Soft Comput manuscript No.
(will be inserted by the editor)

Mesoscopic forecasting of vehicular consumption using neural networks

Michail Masikos · **Konstantinos Demestichas** · **Evgenia Adamopoulou** · **Michael Theologou**

Received: date / Accepted: date

Abstract Accurate forecasting of vehicular consumption is a task of primary importance for several applications. Herein, a vehicular consumption prediction model is proposed, with special emphasis on robustness and reliability. Both features are enabled due to the selection of General Regression Neural Networks (GRNNs) for the implementation of the proposed model. GRNNs are widely used among neural networks because of their capabilities for fast learning and successful convergence to the solution. In particular, the designed GRNN is responsible for approximating the nonlinearities and the specificities between the factors identified as major contributors in vehicular consumption. In order to evaluate its efficiency, a case study involving the application of the introduced model in Fully Electric Vehicles (FEVs) is examined. The performance of the proposed model is successfully validated using real measurements collected during a data acquisition field campaign.

Keywords energy-efficient routing · mesoscopic consumption model · context-aware prediction · FEV

1 Introduction

With the continuous growth of the transportation sector, the amount of consumed energy and the volume of emitted greenhouse gases increase unceasingly. In order to limit the impact of these environmental effects, the scientific community responded with the development of eco-driving and eco-routing systems (Boriboonsomsin et al 2012; Kamal et al 2011). The term 'eco-driving' refers to systems that validate a posteriori the incurred vehicle consumption rate and provide valuable feedback to the driver on how to reduce it. Eco-routing systems, on the other hand, include systems that generate a priori the least 'harmful' routes from an environmental perspective and suggest them to the driver as alternatives. Considering this functional difference, it appears that the development of eco-routing systems is less straightforward than the development of eco-driving systems. The performance of eco-routing systems depends mainly on their efficiency in accurately predicting the consumption or the emissions along the potential routes towards the desired destination.

Our intent in this paper is to provide an innovative model for accurately forecasting the amount of energy required to travel along a road segment. In particular, we propose the adoption of a learning method for the implementation of the forecasting functionality, exploiting the knowledge contained in previously collected travelling experience. As part of our research work, we have designed a General Regression Neural Network (GRNN) algorithm to use as the learning model and we have carefully formulated the set of forecasting variables. Both of these features distinguish our model from previously presented solutions. More specifically, the former provides to our model the abilities of fast learning and successful convergence to the solution, rendering, thus, the forecasting process faster and more accurate, whereas the latter enhances our model's robustness. Indeed, the inclusion of an extensive set of parameters as forecasting variables enables the optimal adaptation of the model's estimation process to the prevailing contextual condi-

M. Masikos(✉) · K. Demestichas · E. Adamopoulou · M. Theologou
School of Electrical and Computer Engineering, National Technical University of Athens, Zografou, GR15780, Greece
Tel.: +30-210-7721478
E-mail: mmasik@telecom.ntua.gr

tions (e.g. traffic conditions, weather conditions etc.). Apart from this, both of the aforementioned design choices are consistent with the aim of developing a model suitable for integration as part of an on-board eco-routing system.

The remainder of this paper is organized as follows: Section 2 concisely reviews the existing vehicular consumption estimation models. Section 3 presents the proposed prediction model. Section 4 analyzes the problem under investigation, justifies the selection of the model's input parameters and reveals the characteristics of the collected measurements. Section 5 analyzes the significance of the selected features in the developed model, and presents comparative evaluation results for the proposed model as well as for a reference model currently applied in routing systems. Section 6 discusses the extracted results and draws conclusions.

2 Literature review

Several techniques have been proposed so far for vehicular consumption estimation. Depending on the level of detail that the existing models incorporate in their calculations, they can be classified into macroscopic, mesoscopic or microscopic approaches. Moving from macro- to micro-scopic approaches the attention to detail grows (replacing averaged values with instantaneous ones), without, however, ensuring a proportional improvement in terms of prediction accuracy. On the contrary, the achieved accuracy is not always consistent with the required amount of computational resources. All these parameters should be considered and evaluated before choosing a specific type of consumption model.

A macroscopic approach for predicting fuel consumption indices in diverse environments that employs a back-propagation neural network is presented in (Wu and Liu 2011). In order to facilitate car manufacturers in designing energy-efficient vehicles that are compliant with emissions standards, Wu and Liu (2011) developed a learning model that forecasts the fuel consumption rate indices in city, highway or combined cycles based on certain vehicle characteristics, i.e. engine type, car weight, vehicle class and transmission type. Although the learning capacity of this model is verified, such a model would not perform efficiently in case of an eco-routing system, where context-specific consumption rates are observed (as opposed to static average values). The aforementioned approach is further enhanced in terms of time and accuracy performance by replacing the back-propagation neural network with a radial basis function neural network (Wu and Liu 2012). Another macroscopic, non-iterative algorithm for estimating the fuel consumption of vehicles is presented in (Ben Dhaou 2011). This algorithm uses Willan's internal combustion engine model (Pachernegg 1969) and requires no instantaneous values of speed or acceleration. The efficiency of the proposed algorithm has been verified against measurement results generated for the following three cycles: motor vehicle expert group (MVEG-95), European driving cycle (ECE), and extra-urban driving cycle (EUDC).

A research tool implementing a mesoscopic approach is presented in (Minett et al 2011). The proposed tool generates synthetic speed profiles based on historical link speed data already stored as digital map attributes. These speed profiles are then used as the basis for estimating the fuel costs per road segment. The generated speed profiles and the corresponding fuel cost estimations are validated against field test data. Yao and Song (2013) attempts to establish a series of mesoscopic models for light-duty gasoline vehicles, mid-duty diesel vehicles, and heavy-duty diesel vehicles separately, based on data collected by a portable emissions measurement system under actual driving conditions. The proposed models consider the influence of the vehicle speed, acceleration and other driving conditions on the vehicle consumption and they are described as functions of average link speed. Furthermore, Yao and Song (2013) proposes the formation of a database that is filled with dynamic traffic information. This information is predicted based on a historical database combined with real-time road traffic information collected from a probe vehicle system and it is used as input in the proposed models. Thus, the proposed models estimate the road segment fuel consumption factor that expresses the amount of fuel required per travelled distance.

Alternatively, a context-specific mesoscopic approach is considered in (Boriboonsomsin et al 2012), where the authors propose a multivariate regression model, based on vehicle characteristics, roadway characteristics, traffic characteristics and other explanatory variables, for concurrent prediction of both energy consumption and gas emissions. Even though the reported validation results suggest a reasonable estimation performance, a regression-related random error due to unexplainable causal factors is identified. In order to overcome such limitations, we suggest in our work the use of a model employing robust learning techniques for predicting vehicular consumption. Due to the learning model's inherent capability of identifying the complex patterns and nonlinearities underlying the relations between the factors that determine a specific variable, we advocate (and prove) that such a model presents outstanding performance in the problem under consideration.

Three instantaneous fuel consumption and emission models are evaluated in (Silva et al 2006), namely EcoGest (Silva et al 2006), Comprehensive Modal Emission Model (CMEM) (Scora and Barth 2006) and the Advanced Vehicle Simulator (ADVISOR) (Brooker et al 2003). These models were selected because they are capable of simulating different vehicles and can account for cold start effects, air conditioning use, and road topography. A key conclusion is that these models are more capable in estimating the variation in fuel consumption and gas emissions attributed to fluctuation in engine power demand than the variation caused by other factors. Thus, the measured accuracy in the examined case studies is limited within 10% to 20%. Another study presented in (Ahn and Rakha 2008) demonstrates comparison results between the micro-models CMEM and VT-Micro (Ahn et al 2002) and a macroscopic emission estimation tool (i.e. MOBILE6 (EPA 2002)). In particular, authors highlight the variations in the instantaneous fuel consumption rates as estimated by the VT-Micro and CMEM models, the deficiency of macroscopic environmental tools (i.e. MOBILE6) to capture them and their significance especially when determining the trip emissions.

Several studies can be found also in the literature that describe microscopic models predicting the brake specific fuel consumption (BSFC) values (i.e. the rate of fuel consumption divided by the power produced). Kara Togun and Baysec (2010) develop an explicit artificial neural network based formulation to predict torque and BSFC of a gasoline engine in terms of spark advance, throttle position and engine speed. The evaluation process is based on experimental measurements and verifies the efficiency of the proposed model. Another neural network based model is presented in (Uzun 2012). The suggested model's predictions are based on engine speed, load and Crankshaft Angel (CA) and they are found to be consistent with the corresponding experimental results. Despite the fact that such models demonstrate sufficient performance results, further calculations are needed in order to render the estimated BSFC values usable to eco-routing systems.

Following a macroscopic approach in order to build an energy consumption prediction model suitable for integration as part of an on-board eco-routing system would not be adequately efficient, as such models ignore valuable transient vehicle behaviour along a route and present inferior performance. Following a microscopic approach, on the other hand, would provide impractical instantaneous consumption information obtained after consuming valuable processing resources and time. Therefore, we propose a model that adopts a mesoscopic approach and estimates the vehicular consumption on a 'per road-segment' basis. The introduced model differs from the rest of the mesoscopic models reported previously since it considers a wider range of factors contributing in energy consumption and it follows a novel approach in identifying the relations amongst the considered factors. In particular, the list of the selected factors is not limited to historical link speed data (Minett et al 2011) and to driving conditions (Yao and Song 2013), but it is expanded with the inclusion of contextual parameters describing the traffic context, the road geometric characteristics, the weather context and the driver profile. Furthermore, the proposed model outperforms the previously reported consumption prediction models because its core functionality is based neither on rigid mapping techniques (Minett et al 2011) nor on error-prone multivariate regression techniques ((Boriboonsomsin et al, 2012)). On the contrary, the proposed model implements a learning technique that is capable of generating forecasts properly adapted to any contextual change and of adequately identifying all the nonlinearities underlying a complex process (like the vehicular energy consumption process). These advantages of the introduced approach are further analyzed and evaluated in the following.

3 Prediction Model

The introduced model is an approximator of a function f that represents the non-linear physical mechanism underlying the vehicular consumption process ($f : C \to \mathbb{R}$) and, thus, calculates the energy to be consumed for travelling through a road segment based on the current value of the contextual parameters affecting the vehicular consumption.

The learning model selected for the representation of the consumption function f is the GRNN (Specht 1991). It constitutes a memory-based network that is capable of providing estimations for continuous variables (like the vehicular consumption) and converging to nonlinear regression surfaces (such as the surface of the mechanism under consideration). A major benefit of designing and developing a GRNN is that it can adequately learn from experience using only a fraction of the training samples needed in case of other learning models (e.g. back-propagation (Rumelhart et al 1986)). Therefore, the required training dataset can be generated after performing just a few passes through the corresponding road segment. Moreover, GRNN does not converge to poor solutions corresponding to local minima of the error criterion (as sometimes happens with iterative learning techniques) and it learns

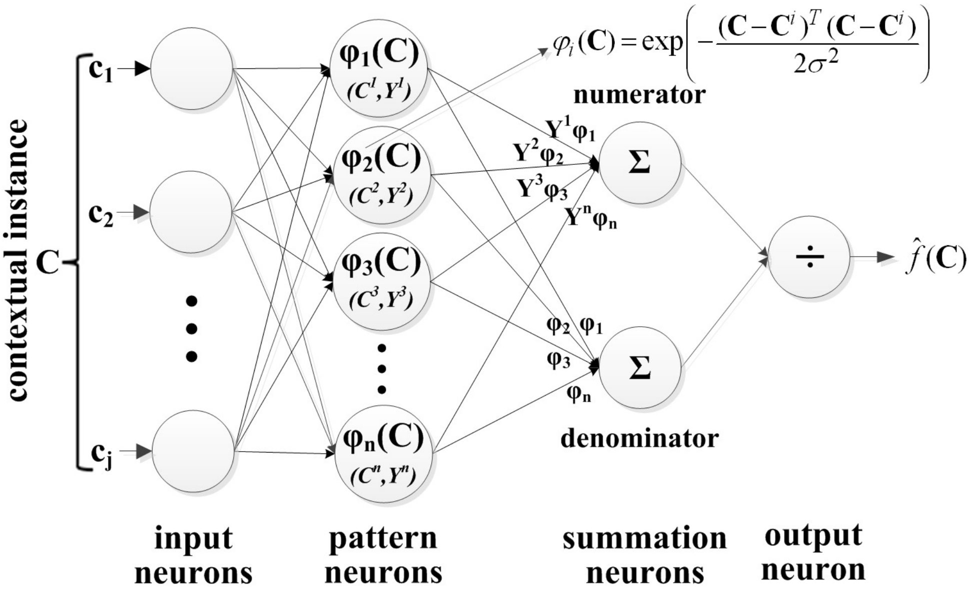


Fig. 1 GRNN architecture

in one pass through the data, thus enabling generalization from examples as soon as these are stored.

The structure of the applied GRNN (Fig. 1) consists of the input units, the pattern units, the summation units and the output unit. In the input layer, one neuron for each predictor variable is assigned with the task of standardising the input values. These standardised values are then being fed to the neurons of the first hidden layer (pattern layer), which has one unit per exemplar (record) contained in the training dataset. When presented with the normalised value of the input $\overrightarrow{c}$, each pattern neuron computes the Euclidean distance of that value from the stored vector (representing the exemplar) and, then, applies the Radial Basis Function (RBF) kernel function. The resulting values are passed to the two summation neurons, i.e. the numerator and the denominator summation units. The denominator summation unit adds up these weight values, while the numerator summation unit performs a dot product between the vector formed by these weight values and the one composed of the signals from the pattern units. The output unit, finally, divides the outputs of the summation units and yields the desired consumption estimate. According to the structure of the applied network, the proposed model is described with the following equations:

$$\hat{f}(\overrightarrow{c}) = \frac{\sum_{i=1}^{n} Y^i exp(-\frac{D_i^2}{2\sigma^2})}{\sum_{i=1}^{n} exp(-\frac{D_i^2}{2\sigma^2})} \tag{1}$$

$$D_i^2 = (C - C^i)^T (C - C^i) \tag{2}$$

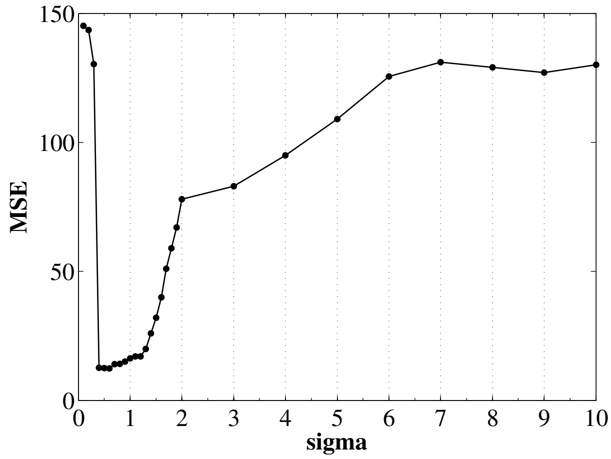


Fig. 2 Choosing the smoothness parameter σ

where $\overrightarrow{c}$ is the input vector containing the instance of the contextual parameters that determine vehicular consumption, (C^i, Y^i) is the i^{th} exemplar (representing the input-output set of the i^{th} training sample), n is the number of training samples, D_i is the distance between the i^{th} exemplar and the input vector $\overrightarrow{c}$, and σ is the smoothness parameter determining the influence range of the RBF kernel.

Choosing the smoothness parameter σ is crucial for the performance of the developed model. Specht suggests in (Specht 1991) the use of the holdout method to select the proper value of σ. For a series of distinctive values of σ, the holdout method consists in removing one sample at a time from the training dataset and constructing a network based on all of the other samples. Then, the constructed network is used to estimate Y for the removed sample. The process is repeated for each sample in the training dataset and the mean-squared error between the actual values Y^i and the corresponding estimations $\hat{Y}^i$ is finally calculated. The value of σ giving the smallest mean-squared error is proposed by the holdout method as the proper σ.

Fig. 2 presents the curve of the calculated mean-squared errors vs. the σ values. It seems that the smallest error can be achieved for a range of σ values between 0.40 and 1.20. Considering that large values of σ result in smoothing out noisy data, while smaller ones allow the estimated regression surface to be as nonlinear as required to approximate more closely the training samples and that the problem's input space is highly nonlinear, the selection of a small σ value is more appropriate. Thus, based on the presented curve the value $\sigma = 0.4$ is finally selected.

4 Measurements

In case a vehicle was moving from point A to point B inside an isolated environment, then the calculation of the vehicular energy consumption would be straightforward. Indeed, the consumed energy would be equal to the energy required to move an object of equivalent mass from point A to point B, a problem that is already solved using the law of Physics. Considering the complexities dominating both in existing road networks and in manufactured vehicles, however, such an approximation is not applicable. Modern vehicles are equipped with additional electric subsystems (e.g. heating system, wipers) that contribute not only in driving safety and user comfort but also in vehicular consumption. In addition, as the engine and the powertrain system degrade with usage, a part of the consumed energy that cannot be determined is transformed into losses. For example, in case of Fully Electric Vehicles (FEVs), the nonlinearly degrading performance of the battery either during a single charge cycle or during its whole life is responsible for such indirect losses. This type of vehicle related constraints, which affect energy consumption calculations, constitute a group of contextual parameters that in this study is called *Vehicle Context*.

Apart from the *Vehicle Context* parameters, a series of other important external factors (not related to the vehicle itself) affect also the vehicular consumption. To start with, the vehicular consumption is changing depending on the morphology of the travelled road (e.g. the consumption rises when moving uphill). Such characteristics together with all the limitations and rules imposed on users of the road infrastructure by road operators and traffic engineers (e.g. speed limits) form the so called *Road Context*. Additional influence on the travelling vehicle's consumption is caused by the presence of other vehicles. Even in the case of FEVs, whose engines consume no energy when coming at a standstill, heavy traffic conditions usually increase total consumption due to the usage of electric auxiliaries for a longer time period as well as because of the extra energy required to start moving after a standstill. The term *Traffic Context* is assigned to the parameters describing the prevailing traffic conditions. At the same time, the travelling vehicle is exposed to the weather conditions (e.g. rain, snow etc.), namely the *Weather Context* that exists in the area of interest. Finally, the driver can also be considered as a factor that affects the energy consuming processes. The term *Driver Profile* is selected for describing the driving attitude (e.g. economy driving) adopted by the user.

Thus, five groups of context parameters are identified and proposed in the present study as contributors to vehicular consumption, namely the *Vehicle Context*, the *Traffic Context*, the *Road Context*, the *Weather Context* and the *Driver Profile*. After searching for appropriate delegates from each group, i.e. parameters that adequately describe the corresponding contextual state and are retrievable by on-board navigation systems, the model's contextual input vector ended up with the following structure:

$$\overrightarrow{C} = (h_b, l_b, \overrightarrow{s}_{aux}, w_v, t_d, t_{mo}, t_{hr}, \theta_{rs}, \kappa_{rs}, T, RH, \bar{c}_d) \quad (3)$$

where h_b and l_b are the battery's state-of-health (SoH) and state-of-charge (SoC), respectively, $\overrightarrow{s}_{aux}$ is the vector describing the status of the vehicles electric auxiliaries, w_v is the vehicle's weight, t_d is the current day of the week, t_{mo} is the current month, t_{hr} is the current hour band of the day, θ_{rs} is the slope of the road segment, κ_{rs} is the class of the road segment, T is the ambient temperature, RH is the relative humidity, and $\bar{c}_d$ is the driver's average consumption rate calculated by the vehicle's trip computer.

The characteristics of the selected delegates are further analyzed in Table 1. SoH and SoC are selected due to their ability to determine the battery's nonlinear performance. Both of these parameters are provided by the FEV's battery management system (BMS) that is responsible for monitoring and controlling the operation of the vehicle's battery pack. It should be noted that the BMS also estimates the incurred energy consumption after applying the proper calculations to the monitored values of voltage (V) and current (I) ($\int_{time} V \times I$). The vector of electric auxiliaries' status ($\overrightarrow{s}_{aux}$) represents in a comprehensive manner the operational sta-

Table 1 Contextual parameters

Group	Parameter	Value range	Source
Vehicle Context	battery SoH (h_b)	0..100%	BMS
	battery SoC (l_b)	0..100%	BMS
	electric auxiliaries status ($\overrightarrow{s}_{aux}$)	vector of strings	vehicle microcontrollers
	weight (w_v)	> 0kg	user input
Traffic Context	day of the week (t_d)	Mo..Su	on-board system's internal clock
	month (t_{mo})	Jan..Dec	on-board system's internal clock
	hour band of the day (t_{hr})	(00:00-01:59)..(22:00-23:59)	on-board system's internal clock
Road Context	road slope (θ_{rs})	-100..100%	digital map
	road class (κ_{rs})	(freeway, arterial, collector, local road)	digital map
Weather Context	ambient temperature (T)	$-30..70^o$C	on-board sensor
	relative humidity (RH)	0..100%	on-board sensor
Driver Profile	avg. consumption rate ($\bar{c}_d$)	> 0Wh/km	trip computer

Table 2 Elements of the vector of electric auxiliaries' status

Parameter	Value
Lights	off-position-driving-high beam
Heating	off-low-mid-high
Air-conditioning	off-low-mid-high
Radio	off-on
Wipers	off-low-mid-high

tus of the vehicle's electrical components (Table 2) that contribute into vehicular energy consumption. Modern vehicles are equipped with several microcontrollers that manage the operation of such components and provide all the relevant information (e.g. their operational status or potential failures). The vehicle weight (w_v) is another important parameter that has a direct influence on vehicular consumption and includes both the net weight and the weight of any load (e.g. passengers, luggage). Today weight sensors are not very popular in vehicles, nevertheless it is assumed that the user can easily enter a rough estimation of the vehicle's load (number of passengers, number of pieces of luggage) into the system running on-board.

The *Traffic Context*, i.e. the traffic conditions on a particular road segment at a specific time slot, can often be retrieved from traffic information providers, i.e. third-party service providers that monitor specific parts of the road network, collect traffic data through advanced (Jianming et al 2012) or traditional techniques (Ki 2011), aggregate them and issue traffic reports based on the performed analysis. In the present study, the periodicity of the traffic conditions is captured by integrating into calculations the time-frame when the vehicle travels across the road segment. Thus, considering that a periodic trend can be detected in traffic conditions on a monthly, weekly or daily basis, the *Traffic Context* is described by the corresponding parameters (Table 1).

There are road characteristics that have a direct influence on vehicular consumption (e.g. road slope) and others that affect the vehicular consumption in an indirect manner (e.g. road class). More precisely, it is obvious that the vehicular consumption increases when moving on roads with higher slopes (more power is needed to overcome vehicle's inertia) or when travelling at very high speeds (road class defines speed limits on a particular segment). Both characteristics, i.e. the road slope and the road class, are usually stored as metadata of commercial digital maps and, therefore, can be easily retrieved by on-board navigation systems.

Although in qualitative terms the influence of the *Weather Context* on vehicular consumption can be easily understood, the development of the corresponding model is not straightforward. In case the weather conditions are characterized by low temperature and increased humidity, the user is forced to turn on several electric auxiliaries (e.g. wipers and heating) and to cause an overhead in vehicular consumption. Concurrently, the fact that the vehicle must travel in lower speeds due to safety reasons contributes into lower vehicular consumption, which might compensate for the usage of electric auxiliaries. In another case when the weather conditions are characterized by high temperature and low humidity, there could be an increase in vehicular consumption as the battery performance might deteriorate and the air-conditioning system might be switched on for cooling. These examples reveal the importance of considering the *Weather Context* as a contributor in vehicular consumption. They also reveal that the use of a deterministic model is not generally possible for vehicular energy consumption prediction and that a learning model is more appropriate instead due

Fig. 3 Field trials' area (Chieri, Turin)

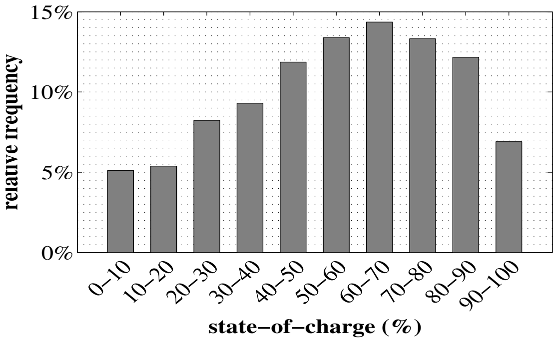
to the uncertainty of how exactly each parameter affects the final outcome (energy consumption level). For the *Weather Context*, the parameters of the ambient temperature and relative humidity have been selected in the present study, and can be easily retrieved from onboard sensors as well as by parsing weather information available on Internet websites.

The fifth group of context parameters, namely the *Driver Profile*, refers to a subjective concept called the user's driving attitude that describes the user's aggressiveness in driving and the user's adaptability in different contextual instances (e.g. driving on wet or dry road). Thus, the *Driver Profile* is detected as one of the determinants of vehicular consumption. As the user's driving attitude is not fixed and evolves over time based on accumulated experiences, driving skills or even mood, the *Driver Profile* should be considered each time the vehicular consumption is estimated. Having all these in mind we concluded that the average consumption rate calculated by the vehicle's trip computer is a proper indicator of the *Driver Profile*. Indeed, the vehicle's trip computer calculates continuously the vehicle's current average consumption rate based on the latest consumption measurements, while the divergence between this calculated average consumption rate and the average consumption rate reported in the vehicle's specifications reveals the user's current driving attitude (e.g. if the trip computer reports an average consumption rate that is considerably higher than the one anticipated from the manufacturer's specifications, then the user can be assumed to have a more aggressive driving style, whereas a more energy-efficient driving style may be presumed in case that the opposite holds).

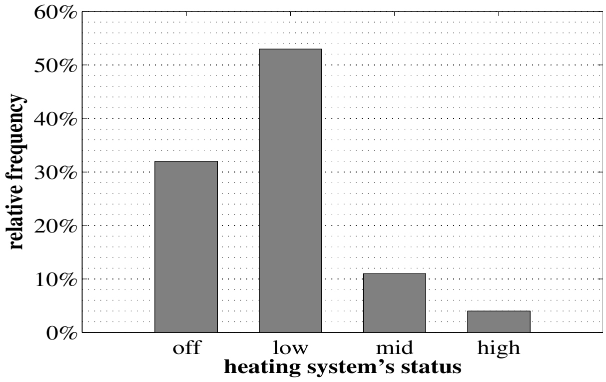
Table 3 Nido EV specifications

Characteristic	Value
Type	City Car 2 seats
Electric drive	Rear
Measurements (L/W/H)	2950/1620/1507 mm
Acceleration 0-60 km	4.4 sec
Top speed (limited electronically)	120 km/h
Range fully charged	140 km
Weight empty	840 kg
Engine	Permanent magneto synchronous
Max power output	60 kW
Peak torque to wheels	90 Nm
Drive batteries	Li-Ion batteries
Rated voltage	350 V
Rated capacity	22 kWh
Charge time	8 hours
Battery weight	150 kg

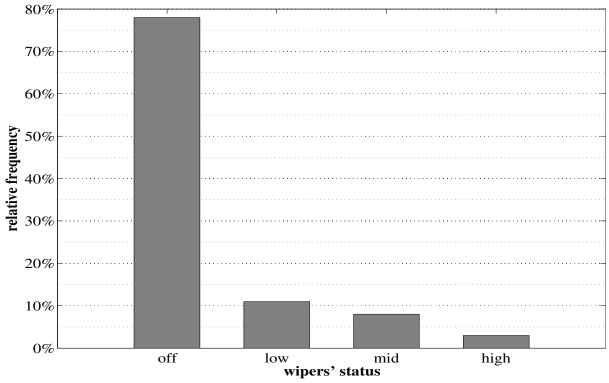
In order to render the proposed model functional, a data acquisition campaign was carefully planned and conducted using a FEV. The employed FEV was developed by the Italian car designer and manufacturer Pininfarina (Pininfarina 2013). The specifications of this vehicle (named Nido EV) are depicted in Table 3. It should be stated that this vehicle is equipped with a regenerative braking system, i.e. an energy recovery mechanism capable of converting vehicle's kinetic energy into electric energy during breaking. The amount of recovered energy is measured as negative current ($I < 0$) consumption by the BMS and this is taken im-



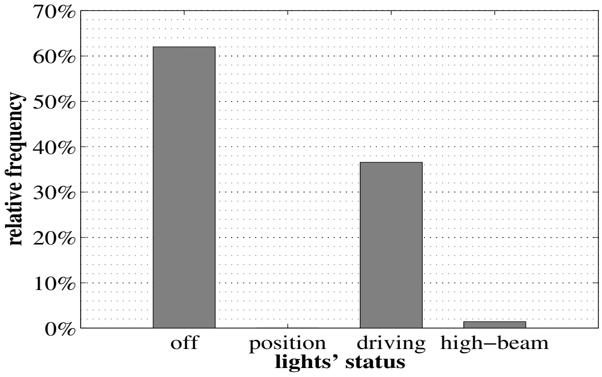
(a) Battery SoC



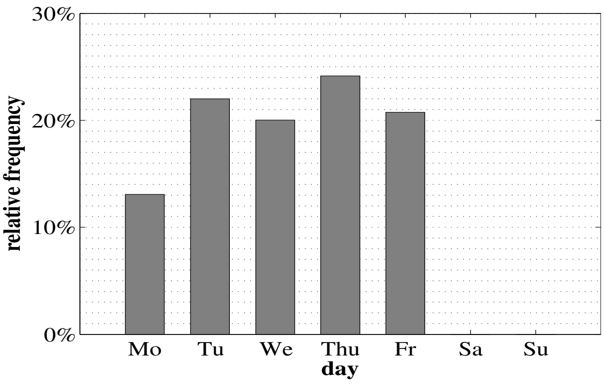
(b) Heating system's status



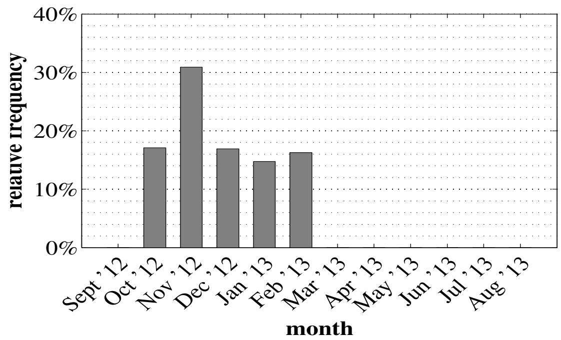
(c) Wipers' status



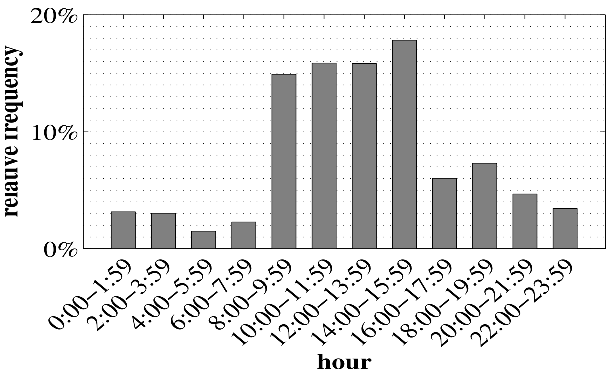
(d) Lights' status



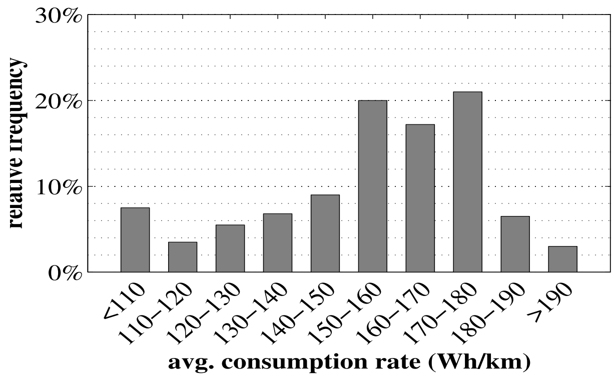
(e) The day of the week



(f) The month of the year

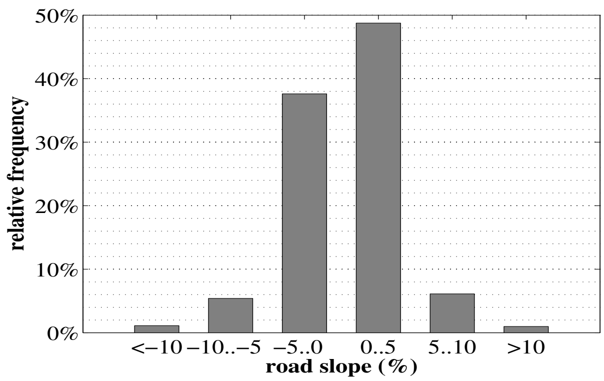


(g) The hour band of the day

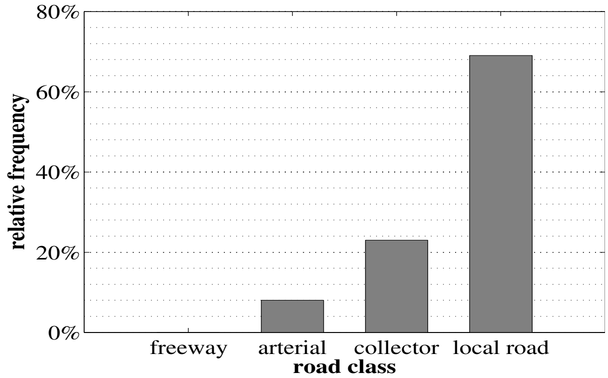


(h) The average consumption rate reported by the trip computer

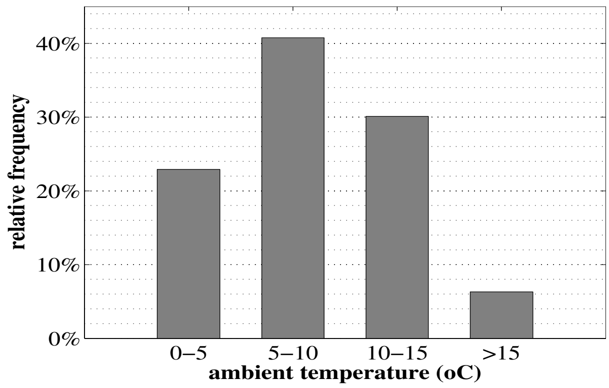
Fig. 4 Relative frequency distribution of contextual parameters' values recorded during the data acquisition campaign (*Vehicle Context*, *Traffic Context* and *Driver Profile*)



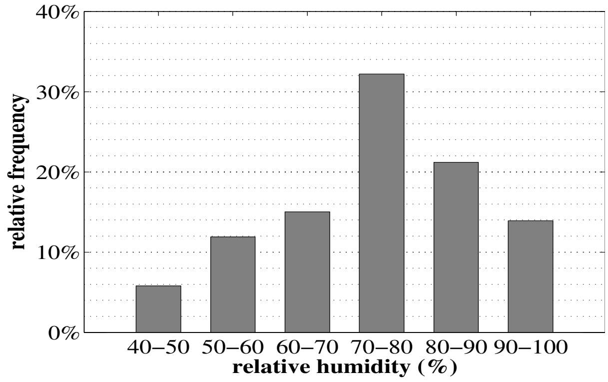
(a) Road slope



(b) Road class



(c) Ambient temperature



(d) Relative humidity

Fig. 5 Relative frequency distribution of contextual parameters' values recorded during the data acquisition campaign (*Road Context* and *Weather Context*)

plicitly into account during the energy consumption calculation ($\int_{time} V \times I$). During this road campaign, which took place in the town of Chieri (Turin, Italy) (Fig. 3), the FEV travelled approx. 1275km and gathered data for training the distinctive consumption models of 2436 road segments. The collected data were recorded by an on-board small form factor computer connected both to external sensors (i.e. GPS, temperature and humidity sensors) and to the vehicle's controller area network bus (CAN-bus) socket. The connection to the CAN-bus socket is realized through a commercially available tranceiver (Vector 2013), providing access to the BMS's, the vehicle microcontrollers' and the trip computer's data (Table 1).

The performed road campaign was not totally random but was rather based on a predefined plan. According to this plan, the FEV would travel during specific time windows through certain parts of the town without needing to follow any specific route. Our purpose was to form an adequate training dataset (containing the most representative exemplars) in a short period, considering that the majority of the selected contextual parameters (Table 1) cannot be controlled and that the availability of resources is limited. The satisfaction of these prerequisites is depicted in the relative frequency distribution diagrams of the selected features that were produced based on the training dataset (Figs. 4 & 5).

Fig. 4 presents the relative frequency distribution of the contextual parameters belonging to the *Vehicle Context*, the *Traffic Context* and the *Driver Profile*. The relevant information is provided for all contextual parameters except for part of the *Vehicle Context*, namely the frequency distribution information regarding the battery SoH, the air-conditioning status, the radio status and the vehicle weight are missing. The reasons for this are as follows: the battery SoH presents an insignificant change during the short-period data acquisition campaign, the air-conditioning was not used due to low environmental temperatures, the specific test vehicle is not equipped with a radio, and the changes of the vehicle weight are negligible (the same driver was always employed with no additional passenger or luggage during the tests). The relative frequency distribution of the rest of the contextual parameters, namely the ones belonging to the *Road Context* and the *Weather Context* are depicted in Fig. 5. Fig. 6, on the

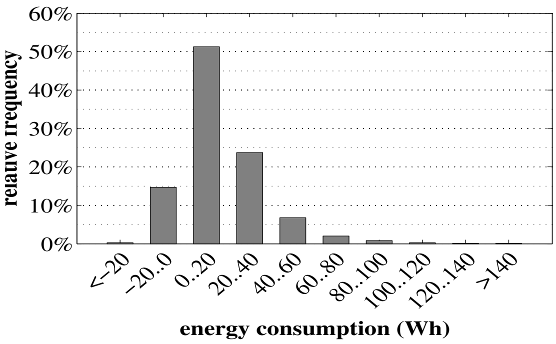


Fig. 6 Relative frequency distribution of the energy consumption's values recorded during the data acquisition campaign

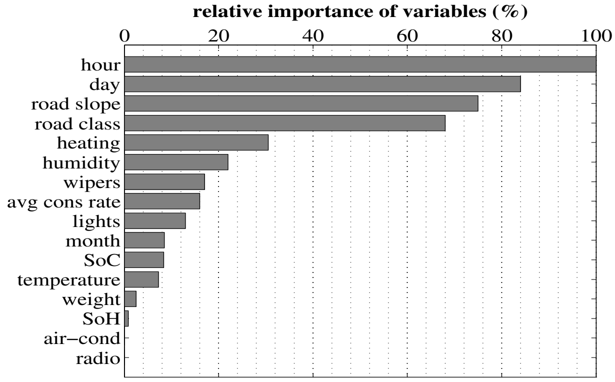


Fig. 7 Relative importance of variables

other hand, presents the relative frequency distribution of the energy consumption's values recorded during the data acquisition campaign. The negative consumption values correspond to measurements recorded while travelling downhill and are due to energy recuperation incurring while breaking. To sum-up, Figs. 4-6 provide a complete and thorough overview of the training dataset collected on-road.

5 Experimental Results

The first road campaign was dedicated to the generation of the training dataset. Based on this dataset a series of GRNN models capable of forecasting the energy consumption across the visited road segments are produced. Before proceeding to the validation of the forecasting models, we attempt to quantify the contributions of the predictor variables in the network. Several methodologies have been developed for determining the significance of input parameters, usually called "the importance of variables", in case of artificial neural networks (Olden et al 2004). The "input perturbation" method initially described by Scardi and Harding Jr. (1999) is applied in the present study. As expected, according to the results of this method, which are depicted in Fig. 7, the parameters of the vehicle weight, the battery SoH, the air-conditioning system and the radio are insignificant in the developed model. Such a conclusion was anticipated considering the distribution of the intensities of these features that were recorded during the data acquisition campaign. More specifically, the vehicle weight was almost the same, the health of the brand new battery was almost intact, the air-conditioning system was not used at all, and there was no radio equipment on-board. However, these features should not be discarded in

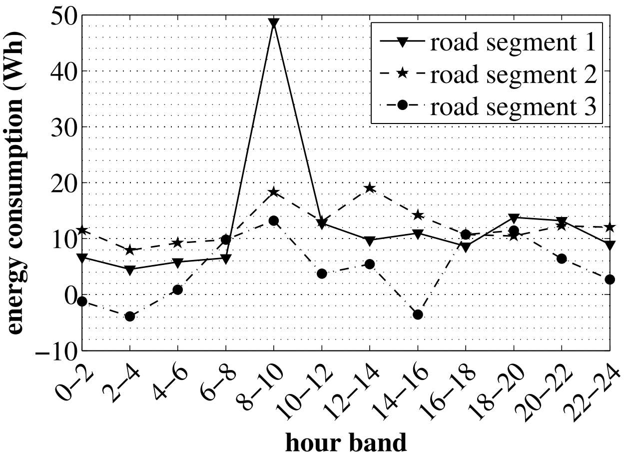


Fig. 8 The change of energy consumption incurred on multiple hour bands

the general model, because their significance could be substantial in the general case (e.g. in other tests or in commercial applications).

The most significant parameter, on the other hand, is the parameter 'hour band'. In order to study the change in energy consumption against the change in that parameter we generate the three line graphs depicted in Fig. 8 based on a subset of the monitored samples of three distinctive road segments. Although the displayed graphs present some trends (higher consumption values during rush hours and lower consumption values during night), the changes in energy consumption values cannot be attributed solely to the change in hour-band as the selected samples present differences in other contextual parameters too. Such differences are expected when the collection campaign is performed in a real road network, namely in an uncontrolled environment that does not allow the enforcement of any restrictions or rules. Table 4 reports the details of the samples forming the line graph of the

Table 4 Recorded contextual instances referring to "road segment 1" (slope=0.4%, class='local road')

hour band	energy cons (Wh)	SoH (%)	SoC (%)	electric aux (lights/ heat/airc/rad/wip)	weight (kg)	day	month	temp (^oC)	humi (%)	avg cons rate (Wh/km)
00:00-01:59	6.71	97	84	driv/mid/off/off/off	923	We	10	13	67	135
02:00-03:59	4.51	97	77	driv/low/off/off/off	923	We	10	12	84	128
04:00-05:59	5.82	97	63	driv/mid/off/off/off	923	We	10	10	85	123
06:00-07:59	6.53	97	42	driv/low/off/off/off	923	We	10	14	67	125
08:00-09:59	48.72	97	99	off/off/off/off/off	923	Tu	10	16	69	194
10:00-11:59	12.77	97	59	off/off/off/off/off	923	Tu	10	16	69	171
12:00-13:59	9.75	97	47	off/off/off/off/off	923	Tu	10	16	69	152
14:00-15:59	10.97	97	18	off/mid/off/off/off	923	Tu	10	5	69	152
16:00-17:59	8.65	97	82	off/mid/off/off/off	923	Thu	10	1	58	144
18:00-19:59	13.83	97	69	driv/low/off/off/off	923	Thu	10	10	75	158
20:00-21:59	13.21	97	55	driv/low/off/off/off	923	Thu	10	10	82	158
22:00-23:59	8.97	97	32	driv/off/off/off/off	923	Thu	10	12	89	134

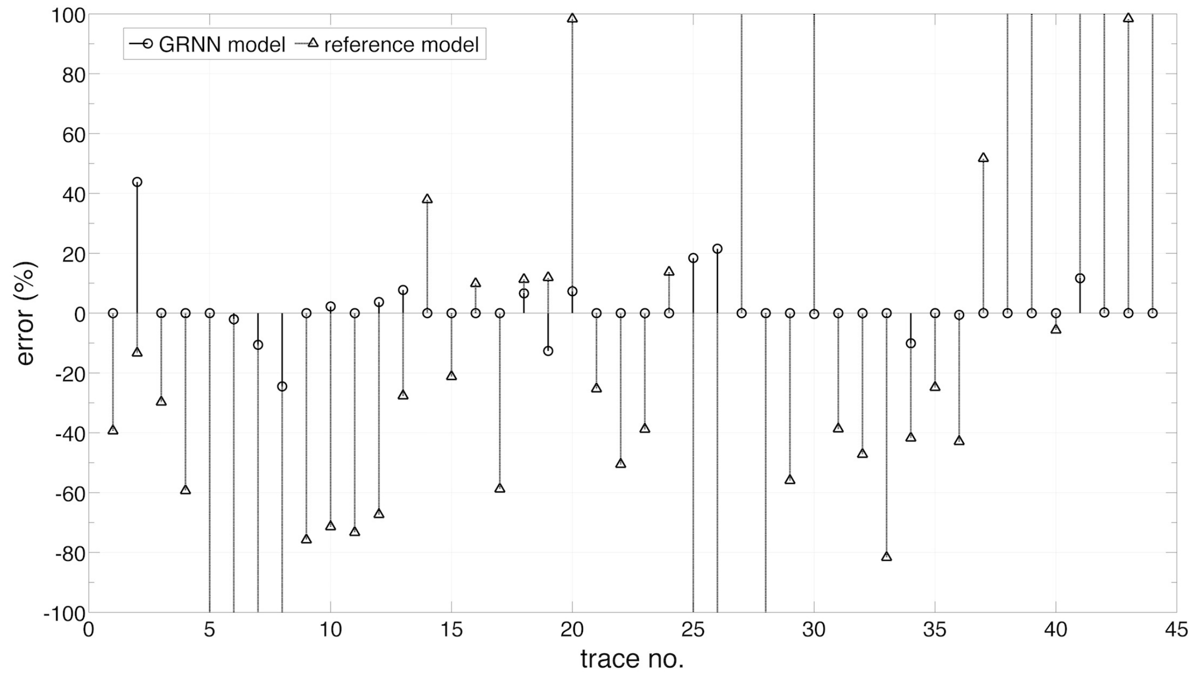


Fig. 9 Comparative diagram of the estimation error for the proposed GRNN model and the reference model

segment called 'road segment 1' and reveals their differences.

A second road campaign took also place in the same area using the same FEV. The purpose of the latter campaign, which spanned approx. 551km, was to perform the validation of the proposed model, i.e. to travel through a random set of validation routes and to record the incurred energy consumption together with the corresponding contextual parameters. The recorded real consumption values were, then, compared, firstly, against the ones estimated by the trained GRNNs, and, secondly, against the ones estimated by a reference model.

As a reference model for comparison purposes we have considered a model that estimates the vehicular consumption per road segment based on the segment length and the vehicle's average consumption rate (per kilometer). The value of the average consumption rate can be defined by the car manufacturer either through field tests specified by relative standards (e.g. (SAE 2012)) or through advanced consumption models (Wu and Liu 2011). In the present study, we consider the average consumption rate that occurs dividing the battery's rated capacity with the vehicle's nominal range ($22kWh \div 140km \approx 157Wh/km$). The values used are the ones provided by the manufacturer in the vehicle's specifications (Table 3). Unfortunately, there are no further details available on the method followed for the vehicle's range estimation.

Based on the collected measurements, the performance indicators of the proposed and the reference

Table 5 Performance indices based on the collected measurements

Model	MPE	MAPE
GRNN model	1.62%	3.96%
reference model	-66.07%	189.59%

model are generated (Table 5). The mean percentage error (MPE) of the proposed model is 1.62%, which means that it is rather unbiased, while the mean absolute percentage error (MAPE) is 3.96%. On the other hand, the MPE of the reference model is -66.07% (severe negative bias, i.e. underestimation of the energy consumption) and the MAPE is 189.59%. This means that our model achieves an improvement for the MPE of more than 50 times over the reference model, and of more than 45 times for the MAPE. Furthermore, by underestimating the amount of energy required to reach the desired estimation might cause energy shortage problems, especially in the case of FEVs.

To further analyze our findings, we have constructed a comparative diagram (Fig. 9), where the horizontal axis represents a series of distinctive validation traces (i.e., different validation routes performed in different contextual conditions) and the vertical axis represents the percentage error of the corresponding predictions performed by the two separate models. Due to page size restrictions the vertical axis range is limited up to $\pm 100\%$ and the diagram depicts only a fraction of the validation results (the estimation errors of 44 traces). According to this diagram, the proposed model provides only few estimations deviating more than $\pm 10\%$ from the real values, in contrast to the reference model, where a significant number of traces exceed even the range of $\pm 100\%$. The superiority of the developed model is justified by the fact that, unlike the reference model, it can predict more accurately both any potential negative consumption (i.e. the FEV may generate energy while braking) as well as any deviation in consumption when travelling through the same route in different contextual conditions (e.g. different traffic conditions).

6 Conclusion

An innovative learning-based context-aware approach to the problem of forecasting a vehicle's consumption along a probable route is discussed in the present paper. The contextual parameters that contribute into vehicular consumption are identified and proper delegates are selected for the model implementation. The learning functionality of the proposed model is implemented through a specifically-designed GRNN, a tool that has been found to be suitable because of its fast learning ability and due to its capability of convergence in a reasonable regression surface using a limited number of samples.

In order to validate the developed model, a road campaign was conducted using a properly equipped FEV. The testing FEV travelled along several routes and collected measurements both for training the model and for validating its performance. The reported training results demonstrate the significance of the selected contextual parameters, while the validation results establish the reliability of the developed model. Indeed, the measured average estimation error of 3.96% proves that the proposed model is quite accurate and provides reliable forecasts. In addition, the proposed model outperforms an existing reference model that was used for comparison, as the measured average estimation error of 3.96% is found to be approx. 45 times better than the corresponding error of the reference model. This discrepancy is attributed to the inability of the reference model to sense the environment (context-unawareness) (e.g. it estimates the same value of energy consumption for two road segments of equal length regardless of their inclination, traffic or weather conditions, etc.) and to perceive the energy recuperation incurred while braking. On the contrary, by implementing a learning mechanism (i.e. a GRNN), the proposed model is capable of incorporating the current contextual conditions in its energy consumption estimations and of dealing with negative values of energy consumption (i.e. values originating from regenerative breaking). Hence, these findings establish the robustness of the proposed model against the reference model.

Acknowledgements This work has been performed under the FP7 314151 project EMERALD, which has received research funding from the European Union's Seventh Framework Programme. This paper reflects only the authors' views, and the Community is not liable for any use that may be made of the information contained therein.

References

Ahn K, Rakha H (2008) The effects of route choice decisions on vehicle energy consumption and emissions. Transportation Research Part D: Transport and Environment 13(3):151–167. doi: 10.1016/j.trd.2008.01.005

Ahn K, Rakha H, Trani A, Van Aerde M (2002) Estimating Vehicle Fuel Consumption and Emissions based on Instantaneous Speed and Acceleration Levels. Journal of Transportation Engineering 128(2):182–190. doi: 10.1061/(ASCE)0733-947X(2002)128:2(182)

Ben Dhaou I (2011) Fuel estimation model for ECO-driving and ECO-routing. In: Intelligent Vehicles Symposium (IV), 2011 IEEE, pp 37–42. doi 10.1109/IVS.2011.5940399

Boriboonsomsin K, Barth MJ, Zhu W, Vu A (2012) Eco-Routing Navigation System Based on Multisource Historical and Real-Time Traffic Information. IEEE Transactions on Intelligent Transportation Systems 13(4):1694–1704. doi: 10.1109/TITS.2012.2204051

Brooker A, Haraldsson K, Hendricks T, Johnson V, Kelly K, Kramer B, Markel T, O'Keefe M, Sprik S, Wipke K, Zolot M (2003) ADVISOR Documentation.

EPA (2002) Users Guide to Mobile: Mobile Source Emission Factor Model. U.S. Environmental Protection Agency

Jianming H, Qiang M, Qi W, Jiajie Z, Yi Z (2012) Traffic congestion identification based on image processing. IET Intelligent Transport Systems 6(2):153–160. doi: 10.1049/iet-its.2011.0124

Kamal MAS, Mukai M, Murata J, Kawabe T (2011) Ecological Vehicle Control on Roads With Up-Down Slopes. IEEE Transactions on Intelligent Transportation Systems 12(3):783–794. doi: 10.1109/TITS.2011.2112648

Kara Togun N, Baysec S (2010) Prediction of torque and specific fuel consumption of a gasoline engine by using artificial neural networks. Applied Energy 87(1):349–355. doi:// 10.1016/j.apenergy.2009.08.016

Ki YK (2011) Speed-measurement model utilising embedded triple-loop sensors. IET Intelligent Transport Systems 5(1): 32–37. doi: 10.1049/iet-its.2008.0084

Minett CF, Salomons AM, Daamen W, Van Arem B, Kuijpers S (2011) Eco-routing: Comparing the fuel consumption of different routes between an origin and destination using field test speed profiles and synthetic speed profiles. In: 2011 IEEE Forum on Integrated and Sustainable Transportation System (FISTS) 32–39. doi: 10.1109/FISTS.2011.5973621

Olden JD, Joy MK, Death RG (2004) An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. Ecological Modelling 178(34):389–397. doi: 10.1016/j.ecolmodel.2004.03.013

Pachernegg S (1969) A Closer Look at the Willans-Line. SAE Technical Paper. doi: 10.4271/690182

Pininfarina (2013) Pininfarina Nido EV. http://www.pininfarina.com/en/nido_ev. Accessed 8 January 2014

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323(6088):533–536. doi: 10.1038/323533a0

SAE (2012) Electric Vehicle Energy Consumption and Range Test Procedure. http://standards.sae.org/j1634_201210/. Accessed 8 January 2014

Scardi M, Harding Jr LW (1999) Developing an empirical model of phytoplankton primary production: a neural network case study. Ecological Modelling 120(23):213–223. doi: 10.1016/S0304-3800(99)00103-9

Scora G, Barth M (2006) Comprehensive Modal Emissions Model (CMEM) User's Guide, version 3.01. University of California, Riverside Center for Environmental Research and Technology. http://www.cert.ucr.edu/cmem/docs/CMEM_User_Guide_v3.01 Accessed 8 January 2014

Silva CM, Farias TL, Frey HC, Rouphail NM (2006) Evaluation of numerical models for simulation of real-world hot-stabilized fuel consumption and emissions of gasoline light-duty vehicles. Transportation Research Part D: Transport and Environment 11(5):377–385. doi: 10.1016/j.trd.2006.07.004

Specht DF (1991) A general regression neural network. IEEE Transactions on Neural Networks 2(6):568–576. doi: 10.1109/72.97934

Uzun A (2012) A parametric study for specific fuel consumption of an intercooled diesel engine using a neural network. Fuel 93(0):189–199. doi: 10.1016/j.fuel.2011.11.004

Vector (2013) Vector Box VN1610 device. http://vector.com/vi_vn1600_en.html. Accessed 8 January 2014

Wu JD, Liu JC (2011) Development of a predictive system for car fuel consumption using an artificial neural network. Expert Systems with Applications 38(5):4967–4971. doi: 10.1016/j.eswa.2010.09.155

Wu JD, Liu JC (2012) A forecasting system for car fuel consumption using a radial basis function neural network. Expert Systems with Applications 39(2):1883–1888. doi: 10.1016/j.eswa.2011.07.139

Yao E, Song Y (2013) Study on Eco-Route Planning Algorithm and Environmental Impact Assessment. Journal of Intelligent Transportation Systems 17(1):42–53. doi: 10.1080/15472450.2013.747822