ are the two most important parameters that determine the model of the

SVM. Therefore, in this paper, the Genetic Algorithm (GA) is used to iteratively search the optimal C and γ that can maximise the classification margin or minimise the fitting error. Genetic algorithm is a specific form of evolutionary algorithm inspired by the process of natural population selection. The major parts of GA contain encoding, selection, crossover, and mutation. Specifically, encoding is a process of genetic representation that describes the SVM parameters using multi-bits binary values like genes. Each individual in the population group represents one possible value of the SVM parameter pair. The selection process chooses the individuals with good fitness and pass their genes to the next generation. To increase the gene diversity, crossover aims to make the genes better by randomly combining two well-fitted individuals and the mutation prevents the optimisation results from being blocked in the local minimum according to the mutation probability. Since the fitting function is difficult to describe using a single equation in this study, the training error is adopted to measure the quality of fitness. Specifically, the candidate C and γ are decoded and fed into the SVM, the training error, which is the average difference between the predicted workload values and the ground truth values. The detailed implementation of this method can be found in [28]. The procedure of model estimation can be summarised by Table 1.

Table 1: Procedure of combining SVR and GA

1. **Input**: train data with label
2. Initialize GA parameters and generate first population
3. **for** i = 1,2,...,Max generation
4.   Decoding chromosomes
5.   Computing fitness using SVR for all population
6.   Select Population according to individual fitness
7.   Crossover and mutation to create offspring
8. **end for**
9. Find best model parameters
10. Train SVR with optimized Parameters
11. Test SVR regression model
12. **Output**: performance Index and optimized parameters for SVR

3. Dataset and Pre-processing

Assessing the drivers' workload in a simulator study is hardly possible because drivers always know that they are navigating through a virtual world. Using the proposed method, this paper analysed a public dataset [29] collected through a real world driving study. The physiological states of the participants, including Skin Conductance Response (SCR), hand temperature and heart rate using ECG, were recorded. The GPS position, brightness level and acceleration were also recorded. Two cameras were used to record the driver's view

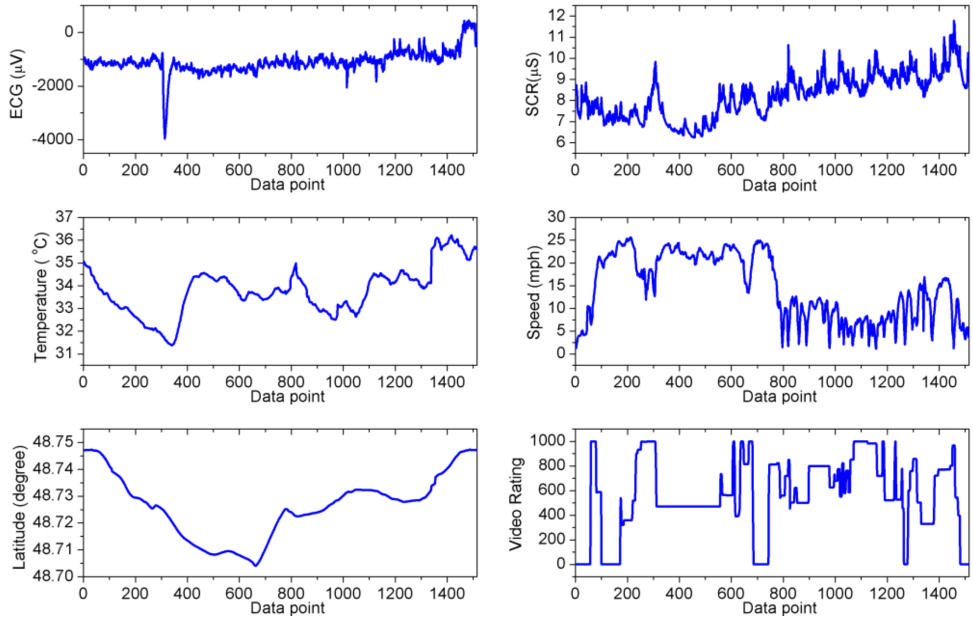

Fig. 2: Selected measures and the video rating representing the workload for the participant 1

Table 2: The correspondence between the variables and the symbols used in the NFIR model.

Symbol	Variable	Symbol	Variable
u_1	ECG	u_9	Lightning
u_2	SCR	u_{10}	Latitude GPS
u_3	Temperature	u_{11}	Longitude GPS
u_4	Heart Rate (HR)	u_{12}	Accuracy GPS
u_5	Heart rate variability (HRV_LF)	u_{13}	Altitude GPS
u_6	Acceleration X	u_{14}	Speed GPS
u_7	Acceleration Y	u_{15}	Bearing GPS
u_8	Acceleration Z	y	Video Rating

onto the road, and the participants were asked to rate the workload offline based on these videos to provide the baseline. The video rating of workload is in the range of 0 to 1000. A value of 1000 indicates a maximum workload. Ten participants (3 females, 7 males) aged between 23 and 57 years took part in these experiments.

A data pre-process step was undertaken before applying the developed method. Since the sampling frequency of each measure is different, the data have been re-sampled at the frequency of 1Hz. There are 1515 (25 min and 15 sec) data points for each variable. An example of the selected measures and the video rating for the participant 1 is shown in Fig. 2.

To save space to present and discuss the results, the corresponding mapping between the name of variables and the symbols in Eq. (3) is shown in Table 2. The mean of each input and output was removed as suggested in Section 2.

4. Results and Discussions

4.1. Identification of important variables

Based on Eq. (3), the NFIR model with quadratic terms to establish the relationship between the 15 input variables and the output is proposed to solve the problem and it can be written as

$$\boldsymbol{y = \theta_0 + \sum_{i=1}^{15} \theta_i u_i + \sum_{i=1}^{15} \sum_{j=i}^{15} \theta_{ij} u_i u_j + \varepsilon} \tag{18}$$

To simplify the model, initially, there is no temporal lag of each input being considered. This model includes 136 candidate terms consisting of 16 linear terms $\{\theta_0, \bigcup_{i=1}^{15} \theta_i u_i\}$ and 120 nonlinear terms $\{\bigcup_{i=1}^{15} \bigcup_{j=i}^{15} \theta_{ij} u_i u_j\}$. The proposed method was then applied to calculate the values of ERRC of input variables and the corresponding SERR for all 10 participants. The results can be illustrated by Fig. 3. The detailed values are shown in Table A1. The results for all 10 participants exhibit a consistent pattern of contribution of each input. It is observed that some common parameters have a low ERRC values, and are therefore less relevant to the driver workload prediction. For example, the values of heart rate, acceleration along different directions, and the accurate GPS signals for all the participants show little contribution to the rating values. However, some input parameters such as ECG, vehicle speed, and latitude GPS signals of most participants show high correlation to the video rating.

To statistically evaluate the importance for each input, the inputs were ranked based on the significance of ERRC values. The second column of Table 3 shows the percentage for each input when the ERRC value is non-zero. The third and fourth column shows the percentage when the ERRC value of the considered input is within the top 3 and the top 5 of all 15 inputs respectively. It can be clearly seen that u_1, u_2, u_3, u_{10}, u_{11}, u_{13}, u_{14} and u_{15} have more than 80% probability that the ERRC values are non-zero, and furthermore, they

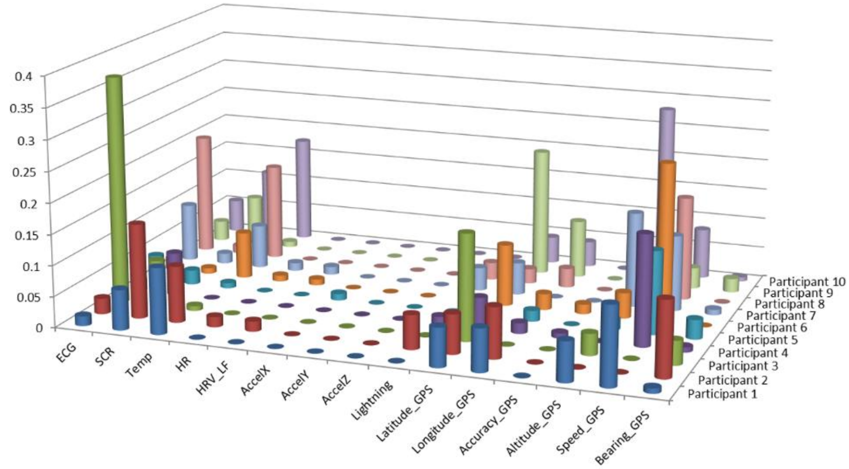

Fig. 3: Illustration of the calculated ERRC indicating contribution of each input to the variation of video rating for all participants.

Table 3: Significant analysis for all variables based on the ERRC values across different participants shown in Table 3.

Symbol	Non-zero	Top 3	Top 5
$\mathbf{u_1}$	100%	40%	60%
$\mathbf{u_2}$	100%	50%	70%
$\mathbf{u_3}$	90%	50%	60%
u_4	30%	0%	0%
u_5	30%	0%	0%
u_6	10%	0%	0%
u_7	0%	0%	0%
u_8	0%	0%	0%
u_9	40%	0%	10%
$\mathbf{u_{10}}$	100%	40%	50%
$\mathbf{u_{11}}$	80%	20%	70%
u_{12}	20%	0%	0%
$\mathbf{u_{13}}$	70%	20%	50%
$\mathbf{u_{14}}$	80%	60%	80%
u_{15}	90%	20%	30%

all have appeared at least once in the top 3. However, u_{15} has only 30% probability that the ERRC values are in the top 5 while others have more than 50% probability. In this study, seven measures (u_1, u_2, u_3, u_{10}, u_{11}, u_{13} and u_{14}), highlighted in Table 3, are therefore considered as the important measures and will be used in the next section to estimate the driver's workload.

The above analysis has not considered the temporal lag of each input. If the past information is considered, the model complexity will be significantly increased. For example, if the past 5 seconds information ($n_i = 5$ in Eq. (2)) is considered, the total number of model terms will be increased from 136 to 2926, and increased to 11476 if the past 10 seconds information is considered. Unless it can significantly improve the performance, including past information should be avoided to save computational time and avoid the model overfitting problem. Table 4 shows the comparison of SERR values without and with considering past information of 7 selected inputs. It is clearly shown that there is no significant improvement for the value of SERR except for the participant 2 and 5. Considering the number of sampling data of 1515, this paper only considers the input at time t to predict the output at time t in the next section.

Table 4: Calculated SERR values indicating the percentage of variation of the output explained by inputs both without and with taking into consideration past information for each participant.

Participant	**SERR**		
	0 second	**5 seconds**	**10 seconds**
1	0.51	0.51	0.53
2	0.47	0.55	0.58
3	0.68	0.71	0.72
4	0.37	0.37	0.37
5	0.28	0.28	0.39
6	0.56	0.56	0.56
7	0.53	0.54	0.62
8	0.57	0.57	0.58
9	0.50	0.52	0.52
10	0.79	0.79	0.80

4.2. Driver workload estimation

SVR models for 10 participants were trained based on the selected 7 indicators which are Latitude GPS, Longitude GPS, Altitude GPS, Speed GPS, ECG, SCR, and Temperature. The boundary of the parameter C was in the range of $[0.1,100]$ and γ was in the range of $[0.01,1000]$. The maximum generation was set as 100, size of population was set as 30, crossover probability was set as 0.4, and mutation probability was set as 0.01. The notion of these parameters can be seen in [28]. For each group of data, 600 data points were randomly selected for the testing purpose and the remaining points were utilized for training. By using the genetic algorithm, the optimal C and γ for 10 participants were estimated and are given in Table 5.

Table 5: The identified parameters for the SVR models.

Participant	C	γ
1	97.18	48.75
2	84.60	64.54
3	76.22	73.57
4	69.59	130.36
5	93.94	73.03
6	90.09	133.23
7	85.79	74.56
8	85.82	145.90
9	24.80	70.63
10	84.42	65.78

Fig. 4. illustrates the estimation of workload based on SVR for all the ten participants. Note 600 randomly selected data points (about 1/3 of the total data) for each participant were used to test the trained SVR model. As shown in the figure, SVR generates precise estimations of workload for all participants, demonstrating that SVR is a robust and reliable estimator for observing workload of different subjects.

To evaluate the performance of different input combinations on the workload estimation, three different scenarios were tested and compared: a) human body features only (ECG, SCR, and temperature); b) GPS signals only (latitude, longitude, altitude, and speed) and c) GPS and human body features. The model performance of each participant, represented by the value of R^2 (Pearson correlation coefficient) between the model prediction and recorded data, is shown in Table 6.

According to the results in Table 6, the average value of R^2 with the workload estimation based on the selected three human body features only is 0.70, while the average value of R^2 under the estimation based on the four GPS signals reaches 0.83. This leads to an interesting conclusion that vehicle state measurements are more relevant to driver workload estimation than those physiological signals. It should be noted that some GPS

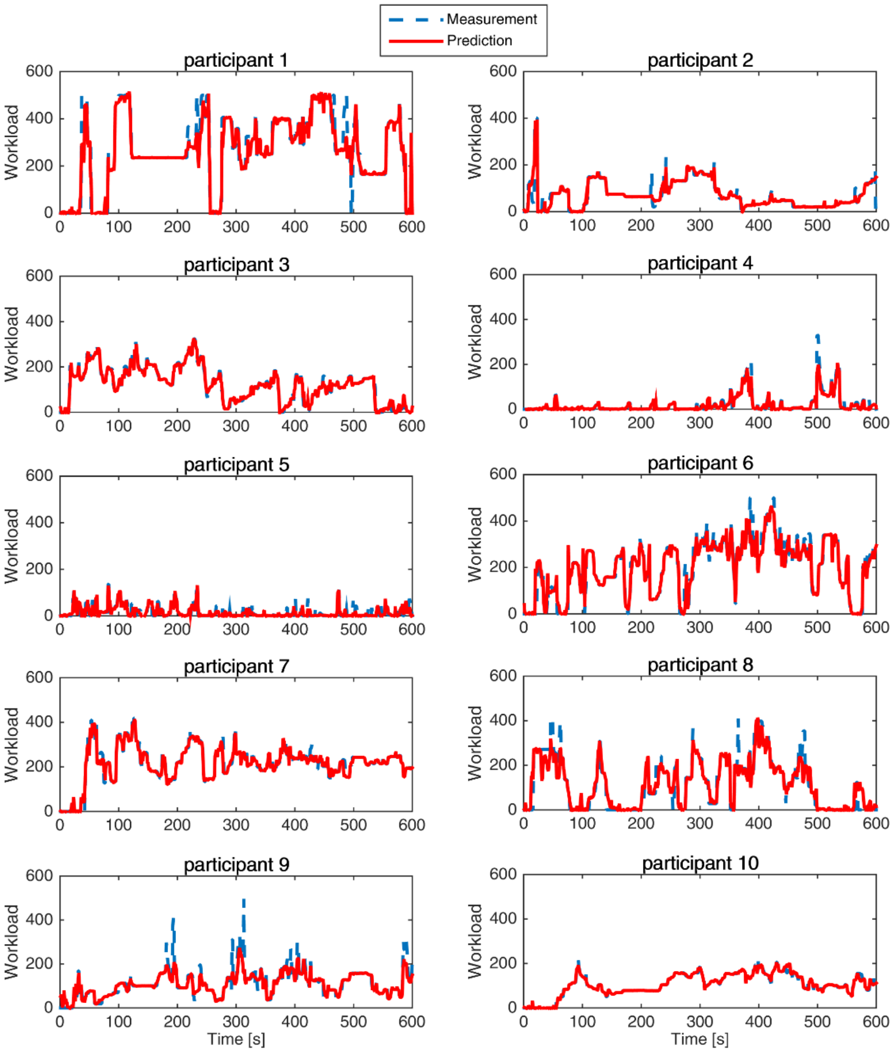

Fig. 4: The comparison between the predicted workload using the trained SVR models and the recorded video rating for 10 participants.

features (e.g. latitude, longitude, altitude) describe the position of the vehicle. In other words, they reflect the road condition, which is subjective. The participants rate the workload based on the video, which captures the road condition and traffic condition. It is therefore not surprising to observe that GPS features perform well in prediction. The final target of this research is to use human body features to estimate the workload. Therefore, this paper is interested in the difference of R^2 between Human Body Features and GPS & Human Body Features. It is observed that the combination of human body features and

Table 6: Comparison of model performance with different input combinations.

Participant	R^2		
	Human Body Features	GPS Features	GPS and Human Body Features
1	0.68	0.85	0.91
2	0.87	0.67	0.92
3	0.81	0.91	0.97
4	0.54	0.86	0.88
5	0.38	0.69	0.74
6	0.72	0.89	0.94
7	0.72	0.83	0.94
8	0.75	0.86	0.90
9	0.62	0.77	0.83
10	0.95	0.96	0.97
Average	0.70	0.83	0.90

vehicle's GPS information would construct more relevant feature vectors for estimating driver workload, with an overall improvement of R^2 of 0.2. It has also been observed that this difference varies between different participants. For the participant 2 and 10, the difference is smaller than 0.05, which indicates that the selected human body features are sufficient to describe the workload. For the participant 4 and 5, the difference is larger than 0.3, which indicates that the selected human body features are not sufficient and more features (e.g. motion, eye gaze) should be considered in order to better estimate the workload.

5. Conclusions

In this paper, a novel hybrid method for measuring driver workload estimation with real-world driving data is proposed. Error reduction ratio causality, a new causality detection approach, is synthesized to quantify the correlation of each measured variable to the variation of workload using a nonlinear finite impulse response model. A full model describing the relationship between the workload and the selected measurements is then identified via a support vector regression model. Real driving data of 10 participants with 15 measured driver's physiological and vehicle state variables are used for the algorithm development, model training, testing and verification. Test results show that the developed error reduction ratio causality method can effectively identify the important variables with the variation of driver workload. Furthermore, the support vector regression based model can successfully and robustly estimate driving workload. The combination of human body features and vehicle's GPS information constructs more relevant feature vectors for estimating driver workload, resulting in an improved performance of the Pearson correlation coefficient at 0.90. The results demonstrate the feasibility and effectiveness of the proposed novel hybrid methodology for driver workload estimation. It is also concluded that more human body features (e.g. motion, eye gaze) should be considered to better estimate the workload if the subjective GPS features are excluded.

Further work will be carried out in the following areas: collecting more data of real world driving tasks under various situations with more human body features, further refinement of the developed hybrid algorithms for driver workload estimation.

Acknowledgements

This work was supported by Through-life Engineering Service Centre at Cranfield University.

Appendix

Table A1: Calculated ERRC indicating contribution of each input to the variation of video rating for all participants.

Input	Participant									
	1	2	3	4	5	6	7	8	9	10
u_1	0.02	0.03	**0.37**	0.02	0.04	0.02	0.10	**0.20**	0.03	0.05
u_2	0.07	**0.15**	0.08	0.07	0.02	0.01	0.02	0.01	0.08	0.11
u_3	0.1	0.09	0.01	0	0.01	0.08	0.07	0.16	0.01	0.17
u_4	0	0.02	0	0	0	0.01	0.01	0	0	0
u_5	0	0.02	0	0	0	0.01	0.01	0	0	0
u_6	0	0	0	0	0.01	0	0	0	0	0

u_7	0	0	0	0	0	0	0	0	0	0
u_8	0	0	0	0	0	0	0	0	0	0
u_9	0	0.06	0	0.01	0	0	0.04	0.03	0.01	0
u_{10}	0.0	0.06	0.17	0.05	0.01	0.10	0.05	0.02	**0.21**	0.04
u_{11}	0.0	0.08	0	0.02	0.02	0.03	0	0.03	0.09	0.04
u_{12}	0	0	0	0.01	0	0.02	0	0	0	0
u_{13}	0.0	0	0.03	0	0.02	0.04	0.15	0	0.02	**0.28**
u_{14}	**0.1**	0	0	**0.18**	**0.13**	**0.26**	**0.12**	0.17	0.03	0.08
u_{15}	0.0	0.12	0.04	0.01	0.03	0	0.01	0	0.02	0.01

References

[1] J. L. Gabbard, G. M. Fitch, and H. Kim, “Behind the Glass: Driver Challenges and Opportunities for AR Automotive Applications,” *Proc. IEEE*, vol. 102, no. 2, pp. 124–136, Feb. 2014.

[2] W. Hernandez, “Optimal estimation of the relevant information coming from a rollover sensor placed in a car under performance tests,” *Measurement*, vol. 41, no. 1, pp. 20–31, Jan. 2008.

[3] J. C. Castellanos and F. Fruett, “Embedded system to evaluate the passenger comfort in public transportation based on dynamical vehicle behavior with user’s feedback,” *Measurement*, vol. 47, pp. 442–451, Jan. 2014.

[4] G. Andria, F. Attivissimo, A. Di Nisio, A. M. L. Lanzolla, and A. Pellegrino, “Development of an automotive data acquisition platform for analysis of driving behavior,” *Measurement*, vol. 93, pp. 278–287, Nov. 2016.

[5] Q. Zhang, Q. Wu, Y. Zhou, X. Wu, Y. Ou, and H. Zhou, “Webcam-based, non-contact, real-time measurement for the physiological parameters of drivers,” *Measurement*, vol. 100, pp. 311–321, Mar. 2017.

[6] M. A. Sotelo, “Electrical, Connected, and Automated Transportation Editor’s Column,” *IEEE Intell. Transp. Syst. Mag.*, vol. 8, no. 2, pp. 2–2, 2016.

[7] V. Faure, R. Lobjois, and N. Benguigui, “The effects of driving environment complexity and dual tasking on drivers’ mental workload and eye blink behavior,” *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 40, pp. 78–90, Jul. 2016.

[8] C. Marina Martinez, X. Hu, D. Cao, E. Velenis, B. Gao, and M. Wellers, "Energy Management in Plug-in Hybrid Electric Vehicles: Recent Progress and a Connected Vehicles Perspective," *IEEE Trans. Veh. Technol.*, pp. 1–1, 2016.

[9] D. Pecchini, R. Roncella, G. Forlani, and F. Giuliani, "Measuring driving workload of heavy vehicles at roundabouts," *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 45, pp. 27–42, Feb. 2017.

[10] P. Choudhary and N. R. Velaga, "Analysis of vehicle-based lateral performance measures during distracted driving due to phone use," *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 44, pp. 120–133, Jan. 2017.

[11] M. Niezgoda, A. Tarnowski, M. Kruszewski, and T. Kamiński, "Towards testing auditory–vocal interfaces and detecting distraction while driving: A comparison of eye-movement measures in the assessment of cognitive workload," *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 32, pp. 23–34, Jul. 2015.

[12] M. Chan, S. Nyazika, and A. Singhal, "Effects of a front-seat passenger on driver attention: An electrophysiological approach," *Transp. Res. Part F Traffic Psychol. Behav.*, vol. 43, pp. 67–79, Nov. 2016.

[13] B. Okumura, M. R. James, Y. Kanzawa, M. Derry, K. Sakai, T. Nishi, and D. Prokhorov, "Challenges in Perception and Decision Making for Intelligent Automotive Vehicles: A Case Study," *IEEE Trans. Intell. Veh.*, vol. 1, no. 1, pp. 20–32, Mar. 2016.

[14] R. D. O'Donnell and F. T. Eggemeie, "Workload assessment methodology," in *Cognitive Processes and Performance*, K. R. Boff, L. Kaufman, and J. P. Thomas, Eds. John Wiley and Sons, 1986.

[15] S. Sega, H. Iwasaki, H. Hiraishi, and F. Mizoguchi, "Verification of driving workload using vehicle signal data for distraction-minimized systems on ITS," in *18th ITS World Congress*, 2011.

[16] Y. Liang, M. L. Reyes, and J. D. Lee, "Real-Time Detection of Driver Cognitive Distraction Using Support Vector Machines," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 340–350, Jun. 2007.

[17] Changxu Wu and Yili Liu, "Queuing Network Modeling of Driver Workload and Performance," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 3, pp. 528–537, Sep.

2007.

[18] E. T. T. Teh, S. Jamson, and O. Carsten, “How does a lane change performed by a neighbouring vehicle affect driver workload?,” in *19th ITS World Congress*, 2012.

[19] J.-B. Lim, S.-B. Lee, K.-H. Kim, S.-Y. Kim, and J.-S. Choi, “A study of the relationship between driver’s anxiety eeg and driving speed in motorway sections,” *J. Korean Soc. Saf.*, vol. 27, no. 3, pp. 167–175, 2012.

[20] Y. Zhao, S. A. Billings, H.-L. W. Wei, and P. G. Sarrigiannis, “A Parametric Method to Measure Time-Varying Linear and Nonlinear Causality With Applications to EEG Data,” *IEEE Trans. Biomed. Eng.*, vol. 60, no. 11, pp. 3141–3148, Nov. 2013.

[21] Y. Zhao, S. A. Billings, H. Wei, F. He, and P. G. Sarrigiannis, “A new NARX-based Granger linear and nonlinear casual influence detection method with applications to EEG data,” *J. Neurosci. Methods*, vol. 212, no. 1, pp. 79–86, Jan. 2013.

[22] P. A. Murtaugh, “Methods of variable selection in regression modeling,” *Commun. Stat. - Simul. Comput.*, vol. 27, no. 3, pp. 711–734, Jan. 1998.

[23] H. L. Wei, S. A. Billings, and J. Liu, “Term and variable selection for non-linear system identification,” *Int. J. Control*, vol. 77, no. 1, pp. 86–110, Jan. 2004.

[24] S. Chen and S. . Billings, “Representations of non-linear systems: the NARMAX model,” *Int. J. Control*, vol. 49, no. 3, pp. 1013–1032, 1989.

[25] S. A. Billings, *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. John Wiley & Sons, 2013.

[26] C. Cortes and V. Vapnik, “Support-Vector Networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[27] A. J. Smola and B. Schölkopf, “A tutorial on support vector regression,” *Stat. Comput.*, vol. 14, no. 3, pp. 199–222, 2004.

[28] F. Rong-en, P. Chen, and C. Lin, “Working set selection using second order information for training support vector machines,” *J. Mach. Learn. Res.*, vol. 6, pp. 1889–1918, 2005.

[29] S. Schneegass, B. Pfleging, N. Broy, A. Schmidt, and F. Heinrich, “A data set of

real world driving to assess driver workload," in *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '13*, 2013, pp. 150–157.

- The contribution of physiological and vehicle measures to workload is quantified
- Seven measures that have high contribution to explain workload are identified
- Drivers' workload can be effectively predicted using GPS and human body features

2017-10-04

Driver workload estimation using a novel hybrid method of error reduction ratio causality and support vector machine

Xing, Yang

Elsevier

Xing Y, Lv C, Cao D, Wang H, Zhao Y, Driver workload estimation using a novel hybrid method of error reduction ratio causality and support vector machine, Measurement, Vol. 114, January 2018, pp. 390-397

https://doi.org/10.1016/j.measurement.2017.10.002