Evaluating Eco-driving Advice using GPS and CANBus data

Karsten Jakobsen
Aalborg University
Institute of Computer Sience
Aalborg, Denmark
Email: karstenjjakobsen@gmail.com

Sabrine C. H. Mouritsen
Aalborg University
Institute of Computer Sience
Aalborg, Denmark
Email: sabrinechm@gmail.com

***Abstract*—Road transportation accounts for about 20% of the emission of CO_2 in EU. This can be reduced by driving more fuel efficient with the added benefit of lower expenditure on fuel. Several advice on fuel efficient driving behaviour exists, but these are very abstract and lacks to be grounded in data. A large set of real, high frequency GPS and CANBus data from four similar vehicles is used to evaluate multiple eco-driving advice. CANBus data provides information about a vehicle such as the fuel consumption and the rounds per minute of the engine. It is estimated that idling costs about $1l/h$ in wasted fuel, and the fuel consumption per kilometer is $76ml/km$ more when driving at $130km/h$ compared to $80km/h$. Examples of stopping at traffic lights indicate an increase in fuel consumption of up to twice as much. The four vehicles are compared and evaluated on their ability to follow the eco-driving advice. There is a 22 % difference between the fuel efficiency of the drivers in this study. Therefore there is potential for saving large quantities of fuel by informing and educating drivers. No single eco-driving advice dominates, but multiple advice are found to have a positive effect on the fuel efficiency when followed.**

I. Introduction

Reducing fuel consumption and greenhouse gas (GHG) emissions is one major approach towards controlling global warming. The focus has mainly been on large consumers as power plants, air planes, etc. but small consumers will have to contribute as well if the goals for GHG reductions shall be met. About one fifth of EU's total emission of CO_2 comes from road transportation and has increased with about 23 % from 1990 to 2010 [5]. Politicians are focusing on setting guidelines and requirements for improving vehicular technologies such as engine performance and alternative fuels, but the individual drivers can also contribute by driving more fuel efficient. This will not only help reduce GHG emissions but also reduce fuel expenditure for the drivers.

Eco-driving aims to change the driving behaviour through simple advice such as maintain a steady speed, accelerate moderately, follow speed limits, anticipate traffic flow and maintain your vehicle [2], [18]. The advice is designed to reduce fuel consumption and hence reduce GHG emissions. Traffic safety and improvement of traffic flow have also shown to be a positive side effects. Feedback on ones eco-driving performance is imperative for improvement. This can either be simple visualisations of fuel consumption or more elaborate observations. These observations could for example be whether the driver accelerates too much or idles. Access to detailed information of driving patterns are needed in order to make these observations. GPS data provides location, speed and direction at a high frequence, and the data is easy and cheap to collect. Additional data is, however, needed in order to evaluate how much the eco-driving advice is followed. CANBus data can provide detailed information about the dynamic state of the vehicle, e.g. rounds per minute, fuel consumption per second and more. As of today, CANBus data is not yet available in the same quantity as GPS data, and the quality varies.

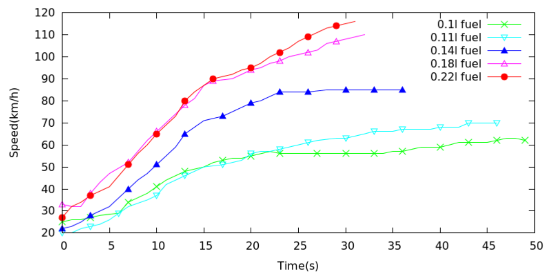


Figure 1. Five trips driving the same road segments with different accelerations

Vehicle id	Average km/l	Standard Deviation
1	6.75	2.63
2	7.46	2.79
3	5.77	2.57
4	6.79	2.73
Average	6.77	2.75

Table I. Fuel efficiency

We have access to GPS and CANBus data from four vehicles over a period from august 2012 to april 2013. Analyses of eco-driving advice have, to the best of our knowledge, not yet been made on such an extensive data set. Table I shows the average km/l driven by the four vehicles (the data will be explain later in Section VI). Vehicle 3 is the least fuel efficient with only $5.77km/l$ and vehicle 2 is the most fuel efficient vehicle with $7.46km/l$. With a gas price of

12 kr/l vehicle 2 spend 2,080 kr a week if he drives 1,000 km where as vehicle 2 spend 1,624 kr and vehicle 1 and 4 only around 1,775 kr a week. Over a year vehicle 3 could save a around 22,800 kr if he is able to drive with the same fuel efficiency as vehicle 2. We aim to identify which factors available through GPS and CANBus data, has an influence on fuel consumption and how one explains the difference in the fuel consumption of different vehicles of the same type. That is, why does vehicle 2 use less fuel than vehicle 3 and can vehicle 3 improve his fuel efficiency using eco-driving strategies? High accelerations are one potential factor for fuel consumption. Figure 1 shows five different acceleration profiles driving the same 800 m road and the amount of fuel consumed in the depicted period. Clearly high accelerations leads to high fuel consumption in this example. The remainder of this article will extend this and other analyses of the driving behaviours and investigate which factors influence their fuel consumption.

It is a basic assumption that the GPS and CANBus data provides ground-truth results, meaning that the recorded values are assumed to be correct. Some instances, e.g. speed and fuel consumption contains values that are obviously wrong, and are either adjusted or disregarded from the data foundation. This is explained later in Section IV-C. It is another basic assumption that one vehicle has one driver. The data provider states that this is correct. The main contributions of this paper are:

- First experimental evaluation of fuel efficiency based on a large set of real CANBus data
- Concrete cost estimates of moderate accelerations versus hard accelerations, of driving fast, of stopping at a red light versus driving through and of idling
- Comparison of vehicles ability to follow eco-driving techniques
- Concrete advice on how to further improve the analysis of eco-driving advice, e.g. minimal requirements on CANBus data

II. Related Work

The authors of [6] investigates the impact of providing real time eco-driving advice to the drivers based on real-time traffic speed, density and flow. They find that a reduction in fuel consumption of 10-20% can be achieved without a significant increase in travel time, and that the effect is greater in severely congested scenarios. The authors of [8] also investigate how real-time feedback affects driving behaviour. From simply displaying the instantaneous fuel economy to 20 sample drivers they show a reduction of 6 % in fuel consumption on city streets and 1% on highways. Most of the drivers were willing to adopt parts of the eco-driving advice after the study. The long-term effects of eco-driving courses are evaluated in [11]. A study on 10 vehicles over 10 months shows a mean reduction in fuel consumption of 5.8 %, but also show that the effect is very different from individual to individual. 20% saw no fuel reduction. Fuel consumption at high speeds and at aggressive accelerations are investigated in [7]. They find that reducing the velocity on highways generally gives about the same reduction in fuel as reducing acceleration on all roads. The paper [9] investigates why the benefit of eco-driving decrease over time. The study finds that group behaviour needs to be taking into account when teaching eco-driving principles.

The exact eco-driving advice vary from reference to reference. Both [2] and [18] details a number of advice. See more in Section III.

Eco-routing is about saving fuel by finding the most fuel efficient routes. GreenGPS [12] is an example of participatory approach to eco-routing and they see a 10 % reduction in fuel consumption. EcoMark [13] is an evaluation framework for evaluating environmental models. Eleven known models for environmental impact are evaluated to investigate whether they can be used to do eco-driving and eco-routing. The evaluation finds that instantaneous models can be used for eco-driving and aggregated models can be used for eco-routing. The INTEGRATION model framework [16] is a model for quantifying environmental impact on a microscopic level. The study shows that the predicted emissions and fuel consumption are consistent with actual data from a field study when the vehicles do not accelerate.

III. Eco-driving Advice

The authors of [2] and [18] list a number of advice for eco-driving that will reduce fuel consumption. The advice is detailed below and will be evaluated in the following sections separately.

1) **Drive in the highest possible gear at lowest possible rounds per minute (RPM).** Fuel consumption is lower at low RPM due to friction. Keep a high load on the engine and shift gear at around 2,500 RPM for gasoline cars and 2,000 RPM for diesel cars.
2) **Maintain a steady speed.** Fuel is primarily consumed when accelerating. Constantly breaking and accelerating will use more fuel than maintaining a steady speed.
3) **Anticipate traffic flow and avoid frequent starts and stops.** Adjusting the speed to traffic lights, turns and other vehicles in good time will make it easier to maintain a steady speed.
4) **Decelerate smoothly.** Slow down by using the engine brake or the neutral gear instead of the actual brakes. Modern vehicles use no fuel when using the engine brake, i.e. the vehicle is in gear and the accelerator is released. It also reduces wear and tear and reduce exhaust emissions and increase traffic safety and flow.
5) **Accelerate moderately.** High accelerations consume much fuel, and one should accelerate in low gears with the RPM below 2,500 (2,000 for diesel) and with the throttle at half position.
6) **Eliminate idling.** It is more fuel efficient to switch off the engine than leaving the engine

running. The average modern vehicle use around 0.9-1.3 liters per hour (l/h) during idling [19].

7) **Drive at or below the speed limit.** Fuel consumption increases at higher speeds.
8) **Do not press the accelerator when switching on the engine.** This is not necessary in modern vehicles and only consumes fuel.
9) **Approach curves at correct speed and in the highest possible gear.** This will reduce the need for acceleration after the curve and improve traffic safety.
10) **Minimise extra weight and air resistance.** Both increase the load on the engine and thereby increase fuel consumption.
11) **Maintain correct tyre pressure.** Incorrect tyre pressure increases the rolling resistance and thereby the fuel consumption.
12) **Avoid fuel consuming accessories.** Air-conditioning and other accessories consume fuel.

IV. DATA FOUNDATION

A number of data sources are available with information about vehicles.

A. Global Positioning System (GPS) Data

GPS's provide spatio-temporal information with high accuracy and reliability at a low cost. GPS data is therefore often used when analysing driving behaviour and patterns because much data exists. GPS data contains a vehicles latitude and longitude position, speed and direction at some UTC time with some frequency. With such information it will for example be possible to analyse how well the driver is at keeping a steady speed (advice 2) and how he accelerates (advice 5). By comparing the positions with a map it will also be possible to analyse how well the driver anticipates the traffic flow for example avoids stopping at traffic lights (advice 3). The coordinates can also be matched to road segments on a map, from which one for example can see the types of roads that are used.

B. CANBus Data

CANBus data gives access to the state of a vehicle providing more detailed information which allows more detailed analysis of driving behaviour. CANBus data always annotates GPS data, but as of today, little GPS data with CANBus information is available. CANBus data can include many different values, here we only mention the most common. The engines RPM will indicate the load on the engine, e.g. if the vehicle is turned off or idling. The current gear can be utilised to understand if the driver is in neutral gear and changes gears at the correct moments. Knowing the driving distance is also useful and can be accessed through the vehicles odometer. Fuel consumption is often available in different formats, e.g. the fuel level in the tank, the instantaneous fuel consumption and the total fuel consumed. Instantaneous fuel consumption is estimated based on other CANBus data such as RPM and fuel flow. This makes it a good estimate of fuel consumption at drive time, but it is too inaccurate at an aggregate level. The total fuel consumption is more accurate when looking at consumption over time. The acceleration can be available in the CANBus data, but can also be calculated from the speed. The position of the throttle indicates how aggressively the vehicle is driven. An engine works best at certain temperatures, and this temperature can also be available through the CANBus.

C. Available Data Foundation

The data set contains records from four real-life diesel vehicles of a minibus type. The area that they operate in can be seen in Figure 31. All vehicles are assumed to be comparable based on statements from the data provider - a statement that will be challenged in the end of this paper. Table II list the number of records of each vehicle and the time the data expands over. About 99% of the data is recorded with 1-2 Hz frequency while the remaining is collected at lower frequencies (See Figure 32). The last 1% of low frequency data is discarded because high frequency data is needed in parts of the analysis. It is likely that the low frequency data was collected when the equipment was installed and then later configured to record with a higher frequency.

Vehicle id	Number of records	Start date	End date
1	1,847,170	2012-08-03	2013-04-30
2	1,917,610	2012-08-07	2013-04-30
3	938,068	2012-08-01	2013-04-30
4	1,862,091	2012-08-02	2013-04-30

Table II. VEHICLE STATISTICS

In the following, we will be using eight data values from the data set. The remaining are either lacking data, is erroneous or not usable in this context. Let **r** be a recording and **r**.**vid** be the vehicle identifier, **r**.**time** be the timestamp of the recording, **r**.**lng** and **r**.**lat** be the longitude and latitude position, **r**.**speed** be the speed in km/h as an integer, **r**.**RPM** be the RPM of the engine, **r**.**kmcounter** be the read-out of the odometer and **r**.**fuel** be the total fuel consumption of the vehicle with a granularity of $10ml$. Fuel consumption can on rare instances be measured incorrectly (see Table XVII). Figure 45 shows the distribution of fuel consumption. A fuel consumption above $20ml/s$ per second is unrealistic and rare and is hence considered an outlier and not included in analysis. Records with a negative fuel consumption are considered errors and not included. If the GPS looses signal, then the speed values might be set to zero or interpolated to smoothly decrease and then smoothly increase. Acceleration is calculated from the speed, and errors in the speed may sometimes lead to very unrealistic acceleration values (See Table XVI in appendix). To account for this we try to even out the speed values if the acceleration is unrealistic. From [10] the maximum deceleration of a passenger car is $-5.61m/s^2$. Figure 2 shows the distribution of recorded accelerations where values above 25,000 are maximised out. The lack of data for some of the acceleration values are due to speed being measured as integers - this is further discussed in Section XIII. The maximum deceleration is set to $-6m/s^2$ to remove data rows with unrealistic decceleration values. Results in article [17] show a maximum acceleration to $3.5m/s^2$ and figure 2 shows that the majority of the accelerations are less then $3.5m/s^2$. This value is hence used as the maximum acceleration and any accelerations above are considered unrealistic. Acceleration is measured in m/s^2 and always registers as positive values, contrary to decceleration which is always registered in negative values.

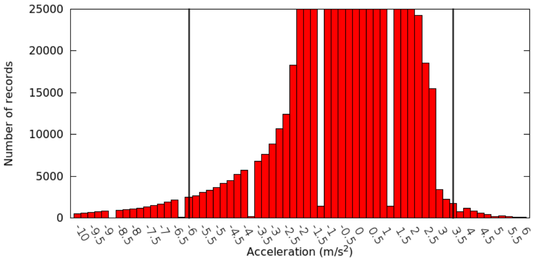


Figure 2. Distribution of acceleration values. Values from -2 to 2 m/s^2 are maximised out for sake of readability

Algorithm 1 shows how the new speed value is calculated from the original values of a trip. Line 1 orders all records of the trip on their timestamp. In line 2-3 we begin looping over all records in the trip except the first as we cannot calculate an acceleration for that. In line 4 we use Algorithm 3 to calculate the acceleration from the previous record to the current record. Line 5 checks if the acceleration is unrealistic. If it is, we continue to calculate the acceleration from the previous record to the next record in line 6. If this acceleration is within the realistic values, then the current speed is interpolated from the previous and next record using algorithm 4 in line 8-9. The value is saved in a temporary column and is hence not used in the next iterations of the algorithm. If $acc2$ is not within realistic values then the record is discarded and R_nSpeed is set to $null$. In line 14-15 we continue with the next record. Finally, in line 16-18, we update speed to the temporary values. Algorithm 4 performs a simple linear interpolation. It uses four arguments: the data set, the index to interpolate and the two indexes to create the interpolation from. Line 1 calculates the slope and line 2 calculate the interpolated value.

The tool M-GEMMA [15] is used to match each record to a road segment from a map from OpenStreetMap (OSM)[3] using the latitude and longitude coordinates. The map-matching process uses a collection of records to match a record to a segment and annotates the data set with a **r**.**segmentkey** referring to a road segment on the map and a **r**.**mdirection** being either **Forward** or **Backward** on the segment. M-GEMMA filters out 33 % of the records partly because some of the records are outside of the map and partly due to various criteria in the map-matching process

Algorithm 1 AdjustSpeed(**trip**)

1: $\mathbf{R} = \mathbf{trip}.orderRecordsOnTime()$
2: $i = 1$
3: **while** $i < len(\mathbf{R})$ **do**
4: $\quad acc = CalculateAcc(\mathbf{R}_{i-1}, \mathbf{R}_i)$
5: $\quad$ **if** $acc < -7.5$ **or** $acc > 3.5$ **then**
6: $\quad\quad acc2 = CalculateAcc(\mathbf{R}_{i-1}, \mathbf{R}_{i+1})$
7: $\quad\quad$ **if** $acc2 > -7.5$ **and** $acc2 < 3.5$ **then**
8: $\quad\quad\quad s = InterpolateSpeed(\mathbf{R}, i, i{-}1, i{+}1)$
9: $\quad\quad\quad \mathbf{R}_i$.**nSpeed** $= s$
10: $\quad\quad$ **else**
11: $\quad\quad\quad \mathbf{R}_i$.**nSpeed** $= null$
12: $\quad\quad$ **end if**
13: $\quad$ **end if**
14: $\quad i{+}{+}$
15: **end while**
16: **for** j **in R do**
17: $\quad \mathbf{R}_j$.**speed** $= \mathbf{R}_j$.**nSpeed**
18: **end for**

(e.g. records too far away or opposite headings) and these records are not map-matched to a road segment.

Some of the advice from Section III can be evaluated with the avaliable data and some cannot. Table III gives an overview of which advice can be evaluated with what data. The category 'Other' indicates that additional information not available in the data set is needed. This includes unavailable CANBus data such as gear, tyre pressure, air conditioning status and weight of the vehicle. 'Other' also includes none-CANBus data such as weather information, traffic information, height data and road pavement. Advice 2, 6 and 7 can be evaluated solely on the provided data set. Advice 1 requires both the state of the gear and the RPM, and as the gears are unavailable, this advice will instead only be evaluated from the speed. Anticipating traffic flow (advice 3) includes several aspects, some of which can be evaluated. One major aspect is avoiding traffic lights or adjusting the speed to the traffic lights, such that the driver avoids stopping. Other aspects, such as adjusting speed to other vehicles is not possible to evaluate. It is, however, possible to estimate the impact of the traffic density by looking at which roads are used and what the corresponding fuel efficiency is. Accelerating moderately (advice 5) can be evaluated on this data set but only with the rate of acceleration and speed. The gears would have provided additional information. Advice 4 about smooth deceleration requires knowledge of the gears and breaks as this advice mainly is about using the gears and breaks to slow down. It can hence not be evaluated. Advice 8 requires the position of the throttle, which is not available. Approaching curves at correct speed and gear (advice 9) is difficult to evaluate with this data set as too much of the required information is lacking. Advice 10, 11 and 12 is not evaluated as they require other data than what is available. It might be possible to evaluate them if enough data to extract the normal behaviour of the vehicle is available.

Advice	GPS	CANBus	Map	Other
1: High gears	-	✓	-	✓
2: Steady speed	✓	-	-	-
3: Anticipate flow	✓	-	✓	✓
4: Deceleration	✓	✓	-	✓
5: Acceleration	✓	-	-	✓
6: Idling	✓	✓	-	-
7: Speed limit	✓	-	✓	-
8: Startup	✓	✓	-	✓
9: Curves	✓	-	✓	✓
10: Weight & air	-	-	-	✓
11: Tyre pressure	-	-	-	✓
12: Accessories	-	-	-	✓

Table III. DATA BASIS FOR EVALUATING ECO-DRIVING ADVICE.

V. DATA GROUPING

Only grouping the data on vehicles will give too few and too diverse groups of data on which it will be difficult to make usable conclusions. The data therefore needs to be grouped into smaller units. In the following we will use two different grouping strategies, periods and trips, explained in Sections V-A and V-B. Figure 3 shows the two concepts. The top lines indicate three trips where the "gaps" represent time between two trips. A trip can thereafter contain smaller periods indicated by the boxes below. A period is a stretch of time where some property holds, e.g. the vehicle drives at the same speed or the vehicle is accelerating.

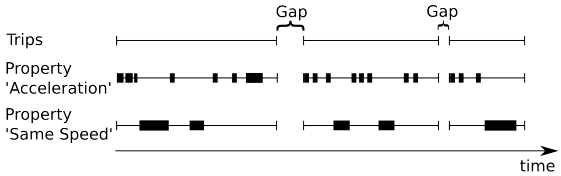


Figure 3. Example of periods and trips

A. Trips

Trips are a temporal and sometimes spatial grouping that primarily looks at the time difference between trips, the length of the trip and sometimes the location of the vehicle. The data set is split into trips by annotating each record with a trip identifier, $\mathbf{r}.\mathbf{tid}$. A trip, $\mathbf{trip}_i$ is defined as a consecutive sequence of at least 30 records with the same vehicle identifier where the engine is turned on and any two consecutive records are within 120 seconds.

$$\mathbf{r}_{j+1}.\mathbf{time} - \mathbf{r}_j.\mathbf{time} < 120$$
$$|\mathbf{trip}_i| \geq 30$$
$$\mathbf{r}_j.\mathbf{tid} = \mathbf{r}_{j+1}.\mathbf{tid} = \mathbf{trip}_i.\mathbf{tid}$$
$$\mathbf{r}_j.\mathbf{vid} = \mathbf{r}_{j+1}.\mathbf{vid} = \mathbf{trip}_i.\mathbf{vid}$$

where j ranges over the records in the trip.

A trip is defined from when the engine is running, that is when $\mathbf{r}.\mathbf{RPM} > 0$, because idle time, i.e. when the engine is running but the vehicle is not moving (see Section XI), is an important factor for fuel consumption. We therefore need to ensure that these records are included in the trips. In order not to split a trip just because the engine stalls, we say that a trips ends when the time gap between two consecutive records is too large. Figure 4 shows the number of trips when varying the time gap from 5 seconds between two trips to 200 seconds. We see that the curve flattens around 120 seconds and we choose this as the gap. Short trips with few records will not give a usable idea of which factors influence fuel consumption. Figure 5 shows the number of trips varying the minimum number of records in a trip. The curve flattens around 30 records, which in most cases will correspond with 30 seconds. A similar test using time as the minimum requirement on trips show no clear result. Table IV shows the total time span on all trips of the vehicles and how many kilometers they have driven. A work day is assumed to be 8 hours.

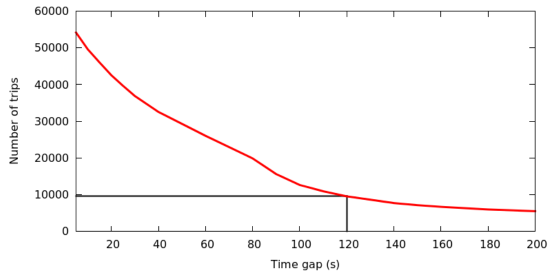


Figure 4. Number of trips at different time gaps

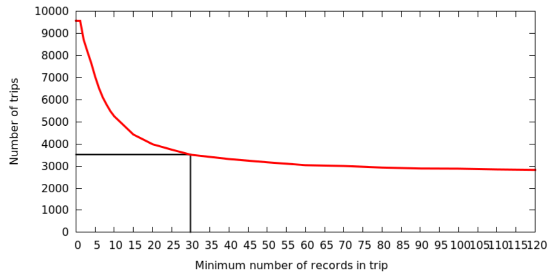


Figure 5. Number of trips at different lengths

B. Periods

It will some times be more interesting to look at records with a similar property than looking solely at trips. Let a period be a sequence of consecutive records within a trip with some similar property. The example in Table V shows two period groupings on a small sample set. The property of period 'Same Speed' is that the speed is exactly the same in all records in the period. The sample set hence contains 2 periods with this property. The property of period 'Acceleration' is that the vehicle is accelerating which results in two different periods.

Vehicle id	Time span of trips s	$work\ days$	Kilometers driven km
1	2,154,336	74.80	18,877.88
2	2,198,987	76.35	23,069.04
3	1,159,756	40.27	17,356.94
4	2,224,390	77.24	28,705.04
Average	1934367.25	67.17	22,002.22

Table IV. TRIP STATISTICS

record id	speed	tid	Period 'Same Speed'	Period 'Acceleration'
$\mathbf{r}_1$	59	0		a
$\mathbf{r}_2$	60	1	1	a
$\mathbf{r}_3$	60	1	1	
$\mathbf{r}_4$	60	1	1	
$\mathbf{r}_5$	63	1		b
$\mathbf{r}_6$	64	1		b
$\mathbf{r}_7$	67	1	2	b
$\mathbf{r}_8$	67	1	2	
$\mathbf{r}_9$	67	2		

Table V. EXAMPLE OF PERIOD GROUPING

VI. CLASSIFYING DATA

Fuel consumption is a measure of effectiveness, and how many kilometers per liter fuel (km/l) a vehicle drives will indicate how fuel efficient the vehicle is. The main goal of a vehicle is transportation and hence, the more kilometers one can drive per liter fuel, the cheaper it is. Let $\mathbf{trip}_i\mathbf{.kml}$ be the total number of kilometers driven divided by the total fuel consumption of $\mathbf{trip}_i$.

Figure 6 plots the distributon of $\mathbf{trip}_i\mathbf{.kml}$ for all trips grouped in intervals of $0.25km/l$. All trips are classified into four classes based on their fuel efficiency which are marked as vertical lines on the figure. The majority of the trips have a $\mathbf{trip}_i\mathbf{.kml}$ between 3.5 to 11.5 and the curve peaks at $8.125km/l$. A class 'outliers' is made of the trips with unusually low fuel efficiency, being those where $\mathbf{trip}_i\mathbf{.kml}$ is less than $3.5km/l$. The trips in class 'outliers' are the irregular trips, e.g. very short trips or idling trips. The distribution of the remaining trips resembles a normal distribution. We split these remaining trips into three equally sized classes, 'low', 'medium' and 'high' each containing one third of the trips. Three classes are chosen because fewer classes makes it difficult to distinqush the more fuel efficient trips from the less, and having more classes makes it increasingly complex to understand for an end-user. The classes contains no upper outliers because this study focuses on identifying fuel inefficiency and because no clear outliers are found in the data set. Constructing the classes from the data itself means that the classes depend on the nature of the data and may be different for other types of vehicles. The four classes are

$$\begin{aligned}
\text{'outliers'}: \quad & 0 && \leq \mathbf{trip}_i\mathbf{.kml} < 3.5 \\
\text{'low'}: \quad & 3.5 && \leq \mathbf{trip}_i\mathbf{.kml} < 7.10 \\
\text{'medium'}: \quad & 7.08 && \leq \mathbf{trip}_i\mathbf{.kml} < 8.41 \\
\text{'high'}: \quad & 8.37 && \leq \mathbf{trip}_i\mathbf{.kml}
\end{aligned}$$

Figure 7 plots $\mathbf{trip}_i\mathbf{.kml}$ for all trips ordered by time and grouped by vehicles with the four classes from Figure 6 marked by horizontal lines. When the trip is 0 km long then $\mathbf{trip}_i\mathbf{.kml} = 0$. Overall, the values are very consistent with what can be expected from minibusses. Some odd values can be seen around 2.5 and $3.33km/l$ where several trips have the exact same $\mathbf{trip}_i\mathbf{.kml}$. This is due to inaccuracies in the measurements of the odometer on very short trips.

Table I lists the fuel efficiency and the standard

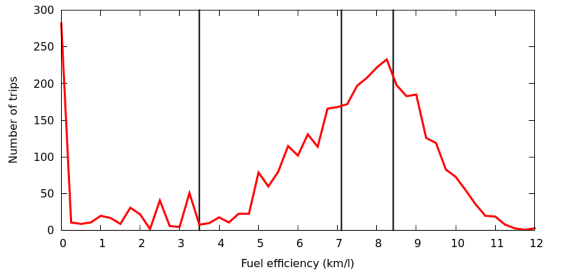


Figure 6. Distribution of all trips grouped in intervals of $0.25km/l$

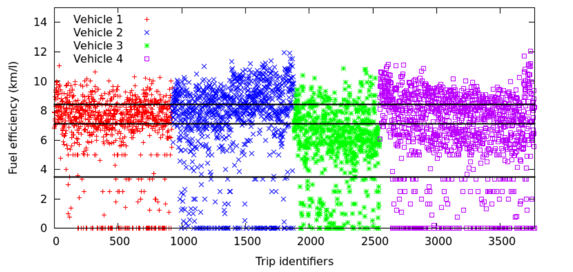


Figure 7. km/l for all trips

sample deviation of the four vehicles and of the total data set. Vehicle 3 is the least efficient and vehicle 2 is the most efficient vehicle with a difference of about 22 %. Vehicle 1 and 4 lie close to the average.

VII. ACCELERATE MODERATELY

Fuel consumption increases when accelerating especially at high velocities [13]. The advice of moderate acceleration says that drivers should accelerate at low gears, with the throttle at 50 % and avoid high RPM above 2000 (2500 for diesel). It is not possible to fully evaluate this as gears is unavailable in the data set, but it is possible to evaluate how hard they accelerate and the RPM. Higher accelerations will result in higher fuel consumption, but the consumption is also dependent on the speed. The starting speed of the accelerations will therefore also be an important factor.

Acceleration is calculated in the standard way from the speed and time stamp.

$$\mathbf{r}_{i+1}\mathbf{.acc} = \frac{\mathbf{r}_{i+1}\mathbf{.speed} - \mathbf{r}_i\mathbf{.speed}}{\mathbf{r}_{i+1}\mathbf{.time} - \mathbf{r}_i\mathbf{.time}} \cdot \frac{1}{3.6}$$

Figure 1 shows the speed of five trips accelerating over time and the associated fuel cost. The road begins right after a roundabout and the measured part shown in the graph is $0.7km$ long. See Figure 38 for details of the area. The starting speed of the trips are almost the same, but the accelerations are quite different, which results in both different end speeds and fuel consumptions. The driver with the highest acceleration uses more than double the amount of fuel than the driver with the lowest acceleration. A figure with all trips driving on this road segment can be seen in Figure 37 in the appendix and a similar example from a different road segment can be seen in Figure 39 and its location in Figure 40. These examples show that a higher acceleration profile leads to a higher fuel consumption.

Looking at the fuel consumption in periods of acceleration results in the same tendency. We define the property of an acceleration period as always having a positive acceleration and that the period is at least 3 seconds long. The minimum time span is added to prevent imprecise measurements. Figure 8 shows the standard deviation on the fuel consumption per second used when accelerating. A small deviation indicates more similar values and hence that most of the outliers have been removed. The curve breaks around a minimum length of 3 seconds, and we choose this as the limit. The worst errors has been removed at 3 seconds, while the risk of removing usable data is minimal. Too high or too low acceleration values might occur because acceleration is calculated from the speed measured by the GPS. This is partially accounted for by interpotating the speed values (see Section IV-C and Algorithm 1), but some errors still exists. To prevent errors in the recorded speed from introducing errors in the acceleration, all acceleration periods containing a speed of $null$ are removed from the analysis of acceleration.

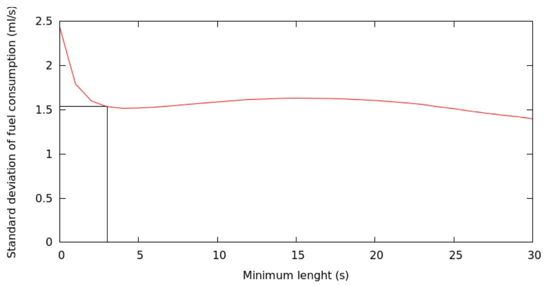


Figure 8. Fuel per second used when accelerating at varying minimum lengths for the periods (s)

Figure 9 shows the fuel consumption at different average accelerations with a plot for different starting speeds. Table VI shows the gradients of the plots and the number of periods with the respective starting speed. Only the regression lines are shown for simplicity, see Figure 42 for data points. The fuel consumption clearly increases as the acceleration increases, which is to be expected due to the extra work required. It could also be expected that the fuel consumption will increase as the speed increases, but this is not supported by the data. The reason might be that the number of data points is very low when the acceleration is above $2m/s^2$.

Figure 10 shows the normalised distribution of acceleration periods with a start speed of $80km/h$ or less grouped by their average acceleration. Periods with speeds higher than $80km/h$ are not comparable to those below $80km/h$ because it becomes too difficult to accelerate highly at these speeds. All vehicles seem to accelerate mostly between 0.373 and $1m/s^2$. Vehicle 3 has slightly more low accelerations than the others and vehicle 1 has slightly more high accelerations.

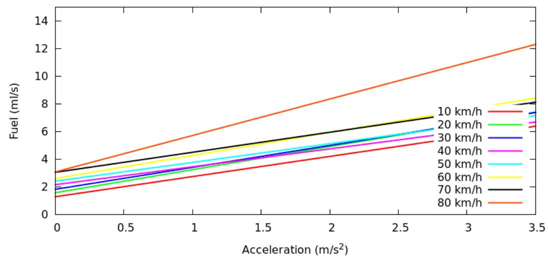


Figure 9. Fuel consumption at different accelerations and starting speeds

Starting speed	Gradient	Number of data points
10	1.5	26,050
20	1.7	19,550
30	1.6	10,699
40	1.3	7,241
50	1.4	7,241
60	1.7	5,694
70	1.4	4,338
80	2.6	1,979
90	2.6	576
100	3.0	410
110	0.3	289
120	1.5	237
130	4.0	86
140	4.3	22

Table VI. GRADIENTS OF GRAPHS IN FIGURE 9

Vehicle 2 has least high accelerations. Generally, there is not any significant difference in their acceleration profiles, making it difficult to evaluate their driving behaviour from it. The speed at which these accelerations are made does, as stated previously, have an effect. Figure 11 shows the average fuel consumption per second of the acceleration periods at various start speeds and accelerations for all vehicles. The x-axis shows the starting speed in ranges of $10km/h$, the y-axis shows the average acceleration of the period in ranges of $0.25m/s^2$ and the z-axis shows the fuel consumption in ml/s of the acceleration period. The individual graphs for the four vehicles can be seen in Figure 43. Values at $80km/h$ and below or at low accelerations are most trustworthy, based on Figure 43e of the number of periods in each range and Figure 43f of the standard deviation on fuel consumption. I can be seen from Figure 11 that there is a strong correlation between speed, acceleration and fuel consumption. The higher the speed, the higher the fuel consumption per second. It also shows that a higher acceleration increase the fuel consumption per second - a steep increase is especially seen at low speeds. This may also be the case at higher speeds, but there is not enough data to support this. Vehicle 3 generally have much higher fuel consumptions (see Figure 43c) than the others and especially vehicle 2 (see Figure 43b) use little fuel on average. This may indicate that vehicle 2 accelerates more moderately than vehicle 3 and that may be a contributing factor for vehicle 2 having the best fuel efficiency.

The second part of the advice is to avoid RPM over 2,500 (2,000 for diesel). Figure 12 shows the average fuel cost per second when accelerating over the RPM. It is clear that a higher RPM leads to a higher fuel consumption. Table VII shows the percentage of times where the vehicles have a RPM greater then 2,000 when accelerating. It is clear that vehicle 3 often breaks this advice and it can explain why the fuel consumption of vehicle 3 is higher than the other vehicles when accelerating. This can be seen by comparing Figure 43a 43b 43c and 43d.

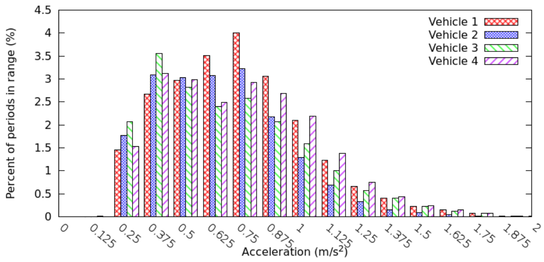


Figure 10. Number of accelerations for all vehicles

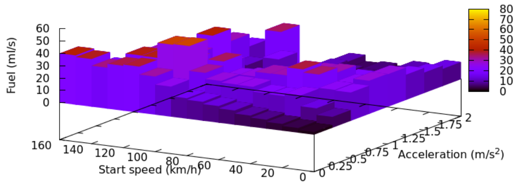


Figure 11. Fuel efficiency at start speed and acceleration

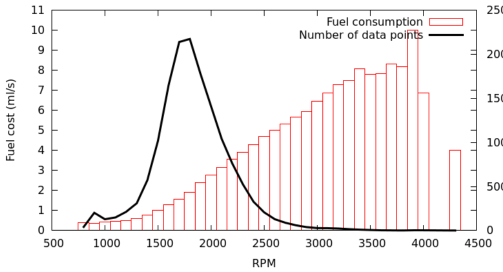


Figure 12. Average fuel cost per second when accelerating with varying RPM

VIII. Maintain a Steady Speed

Driving at a steady speed is more fuel economic than a fluctuating speed, as more fuel is consumed when accelerating [13]. It is therefore interesting to evaluate how good the drivers are at maintaining a steady speed.

It is very difficult to drive at a constant speed for longer periods of time. The speed will always fluctuate with a few km/h for example due to changes in the road conditions or weather conditions. From observing a small data set where cruise control has been used, we see that the speed only varies with $\pm 1 km/h$ (See Figure 13 for an example). Observing data where cruise control has not been used, we see that an experienced driver is able to maintain a steady speed but not over a longer time period.

Vehicle	RPM exceeding 2000
1	10.41%
2	14.03%
3	40.69%
4	28.57%

Table VII.

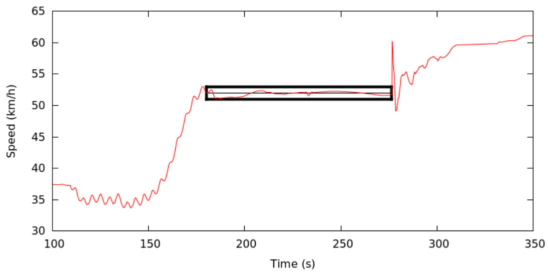


Figure 13. Real example of a period of steady speed

The property of a period of steady speed is that the speed does not vary with more that $\pm 1 km/h$ from the speed at the beginning of the period (the *cruise speed*) for at least 20 seconds. The property is illustrated in Figure 13 where the black box marks the steady speed period, and the thin black line marks the cruise speed.

Figure 14 plots the number of records that maintain a steady speed over the minimum duration of the period with a plot for different speed variations. All four plots break at 20 seconds why this is chosen as the minimum duration. There are more than twice as many records with a steady speed when a speed variation of $1 km/h$ is allowed as opposed to requiring a constant speed with no fluctuations. Increasing the allowed variation to $2, 3$ and $4 km/h$ does not increase the number of records as significantly. A speed variation of $1 km/h$ is therefore chosen.

Figure 15 plots the class distribution of how much of all trips are at a steady speed in percent. It is clear that the trips that often maintain a steady speed primarily belongs to class 'high', and that all trips in class 'low' rarely maintain a steady speed. This indicates that the advice of maintaining a steady speed will reduce the fuel consumption.

Table VIII shows how long the vehicles maintain a

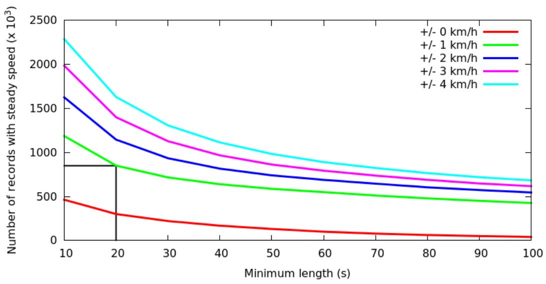


Figure 14. Steady speed at varying minimum duration and speed variations

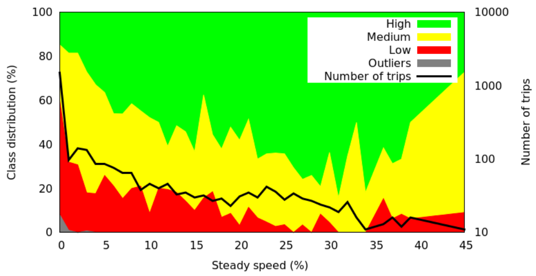


Figure 15. Class distribution of percentage in periods of steady speed

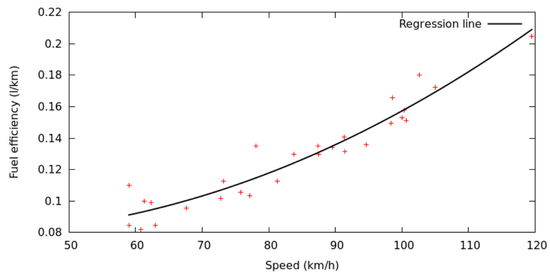


Figure 16. Fuel efficiency at different speeds for 1.1 km long trajectory

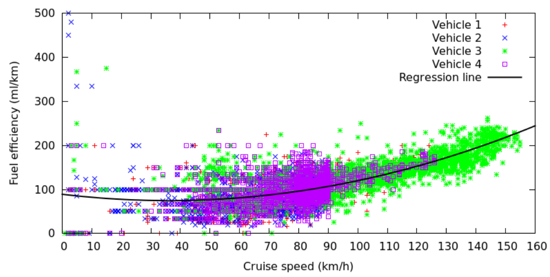


Figure 17. Fuel efficiency at different cruise speeds

steady speed and what that corresponds to in minutes per work day where the driving times from Table IV are used. The numbers should therefore be read as if the vehicles drives 8 hours a day. Vehicle 3 is the vehicle with the worst fuel efficiency ($5.77km/l$) and the one that drives second rarest with at steady speed. This may indicate a connection, but vehicle 2, on the other hand, is the most fuel efficient vehicle but is not the one that drives most often with a steady speed being vehicle 4.

IX. Drive at or below speed limit

The fuel consumption depends on how fast one drives and driving too fast will use more fuel. Figure 16 shows the fuel efficiency of 27 trajectories driving on a 11.1 km road section as a function of their average speed. The number of litres per km increase as the speed increases, supporting the claim that more fuel is consumed at higher speeds. Figure 17 shows ml/km for all periods of steady speed at different cruise speeds. The oddly similar values at for example $100ml/km$ are due to speed only being recorded as integers. Again, the fuel consumption per km increases as the speed increases but instances at low speeds also show a high fuel consumption per km. Accelerating at high speed is also expensive in fuel as seen in Figure 9. The higher the starting speed, then higher the fuel consumption per second. High speeds hence results in higher fuel consumption. Using the regression line, it can be estimated that driving at $130km/h$ costs $172ml/km$ while it only costs $96ml/km$ to drive at $80m/h$.

Figure 18 shows how often the four vehicles breaks the speed limit and by how much. The y-axis plots the number of records with the corresponding speed breach normalised by the total number of records for that vehicle. Table IX lists the percentage of records where the vehicles breaks the speed limit. Vehicle 2 and 4 breaks the speed limit in about 13 % of the records but most of these are with small breaches below $10km/h$. Vehicle 3, on the other hand, breaks the speed limit on more than 25 % of the records and at much higher speeds. The highest breach is at 79 km/h above the speed limit. Vehicle 1 rarely breaks the speed limit, and when he does it, it is with small breaches.

As seen, driving too fast will reduce ones fuel efficiency, and the least fuel efficient vehicle breaks the speed limit much more than the other vehicles. High speeds may contribute to this.

X. Anticipate Traffic Flow

Anticipating traffic flow is a key factor for maintaining a steady speed and avoiding unnecessary ac-

Vehicle id	Steady speed s	%	$\frac{min}{work\ day}$
1	113,751	5.28	25.62
2	304,594	13.85	66.80
3	108,656	9.37	45.27
4	459,742	20.67	99.51
Average	246,685.75	12.29	59.30

Table VIII. Time spend maintaining a steady speed

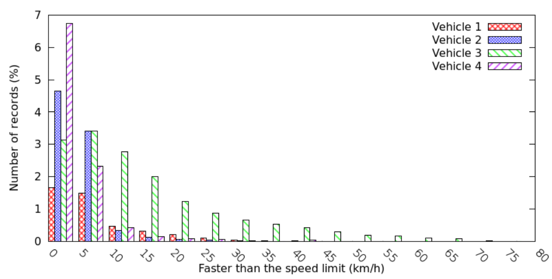


Figure 18. Breaking the speed limit

vid	Speed breach(%)
1	6.96
2	13.13
3	25.71
4	13.11

Table IX. Speed breaches

celerations. Avoiding stopping at traffic lights and adjusting the speed to the expected phases of the light is one part of anticipating the traffic flow [14].

A. Avoid Traffic Lights

Traffic light data is collected from OpenStreetMap, that contains a fair number of traffic lights but not all, especially the smaller traffic lights are missing. Algorithm 2 in appendix details the procedure for counting traffic lights in a trip. Figure 19 plots the percentage of trips in the four classes driving through different concentrations of traffic lights. From this graph it is clear that the trips in class 'high' with a high km/l are mostly driving in areas with few traffic lights, where as trips driving in areas with many traffic lights mostly are in class 'medium' or 'low'.

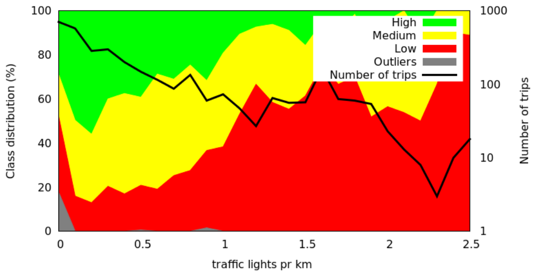


Figure 19. Traffic lights per kilometer

Figure 20 shows all 16 trips driving on a $0.7km$ straight road containing a single traffic light and few small side roads see Figure 33. All trips are driving in the time span from 9:00 to 15:00. From 20 to 46 seconds some of the trips stop, presumably because of a red light. The graph shows the time each trip takes to drive the $0.7km$ road, their speed and the fuel cost. It is clear that the trips that stop are using more fuel than the other trips. Defining a traffic light related stop as in Algorithm 2 in Appendix we get the average fuel consumption for the $0.7km$ road section from these 16 trips as in Table X. A similar example can be seen in Figure 34.

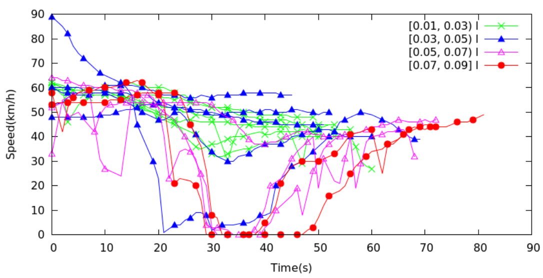


Figure 20. Cost of stopping at a traffic light

Figure	With stop		Without stop	
20	68 ml	$73s$	34 ml	$54s$
34	60 ml	$48s$	31 ml	$21s$

Table X. FUEL CONSUMPTION AND TRAVEL TIME AT A TRAFFIC LIGHT

B. Traffic Density

Is is difficult to evaluate the vehicles relation to other vehicles because no data about them is available. There is, however, often a connection between which types of roads are used and traffic flow and density. Roads in the city tend to have more obstacles, e.g. vehicles and intersections than motorways. It will therefore be interesting to investigate whether there is a connection between road types and fuel efficiency. As explained in Section IV-C, each records is matched to a road segment from OSM. Each segment has a category depending on the type of road[1].

- 'Motorways' are unidirectional roads with no traffic lights and no crossing traffic. They are categorised as motorways in OSM.
- 'Main roads' contains all main roads both in and outside of the cities categorised in OSM as trunk, primary, secondary and tertiary highways. They usually have a speed limit from 50-80km/h.
- 'Small roads' are mostly found in residential areas or connecting smaller cities and are categorised in OSM as residential, road, service, living streets or unclassified. The speed limit can vary from 10 to 80 km/h.

Figure 21 shows the class distribution of trips at different percentage driving the small roads. The trips from class 'high' drives less on small roads than trips in class 'medium' and 'low' and vice verse. In contrast, Figure 22 shows the class distribution of driving on main roads where it is clear that the trips from class 'high' drives more often than those in class 'medium' and 'low'. Too little data is available for motorways, but the class distribution can be seen in Figure 44. These results suggest that vehicles with a good fuel efficiency tend to drive less on small roads and more on main roads and vice versa. Several reasons could be the cause, such as more intersections that requires a slower speed and higher traffic density that makes it difficult to drive at optimal speed.

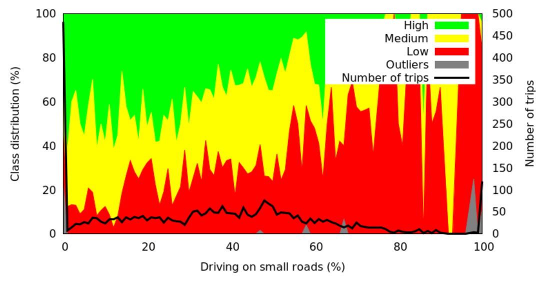


Figure 21. Class distribution of driving on small roads

Table XI shows the average percentage of records recorded on the tree types of roads over all trips. The values do not add up to 100 % as they are averages over the vehicles trips. Vehicle 3 drives much more on motorways than shown, but this has not been recorded as some of the motorways are not on the available map. All vehicles spend about a third of their time on small roads and vehicle 1, 3 and 4 spend about 35 % of

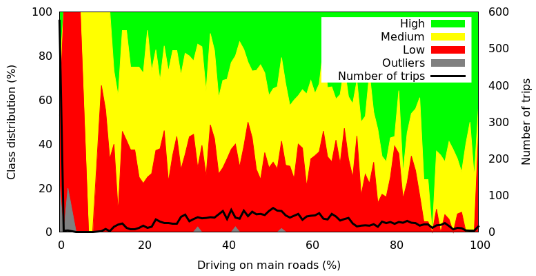


Figure 22. Class distribution of driving on main roads

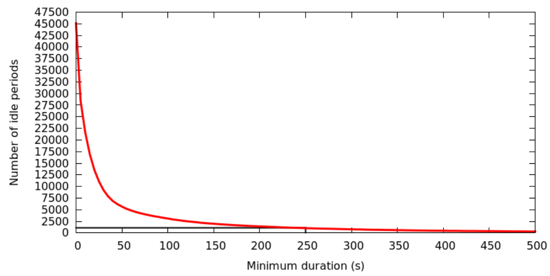


Figure 23. Minimum idle duration

their time on main roads. Only vehicle 2 spends about half of the time on main roads. As Figure 21 and 22 showed, main roads tend to be more fuel efficient than small roads and the type of road may contribute to the fact that vehicle 2 is more fuel efficient than the others.

Vehicleid	Small (%)	Main (%)	Motorway (%)
1	34.95	35.59	3.00
2	32.97	48.67	0.75
3	31.58	35.38	6.38
4	28.68	35.42	12.67

Table XI. AVERAGE PERCENTAGE OF TRIPS ON THE THREE ROAD TYPES

XI. AVOID IDLING

Avoiding idling or minimising idle time is a factor in eco-driving as fuel is still consumed when the engine is running even though the vehicle is not moving. The driver is hence consuming unnecessary fuel when idling.

Say a vehicle is *stopped* iff. the RPM is above zero and the speed is zero for at least 2 consecutive recordings.

$$\begin{aligned}\mathbf{r}_{i+1}.\mathbf{stop} = true \text{ iff.}\\ &\mathbf{r}_i.\mathbf{RPM} > 0 \wedge \mathbf{r}_{i+1}.\mathbf{RPM} > 0 \wedge\\ &\mathbf{r}_i.\mathbf{speed} = 0 \wedge \mathbf{r}_{i+1}.\mathbf{speed} = 0 \wedge\\ &\mathbf{r}_{i+1}.\mathbf{kmcounter} - \mathbf{r}_i.\mathbf{kmcounter} = 0 \qquad (1)\end{aligned}$$

A vehicle is hence stopped when for example waiting at a red light, in a queue or parked with the engine turned on. A *stopped period* is a sequence of records where $\mathbf{r}.\mathbf{stop}$ is true as per Section V-B. But when looking at idling, we are not interested in all stopped periods. All stops near traffic lights are not idling and short stops are neither.

Figure 23 shows the total number of stopped periods with different minimum durations of the periods. The curve flattens around $250s$, which is a little over 4 minutes. An *idle period* is therefore a stopped period that is longer than $250s$ and not near a traffic light. Involuntary stopped periods such as queues, duty to give way and alike are thereby reduced.

Figure 24 and Figure 25 shows the total sum of how long the vehicles are idling. The periods are combined into ranges of 50 and 100 seconds and displayed on the x-axis, e.g. 250 indicates the time range 250-299 seconds in Figure 24. The height of the columns are the total sum of time spend idling in the given range, because this makes the few but long idle periods comparable with the many but short periods. If a column has the same height as its position on the x-axis, for example at $4400s$, then only one occurrence of the idle period exist. Vehicle 3 and 4 idles less in the short periods than the other vehicles, but vehicle 3, on the other hand, idles in much longer periods than the others.

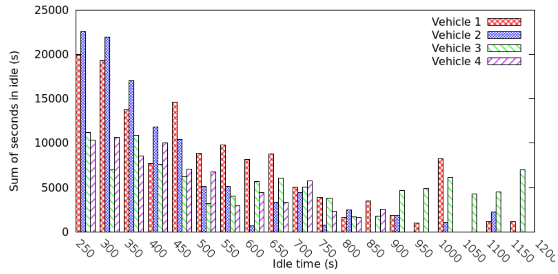


Figure 24. Number of idle periods at different ranges below $1200s$

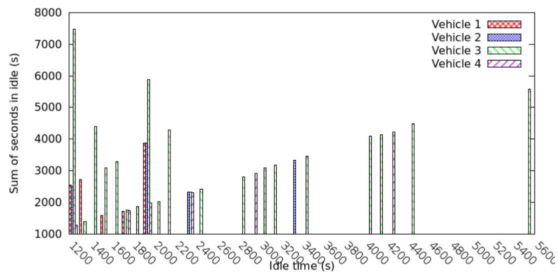


Figure 25. Number of idle periods at different ranges above $1200s$

Figure 26 shows how much fuel is consumed in the idle periods. Fuel consumption is shown on the y-axis and the number of seconds of the idle period is shown on the x-axis. A linear regression line has been plotted for each vehicle. We see that few idles for more than 1,300 seconds ($\sim$ 21 minutes) but that these use between 0.5 and 2 litres of fuel each time. Following the regression lines it can be estimated that between 0.85 and 1.25 litres of fuel are used per idling hour. Table XII shows how much each vehicle idles and how much fuel is used using the driving times from Table IV assuming a work day of 8 hours. Vehicle 3 idles the most (15 % of the time) and consumes about one and a half litre of fuel a day on idling. This corresponds well with the fact that vehicle 3 is the least fuel efficient vehicle. Vehicle 4 idles the least and consumes least fuel on idling, but vehicle 2 and 3 are close. Vehicle 2 is the most fuel efficient vehicle of the four, but these results suggest that other factors than idling contributes to this. Based on the regression lines and the values from Table XII, it can be estimated that idling costs about $1l/h$.

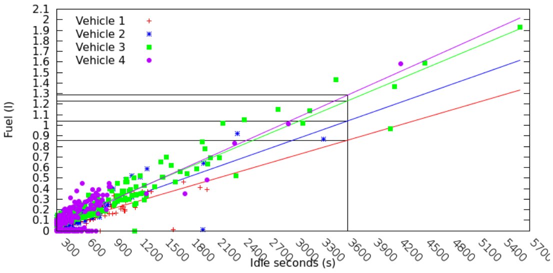


Figure 26. Fuel consumption when idling

Vehicle id	Idle			Fuel	
	s	%	$\frac{min}{work\ day}$	l	$\frac{l}{work\ day}$
1	150,763	7.00	33.96	35.12	0.47
2	122,890	5.59	26.95	35.12	0.46
3	176,522	15.22	73.55	59.21	1.48
4	90,919	4.09	19.68	26.27	0.34
Average	135,273.50	7.98	38.54	38.93	0.69

Table XII. TIME AND FUEL CONSUMPTION IN IDLE

It will also be interesting to investigate whether how much of a trip is in an idle state has an influence on the trips km/l. Figure 27 shows the class distribution of all trips grouped by how much of the trip is in an idle period. The black line indicates the number of trips on a logarithmic scale. We clearly see that of the trips with a small idle percentage are mostly in the class 'high' and vice versa. The classes 'low' and 'outliers' consists of trips that idles much of the time. This strongly suggests that there is a correlation between idle and fuel efficiency. Figure 28 shows the class distribution over the length of the idle periods split into ranges. We do not see the same clear distinction in this plot as for Figure 27. This might indicate that the trips from class 'outliers' and 'low' do not necessarily idle for long periods but idles more often.

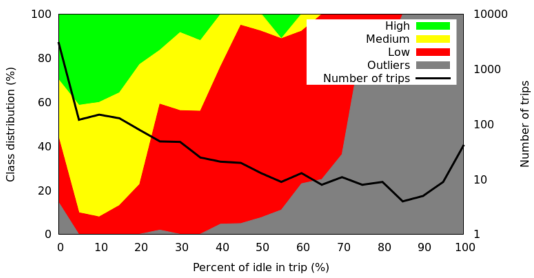


Figure 27. Class distribution of percentage in idle periods

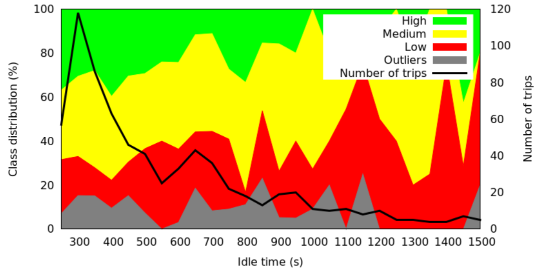


Figure 28. Class distribution of time in idle periods

XII. DISCUSSION

In this section we will discuss how well the eco-driving advice is followed based on the the findings in the previous sections and estimate the expected savings if the advice is followed more. The section ends with a short discussion on whether the assumption that the vehicles are comparable is true.

A. Evaluating Driving Behaviour

Figure 29 is a radar chart over seven aspects of the evaluated eco-driving advice. It is designed to compare the drivers as fair as possible and such that a vehicle that follows the advice will have a small area. The axes are as follows: 'Not at steady speed' is how much of the time is spend not at steady speed (Advice 2), 'RPM' is how many of their acceleration periods have an average RPM above 2000 (Advice 5), 'Acceleration' is the average fuel efficiency over all acceleration periods (Advice 5), 'Idle' is how much of the time is spend idling (Advice 6), 'Traffic lights per km' is the average number of traffic lights crossed per km (Advice 3), 'Not on main roads' is how many records have not been recorded on main roads (Advice 3) and 'Speed limit' is how many records have been recorded with a higher speed than the speed limit (Advice 7). Vehicle 3 idles more than the others, breaks the speed limit more often, accelerates at higher RPM and use more fuel on accelerating. A driver has, to some extent, an influence on all of these factors. Vehicle 3 does, however, cross less traffic lights than the others and drives much on main roads as vehicle 1 and 4. Both of these two factors are correlated with driving in city areas, which the driver may not have influence over. Vehicle 2 is the most fuel efficient vehicle, but is not distinctively better than the other vehicles on any of the factors in Figure 29. Vehicle 2 is always better than either vehicle 1 or 4 except on speed limit and only once worse than vehicle 3. Vehicle 1 is best at driving below the speed limit and accelerating at low RPM, but drives much outside main roads, through many traffic lights and is the vehicle that rarest drive at a steady speed. Vehicle 4 is the vehicle that idles the least and most often drives at a steady speed, but he always takes the third place on all other factors. Table XIII ranks each vehicle from 1 to 4 based on their position in Figure 29. Vehicle 4, for example, receives a 1 on 'Not at steady speed' because he drives the most at a steady speed, while vehicle 1 gets a 4 because he drives the least at a steady speed. By summing all the ranks, it can be seen that vehicle 2 has the best overall rank which corresponds with the fact that vehicle 2 is the most fuel efficient vehicle. Vehicle 3 gets the highest rank, which also corresponds with him being least fuel efficient. Vehicle 1 and 4 gets a similar score and also has about the same fuel efficiency. It can therefore be be said that these seven factors, in this situation, gives a collective representation of the driving behaviour that corresponds with the fuel efficiency.

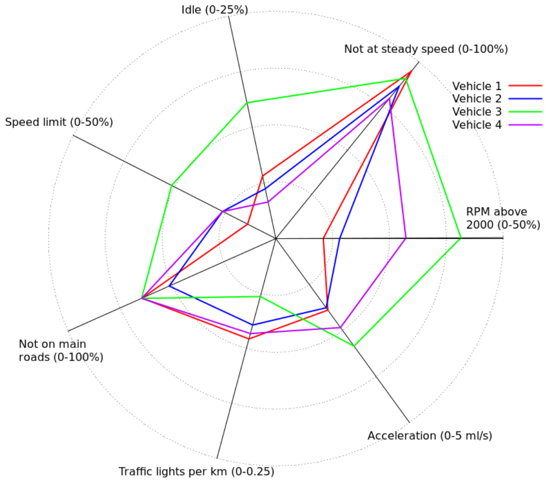


Figure 29. Summary of eco-driving behaviour

Advice	Vehicle 1	Vehicle 2	Vehicle 3	Vehicle 4
Not at steady speed	4	2	3	1
RPM	1	2	4	3
Acceleration	2	1	4	3
Traffic lights per km	4	2	1	3
Not on main roads	2	1	2	2
Speed limit	1	2	3	2
Idle	3	2	4	1
Sum	**17**	**12**	**21**	**15**

Table XIII. Ranking vehicles based on Figure 29

B. Expected Savings

Vehicle 3 can improve his fuel efficiency on several factors. Idling costs about $1l/h$ and vehicle 3 idles about 15 % of his time consuming a about $1.48l$ a day assuming a 8 hour work day. By removing all idle time, vehicle 3 could save around $550, 56kr$ a month with a gas price of $12kr/l$. Vehicle 3 could also consider driving slower in order to improve his fuel efficiency. Driving at $80km/h$ in stead of $130km/h$ is estimated to save about $76ml/km$ on average. Vehicle 3 can then save around $1.52l$ a day, assuming a daily trip of $20km$. This adds up to $565, 44kr$ a month. It remains to be investigated whether the reduced travel time outweighs the higher fuel consumption. Generally it can be said that drivers should avoid cities if possible as seen in Figure 19 and Figure 21. They should also try to maintain a steady speed as mush as possible (See Figure 15) and avoid high accelerations.

An interesting remark about Figure 7 is that vehicle 2 tends to become more fuel efficient over time and that vehicle 4 becomes slightly less fuel efficient over time. This could be due to a change in weather that changed the driving conditions or need for air conditioning, the driver becoming more/less aware of his driving behaviour or service being performed on the vehicle.

C. Are the Vehicles Comparable?

It has been assumed until now that the four vehicles are comparable. This assumption is challenged in this section. Statistics over the RPM of the four vehicles when idling can be see in Table XIV. RPM when idling is not affected by driving behaviour and should be similar if the engines are comparable. The RPM is around 852 and we believe that, even though the deviations are different, the vehicles are comparable on this parameter. The fuel consumption per second when idling can be seen in Table XV. These values vary more which may be because they use other fuel consuming accessories e.g. air conditioning. The avgerage fuel efficiency during periods of steady speed at different cruise speeds for the four vehicles can be seen in Figure 30. There is not enough data for all vehicles below $50km/h$ and above $90km/h$ (See Figure 17). The fuel efficiencies are, however, much similar between $50km/h$ and $90km/h$ and also up to $120km/h$ looking aside from vehicle 2. These findings support the fact that the vehicles are comparable.

Vehicle id	RPM			
	Min	Avg.	Max	Std. deviation
1	79	852.67	3067	54.81
2	54	851.95	3413	37.10
3	42	851.86	4149	29.95
4	51	852.34	3204	36.57

Table XIV. RPM when idling

Vehicle id	Fuel consumption (ml/s)			
	Min	Avg.	Max	Std. deviation
1	0	0.23	0.41	0.07
2	0	0.28	0.56	0.10
3	0	0.33	0.59	0.09
4	0	0.26	0.83	0.19

Table XV. Fuel consumption when idling

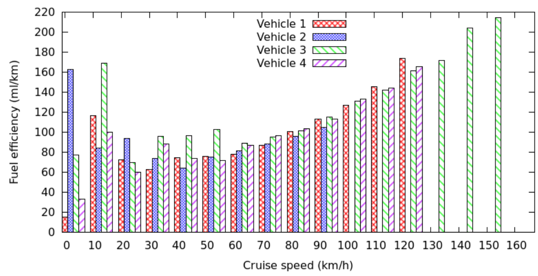


Figure 30. Fuel efficiency at different cruise speeds

XIII. Experience with Data

We have found that there are certain requirements to the GPS and CANBus data before usable and precise analysis of the driving behaviour can be made.

Most of the data set is recorded with a frequency of 1-2 seconds. This is sufficient to calculate and analyse accelerations. A minimum frequency of 2 seconds are hence used and remaining data is removed from the analysis as is provide the best compromise between high frequency and discarding too much data. Lower frequencies greatly reduce the accuracy and precision as many acceleration periods are only a few seconds long (see Figure 41). A higher frequency could improve the results in acceleration. It would allow more precise periods of acceleration and the fuel cost. This in turn will improve the precision in Figure 9, 42, 11 and 43. With a higher frequency it might be possible to detect when the vehicle changes gear by looking at the changes in acceleration and RPM. This is difficult with a frequency of 1-2 seconds and therefore not investigated further in this study.

Speed data is used in many connections and has a great influence on the results. Both acceleration and steady speed use speed in their calculations. The speed values are only recorded as integers, which results in imprecise calculations of for example acceleration. Figure 2 shows how the use of integers results in holes in the acceleration ranges. The speed values are sometimes also incorrect when the GPS device looses signal (See Table XVI). The speed can be set to zero or interpolated. The fact that speed erroneously can be set to zero has resulted in the last requirement of the property of stopped periods (See Definition 1). It can be difficult to detect some of these errors, as they could in some cases resemble an emergency braking followed by a fast acceleration, and the errors hence create a false acceleration period. These errors are mitigated by the minimum and maximum accelerations put in place. A better way to handle these errors could be to compare speed values from the GPS with the read-out from the speedometer which is a CANBus value. If the two values deviate by some threshold, then the data can be disregarded. The GPS speed will on many occasions be the most correct value, but errors will be easier to detect if data from the speedometer is available as well. This was not the case in the available data set, but it would be recommended to add them if these errors are to be avoided.

The fuel consumption will have to have a certain accuracy and precision. The original data set contains six vehicles of which two only has a precision of $500ml$ in fuel consumption. The remaining four has a precision of $10ml$ which is adequate for the analysis. A precision of $500ml$ will result in very inaccurate calculations on small periods as fuel consumption often is a few litres per kilometre, especially when idling or driving slowly. The fuel consumption of small periods will often randomly either be $0ml$ or $500ml$ which is not usable in most of the analyses made. We therefore choose to discard the two vehicles with too imprecise fuel measurements. Data with a low precision can be used on aggregated levels such as trips. However, these trips will still need to be long enough for the fuel consumption to be sufficiently higher than the precision of the fuel measurement. A trip with a length of $50km$ can have its fuel consumption vary by up to 6.7% given a $500ml$ precision and an average fuel consumption of $6.73km/l$. A higher precision than $10m$ could improve the results, especially when working with acceleration as this is often measured in short periods. It would improve the results in Figure 9, 42, 11 and 43.

The odometer only measures kilometres with one decimal point, i.e. $100m$ precision. This will occasionally result in inaccurate calculations when driving short distances. A trip with a distance of $100m$ can have driven anywhere between $0m$ to $100m$. The precision is hence dependent on the length of the trip. This problem was reduced by saying that a trip should have a minimum number of 30 records. A more precise distance could improve the precision of the Figure 6 where some of the values appears to be on a line because of the odometer. If higher precision on the odometer is not available, it might be possible to improve it by using the GPS latitude and longitude to calculate distance in stead.

The gears are not available and prevented analysis of some of the eco-driving advice. Specifically advice 1, 4, 5 and 9 directly refers to the gears. It is, however, possible to partially evaluate some of the advice. For more detailed analysis, gears and other measurements will be required. These might be the position of the throttle, road gradients, make and model of the vehicle, and others.

The available data set contains an acceleration value which is not used in the analysis. The nature of the values indicate that it might be measurements from an accelerometer, but the units are unknown. An acceleration of 50 were also often recorded even though the vehicle was maintaining the same speed. Due to this undocumented nature of this value and the large unexplained fluctuations, we chose to disregard it in this study.

XIV. Conclusion

This paper analysed the foundations of a number of eco-driving advice based on a large set of CANBus data from four vehicles over a period of nine months. The purpose was to identify why one vehicle drives $0.66km/l$ better than the average and why another vehicle drives $0.96km/l$ worse than the average. It is found that high accelerations especially at high speeds and driving at unsteady or at high speeds leads to a higher fuel consumption. It is also found that the fuel consumption is higher when driving in areas with many traffic lights or on small roads and that idling consumes a considerable amout of fuel. Not one single of these factors has been found to lead to a good fuel economy on its own but several factors have been found where the least fuel efficient vehicle can improve. This vehicle idles much more than the others, and it drives at much higher speeds. To the contrary, it drives through fewer traffic lights. This leads to the conclusion that no single factor is all-important, however, it is the many small improvements that leads to a good fuel economy.

The next steps in this process will be to expand the analysis to a larger data set with more vehicles recorded over a longer time period. This will allow more detailed analyses. More data values will also be necessary in order to evaluate all of the eco-driving advice. Gears, for example, are essential for several parts of the advice, but other information such as the gradient of the roads might also provide usable information. Much of the CANBus data is available, but is not yet documented by car manufacturers. Hence, it must be reverse-engineered for every type of vehicle which is time consuming and expensive. It will also be interesting to track the changes in fuel economy for drivers over time. This will require data from a longer time period or perhaps data before and after the drivers have been educated about their driving behaviour.

XV. Acknowledgement

We would like to give our thanks to ProTracking[4] for providing the data for this project. Ove Andersen shall also receive out thanks for map-matching the data set.

XVI. Appendix

Figure 31 shows where the vehicles operate in Denmark. Each dot represents a data record. Most data is hence in central Jutland with a few trips to Germany to the south and Copenhagen to the east.

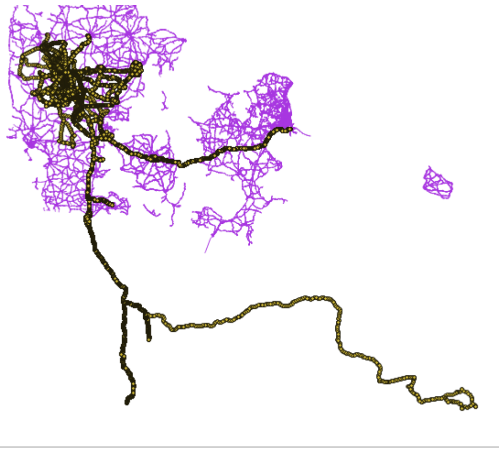


Figure 31. Data location

Figure 32 shows the data frequency of the records where 99% of the data is recorded with 1-2 seconds frequency.

Figure 33 shows a map of a 0.8 km long road segment indicated by the arrow. It is located in Tjørring 1 km north of Herning, Denmark and is straight. There is only one traffic light on the road segment and two small dead end side roads. The remaining side roads are closed off. The speed limit on the road is 50 km/h.

Figure 34 shows 14 trips and their fuel consumption of driving on a $0.4km$ straight road with a single traffic light (see Figure 35).

Figure 35 shows a map of a 0.4 km long road segment indicated by the arrow. It is located 1 km west of Herning, Denmark and is straight. There is only one traffic light on the road segment and a few small side roads are connected. The speed limit on the road is 60 km/h before the traffic light and 80 km/h after.

Algorithm 2 details the procedure for counting traffic lights in a trip where $TLInRange$ finds the closest traffic light within a range of $25m$ and returns

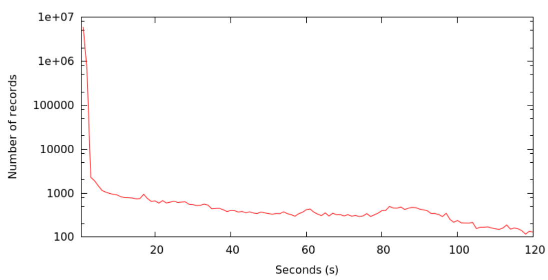


Figure 32. Data frequency

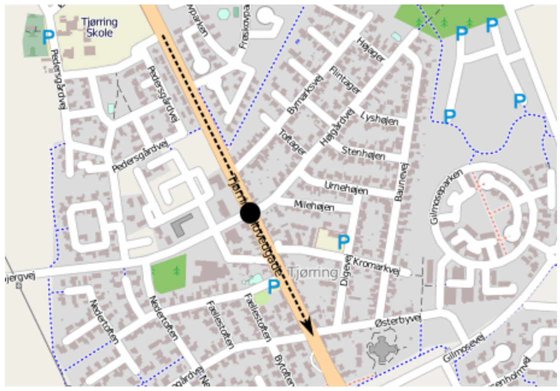


Figure 33. Road segment where trips are driving through a traffic light represented by the black dot. 1 km north of Herning

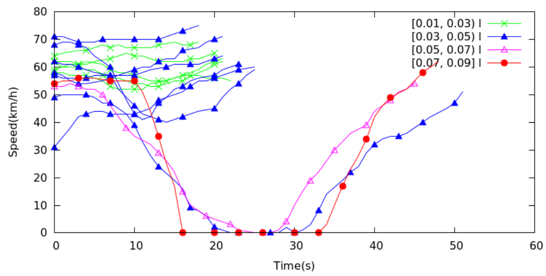


Figure 34. Cost of stopping at a traffic light

null if none is found. Line 1 orders the records of the trip on its timestamp. Line 2-4 set up the initial variables. Line 5 loops though the individual records in chronological order on the timestamp. Line 6-8 check if the record is within 25 meters of a traffic light and if so, adds one to the counter and sets the variable inL indicating that the vehicle is inside the area of a traffic light. Line 9-11 check if the vehicle has left the traffic light and then resets the variable inL. Line 12-15 check if the record has a speed of zero while the vehicle is inside a trafficlight. If so, we count it as a full stop and increment the counter $redCounter$. To avoid favouring shot trips, the number of traffic lights visited is divided with the length of the trip. This is done on line 17. The 25 meter radius is determined from the graph in Figure 36. It is clear that the average number

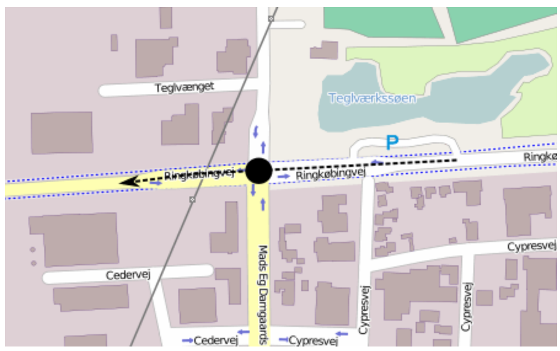


Figure 35. Road segment where trips are driving through a traffic light represented by the black dot. 1 km north of Herning city

of traffic lights a vehicle crosses stabilises after 25 meters. Hence, vehicles crossing a traffic light without being counted is very unlikely.

Algorithm 2 $countTrafficLights(\textbf{trip})$

$R = \textbf{trip}.timeOrderRecords$
$TL =$ all traffic lights
$inL, red = False$
$counter, redCounter = 0$
while $r = R.popFirst()$ **do**
  if not inL **and** $TLInRange(TL, r, 25)$ **then**
    $inL = True$
    $counter+ = 1$
  else if not $TLInRange(TL, r, 25)$ **then**
    $inL = False$
  end if
  if not red **and** inL **and** $r_{speed} = 0$ **then**
    $red = True$
    $redCounter+ = 1$
  end if
end while
return $counter/\textbf{trip}.\textbf{kmcounter}$

Algorithm 3 CalculateAcc(r1,r2)

$acc = ((r2_{speed} - r1_{speed})/(r2_{time} - r1_{time}))/3.6$
return acc

Algorithm 4 InterpolateSpeed(R,r,r1,r2)

$h = (R[r2]_{speed} - R[r1]_{speed})/(R[r2]_{time} - R[r1]_{time})$
$g = R[r1]_{speed} + (h * (R[r]_{time} - R[r1]_{time}))$
return g

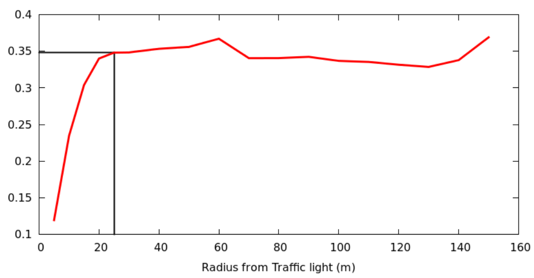


Figure 36. Size of traffic lights

Figure 37 plots the speed of 39 trips while they accelerate and the associated fuel consumption as a function over time. The road segment can be observed in Figure 38.

Figure 38 shows a map of a 0.7 km long road segment indicated by the arrow. It is located 12 km south of Herning, Denmark and has a slight curve. Vehicles first exit the roundabout and then accelerate along the road. The speed limit on the road is 80 km/h.

Figure 39 plots the speed of 85 trips while they accelerate and the associated fuel consumption as a function over time. The road segment can be observed in Figure 40.

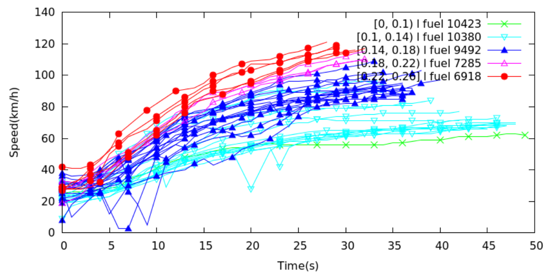


Figure 37. Trips driving the same road segments with diffrent accelerrations

Figure 38. Road segment where trips are accelerating. 12 km south of Herning

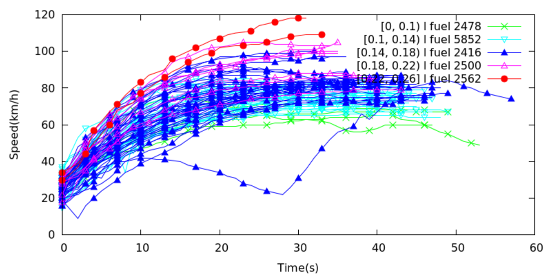
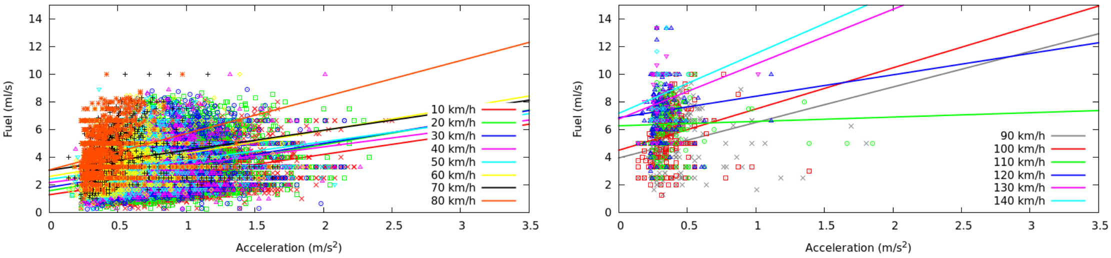


Figure 39. Trips driving the same road segments with different accelerations

Figure 40 shows a map of a 0.8 km long road segment indicated by the arrow. It is located 2 km east of Herning, Denmark and is straight. Vehicles first exit the roundabout and then accelerate along the road. The speed limit on the road is 80 km/h.

Figure 41 plots the number of acceleration periods with different lengths. Note that in the analysis only use periods of three seconds or longer.

Figure 42 shows the fuel consumption at different average accelerations with a plot for different starting speeds including the regression lines. Values above



(a) Speed less or equal 80 km/h (b) Speed above 80 km/h

Figure 42. Fuel consumption at different accelerations and starting speeds

Figure 40. Road segment where trips are accelerating. 2 km east of Herning

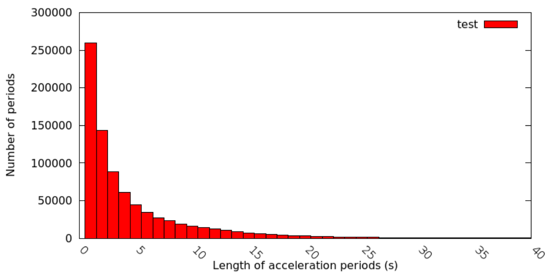


Figure 41. Number of acceleration periods with different length

$90 km/h$ are based on very few data points which might be the reason for their regression lines being different from the rest.

Figure 43a, 43b, 43c and 43d plots the fuel consumption per second over acceleration and the start speed for vehicle 1, 2, 3 and 4 respectively. The number of periods in every range can be seen in Figure 43e where the x- and y-axis has been reversed. Most of the periods are at low speeds and low accelerations. The standard deviation over all periods are plotted in Figure 43f. The largest deviations are at high speeds.

Figure 44 shows the distribution of trips with from 0% to 10% trips driving on the motorways. There is insufficient data on trips driving on motorways.

Table XVI lists an example of an error in the recorded speed data and how this is adjusted. It can be seen how speed is adjusted to the modified speed

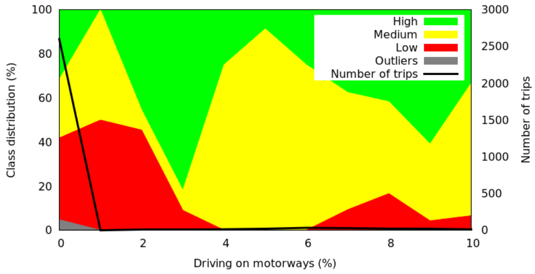


Figure 44. Class distribution of driving on motorways

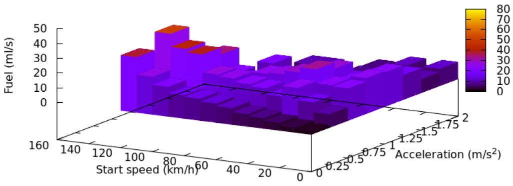
to repair some of the errors caused by the recording equipment.

Timestamp	Speed	Modified speed	RPM
2013-02-01 12:13:14	89	89	2253
2013-02-01 12:13:16	89	89	2248
2013-02-01 12:13:17	89	89	2251
2013-02-01 12:13:18	6	$null$	2247
2013-02-01 12:13:19	30	$null$	2248
2013-02-01 12:13:20	47	30	2251
2013-02-01 12:13:21	30	30	2247
2013-02-01 12:13:22	30	30	2246
2013-02-01 12:13:23	74	$null$	2245
2013-02-01 12:13:24	83	83	2250
2013-02-01 12:13:25	87	87	2251
2013-02-01 12:13:26	88	88	2251
2013-02-01 12:13:28	89	89	2249
2013-02-01 12:13:29	89	89	2252
2013-02-01 12:13:30	89	89	2244

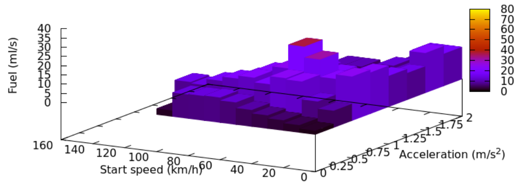
Table XVI. EXAMPLE OF ADJUSTING SPEED VALUES

Figure 45 shows the distribution of fuel consumption per second. For obvious reasons, records with a negative fuel consumption is considered errors. As most of the records use less then $20ml/s$ any record above this is considered erroneous and is removed.

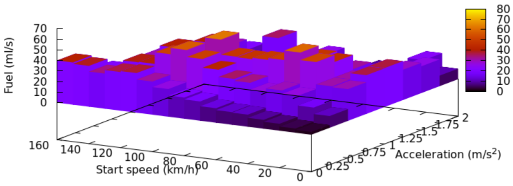
Table XVII lists an example of errors in the recorded fuel measurement.



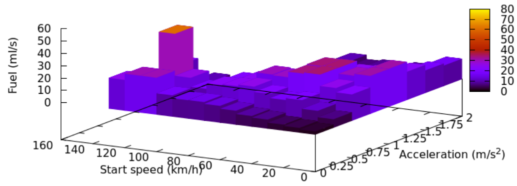
(a) Vehicle 1



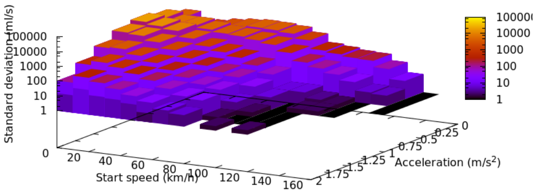
(b) Vehicle 2



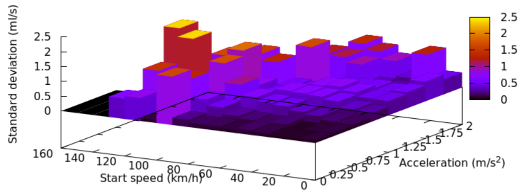
(c) Vehicle 3



(d) Vehicle 4



(e) Number of periods in range. Notice that the x- and y-axes have been reversed



(f) Standard deviation on fuel consumption

Figure 43. Fuel efficiency at start speed and acceleration

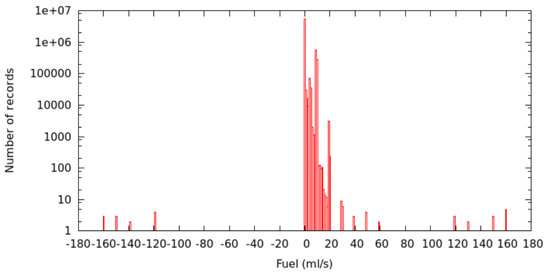


Figure 45. Number of records over fuel used per second when accelerating

Timestamp	Speed	Fuel
2013-02-21 16:28:56	46	2845.41
2013-02-21 16:28:57	45	2845.41
2013-02-21 16:28:58	39	2845.56
2013-02-21 16:28:59	45	2845.41
2013-02-21 16:29:00	45	2845.41
2013-02-21 16:29:01	46	2845.41
2013-02-21 16:29:02	48	2845.57
2013-02-21 16:29:03	50	2845.57
2013-02-21 16:29:04	52	2845.57
2013-02-21 16:29:05	52	2845.57
2013-02-21 16:29:06	53	2845.57
2013-02-21 16:29:07	54	2845.58
2013-02-21 16:29:08	54	2845.43
2013-02-21 16:29:09	55	2845.43
2013-02-21 16:29:10	56	2845.43
2013-02-21 16:29:11	51	2845.58
2013-02-21 16:29:12	56	2845.44
2013-02-21 16:29:13	55	2845.44

Table XVII. EXAMPLE OF ERRONEOUS FUEL MEASUREMENTS

REFERENCES

[1] http://wiki.openstreetmap.org/wiki/Key:highway.

[2] Ecodriving - the smart driving style. Produced by SenterNovem, Utrecht for the EC TREATISE project, September 2005.

[3] Openstreetmap. http://www.openstreetmap.org/.

[4] Protracking.dk. http://www.protracking.dk/.

[5] Road transport: Reducing co2 emissions from vehicles. http://ec.europa.eu/clima/policies/transport/vehicles/index_en.htm. Visitied 9. april 2013.

[6] M. Barth and K. Boriboonsomsin. Energy and emissions impacts of a freeway-based dynamic eco-driving system. *Transportation Research Part D: Transport and Environment*, 14(6):400–410, Aug. 2009.

[7] I. M. Berry. The effects of driving style and vehicle performance on the real-world fuel consumption of u.s. light-duty vehicles. Master's thesis, Mechanical Engineering, Virginia Polytechnic Institute and State University, 2007.

[8] K. Boriboonsomsin and M. Barth. Eco-driving: Pilot evaluation of driving behavior changes among u.s. drivers. 2010. University of California, Riverside.

[9] K. Dzenisiuk. Eco-driving - changing truck driver behavior to achieve long-term sustainability results. Master's thesis, 2012.

[10] D. B. F. et al. *NCHRP Report 400: Determination of Stopping Sight Distances*. Transportation Research Board, 1997.

[11] B. B. et.al. Using on-board logging devices to study the longer-term impact of an eco-driving course. *Transportation Research Part D: Transport and Environment*, 14, 2009.

[12] R. K. Ganti, N. Pham, H. Ahmadi, S. Nangia, and T. F. Abdelzaher. Greengps: a participatory sensing fuel-efficient maps application. In *Proceedings of the 8th international conference on Mobile systems, applications, and services*, MobiSys '10, pages 151–164, New York, NY, USA, 2010. ACM.

[13] C. Guo, Y. Ma, B. Yang, C. S. Jensen, and M. Kaul. Ecomark: evaluating models of vehicular environmental impact. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '12, pages 269–278, New York, NY, USA, 2012. ACM.

[14] K. Jakobsen and S. Mouritsen. Greenflow: Improving traffic by adjusting velocity to traffic lights. 2012.

[15] F. C. Pereira, H. Costa, and N. M. Pereira. An off-line map-matching algorithm for incomplete map databases. 2009.

[16] H. Rakha and K. Ahn. Integration modeling framework for estimating mobile source emissions.

[17] M. C. Snare. Dynamics model for predicting maximum and typical acceleration rates for passenger vehicles. 2002.

[18] J. Stæhr. Økonomikørsel - kør grønt - for klimaets skyld. February 2012. http://www.oekonomikoersel.dk/. Visitied 9. april 2013.

[19] Vejdirektoratet. Bedre trafiksignaler - konsekvenser for trafikanter og samfund. 2012.