

## Article

# Economic Impact of the Use of Inertia in an Urban Bus Company

Alejandro G. Tuero <sup>1,\*†</sup>, Laura Pozueco <sup>1</sup>, Roberto García <sup>1</sup>, Gabriel Díaz <sup>2</sup>, Xabiel G. Pañeda <sup>1</sup>, David Melendi <sup>1</sup>, Abel Rionda <sup>3</sup> and David Martínez <sup>3</sup>

<sup>1</sup> Informatics Department, University of Oviedo, 33204 Gijón, Asturias, Spain; pozuecolaura@uniovi.es (L.P.); garciaroberto@uniovi.es (R.G.); xabiel@uniovi.es (X.G.P.); melendi@uniovi.es (D.M.)

<sup>2</sup> Electrical, Electronic, Control and Telematics Engineering, and Engineering applied Chemistry Department, Spanish University for Distance Education (UNED), 28040 Madrid, Spain; gdiaz@ieec.uned.es

<sup>3</sup> ADN Context-Aware Mobile Solutions S.L., 33394 Gijón, Asturias, Spain; abel.rionda@adnmobilesolutions.com (A.R.); david.martinez@adnmobilesolutions.com (D.M.)

\* Correspondence: garciatalejandro@uniovi.es; Tel.: +34-985-182-000 (ext. 6565)

† This author is an Industrial Technologies Ph.D. Candidate, at Spanish University for Distance Education (UNED), 28040 Madrid, Spain.

Academic Editor: Xiaosong Hu

Received: 24 April 2017; Accepted: 11 July 2017; Published: 19 July 2017

**Abstract:** Public transport companies face great expenses. In order to be profitable companies, they need to reduce costs in all aspects. In addition, many countries have increased taxes to force transport companies to reduce their fuel consumption and, thus, greenhouse effect emissions. Efficient driving is the cheapest way to achieve this goal. In this paper, we analyze the economic impact of one of the most influential efficient driving techniques in an urban bus company. We calculate the difference of consumption between a well-performed deceleration using the inertia of the vehicle and a loss of speed obtained by using the brakes. For this purpose, we compare a real track performed by any driver in a precise vehicle with a simulated ideal track. To obtain the deceleration of the simulation, we develop a linear regression model on over 170,000 samples captured from the same vehicle over a period of one month. The results show that the costs of the vehicle under test in one month may be reduced by more than 2500 € by using this technique. With these results, we can see that it is worth trying to anticipate different situations while driving only by applying inertia adequately.

**Keywords:** efficient driving; fuel consumption; public transport; economic impact; urban bus

## 1. Introduction

Fuel consumption, in terms of economical reliability, is one of the biggest challenges for transport companies. The fact that road transport is one of the most polluting sectors in terms of greenhouse effect gas emissions [1] has led to a resolution by the Conference of the Parties to reduce Global Warming. This resolution motivates the countries to minimize emissions. Accordingly, transport companies are forced to apply certain measures to reduce costs such as fleet renovation, changing from conventional fuel to bio-fuel [2], preventive and predictive maintenance or efficient driving training for their drivers, the last being the cheapest way to obtain early results.

Many methods lead to efficient driving, depending on the final application of the learned techniques. In the particular case of urban public transport companies, the most common method consists of an initial training day, including one practical individual session. In that session, each driver drives a route and the trainer observes his or her mistakes and strengths. After that, all drivers attend a theoretical session in which the national or international recommendations on efficient

driving are explained [3]. Finally, in another practical individual session, the drivers apply the learned techniques following the indications of the trainers. Monitoring the vehicles and allowing trainers to check the evolution of the performance of the drivers and to prepare periodic reinforcement sessions based on a more objective evaluation may complete this initial training day.

Regarding the evaluation of the driver's performance, there is more uniformity. Most companies check the evolution of consumption as the evaluation metric. Nevertheless, it has been shown that this is not a good way to assess the driver's performance, as fuel consumption depends on too many aspects that are beyond the control of the driver such as the weather or the density of traffic and not only on his or her driving skills [4].

Our process dismisses fuel consumption as an evaluation method. Thus, in our previous research we developed an evaluation method based on some efficient driving patterns [5]. This method allows us to perform an objective evaluation of the skills of the drivers because it is based on real driving data. Furthermore, it is also able to facilitate adaptive training as we can see the evolution of the performance of the driver in each pattern and focus the reinforcement sessions on his or her mistakes [6]. All the real data of our studies are obtained by a system called Cated Box, which receives information directly from the vehicle's engine control unit (ECU) [7]. This is already under application in 16 professional fleets in Spain and Morocco, involving 880 professional drivers.

In the way that our system obtains the data, it is impossible to know why a driver behaves in a particular way. Nevertheless, observation of data and experience allow us to make assumptions. Our training method promotes safe and efficient driving. That includes anticipation, and keeping the safety distance is paramount to anticipation as it minimizes the influence of other vehicles on the performance of a driver. Thus, every driver, no matter the density of the traffic, must strive to maximize the use of inertia.

In this paper, our main goal is to show the economic impact, in the form of money saving, that the application of some of the efficient driving recommendations may generate in an urban public transport company. To do so, we analyze real driving data gathered from one single vehicle during a complete month. This means analyzing a total amount of over one million samples. We compare a real track in which the driver decelerates in an inefficient way with an ideal theoretical track obtained from the linear regression. This ideal track simulates a perfect inertia deceleration with null consumption. The real track is conformed by data gathered during the normal activity of the vehicle. As we explain in [8], these data are sent to the central system and processed in a learning analytics module. In this case, we represent an example of real driving in terms of speed and distance.

The paper is structured as follows; in Section 2, we make a review of related work. Section 3 gives a background to the study while Section 4 shows the methodology applied for the calculations. Section 5 gives the results of a brief case study, and Section 6 includes the discussion and conclusions.

## 2. Related Work

There is ample literature on the influence of efficient driving on fuel consumption. A good overview can be found in [9] as it shows the methodology, type of study, and results for the most relevant studies made in recent years.

Different types of techniques are applied in the literature to quantify the problem of fuel consumption. In field tests, the authors usually apply their own methodology to actual vehicles and drivers and obtain results that are real or, at least, estimated from real data. The dimension of the experiments in the overviewed literature varies from three drivers performing a training route of 15 km [10] to 54 drivers who perform routes of 16 km over a period of six weeks [11]. All these studies show that the application of efficient driving techniques increases savings in fuel consumption. However, these studies are focused on a set of basic recommendations on efficient driving with the aim of reducing fuel consumption. Our approach is different as we use efficient driving patterns as the basis of an efficient driving learning methodology, the main goal of which is the modification of driving habits of professional drivers to perform more efficient and safe driving, but it also allows noteworthy fuel savings. These patterns have the particularity of having a temporal dimension instead of being a combination, simple or complex, of instantaneous events. Therefore, we base the

potential fuel savings on the application of an efficient driving pattern, inertia, that describes the driver's behavior in a temporal window. In addition, our dataset is composed by real data, obtained during the complete working day activity of the professional drivers, while other related works are constrained to prepared scenarios. In this sense, the work presented in [12] also highlights that drivers should be evaluated during their normal work activity because fuel saving goals could be affected by other external factors. Furthermore, the deceleration during inertia is the result of a linear regression model, which is also a novelty.

Nevertheless, the range of fuel saving varies between the 1% obtained in highway testing with the help of a real time on-board assistance device [13] and the 18.7% obtained in 1 km long sections with three intersections [14].

The authors in [15] compare two different approaches in two public transport companies. It is very relevant to our research as they obtain data from an on-board device from vehicles and drivers during their professional activity. One of them provides data to illustrate in-class regular training and the other sends real-time alerts. Both systems use isolated undesired events such as excessive speed or prolonged idle time to advise the drivers on their behavior. Our pattern method is not attached to single-instant events but is opened in a temporal window, which provides a wider overview of behavior and context.

There are also some review papers or reports by organizations in which the authors give their conclusions based on the material available that fits their requirements. Of these papers, [4] provides the greatest improvement in consumption (up to 30%) by comparing efficient and aggressive driving and using a speed control device. The lowest savings are shown in [16], ranging from 5% after training to 15%, depending on the driver. The authors ensure 10% if a real time assistance device is used.

Finally, a last type of paper is that in which the authors develop a mathematical model that seeks to calculate the benefits of efficient driving. Although not comparable as they use different metrics, in [17] the authors achieve maximum savings of 34% with an algorithm tested under standard urban traffic. Additionally, [18] obtained a saving of 1.5 l/100 km with an algorithm based on micro-economic theory. They performed tests in five stages, considering different levels of density and traffic accidents. For a mixed consumption of 10.24 l/100 km, their results achieve savings of 14.6%.

To our knowledge, there are no references in the literature in which inertia is used in the same way as in our work. We apply a linear regression model based on real driving data in order to obtain an ideal deceleration without fuel consumption by using inertia as an efficient driving technique. Then we compare this ideal function with real data using the brakes to see the differences in fuel consumption. Starting at the same point and speed, a vehicle using the brakes will only reach the same point and the same speed as another using inertia if the former accelerates during a certain period of time. We calculate the cost of using the brakes by checking the consumption caused by that acceleration in contrast with the null expense of fuel required by inertia driving. It is an effort to quantify the impact of anticipation separately from the other efficient driving techniques.

### 3. Background

The present study is based on the theoretical good usage of inertia when a vehicle is approaching a stopping point or when speed needs to be adapted to slow traffic ahead. It has much to do with the anticipation of traffic and service situations.

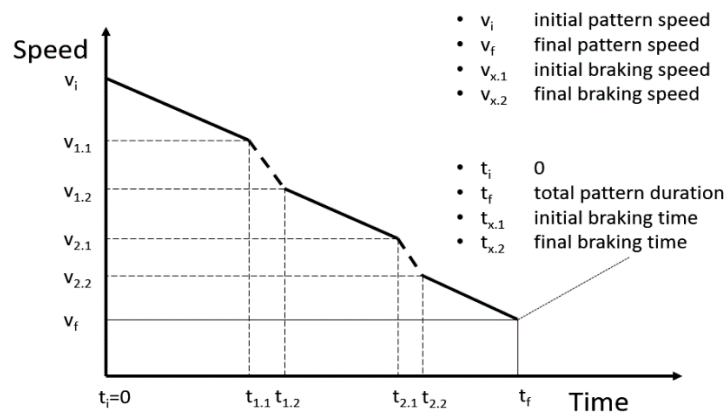
We initially developed the inertia pattern [5], in which we looked for periods of time in which the vehicle was in motion and gear engaged without consumption. Its key performance indicator (KPI) is the percentage of time in which the vehicle runs with no consumption. This was a first approach to inertia as an efficient driving pattern. At that time, we could only obtain certain parameters from the vehicle to define inertia. The use of the brakes was not included among those parameters, although it is a key element to determine the way inertia is conformed.

With an extension in the amount of data registered by the system, we started to obtain the parameter of the use of the brakes by the driver, and thus we saw the chance to study how that initial

inertia pattern was conformed. Thus, we developed a set of complex inertia patterns that improve our detailed knowledge of driving conditions. The definition of these patterns follows.

### 3.1. Pattern of Corrected Inertia without Stop (#1)

The driver starts decelerating in inertia and corrects the speed with the brakes so that stopping is not finally necessary, as Figure 1 shows. It may appear next to a traffic light that is initially red that becomes green before the vehicle arrives at the stopping point. The continuous lines show the periods in which inertia is used, and the dashed lines show those in which the driver uses the brakes. The  $v$  parameters show the relevant speed values along the pattern such as the initial speed ( $v_i$ ), braking start ( $v_{1.1}, v_{2.1}$ ), or the speed at which the driver retains inertia ( $v_{1.2}, v_{2.2}$ ).  $v_f$  represents the final speed of the pattern and the point at which the driver accelerates again.



**Figure 1.** Pattern of Corrected Inertia without Stopping.

We can ascertain the total loss of speed of the pattern just by subtracting the initial speed from the final speed as follows:

$$v_{loss} = v_i - v_f, \quad (1)$$

In addition, we can calculate the speed loss due to the use of brakes as follows:

$$v_{B.loss} = \sum_{j=1}^n (v_{j.1} - v_{j.2}), \quad (2)$$

where  $n$  is the number of braking periods in the whole pattern and  $j$  is the order number. Finally, we can calculate the loss of speed due to inertia as follows:

$$v_{I.loss} = (v_i - v_{1.1}) + \left( \sum_{j=1}^{n-1} (v_{j.2} - v_{(j+1).1}) \right) + (v_{n.2} - v_f), \quad (3)$$

The calculations to determine the time of use of the brakes ( $t_B$ ) and the time of inertia ( $t_I$ ) are equivalent.

The KPI for this pattern is the percentage of time of use of the brakes ( $t_B$ ) against the total duration of the pattern ( $t$ ).

$$KPI(\%) = \frac{t_B}{t} \cdot 100 = \frac{\sum_{j=1}^n (t_{j.2} - t_{j.1})}{t_f - t_i} \cdot 100, \quad (4)$$

The rest of the patterns have the same structure, the only difference being what happens after the pattern finishes. While in this case the vehicle does not stop (there is a long acceleration after the inertia period), the other two patterns result in the stopping of the vehicle but in different circumstances.

### 3.2. Pattern of Corrected Inertia Long (#2)

As we have explained previously, this pattern (Figure 2) has the same structure as the other complex patterns of inertia but with a necessary stop at the end, which is probably a red traffic light or a bus stop.

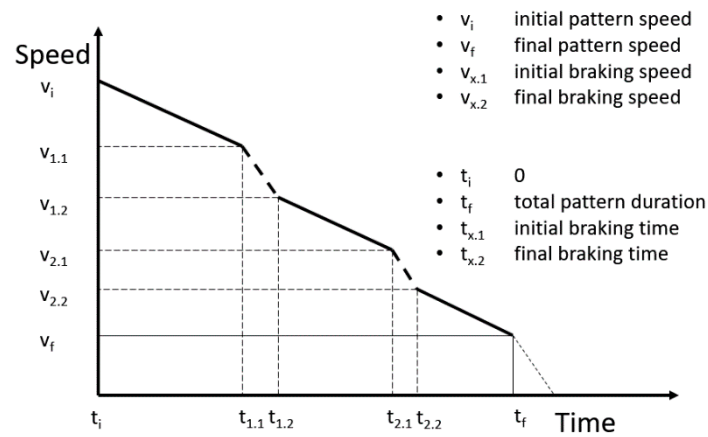


Figure 2. Long Pattern of Corrected Inertia.

The KPI is the percentage of time using the brakes compared with the total duration of the pattern, exactly as in the case of the first pattern.

### 3.3. Pattern of Corrected Inertia Short (#3)

This pattern (Figure 3) again has the same structure but with a short acceleration period at the end, represented by a line of dots, caused by a bad estimation of the driver of the distance that the vehicle is able to run with its inertia.

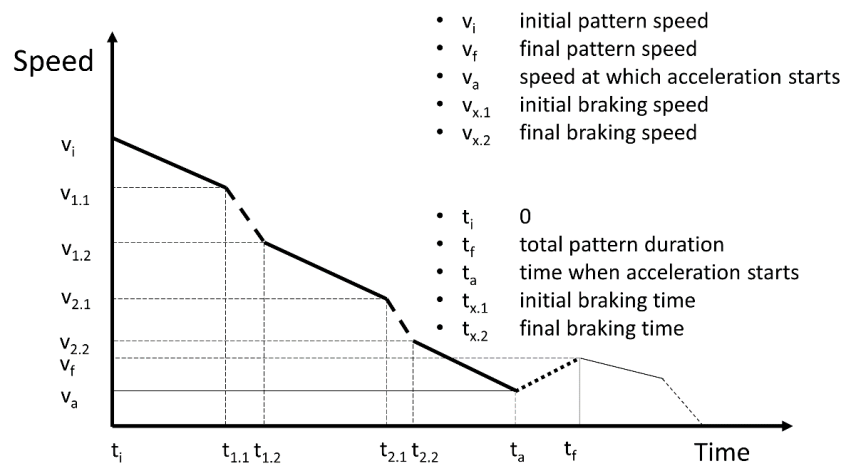


Figure 3. Short Pattern of Corrected Inertia.

This pattern has two KPIs. On the one hand, we have the percentage of time of use of the brakes compared with the total deceleration time ( $KPI_B$ ), as in the previous patterns.

$$KPI_B(\%) = \frac{t_B}{t} \cdot 100 = \frac{\sum_{j=1}^n (t_{j,2} - t_{j,1})}{t_a - t_i}, \quad (5)$$

On the other hand, we have the percentage of time of the final acceleration compared with the total duration of the pattern ( $KPI_a$ ).

$$KPI_a(\%) = \frac{t_f - t_a}{t_f - t_i} \cdot 100, \quad (6)$$

The goal of designing these patterns is to determine how drivers perform inertia. The patterns show how drivers behave in terms of performance and anticipation. It is not the same trying to stop the vehicle (patterns #2 and #3) or just trying to adapt the vehicle speed to that of the traffic (pattern #1).

#### 4. Methodology

The objective of the methodology is to find out how much fuel is spent prior to one of our patterns from the point at which inertia could have been started to be used to arrive at the same point at the same speed. This amount of fuel is what the vehicle consumes during the acceleration.

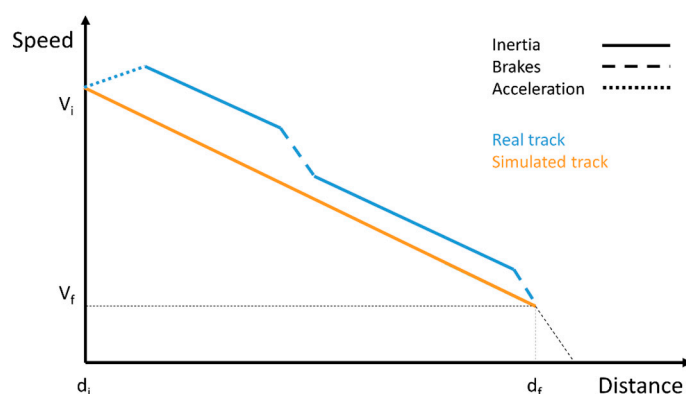
We estimate that a month provides a sufficiently large enough set of data for the methodology to be valid. From these data, we obtain the different cases in which the patterns explained in the Background section appear and select an appropriate one for our purpose. To be more representative, the selected pattern should be as long as possible to give a wide amount of results. Furthermore, the KPI should be similar to the mean KPI of the fleet, which is registered in our records from the company. We then compare the selected pattern with an ideal one. In this ideal pattern, the driver does not use the brakes until the rpm (revolutions per minute) are close to idle. While the ideal curve only shows a long inertia period, the real one alternates inertia with the use of the brakes. Thus, an acceleration period appears in the real track curve, so the initial and final speeds and positions are the same.

To find the point in which the driver should have started to exploit the vehicle's inertia, we develop a linear regression model based on the same real data that includes the studied track. Other modeling techniques could be useful to improve the accuracy of our predictions such as classification and regression trees, Knn (K-nearest-neighbour) Clustering, and Neural Networks. Nevertheless, in [19] the authors explain that when working with real world situations, variables do not change independently. Thus, multivariate regression is the appropriate method since it can assess the independent effect of one variable controlled by the others.

Furthermore, stepwise linear regression removes the weakest correlated variables after considering all candidates. Authors in [20] explain how the models resulting after using this method only include the variables that best explain the distribution of the dependent variable, according to their statistical significance ( $\leq 0.05$ ) and efficiency (increase in the *r-squared* and the *F-statistic* above a predefined threshold).

The result of this linear regression model is theoretical instant acceleration (deceleration in this case) considering speed, engine regime, engaged gear, and use of brakes at each particular point, starting with the final one according to the real curve. This final point is identified as  $(t_f, v_f)$  in the graphical representation of the patterns shown in Figures 1–3. Once the acceleration value at the final point is obtained, and knowing the elapsed time for the corresponding sample, we can calculate the speed in the instant of the previous sample and thus the travelled distance. Repeating this process, we confirm the simulated track curve.

Figure 4 shows a schematic representation of the ideal and real tracks. While the ideal curve shows how deceleration would be subject to a single inertia period, the real curve, after an acceleration period, alternates inertia with the use of the brakes, which implies harder decelerations. Both curves start and end at the same point.



**Figure 4.** Pattern of Perfect Inertia (Simulated) Compared with the Use of Brakes (Real).

#### 4.1. Data Set

For the model to be valid, we must select a large amount of real data for a single bus, regardless of the driver. The reason is that, in this case, we are not evaluating the driving skills of the driver. We only consider the difference between applying the efficient driving recommendations or not, and we want to minimise the dispersion of the vehicle's mechanics in the results. As we only use one vehicle, we have the same response if the conditions are the same.

The data must be filtered as we are only interested in inertia; that is, data with null fuel consumption. All data were captured by the cated box system [7]. The structure of the data is shown in Table 1.

**Table 1.** Parameters obtained from the vehicle.

|                                      |   |
|--------------------------------------|---|
| id_sample                            | Unique sample identifier  |
| id_vehicle                           | Unique vehicle identifier   |
| time_stamp                           | Date and hour (Day Month Year hh:mm:ss.sss)   |
| epoch_time                           | Time elapsed from 1 January 1970 at 00:00:00.000 (in ms)                            |
| distance                             | Distance travelled (in km). Low refresh ratio (at least 100 m)                      |
| fuel_rate                            | Instant fuel consumption (L/h)  |
| rpm                                  | Engine regime (rpm $\rightarrow$ min <sup>-1</sup> )                                |
| speed                                | Speed from the CAN (Controller Area Network) (km/h)                                 |
| gps_speed                            | Speed from the GPS (Global Positioning System). Eventually coverage problems (km/h) |
| cated_acc                            | Acceleration processed at the Cated Box device                                      |
| accel_pedal_pos                      | Position of the accelerator (in %)  |
| driver_demand_engine                 | Instant driver demand (in %)  |
| actual_engine_percent_torque         | Instant engine torque (in %)  |
| engine_percent_load_at_current_speed | Instant engine load (in %)  |
| trans_current_gear                   | Gear currently engaged  |
| trans_selected_gear                  | Gear selected by the ECU  |
| instant_fuel_economy                 | Instant average distance travelled with 1 L of fuel (in km)                         |
| total_vehicle_distance               | Total distance since the first day (in m)   |
| total_fuel_used                      | Total consumption since the first day (in L)  |
| latitude                             | Latitude in (°)   |
| longitude                            | Longitude in (°)  |
| brakeSwitch                          | Brakes on (1) or off (0)  |

This makes 22 parameters that we need to filter so they become adequate to the peculiarities of our study. In this case, we are only interested in the samples showing an inertia pattern. This means that all the samples we use have:

- fuel rate = 0
- accel pedal pos = 0
- driver demand engine = 0
- actual engine percent torque = 0
- engine percent load at current speed = 0
- instant fuel economy = 125.4980469 (maximum value accepted by the system)

Therefore, the rest are not relevant in the linear model, and they can be discarded. We can also discard other parameters that are not of use in this case:

- id sample: This is only an identification number and is of no use in a linear model.
- id vehicle: We only monitor one vehicle.
- time stamp: The text format makes this difficult to manage. It is better to use epoch time.
- distance: The low refresh ratio makes this useless for us.
- gps speed: This parameter is of use in those vehicles without controller area network (CAN).
- trans selected gear: This shows the same number as the trans current gear parameter with a delay of one sample.
- total vehicle distance: This has the same issue as distance.
- total fuel used: This has very low refresh ratio (0.5 liters).
- latitude: There is no need for geolocation.
- longitude: There is no need for geolocation.

Taking all of the above into account, our final set of data will be composed by:

- epoch time
- rpm
- speed
- cated acc
- trans current gear
- brakeSwitch

#### 4.2. Linear Regression Model

The objective of this linear regression model is to determine a function that allows us to verify the deceleration of the vehicle depending on the speed, rpm, engaged gear, and the use of brakes in every single moment, and the initial expression structure is:

$$y \sim 1 + x_1 + x_2 + x_3 + x_4, \quad (7)$$

where:

- $y = \text{cated acc}$
- $1 = \text{independent term (intercept)}$
- $x_1 = \text{rpm}$
- $x_2 = \text{speed}$
- $x_3 = \text{trans current gear}$
- $x_4 = \text{brakeSwitch}$

The variables  $y$ ,  $x_1$ , and  $x_2$ , are continuous, while  $x_3$  and  $x_4$ , are categorical. This means that  $x_3$  and  $x_4$  can only take exact values. In this case  $x_3 = [2,3,4,5,6]$  and  $x_4 = [0,1]$ . The reason why there is no  $x_3 = 1$  term is that, in decelerations, the ECU (Engine Control Unit) always engages the first gear in idle conditions, so fuel consumption is never null.

We used MatLab (The Mathworks, Inc. Natick, MA, USA; Matlab R2015b: Statistics and Machine Learning Toolbox) to develop the linear model we have just described.

### 5. Case Study

We have applied our method on an 18 m long articulated bus that belongs to the municipal transport company of the city of Gijón in the north of Spain. This bus is driven by different employees of the company and always does the same route. The raw data for this case study were obtained during December 2015, and the total amount is over one million samples. The bus completed the route 611 times, considering its two directions. A first filter on the raw data gives us a full set of 174,344 samples with null consumption. This filtered set conforms the data to undertake our linear regression model.

In addition, a glance at the weather internet portal <http://www.wunderground.com> shows that conditions in Gijón were very variable that month. There were 12 rainy days out of 31 and two foggy



days (low visibility), and the temperature varied from a maximum of 24 °C (15 December 2015) to a minimum of 3 °C (1 December 2015). Thus, a wide range of weather conditions is covered.

The results of a first iteration are as shown in Table 2. First we can see the function of the model and then the chart of results. The coefficients can be seen in the first column, where the (Intercept) is the independent term and the categorical variables have their different values by their sides. The Estimate column shows the estimated coefficients for the different parameters, SE shows the standard error, and  $tStat = Estimate/SE$ . Finally, the  $p$ -value column has values below 0.05, meaning that the Estimate coefficient is not valid in less than 5% of the cases. If it were higher, we would dismiss the parameter.

The second part of Table 2 shows the statistics of the model as the Number of observations or the Error degrees of freedom.

$$Error\ degrees\ of\ freedom = N^{\circ}\ of\ observations - N^{\circ}\ of\ variables$$

The Root Mean Square Error estimates the standard deviation of the error distribution. The  $R$ -squared (coefficient of determination) suggests the level of explanation of the variability in the response variable. The  $F$ -statistic versus constant model tests if there is a significant linear regression relationship between the response variable and the predictor variables. Finally, we have the  $p$ -value for the  $F$ -test of the model that shows if the model is significant or not.

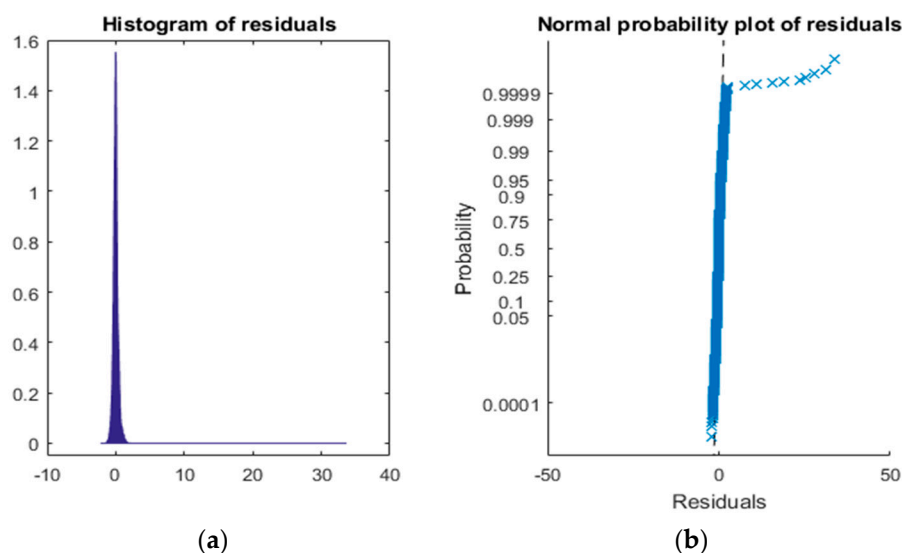
**Table 2.** Results of the first iteration.

| Linear Regression Model   |                        |                        |          |          |         |
|---|------------------------|------------------------|----------|----------|---------|
| $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4$ |                        |                        |          |          |         |
| Estimated Coefficients  |                        |                        |          |          |         |
| Variable  | Parameter              | Estimate ( $\beta_i$ ) | SE       | tStat    | p-Value |
| (Intercept)   | independent term       | −0.20829               | 0.014698 | −14.171  | < 0.001 |
| $x_1$   | rpm                    | 0.00060                | 0.000009 | 70.393   | < 0.001 |
| $x_2$   | speed                  | 0.00555                | 0.000349 | 15.943   | < 0.001 |
| $x_{3,2}$   | trans_current_gear (2) | −0.66320               | 0.014095 | −47.051  | < 0.001 |
| $x_{3,3}$   | trans_current_gear (3) | −0.89174               | 0.014336 | −62.201  | < 0.001 |
| $x_{3,4}$   | trans_current_gear (4) | −0.90528               | 0.015655 | −57.829  | < 0.001 |
| $x_{3,5}$   | trans_current_gear (5) | −0.88652               | 0.017681 | −50.140  | < 0.001 |
| $x_{3,6}$   | trans_current_gear (6) | −0.91776               | 0.021989 | −41.738  | < 0.001 |
| $x_{4,1}$   | brakeSwitch (on)       | −0.47282               | 0.001971 | −239.900 | < 0.001 |
| Number of observations  |                        | 174,344                |          |          |         |
| Error degrees of freedom  |                        | 174,335                |          |          |         |
| Root Mean Squared Error   |                        | 0.398                  |          |          |         |
| $R$ -squared  |                        | 0.358                  |          |          |         |
| $F$ -statistic vs. constant model                                 |                        | 12,128                 |          |          |         |
| $p$ -value  |                        | < 0.001                |          |          |         |

We can see how the most influential term is the engaged gear (trans\_current\_gear) as the Estimate value is higher than any other parameter coefficient for all of its categorical values. The influence of the use of the brakes is also important, and the rpm, as the result of its product by the corresponding Estimate coefficient, has a similar value. On the contrary, speed has a lower influence on the results of the regression, although it is also relevant. tStat shows that the Estimate is significantly higher than the standard error (SE) for all parameters, reflected in the  $p$ -value column, which is in all cases far lower than 0.05.

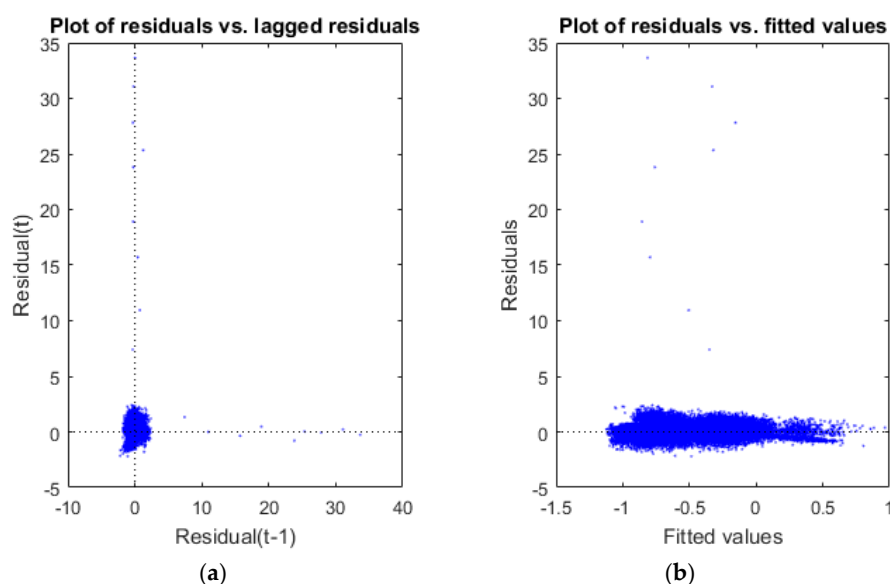
In accordance with the tStat value, the brakeswitch (on) variable would be very important. Obviously, the use or not of the brakes dramatically affects the value of the deceleration of the vehicle. In our study, we aim to characterize the deceleration when the driver uses the inertia of the vehicle. Thus, the brakeswitch (on) does not affect the results. As the engaged gear during all the deceleration process in this case is trans\_current\_gear(3), the other trans\_current\_gear variables do not have any effect on the results.

To examine more closely the results, we launched an ANOVA analysis to check the variance. Our model shows good results regarding the  $p$ -values of the coefficients and F-tests, but the coefficient of determination  $R$ -squared is too low, so we will try to improve it. If we display the residuals of the model as in Figure 5, we can see how the histogram has a very long tail. This means that the distribution of observations is not normal, so we need to adapt it to make it useful. To that end, the first step is to find and eliminate the outliers.



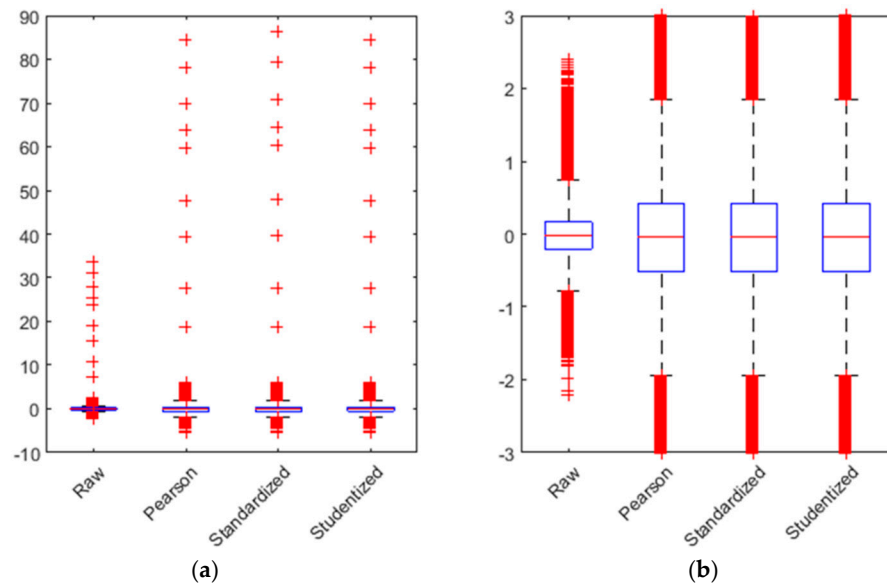
**Figure 5.** Plot of residuals for the linear model. (a) Histogram. (b) Normal Probability.

The scatter plot of residuals in the first iteration showed evidence of the existence of outliers (Figure 6).



**Figure 6.** Scatter plot of residuals. (a) Residuals versus lagged residuals. (b) Residuals versus fitted values.

In the boxplot, we can see the representation of all outliers (in the left image). We have zoomed in so the values of the valid residuals can be better appreciated (Figure 7).



**Figure 7.** Boxplot of residuals. (a) Boxplot of residuals (automatic axis). (b) Boxplot of residuals (zoomed in).

All the outliers are shown in red in the figure. We use a whisker  $w = 1.5 \sim \pm 2.7\sigma$ , which is plus/minus 2.7 times the typical deviation. With this value, the outliers are defined in Equations (8) and (9).

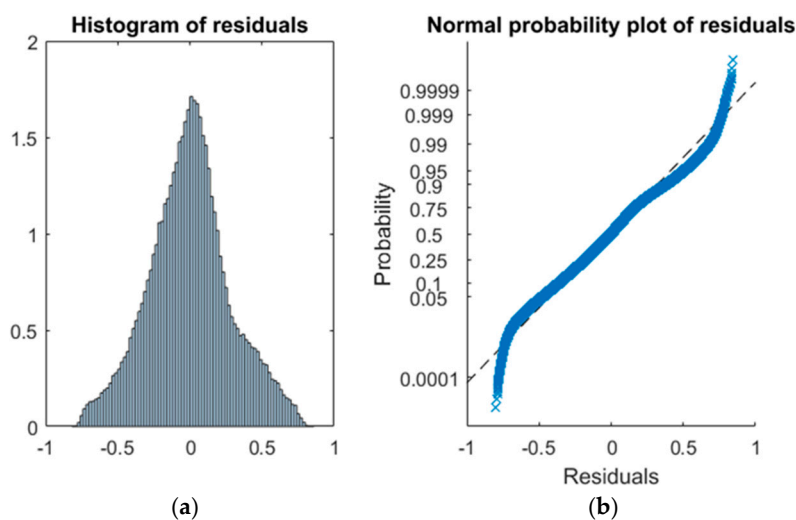
$$\text{Outliers} > q_3 + w(q_3 - q_1), \quad (8)$$

$$\text{Outliers} < q_1 - w(q_3 - q_1), \quad (9)$$

where  $w = 1.5$  is the whisker,  $q_1$  is 25th percentile, and  $q_3$  is the 75th percentile. With this approach, we eliminate 8618 outliers, which is less than 5% of the total data population, as shown in Equation (10).

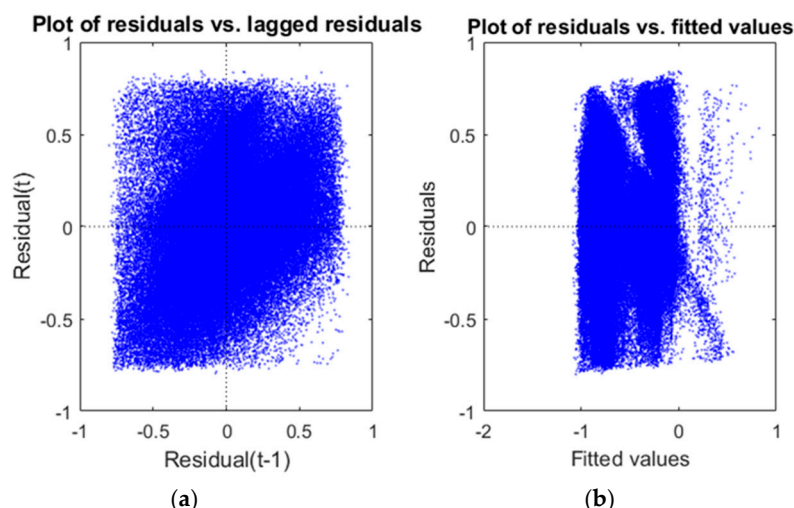
$$\frac{8618}{174,344} = 0.0495 < 0.05, \quad (10)$$

In this case, we eliminate 8618 outliers in order to improve the normality of the residuals, so the new distribution is as shown in Figure 8. Although we have removed the residuals, it is still necessary to analyse the correlations between them so we can check the quality of the model.



**Figure 8.** Plot of residuals for the linear model after the removal of outliers. (a) Histogram. (b) Normal Probability.

The scatter plot of residuals in Figure 9 shows that there are neither positive nor negative relationships between residuals and lagged residuals. That means that there is no serial correlation among residuals. Furthermore, when a model is not well fitted, a clear tendency to have bigger residuals for bigger fitted values can be seen. In our case, the fitted values figure shows that the residuals are practically independent of the measured values.



**Figure 9.** Scatter plot of residuals. (a) Residuals versus lagged residuals. (b) Residuals versus fitted values.

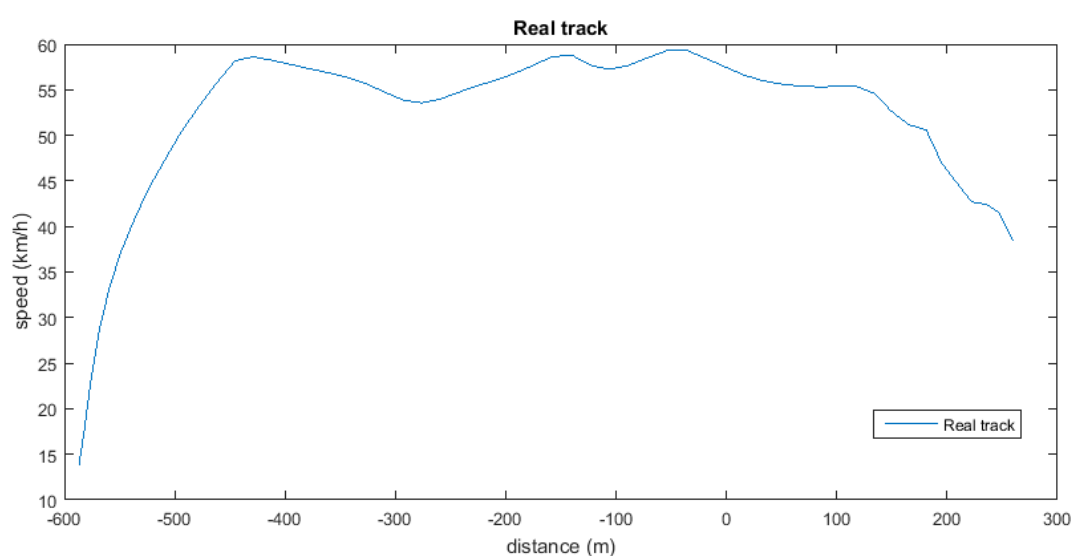
Anyway, we perform a new phase of elimination of outliers, as *R-squared* is still too low. Once the outliers have been removed, we still have 163,179 observations. In this new phase, the *R-squared* = 0.516, which is big enough for the model to be considered reliable. Table 3 shows how the *p*-values keep well below 0.05.

**Table 3.** Results of the second iteration.

| Linear Regression Model   |                        |                        |                         |         |                 |
|---|------------------------|------------------------|-------------------------|---------|-----------------|
| $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$ |                        |                        |                         |         |                 |
| Estimated Coefficients  |                        |                        |                         |         |                 |
| Variable  | Parameter              | Estimate ( $\beta_i$ ) | SE                      | tStat   | <i>p</i> -Value |
| (Intercept)   | independent term       | −0.12578               | 0.011149                | −11.282 | < 0.001         |
| $x_1$   | rpm                    | 0.00044825             | $6.1339 \times 10^{-6}$ | 73.078  | < 0.001         |
| $x_2$   | speed                  | 0.0063281              | 0.00024846              | 25.469  | < 0.001         |
| $x_{3,2}$   | trans_current_gear (2) | −0.70073               | 0.010772                | −65.053 | < 0.001         |
| $x_{3,3}$   | trans_current_gear (3) | −0.87035               | 0.010921                | −79.693 | < 0.001         |
| $x_{3,4}$   | trans_current_gear (4) | −0.86129               | 0.011798                | −73.003 | < 0.001         |
| $x_{3,5}$   | trans_current_gear (5) | −0.85393               | 0.01315                 | −64.937 | < 0.001         |
| $x_{3,6}$   | trans_current_gear (6) | −0.87933               | 0.016096                | −54.631 | < 0.001         |
| $x_{4,1}$   | brakeSwitch (on)       | −0.45049               | 0.0013913               | −323.79 | < 0.001         |
| Number of observations  |                        | 163,179                |                         |         |                 |
| Error degrees of freedom  |                        | 163,170                |                         |         |                 |
| Root Mean Squared Error   |                        | 0.271                  |                         |         |                 |
| <i>R-squared</i>  |                        | 0.516                  |                         |         |                 |
| <i>F</i> -statistic vs. constant model                                |                        | 11,800                 |                         |         |                 |
| <i>p</i> -value   |                        | < 0.001                |                         |         |                 |

In further attempts for a more complex model, we tried quadratic regression and log linear regression but did not produce coherent results as we obtained decelerations of the order of  $-5 \text{ m/s}^2$ , so we chose the model in Table 3 to calculate the increase in the costs of fuel.

From the set of data of December 2015, we selected a long complex pattern based on inertia. The ideal pattern should have many phases of pure inertia, that is, without the use of brakes and braking periods. We selected one according to the number of inertia and braking phases and with a KPI close to the mean of all the patterns during the studied period. We chose one particular situation in which the driver performed a type #1 pattern with three phases of pure inertia and two braking periods. The total time of the pattern is 20 s, and during seven of those seconds the driver used the brakes. This would make  $KPI = 35\%$ ; this is a 35% of use of the brakes. The mean KPI for this type of pattern on this bus route and in the month of study is 36%, which makes the selected track an average one. The selection of this track fulfills the aim of this paper to estimate the potential savings of the use of inertia in the company as it compensates the peaks and troughs of performance by each individual driver. Figure 10 shows the representation of the pattern in terms of speed ( $y$ -axis) and distance ( $x$ -axis). Although the bus belongs to an urban transport company, speeds of over 50 km/h implies that it is a non-urban road, which is quite common near the start and finish points of a bus route. In addition, the long distance indicates that the traffic is very light.



**Figure 10.** Real track curve of speed versus distance.

We have forced the 0 m point at the start of the pattern, so the “positive distance” is the non-consumption part of the track and the “negative distance” is where the vehicle consumes fuel. It could seem that inertia started before the “0” point, but we only have a reduction of speed due to the progressive release of the accelerator pedal by the driver. That makes a difference between deceleration, with consumption in this case, and inertia, always without consumption. Although in the part of the track prior to the start of the pattern there might be some null-consumption periods, as there are some decelerations, they are not relevant for the results as we will only add up the instant fuel consumption of all the samples included in the track.

To check the inefficiency of the real track, we apply the linear regression model, following these steps:

- (1) Starting with the speed, engaged gear, engine regime, and position data at the last sample of the real track, we apply the linear regression model. The result is the simulated instant acceleration (negative) of the vehicle if the vehicle was using inertia.
- (2) We calculate the speed of the vehicle in the instant of the previous sample.
- (3) From the result of Step 2, and knowing which gear was engaged, we can obtain the engine regime and the distance travelled in the time elapsed between the last sample and the previous one.

Now we have the data to apply the linear model, starting with Step 1 again. Therefore, we go back, calculating the ideal points of the simulated track. When we reach a point at which the ideal curve cuts the real one (ideal speed is higher than real speed), then we have the ideal starting point of inertia, as shown in Figure 11. In that point, the position and speed are the same for both the real and ideal tracks. Therefore, we now know the domain of samples that we have to analyse. We know that the consumption in the simulated track is null, as the driver would be applying inertia. In order to quantify the inefficiency of the real track, we only need to calculate the total consumption of the bus in the real track. We have the instant consumption of each sample in  $(l/h)$ , and the elapsed time between every two samples in  $(ms)$ , so for  $n$  samples we just need to apply:

$$Total\ consumption\ (l) = \left[ \sum_{i=1}^n Inst. Cons_i \left( \frac{l}{h} \right) \cdot (t_{i+1} - t_i)(ms) \right] \cdot \frac{1(h)}{3.6 \cdot 10^6(ms)}, \quad (11)$$

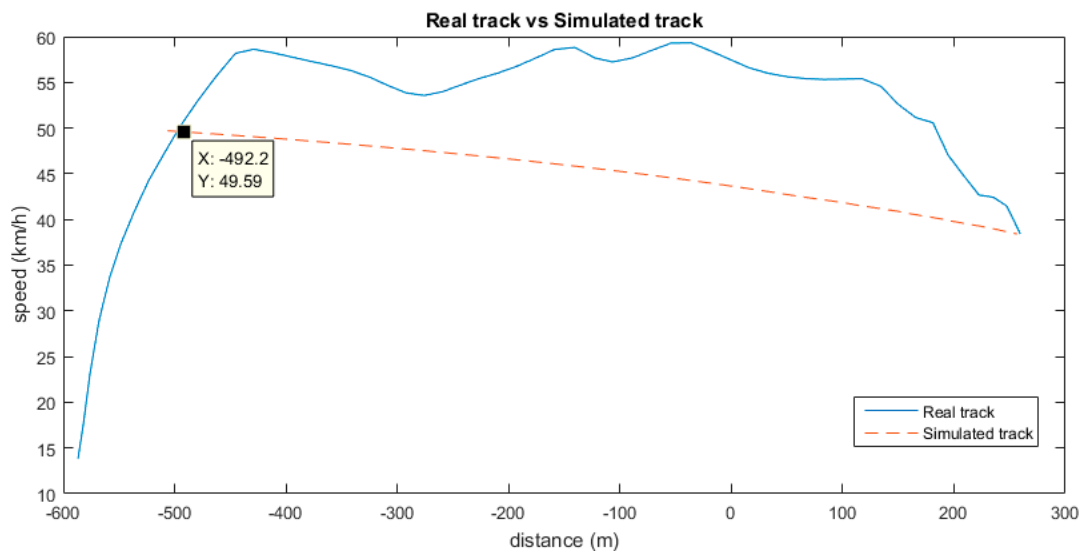


Figure 11. Real versus Simulated track curves.

With this approach, we could say that the driver would have saved 0.184 l of fuel if inertia had been used as proposed in our simulation. The case shown by the simulation is an extreme one, in which we do not consider the loss of time (11.3 s over the whole track), but it shows a good estimation of the potential savings if we managed to maximize inertia against continued accelerations and the use of brakes. The loss of time over a whole route of 12.4 km, which is this case, would be of 2 min and 44 s, extrapolating the data of the studied track.

According to the web site [geoportalgasolineras.es](http://geoportalgasolineras.es) of the Spanish Ministry of Energy, Tourism, and Digital Agenda, the fuel price in Gijón, the city where the case study takes place, varies from 0.989 €/l to 1.149 €/l (prices for the 5th of December 2016). If we use an average price of 1.069 €/l the saving in the track referred to in the case study would have been of 0.197 €. This result could be extrapolated to a longer period of time, for example, a month.

Considering that the KPI of the studied pattern is 35% of use of the brakes, we could estimate the results for the whole route, which is 12.4 km long. Also, we consider the relation between the distance run during the pattern and the total length of the studied track. In this case, these values are 260 m and a total distance of the track of 850 m.

$$Inertia\ distance\ ratio = \frac{260\ m}{850\ m} = 0.306, \quad (12)$$

The total distance run in inertia in the studied route was of 3350.81 m, making an inertia distance ratio of 0.270 and a brakes use percentage of 58.60%. With this data, the total amount of money saved by a single vehicle on a complete route with an ideal usage of the brakes would be:

$$Fuel\ saving = 0.184\ l \cdot \frac{58.60\%}{35\%} \cdot \frac{0.270}{0.306} \cdot \frac{12,400\ m}{850\ m} = 3.965\ l, \quad (13)$$

$$\text{Money saving} = 4.239\text{€/route}, \quad (14)$$

Considering that, according to our records, this vehicle completed 611 routes during the selected month, the company would have saved:

$$\text{Money saving} = 4.239 \frac{\text{€}}{\text{route}} \cdot 611 \frac{\text{routes}}{\text{month}} = 2590\text{€}, \quad (15)$$

This would be the money saved by the vehicle that has been monitored for our study. On that route there is one bus every 15 min so, considering that the time spent to complete it is slightly over 45 min, there are four buses going from header 1 to header 2 with a break at the end of every run and another four buses going in the opposite direction. If we consider that those eight buses have the same characteristics, the total monthly saving for the company on the studied route would be of 20,720 €.

## 6. Discussion and Conclusions

With these results, we can see the importance of trying to anticipate different situations while driving. According to our records, only by applying inertia adequately can savings of 59.97% of the 4318 € total expenditure on fuel that the vehicle had during the month of study be achieved. Needless to say, this is an extreme top saving case, but it is a good theoretical estimation of the potential that efficient driving has in urban transport companies. It is important to point out that we used real consumption data directly obtained from the vehicle during working time in real conditions to make the calculations of the case study. It was not possible to verify the claims in a prepared test scenario as the vehicles are not available for off-duty activities, but the deceleration values obtained with the linear regression are very realistic, given the real data used to develop them.

These results are obviously limited to the conditions defined in the case study. The bus route, the weather conditions, which were very variable throughout the month, or the type of bus could also influence the final figures. Of course, the drivers using the monitored vehicle could have made a difference as well. Nevertheless, we consider that the results may be generalized since we did not apply the study to one particular driver and the objective of the paper is not the evaluation of drivers. In addition, we are aware of the seasonality of consumption. The selected month has the most similar consumption to the monthly average consumption, the consumption in July being the highest (69.70 l/100 km) and in January the lowest (61.26 l/100 km). Apart from that, the gathered data include peak hours on the one hand and weekends, with a very low density of passengers and traffic, on the other, so the whole spectrum of weight and traffic density range is considered.

The reliability of the mathematical model resides in the repeatability of results when conditions are the same. Thus, the use of several buses would have reduced this repeatability due to mechanical differences caused by each bus being a different model of different age or because each is involved in a different stage of maintenance. Our saving estimation for all the vehicles working in this route is still valid as they are all the same model, but if we apply the study on each bus, more precise results would be obtained.

We based the comparison of the study on the inertia patterns because they give a clear base (null consumption) to calculate the savings. In our previous work [5], we defined other efficient driving patterns that also affect fuel consumption, but they are not so easily quantifiable as inertia. The reason is that we would have to compare their absence instead of their presence, as they represent inefficiency.

Regarding the complex inertia patterns, they showed very little variability in terms of consumption. They were originally designed to describe the behaviour of the driver, as it is not the same to exceed the use of the brakes while stopping as to do it when trying to adapt your speed to that of the traffic. However, all of the patterns and their KPIs show a similar structure when we are talking about expending fuel. This paper demonstrates that the exploitation of the vehicle's inertia is enough to save fuel, but, depending on how well you use the previously gained energy, the savings could be much higher.

Any company applying an efficient driving methodology will achieve savings in fuel and money, as well as improving safety and comfort. Although our evaluation method is not directly based on fuel consumption, we conclude that it has proved to be very effective when it comes to reducing costs and improving the productivity of the company. In addition, this behavior based evaluation method could be a fair starting point for a reward system to incentivize the correct use of efficient driving techniques, as it would minimize dependence on the environmental parameters that affect fuel consumption such as seasonality or weight variations due to changes in passenger density when it comes to evaluating drivers' performances. The findings in this paper would enable the companies and their drivers to envisage which techniques need revising and improving.

Future work aims to develop an automatic recommendation system that allows drivers to receive proper information about how to approach the ideal inertia performance indicated in this paper. We will also work with real data to see which combination of actions of the driver and efficient driving patterns would allow him/her to approach the ideal, simulated track without a loss of time.

**Acknowledgments:** This work has been supported by the General Direction of Traffic (DGT) of Spain through the project SPIP2015-01686, the the Spanish National Research Program within the project MINECO-13-TIN2013-41749-R and the Science, Technology, and Innovation Plan of the Principality of Asturias, within the project GRUPIN-14-065.

**Author Contributions:** Xabiel G. Pañeda conceived the idea of the paper and coordinated all the works; Alejandro G. Tuero wrote the paper, conceived the experiment, and analyzed the data; Laura Pozueco extracted and organized the data from the database; Roberto García, Gabriel Díaz, and David Melendi revised the text and contributed to the design and quantification of the experiment with the support of Laura Pozueco; and Abel Rionda and David Martínez obtained the data from real vehicles and stored them in the database. All authors have read and approved the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Jochem, P.; Rothengatter, W.; Schade, W. Climate change and transport. *Transp. Res. Part D Transp. Environ.* **2016**, *45*, 1–3.
2. Kannan, N.; Saleh, A.; Osei, E. Estimation of Energy Consumption and Greenhouse Gas Emissions of Transportation in Beef Cattle Production. *Energies* **2016**, *9*, 960.
3. Moss, D. *The Efficient Driver's Handbook—Your Guide to Fuel Efficient Driving Techniques and Car Choice*; Veloce Publishing Ltd: Dorchester, UK, 2010.
4. Sivak, M.; Schoettle, B. Eco-driving: Strategic, tactical, and operational decisions of the driver that influence vehicle fuel economy. *Transp. Policy* **2012**, *22*, 96–99.
5. Pañeda, X.G.; Garcia, R.; Diaz, G.; Tuero, A.G.; Pozueco, L.; Mitre, M.; Melendi, D.; Pañeda, A.G. Formal characterization of an efficient driving evaluation process for companies of the transport sector. *Transp. Res. Part A Policy Pract.* **2016**, *94*, 431–445.
6. Pozueco, L.; Tuero, A.G.; Paneda, X.G.; Melendi, D.; Garcia, R.; Paneda, A.G.; Rionda, A.; Diaz, G.; Mitre, M. Adaptive learning for efficient driving in urban public transport. In *Proceeding of the 2015 International Conference on Computer, Information and Telecommunication Systems (CITS)*, Gijon, Spain, 15–17 July 2015; pp. 1–5.
7. Rionda, A.; Pañeda, X.G.; García, R.; Díaz, G.; Martínez, D.; Mitre, M.; Arbesú, D.; Marín, I. Blended learning system for efficient professional driving. *Comput. Educ.* **2014**, *78*, 124–139.
8. Pañeda, A.G.; Pozueco, L.; Melendi, D.; Pañeda, X.G.; García, R.; Tuero, A.G.; Rionda, A.; Orueta, G.D.; Arciniegas, J.L. An Architecture for a Learning Analytics System Applied to Efficient Driving. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje* **2016**, *11*, 137–145.
9. Alam, M.S.; McNabola, A. A critical review and assessment of Eco-Driving policy & technology: Benefits & limitations. *Transp. Policy* **2014**, *35*, 42–49.
10. Zarkadoulas, M.; Zoidis, G.; Tritopoulou, E. Training urban bus drivers to promote smart driving: A note on a Greek eco-driving pilot program. *Transp. Res. Part D Transp. Environ.* **2007**, *12*, 449–451.
11. Strömberg, H.K.; Karlsson, I.C.M. Comparative effects of eco-driving initiatives aimed at urban bus drivers—Results from a field trial. *Transp. Res. Part D Transp. Environ.* **2013**, *22*, 28–33.



12. Dogan, E.; Steg, L.; Delhomme, P. The influence of multiple goals on driving behavior: The case of safety, time saving, and fuel saving. *Accid. Anal. Prev.* **2011**, *43*, 1635–1643.
13. Boriboonsomsin, K.; Vu, A.; Barth, M. *Eco-Driving: Pilot Evaluation of Driving Behavior Changes among U.S. Drivers*; UCTC Research Paper 2010; University of California Transportation Center: Berkeley, CA, USA, 2010.
14. Qian, G.; Chung, E.; Horiguchi, R. Assessing eco-driving performance of a traffic platoon at urban intersections: An experimental study. In Proceedings of the International Symposium on Recent Advances in Transport Modelling (OPTIMUM), Kingscliffe, Australia, 21–23 April 2013.
15. Duarte, G.O.; Gonçalves, G.A.; Farias, T.L. Vehicle monitoring for driver training in bus companies—Application in two case studies in Portugal. *Transp. Res. Part D Transp. Environ.* **2013**, *18*, 103–109.
16. Klunder, G.A.; Malone, K.; Mak, J.; Wilmink, I.R.; Schirokoff, A.; Sihvola, N.; Holmén, C.; Berger, A.; de Lange, R.; Roeterdink, W.; et al. *Impact of Information and Communication Technologies on Energy Efficiency in Road Transport—Final Report*; TNO: The Hague, The Netherlands, 2009.
17. Mensing, F.; Trigui, R.; Bideaux, E. Vehicle trajectory optimization for application in ECO-driving. In Proceedings of the 2011 IEEE Vehicle Power and Propulsion Conference, Chicago, IL, USA, 6–9 September 2011; pp. 1–6.
18. Saboohi, Y.; Farzaneh, H. Model for developing an eco-driving strategy of a passenger vehicle based on the least fuel consumption. *Appl. Energy* **2009**, *86*, 1925–1932.
19. Walnum, H.J.; Simonsen, M. Does driving behavior matter? An analysis of fuel consumption data from heavy-duty trucks. *Transp. Res. Part D Transp. Environ.* **2015**, *36*, 107–120.
20. De Abreu e Silva, J.; Moura, F.; Garcia, B.; Vargas, R. Influential vectors in fuel consumption by an urban bus operator: Bus route, driver behavior or vehicle type? *Transp. Res. Part D Transp. Environ.* **2015**, *38*, 94–104.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).