Longitudinal dynamics analysis of a city bus in traffic

Master of Science Thesis

DAVID ELEBRING

Department of Signals and Systems
Division of Automatic Control, Automation and Mechatronics
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden, 2014
Report No. EX003/2014

ABSTRACT 5

1 INTRODUCTION 7

1.1 Background 7

1.2 Method 7
1.2.1 Variables 8

2 MEASUREMENTS 9

2.1 Equipment 9
2.1.1 Planning and implementation 9
2.1.2 Components 9
2.1.3 Software 10

2.2 Data Collection 10

3 DATA ANALYSIS 11

3.1 Pre-processing 11
3.1.1 Position 11
3.1.2 Speed 12
3.1.3 Acceleration and brake pedals 13

3.2 Route description 15
3.2.1 Passenger data 17
3.2.2 Stop time distribution 18
3.2.3 Correlation analysis 21
3.2.4 Distributions 24

3.3 Positions and sections along the route 26
3.3.1 Bus stops 26
3.3.2 Intersections 31
3.3.3 Intersection lights 34
3.3.4 Crossing lights 35
3.3.5 Road bumps 37
3.3.6 Roundabout 38
3.3.7 Bridge 39
3.3.8 Other sections 39

4 CONCLUDING REMARKS 41

REFERENCES 42

Abstract

The interaction between the driver and the traffic environment is a central part in many powertrain control problems. Good prediction of the driver's actions gives the possibility to improve the powertrain control if predictive control is used. By predicting what the driver will do next, the control units in the vehicle can reduce fuel and emissions by avoiding unnecessary or unintentional accelerations. Predictive powertrain control also plays a central part in new hybrid technology, where the energy strategy benefits a lot from predictive control, so that unnecessary start up or shutdown of the internal combustion engine could be avoided. Also automated transmission could use predictive control to reduce the number of gearshifts, which costs torque and thus fuel.

By comprehensive data collection and statistical analysis of the position, speed and pedal action of a city bus in Gothenburg, a description of the bus driver behaviour is given in this work. This can be further developed to a more advanced driver model, a drive cycle generator or other applications that are based on driver behaviour. This thesis gives the data basis for such a model, where the driver behaviour at different kinds of common traffic situations is presented and analysed. For example, the driver behaviour at bus stops could be modelled as one of a number of types, depending on where it is located on the route, time of day, etcetera. This requires a more detailed analysis, preferable on several types of routes and where the CAN network in the bus is used for data collection.

1 Introduction

1.1 Background

The interaction between the driver and the traffic environment is a central part in many powertrain control problems (Sciarretta and Guzzella, 2007). Good prediction of the driver's actions gives the possibility to improve the powertrain control if predictive control is used. The number of gear shifts in an automated gearbox can be reduced and smart throttles can avoid unintentional acceleration and thereby reduce fuel. Predictive control requires a good prediction model. Many powertrain control problems require a time horizon of a couple of seconds, but for the powersplit problem in hybrid vehicles a longer horizon is necessary to obtain an efficient energy management strategy. Other interesting applications for predictive are control adaptive cruise control and predictive fuel-saving functions in automatic transmissions. (Andersson and Axelsson, 2009)

The objective of this thesis is to collect and analyse data for a city bus in traffic, to give a starting point for a longitudinal prediction model of the driver behaviour. Longitudinal prediction aims to predict the vehicle movement in the tangential direction, as opposed to the lateral direction which describes the change of the tangential direction. The collected data will be the basis for a model where the vehicle's longitudinal motion and the driver's pedal action will be modelled by differential equations and stochastic processes.[1] The data can also be used as input in drive cycle generators. A drive cycle is a sequence of torque and engine speed values, representing different kinds of driver behaviour or application areas. It can for example represent a city bus on a specific route in London or a long hauler truck in North America. The risk with fixed drive cycles is that the vehicle is optimized for a specific cycle, which might not be representative enough. By developing stochastic drive cycle generators, an infinite number of drive cycles could be generated and thus enable engineers to simulate different kinds of behaviour at the desk and reduce expensive time in vehicle.

A fundamental part of this thesis is to manage a measurement study. A typical bus route in central Gothenburg is chosen for the study, with a varying road profile and traffic density. A measurement system is developed to log the vehicle's position and the driver's pedal commands. The passenger statistics from the bus operator will be used as input for analysis of stop times at different bus stops and passenger density. The route is studied in detail where the position of each bus stop, crossing, traffic light etc will be logged.

1.2 Method

There are a large number of aspects which decide the driver's actions and which are interesting to measure and used as input to a prediction model, but for practical reasons some delimitation is necessary. For example the vehicle's position and velocity is necessary to measure to predict the future demanded torque. The vehicle's velocity in turn is affected by many factors. Some factors are dependent on the vehicle's position, such as bus stops, traffic lights, speed limits and the road slope. Other factors are somewhat dependent on time, such as traffic and the number of passengers, while weather and road works are regarded as noise in this thesis. Furthermore, the actual driver's action is of highest importance, resulting in accelerator- and brake pedal positions as input signals.

The data collection was conducted in cooperation with Göteborgs Spårvägar AB, the largest city bus operator in Gothenburg. It was important that traffic safety and passenger comfort was not affected and that the data could not be traced back to a specific driver, in case of accidents or speeding.

---

[1] The collected data has been further used in Johannesson, Lars (2009): *Predictive Control of Hybrid Electric Vehicles on Prescribed Routes*. Chalmers university of technology, Göteborg.

For practical reasons, one specific bus route in Gothenburg was chosen for the study, namely route 17 which is a so called trunk bus line from Östra Sjukhuset to Hinnebäcksgatan in Tuve. Being a trunk bus line, the route has a high passenger capacity with fewer stops than a regular city bus and with a dense timetable. It has a varying road profile with both hills and flat land, with all kinds of traffic situations and with different speed limits. The route is described in more detail under the section "Data analysis". The route is operated by a number of bus models and a Volvo B9S was chosen for this study. It is an articulated low-entry bus with a side-mounted 9 litre 340 hp diesel engine.

To collect sufficient data, the goal was to collect at least 30 hours of data, or more than 20 runs in each direction of the route, spread round the clock to cover as many events as possible, like traffic congestions, the passenger load, weather conditions and so on.

1.2.1 Variables

To conduct the study a number of variables were measured. Because there was no access to the bus internal network, the CAN-bus, sensors had to be installed in the vehicle.

Table 1 lists the variables that were found necessary to measure:

- longitude and latitude
- height above sea-level
- speed
- accelerator pedal position
- brake pedal position
- passenger flow
- time

Table 1: Variables that were measured.

In addition, every traffic situation along the route was logged and defined as one of the following types:

- bus stop
- pedestrian crossing
- pedestrian crossing with lights
- intersection
- intersection with lights
- road bump
- roundabout
- speed limit 30 km/h
- speed limit 50 km/h
- speed limit 70 km/h

Table 2: Traffic situation types along the route.

This was done by driving along the route with a car and logging every situation with a GPS-receiver and specially developed software.

2 Measurements

2.1 Equipment

2.1.1 Planning and implementation

Because there was no access to the vehicle's internal network, the CAN-bus, sensors had to be installed. A GPS was used for position and velocity measurements and two optical distance sensors were used for pedal position measurements. A laptop computer with a data acquisition device collected and stored the data. The computer needed stable power supply. The equipment needed to be stored safely in the vehicle, both in regards to traffic safety aspect and so it would not be stolen.

The bus's electronics box, placed over the front left wheelhouse, was large enough to hold the equipment needed, had a lock and was easily accessible in the bus. 30 V voltage supply was available in the box and the box was situated close to a window, which was necessary for the GPS antenna cable. Therefore the measurement equipment setup was built within the electronics box. To ensure operation stability, the equipment's effect on the temperature in the box had to be calculated. The heat flow through the box walls is calculated with the following formula,

$$\Delta T = \frac{1}{A\lambda} \cdot Q \cdot dx \approx 6.6°C$$

Equation 1: Calculation of heat transfer through the box.

where A is the box mantle surface, dx the thickness of the box, λ the thermal conductivity for pine wood and Q is the power of the computer (65 W). The heat transfer is indirectly related to the volume inside the box, since the mantle surface increases with a larger volume. The added heat was well within the operating temperature interval for the computer.

The two optical distance sensors used for pedal position were installed behind the pedals on a flexible chase holder enabling easy angle adjustment. The sensors were placed 12 cm from the pedal.

2.1.2 Components

GPS

The GPS receiver used for the measurements was a GARMIN GPS 18 5Hz, which samples data at 5 Hz, as the model name indicates. It communicates with the computer via a USB interface, developed by Fägerlind, Hermansson and Rönnberg (2007). The antenna was easily affixed to the bus roof with a strong magnet and the antenna cable was 5 meters long, which was more than enough.

NI-DAQ

A National Instrument data acquisition device (NI-DAQ) was needed to sample data from the pedal sensors. It has 8 analog input connections, 2 analog output connections and 12 digital I/O connections, in addition to 5 V, 2 V and ground outputs. The NI-DAQ is compatible with both LabVIEW and Matlab.

SHARP GP120D

Two SHARP GP120D infrared distance sensors were used as pedal position sensors. The sensor can measure distances from approximately 4 cm up to 30 cm. The sensor operating supply voltage is 4.5 to 5.5 V, with an output voltage proportional to the distance from the object. Unfortunately no cables were delivered with the sensors and no retailer in Sweden sold cables that fitted into the housing on the sensors. Since the shipping time for the sensors was a couple of weeks no supplementary order was placed. Instead contacts were built with connector pins and duct tape, since lasting performance was not a priority in comparison to the delivery time.

2.1.3 Software

LabVIEW was used to interpret and store data. The program for the measurement system developed by Du (2006) and Fägerlund, Hermansson and Rönnberg (2007) was adjusted to suit the project. More or less, the communication interface with the NI-DAQ and GPS was kept intact, but the interpretation of the signals was modified. The data was stored in two main logfiles, one for GPS data and one for driver interaction data. No pre-processing filtering of the data was done in LabVIEW.

2.2 Data Collection

Data was collected during two weeks in November 2007, spread from 08.00-02.00 hours. The time intervals and days were chosen to cover as many situations as possible, such as traffic congestion, passenger flow and different driver behaviour. The sampling frequency was 5 Hz for all signals except for the passenger flow which is one sample every time a passenger pays the fare in the machine on the bus. The time for the sample was then matched with the position signals to define between which stops the fare was paid.

3 Data Analysis

3.1 Pre-processing

The raw data had to be pre-processed before analysis of the data was possible. No filtering was made in the data acquisition system, because of the risk to loose useful data.

3.1.1 Position

Fägerlind, Hermansson and Rönnberg (2007) pointed out a couple of problems associated with GPS measurements, derived from the error sources explained in Koff et al (2005). At low speed the position error is notably, but even at high speeds some samples are incorrect.

Distance calculations are used to identify which samples are nearest to a specific traffic situation, the starting point for all sorting of the data. Since a passing vehicle will not pass exactly at the longitude and latitude for the traffic situation, the closest sample to the desired position has to be calculated. It is therefore important that a sequence of samples are not scattered too much, otherwise a sporadic sample with large errors might be registered as the position in question.

After a position sample has been defined as the sample closest to the desired position, the position, velocity and pedal signals are transformed from the time domain to the distance domain, by calculating how many samples before and after the position that represent a desired distance boundary, usually ±200 meters.

Errors at low speeds looks like as if the bus is straying in an impossible manner, seen in Figure 1 at the first turn. Errors at high speeds are sporadic samples out of place, like the sample east of the third turn. The high speed errors are corrected by filtering the longitude and latitude signals with a moving average filter, since the position for natural reasons can not change abruptly from one sample to another. The average is calculated over a span of 30 samples, around 6 seconds. The low speed errors remain to some extent, but since they are in line they don't affect this application, since an error at low velocities has less impact because of the transformation to the distance domain. All samples where the velocity is zero are thus calculated as the same distance.

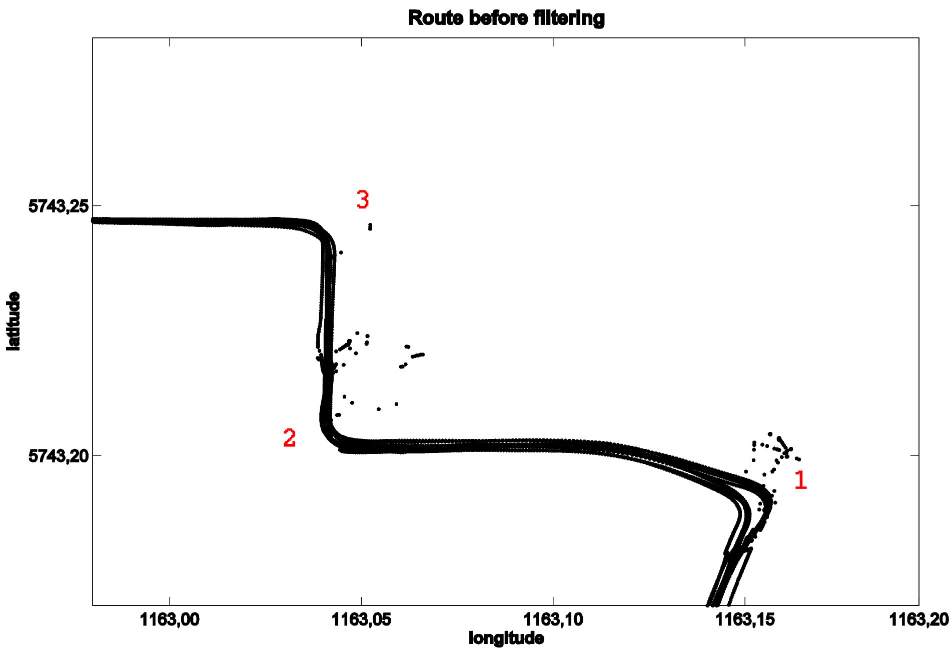

Figure 1: Raw position signal for several runs past Östra Sjukhuset. Erroneous samples are seen at turn 1, 2 and 3.

The importance of removing high velocity errors is illustrated in Figure 1 and Figure 2. Suppose the number "3" in Figure 1 marks the logged position of an intersection. The closest sample would then be an erroneous sample. The impact of this is shown in the filtered signal in Figure 2. The position of this erroneous sample, marked with the red arrow, is now located several meters away. As seen in Figure 2, the moving average filter gives a satisfying improvement, although there are still errors.

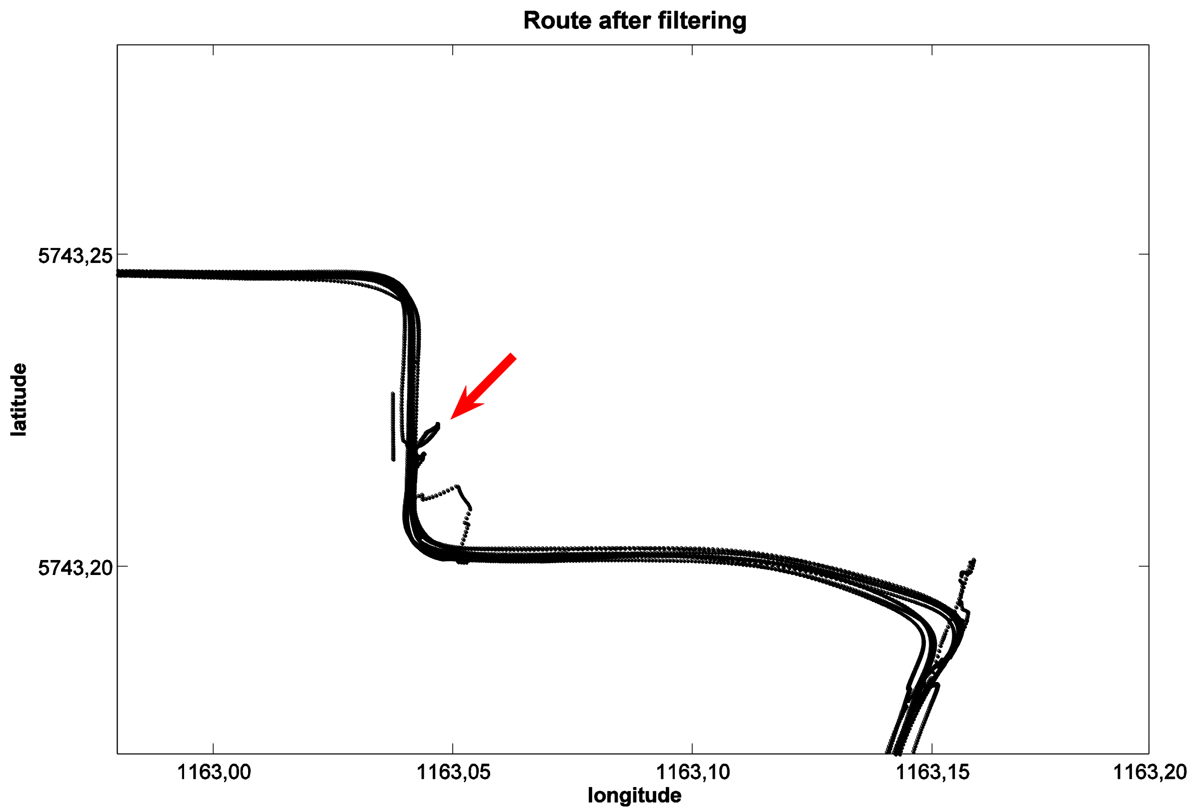

Figure 2: Filtered position signal for several runs past Östra Sjukhuset. The erroneous samples at turn three are now removed and located marked by the red arrow. The erroneous samples at turn one and two are still erroneous, but their impact when transforming into a distance separated domain is low since the velocity is low.

Since the geographic coordinate system with longitudes and latitudes is defined in degrees, minutes and seconds, there is a numerical offset when one minute changes to another. The GPS format for longitude and latitude is called decimal degree (DDMM.MMMM D=degree and M=minute) so the minutes are decimal separated but one minute is 1/60 of a degree. Therefore the longitudes' numerical value leaps from 1159.9999 to 1200.0000 close to Redbergsplatsen in Gothenburg, which leads to computational complications. For practical reasons this numerical offset is removed prior to data analysis. This give erroneous absolute longitude values in the eastern parts of the route, but it does not affect the analysis.

3.1.2 Speed

The vehicle speed was measured by GPS which according to the manufacturer GARMIN's specifications is quite accurate, with an error less than 0.2 km/h, which is well within the requirements for this project.. The GPS sample frequency was 5 Hz.

The speed signal is filtered with a zero-phase digital filter (filtfilt in Matlab) to remove the measurement noise, seen in Figure 3(a) by filtering in both forward and backward direction with the following filter:

$$y(n) = b_0 x(n) + b_1 x(n-1) + b_2 x(n-2) + \ldots + b_9 x(n-9) - a_1 y(n-1) - a_2 y(n-2) - \ldots a_9 y(n-9)$$

Equation 2: Zero-phase filtering of the speed signal

Unfortunately, the GPS sometimes loose connection with the satellites and no speed is measured for a number of samples, as seen in Figure 3(a). If such a signal is filtered, then the signal will be interpreted as a deceleration zone, as seen in Figure 3(b). For this reason constant acceleration is assumed where the signal is erroneous, as in Figure 3(c), which is then zero-phased filtered, Figure 3(d).

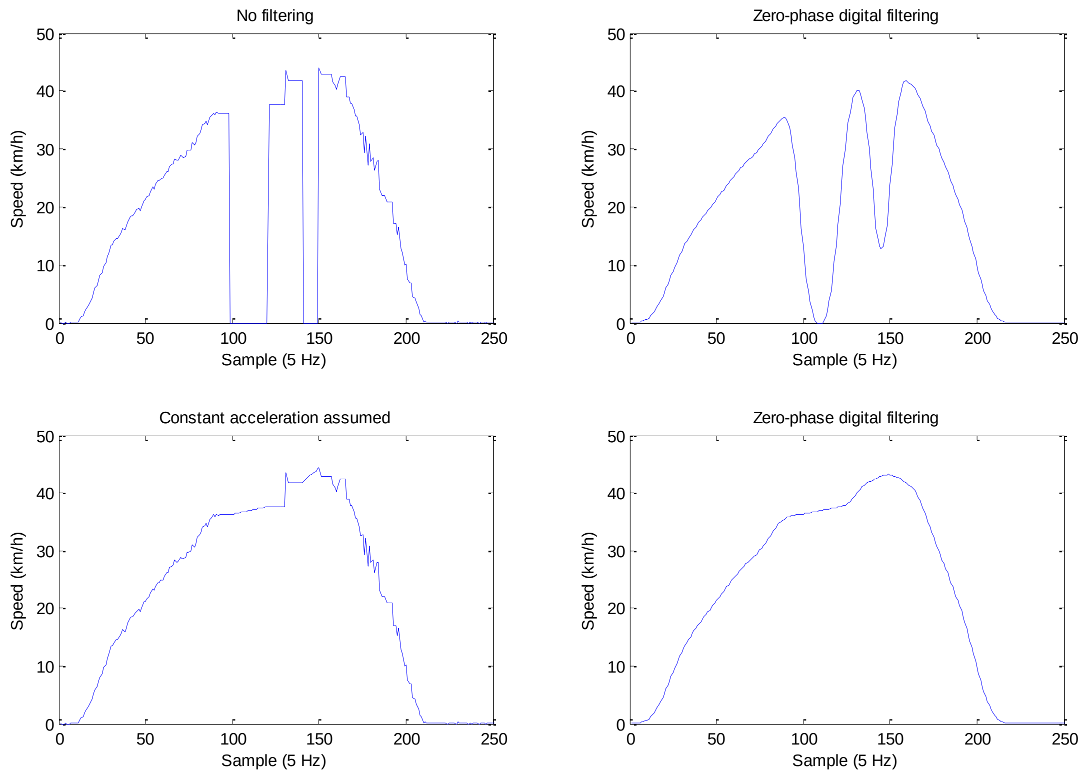

Figure 3: Two-step filtering of speed signal. 3a is upper left corner, 3b is upper right corner, 3c is bottom left corner and 3d is bottom right corner.

3.1.3 Acceleration and brake pedals

The driver's pedal actions are measured with two infrared distance sensors behind the accelerator- and brake pedals. The sensors output voltage is proportional to the distance between the pedal and the sensor and is sampled at 10 Hz with the NI-DAQ. The signal is then filtered with a zero-phase digital filter and downsampled to 5 Hz for comparison with the other signals.

The accelerator pedal has a mechanical maximum position, whereas the brake pedal is pneumatic which means that the true maximum position is rarely used. Instead, calibration of the sensors has been done by analysing the signal afterwards. The mapping from distance to position is not exactly linear, but the effect of the nonlinearity is too small to be regarded.

To simplify interpretation of the accelerator and brake commands, three modes of commands are defined – not used, used and floored. Hence a signal is constructed by thresholding of the accelerator and brake signals. Each signal is quantized from the table below, where the brake signal has negative numbers and the accelerator signal has positive numbers.

Percentage of max	Mode	Level
< 20 %	Not used	0
20% - 66 %	Used	± 1
> 66 %	Floored	± 2

Table 3: Threshold levels for driver signal.

It is primarily the driver's combined actions that are interesting for analysis. The level is thus added into one driver signal, with values ranging from -2 to 2. The result is seen in Figure 4 below. The interpreted signal gives a good representation of the drivers' pedal actions and transitions between acceleration and brake modes.

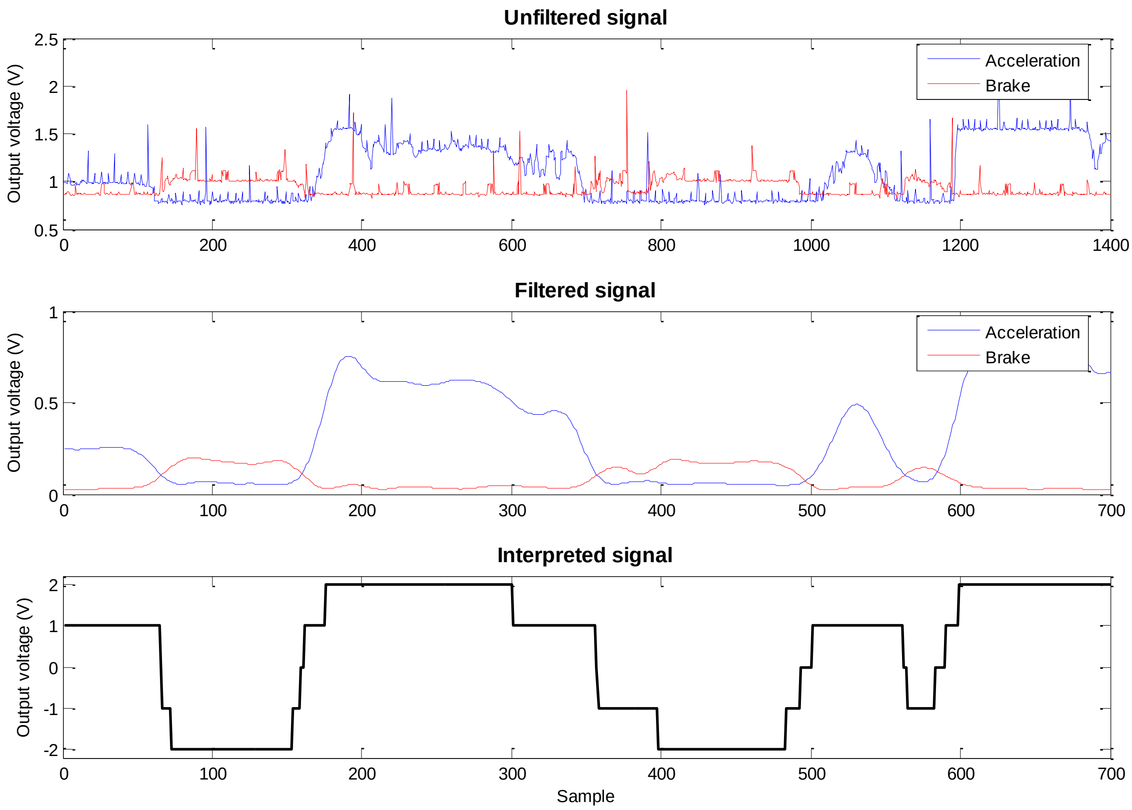

Figure 4: Filtering and thresholding of the driver signal. The brake and accelerator pedal signal is combined into one driver signal, ranging from +2 indicating a floored accelerator pedal to -2 indicating a floored brake pedal.

3.2 Route description

As stated above, the route in question is the trunk bus line number 17 in Gothenburg, Östra Sjukhuset – Tuve – Östra Sjukhuset. It is about 16 kilometres long in each direction.

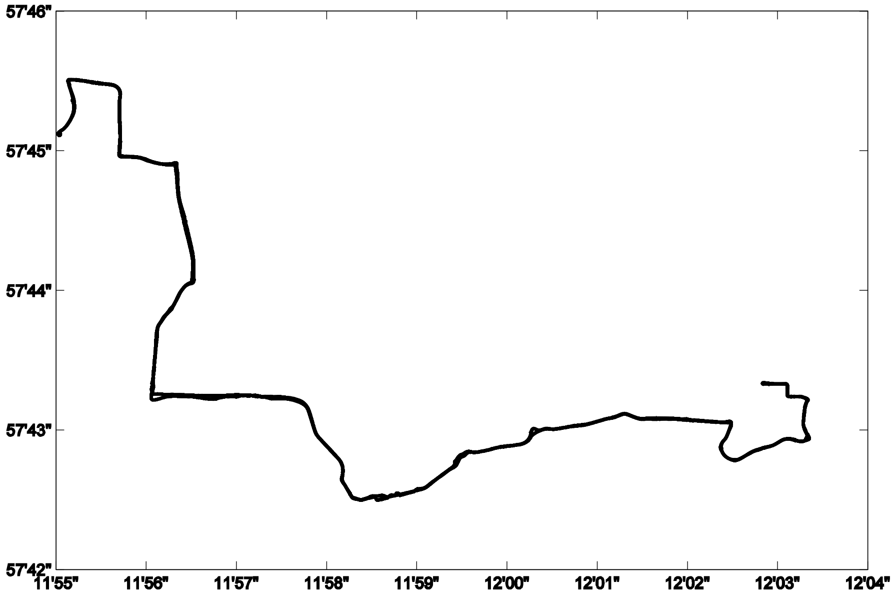

Figure 5: The studied trunk bus line number 17 in Gothenburg, Sweden.

Fägerlind, Hermansson and Rönnberg (2007) showed that the accuracy in the GPS height measurements were poor. Still the approach was that with many drive cycles, the topography could be estimated. Unfortunately, it turns out that the height measurements are so poor that no estimation can be done, as seen in Figure 6 below. The height difference at a given position can be up to 60 meters.

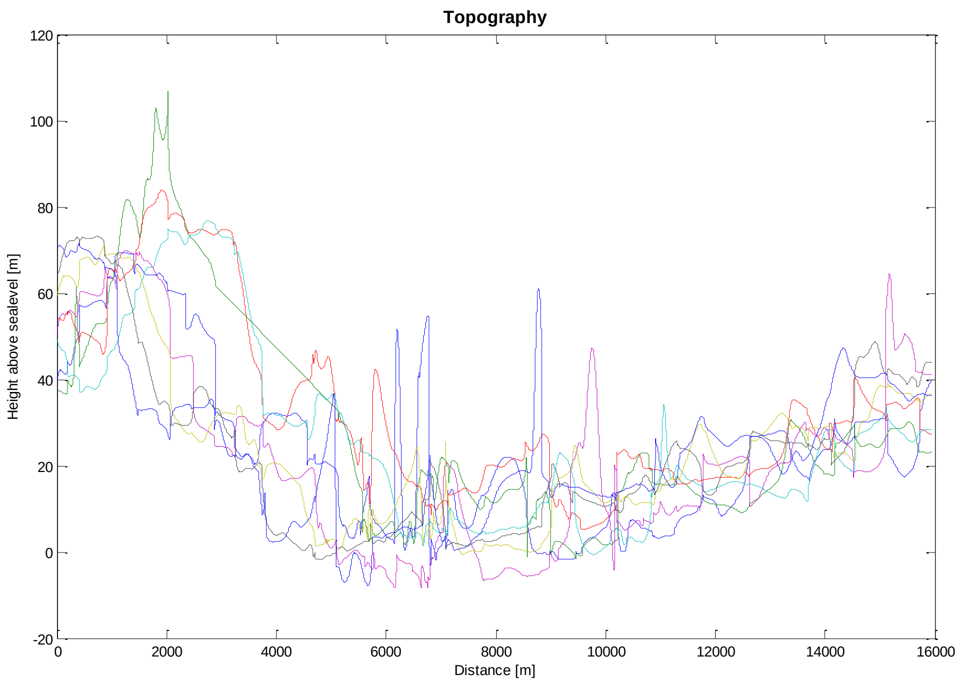

Figure 6: Height measurements along the route.

Under the assumption that there is no systematic source of errors, the mean height could be used. Figure 7 clearly shows that there must be an error source, since the mean height clearly can not be true with instant steep inclinations if the road is drivable. The mean height could possibly be improved with several more measurements. The route height according to Google maps[2] data is shown in **Figure** 8.

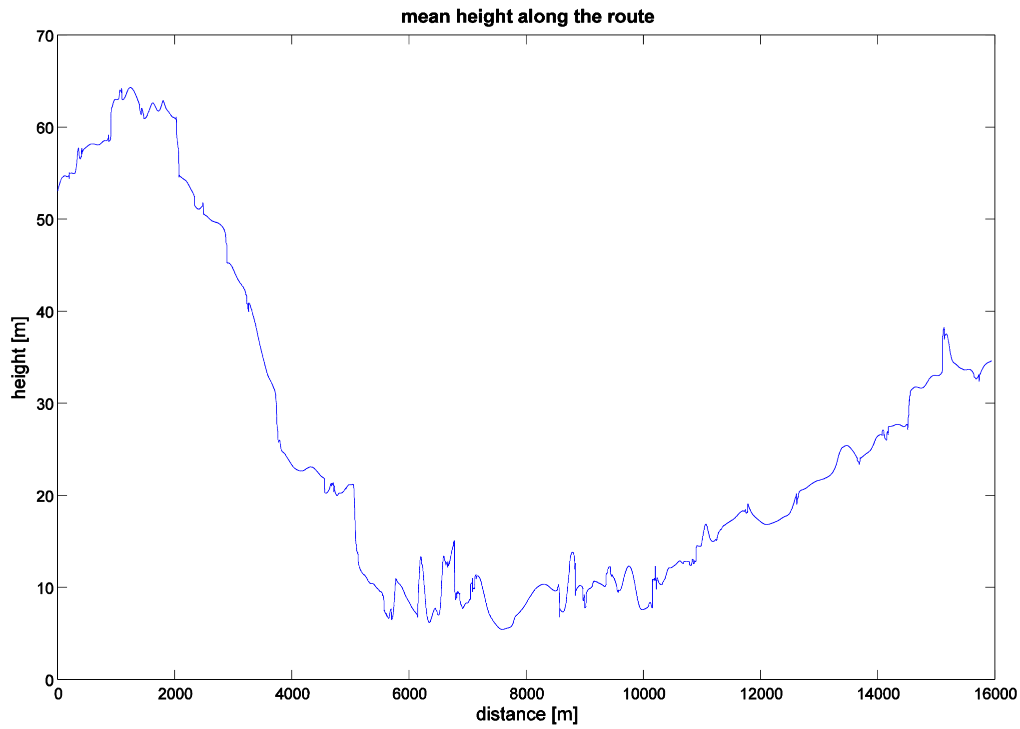

Figure 7: The mean height along the route in the Tuve direction.

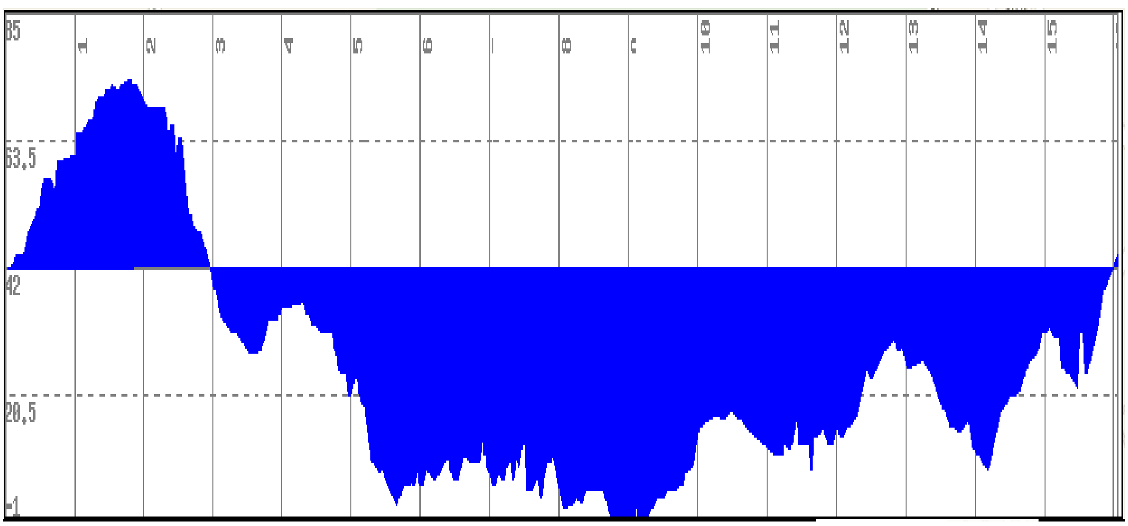

Figure 8: The altitude profile of the bus route according to Google maps data

[2] http://www.gmap-pedometer.com/?r=6197200

The velocity trajectories along with the probability mass functions are shown in Figure 9 below. There are a total of 18 trajectories in each direction. As one can see, the trajectories coincide a lot, except for longer sections without bus stops or intersection lights, which will be shown later in the report.

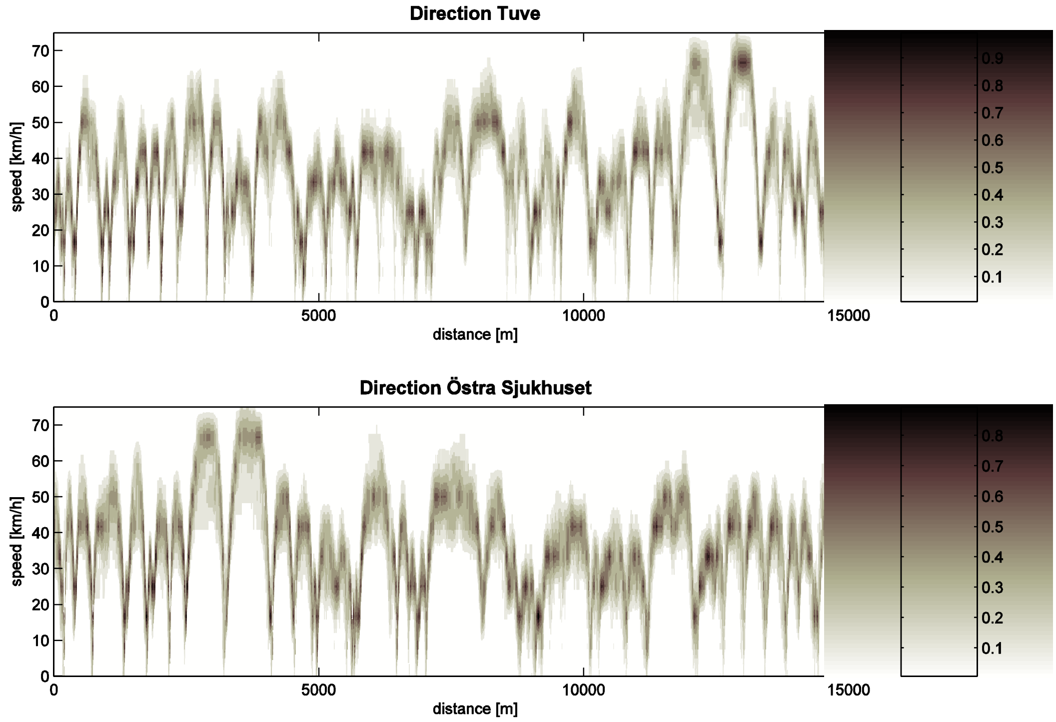

Figure 9: The probability mass function along the route. Darker color means higher probability for a velocity. For each sample the speed was quantized into ten levels.

In total 65 positions along the route are analysed, consisting of 29 bus stops, 4 intersections, 18 intersection lights, 2 crossing lights, 8 road bumps, 3 roundabouts and one bridge. The last position type, crossings, did not affect the driver behaviour in any high degree and was skipped in the analysis. One reason behind this might be that crossings are often placed close to intersections or bus stops and those types affect the driver behaviour more than crossings.

3.2.1 Passenger data

Passenger statistics shown in Figure 10 was received from Västtrafik AB, coordinators of public transportation in the Gothenburg region. The passenger data consists of time and date for every paying passenger for the given bus and days when the data has been logged. It is then broken down for each bus stop by comparing the time stamps for registered fare payments with the time stamps in the position log files and under the assumption that a passenger boarding the bus at one stop pays the fare before the next stop.

The stops with most waiting passengers are in central Gothenburg, around stops 14 to 17. There are few paying passengers in the end of the route in both directions, because travellers seldom board the bus in the end of the route.

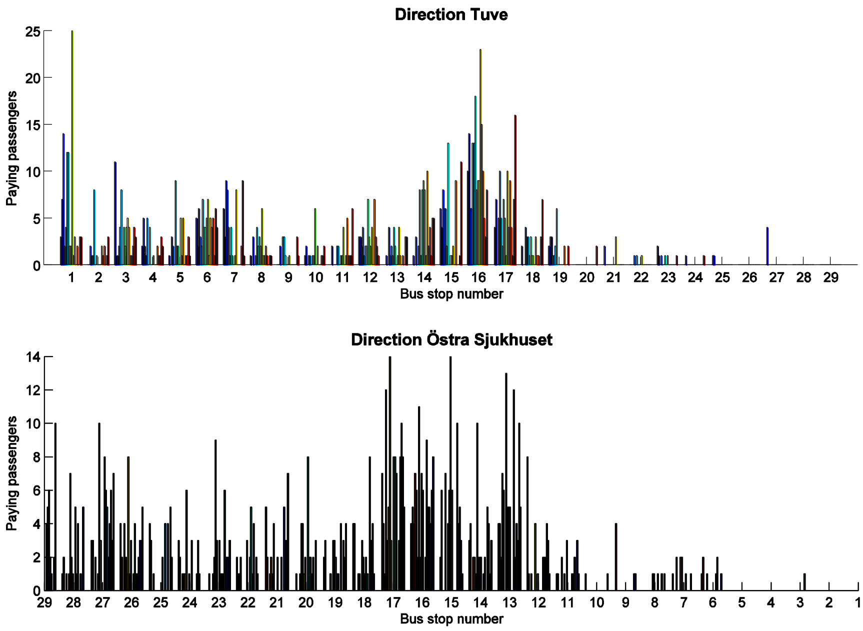

Figure 10: Number of paying passengers for each bus stop and runs. Each bar represents a stop at the bus stop number given by the x axis. Most passengers boards the bus in central Gothenburg (stops 11-18) and there are more paying passengers in the beginning of the route, which is natural since it is then more bus stops left to go to.

The passenger data is of course for paying passengers. Unfortunately not all passengers with period cards register their card upon boarding of the bus, and some passengers do not pay at all. Still, the relative change between bus stops is relevant.

3.2.2 Stop time distribution

Figure 11 shows statistics for the stop time along the route. The average stop time does not change much during the day, although it reaches its max during the afternoon.

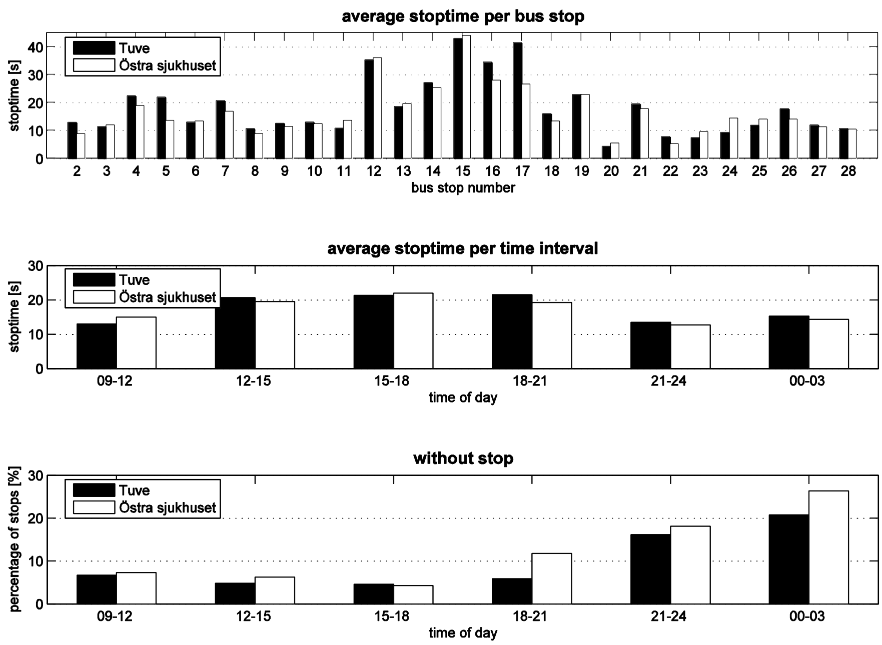

Figure 11: Stop time per stop, the average stop time depending on time of day and the share of passed bus stops without stop.

The number of skipped bus stops is at its low at rush hours while peaking at late hours. The average stop time per bus stop is at its high in central Gothenburg from bus stop 12 (Redbergsplatsen) to bus stop 17 (Hjalmar Brantingsplatsen). The stop time is dependent on both the boarding and deboarding passengers, and therefore not as dependant on whether the bus stop is in the beginning or in the end of the route as for the paying passenger statistics. The somewhat higher stop time in the lower bus stop numbers could be explained by the denser population around Björkekärr and Härlanda (stops 1 to 11) than around Tuve (stops 22 to 29). The stop times for the start and end of the route has been excluded from the data, since these stop times are more dependent on schedules than on boarding passengers.

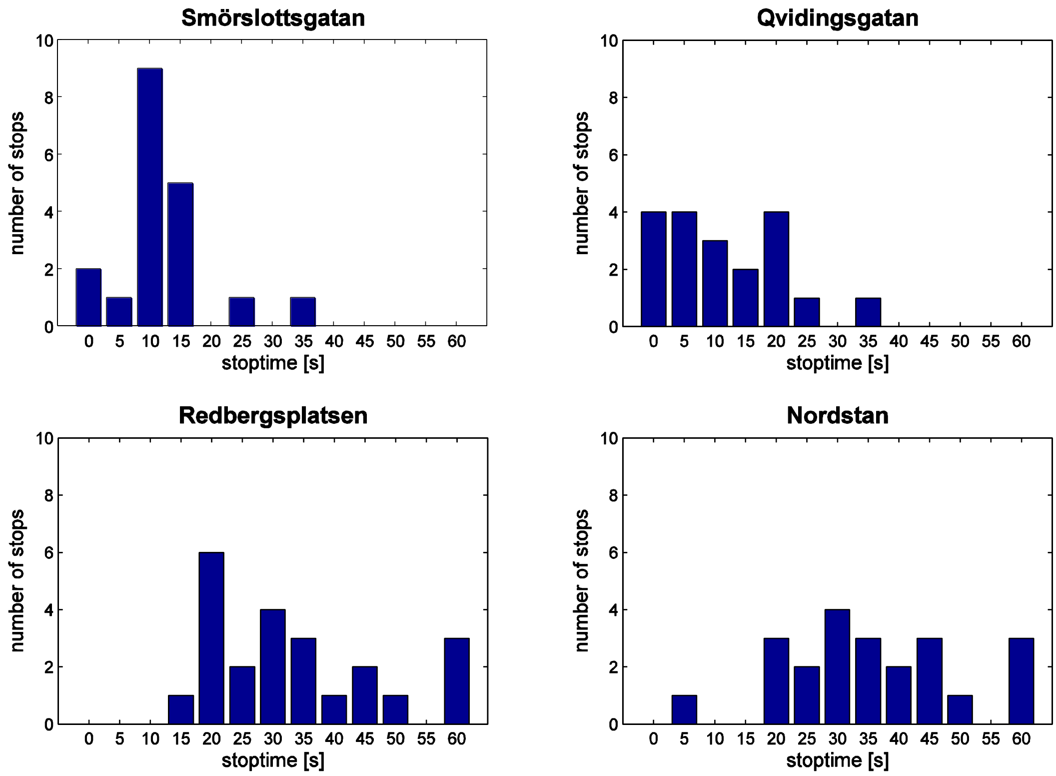

Figure 12: Histograms of the stoptimes for four different bus stops along the route.

The distribution of the stop times is interesting to look at for modelling purposes. Figure 12 shows the stop time histograms for four chosen bus stops. A general distribution of the stop times for all bus stops can not be given, although a few different distributions depending on the type of stop could perhaps be possible. For example the stops Redbergsplatsen and Nordstan, being commuter hub bus stops, have a similar distribution of stop times. The stop time for two of the chosen bus stops above, Qvidingsgatan (stop 9) and Nordstan (stop 16), during rush hours and late hours are shown in Figure 13.

As shown below, the stop time varies both depending on the bus stop characteristics and on the time of day. The numbers of stops are too few to form a distribution. A box plot for the stop time is shown in Figure 14. The box has lines at the lower quartile, the median value and the upper quartile values. The broken line show the extent of all of the stop times, except for outliers which are indicated with the red plus sign.

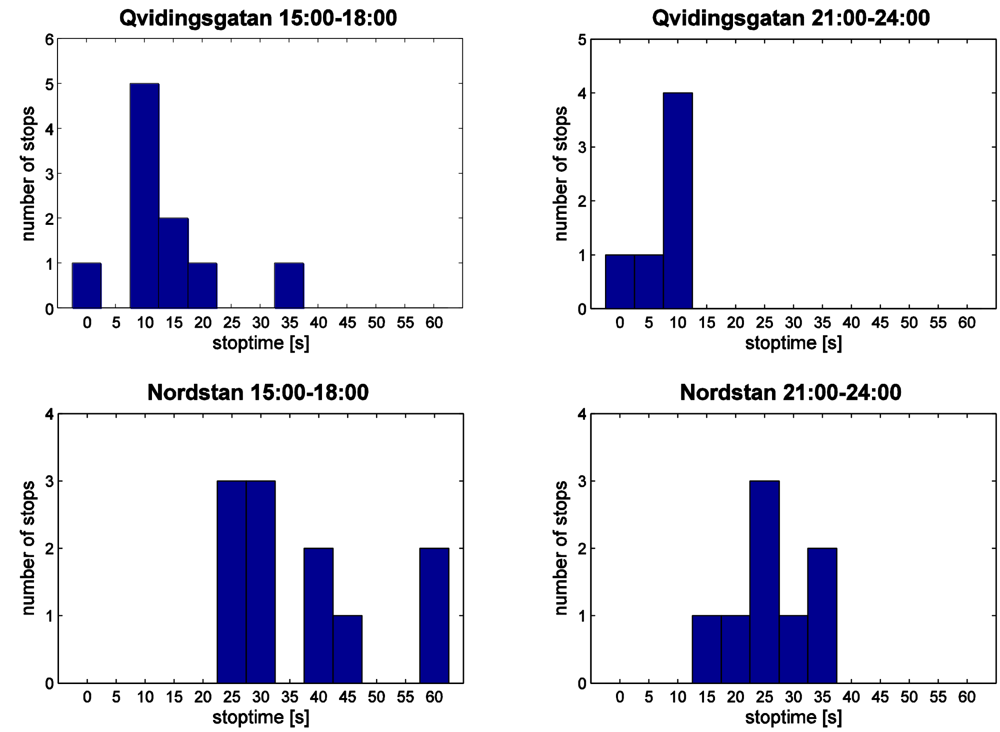

Figure 13: Stop time histograms for Qvidingsgatan and Nordstan at rush hours and late hours.

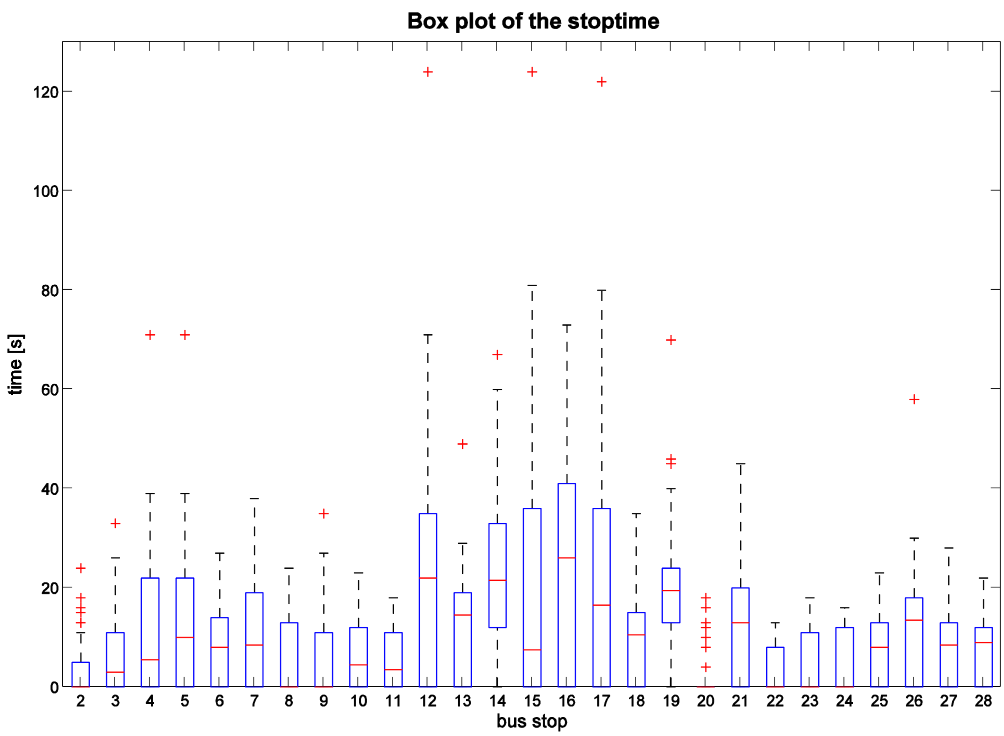

Figure 14: Box plot of the stop time for each bus stop, where the boxes indicate the lower and upper quartiles along with the median stop time in red. The broken line gives the range of the stop times except for outliners indicated with a red plus sign.

3.2.3 Correlation analysis

In modelling and simulation the statistic feature correlation can be a powerful tool, as it indicates relationship between two variables. A strong relation between the speed prior to a bus stop and the speed at the bus stop could indicate whether the bus driver will stop or not. The correlation coefficient can indicate such a relationship, but its value alone is not enough to determine a relation, which will be shown below.

Correlation is a statistic feature that indicates a supposed linear relationship between two random variables X and Y. The correlation coefficient is defined as

$$\rho_{x,y} = \frac{Cov(X,Y)}{\sqrt{(VarX)(VarY)}}, -1 \leq \rho_{x,y} \leq 1$$

Equation 3: Definition of correlation coefficient

If $\rho = 1$ there is a perfect positive correlation between X and Y, which implies that small values of X is associated with small values of Y and large values of X is associated with large values of Y. If $\rho = -1$ there is a perfect negative correlation between X, and Y, which implies that small values of X is associated with large values of Y and large values of Y is associated with small values of Y. If $\rho = 0$, X and Y are uncorrelated, but that does not imply that they are unrelated. If a relationship between X and Y exists, it is just not linear.

The computational formula for the estimated Pearson correlation coefficient r is

$$\hat{\rho} = r = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left[n\sum x^2 - \left(\sum x\right)^2\right] \cdot \left[n\sum y^2 - \left(\sum y\right)^2\right]}}$$

Equation 4: The Pearson correlation coefficient formula

The interpretation of the correlation coefficient is dependent on the area of application, but $|r|>0.5$ is often interpreted as a moderate to strong correlation.

The correlation coefficient gives a hint of the linear relationship between two variables, but its value alone may not be enough to state a relationship. Anscombe's quartet in Figure 15 consists of four different datasets, but with the same mean and variance for each x and y variable, as well as the same correlation.

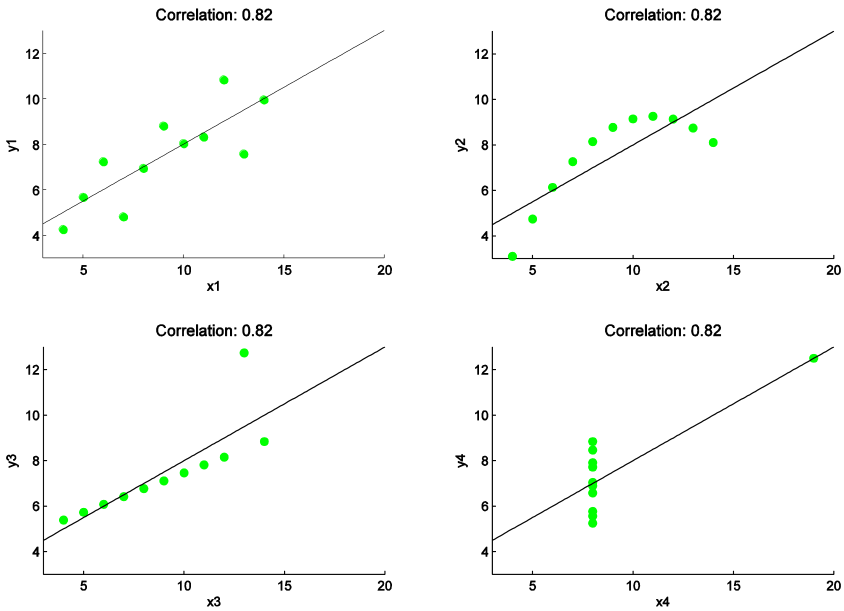

Figure 15: Anscombe's quartet of series with the same correlation

3.2.3.1 Correlation of velocities at different positions

For correlation between the velocity trajectories at one position and another position, the velocity values at the two positions give the random variables X and Y. An example is given for the bus stop Gunnesgärde in both directions in Figure 15. As one can see, there is a slight negative correlation between -200 m and the stop in the Tuve direction, but the same position is almost uncorrelated in the other direction. On the other hand, there is a strong correlation between -200 m and 200 m for both directions. This kind of autocorrelation is mostly useful for analysing the stringency of the driver behaviour, i.e. if a driver's speed at two positions has the same relation to the other driver's speed. A negative correlation between a stop and a position prior to the stop can indicate passing without stop, as for Gunnesgärde in the Tuve direction in Figure 16 below, but if the passing velocity trajectories are higher than the other trajectories in both positions, the correlation is positive. To conclude, the autocorrelation of the velocity trajectories does not give much useful information. Figure 17 show the correlation between the stop and positions up to 200 m prior to the stop.

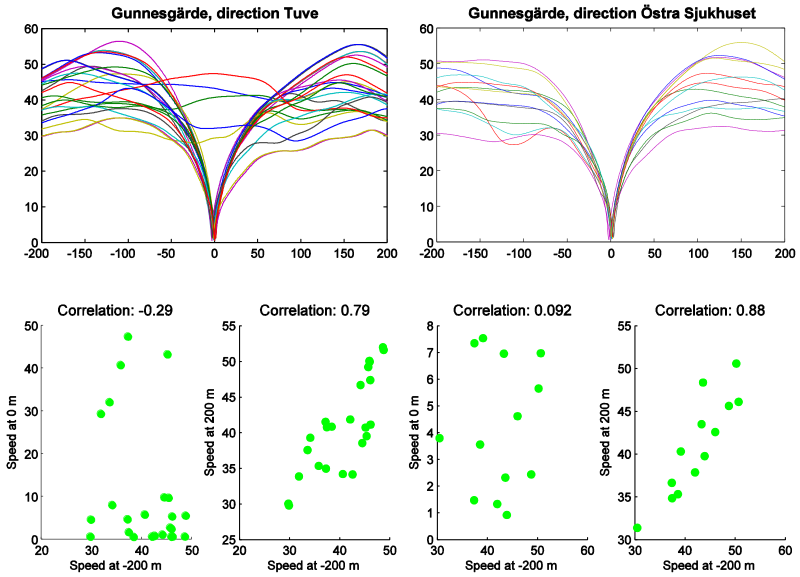

Figure 16: Correlations for velocity trajectories at positions from -200 m to 0 m and 200 m respectively.

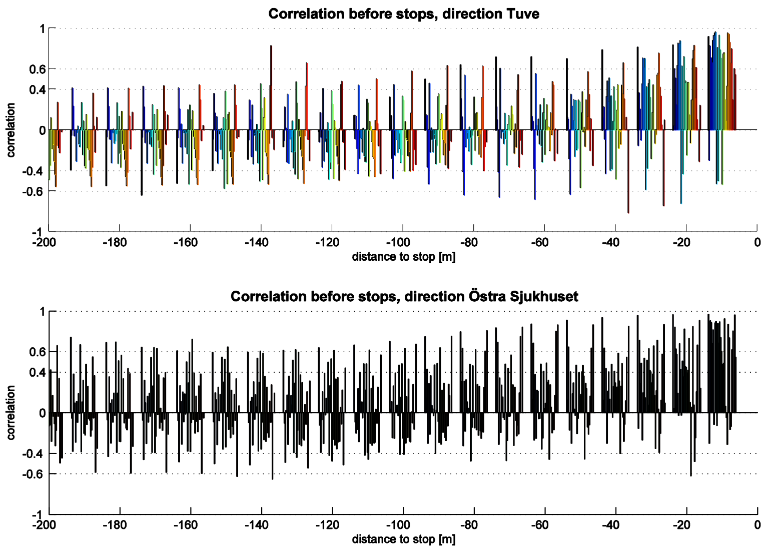

Figure 17: Autocorrelation between the bus stop and positions, from 10 meters prior to the stop and up to 200 meters prior the stop. Each bar represents one of 29 bus stops.

3.2.3.2 Correlation of driver input vs. velocity

The cross correlation between driver behaviour and velocity indicates the brake zone prior to a bus stop. At first, both the speed and driver behaviour signals are high, which gives a positive correlation. When the driver starts to brake, the driver behaviour signal is low while the speed is still high, which gives a negative correlation. When the speed decreases, the correlation is positive once again. This trend can be seen in Figure 18.

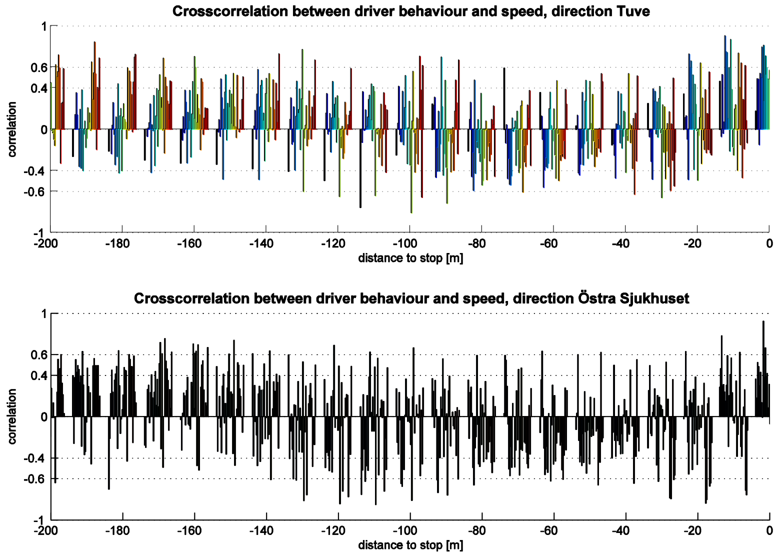

Figure 18: Cross correlation between driver behaviour signal and speed at positions from the bus stop up to positions 200 meters prior to the bus stop. Each bar represents one of 29 bus stops.

3.2.3.3 Correlation of driver input vs. acceleration

There is of course a strong positive cross correlation between driver behaviour signal and acceleration, as seen in Figure 19. The exceptions could be explained from a small offset between the driver behaviour signal and velocity/acceleration signal, which lead to negative correlation at samples close to abrupt changes in the acceleration, where the driver signal might still be low while the acceleration signal is high.

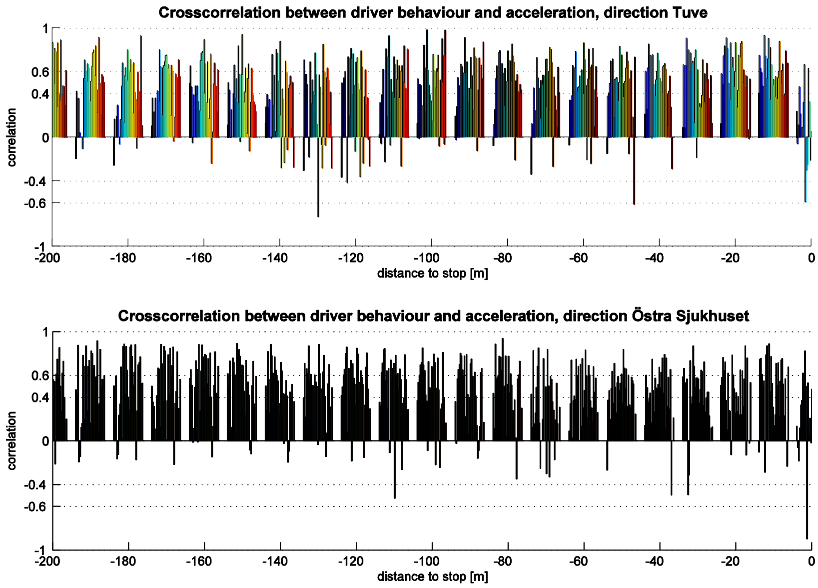

Figure 19: Cross correlation between driver behaviour signal and acceleration at positions from the bus stop up to positions 200 meters prior to the bus stop. Each bar represents one of 29 bus stops.

3.2.4 Distributions

Markov processes can be used to model the longitudinal dynamics of a vehicle, the average distribution of acceleration and velocity (Johannesson 2006). The markov process is defined by the conditional distribution function of the acceleration A over the next time step, given the markov state x at time k.

$$P(A_{k+1} \leq a_k \mid \mathbf{x_k})$$

Equation 5: The markov process for acceleration A over the next time step.

After choosing markov state variables a suitable conditional distribution function of the acceleration needs to be chosen. The conditional distribution function has a limited number of parameters, in turn dependent on the markov state X.

Johannesson (2006) use velocity and acceleration as markov states and a beta distribution as the distribution function, giving α and β as parameters in the markov process. The formula for the probability density function of the beta distribution is

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$$

Equation 6: The probability density function of the beta distribution.

where $B(\alpha, \beta)$ is the Beta function given by

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx$$

Equation 7: The Beta function.

In the application described in this report, modelling of longitudinal dynamics of a city bus, the acceleration state in the markov process could be replaced by the pedal position, giving the markov states vector

$$\mathbf{X_k} = (V_k, P_k)$$

Equation 8: Markov state based on velocity and pedal position.

where V_k is the velocity and P_k is the pedal position.

The velocity-acceleration diagram in Figure 20 gives an idea how the vehicle is operated. The vehicle is mostly driven at speeds ranging from 25 km/h to 45 km/h with modest acceleration and with quite heavy acceleration and retardation at low speeds. The diagram was constructed by dividing the acceleration into intervals of 0.1 m/s^2 and the velocity into intervals of 1 m/s. It was then quantized into levels of 5 % of the maximum count.

The next task is to find a suitable conditional distribution for the acceleration. The beta distribution is easily shaped with only two parameters α and β and since it is only valid in the interval [0 1] it is suitable for vehicle applications, because limitations in the driveline are considered if the variables are normalized with respect to the limitation. Figure 21 shows the histograms for the next accelerations at centre velocities of 30, 40 and 50 km/h and centre accelerations at -0.5 m/s^2, 0 m/s^2 and 0.5 m/s^2.

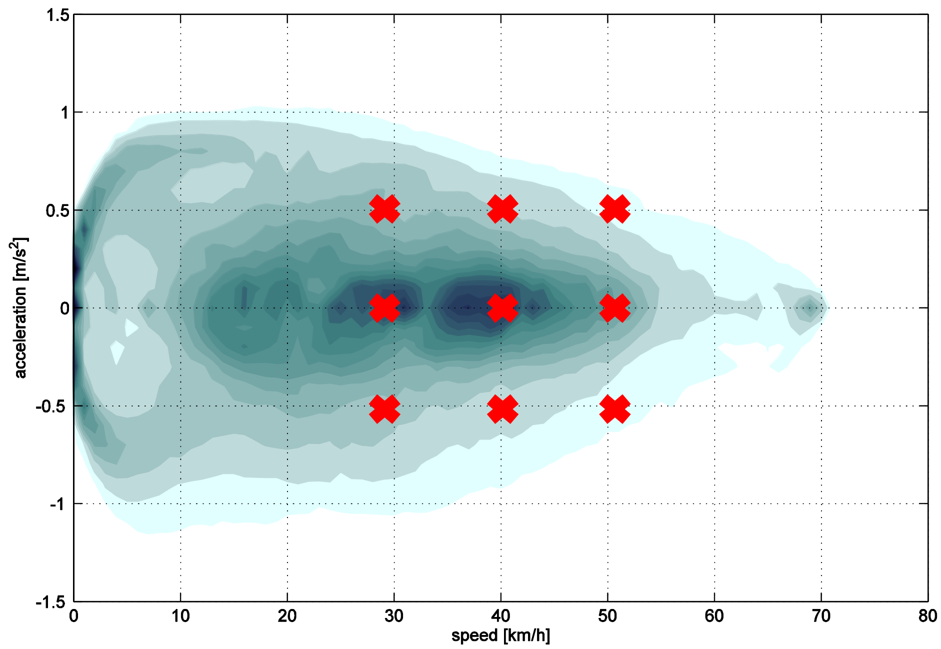

Figure 20: Velocity-acceleration density diagram of the data. Darker colour means higher frequency of the given speed/acceleration combination. The red crosses indicate the chosen points for investigation of suitable conditional distribution function.

The velocities and accelerations have been chosen from Figure 20 to give a good representation of the velocity-acceleration density of the studied route. The parameters α and β in the beta distribution have been computed to maximize likelihood for each histogram in Figure 21 with the Matlab function *betafit*. The beta distribution represents the histograms quite well and is therefore chosen as probability density function in this report.

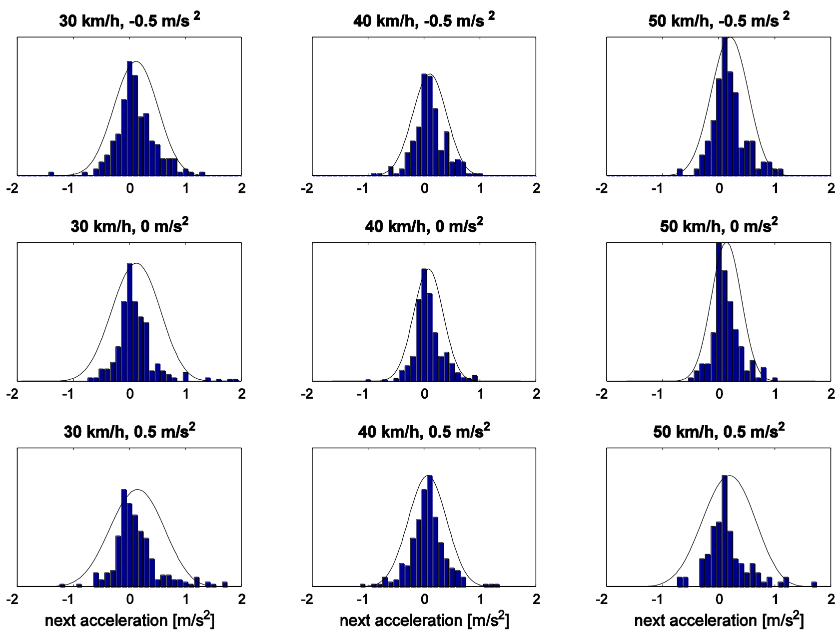

Figure 21: Histograms of next acceleration for different combinations of current velocity and acceleration, along with the corresponding beta distribution.

The histograms for the next pedal position are shown in Figure 22 below. It shows the next pedal position at the next second for the current velocities and pedal positions shown above the plot. In contrary to the histograms of the next acceleration the histograms are not centred.

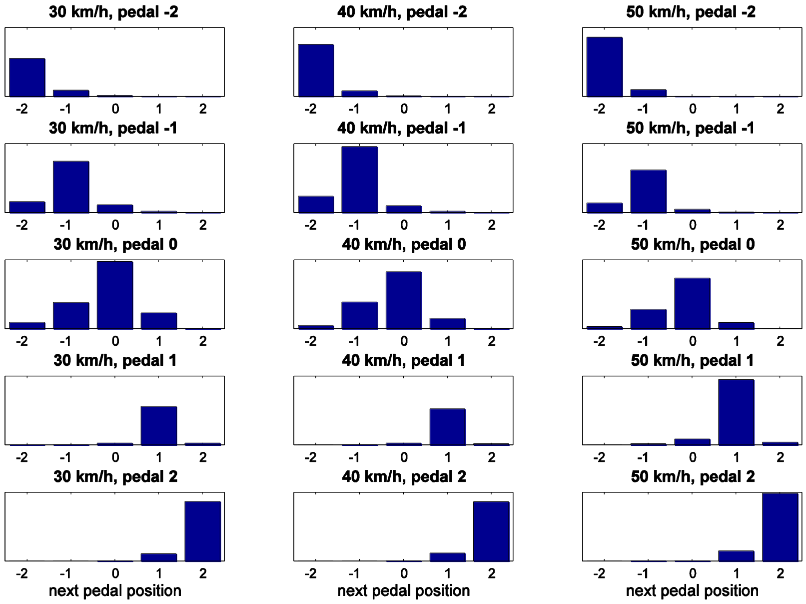

Figure 22: Histograms of the next pedal position for different combinations of current velocity and pedal position.

3.3 Positions and sections along the route

3.3.1 Bus stops

There are in total 29 bus stops along the route. The stops along with the stop numbers used consistently in the report are plotted in Figure 23 below.

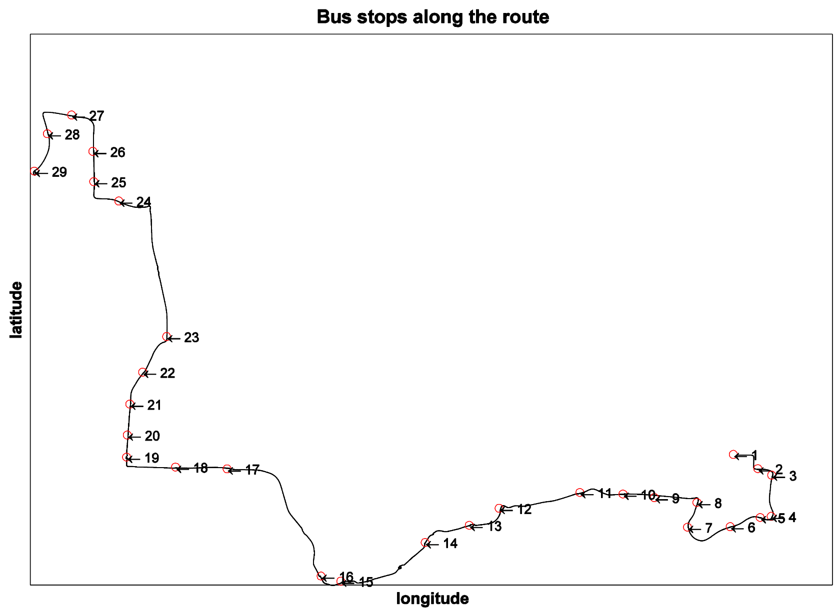

Figure 23: The bus stops along the route.

Table 4 lists the bus stop numbers along with their names.

Nr	Name	Nr	Name	Nr	Name
1	Östra Sjukhuset	11	Härlanda	21	Taklöksvägen
2	Barnkliniken	12	Redbergsplatsen	22	Bäckedalsvägen
3	Smörslottsgatan	13	Olskrokstorget	23	Grimbo
4	Spåntorget	14	Svingeln	24	Gunnesgärde
5	Trätorget	15	Centralstationen	25	Brunnehagen
6	Stabbetorget	16	Nordstan	26	Tuve Centrum
7	Studiegången	17	Hjalmar Brantingsplatsen	27	Norumsgärde
8	Bovallstrandsgatan	18	Vågmästareplatsen	28	Norumshöjd
9	Qvidingsgatan	19	Wieselgrensplatsen	29	Hinnebäcksgatan
10	Munkebäckstorget	20	Wieselgrensgatan		

Table 4: Bus stop numbers and names.

For practical reasons only four stops are presented in the report; they are chosen to show the diversity among the stops. Only stops in the Tuve direction are considered here. Statistics for all stops are presented in the appendix.

Stop number 9, Qvidingsgatan, is chosen as a bus stop in the beginning of the route with fewer stops and waiting passengers. Svingeln in central Gothenburg, stop number 14, is a hub for commuters with an uncomplicated traffic situation. Nordstan, stop number 16, is the main commuter hub along the route with sometimes queuing vehicles, leading to multiple stops. The last stop to be presented is stop number 26, Tuve Centrum, found in the end of the route and characterized by many stops and spread velocity trajectories. These characteristics are shown in Figure 24.

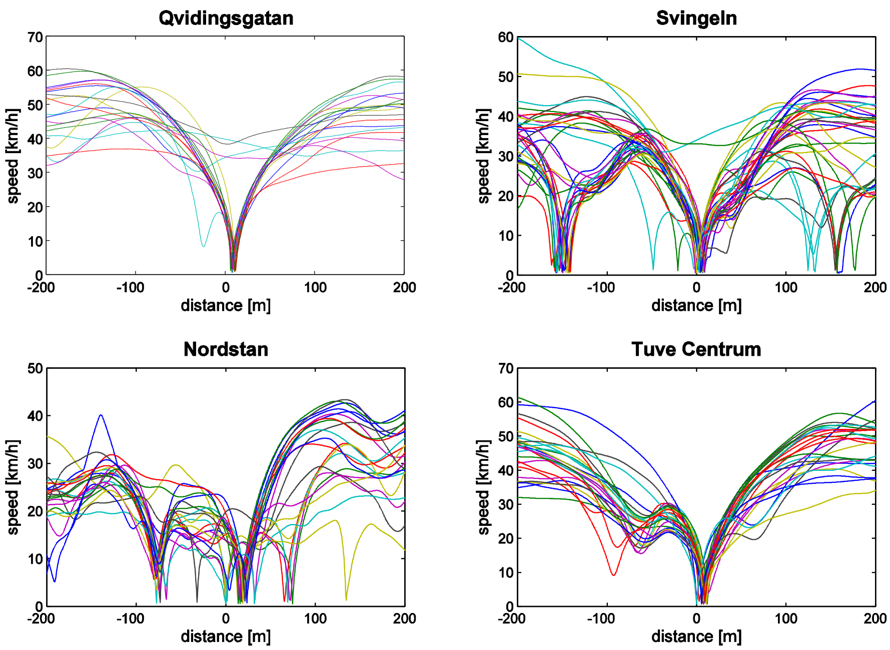

Figure 24: Velocity trajectories for the four specific bus stops.

As defined in Table 3, the pedal signals are thresholded and constructed into one signal of integer values from -2 to 2, representing full brake, brake, idle acceleration and full acceleration. In Figure 25 the driver behaviour at the different bus stops is presented. The driver's actions varies among bus stops, but since a bus stop is often characterized by a stop, there is a brake zone prior to the stop and an acceleration zone after the stop. The driver could be described as a regulator trying to hold a reference speed by accelerating and braking. The reference speed is defined by the driver from a lot of different parameters such as road speed limit, position, traffic and weather.

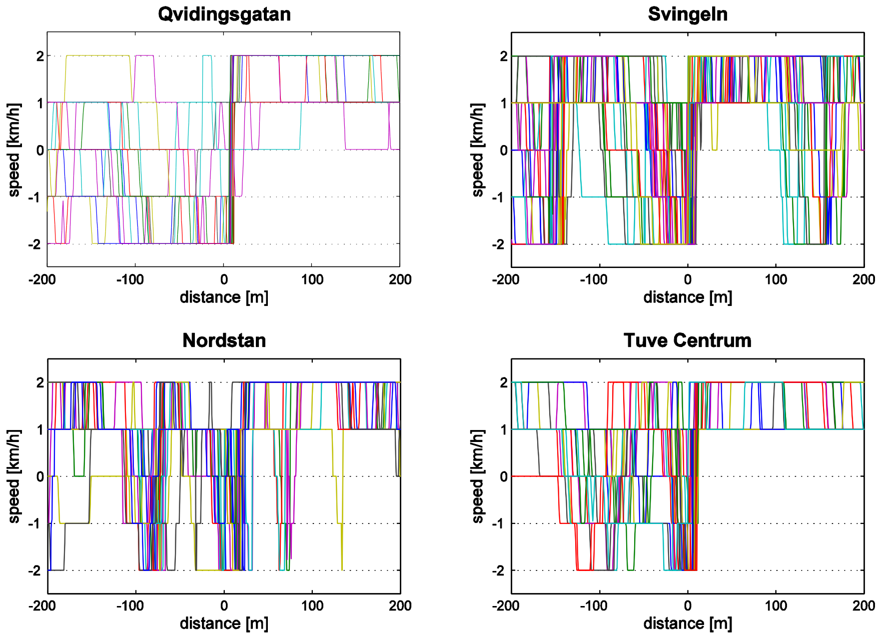

Figure 25: Driver behaviour signals for the four specific bus stops.

The brake behaviour is much more diverse than the acceleration behaviour. The brake dynamics are much faster in regards to the response to demanded speed, which leaves more for the driver to decide. The acceleration on the other hand is much more dependent on the limitations in the powertrain which in part explains why the acceleration zone is much more coherent. The bus stop number 27, Norumsgärde, explains this in a good way in Figure 26.

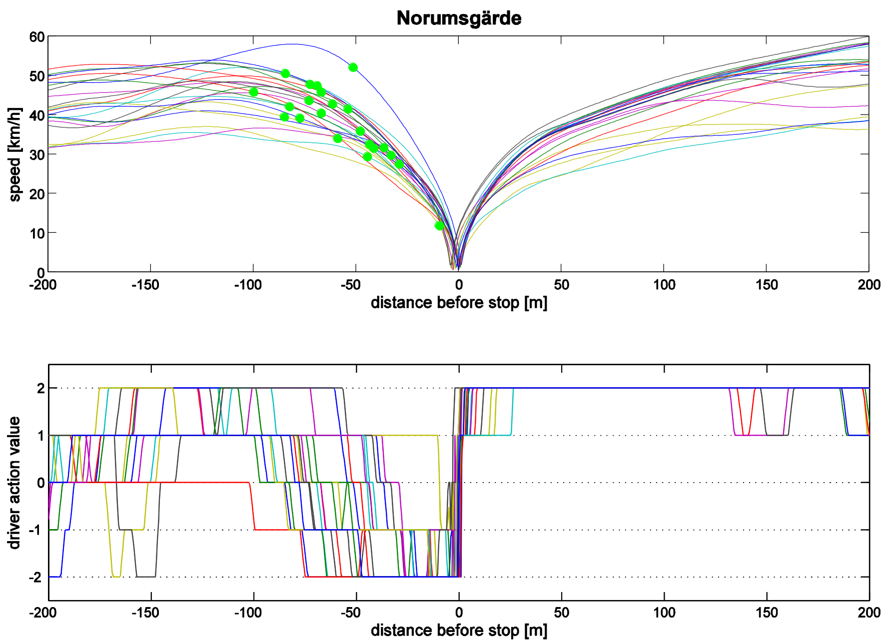

Figure 26: The upper figure shows the velocity trajectories for the Norumsgärde bus stop, along with the brake points marked with a green dot. The lower figure shows the driver action signal for the bus stops.

The acceleration zone is characterized by maximum acceleration for a long time, which can also be seen in the velocity figure. The brake point is defined as the position where the brake pedal is used for the first time, i.e. where the driver action signal is negative. Usually a higher velocity leads to an earlier brake point, but it is also influenced by the aggressiveness of the driver. The brake points are spread from 100 meters to 10 meters before the stop.

Another way to describe the driver actions for a bus stop is by producing histograms for the different driver action levels. Figure 27 shows the dominance of the acceleration pedal in comparison to the brake pedal. Only Qvidingsgatan has a somewhat spread probability for the different pedal levels.

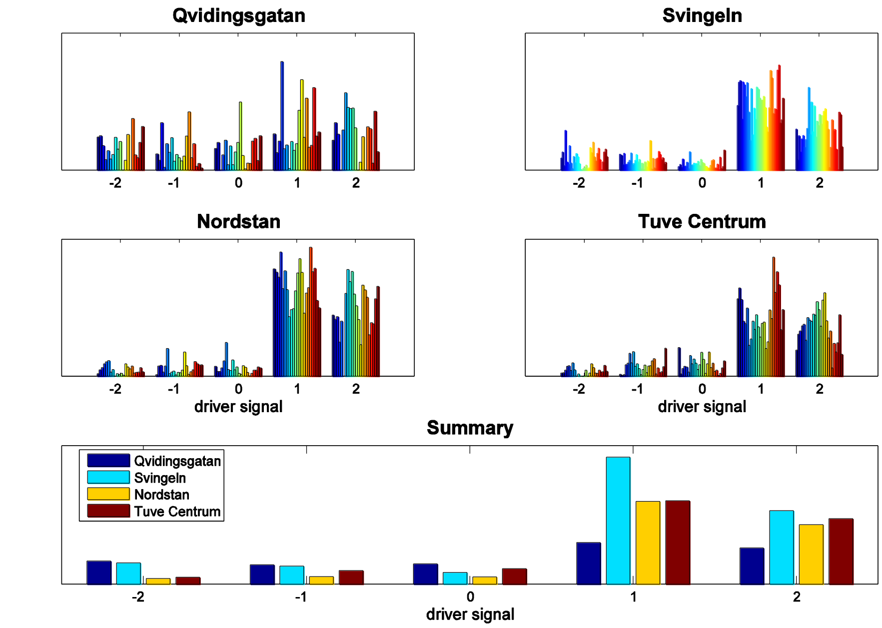

Figure 27: Histograms of the driver action signal for the four bus stops. Each bar represents a trajectory. The bottom figure shows a summary of the probability for each driver mode for the four stops.

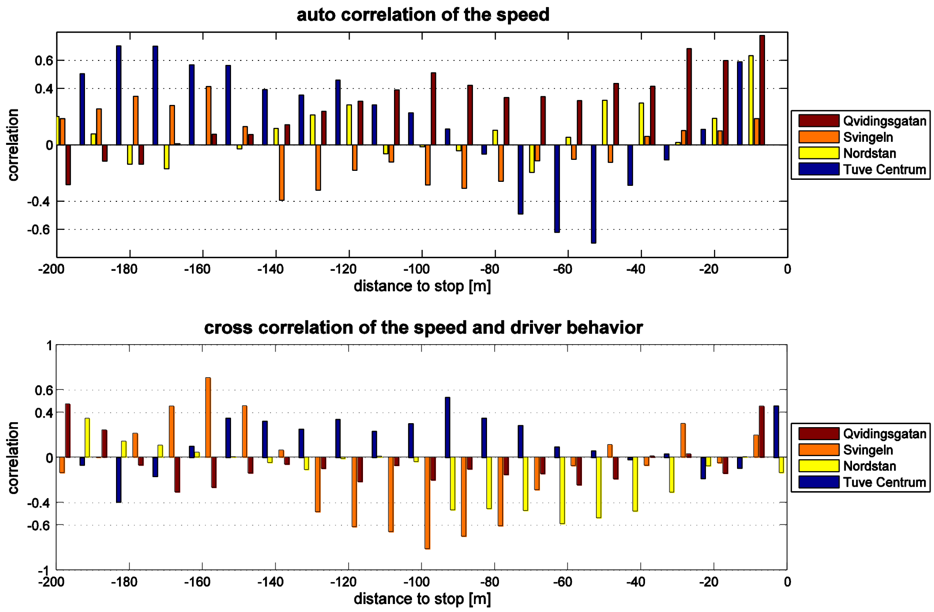

Figure 28: The upper figure shows the speed auto correlation between the stop (distance 0) and every 10 meters prior to the stop. The bottom figure shows the cross correlation between the speed and driver behaviour, calculated at every 10 meters prior to the stop.

The correlations for the bus stops do not give much useful information. The auto correlation for the speed is calculated between the stop and every 10 meters prior to the stop, giving a distance lag from -10 meters to -200 meters. The cross correlation between the speed and the driver behaviour signal is calculated at every 10 meters prior to the stop. The correlations are presented in Figure 28. As seen in the upper figure only Qvidingsgatan and Tuve Centrum has strong positive or negative auto correlation, except for the last few meters prior to the stop. For example, the trajectories without stop at Qvidingsgatan and Svingeln are not predicted by a strong negative auto correlation.

In the correlation section above there was a trend in the cross correlation between speed and driver behaviour that was thought to indicate the start of the brake zone. When the driver behaviour signal switched from high to low while the speed was still high, was said to lead to a change from positive correlation to negative correlation. But the velocity trajectories for the bus stops with the brake point marked as in Figure 26 have no such relation. According to Figure 28 the start of the brake zone for the Svingeln stop would be somewhere around 140 meters prior to the stop according to this theory, but in Figure 29 we see that most brake points are found in the last 50 meters prior to the stop.

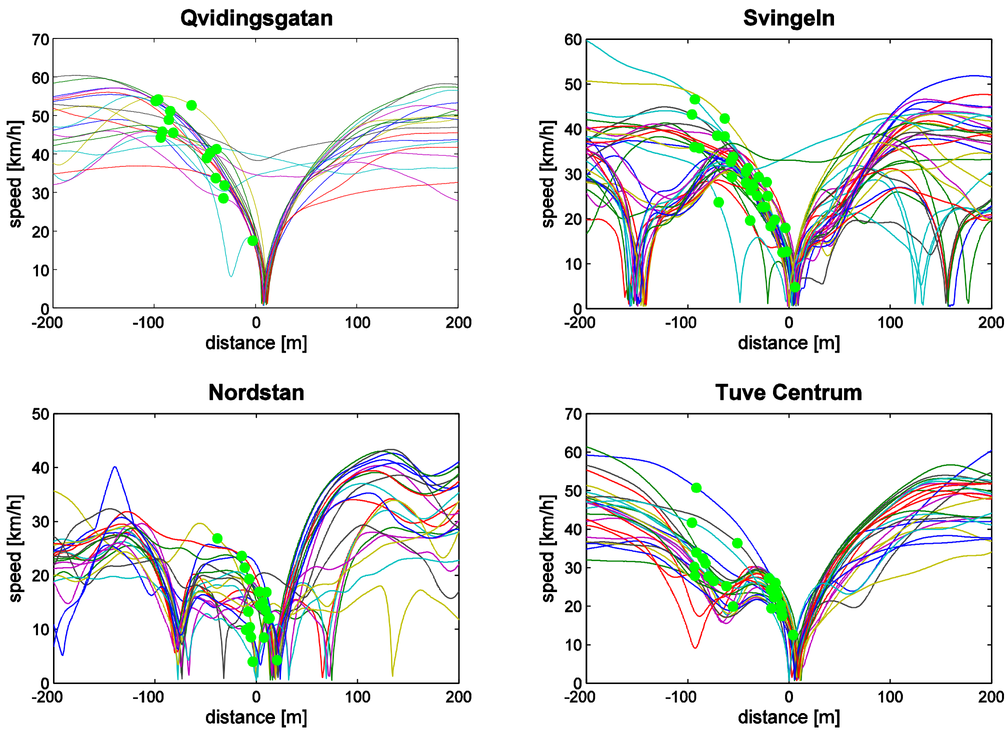

Figure 29: The four bus stops with the position where the brake pedal is first pushed prior to the stop marked with a green dot.

The histograms of the speed for the four stops are shown in Figure 30. Nordstan has much lower speeds, because of its multiple stops and heavy traffic. Qvidingsgatan on the contrary has a higher mean speed since the bus sometimes pass by without stopping. Svingeln and Tuve centrum has similar speed profiles, but Tuve centrum has a higher mean speed because of its location outside of the city centre with less traffic.

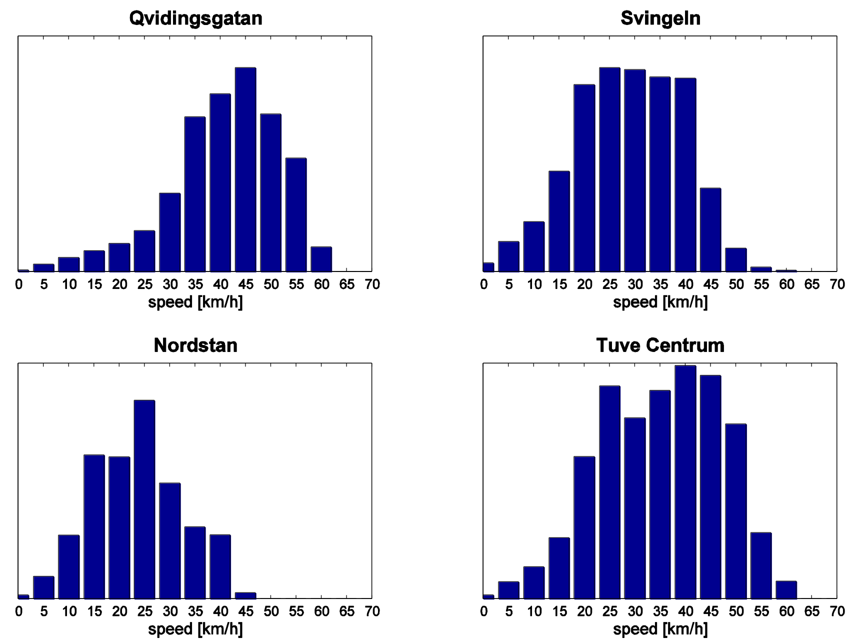

Figure 30: Histograms of the speed for the four bus stops.

3.3.2 Intersections

In total six intersections were logged along the route, but two of them were discarded because they do not affect the vehicle behaviour. Intersections where the bus has precedence were not logged at all. Figure 31 shows where the intersections are situated. As for bus stops, only data for the Tuve direction is presented.

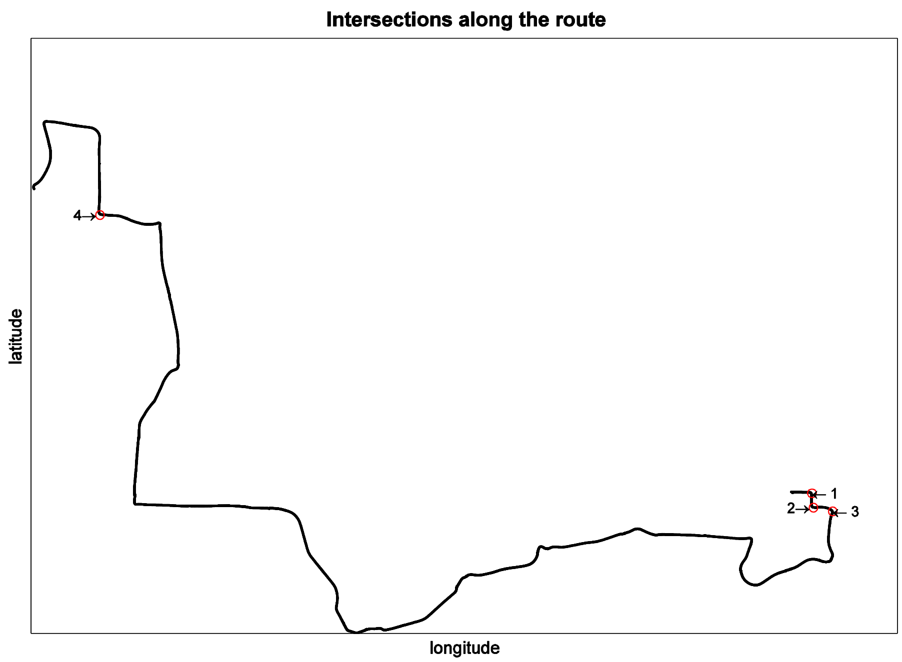

Figure 31: Location of the four intersections along the route.

An intersection is characterized by a decreased speed which sometimes leads to a stop. The speed is affected by both the traffic and physical limitations imposed by the maximum possible velocity in the bend of the road. Figure 32 shows the velocity trajectories for the four intersections. All four intersections have a distinct speed dip around the intersection, and only the third intersection leads to a stop, except for a few trajectories at the first intersection. Notably the third intersection is the only intersection where the bus leaves a smaller road for a larger road.

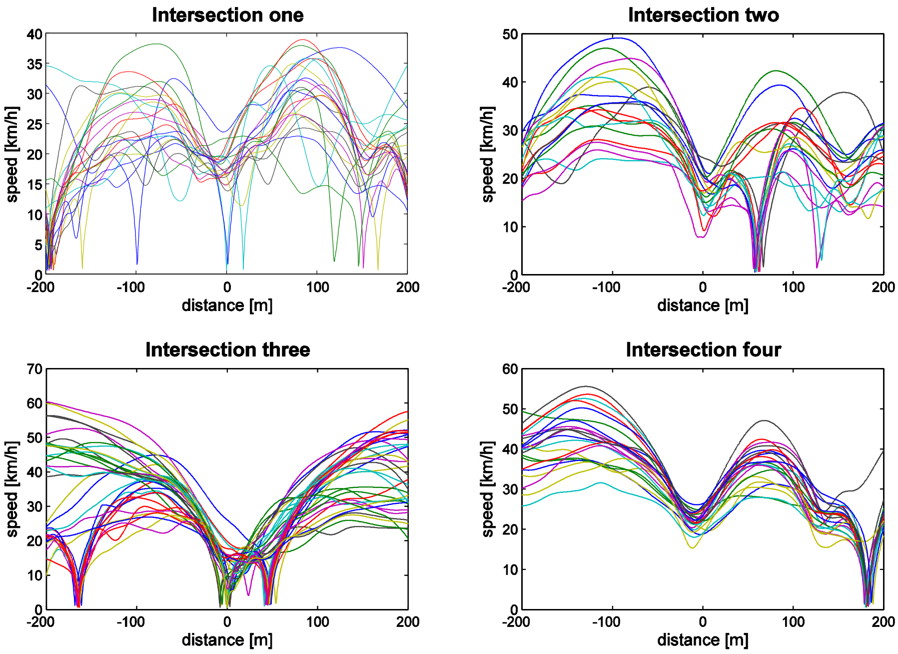

Figure 32: Velocity trajectories for the four intersections along the route.

As for the bus stop, the driver's actions are characterized by a brake zone prior to the intersection and an acceleration zone after the intersection. Since a complete stop is not necessary most of the time, the brake zone is not as distinct as for the bus stop. The third intersection leads to a halt most of the time, but the brake zone is still very short. This is explained by the intersection brake zone being placed in an uphill slope. The driver behaviour is shown in Figure 33.

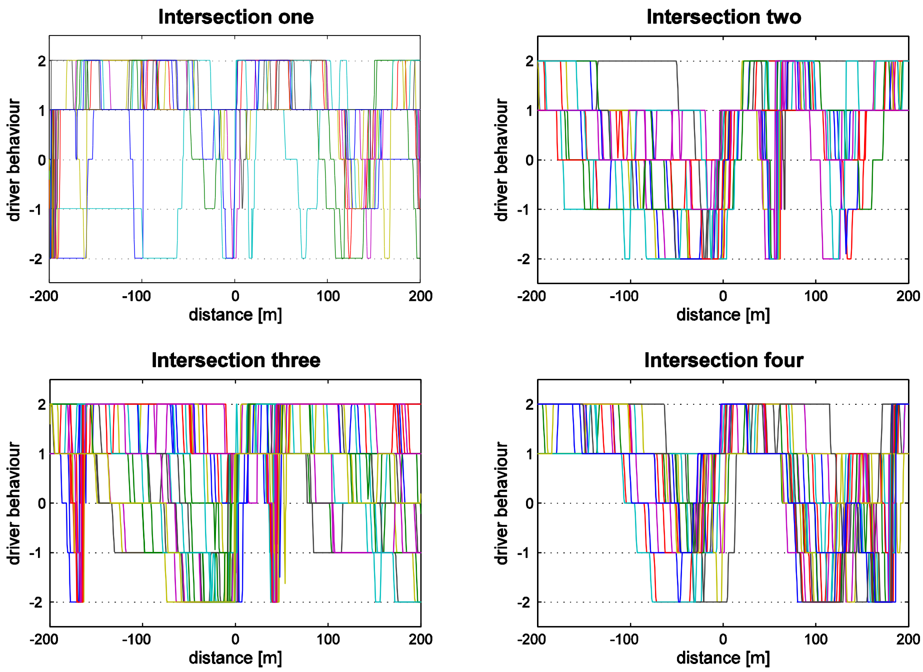

Figure 33: Driver behaviour signals for the four intersections.

Because of the lack of useful information from the correlation plots in the bus stop section, no correlation data is presented for the other positions. The speed histograms for the intersections are presented in Figure 34. Because of adjacent bus stops are found within the distance interval the mean speed is lower than otherwise. This is seen in Figure 35, where the intersection four's histogram is shown with two different distance intervals, such that the shorter interval of +/- 150 meters excludes the bus stop situated about 180 meters after the intersection.

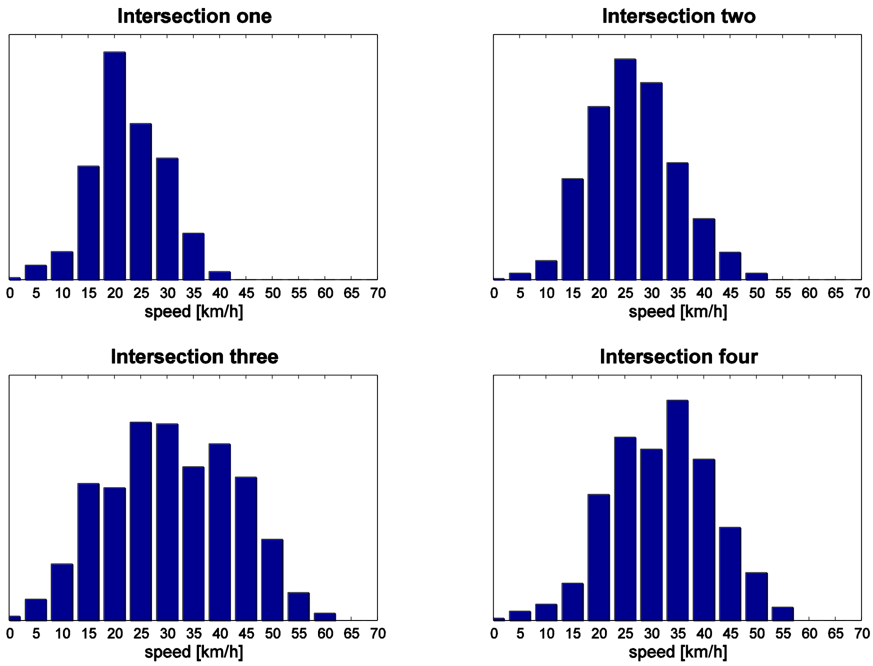

Figure 34: Speed histograms for the four intersections.

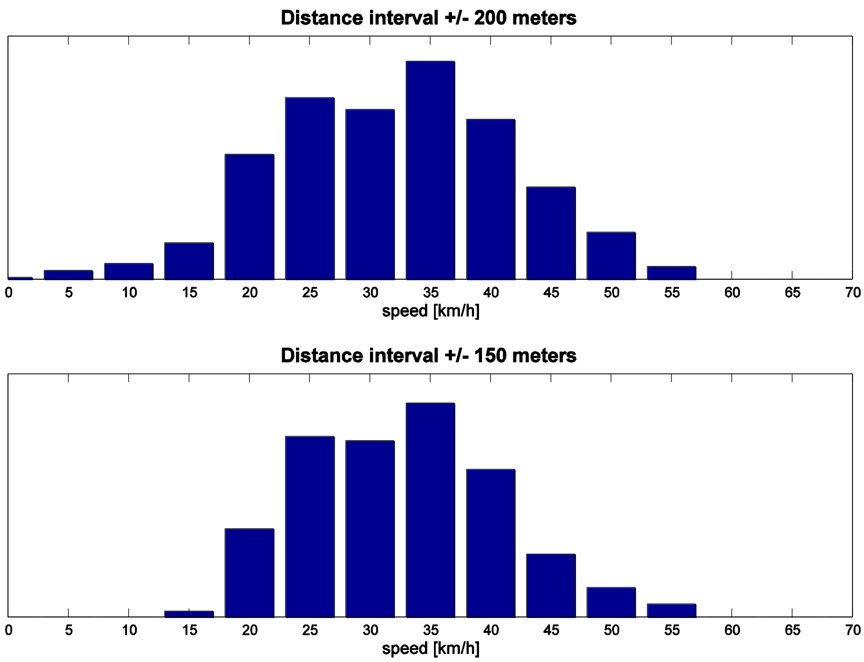

Figure 35: Speed histogram for the fourth intersection, where the shorter distance interval excludes the bus stop situated 180 meters after the intersection, thus giving a higher mean speed.

3.3.3 Intersection lights

There are in total 18 intersection lights along the route. Intersection lights can lead to a halt, but the position of the stop is dependent on the number of queuing vehicles at the intersection. Some lights are operated by the bus driver, leading to few stops. Four of the intersection lights were chosen to show the diversity among the stops, as seen in Figure 36.

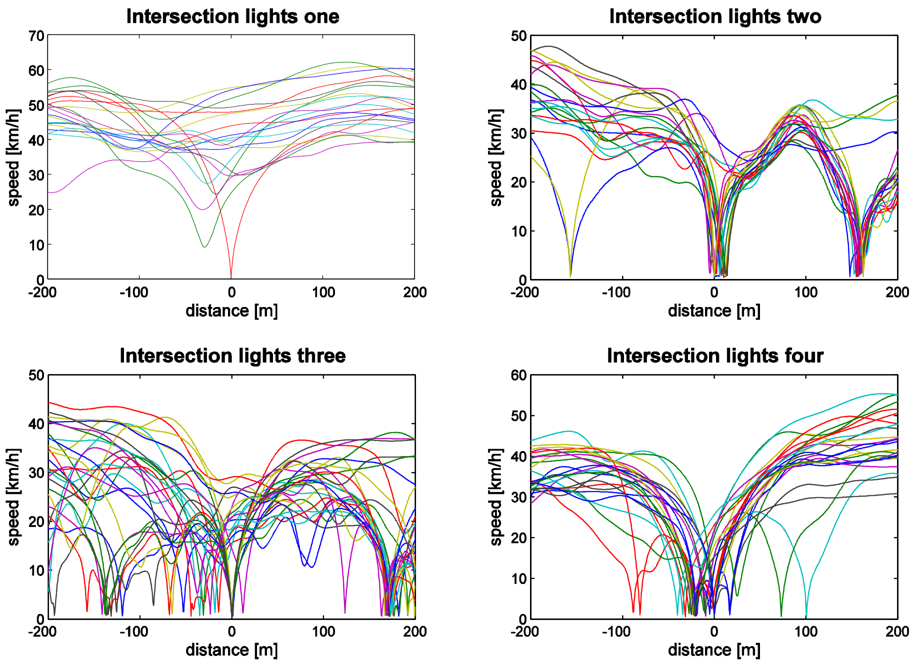

Figure 36: Velocity trajectories for four intersection lights along the route.

The first intersection light seldom leads to a stop, but sometimes to a slight brake in prior to the stop. The second intersection light leads to a halt most of the time and at the same position. Intersection lights three and four are complex intersections with heavy traffic, leading to stops at different positions due to queuing vehicles.

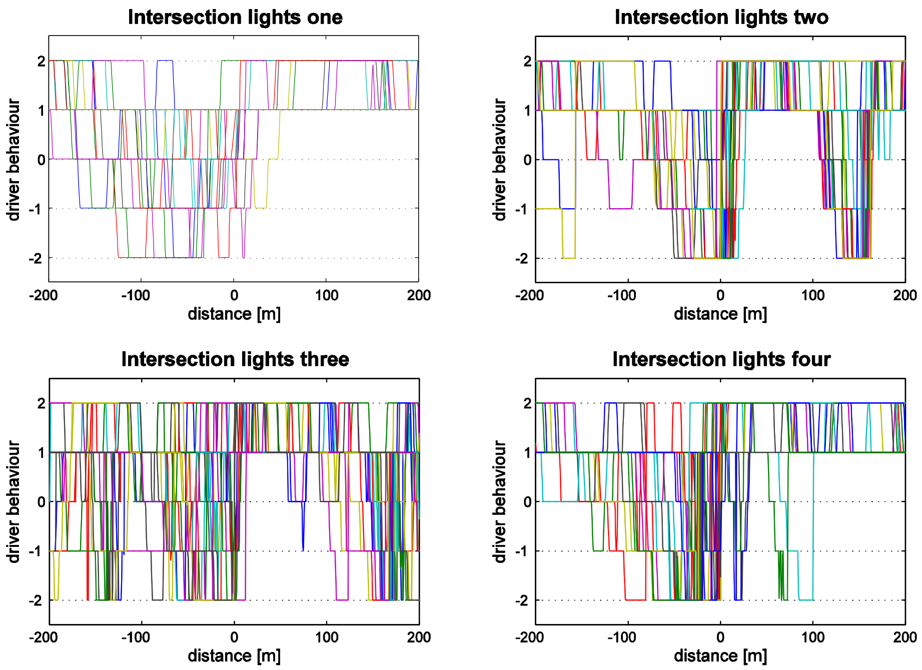

Figure 37: Driver behaviour signals for the four intersection lights.

The driver behaviour at the intersection lights are shown in Figure 37. Both the first and second intersection has an acceleration zone after the intersection and a brake zone with no full acceleration the last 50 meters before the intersection lights. Intersection three and four are more complex, since intersection sometimes does not lead to a halt and intersection four leads to a stop at different positions. Both intersection lights three and four have a distinct acceleration zone after the intersection lights, although the acceleration zone for intersection three is much shorter than the acceleration zone for intersection four.

The speed histograms for the intersection lights are seen in Figure 38, where intersection one sticks out as an exception, with its high velocities due to the few stops at the intersection lights.

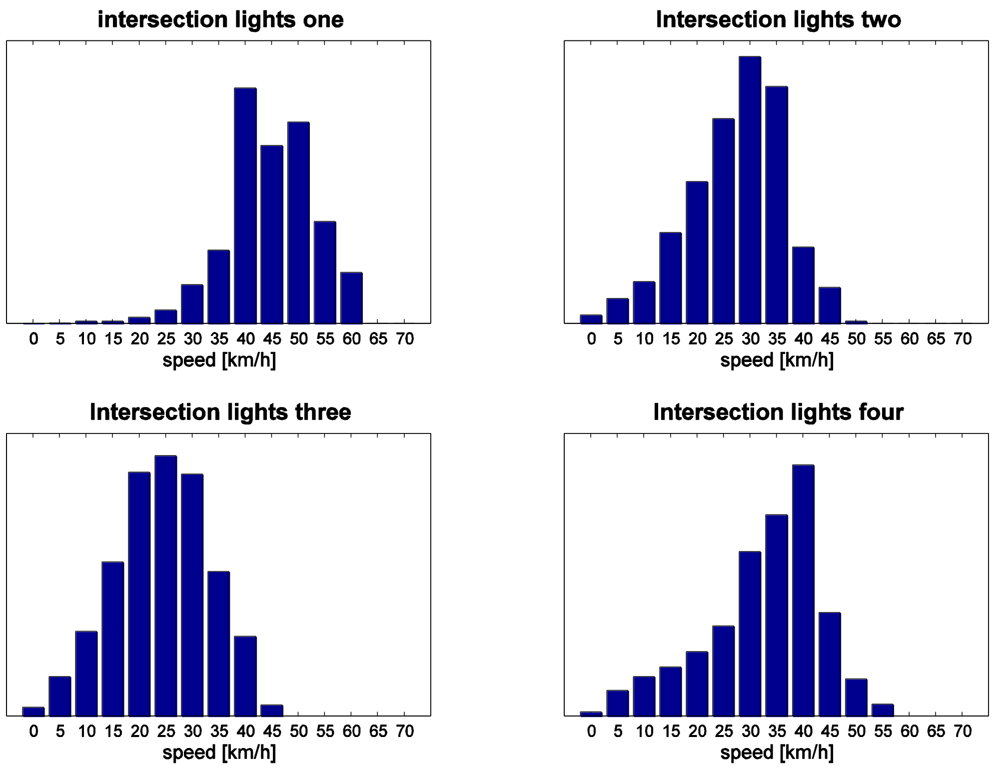

Figure 38: Speed histograms for the four intersection lights.

3.3.4 Crossing lights

There are only two crossing lights along the route, which are not in conjunction with an intersection. The velocity trajectories along with the driver behaviour signals are shown in Figure 39. There is a clear difference in the driver behaviour between the two stops, where no full acceleration takes place the last 50 meters prior to the second crossing light, but at the first crossing lights there is no clear brake zone. This might be because of the higher speed prior to the second crossing lights, which forces the driver to slow down to be aware of sudden crossing light signal changes. The speed histograms for the two crossing lights are shown in Figure 40, where the second crossing lights has a much higher mean speed that the first crossing.

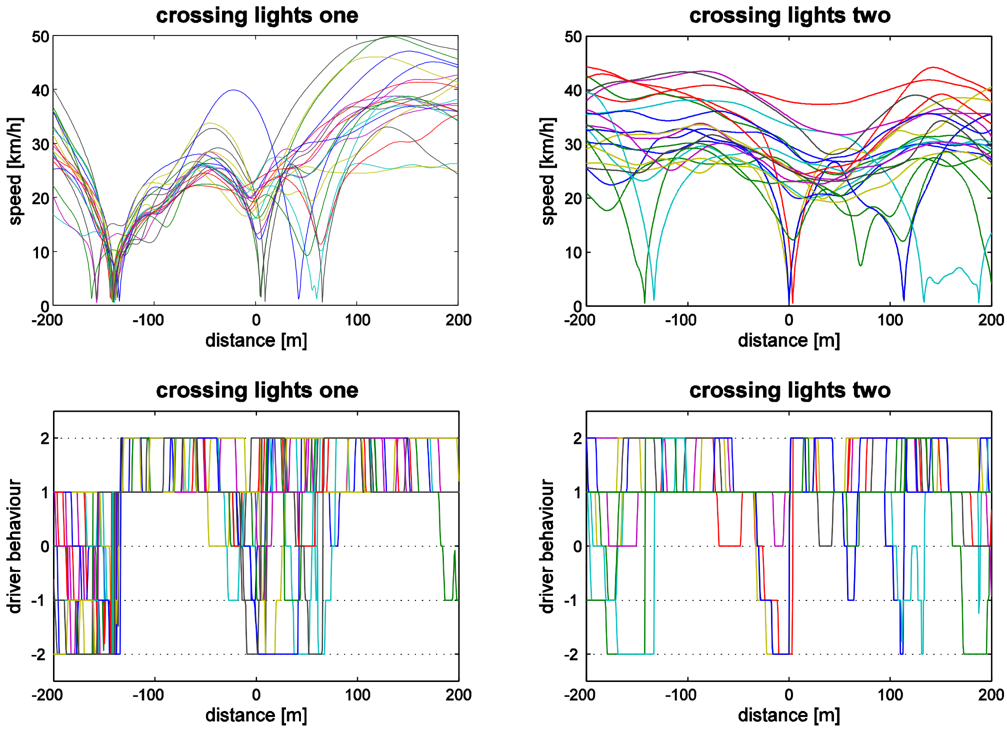

Figure 39: Velocity trajectories and driver behaviour signals for the two crossing lights along the route.

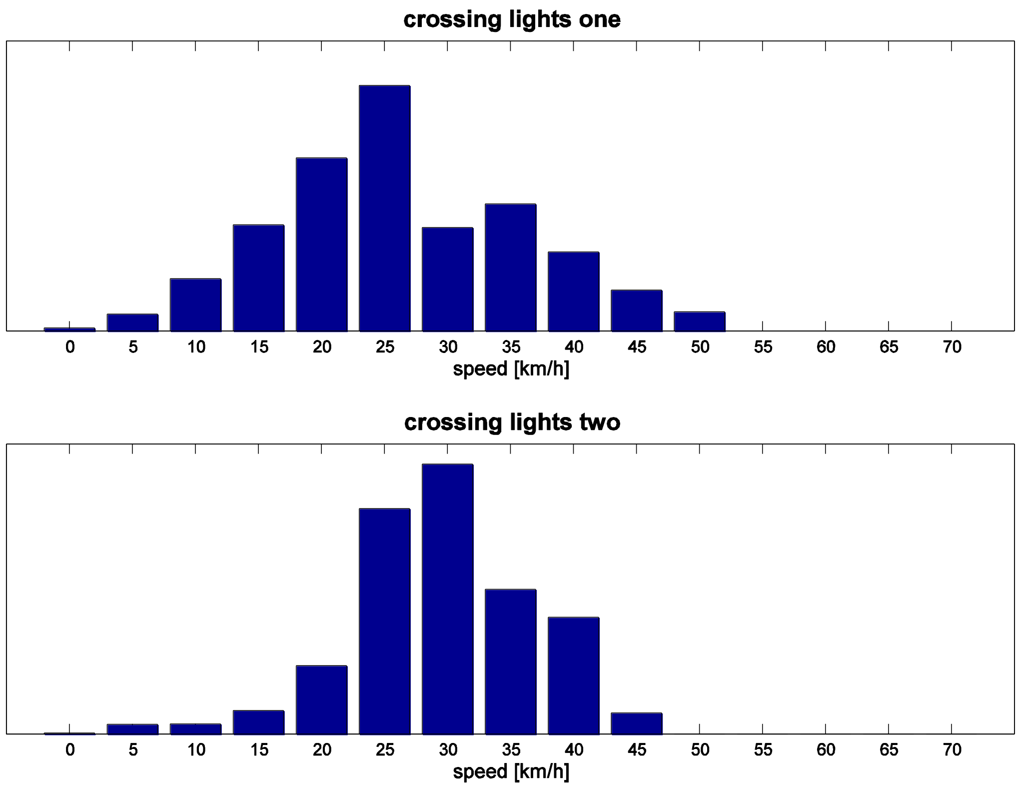

Figure 40: Speed histograms for the two crossing lights along the route.

3.3.5 Road bumps

There are in total eight road bumps along the route, where two of them are presented in the report. The whole idea with a road bump is to reduce the speed. The first road bump is situated with no other traffic situation close to it, whereas the second road bump is situated 75 meters prior to a bus stop. The velocity trajectories and the driver behaviour signals are shown in Figure 41.

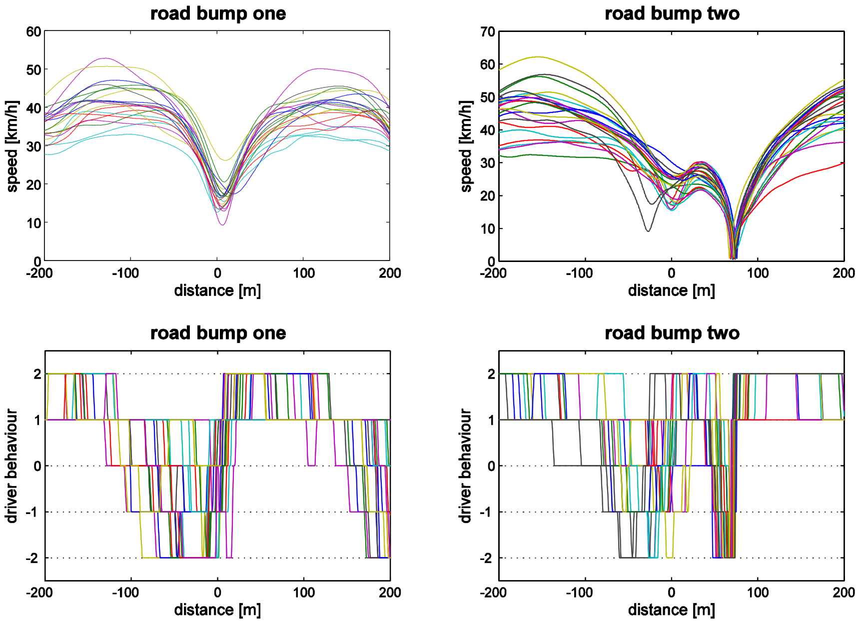

Figure 41: Velocity trajectories and driver behaviour signals for two road bumps along the route.

The first road bump is very typical where otherwise spread velocities are forced together into a narrow span of velocities at the road bump. The second road bump has the same characteristics, although its acceleration zone is very affected by the bus stop 75 meters after. The brake zone is longer for the road bumps than for example bus stops.

The histograms for the two stops are shown in Figure 42, along with the speed over the road bump. It is clear that the possible speed over the bump is in a narrow interval.

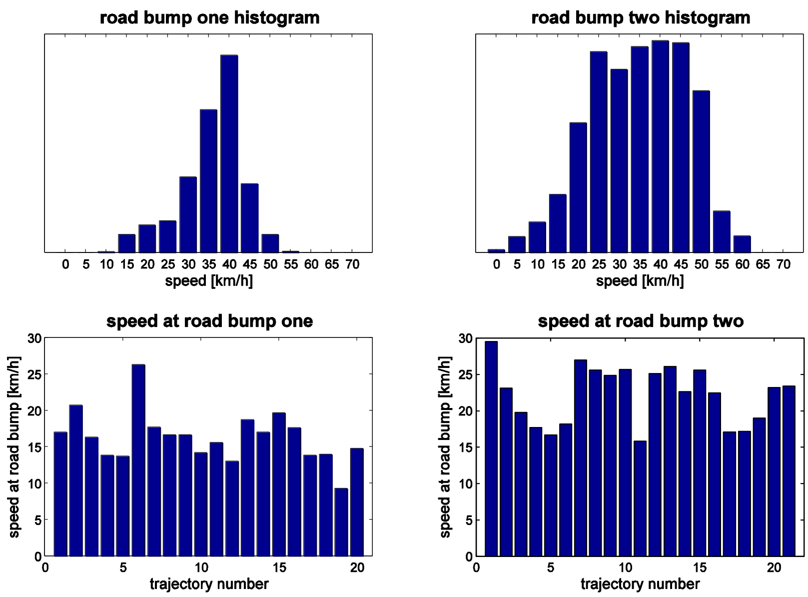

Figure 42: Histograms and the speed over the road bump for two road bumps along the route.

3.3.6 Roundabout

There are three roundabouts along the route; all situated in the second half of the route in the Tuve direction. Two of the roundabouts are presented in the report. The third roundabout is located just after a bus stop, with just a small street connecting to it, thus acting more as a road bump dominated by the bus stop. From a velocity point of view the roundabout is an effective way to lower the speed, much like the road bump. The roundabout sometimes leads to a halt, but not as often as for a typical intersection with traffic lights. The velocity trajectories and the driver behaviour signals are shown in Figure 43.

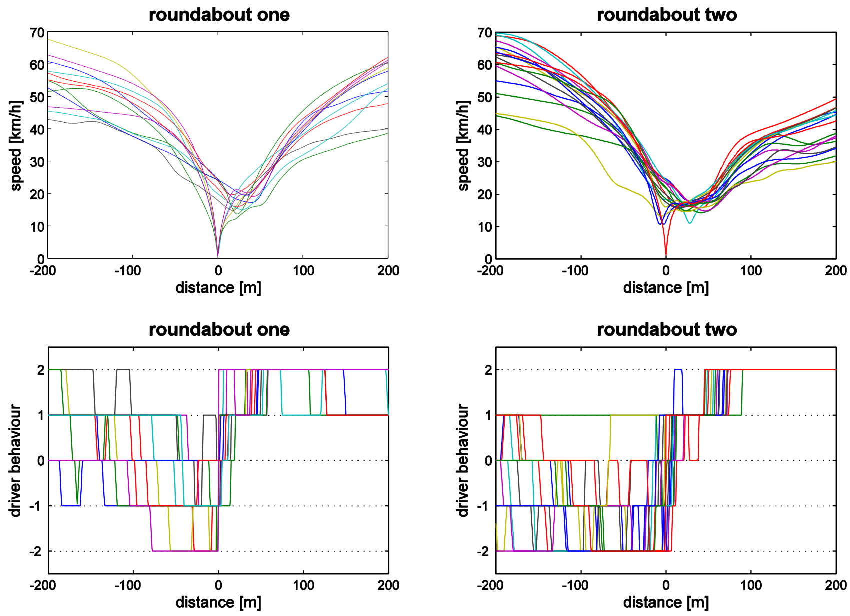

Figure 43: Velocity trajectories and driver behaviour signals for two roundabouts along the route.

The driver behaviour signals indicate a distinct acceleration zone after the roundabout, with a smoother brake-in zone prior to the roundabout. The softer acceleration after the second roundabout could be explained by the speed limit change from 70 km/h to 50 km/h after the roundabout, with an uphill slope. The slope explains the speed profile the last 100 meters at the second roundabout, where full acceleration result in a modest constant acceleration.

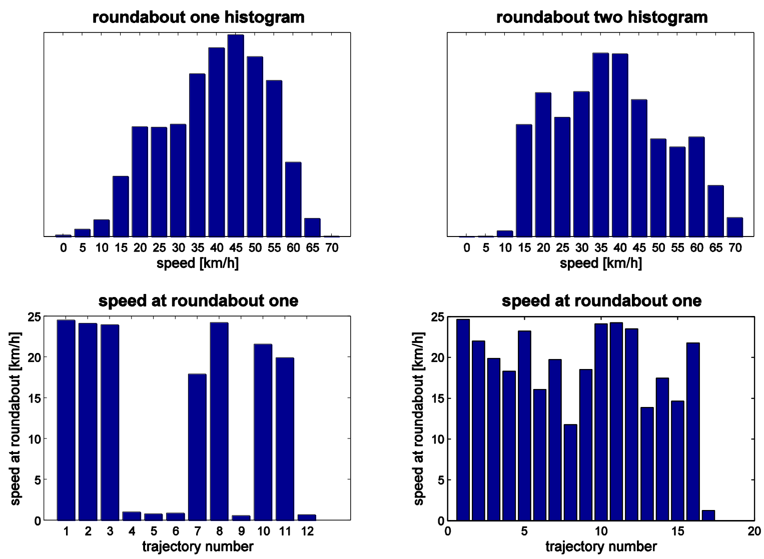

Figure 44: Histograms and the speed at the roundabout for the two roundabouts.

The velocity histograms along with the speeds at distance zero, i.e. the speed at the roundabout, are shown in Figure 44. The histogram for the second roundabout is mote spread due to the 50 km/h speed limit after the roundabout, as explained above. The speed at the roundabout is the range 20 km/h ± 5 km/h, except for the trajectories which leads to a stop.

3.3.7 Bridge

A bridge might not be a typical feature along any bus route, but there is actually one along the bus route in question, in the middle of Gothenburg halfway between Östra Sjukhuset and Tuve. It's about 500 meters long and it's possible to open it to let big ships through. The velocity trajectories and the driver behaviour signals for the bridge in both directions are presented in Figure 45.

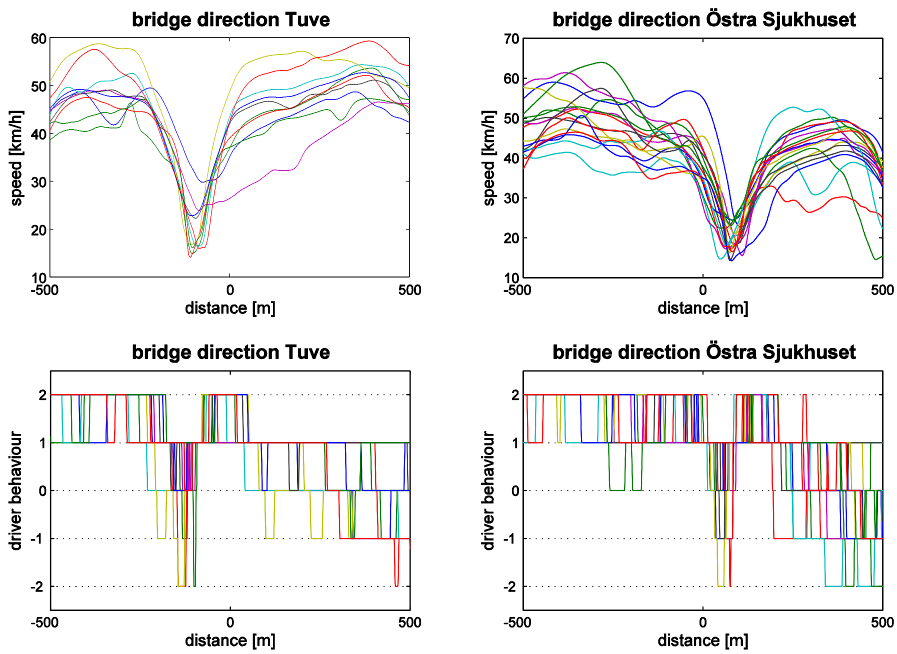

Figure 45: Velocity trajectories and driver behaviour signals for the bridge in both directions.

Since the bridge is possible to open, there is a risk that the bus has to stop for several minutes waiting for the bridge to close. There is a dip in velocity, but not a complete stop. It is also interesting to note that this dip is not prior to the stop in both of the direction, thus not related to signalling, bridge openings or traffic congestions. Actually the trajectories reach its low at the same position in both directions, 85 meter south of the middle of the bridge. Although it hasn't been verified, the reason to the decreased speed is probably road works or a temporary speed limit due to maintenance work. It is interesting to note how spread the velocities are at the bridge, with a range of about 25 km/h between the highest and lowest speed.

3.3.8 Other sections

Sections where no particular traffic situation affects the driver have been sorted according to the speed limit. There are only two different speed limits along the route, 50 km/h and 70 km/h respectively. Because of the many traffic situations along the route, there are few such sections. Except for the bridge which is presented above there is only one section with a speed limit of 50 km/h, located between the bus stops Vågmästareplatsen and Wieselgresplatsen. There are two sections with a speed limit of 70 km/h, located before and after the first roundabout presented above. The data for both directions are plotted in the same graphs, and the two sections with a

70 km/h speed limit are combined. The velocity trajectories and driver behaviour signals for the two speed limits are shown in Figure 46.

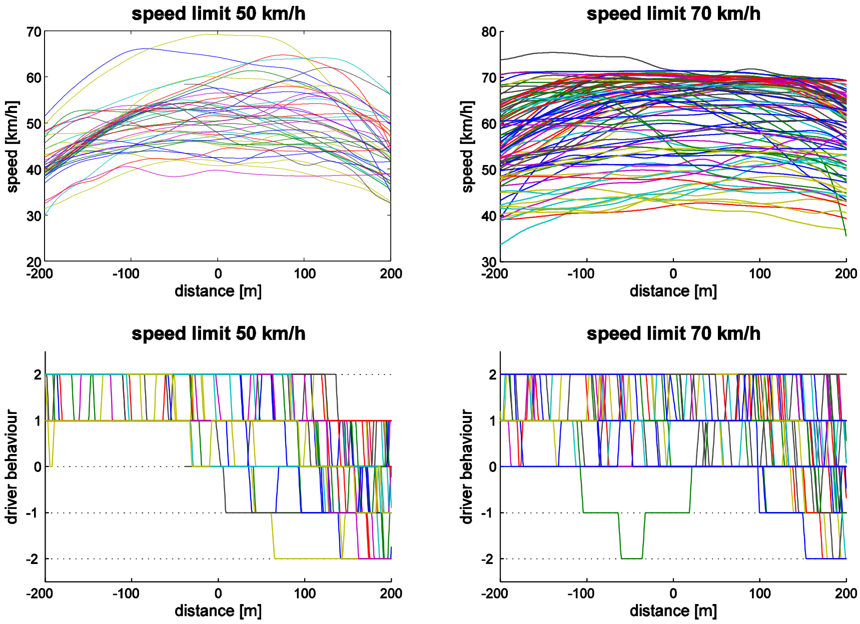

Figure 46: Velocity trajectories and driver behaviour signals for sections with speed limits of 50 km/h and 70 km/h respectively.

The driver behaviour signal is quite spread for the two speed limits, but there is an indication of a brake zone the last fifty meters of both sections. The signal toggles between no action, acceleration and full acceleration in the middle of the sections, indicating constant speed regulation. The range of velocities is large, about 20 km/h for the lower speed limit and 30 km/h for the higher speed limit.

Figure 47 shows the histograms for both the velocity trajectories and the driver behaviour signals for the two speed limits. The speed limit, or just under it, is the most frequent speed for both limits and the acceleration modes are dominating the driver behaviour.

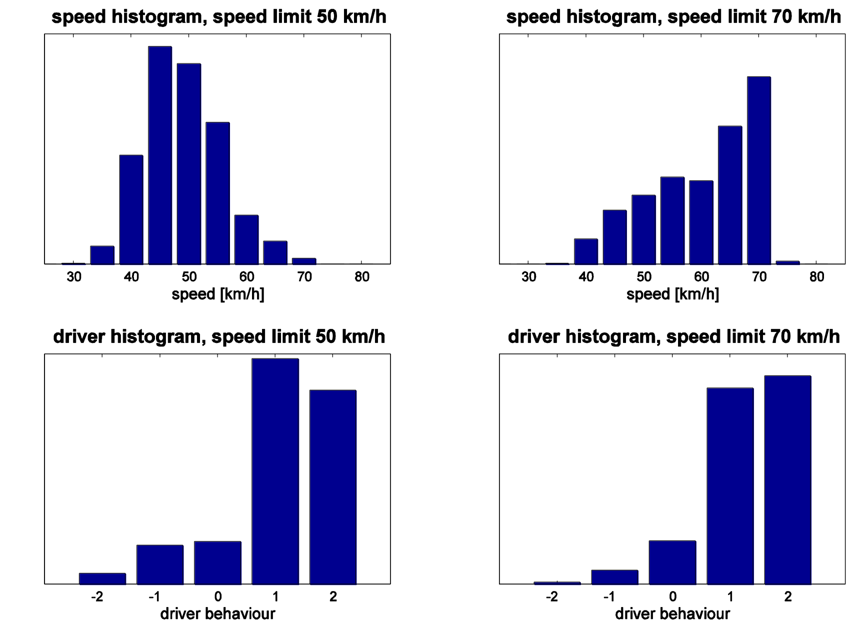

Figure 47: Histograms of the speed and the driver behaviour for the sections with speed limits 50 km/h and 70 km/h respectively.

4 Concluding remarks

By comprehensive data collection a statistical basis for a predictive model of the driver behaviour is provided, that could be used in a number of applications – such as gear selection control, hybrid vehicle energy strategy or smart torque management to reduce fuel consumption. The vehicle position and velocity along with the pedal positions gives a good description of the driver's propulsion of the vehicle along the given route and could be used as input to different driver modelling approaches such as markov processes.

The histograms of velocities and pedal actions at different traffic situations describe the nature of the positions in a good way, but correlation analysis failed to indicate relationships between signals. For example, the driver behaviour at bus stops could be modelled as one of a number of types, depending on where it is placed on the route, time of day, etcetera.

The data could also be the basis for a drive cycle generator application, where different traffic situation can be combined to simulate driver behaviour in different environments. A drive cycle is a sequence of torque and engine speed values, representing different kinds of driver behaviour or application areas. It can for example represent a city bus on a specific route in London or a long hauler truck in North America. The risk with deterministic drive cycles is that the vehicle is optimized for a specific cycle, which might not be representative enough. By developing stochastic drive cycle generators an infinite number of drive cycles could be generated and thus enable engineers to simulate different kinds of behaviour at the desk and reduce expensive time in vehicle.

The work could be further improved by more comprehensive data collection on additional bus routes and by using the vehicles' internal sensors and log data.

References

Andersson Henrik, Axelsson Johan (2009). *On-board collection and storage of road data and implementation of predictive transmission fuel-saving functions*. Chalmers tekniska högskola, Göteborg.

Anscombe, FJ (1973). Graphs in statistical analysis. *The American statistican*, vol 27, no 1 (Feb 1973), pp 17-21.

Bruce, Maria (2005). *Model based powertrain control. Damping of oscillations in a heavy duty vehicle.* Chalmers tekniska högskola, Göteborg.

Du, Ping (2006). *Thesis work: Development of a modular vehicle dynamics measurement system.* Chalmers tekniska högskola, Göteborg.

Fägerlind Gustav, Hermansson Mattias, Rönnberg Johan (2007). *Bachelor thesis work: Skattning av en vägsträckas fart- och höjdprofil med GPS och accelerometer.* Chalmers tekniska högskola, Göteborg

Gillespie, Thomas (1992). *Fundamentals of vehicle dynamics.* SAE, Warrendale, USA.

Johannesson, Lars (2006). *On energy management strategies for hybrid electrical vehicles.* Chalmers tekniska högskola, Göteborg.

Johannesson, Lars (2009). *Predictive control of hybrid electric vehicles on prescribed routes*, Chalmers tekniska högskola, Göteborg

Pettersson Stefan, Egardt Bo, Bruzelius Fredrik (2006). *Powertrain modelling & control.* Gröna Bilen, Göteborg.

Rice, J A (1995). *Mathematical statistics and data analysis.* Duxberry Press.

Sciarretta Antonio, Guzzella Lino (2007). *Control of hybrid electrical vehicles.* IEEE Control systems, Vol 27, no 2, p 60.