Trajectory Data and Flow Characteristics of Mixed Traffic

Venkatesan Kanagaraj*
Post-Doctoral Fellow
Transportation Research Institute
Technion – Israel Institute of Technology
Haifa 32000, Israel
Email: vkanagaraj.iitm@gmail.com

Gowri Asaithambi
Assistant Professor
Department of Civil Engineering
National Institute of Technology Karnataka,
Surathkal, Mangalore, India
Email: gowri_iitm@yahoo.co.in

Tomer Toledo
Associate Professor
Faculty of Civil and Environmental Engineering
Transportation Research Institute
Technion – Israel Institute of Technology
Haifa 32000, Israel
Email: toledo@technion.ac.il

and

Tzu-Chang Lee
Assistant Professor
Department of Urban Planning
National Cheng Kung University
Tainan 701, Taiwan
Email: jtclee@mail.ncku.edu.tw

*Corresponding author
Word Count: 6978 (4478 words +2500 (8 Figures and 2 Tables))
ABSTRACT

Driving behavior (e.g. car following and lane changing) models describe the longitudinal and lateral movements of vehicles in the traffic stream. Calibration and validation of these models require detailed vehicle trajectory data. To the best of our knowledge, trajectory data are not publicly available in the context of traffic in developing world cities, which is characterized by a heterogeneous mix of vehicle types and by a lack of lane discipline.

This paper reports on the various steps of an effort to create a dataset of vehicle trajectory data in mixed traffic and on the analysis of these data. The data was collected using video photography in an urban midblock road section in Chennai, India. The trajectory data was extracted from the video sequences using specialized software and processed using the locally weighted regression method in order to reduce measurement errors and to obtain continuous position, speed and acceleration functions. The traffic flow characteristics of these trajectories such as speed, acceleration and deceleration, and longitudinal spacing are investigated. The results show statistically significant differences among the various vehicle types in travel speeds, accelerations, distance keeping and selection of lateral positions on the roadway. The results also indicate that vehicles, in particular motorcycles, move substantially also in the lateral direction, and that in a substantial fraction of the observations drivers are not strictly following their leader. The results suggest directions for development of driving behavior model for mixed traffic streams.
INTRODUCTION

The study of vehicle-to-vehicle interactions is necessary for the understanding of traffic flow and safety problem. Driving behavior models describe drivers’ maneuvering of their vehicle movement in different traffic situations. The core behavioral models, namely car following/acceleration and lane changing models have been studied for several decades. A large number of theories regarding the functional forms governing these behaviors have been proposed. Comprehensive reviews of these models can be found, for example, in Brackstone and McDonald (1) and Toledo (2). However, far less research has focused on calibration and validation of driving behavior models using observational data. One reason for this gap in the literature is the difficulty to obtain the required data, which consists of time-space trajectories of the vehicles in a section of road. From these data, time series of the variables that are used in driving models (e.g. positions, speeds, and accelerations of the various vehicles, relative speeds, time and space headways) are extracted. The validity of driving behavior models depends on the availability and quality of these data. FHWA’S Next Generation Simulation project (3) collected and shared several datasets of vehicle trajectories on expressway and urban arterials in the US. These have been used extensively to calibrate and validate driving behavior models (4, 5, 6, 7, 8 among others).

In the context of mixed traffic modeling, calibration and validation of driving behavior have mostly been based on macroscopic flow characteristics, such as flows, speeds and densities (9,10,11,12). This approach limits the level of detail that can be captured in the developed models. Few studies utilized trajectory data, but these are often small samples collected for a specific study. For example, Kanagaraj et al. (13) collected the trajectories of the subject and lead and lag vehicles in a merging situation. Sangole and Patil (14) selectively collected trajectories for the involved vehicles in group gap acceptance behavior at uncontrolled intersection. To the best of our knowledge, vehicle trajectory data are not publicly available in the context of mixed traffic. This may, to a large extent, be due to the difficulty and high cost involved in data collection and extraction, and the technical complexities associated with having a wide mix of vehicles types with varying physical dimensions and dynamics characteristics (speed and acceleration capabilities) and non-lanes based movements (15).
This paper reports on the various steps of an effort to create a dataset of vehicle trajectory data in mixed traffic and analysis of these data. The data was collected using video photography in an urban midblock road section in Chennai, India. The trajectory data was extracted from the video sequences using the Trajectory Extractor (16). Then, the data was processed using the locally weighted regression method (17,18) in order to reduce measurement errors and to obtain continuous position functions that may be differentiated once or twice in order to obtain speed and acceleration functions, respectively. Finally, the traffic flow characteristics of these trajectories are investigated.

DATA COLLECTION

Site
The video data was collected on an urban arterial road at the Maraimalai Adigalar Bridge in Saidapet, Chennai, India. This is a six-lane separated arterial. Collection took place on the northbound approach, which is shown in FIGURE 1. The section is on a bridge. This ensures that the road geometry is uniform and that there are no nearby intersections, bus stops, parked vehicles and other side factors that may affect drivers’ behavior. Furthermore, there is no interaction between the vehicle traffic and pedestrians, as the pedestrian walkway is segregated by a barrier. The video data were recorded over the period 10.00 AM- 3.30 PM on February 13, 2014. The video camera was located on a roof of a building adjacent to the section.

FIGURE 1 Data collection site in Chennai, India
Trajectory Extraction

Over the years, several automated or semi-automated tools to extract trajectory data from video sequences have been developed. Some examples include VEVID (19), the NGSIM-Video (20) and Trajectory Extractor (16). The latter was used to collect motorcycle and bicycle trajectories in London, UK. TRAZER (21) and Traffic Data Extractor (15) have been developed and used for trajectory extraction in mixed traffic.

In this study, Trajectory Extractor (16) was used to obtain the coordinates, dimensions and vehicle class of all vehicles that appeared in the video sequences during a period of 30 minutes between 2:45 PM and 3:15 PM. This time period represents medium level traffic flows, which exhibit both vehicle following and lateral shift behaviors. The trajectories were extracted at a time resolution of 0.5 second. The extraction is semi-automated. A Windows-based graphical user interface allows a human operator to use the mouse pointer to identify the edges of vehicles on the screen. The system converts these into real-world coordinates and calculates vehicles’ positions, speeds and accelerations. The coordinate conversion relies on four reference points in the video images and their coordinates in the real-world. FIGURE 2 shows the software’s graphical user interface showing an image from the road section of this study and the rectangle created by the four reference points.

FIGURE 2 Trajectory Extractor user interface showing the road section and reference points
Data Smoothing

Once the position data has been extracted, it needs to be smoothed in order to overcome missing observations (e.g. due to occlusions), reduce measurement errors and calculate other variables of interest, such as speeds, accelerations and inter-vehicle relations. Several studies (22, 23, 24) have shown that this step is necessary in order to obtain unbiased and internally consistent trajectories. Methods for the trajectory data processing relied on signal filtering (25, 26), smoothing methods (18, 27) or moving average techniques (23, 28).

The locally weighted regression approach proposed by Toledo et al. (18) was used for data smoothing. The method uses a set of \( N \) (window size) observations before and after the measurement point of interest, \( t_0 \). The trajectory function around this point is assumed to be a polynomial function of time:

\[
x(t) = f_{t_0}(t, \beta_{t_0}) + \varepsilon_{t_0,t} = \sum_{m=0}^{M} \beta_{t_0,m} (t)^m + \varepsilon_{t_0,t}
\]

(1)

Where, \( f_{t_0}(t, \beta_{t_0}) \) is the fitted position at time \( t \) estimated by a local regression function centered at time \( t_0 \). \( \beta_{t_0} = \begin{bmatrix} \beta_{t_0,0} & \beta_{t_0,1} & \beta_{t_0,2} & \cdots & \beta_{t_0,M} \end{bmatrix} \) is a vector of the \( M + 1 \) parameters of the polynomial function estimated around time \( t_0 \). \( \varepsilon_{t_0,t} \) are normally distributed error terms.

The parameters of this local function are estimated using \( N \) observations in the window around \( t_0 \) with a weighted least-squares estimator:

\[
\beta_{t_0} = \arg \min_{\beta} \left[ X_{t_0} - f_{t_0}(t, \beta) \right]' W_{t_0} \left[ X_{t_0} - f_{t_0}(t, \beta) \right]
\]

(2)

Where, \( X_{t_0} \) is the column vector of \( N \) position observations used to estimate a trajectory function centered on \( t_0 \). \( f_{t_0}(t, \beta_{t_0}) \) is the corresponding vector of fitted values. \( W_{t_0} \) is an \([N \times N]\) diagonal matrix, with elements corresponding to the weights of the observations.
The observation weights are a tricube function of its distance from the point of interest $t_0$:

$$w(t_0,t) = \left(1 - \left(\frac{|t-t_0|}{d}\right)^3\right)^3$$

(3)

$w(t_0,t)$ is the weight assigned to the observation at time $t$ in fitting a curve centered at $t_0$. $d$ is the distance from $t_0$ to the nearest point outside the window of $N$ points to be considered in fitting the curve.

Upper and lower bound constraints on the estimated speeds and accelerations are also added to the optimization problem in Equation 2. Instantaneous speeds and accelerations are calculated as the first and second derivatives with respect to time of the fitted position polynomial function.

The longitudinal and lateral positions were smoothed independently of each other. Following the results in Toledo et al. (18) and some experimentation with the current data set, a window size of $N = 7$ and polynomial order $M = 4$ were used in both cases. In the longitudinal direction, the smoothed data has a mean average error (MAE) of 0.544 meters and a root mean squared error (RMSE) of 0.788 meters, compared to the raw data. In the lateral dimension the errors are MAE=0.062 meters and RMSE=0.082 meters.

Punzo et al. (24) defined consistency conditions that need to be met for trajectory data to be useful for study of driving behavior. The internal consistency condition guarantees the agreement between position, speed and acceleration values. The smoothing method used above defines the position as a continuous function, and speeds and accelerations as its derivatives. Thus, the internal consistency condition is satisfied by definition. The platoon consistency condition guarantees that there are no overlaps between the trajectories of different vehicles, which imply collisions. 4,107 position points, out of 111,629 in the dataset, violated this condition. In these cases, the two overlapping vehicles were moved to eliminate the collisions, and the smoothing process repeated. These modified observations were flagged in the dataset, and should be used with caution for microscopic-level analysis.
TRAFFIC FLOW CHARACTERISTICS

The collected dataset includes 3005 vehicle trajectories. The positions are observed at a resolution of 0.5 second generating a total of 111,629 observations. Mixed traffic flow has distinct characteristics that distinguish it from homogeneous traffic. The first is a more varied vehicle mix. This is well demonstrated in the collected data: only 26.6% of the vehicles in the traffic flow were passenger cars, while 56.4% were motorcycles, 12.2% were auto-rickshaws and 4.8% heavy vehicles including light and heavy trucks and buses.

Longitudinal Movement

TABLE 1 presents summary statistics of the traffic flow characteristics in the longitudinal direction. Traffic flows and densities are 1-minute averages. The reported speed and acceleration statistics are for instantaneous values. The total traffic flow observed in the study section is 6010 veh/hr. Instantaneous speeds vary from 0 to 15.22 m/sec with a mean of 5.88 m/sec. The average speeds of the various vehicle types in the stream differ. The mean speed of cars is the highest (6.13 m/sec), followed by motorcycles (6.01 m/sec). Heavy vehicles (5.64 m/sec) and especially auto-rickshaws (5.06 m/sec) travel at lower speeds. ANOVA tests were conducted for the average speeds of individual vehicles. These test shows that the differences among the vehicle types are statistically significant ($F(3,3001)=114.93$, $p$-value$<0.001$). Pairwise comparisons show the mean speeds of auto-rickshaws and heavy vehicles are each statically significant compared to those of motorcycles and cars. All for differences are significant with $p$-value$<=0.001$. The mean speeds of motorcycles and cars are not statistically significant ($p$-value$=0.590$). These differences may stem from the higher operating capabilities of cars and the higher maneuverability of motorcycles within the traffic stream, compared to the lower maneuverability of heavy vehicles and poor dynamics characteristics of auto-rickshaws.

The acceleration rates applied by the various vehicle types also differ for both acceleration and deceleration ($F(3,3001)=51.76$, $p$-value$<0.001$) and ($F(3,3001)=64.21$, $p$-value$<0.001$, respectively). The mean deceleration and acceleration rates of motorcycles (-0.731 m/sec$^2$ and 0.761 m/sec$^2$ and, respectively) are higher compared to other types of vehicles. The post-hoc analysis shows that these are statistically different from all the other types in both acceleration and deceleration ($p$-value$<0.001$ in all cases). This again may be due to the greater
maneuverability of motorcycles, which allows them to apply higher acceleration and deceleration rates as they weave through the traffic. Other vehicle types are more constrained by the vehicles surrounding them due to their size and lesser maneuvering capabilities.

**TABLE 1 Longitudinal Traffic Flow Characteristics of the Collected Data**

<table>
<thead>
<tr>
<th>Vehicle Type</th>
<th>Mean</th>
<th>Std dev.</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flow (veh/hr)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motorcycle</td>
<td>3390</td>
<td>643.5</td>
<td>3300</td>
<td>2040</td>
<td>4920</td>
</tr>
<tr>
<td>Car</td>
<td>1600</td>
<td>454.3</td>
<td>1500</td>
<td>960</td>
<td>2880</td>
</tr>
<tr>
<td>Auto-Rickshaw</td>
<td>732</td>
<td>220.3</td>
<td>720</td>
<td>240</td>
<td>1080</td>
</tr>
<tr>
<td>Heavy Vehicles</td>
<td>288</td>
<td>124.5</td>
<td>270</td>
<td>60</td>
<td>540</td>
</tr>
<tr>
<td>All types</td>
<td>6010</td>
<td>1004.5</td>
<td>5940</td>
<td>3960</td>
<td>7860</td>
</tr>
<tr>
<td>Density (veh/km/lane)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motorcycle</td>
<td>67.1</td>
<td>13.2</td>
<td>64.1</td>
<td>40.8</td>
<td>91.9</td>
</tr>
<tr>
<td>Car</td>
<td>31.9</td>
<td>11.1</td>
<td>31.4</td>
<td>12.2</td>
<td>59.5</td>
</tr>
<tr>
<td>Auto-Rickshaw</td>
<td>18.0</td>
<td>5.2</td>
<td>18.5</td>
<td>4.2</td>
<td>26.5</td>
</tr>
<tr>
<td>Heavy Vehicles</td>
<td>6.5</td>
<td>2.7</td>
<td>6.8</td>
<td>0.7</td>
<td>11.2</td>
</tr>
<tr>
<td>All types</td>
<td>123.5</td>
<td>24.0</td>
<td>120.4</td>
<td>77.7</td>
<td>186.9</td>
</tr>
<tr>
<td>Speed (m/sec)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motorcycle</td>
<td>6.01</td>
<td>1.44</td>
<td>5.94</td>
<td>0.02</td>
<td>15.22</td>
</tr>
<tr>
<td>Car</td>
<td>6.13</td>
<td>1.29</td>
<td>6.06</td>
<td>0.37</td>
<td>13.96</td>
</tr>
<tr>
<td>Auto-Rickshaw</td>
<td>5.06</td>
<td>1.19</td>
<td>5.03</td>
<td>0</td>
<td>11.51</td>
</tr>
<tr>
<td>Heavy Vehicles</td>
<td>5.64</td>
<td>1.13</td>
<td>5.67</td>
<td>0</td>
<td>10.40</td>
</tr>
<tr>
<td>All types</td>
<td>5.88</td>
<td>1.40</td>
<td>5.82</td>
<td>0</td>
<td>15.22</td>
</tr>
<tr>
<td>Acceleration (m/sec²)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motorcycle</td>
<td>0.761</td>
<td>0.748</td>
<td>0.519</td>
<td>0</td>
<td>4.734</td>
</tr>
<tr>
<td>Car</td>
<td>0.646</td>
<td>0.653</td>
<td>0.431</td>
<td>0</td>
<td>4.436</td>
</tr>
<tr>
<td>Auto-Rickshaw</td>
<td>0.692</td>
<td>0.712</td>
<td>0.459</td>
<td>0</td>
<td>4.501</td>
</tr>
<tr>
<td>Heavy Vehicles</td>
<td>0.672</td>
<td>0.652</td>
<td>0.465</td>
<td>0</td>
<td>3.981</td>
</tr>
<tr>
<td>All types</td>
<td>0.717</td>
<td>0.717</td>
<td>0.484</td>
<td>0</td>
<td>4.734</td>
</tr>
<tr>
<td>Deceleration (m/sec²)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motorcycle</td>
<td>-0.731</td>
<td>0.714</td>
<td>-0.503</td>
<td>-4.659</td>
<td>0</td>
</tr>
<tr>
<td>Car</td>
<td>-0.605</td>
<td>0.608</td>
<td>-0.407</td>
<td>-4.371</td>
<td>0</td>
</tr>
<tr>
<td>Auto-Rickshaw</td>
<td>-0.654</td>
<td>0.668</td>
<td>-0.426</td>
<td>-4.340</td>
<td>0</td>
</tr>
<tr>
<td>Heavy Vehicles</td>
<td>-0.630</td>
<td>0.623</td>
<td>-0.420</td>
<td>-4.208</td>
<td>0</td>
</tr>
<tr>
<td>All types</td>
<td>-0.681</td>
<td>0.679</td>
<td>-0.460</td>
<td>-4.659</td>
<td>0</td>
</tr>
</tbody>
</table>

**Lateral Movement**

An important characteristic of mixed traffic is the existence of substantial lateral movement and lack of lane discipline. **TABLE 2** presents the summary statistics of the lateral movements in the collected data. Speeds and accelerations/decelerations in this direction are obviously much lower
compared to the longitudinal direction. But, the differences among vehicle types remain similar. Motorcycles and cars have higher average lateral speeds compared to auto-rickshaws and heavy vehicles.

### TABLE 2 Lateral Traffic Flow Characteristics of the Collected Data

<table>
<thead>
<tr>
<th>Vehicle Type</th>
<th>Mean</th>
<th>Std dev.</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed (m/sec)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motorcycle</td>
<td>0.116</td>
<td>0.107</td>
<td>0.087</td>
<td>0</td>
<td>1.458</td>
</tr>
<tr>
<td>Car</td>
<td>0.095</td>
<td>0.089</td>
<td>0.071</td>
<td>0</td>
<td>1.209</td>
</tr>
<tr>
<td>Auto-Rickshaw</td>
<td>0.082</td>
<td>0.075</td>
<td>0.062</td>
<td>0</td>
<td>1.215</td>
</tr>
<tr>
<td>Heavy Vehicles</td>
<td>0.088</td>
<td>0.079</td>
<td>0.067</td>
<td>0</td>
<td>0.798</td>
</tr>
<tr>
<td>All types</td>
<td>0.104</td>
<td>0.098</td>
<td>0.078</td>
<td>0</td>
<td>1.458</td>
</tr>
<tr>
<td>Acceleration (m/sec^2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motorcycle</td>
<td>0.090</td>
<td>0.075</td>
<td>0.072</td>
<td>0</td>
<td>0.648</td>
</tr>
<tr>
<td>Car</td>
<td>0.083</td>
<td>0.069</td>
<td>0.066</td>
<td>0</td>
<td>0.606</td>
</tr>
<tr>
<td>Auto-Rickshaw</td>
<td>0.077</td>
<td>0.065</td>
<td>0.061</td>
<td>0</td>
<td>0.639</td>
</tr>
<tr>
<td>Heavy Vehicles</td>
<td>0.084</td>
<td>0.069</td>
<td>0.067</td>
<td>0</td>
<td>0.548</td>
</tr>
<tr>
<td>All types</td>
<td>0.086</td>
<td>0.072</td>
<td>0.068</td>
<td>0</td>
<td>0.648</td>
</tr>
<tr>
<td>Deceleration (m/sec^2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motorcycle</td>
<td>-0.091</td>
<td>0.078</td>
<td>-0.070</td>
<td>-0.639</td>
<td>0</td>
</tr>
<tr>
<td>Car</td>
<td>-0.082</td>
<td>0.072</td>
<td>-0.064</td>
<td>-0.592</td>
<td>0</td>
</tr>
<tr>
<td>Auto-Rickshaw</td>
<td>-0.077</td>
<td>0.069</td>
<td>-0.059</td>
<td>-0.652</td>
<td>0</td>
</tr>
<tr>
<td>Heavy Vehicles</td>
<td>-0.080</td>
<td>0.070</td>
<td>-0.062</td>
<td>-0.561</td>
<td>0</td>
</tr>
<tr>
<td>All types</td>
<td>-0.086</td>
<td>0.075</td>
<td>-0.066</td>
<td>-0.652</td>
<td>0</td>
</tr>
</tbody>
</table>

Motorcycles and cars have higher average lateral speeds (0.116 m/sec and 0.095 m/sec, respectively) compared to auto-rickshaws and heavy vehicles (0.082 m/sec and 0.088 m/sec, respectively). ANOVA results shows that these differences are statistically significant (F(3,3001)=109.27, p-value<0.001), both overall and in the comparison of pairs of vehicle types (p-value<0.001 in all cases). This may be due to swerving or weaving in traffic by the motorcycles due their size and higher maneuverability compared to other vehicle types. Motorcycles also have higher values of mean lateral deceleration and acceleration (-0.091 m/sec^2 and -0.090 m/sec^2, respectively). Auto-rickshaws have the lowest mean values (-0.077 m/sec^2 and 0.077 m/sec^2, respectively). The inequality of mean lateral deceleration and acceleration values among the vehicle classes are statistically significant (F(3,3001)=90.33, p-value<0.001 F(3,3001)=65.73, p-value<0.001, respectively). The post hoc test shows that lateral decelerations and accelerations of motorcycles are different from all other vehicle types (p-
value < 0.001 in all cases). The values of cars and auto-rickshaws are also statistically significant in both decelerations and accelerations (p-value = 0.006 and p-values < 0.001, respectively).

**Lateral Position**

The lack of lane discipline affects drivers’ choices related to the lateral positions of their vehicles. FIGURE 3 shows the distribution of lateral positions by vehicle type. It is useful to note that driving in India is on the left-hand side. The lateral position 0.0 is on the left-most (near) side of the roadway, and 11.2 on the right-most (far) side. The mean lateral positions of motorcycles (4.39 meters) and auto-rickshaws (4.51 meters) are to the left of those of heavy vehicles (5.85 meters) and cars (7.16 meters). The lateral position distributions of motorcycles and auto-rickshaws are skewed to the near side. Nearly, 74% of the motorcycles and 62% of the auto-rickshaws are observed on the near third of the roadway (0.0 to 3.73 meters) and only 18.82% and 19.32% of motorcycles and auto-rickshaws are observed on the far third (7.47 to 11.20 meters), respectively. In contrast, the lateral position distribution of cars is skewed to the far side of the roadway. 56% of the cars are observed in the far third and only 6% are observed in the near third. Heavy vehicles tend to be in the middle third of the roadway (3.74 to 7.46 meters). 55% of their observations are in this part and about 16% and 29% are on the near and far sides, respectively. ANOVA tests on the positions of different types of vehicles show that they are significantly different (F(3,3001)=241.33, p-value < 0.001). All pairwise comparisons are statistically significant (p-value < 0.001 in all cases), except those of motorcycles and auto-rickshaws. A possible explanation may be that cars tend to prefer the higher speeds and lesser friction with other vehicle types and obstructions (e.g. parked vehicles, bicycles and pedestrians) offered by the far side of the roadway. The maneuverability of motorcycles enables them to obtain higher speeds even on the near side. Therefore, they may prefer to avoid interacting with the larger cars and heavy vehicles and keep to the near side. Auto-rickshaws stop for passengers on the near side, which may further affect their tendency to stay on this side of the roadway. Most of the heavy vehicles in the section are buses. They need to make stops on the near side, but at the same time may prefer to avoid interacting with the smaller vehicle types and other obstructions that are also more present on the near side.
Lateral Movement Variation

In order to evaluate the extent of lateral movement that vehicles undertake within the section, the standard deviation of the lateral positions within the section was calculated. FIGURE 4 presents the distributions of these standard deviations. Overall, auto-rickshaw make the least lateral movements (0.43 meters) followed by heavy vehicles (0.49 meters), and cars (0.51 meters). Motorcycles have higher values (0.62 m). The differences among these values are statistically significant ($F(3, 3001)=34.01$, $p$-value <0.001). Again, it is plausible that their smaller size and higher maneuverability allow motorcycles easier lateral movement compared to other vehicles. This is supported by the lesser lateral movement of heavy vehicles and auto-rickshaws that are characterized by large sizes (heavy vehicles) and poor maneuverability.

Longitudinal Spacing

Longitudinal spacing is a variable that captures the relations between vehicles in the stream. It is an important explanatory variable in car following models. In non-lane based traffic the definition of the leader and follower is not trivial. For this purpose a leader is defined as the nearest vehicle in front of the subject that laterally overlap with it, and that the spacing between
FIGURE 4 Distributions of standard deviations of lateral positions of various vehicle types

the two vehicles is less than 30 meters (roughly 2 seconds). FIGURE 5 and FIGURE 6 present
the distributions of longitudinal spacing by the vehicle type of the leader and the follower,
respectively. With respect to the leader type, vehicles maintain larger spacing with heavy
vehicles (16.19 meters) followed by auto-rickshaws (15.12 meters), and lower spacing with
motorcycles (14.77 meters) and cars (14.52 meters). ANOVA analysis shows that the effect of
the leader type is statistically significant (F(3,1503)=3.77, p-value=0.010). But, in pairwise
comparisons, only the differences between heavy vehicles and motorcycles (p-value=0.046) and
cars (p-value=0.017) are significant. A similar trend is also observed with respect to the follower
type. Heavy vehicles (15.83 meters) and auto-rickshaws (15.64 meters) maintain larger spacing
compared to motorcycles (14.40 meters) and cars (14.63 meters). These differences are
statistically significant (F(3,1504)=6.43, p-value<0.001).
FIGURE 5 Distributions of longitudinal spacing based on the leader type

FIGURE 6 Distributions of longitudinal spacing based on the follower type
Lateral Overlap

In homogeneous traffic, vehicles are predominantly cars which adhere to lane discipline. Hence, most of the time, a vehicle strictly follows the leader in the same. In mixed traffic, the lack of lane discipline and mix of vehicles with different width characteristics causes different types of relations between vehicles in the stream. In order to investigate the following scenario, the extent of lateral overlap between a leader and follower is defined as the percentage of the follower width (W) that laterally overlaps (LD) with the leader. The widths of the leader and follower may differ substantially. Therefore, the overlap is defined by the smaller distance between opposite edges of the two vehicles, as shown in FIGURE 7. With this definition, the lateral overlap may exceed 100%, if the follower is narrower and entirely overlaps with the leader.

![FIGURE 7 Definition of lateral overlap](image)

The distributions of lateral overlap are shown in Figure 8. The mean value for lateral overlap is higher for motorcycles (61%) followed by cars (51%). Auto-rickshaws (45%) and heavy vehicles (40%) have lower overlap values. These differences among the vehicle types are statistically significant (F(3,1503) =38.78,p-value<0.001). All pairwise comparisons, except that of heavy vehicles and auto-rickshaws (p-value=0.804) are statistically significant. Overall, the figure shows a wide range of overlap values. In 45% of the observations the overlap between the leader and follower is less than half of the follower width. Thus, it is not evident that car following behavior is applicable to these situations, and suggests a need to study other types of following behavior (e.g. staggered following) in mixed traffic streams.
SUMMARY

This paper focused on the study of the traffic characteristics of mixed traffic. For this purpose, a detailed set of vehicle trajectory data were collected in an urban midblock road section in Chennai, India. These data were processed and smoothed to reduce the effects of measurement errors and to estimate instantaneous position, speed and acceleration values. The data is freely available from the authors.

The resulting data was studied with respect to aggregate traffic flow characteristics and variables related to the longitudinal and lateral movement of the vehicles. Several insights arise from this analysis:

1. A main characteristic of mixed traffic is the presence of significant numbers of vehicles of various types in the stream. There are substantial differences in the flow characteristics among these vehicle types. The results show differences in travel speeds, accelerations, the choice of lateral position on the roadway, and practically all other measures that were studied.

2. Car following is a critical component of driving behavior. Analysis of the relations between leaders and followers shows differences in distance keeping between the various types of
vehicles. Furthermore, in almost half the observations, strict following (in which a vehicle follows almost exactly behind another one) is not present.

3. Another characteristic of mixed traffic is the weak lane discipline. The study showed that vehicles in the stream, and in particular motorcycles, move substantially in the lateral direction.

Driving behavior models for mixed traffic streams should account for the characteristics described above. Models need to account for the different capabilities and preferences of various vehicle types. Specific attention should be given to modeling of lateral movements, and in particular those of motorcycles. Longitudinal movement models should focus not only on car following, but also consider other situations in which a vehicle does not strictly follow another one. The trajectory data collected in this study may be useful in studies in these directions.

ACKNOWLEDGMENTS
The first author was partly supported by a fellowship of the Israel Council for Higher Education at the Technion – Israel institute of Technology.

REFERENCES


