Machine-Learning Methodology for Energy Efficient Routing

M. Masikos[1*], K. Demestichas[1], E. Adamopoulou[1], M. Theologou[1]

[1] Institute of Communication and Computer Systems, Heroon Polytechneiou 9, Zografou–Athens, Greece, Tel: +30 210 772 1478, email: mmasik@telecom.ntua.gr

Abstract: Eco-driving assistance systems encourage economical driving behaviour and support the driver in optimizing his/her driving style to achieve fuel economy and, consequently, emission reductions. Energy efficiency is also one of the most pertinent issues related to the autonomy of Fully Electric Vehicles (FEVs). This paper introduces a novel methodology for energy efficient routing, based on the realization of dependable energy consumption predictions for the various road segments constituting an actual or potential vehicle route, performed mainly by means of machine-learning (ML) functionality. This proposed innovative methodology, the functional architecture implementing it, as well as demonstrative experimental results are presented in this paper.

Keywords: energy efficiency, routing, machine-learning, consumption prediction

1. Introduction

Eco-driving is a widely known mechanism used to promote fuel (and, generally, energy) efficiency through the use of driver behaviour feedback [1]-[2]. The efficiency of such systems has been, recently, enhanced through the adoption of vehicle-to-vehicle and vehicle-to-infrastructure (V2X) technologies [3]-[4], and the consequent exploitation of information regarding the state of surrounding vehicles and road conditions. These on-trip eco-assisting systems have been, recently, complemented by the "Eco-Routing" concept; described as a pre-trip feature that produces route-based energy consumption estimations between an origin and a destination, while searching for the optimal route solution that minimizes the total route energy consumption [5]. It is argued that systems supporting Eco-Routing could be an advantageous addition to existing eco-driving feedback mechanisms by providing preemptive route-based energy consumption estimations, therefore contributing to an overall improved driving behaviour. The results of the survey performed in [6] strengthen this perspective, as it is reasoned that the majority of European car drivers welcome the deployment of both pre-trip and on-trip eco-assistance systems.

This paper introduces a novel approach for energy efficient routing that is based on experience collection and on the application of ML methodology [7], which can be integrated in modern Advanced Driver Assistance Systems (ADASs). In Section 2, an overview of the scientific state-of-the-art is presented. Section 3 discusses the formulation of the proposed solution, while Section 4 proposes an appropriate system architecture. A prototype was further developed, based on this architecture, in order to perform appropriate validation tests. The results of these tests are reported and analyzed in Section 5. Finally, Section 6 summarizes the achievements described in the present paper and concludes with future work plans.

2. State-of-the-art

The effects of route choice in reducing fuel consumption and emissions consists an issue addressed by several authors. In [8], the impact of digital map attributes on the calculation of an eco-route is investigated through the development of a research tool in which historical link speed data are used as a basis for replicating vehicle speed profiles, enabling the calculation of fuel costs per link. Results show that it is worth taking into account map attributes for the generation of synthetic speed profiles. In [9]-[10], the authors propose an ecological route search system that advises drivers on fuel-minimizing routes. The ecological route search system consists of both fuel consumption prediction technology and route search technology. The first step was to build a fuel consumption model, as this is essential for performing an ecological route search. The results from this model were compared with actual fuel consumption for different traffic conditions and geography. The fuel consumption model proposed in [11] considers also the same consumption explanatory variables; i.e., geography is considered in estimating the actual power needed to overcome driving resistance of each link and traffic is considered through the link travel speeds and the congestion volume/capacity ratios. This set of explanatory variables is further populated in [12], where the authors also consider vehicle and driver characteristics for fuel consumption estimation. However, as stated by the authors, there is an upper limit in their estimation accuracy due to the error aggregated during the construction of their model.

In [5], [13]-[14], [15]-[17], field experiments were conducted and a wide range of models (macroscopic, microscopic, mesoscopic) were applied to evaluate the impact of route selection in terms of emissions and energy use, over several case-studies. The majority of these studies have concluded that route choice has a significant impact on both emissions and energy use. On one hand, certain studies point out that time minimization paths, often, also minimize energy use and emissions [10], [15], while, on the other hand, some other research activities have demonstrated that, frequently, the faster alternatives cannot be considered as best from the environmental perspective [8], [13]-[14].

All these considered, the present paper provides a valuable addition to the scientific and technical state-of-the-art, since: (i) It describes a new ML based methodology for energy consumption prediction, based on actual consumption experience accumulated by the vehicle in an ongoing manner. In contrast, most of previous relevant research studies have typically relied on the definition and verification of consumption models (sets of equations), as opposed to actual experience and learning. (ii) The exploitation of context information is carried out through learning, and not through a deterministic, equation-based manner; having the advantage of being able to better infer hidden consumption patterns that are otherwise complex to detect and represent through sets of equations. (iii) Focus is given on the energy consumption prediction and eco-routing of FEVs, while previous research has typically concentrated on internal combustion engine vehicles.

3. Machine-Learning for Energy Efficiency Routing

This paper proposes an alternative way to extend the default range of a FEV (as defined by its manufacturer), by selecting the most energy efficient route. Assuming that the driver wants to travel from the origin point *O* to the destination point *D*, there are $|R^{OD}|$ possible routes connecting these two points. The amount of energy required to travel the route $r^{OD} \in R^{OD}$ is calculated by the function *en(r)*. Our goal is to find the optimal route r^{OD}_{OPT} that involves the minimum energy consumption while travelling from *O* to *D*:

$$en(r) = \min\left\{en(r^{OD}_{OPT}) | r^{OD}_{OPT} \in R^{OD}\right\} \qquad (1)$$

The road network can be represented by a digraph $\boldsymbol{G}=\{\boldsymbol{V}, \boldsymbol{S}, \boldsymbol{w}\}$ where $\boldsymbol{V}$ is the set of the road network junctions, $\boldsymbol{S}$ is the set of the road segments (delimited by consecutive junctions) and $\boldsymbol{w}$ is the table of the costs assigned to these road segments. Herein, we assume that there is no cost or utility associated with the vertices of the network. As a result, the energy consumption of a route can be further decomposed as the sum of the energy consumption needed to travel through each road segment $s_i(r^{OD})$ of the route r^{OD}:

$$en\left(r^{OD}\right) = en\left(\sum_{i=1}^{N\left(r^{OD}\right)} s_i\left(r^{OD}\right)\right) = \sum_{i=1}^{N\left(r^{OD}\right)} en(s_i\left(r^{OD}\right)) \qquad (2)$$

where $N(r^{OD})$ returns the number of elements in set r^{OD}. Furthermore, the energy consumption needed to travel through a road segment can be considered as the sum of two factors. The first factor (en_T) is the expected part of the energy consumption that can be predicted based on the previous knowledge of the amount of energy needed to cross this road segment and on the current context (c_0). The second factor (en_R) is the energy overhead due to unexpected traffic events that may occur at time t_0. Based on this assumption, equation (2) can be further analyzed as:

$$en\left(r^{OD}\right) = \sum_{i=1}^{N\left(r^{OD}\right)} \left(en_T(s_i\left(r^{OD}\right), c_0\right) + en_R(s_i\left(r^{OD}\right), t_0)) \qquad (3)$$

According to this analysis, the problem at hand is formulated as an instantiation of the famous shortest path problem (eq. (1)). Several efficient algorithms have been proposed for solving this problem; thus, the present paper does not envision to work on altering these algorithms. Instead, this paper's contribution -partially described in formal terms by equation (3)- is about a new approach in estimating the energy consumption costs that are associated with the road segments of a route under investigation. The proposed approach intends to take into account all factors that affect the energy cost associated with a road segment, by integrating the influence of both expected and unexpected factors. By being able to reliably predict the total energy costs of individual road segments, it is then possible to find the optimal route based on existing shortest path algorithms (e.g., Dijkstra, Bellman-Ford).

This paper does not elaborate on the estimation process of the contribution of the unexpected traffic events on the energy consumption (en_R). Specifically, the percentage of the energy consumption that can be attributed to expected conditions is significantly larger than the part related to unexpected traffic events (trial results presented in Section 5 indicate that this percentage is approximately 87%). It can also be assumed that the value of the unexpected part can be estimated and made known by a third party Traffic Information Provider. Such an estimation mechanism can be implemented in a manner similar to that for detecting travelling time deviations. A central traffic information system may process energy and travelling time "exception" reports received from travelling vehicles. In particular, in case a vehicle detects an overspending of energy after passing through a road segment (i.e., an actual expenditure on the segment significantly higher than the predicted one), it raises an exception with the overspending quota. The statistical processing of the quotas (at the central traffic information system) corresponding to a particular road segment may reveal the existence or not of a traffic event and their values can provide an estimation of the additional energy required. The following analysis, however, focuses mainly on developing a mechanism suitable for on-board installations and efficient in accurately estimating the expected part of the energy consumption (eq. (3)).

Trying to analyze the factors that determine the expected energy consumption of a vehicle while travelling through a road segment, five vectors of context variables can be identified as the major contributors to the consumed energy, namely the *Vehicle Context* ($\boldsymbol{c^V}$), the *Weather Context* ($\boldsymbol{c^W}$), the *Traffic Context* ($\boldsymbol{c^{Tr}}$), the *Road Segment Context* ($\boldsymbol{c^S}$), and the *Driver Profile* ($\boldsymbol{c^{Dr}}$). Vector $\boldsymbol{c^V}$ describes the vehicle characteristics that seriously affect the energy consumption, like the ones reported in Table 1. Vector $\boldsymbol{c^W}$ includes the weather metrics that have an impact on the consumed energy, such as temperature and humidity. Regarding vector $\boldsymbol{c^{Tr}}$, a number of methodologies have been developed [18]-[19], in order to quantify the traffic status being encountered on a road segment in a specific time slot. The study of the observed traffic volume on a specific road segment (after removing the effects of unexpected traffic events) versus the time usually reveals the existence of a periodic trend. In order to take this traffic periodicity into account, we identify several time slots (Monday-Sunday, from 00:00 to 24:00 o'clock, January-December) and we consider the Traffic Context at a particular time point as an unknown function of the corresponding time slot. Vector $\boldsymbol{c^S}$ includes the characteristics of the road segment affecting the energy consumption in a direct (slope, traffic lights) or indirect manner (speed limit, number of lanes). Finally, regarding $\boldsymbol{c^{Dr}}$, the metric of the actual average consumption (which can be retrieved from the vehicle's trip computer) is used to describe the dependency of the energy consumption on the driver's profile. This metric is suitable for revealing the user's driving attitude (more or less aggressive) when compared to the corresponding average consumption (nominal) value provided by the vehicle's manufacturer.

Full identification of the relations and dependencies among these five groups of context variables is probably impossible to reach. As an indicative example, in the case of the Vehicle Context, the use of electric auxiliaries, like the heating system or the windshield wipers, affect the energy consumption. The use of these auxiliaries implies the existence of a specific Weather Context, i.e. low temperature and high humidity (rain). Such a Weather Context, in turn, usually implies that vehicles move in low speeds for safety reasons and that the driver's attitude becomes less aggressive. Another indirect effect of such a context is the occurrence of traffic congestion, which depends also (although not exclusively) on the class of the road segment (e.g. highway, urban street, or other type of road segment). This example demonstrates in a comprehensible way some of the complex relations that exist among the variables of the context vectors defined above.

It becomes obvious that the identification and modeling of the aforementioned relations and dependencies, so as to build an efficient and accurate mathematical formula, is not straightforward. Thus, this paper proposes the application of ML functionality for predicting the energy costs of the various road segments constituting an actual or potential vehicle route (candidate segments). ML algorithms are capable of automatically detecting patterns or regularities that underlie in a complex process like the energy consumption estimation.

ML functionality is applied through the use of so-called "Machine-Learning Engines" (MLEs). The operation of these engines typically consists of two steps: the learning or training process, and the scoring or decision process. During the first step, the MLE is provided with a set of historical data so as to "learn" how to produce reliable predictions for -yet unseen- situations.

Based on this analysis, the prediction function of the expected energy consumption (eq. (3)) could be expressed as:

$$E_T = en_T(s_i, \boldsymbol{c}_0) = en_T(s_i, \boldsymbol{c}_0^{\boldsymbol{V}}, \boldsymbol{c}_0^{\boldsymbol{W}}, \boldsymbol{c}_0^{\boldsymbol{Tr}}, \boldsymbol{c}_0^{\boldsymbol{Dr}}) \qquad (4)$$

Term $\boldsymbol{c}^{\boldsymbol{S}}$ is taken implicitly into account through term s_i (denoting the road segment identification in a map database) so it is omitted from the function parameters. What we need is to feed a MLE with the appropriate dataset (s_i, $\boldsymbol{c}^V$, $\boldsymbol{c}^{\boldsymbol{W}}$, $\boldsymbol{c}^{\boldsymbol{T}}$, $\boldsymbol{c}^{\boldsymbol{Dr}}$, E_T) and train it, so as to identify the process *$en_T()$*. Table 1 provides an analysis of the variables of the contextual vectors of eq. (4), while Table 2 depicts a small part of an example dataset with field measurements.

Table 1. Analysis of the employed contextual variables

Vector	Variable	Abbreviation	Type	Range
$\boldsymbol{c}^S$	Road segment ID	LinkId (s_i)	Unsigned integer	>0
$\boldsymbol{c}^V$	Battery State-of-Health	SoH	Integer	0…100%

	Battery State-of-Charge	SoC	Integer	0…100%
	Battery Capacity	Capacity	Integer	>0Wh
	Battery Technology	Technology	String	e.g. Li-Ion
	Lights	Lights	String	off-park-low-high
	Heating	Heat.	String	off-low-mid-high
	Air-conditioning	Airc.	String	off-low-mid-high
	Radio	Radio	String	on-off
	Wipers	Wipers	String	off-low-mid-high
	Motor's maximum power output	maxPowerOutput	Integer	>0kW
	Vehicle mass plus load weight	Weight	Integer	>0kgr
c^W	Temperature	Temperature	Integer	-30...50°C
	Humidity	Humidity	Integer	0…100%
c^{Tr}	Weekday	Weekday	String	Mo…Su
	Time Band	Time Band	String	(00:00-01:59) … (22:00-23:59)
	Month	Month	String	Jan…Dec
c^{Dr}	Average vehicle consumption as reported by the trip computer	Avg. Consumption	Numerical	>0Wh/km
	Road segment energy cost	Recorded Consumption	Numerical	(Wh)

Table 2. Training dataset example

Instance attributes											
LinkId	**SoH**	**SoC**	**Capacity**	**Techno-logy**	**Lights**	**Heat.**	**Airc.**	**Radio**	**Wipers**	**maxPower-Output**	**Weight**
565534258	95	98	21500	Li-Ion	off	off	off	off	off	60	1070
539099812	95	98	21500	Li-Ion	off	low	off	off	off	60	1070
…											

						Target attribute
Temperature	**Humidity**	**WeekDay**	**Time Band**	**Month**	**Avg. Consumption**	**Recorded Consumption**
3	65	Mo	08:00-10:00	Dec	150	57.278
3	65	Mo	08:00-10:00	Dec	158	12.754
…						

Our intention is to build a robust and autonomous prediction system suitable for on-board systems. Having this in mind, we need to optimize the ML functionality so as to be efficient for on-board installations, i.e. in systems with limited memory and computational resources. An important step towards this approach is to minimize the dataset parameters without losing any important information. This also helps decrease the training and scoring times of the MLEs. For scalability reasons, we propose the deployment of a MLE per road segment instead of having a MLE for the entire road network. The benefits of adopting this approach focus on the reduced size of the required dataset, since less training data are needed, and on the increased accuracy of the energy cost predictions, as the underlying process of a dataset corresponding to a specific road segment is of lower complexity than the one corresponding to a road network. Additionally, no specific knowledge on the particular characteristics (e.g. inclination, geometry, etc.) of road segments is needed, since a different MLE is used for each segment. An unavoidable requirement of the proposed approach is the additional space needed to store the various MLEs. However, scalability studies on this topic have shown that the effect on storage is manageable [20]. Therefore, equation (4) can be expressed as:

$$E_T = en_{T,s_i}(\boldsymbol{c}_0^{\boldsymbol{V}}, \boldsymbol{c}_0^{\boldsymbol{W}}, \boldsymbol{c}_0^{\boldsymbol{Tr}}, \boldsymbol{c}_0^{\boldsymbol{Dr}}) \qquad (5)$$

It has been explained that the training process of a specific MLE takes place on-board. But how is the corresponding dataset generated? In case the vehicle travels regularly over a specific road segment, the dataset may easily be populated with measurements retrieved by the on-board system. If the vehicle has not passed through the road segment before, the dataset may be populated with measurements retrieved by other vehicles and shared indirectly (with the mediation of a central synchronization server) through Vehicle-to-Infrastructure and Infrastructure-to-Vehicle communication.

Table 3. Terminology

Term	Explanation
$E_{T,N}$	Normalized Energy Cost Factor, i.e. a cost factor that corresponds to the energy cost of a particular road segment and can be used directly by vehicles of different characteristics (e.g. engine type, air drag coefficient, purpose of usage etc.).
$\hat{E}_{T,N}$	Predicted Normalized Energy Cost Factor.
E_T	Actual Energy Cost.
$\hat{E}_T$	Predicted Actual Energy Cost.
λ	The lamda factor represents the relation between the monitored energy cost of a vehicle and the corresponding energy cost of a standard vehicle driven by a driver with "neutral" driving attitude, namely the Normalized Energy Cost. The term "neutral" refers to a reference stylized driving speed pattern as defined by the European standard New European Driving Cycle (NEDC). NEDC consists of an urban part called ECE, which is repeated four times, and an extra-urban part, the EUDC (Extra-Urban Driving Cycle) [21].
MLE	Machine-Learning Engine.

In order to enable the sharing of experiences between vehicles of different characteristics, a complementary normalization scheme, based also on ML functionality, is proposed. This presents the advantage that parameters containing somewhat privacy-sensitive information (such as part of the Vehicle Context $\boldsymbol{c}^{\boldsymbol{V}}$ or the Driver Profile $\boldsymbol{c}^{\boldsymbol{Dr}}$) do not need to be shared. For this purpose, a normalization MLE is generated and used to transform the energy consumed by a vehicle on a road segment (denoted as E_T) into a normalized (neutral) energy cost factor (denoted as $E_{T,N}$). The various road segment MLEs are trained so as to produce reliable predictions about the road segments' normalized energy cost factors ($\hat{E}_{T,N}$). Whenever such a prediction is performed by a MLE, a "reverse" process must subsequently follow, transforming the predicted normalized energy cost factor into the corresponding predicted actual energy value ($\hat{E}_T$). The normalization/denormalization procedures are based on the use of the lamda factor (λ) defined in Table 3.

In order to support the normalization/denormalization process, the Vehicle Context ($\boldsymbol{c}^{V}$) may be analyzed in the following sub-groups of context variables:

- Battery State ($\boldsymbol{c}^{bs}$), which includes the SoH and the SoC.
- Battery Type ($\boldsymbol{c}^{bt}$), which includes the capacity and the technology.
- Electric Auxiliaries ($\boldsymbol{c}^{aux}$), which includes variables that describe the electric auxiliaries' status.
- Vehicle Type ($\boldsymbol{c}^{vt}$), which consists of the maxPowerOutput and the weight.

Taking this categorization into account, by adopting the proposed normalization scheme equation (5) can be replaced by the following set of equations:

$$E_{T,N} = en_{T,s_i}\left(\boldsymbol{c}_0^{W}, \boldsymbol{c}_0^{Tr}, \boldsymbol{c}_0^{aux}, \boldsymbol{c}_0^{bs}\right) \tag{6}$$

$$\hat{\lambda} = norm(\boldsymbol{c}_0^{Dr}, \boldsymbol{c}_0^{vt}, \boldsymbol{c}_0^{bt}) \tag{7}$$

$$\lambda := \frac{E_T}{E_{T,N}} \tag{8}$$

Equation (6) is derived from equation (5) after transferring some contextual parameters $\left(\boldsymbol{c}_0^{Dr}, \boldsymbol{c}_0^{vt}, \boldsymbol{c}_0^{bt}\right)$ to equation (7), which, in turn, estimates the lamda factor. The lamda factor is used for converting the actual energy values into the normalized ones (during the experience collection process used for training) and the inverse (during the energy costs estimation process used for routing). The benefits stemming from the proposed normalization scheme are: Firstly, the sharing of experiences is enabled, and user privacy is better protected as exchanged measurements are depersonalized. Secondly, the volume of the exchanged measurements is reduced. The same reduction applies also to the required on-board storage capacity, as the volume of stored experiences is decreased. Finally, the segmentation of the initial procedure in successive calculation stages facilitates the efficient use of the limited processing and memory resources available in on-board computer systems.

4. Machine-Learning Routing System Architecture

The proposed methodology is implemented through a functional architecture comprising four main application components: the Prediction System, the Routing Cost Conversion System, the ML Factory, and the Training System. The specifications for these components (i.e. application functions, application services and interfaces that they realize or make use of) are presented in detail in the following (using the Archimate 2.0 modeling language [22]).

Figure 1 provides the specifications for the *Prediction System*, which is responsible for the ML-based prediction of a road segment's energy cost. In particular, it retrieves the corresponding MLE via the *Machine-learning engine extraction* service (see also Figure 3), as well as the current context via the *IContextInfo* interface. The application feeds the context to the MLE and returns the predicted energy cost.

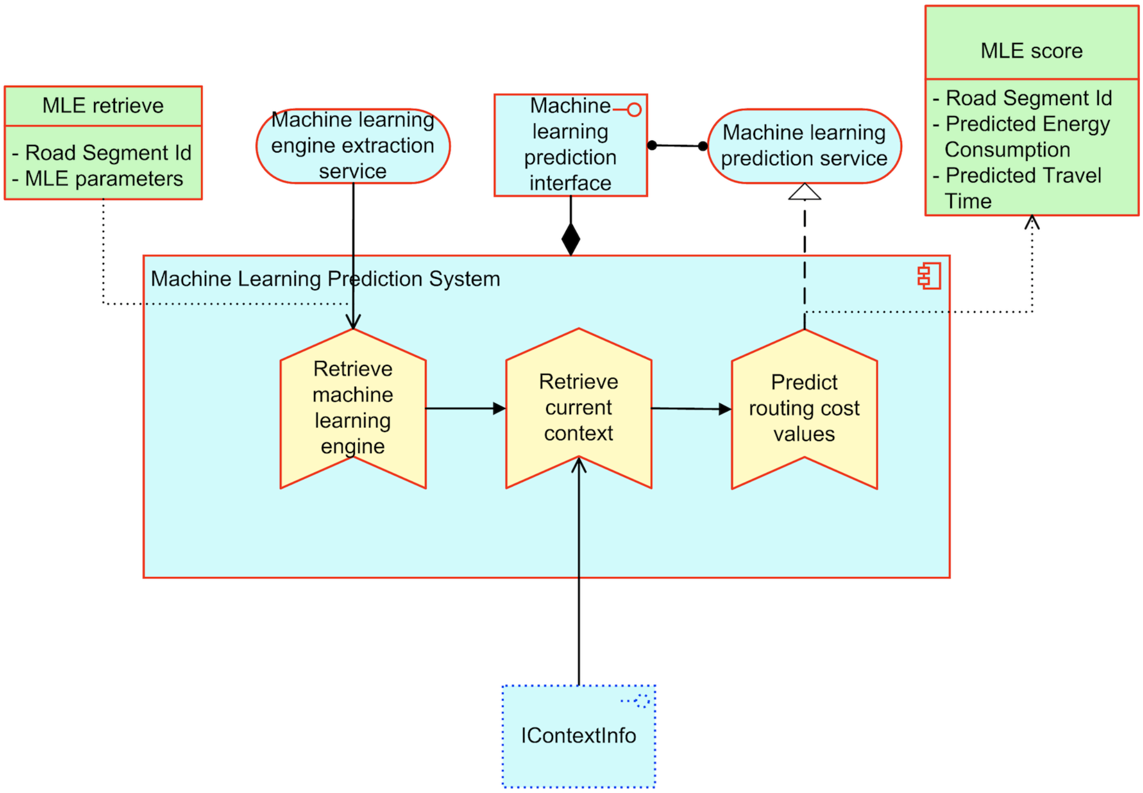


Figure 1. Prediction System

Figure 2 provides the specifications for the *Routing Cost Conversion System*, which is responsible for converting actual energy costs of road segments into normalized values, and vice versa. This application comprises two sub-components, namely the *Normalization System* and the *Denormalization System*, whose functionality has been explained in Section 3.

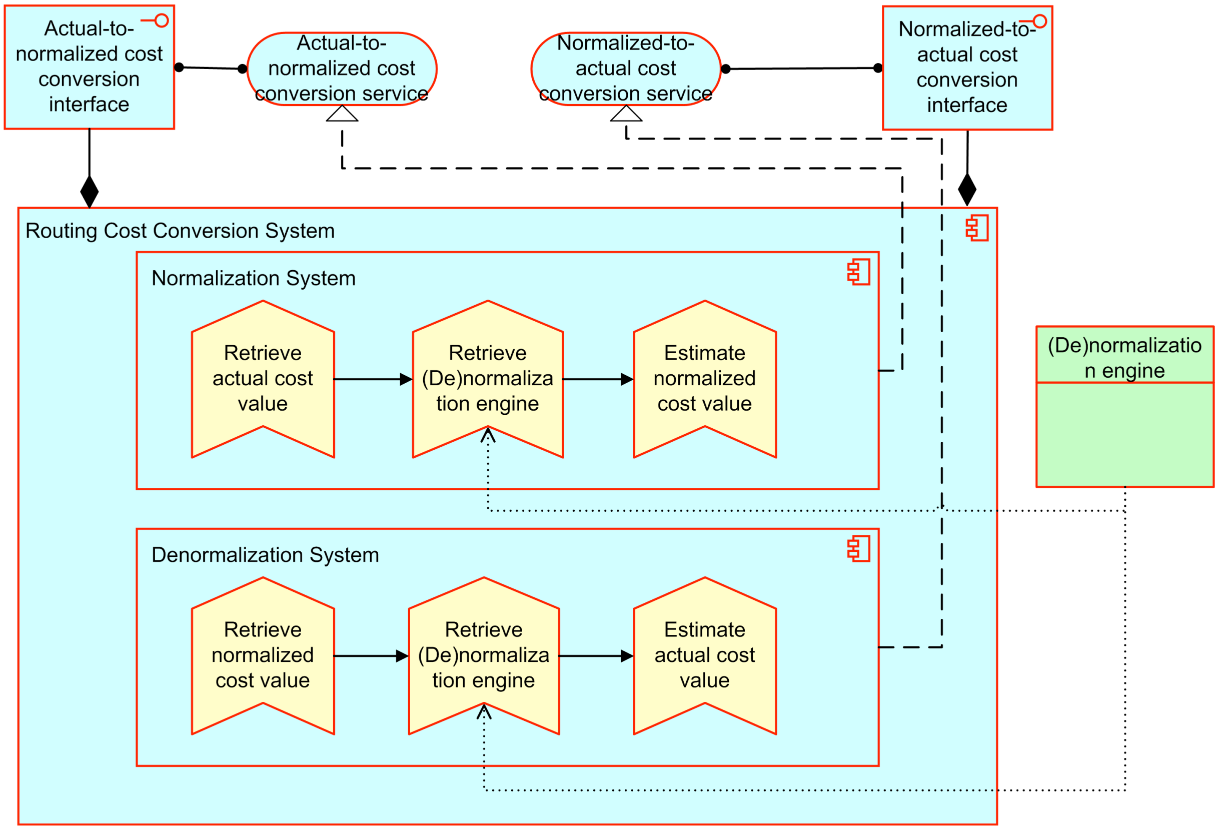


Figure 2. Routing Cost Conversion System

Figure 3 provides the specifications for the *ML Factory*, which is responsible for retrieving the corresponding MLE upon request, or for producing a new one in case there is yet no MLE associated to the road segment in question. For this reason, this application first checks if the requested MLE already exists in the *Machine-Learning Engine database*. If such engine exists, it means that it has been produced at some time in the past, and so it is returned. Otherwise, the particular road segment does not yet have any MLE associated to it. In this case, the application produces a stub engine, stores it and returns it to the requestor.

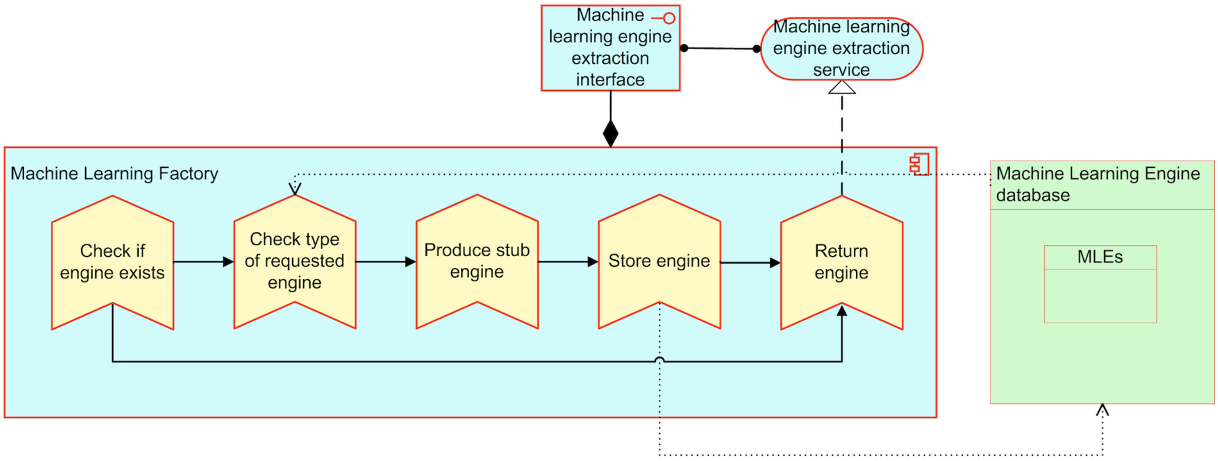


Figure 3. ML Factory

Figure 4 provides the specifications for the *Training System*, which is responsible for performing the MLEs' training. Periodically, the MLEs need retraining, since new historical data are being recorded. This application is also responsible for performing the scheduling of training sessions. For this reason, its functions are grouped internally into two sub-components, the *Training Scheduler* and the *Training Execution Module*. The first one implements the training policies (e.g. perform training when battery level is above a threshold, or during specific time windows, etc.) and initiates training or retraining based not only on the availability of new data (retrievable through the *ISynchInfo* interface), but also on the frequency or probability that these segments will be visited by the user. This is the reason why the application makes use of the driving history data archive and of the external *Electronic horizon* service. After making this decision, the application may initiate the training of the chosen road segments (i.e., subsequently, of the chosen MLEs, retrieved via the *Machine-learning engine extraction* service). The execution of the training itself, on a particular MLE, is handled by the second application sub-component, which must also make use of the *IHdbIO* interface, for retrieving the corresponding historical data.

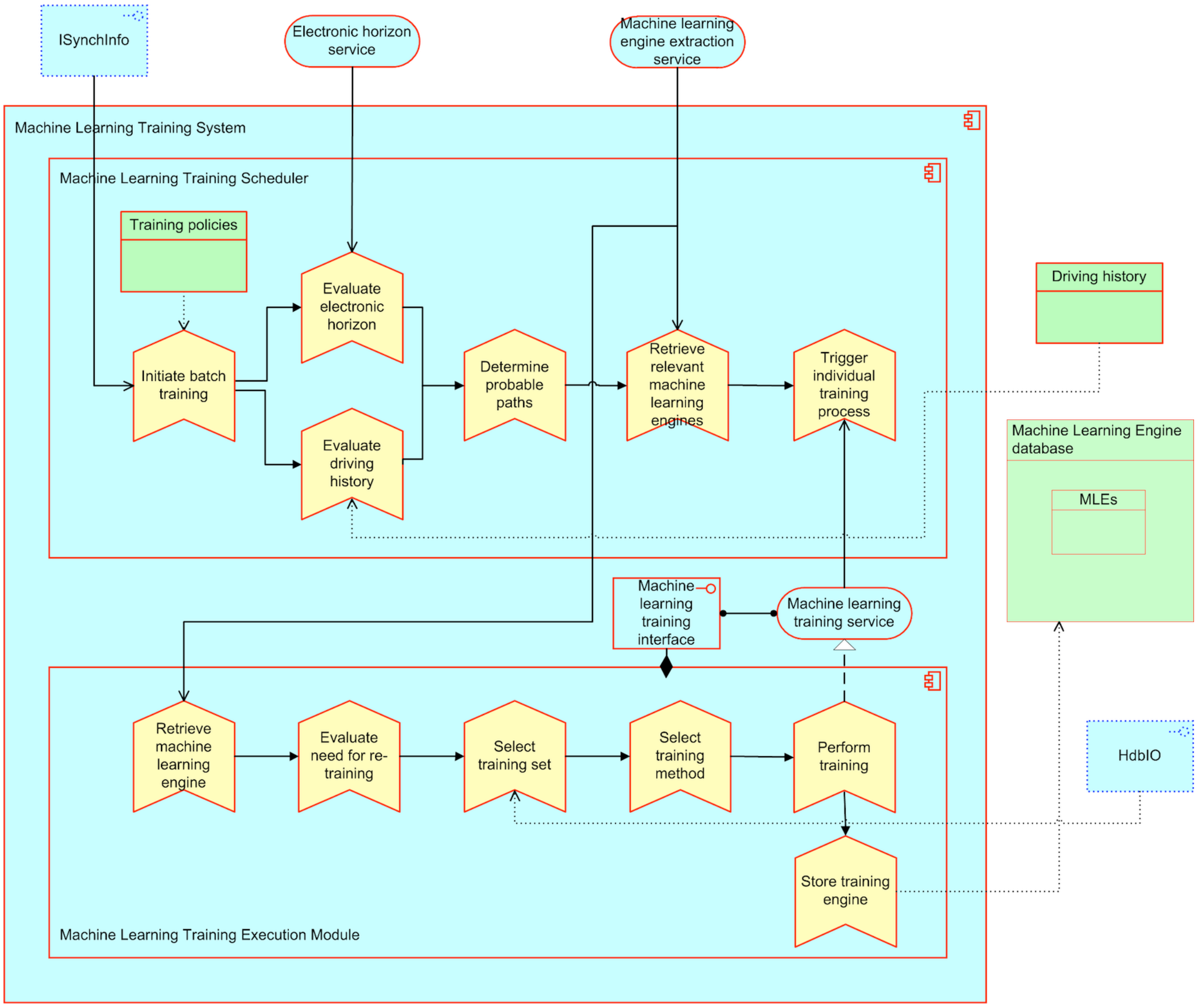


Figure 4. Training System

5. Experimental Results

This section presents the experimental results generated during a field trial campaign that took place in Chieri (near Turin), Italy. The study of these results aims at validating the proposed ML methodology used for predicting energy consumptions. The experimental setup includes:

- an ADAS prototype equipped with 3G connectivity as well as with the proposed ML functionality, implemented in C++.
- a FEV equipped with a Controller Area Network (CAN) bus interface. The test FEV is powered by a battery of 21.5kWh and has a nominal range of 140km.
- a CAN bus adapter: The equipment used is CANcaseXL [23].
- the test area, extended in a region of about 6x6km, comprises the town of Chieri (Figure 5) and its suburbs. This area is suitable for tests, since it contains a meshed road network including segments with a variety of characteristics, in terms of speed limits, slopes, number of lanes, road class, etc., and presenting variable traffic conditions during the day.

Figure 5. Test Area (Chieri, Italy)

A 4-week data collection campaign was performed with the test FEV, in order to generate an adequate dataset, i.e. representative of different sets of contextual conditions, for training the MLEs. A specific parking lot was selected as the starting point of each trip. The FEV started each time with a fully charged battery and returned when the SoC was approaching a critical level (after travelling for about 100-140km). Each trip was followed by a recharging process lasting from 8 to 10 hours and then a new cycle started.

Table 4. Contextual data sources

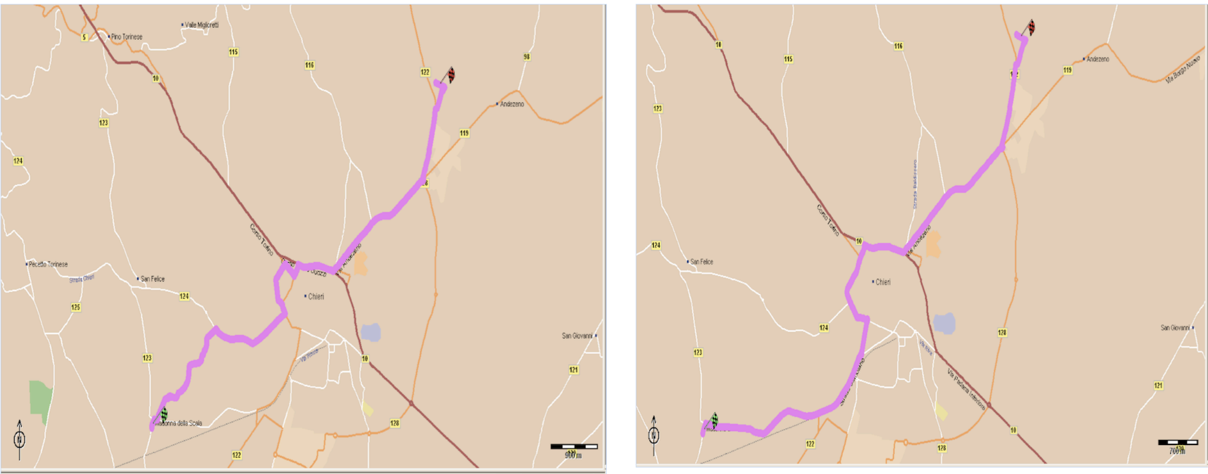
Source	Connectivity	Variables measured
Vehicle microcontrollers	Wired through CAN bus interface	Electric auxiliaries' status, avg. consumption
Internet web services	3G	Temperature, humidity
Configuration file	Local	maxPowerOutput, weight, capacity, battery technology
Battery Management System (BMS)	Wired through CAN bus interface	SoC, SoH, instantaneous voltage and current (used for energy consumption cost estimation through $\int_{time} Voltage * Current$)
ADAS internal clock	Local	Weekday, month, time band

The contextual data collected during the campaign were retrieved through several sources (Table 4). After the completion of the data collection campaign, the generated dataset was used to train the MLEs of the road segments that had been visited. The ML method employed for this campaign is the Artificial Neural Networks (ANNs). An ANN represents a supervised learning algorithm, suitable for regression problems, and features superior performance in a wide range of problems according to empirical studies [24]. Regarding the learning process, the scaled conjugate gradient algorithm [25] was preferred instead of the traditional gradient descent, as it follows a more direct path to the optimal set of weight values. Table 5 contains an analytical description of the ANN engines employed by the proposed methodology. The training process resulted in a set of 2436 estimation engines corresponding to the road segments of the testing area and in one normalization engine.

Table 5. Description of ANN engines

	Estimation engine	**Normalization engine**
Architecture	multilayer feed-forward with one hidden layer	multilayer feed-forward with one hidden layer
Hidden layer activation function	Logistic	Logistic
Output layer activation function	Linear	Linear
Learning algorithm	Scaled conjugate gradient	Scaled conjugate gradient
Network structure (x is computed by an optimization process applied to each ANN separately)	12-x-1	5-x-1
Input variables	(temperature, humidity, weekday, time band, month, lights, heat., airc., radio, wipers, SoC, SoH)	(avg. consumption, maxPowerOutput, weight, capacity, battery technology)
Output variable	Normalized value of the road segment energy consumption	λ

The next step of the validation process included the generation of routing results based on the previously trained MLEs. Firstly, 80 sets of origins and destinations were selected randomly in a way ensuring that the generated routes would be included in the testing area. Then, for each of the 80 sets, the fastest and the most energy efficient routes were calculated and driven, and the corresponding results (in terms of energy consumption) were extracted. The most energy efficient route corresponds to the one proposed by the developed ML algorithm, while the fastest one corresponds to the "standard" route as suggested by a typical navigator (Nokia HERE [26]). Of course, for each origin-destination pair both the fastest and the most energy efficient route were calculated after ensuring the same contextual conditions. Figure 6 illustrates an example, i.e. the most energy efficient and the fastest routes for the same pair of origin ("Strada Madonna della Scala") and destination ("Via del Ponte Vecchio 15") points. The two generated routes refer to the same contextual conditions (Table 6), while their comparison results are presented in Table 7. According to this example, there is a 12.24% energy saving in case the FEV follows the energy efficient route, and also a 9.15% loss in the time required to travel from the origin to the selected destination.



(a) **(b)**

Figure 6. Example of Energy Efficient Routing (a) and Fast Routing (b) for the same origin-destination pair and contextual conditions

Table 6. Contextual Instance when generating routes for the example of Figure 6

<table>
<tr><td colspan="11">CONTEXTUAL CONDITIONS</td></tr>
<tr><td colspan="11">Vehicle Context</td></tr>
<tr><td>SoH (%)</td><td>SoC (%)</td><td>Capacity (Wh)</td><td>Techno-logy</td><td>Lights</td><td>Heat.</td><td>Airc.</td><td>Radio</td><td>Wipers</td><td>Power out-put max (kW)</td><td>Weight (kgr)</td></tr>
<tr><td>95</td><td>80-90</td><td>21500</td><td>Li-Ion</td><td>off</td><td>mid</td><td>off</td><td>off</td><td>off</td><td>60</td><td>1070</td></tr>
<tr><td colspan="2">Weather Context</td><td colspan="3">Traffic Context</td><td>Driver Profile</td></tr>
<tr><td>Temperature (°C)</td><td>Humidity (%)</td><td>WeekDay</td><td>Time Band</td><td>Month</td><td>Avg. Consumption (Wh/km)</td></tr>
<tr><td>5</td><td>74</td><td>We</td><td>08:00-09:59</td><td>Dec</td><td>162</td></tr>
</table>

Table 7. Comparison of the Energy Efficient routing and Fast routing results for the example of Figure 6

	Routing Algorithm		**Comparison**
	Energy Efficient	*Fastest*	*Energy Efficient vs. Fastest*
Energy Consumption (Wh)	1238.2	1410.95	-12.24%
Travel Time (s)	732	665	9.15%
Route Length (m)	8944	9046	-1.13%
Number of Links	70	82	-14.63%

Comparison results between all of the proposed and the “standard” routes were collected and aggregated, and the final outcome can be observed in the diagrams depicted in Figure 7. Figure 7 includes a frequency graph (primary axis) and a cumulative frequency graph (secondary axis) of the energy savings that were achieved when applying the proposed routing scheme. According to these graphs, the routes generated by the proposed scheme are always more energy efficient than the corresponding “standard” routes and, therefore, the proposed routing scheme can be regarded as fully verified. Furthermore, the achieved energy savings are greater than 10% in almost two thirds of the cases, and greater than 15% in half (50%) of the cases, which is quite encouraging regarding the efficiency of the proposed solution. In general, during the conducted field trials, the energy savings enabled were found to be up to 36.88% (compared to the energy spent in the “standard” route).

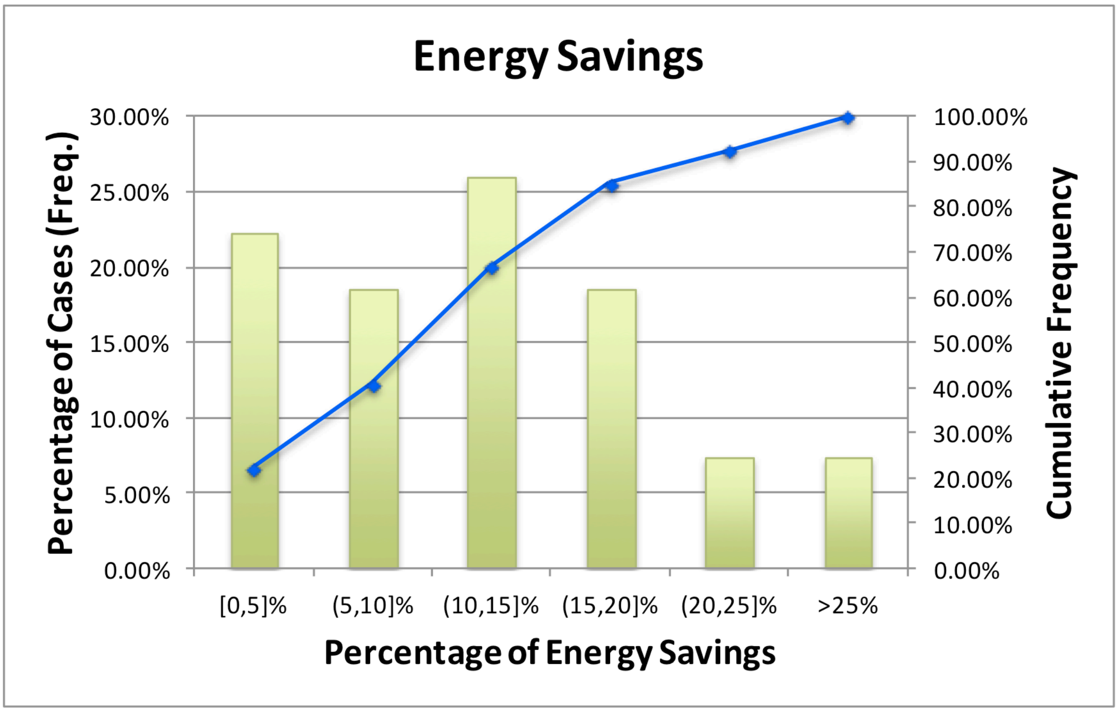


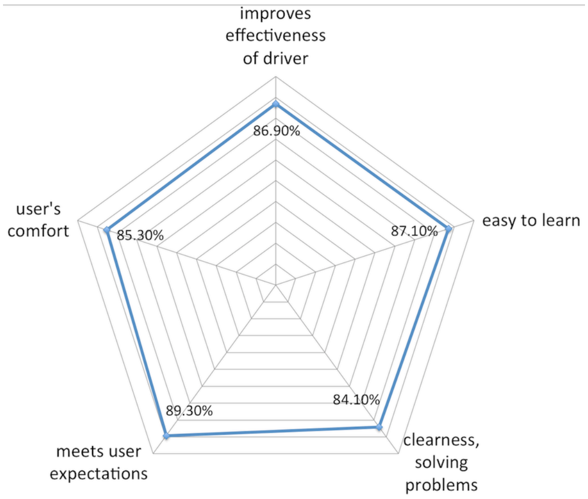
Figure 7. Frequency and Cumulative Frequency diagrams of the Energy Savings

Energy consumption prediction accuracy: It should be noted that, for non-linear estimation models as in the proposed one, the percentage of energy consumption caused by expected conditions may be calculated through the statistic metric pseudo-R^2 ($\tilde{R}^2$) [27], which is the closest metric to the co-efficient of determination R^2 used in linear models.

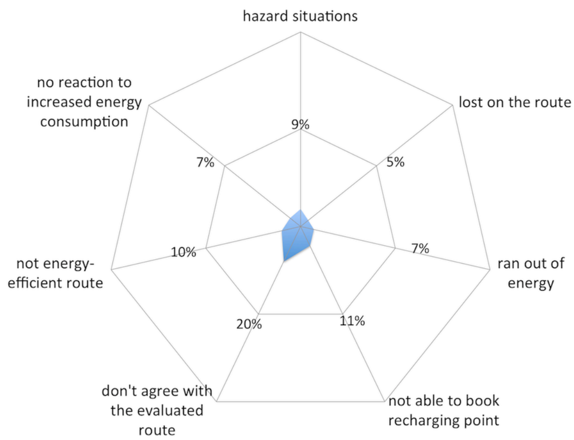
$$\tilde{R}^2 = 1 - \frac{\sum_{i=1}^{N}\left(E_i - \hat{E}_{T,i}\right)^2}{\sum_{i=1}^{N}\left(E_i - \bar{E}\right)^2} \qquad (9)$$

where E_i denotes the measured value of the real energy consumption for a particular case i (combination of road segment and context), $\hat{E}_{T,i}$ represents the corresponding predicted value of the expected part of the energy consumption, and $\bar{E}$ expresses the mean value of the real energy consumption over N different cases. For the aforementioned experimental campaign, the pseudo-R^2 metric is calculated equal to 87%, meaning that the energy consumption attributed to predictable (expected) conditions, which is estimated by the proposed model, composes the largest part of the total actual energy consumption.

User acceptance: Finally, although the end-user perspective is not the direct focus of the present paper, it is worth presenting shortly the outcomes of a relevant survey carried out in order to assess user acceptance of the proposed system. A specifically designed questionnaire was prepared and distributed to 35 drivers, who may be perceived as end-users of the system. The questionnaire consisted of twenty questions regarding the system's effectiveness, easiness, clearness, comfort, etc. The degree of agreement with each of the statements of the questionnaire was expressed using a number from Likert scale.



(a)



(b)

Figure 8. (a) System usability results, (b) Issues faced by end-users

Based on the results of this survey (Figure 8), the system vastly meets user expectations, addressing the issues of user comfort and driver efficiency. The percentage of drivers disagreeing with the evaluated route (either for reasons of time efficiency or other) ranges at low levels, which provides good first indications with regards to user acceptability.

6. Conclusions

To sum up, this paper presented an innovative methodology for energy efficient routing based on ML techniques. By applying these methods, vehicles are rendered capable of learning, thus predicting energy consumption along road segments. The proposed techniques together with the functional architecture that implements them were discussed in detail, accompanied by figures of the main functional blocks, as well as experimental results. Results show that the proposed solution succeeds in guiding FEVs through more energy efficient routes, enabling energy savings up to 36.88%. Particularly, savings of more than 15% were recorded in half (50%) of the test cases conducted in the field. Future work includes a detailed study on estimating the unexpected part of the energy needed to travel through a road segment at a specific time slot, as well as further field trials of the proposed system.

Acknowledgement: This work has been performed under project EcoGem [28], which has received research funding from the European Community's Seventh Framework Programme.

References

[1] Wada, T., Yoshimura, K., Doi, S.-I., Youhata, H., and Tomiyama, K.: 'Proposal of an eco-driving assist system adaptive to driver's skill'. Proc. 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), October 2011, pp. 1880-1885

[2] Kamal, M.A.S., Mukai, M., Murata, J., and Kawabe, T.: 'Ecological Vehicle Control on Roads With Up-Down Slopes'. IEEE Transactions on Intelligent Transportation Systems, 2011,12, (3), pp.783-794

[3] Yuxiao C., Dezhao Z., and Keqiang L.: 'Enhanced eco-driving system based on V2X communication'. Proc. 15th International IEEE Conference on Intelligent Transportation Systems (ITSC), September 2012, pp. 200-205

[4] Rakha, H., and Kamalanathsharma, R.K.: 'Eco-driving at signalized intersections using V2I communication'. Proc. 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), October 2011, pp.341-346

[5] Ericsson, E., Larsson, H., and Brundell-Freij, K.: 'Optimizing route choice for lowest fuel consumption - Potential effects of a new driver support tool', Journal of Transportation Research Part C: Emerging Technologies, 2006, 14, (6), pp. 369-383

[6] Trommer, S., and Höltl, A.: 'Perceived usefulness of eco-driving assistance systems in Europe', IET Intelligent Transport Systems, 2012, 6, (2), pp. 145-152

[7] Mitchell, T.: 'Machine Learning' (McGraw-Hill, 2002)

[8] Minett, C.F., Salomons, A.M., Daamen, W., van Arem, B., and Kuijpers, S.: 'Eco-routing: comparing the fuel consumption of different routes between an origin and destination using field test speed profiles and synthetic speed profiles'. Proc. IEEE Forum on Integrated and Sustainable Transportation System (FISTS), Vienna, Austria, June 2011, pp. 32-39

[9] Kono, T., Fushiki, T., Asada, K., and Nakano, K.: 'Fuel consumption analysis and prediction model for 'eco' route search'. Proc. 15th World Congress on Intelligent Transport Systems and ITS America's 2008 Annual Meeting, New York, November 2008

[10] Barth, M., Boriboonsomsin, K., and Vu, A.: 'Environmentally-Friendly Navigation'. Proc. IEEE Intelligent Transportation Systems Conference, Seattle, September 2007, pp. 684-689

[11] Kraschl-Hirschmann, K., and Fellendorf, M.: 'Estimating energy consumption for routing algorithms'. Proc. IEEE Intelligent Vehicles Symposium (IV), June 2012, pp. 258-263

[12] Boriboonsomsin, K., Barth, M.J., Weihua, Zhu, and Vu, A.: 'Eco-Routing Navigation System Based on Multisource Historical and Real-Time Traffic Information', IEEE Transactions on Intelligent Transportation Systems, 2012, 13, (4), pp. 1694-1704

[13] Frey, H.C., Zhang, K., and Rouphail, N.M.: 'Fuel Use and Emissions Comparisons for Alternative Routes, Time of Day, Road Grade, and Vehicles Based on In-Use Measurements', Journal of Env. Science & Technology, 2008, 42, (7), pp. 2483-2489

[14] Ahn, K., and Rakha, H.: 'The effects of route choice decisions on vehicle energy consumption and emissions', Transportation Research Part D, Transport and Environment, 2008, 13, (3), pp. 151-167

[15] Vlieger, I. De, Keukeleere, D. De, and Kretzschmar, J.G.: 'Environmental effects of driving behaviour and congestion related to passenger cars', Atmospheric Environment, 2000, 34, (27), pp. 4649-4655

[16] Bandeira, J., Carvalho, D., Khattak, A., and Coelho, M.C: 'A Comparative Empirical Analysis of Eco-friendly Routes during Peak and Off-peak Hours'. Proc. 91st Annual Meeting of the Transportation Research Board, Washington, D.C., 2012

[17] Ben Dhaou, I.: 'Fuel estimation model for ECO-driving and ECO-routing'. Proc. IEEE Intelligent Vehicles Symposium (IV), June 2011, pp. 37-42

[18] Jianming, H., Qiang, M., Qi, W., Jiajie, Z., and Yi, Z.: 'Traffic congestion identification based on image processing', IET Intelligent Transport Systems, 2012, 6, (2), pp. 153-160

[19] Ki, Y.-K.: 'Speed-measurement model utilising embedded triple-loop sensors', IET Intelligent Transport Systems, 2011, 5, (1), pp. 32-37

[20] EcoGem Project, Deliverable D3.1, "Scalability and Performance Analysis of EcoGem Applications", June 2011

[21] OICA Definition Test Procedure WLTP-DTP-02-04, V.: 28.05.2010

[22] THE Open Group: 'ArchiMate 2.0 Specification' (Van Haren Publishing, Jan. 2012)

[23] CANcaseXL product Webpage, http://www.vector.com/vi_cancase_xl_log_en.html, accessed December 2012

[24] Caruana, R., and Niculescu-Mizil, Al.: 'An empirical comparison of supervised learning algorithms'. Proc. 23rd International Conference on Machine learning, New York, USA, 2006, pp. 161-168

[25] Lawrence, S., and Giles, C.L.: ‘Overfitting and Neural Networks: Conjugate Gradient and Backpropagation’. Proc. International Joint Conference on Neural Networks, 2000, 1, pp. 114-119

[26] NOKIA here Website, http://here.com, accessed December 2012

[27] http://statistics.ats.ucla.edu/stat/mult_pkg/faq/general/Psuedo_RSquareds.htm, accessed April 2013

[28] EcoGem Project Official Website, http://www.ecogem.eu, accessed December 2012