Proceedings of the 11th International IEEE
Conference on Intelligent Transportation Systems
Beijing, China, October 12-15, 2008

Speed and acceleration distributions at a traffic signal analyzed from microscopic real and simulated data

Francesco Viti, Serge P. Hoogendoorn, Henk J. van Zuylen, Isabel R. Wilmink, Bart van Arem

Abstract— **Modeling realistic driving behavior at signalized intersections is crucial for many applications, for instance to determine the traffic signal performance, to assess the effect of different control strategies, or to estimate traffic emissions. In these applications, often microscopic models are used to simulate the trajectory of each vehicle. Despite the possibility to model vehicles with great detail and at fractions of a second, speed, acceleration and deceleration characteristics are determined by parameters that are rarely calibrated using real data, and default parameters are often chosen. This is because collecting real vehicle trajectories near traffic signals is a challenging task.**

This paper presents a method to collect such dataset using image processing techniques. This methodology allows one to obtain vehicle trajectories near a signal control, and to measure individual vehicles speeds and accelerations at a microscopic level. We focus on the analysis the empirical distributions of speeds and accelerations observed with this unique dataset near and up to a few meters upstream of the stop-sign. We compared these distributions with the results of repeated simulations of two microscopic software programs, using default parameters. Some inconsistencies were found with this comparison, which suggests that the two analyzed microscopic simulation programs run with default parameters do not provide realistic results for this type of road sections.

I. Introduction

Driving behavior observed on the road changes depending on the type of road section. At signalized intersections, drivers are taught to moderate their speed, and to comply with the priority rules set by the traffic light. Therefore, vehicles stop and queue up during the red phase, and they are allowed to pass the junction only during green or amber phases. During these operations, vehicles driving modes vary considerably. The way they decelerate, stop and move in the queue and then accelerate to leave the intersection depends on each individual driver's behavior. Aggressive drivers may show stronger accelerations, others may decide to start decelerating very early, once detected that there is no right of way at the signal. During the amber phase aggressive drivers may accelerate to clear the intersection faster, while risk-averse drivers may decide to decelerate, or even break hard to avoid passing the stop-sign after the start of the red phase. Even when vehicles do not stop at the traffic control, some may decide to reduce their speed, for instance when different traffic streams share the same green phase.

Therefore, individual vehicle trajectories are found very variable at these sections. In applications that require good estimates of vehicle driving modes, like when estimating concentration levels of pollutants, it is fundamental to provide realistic estimates of these trajectories, thus realistic speeds and speed variations. For instance, these measures are used in microscopic instantaneous emission models like VT-Micro [1], MOVES [2], or EnViVer [3]. Emission models are often combined with traffic flow models, e.g. microscopic simulation programs, to obtain estimates of emission levels, for instance when assessing ITS measures or new planning strategies. However, these models simulate individual vehicles trajectories using simplified driving behavior, characterized by a limited number of parameters. Although microscopic programs can simulate vehicle movements at fractions of a second and account for many different behavioral aspects (vehicle acceleration and deceleration capabilities, car-following behavior, degree of aggressiveness, etc.), speed and speed variations are often found quite unrealistic. Moreover, driving behavior parameters are often kept with their default values, or chosen using rules-of-thumb or they are taken from other studies and other types of road sections, typically on motorways.

In this paper we present a method to collect individual trajectory data near a traffic light, through which we can analyze speed and acceleration behavior of drivers and compare it with data simulated with microscopic programs. We collect microscopic data using image processing techniques, which allow one to obtain vehicles trajectories at very high definition and to measure, under some degree of reliability, actual speeds and speed variations.

This paper is structured as follows. Section II describes how drivers' behavior is modeled at signals. Section III gives an overview of the collected dataset, and the

Manuscript received September 1st, 2008. This project is co-financed by the Delft University of Technology, the TNO Institute of Applied Sciences and the TRAIL Research School.

Francesco Viti is with the Delft University of Technology, section Transport and Planning, and with the Katholieke University of Leuven, Transport and Logistics section. Mailto: Stevinweg 1, P.O.Box 5048, 2600 GA Delft, Phone NL +31 15 2786909 Fax +31 15 2783179, f.viti@tudelft.nl, Phone BE +32 16 321673 Fax +32 16 322986, Francesco.viti@cib.kuleuven.be

Serge P. Hoogendoorn and Henk J. van Zuylen are with the Delft University of Technology, section Transport and Planning. S.p.hoogendoorn@tudelft.nl, H.j.vanzuylen@tudelft.nl.

Isabel R. Wilmink is with the TNO Institute of Applied Sciences. Isabel.wilmink@tno.nl

Bart van Arem is with the TNO Institute of Applied Sciences and with the Technical University of Twente. Bart.vanarem@tno.nl

description of how it has been processed. Section IV analyzes the microscopic data and investigates the drivers' interactions with the traffic control operations, looking especially at their speed and speed variations upstream and when passing the traffic control. Finally Section V concludes this paper and presents future developments in this research.

II. Driving behavior at traffic controls

Traffic flow models are developed at macroscopic, mesoscopic and microscopic levels. Macroscopic models provide direct relationships between macroscopic variables (e.g. average speeds, flows, densities), thus they can capture traffic state variations only to certain extents. In mesoscopic models traffic flow and performance variables are instead represented through probability distributions. Examples of widely used mesoscopic simulation programs are Dynasmart [4], DynaMIT [5] and the Cellular Automata model [6]. Like the macroscopic models, these models have also a limited application range and they simplify the behavior of vehicles by using specific distributions to model individual trajectories. Therefore these two modeling levels are not suited for estimating characteristics like individual vehicle speeds, accelerations and decelerations.

Microscopic models have been developed with the main aim of simulating the movement of vehicles on the roads at the individual level, i.e. each vehicle movement is determined through the simulated network infrastructure at fractions of a second, and interactions with other vehicles are simulated on the basis of a few driving rules (e.g. car-following, overtaking etc.), modeled through mathematical relationships. Each vehicle is thus characterized by a speed profile that depends on a limited number of parameters describing individual drivers, and the interaction with the other vehicles and the road infrastructure. However, not all aspects of these models are completely realistic. Different driving behaviors and vehicle types are aggregated into a limited number of classes, the simulated vehicles enter the system according to controlled probability distributions, which often do not resemble the real ones, and the movements and interactions are determined via analytical relationships, which do not always catch heterogeneity of human behavior.

Most of driving behavior variability modeled in microscopic programs is determined by the inter-driving variability in free driving (for instance by assigning different desired speeds to each vehicle) and the car-following model in constrained driving (through different preferred safety distances, or reaction times, etc.), as for example it is done in the widely used microscopic software programs Paramics [7], Aimsun [8], VISSIM [9], MITSIM [10] and INTEGRATION [11]. All these parameters need to be chosen carefully in order to obtain realistic estimates.

The calibration and validation procedures in a microscopic model are in this sense of crucial importance. Traditionally, calibration of microscopic models is done using macroscopic traffic performance measures, e.g. average flow, speed, density etc. However, many parameters in these programs, i.e. desired safety distance, reaction time, degree of aggressiveness, etc., should be measured directly, since their effect is hardly captured with aggregated data. Despite the importance of calibration and validation procedures, many practical studies neglect these modeling phases, and default parameters in commercial software programs are often kept by the users. These parameters may fit the context in which they have been calibrated, but they may change considerably in other road sections. AIMSUN for example has been calibrated using observed flows and speeds of the Barcelona's Ring Roads and main accesses to the city, while in VISSIM the default values have been set after calibration on German highways.

The use of macroscopic data for calibrating microscopic simulation models has been justified in the past with the extreme difficulty of collecting microscopic trajectory data. Exception is a study that used GPS signals in California to match the variability of driving patterns observed on urban corridors with the ones simulated with Paramics [12]. The study for the first time discusses the importance of calibrating a microsimulation program to estimate emission rates in areas where flow interruptions are determinant. However, since GPS positions were recorded for only a subset of vehicles in the platoons comparison could be made only using mean travel times.

The use of macroscopic data for calibrating microscopic simulation models has been justified in the past with the extreme difficulty of collecting microscopic trajectory data. Thanks to image processing techniques behavioral aspects that play fundamental role especially at microscopic levels could be analyzed on motorways ([13], [14]). Many studies nowadays use the NGSIM dataset, which is a large dataset of individual vehicle trajectories collected in several freeway and arterial roads in the US, obtained from processing video images automatically [13]. In other studies the principle trajectory data was used to calibrate and compare car-following models using snapshots taken from a helicopter [14]. Another research group used the same technique but from high-rise buildings [15]. Despite the large dataset collected with these projects, no study has yet focused on collecting microscopic data specifically at signalized intersections. This is done in the research described in this paper and presented in the next section.

III. Data collection and process

In this project we focus exclusively on the longitudinal driving behavior, to reduce the degree of complexity of the problem; thus we focus on free driving and car-following behavior. Therefore vehicle trajectories must be collected in an area where lane changing is rarely observed or prohibited. This way, vehicle speed and speed variations can be studies in a limited set of traffic situations, and related in a more direct way to the parameters available in microscopic software programs.

To obtain a dataset with these properties we need to: 1) choose a study area to analyze the (longitudinal) driving behavior near signalized intersections, 2) collect (microscopic) trajectory data and 3) process this data to obtain actual individual trajectories.

We describe in detail how we have undertaken these phases in the following of this document.

A. Data requirements and choice of the study area

To study exclusively the longitudinal behavior near and passing a signalized intersection we need a location and a dataset with the following requirements:

- the road section should not have any significant change in the road geometry (e.g., in lane width, speed limit, etc.), and
- it should mainly contain longitudinal movements. An appropriate study area has to be selected, possibly where overtaking is not allowed;
- the dataset should contain a sufficient number of (consecutive) individual trajectories, or driving cycles;
- it should contain trajectories that can be traced back to a sufficient number of meters upstream the stop-sign to study the passing and deceleration behavior,
- it should contain trajectories that can be traced forward up to a sufficient number of meters after the stop signal to evaluate the acceleration behavior;
- the recorded time of individual vehicle positions should be as short as possible, at least the same updating time of microscopic software programs (e.g., 1 second);
- it should contain a small number of measurement errors and missing data to be able to derive accurate estimates of individual speeds and accelerations;
- the dataset should contain a complete picture of the whole study area at each time step.

A place that fits the above requirements has been identified in Rotterdam, the Netherlands. The Euromast Space tower is a high rise building, where at 100m from the ground is located a restaurant with a view to the Maas River. Aside of the tower, a traffic light controls the flows coming from two two-lane roads that merge into one two-lane section. The traffic light is vehicle-actuated, thus green light is assigned alternatively to the two carriageways dynamically. Maximum and minimum green times are set to respectively 6 and 30 seconds. Overall the study area speed limit is set to 50km/h and a speed control camera is installed aside of the road. The two carriageways run in parallel and have small difference in terms of geometry. Overall the study area lane changing is prohibited. Vehicles can be seen for more than 140m upstream of the stop sign and for more than 60m downstream. The view from the tower is quite limited downstream the traffic signal, but it is sufficient to analyze the acceleration behavior of vehicles.

High resolution pictures have been taken from the tower at 15Hz frequency with a digital camera, as shown in Fig. 1 and for nearly 2 hours (between 8:00AM and 10:00AM).

Fig. 1: View of the traffic control

The choice of this study area has inevitably a number of limitations for the image processing method. The first is that pictures had to be taken from behind a safety glass. Shadows sometimes affect the light balance of some picture areas. A second problem is that brightness changes very often and this is a major source of error especially to obtain the background image upon which vehicle positions are automatically detected, as it will be described in the next section. We explain in the filtering process how we solved these issues.

B. Data conversion process

To detect automatically vehicles as moving objects we use image processing technique: a background image is generated through superposition of a number of frames. The background image is obtained by finding the color that is most likely to be observed at each pixel, i.e. the color of the road pavement on the carriageways. In this way, moving objects are removed. The process was done using a code programmed in MATLAB®. A limited number of frames can be loaded contemporarily to obtain the background image so it is important to choose 1) the time window for which the background image is calculated and 2) the number of frames used for this calculation. Differently from the highway system, vehicles may in fact stand still for many frames near the stop-line. In saturated green phases this can occur for more than 40s. To avoid errors in the background image we loaded a number of frames randomly from 2 minutes intervals. We could not use longer time periods because ambient light changes and the color of the background image could not match sufficiently the color of the loaded frames, resulting in a very noisy tracking. Images are later compared one by one with the background images to find the consistent color differences, and through a clustering method, moving objects are tracked.

The method described enables one to get a complete picture of the area. On the other hand many vehicles sometimes tend to be grouped into one cluster. This occurs often in two cases: 1) when they are standing in queue, thus they wait in too short distances to be detected by the tracking method and 2) when driving in parallel on two adjacent lanes. To overcome this second limitation we process the four lanes separately. After this process we have obtained the geographical vehicle positions at each time frame.

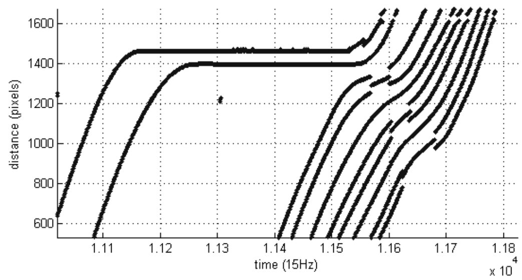

Fig. 2: Examples of vehicles positions recorded with image processing

Fig. 2 shows an example of the processed data in a space-time diagram for the rightmost lane. Individual vehicle trajectories can be clearly identified from visual inspection. As one can also notice, especially when vehicles are stopped, their trajectories tend to be missing or misplaced. This is mainly due to the clustering of vehicles that are too close to be identified by the clustering algorithm. Before obtaining individual vehicle trajectories we therefore need to clean the data from these errors.

C. Data cleaning

Looking back at Fig. 2, some isolated points can be clearly identified. These are due to, e.g., errors when calculating the background image, or because of shadows, etc. Some of these errors can be easily identified by filtering the dataset using the area size. If clusters are too small or too large they are removed from the dataset, as they cannot be vehicles but they are clearly errors.

Moreover, as one can see, not all trajectories can be identified completely. Especially when vehicles queue up, sometimes they are clustered together. The tracking of individual trajectories is done in two steps. In the first part of the process we look at each moving object at one time frame and find the next point in time and space. A trajectory is therefore identified if a point is found assuming a minimum (zero) and a maximum speed at which a vehicle may have moved forward. We define a maximum speed v=80km/h. To allow small errors at low speeds due to the object clustering we consider valid also small negative speeds. These errors will be later corrected by smoothing the data.

Detecting trajectories by looking at one time step interval may result in finding many partial trajectories. For example, there can be only one time step where a vehicle has not been detected correctly (due to, e.g., aggregation with another vehicle), and later this error does not occur. This single error results into two separated trajectories of actually the same vehicle. We thus estimate the speed during the last second before losing the trajectory and the speed at the starting of the next trajectory. By using quadratic interpolation we can thus find, by comparing partial trajectories in pairs, partial trajectories belonging to the same vehicle. Fig. 3 gives a visual example of the result of this procedure.

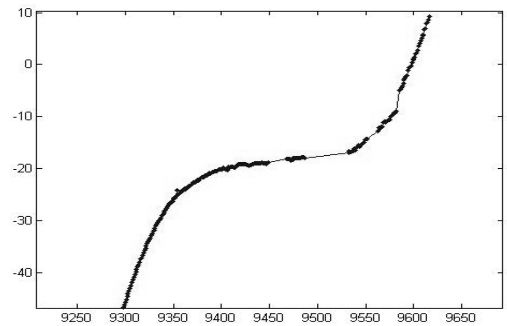

Figure 3: Example of a few partial trajectories of the same vehicle connected via quadratic interpolation

The processed trajectories still contain errors due to the clustering algorithm. If these errors look small on a space-time diagram, they become sensibly larger when looking at first and second derivatives, i.e. at speed and acceleration profiles. Smoothing the data is therefore needed to obtain realistic speeds and acceleration profiles. However, traditional smoothing algorithms (like linear or quadratic fitting) may far-fetch the real behavior of the vehicle. Some actual variations of speeds and acceleration may be smoothed out, simplifying the real vehicle behavior. We choose to apply Locally Weighted Regression, which has been demonstrated to be an effective method to smooth trajectories reliably [16]. Locally Weighted Regression is a method that corrects any data point by looking at its neighboring points. The principle is based on Robust Fitting, i.e. Least Square estimation where for each data point a weight is assigned. In [16] a tri-cubic function is recommended as weight function, thus the information level of each point in the neighborhood is simply based on its (temporal) distance.

By applying clustering technique to process the data we obtained information also on the area of each cluster, and this estimated area is determinant of part of the error contained in the positions recorded at each time frame (i.e., the position of the centroid of these areas). We can therefore use this information as measure of the reliability of each point in the neighborhood. Moreover, by applying image transformation and rotation, the recorded movements of each vehicle are mainly on the x-axis, thus, assuming that the vehicle cannot sway considerably, the position on the y-axis can be also used in as measure of reliability.

These measures are calculated with the following functions:

$$Ra = \frac{\left|A(t) - \tilde{A}(t_0)\right|}{2 \cdot \sigma_A}, Ry = \frac{\left|y(t) - \tilde{y}(t_0)\right|}{2 \cdot \sigma_y}$$

Where $A(t)$ and $y(t)$ are respectively the area and the y-position of the point recorded at time t in the neighborhood of t_0, $\tilde{A}(t_0)$ and $\tilde{y}(t_0)$ are respectively the average value of the area and of the y-position in the neighborhood, and σ_A and σ_y are respectively the standard deviation of the area and of the y-position in the neighborhood The weights are thus computed by $W_0 \cdot \sqrt{R_A \cdot R_y}$, where W_0 is the value of calculated with simply the tri-cubic function. Figure 4 gives an example of computed weight function.

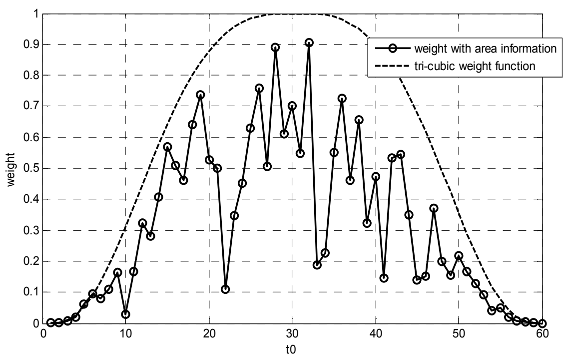

Figure 4: Example of weight function used to correct the position of a vehicle and the tri-cubic weight function used in [16].

Figure 5 shows how the correction of the vehicle positions results into smoother speeds. In another paper we have tested this algorithm on floating car data, showing that it successfully corrects highly corrupted data [17].

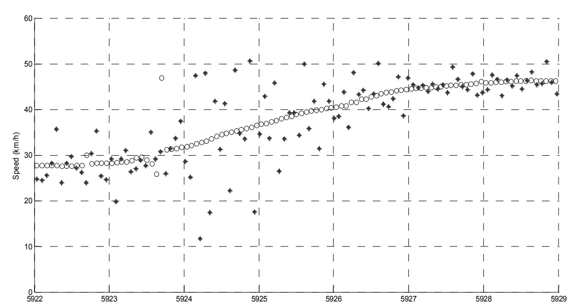

Figure 5: smoothed speeds calculated after the correction of each point using Extended Locally Weighted Regression

IV. SPEED AND ACCELERATION DISTRIBUTIONS

The above described dataset has been used to analyze the speed and acceleration profiles of each vehicle. In this section we analyze the observed distributions of these driving behavior characteristics.

Figure 6 shows a contour plot of the speed distribution of vehicles from the real data in comparison with the result of microscopic simulations done with the programs VISSIM and AIMSUN under the same flow rates. In VISSIM vehicles tend to decelerate gradually while approaching the signal, and they start doing that already over 100m from the stop-line. On the contrary, in AIMSUN the distribution of speeds is quite concentrated on high speeds until 40m from the stop-line. This means that in this program they decelerate within shorter times. The collected data shows that the reality lies in between. As one can also clearly see, the distribution of speeds seems much wider both far upstream and near the stop-line with respect to the ones estimated with the microscopic models. This implies that the latter models will underestimate the effects of traffic, for example, if used to estimate traffic emissions where speed variations are determinant.

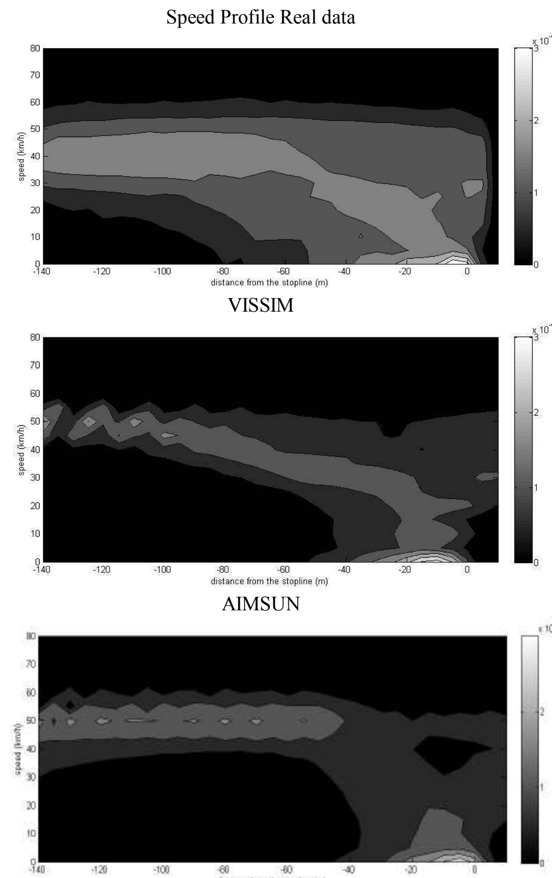

Figure 6: Speed profiles for different road sections before and right after the stop-line

It seems that the largest inconsistency between reality and microscopic models can be attributed to the passing vehicles, as these programs do not consider intra-driving variability. This is also confirmed by Figure 7, where we subdivided the dataset into vehicles stopping at the traffic light and vehicle passing without a full stop. As one can see from the speed distributions of passing vehicles, much higher speeds are estimated with the microscopic models.

The peaks in the acceleration distributions are in the same place (at 0 m/s^2), but the distributions are much narrower in the simulations than in reality. There are clearly more accelerations below -0.5m/s^2 and above 0.5m/s^2 in the real data than in microscopic models.

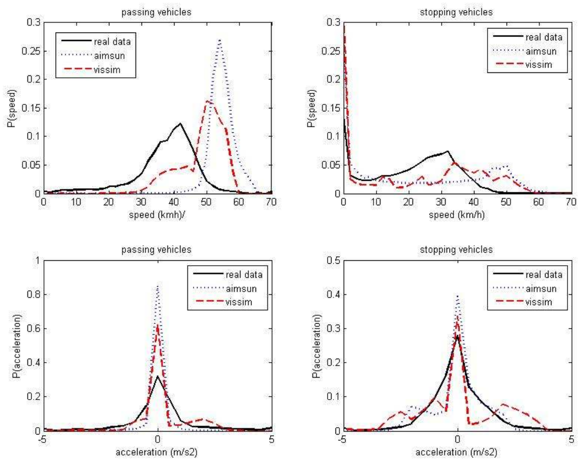

FIGURE 7 Probability density functions of speeds (top) and accelerations (bottom) for passing and stopped vehicles

Figure 5 shows the distribution of speeds at different sections, i.e. 30m before the stop-line, near the stop-line and 20m after the stop-line for both passing vehicles and stopped vehicles. As it was observed with the speed profiles, the speeds in AIMSUN are considerably higher before the signal, also when vehicles have to make a full stop. Looking at the distribution of stopped vehicles after the signal, is seems that in both simulation models vehicles reach high speeds in less time, pointing at harder accelerations from zero speed.

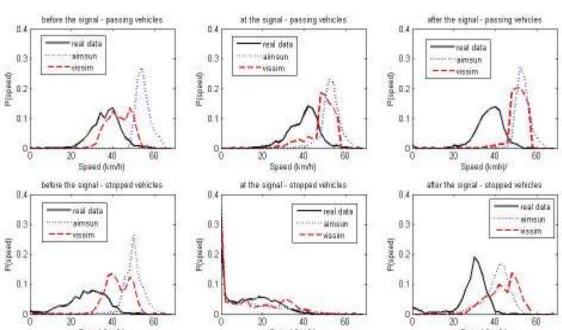

FIGURE 5 Speed distribution observed before the signal (left), at the signal (middle column) and after the signal (right) for passing (top) and stopped vehicles (bottom).

V. Conclusions and Future Research

This paper has presented the first results of a study of driving behavior at signalized intersection using microscopic real data. Substantial differences were found between the driving patterns of the real world data and the data from the simulations, using the default settings, if looking at speed and acceleration distributions. Calibration of these models using microscopic data like the one presented in this paper is therefore necessary and will be done in further research.

The dataset presented in this paper will be very valuable to study the variability of drivers' behavior at these signals, to gain insight into which factors determine their actual trajectories and to estimate emissions from traffic. These applications will be done and presented in future works.

Acknowledgments

The authors would like to thank Minwei Li, Janneke van Baalen, Jie Li for the contributions made within this project.

References

[1] Rakha H., Ahn K., Trani A., 2004. Development of VT-Micro model for estimating hot stabilized light duty vehicle and truck emissions. *Transportation Research Part D* **9**(1): p. 49-74.

[2] Environmental Protection Agency (EPA), 2008. MOVES (Motor Vehicle Emission Simulator). http://www.epa.gov/otaq/ngm.htm accessed on May 2008.

[3] TNO. EnViVer, 2008. http://www.tno.nl/content.cfm?context=markten&content=case&laag1=196&item_id=648 access May 2008.

[4] Mahmassani H. S. (1997). DYNASMART-X System, Part 1: Overview, Structure, Functions, Key Components. Sessions on Dynamic Traffic Assignment II and III, INFORMS Annual Meeting, San Diego, May 5, 1997.

[5] Qi Y., 1997. A Simulation Laboratory for Evaluation of Dynamic Traffic Management Systems. PhD Thesis, Massachusetts Institute of Technology, Boston.

[6] Nagel K. and Schreckenberg M., J. Phys. I (France) , 2:2221, 1992.

[7] http://www.paramics-online.com/.

[8] Barceló, J., Ferrer, J.L., Montero, L., Perarnau, J. (2004). Microscopic Simulation With Aimsun For The Assessment Of Incident Management Strategies. European Transport Conference, ETC2004, Strasbourg, France.

[9] PTV (2003). VISSIM v. 3.70 user manual, Karlsruhe, Germany.

[10] MITSIMLab (2000). MITSIM http://web.mit.edu/its/papers/MITSIM2001

[11] Van Aerde, M., and Associates (2005). INTEGRATION Release 2.30 for Windows, User's Guides. Kingston, Canada, M.Van Aerde and Associates.

[12] Nesamani K.S., Chu L., McNally M. G., Jayakrishnan R. (2005). Estimation of Vehicular Emissions by Capturing Traffic Variations. Internal report UCI-ITS-AS-WP-05-2, Univ. of California, Irvine.

[13] Brockfeld E., Kuehne R., Liebe C., Wagner P. (2005). Calibration of Microscopic Car-Following Models using NGSIM Data. Paper presented at the workshop Traffic Modeling: Simulation Models: from the Labs to the Trenches, September 18-21, Sedona, Arizona.

[14] Ossen, S. J. L., Hoogendoorn, S.P. (2007). Calibrating car-following models using microscopic trajectory data. Internal Report, Delft University of Technology, T&P 2006.010.

[15] Kovvali, V., Alexiadis V. and Zhang L. (2007). Video-based vehicle trajectory data collection. Presented at the 86th Meeting of the Transportation Research Board, 2007, Washington D.C.

[16] Toledo T., Koutsopoulos H.N., Ahmed K.I. (2007). Estimation of Vehicle Trajectories with Locally Weighted Regression. In Proceedings of the 86th TRB Annual Meeting, CD-ROM

[17] Viti F., Li J., Wilmink I.R., van Zuylen H.J., Hoogendoorn S.P., van Arem B (2009). Using image processing to analyze Driving behavior at traffic signals. Submitted to the 88th TRB Annual Meeting, January 11-15 Washington D.C.