# IMPROVED SIMULATION OF STOP BAR DRIVER BEHAVIOR AT SIGNALIZED INTERSECTIONS

**Final Report**

**KLK712**

**N10-07**

**National Institute for Advanced Transportation Technology**

**University of Idaho**

NIATT

**Michael Dixon, Ahmed Abdel-Rahim, Michael Kyte**

**April 2010**

## DISCLAIMER

| | | |
|---|---|---|
| 1. Report No. | 2. Government Accession No. | 3. Recipient's Catalog No. |
| 4. Title and Subtitle<br>Improved Simulation of Stop Bar Driver Behavior at Signalized Intersections | | 5. Report Date<br>April 2010<br>6. Performing Organization Code<br>KLK712 |
| 7. Author(s)<br>Dixon, Dr. Michael; Abdel-Rahim, Dr. Ahmed; Kyte, Dr. Michael | | 8. Performing Organization Report No.<br>N10-07 |
| 9. Performing Organization Name and Address<br>National Institute for Advanced Transportation Technology<br>University of Idaho<br>PO Box 440901; 115 Engineering Physics Building<br>Moscow, ID 83844-0901 | | 10. Work Unit No. (TRAIS)<br>11. Contract or Grant No.<br>DTRT07-G-0056 |
| 12. Sponsoring Agency Name and Address<br>US Department of Transportation<br>Research and Special Programs Administration<br>400 7th Street SW<br>Washington, DC 20509-0001 | | 13. Type of Report and Period Covered<br>Final Report: August 2007 – August 2009<br>14. Sponsoring Agency Code<br>USDOT/RSPA/DIR-1 |
| 15. Supplementary Notes: | | |

16. Abstract

The Federal Highway Administration's Next Generation Simulation (NGSIM) program has identified a set of priority algorithms that are yet to be developed. These algorithms include starting/stopping behavior and permitted left turns at signalized intersections. The project has two objectives: 1. Document characteristics of stop bar behavior at both the beginning and ending of green and 2. Determine the level of precision with which car following algorithms can be expected to represent vehicle headways and speed.

Some errors were found and corresponding corrections were made, resolving most of the errors. However, problems exist for individual driver stop bar behavior data. It was found that the primary cause for this shortcoming is the accuracy of the phase status time stamps. Efforts to correct these time stamps did not remove errors and, as a result, driver behavior measurements, such as the response to the onset of yellow, were unrealistic. Overall, for each of the errors found, suggestions are given to control or remove these errors in future data collection efforts. Fortunately, the data do seem to relate vehicle-vehicle interactions realistically. This is especially the case after filtering out spurious changes in velocity.

| | |
|---|---|
| 17. Key Words<br>Microscopic traffic flow; traffic simulation; traffic characteristics | 18. Distribution Statement<br>Unrestricted; Document is available to the public through the National Technical Information Service; Springfield, VT. |

| | | | |
|---|---|---|---|
| 19. Security Classif. (of this report)<br>Unclassified | 20. Security Classif. (of this page)<br>Unclassified | 21. No. of Pages<br>56 | 22. Price<br>… |

Form DOT F 1700.7 (8-72)          Reproduction of completed page authorized

# Table of Contents

# List of Figures

## List of Tables

# 1 INTRODUCTION

The Federal Highway Administration's Next Generation Simulation (NGSIM) program has identified a set of priority algorithms that are yet to be developed. These algorithms include starting/stopping behavior and permitted left turns at signalized intersections. To support the FHWA NGSIM program objectives, the University of Idaho Transportation Center started a project titled "Improved Simulation Stop Bar Driver Behavior at Signalized Intersections." The project developed new insights on NGSIM dataset description of the operation of traffic along arterials, particularly driver behavior at the beginning and end of green. The project used high resolution datasets from Atlanta, Georgia.

Originally, the project had two objectives, and these are listed below, where each objective has a list of driver behavior areas or issues. Significant effort was needed to assess dataset shortcomings, determine corrective measures, and execute all possible corrective measures. However, the original project scope did not account for these assessment and corrective actions. As a result, resources did not allow work to continue to conclusive results related to the second objective.

1) Document characteristics of stop bar behavior at both the beginning and ending of green.
    a. Driver response time
    b. Start-up lost time
    c. Probability of stopping given distance to stop bar and beginning of yellow
2) Determine the level of precision with which car following algorithms can be expected to represent vehicle headways and speed.
    a. Mid-block driving
    b. Decelerating to stop in queue
    c. Accelerating out of queue

This report is organized into six chapters. Chapter 1 is the introduction. Chapter 2 describes the data that are required to measure driver behavior characteristics at the stop bar. Chapter 3 discusses the NGISM dataset characteristics. This chapter is categorized into six sections (section 3.1 to section 3.6). In this chapter, NGSIM dataset characteristics along with examples of measuring the parameters from the vehicle trajectory data are explained thoroughly. Chapter 4 discusses integrating dataset tables. Chapter 5 explains the corrections that were made to clean the dataset for this research. This chapter is divided into six sections (section 5.1 to section 5.6). These sections discuss corrections in the following area dataset continuity, stop bar locations, phase status time synchronization, unnumbered vehicles, vehicle location correction, and spurious stops. The next chapter (chapter 6) describes the analysis of the dataset. This chapter is organized into three sections (section 6.1 to section 6.3) which include the beginning of green, end of green, safety clearance distance, and vehicle velocity and acceleration.

# 2 DATASET REQUIREMENTS FOR TARGETED DRIVER BEHAVIOR

Stop bar driver behavior research requires certain types of traffic information in order to accurately measure driver behavior. To accomplish this, researchers need to know vehicle location, phase status, and the corresponding times. From this information, the circumstances of the driver/vehicle were assessed, such as whether or not the vehicle was in a queue or stopped, the vehicle's location relative to the stop bar, or the status of the phase controlling the vehicle's right-of-way.

To accurately extract the above information, the vehicle trajectory data needed to be time synchronized with the phase status data and location synchronized with the stop bar location data. Once synchronized, measurements relevant to the targeted driver behavior were taken and analyzed.

# 3 NGSIM DATASET CHARACTERISTICS

This section presents an overview of the NGSIM Atlanta dataset. The NGSIM dataset holds 24 different fields whose descriptions are given in a data dictionary subsection. The geometric configuration and intersection lengths are described in the subsection paragraphs. The total numbers of numbered vehicles and their records and parameter ranges are expressed in the following sections. The different types of parameters that can be obtained from dataset trajectory data are also grouped together in this section. The parameters are:

i. End-of-green-gain
ii. Safety clearance distance
iii. Start-up response time

iv. Start-up lost time
v. Response time
vi. Follow-up headway
vii. Velocity, acceleration, deceleration rates of vehicles

The summary of the above car following model parameters that can be obtained from the Atlanta NGSIM vehicle trajectory data are also presented in this section in a tabular format.

## 3.1 Overview

Only the NGSIM Atlanta dataset is used in the analysis. The dataset represents vehicle trajectories on a segment of Peachtree Street in Atlanta, Georgia collected between 12:45 p.m. and 1:00 p.m. and between 4:00 p.m. and 4:15 p.m. on November 8, 2006. In addition to the vehicle trajectory data, signal indication information is available for the dataset. The geometric configuration of the corridor of the Atlanta dataset is shown in Figure 3-1.

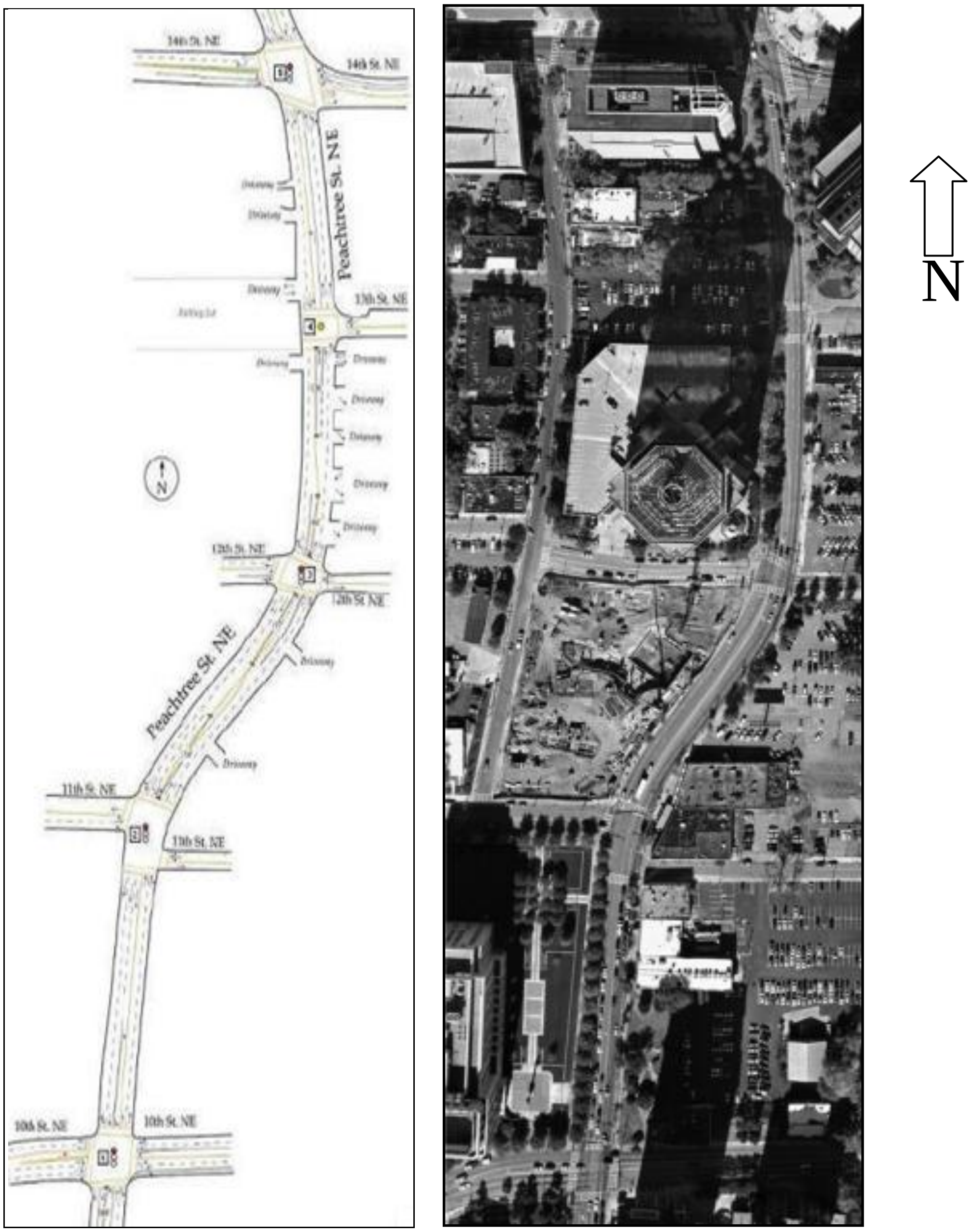



**Figure 3-1: Peachtree Street - Atlanta, Georgia.**

## 3.2 Data Dictionary

Each row in the NGSIM dataset represents the temporary and spatial data for a single vehicle reported in 24 different fields. Description of the data included in each field is provided in Table 3-1.

**Table 3-1: Description of the NGSIM Vehicle Trajectory Dataset**

| Column | Name | Unit | Description |
|---|---|---|---|
| 1 | Vehicle ID | Number | Vehicle identification number (ascending by time of entry into section). |
| 2 | Frame ID | 0.1 sec. | Frame identification number (ascending by start time). |
| 3 | Total Frames | 0.1 sec. | Total number of frames in which the vehicle appears in this dataset. |
| 4 | Global Time | 0.001 sec | Epoch time, elapsed time since Jan. 1, 1970. |
| 5 | Local X | Feet | Lateral (X) coordinate of the front center of the vehicle - perpendicular to the median of the Peachtree Street. Vehicles traveling on the east side of the median have positive local X values, while those traveling on the west side of the median have negative local X values. |
| 6 | Local Y | Feet | Longitudinal (Y) coordinate of the front center of the vehicle along the median of the Peachtree Street. The start point is at the southern boundary of the study area. |
| 7 | Global X | Feet | X coordinate of the front center of the vehicle based on NAD83. |
| 8 | Global Y | Feet | Y coordinate of the front center of the vehicle based on NAD83. |
| 9 | Vehicle Length | Feet | Length of vehicle. |
| 10 | Vehicle Width | Feet | Width of vehicle. |
| 11 | Vehicle Class | Number | Vehicle type: 1 - motorcycle, 2 - auto, 3 – truck. |
| 12 | Vehicle Speed | ft/sec | Instantaneous speed of vehicle. |
| 13 | Vehicle Acceleration | $ft/sec^2$ | Instantaneous acceleration of vehicle. |
| 14 | Lane Identification | Number | Current lane position of vehicle. Lane numbering is incremented from the left-most lane, except for locations where left-turn or right-turn bays exist. Left-turn bays are numbered starting from 11 and are incremented from the left-most left-turn bay. |
| 15 | Origin Zone | Number | Origin zones of the vehicles, i.e., the place where the vehicles enter the tracking system. There are 21 origins in the study area, numbered from 101 through 123. Destination 204 and 209 are a one-way off-ramp; hence, there are no associated origin number 104 and 109. Please refer to the data analysis report for more detailed information. |
| 16 | Destination Zone | Number | Destination zones of the vehicles, i.e. the place where the vehicles exit the tracking system. There are 22 destinations in the study area, numbered from 201 through 223. Origin 119 is a one-way off-ramp; hence there is no associated destination |

| | | | |
|---|---|---|---|
| | | | number 219. Please refer to the data analysis report for more detailed information. |
| 17 | Intersection | Number | Intersection in which the vehicle is traveling. Intersections are numbered from 1 to 5, with intersection 1 at the southernmost, and intersection 5 at the northernmost section of the study area. |
| 18 | Section | Number | Section in which the vehicle is traveling. Peachtree Street is divided into six sections (south of intersection 1; between intersections 1 and 2, 2 and 3, 3 and 4, 4 and 5, 5 and 6; and north of intersection 6). Please refer to the data analysis report for more detailed information. |
| 19 | Direction | Number | Moving direction of the vehicle. 1 - east-bound (EB), 2 - north-bound (NB), 3 - west-bound (WB), 4 - south-bound (SB). |
| 20 | Movement | Number | Movement of the vehicle. 1 - through (TH), 2 - left-turn (LT), 3 - right-turn (RT). |
| 21 | Preceding Vehicle | Number | Vehicle ID of the lead vehicle in the same lane. A value of '0' represents no preceding vehicle. |
| 22 | Following Vehicle | Number | Vehicle ID of the vehicle following the subject vehicle in the same lane. A value of '0' represents no following vehicle. |
| 23 | Spacing | Feet | Space headway; spacing provides the distance between the front-center of a vehicle to the front-center of the preceding vehicle. |
| 24 | Headway | Second | Time headway; headway provides the time to travel from the front-center of a vehicle (at the speed of the vehicle) to the front-center of the preceding vehicle. A headway value of 9999.99 means that the vehicle is traveling at zero speed (congested conditions). |

## 3.3 Intersections

In the current NGSIM (Peachtree Arterial, Atlanta) dataset, there are five intersections (see Figure 3-1). Of them, four are signalized and one is unsignalized. The intersections are denoted as number 1, 2, 3, 4 and 5 in the dataset, increasing in the north direction. Intersections 1, 2, 3 and 5 are signalized intersections and intersection 4 is unsignalized. All four signalized intersections have major street approaches with two through traffic lanes and a left turn bay except for intersection 2 which has no left turn bay. Also, there is no separate right turn lane for any of the major street approaches of the four signalized intersections. All intersections have protected left turns on the major street approaches with protected-permitted or permitted left turn on minor streets. Intersection 4 is a two-way stop control intersection, requiring minor street traffic to stop.

All lane widths are 12 feet at the intersections and lane/pedestrian crossing markings are clearly marked. Pedestrian phases run concurrent with their adjacent through traffic movements. There are pedestrian crossings on each intersection approach.

There are six sections denoted as sections 1 through 6 in the dataset. Section 2, 3, 4 and 5 represent the roadways between the intersections. The intersection and section lengths are shown in Table 3-2 and Table 3-3 below.

**Table 3-2: Intersection Lengths**

| Dataset | Intersection | Length, ft |
|---|---|---|
| A1 | 1 | 99.645 |
| | 2 | 129.795 |
| | 3 | 73.815 |
| | 4 | 66.979 |
| | 5 | 117.411 |
| A2 | 1 | 99.732 |
| | 2 | 129.875 |
| | 3 | 73.513 |
| | 4 | 66.602 |
| | 5 | 121.319 |

**Table 3-3: Section Lengths**

| Dataset | Section | Length, ft |
|---|---|---|
| A1 | 1 | 127.391 |
| | 2 | 441.437 |
| | 3 | 412.070 |
| | 4 | 353.727 |
| | 5 | 343.922 |
| | 6 | 11.730 |
| A2 | 1 | 143.619 |
| | 2 | 417.976 |
| | 3 | 412.172 |
| | 4 | 351.511 |
| | 5 | 344.427 |
| | 6 | 28.636 |

## 3.4 Number of Vehicles

There are a total number of 1115 and 1222 numbered vehicles in dataset 1 and dataset 2, respectively. Also there are some unnumbered vehicles along the corridor but not included in the original datasets. Vehicles widths vary from 5ft to 8.5ft. Vehicles lengths vary from 3.5ft to 42.6ft. Figure 3-2 shows the vehicles by length for both datasets. As expected, the largest segment of the vehicle lengths is in the 15 to 20 ft range, which reflects the large majority of passenger cars. Very few heavy vehicles exist in the datasets.

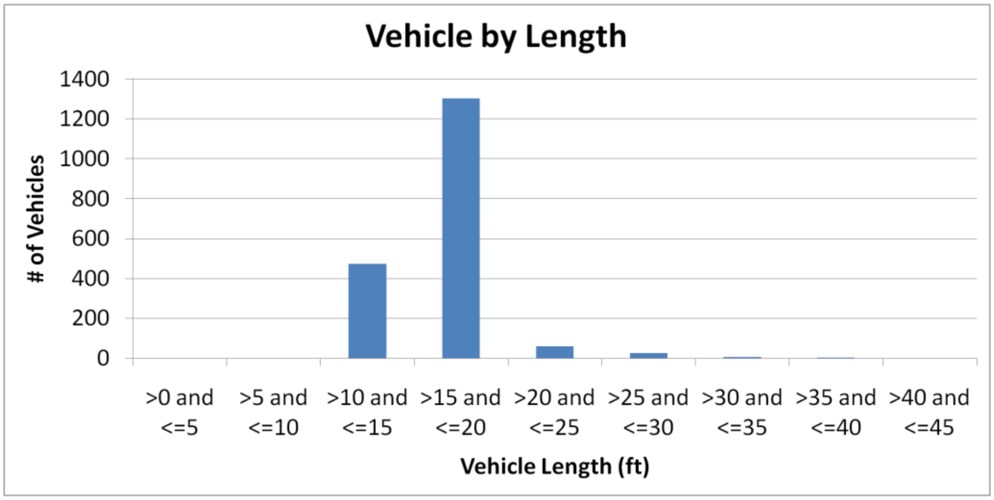



**Figure 3-2: Vehicles by length.**

## 3.5 Records and Parameter Ranges

In the current NGSIM dataset, there are 873889 numbers of records available for both the datasets. All the vehicles in the dataset are tracked through the system with their information recorded every 10$^{th}$ of a second.

The data for the datasets were collected for 15 minute time periods for both the AM and PM slots. Because data were sampled in 0.10 sec intervals, the total number of frame IDs' for each 15 minute period is (15*60*10=) 9000.

A column in each dataset represents the total number of frames that each of the vehicles occupied on the dataset. The local Y is increasing for the vehicles that are traveling in the northbound direction, whereas local Y is decreasing for the southbound vehicles.

The velocity of the vehicle's ranges from 0~56 ft/sec for both datasets even though the posted speed limit of the corridor is 35mph (51 ft/sec). The acceleration/deceleration rate ranges from -12.27 ft/sec$^2$ to 12.27 ft/sec$^2$ for both the datasets.

Numbers in the signal indication field 1, 2, 3, 4, 5, and 6 denote Red, Green, Yellow, Green for Left, Yellow for Left and Red for Left, respectively.

## 3.6 Examples of Parameters Potentially Observed from Vehicle Trajectory Data

Data collected in a fashion similar to the NGSIM Atlanta dataset has the potential to describe a wide variety of parameters that can be obtained from the vehicle trajectory. These are end-of-green-gain, safety clearance distance, start-up response time, start-up lost time, response time, follow-up headway, velocity, and acceleration rates. Illustrations and their corresponding discussions are given in the following subsections.

### 3.6.1 End-of-Green-Gain

End-of-green-gain ($t_{eg}$) is defined as the duration of the interval between the end of the displayed green period and the end of the effective green period for a movement. This parameter is measured by observing the time the green terminates ($t_g$) and the time the last queued vehicle discharges after the green ($t_{lv}$). The parameter $t_{eg}$ is the difference ($t_{lv} – t_g$). This is used in signal timing and performance analysis to allow for additional departures after the end of the green period (Akcelik, 2004) and is illustrated in Figure 3-3.

Notice leader lines and the arrow in the figure. The leader lines define the beginning and ending of the end-of-green-gain. It should also be noted that the end-of-green-gain can be related to a more common term, clearance lost time ($t_{cl}$), using the equation, $t_{cl} = r - t_{eg}$.

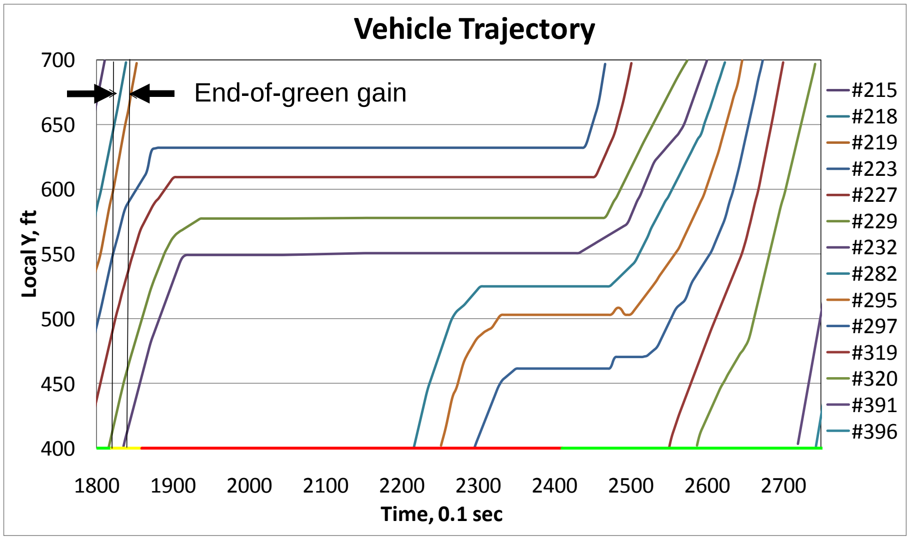



**Figure 3-3: Illustration of the end-of-green-gain.**

### 3.6.2 Safety Clearance Distance

Distance between the leader and the follower when they are stopped due to red is defined as safety distance. It can be illustrated as shown in Figure 3-4.

Note that vehicles 223 and 227 are both stopped on red. The safety clearance distance between them is (630-610=) 20ft. Similarly, the safety distance between vehicles 227 and 229 is (610-575=) 35 ft.

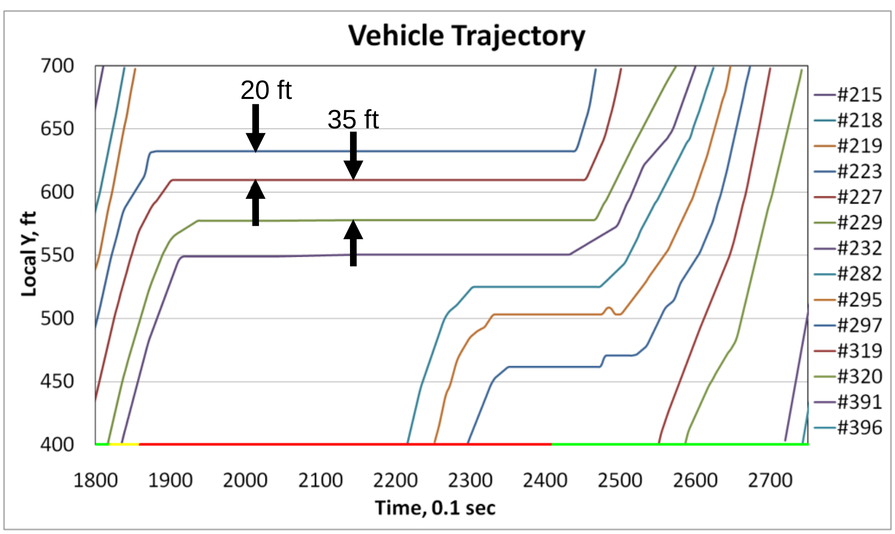



**Figure 3-4: Illustration of the safety clearance distance.**

### 3.6.3 Start-Up Response Time

Start-up response time specifies the first stopped vehicle's reaction time at the beginning of green. It is illustrated as shown in Figure 3-5. Vehicle 223 is the leader of the queue. At frame ID 2410 the green begins, but the vehicle starts to move at frame ID 2435, which implies that the start-up response time for this vehicle is (2435-2410=) 25 frames, which is (25*0.1=) 2.5 sec.

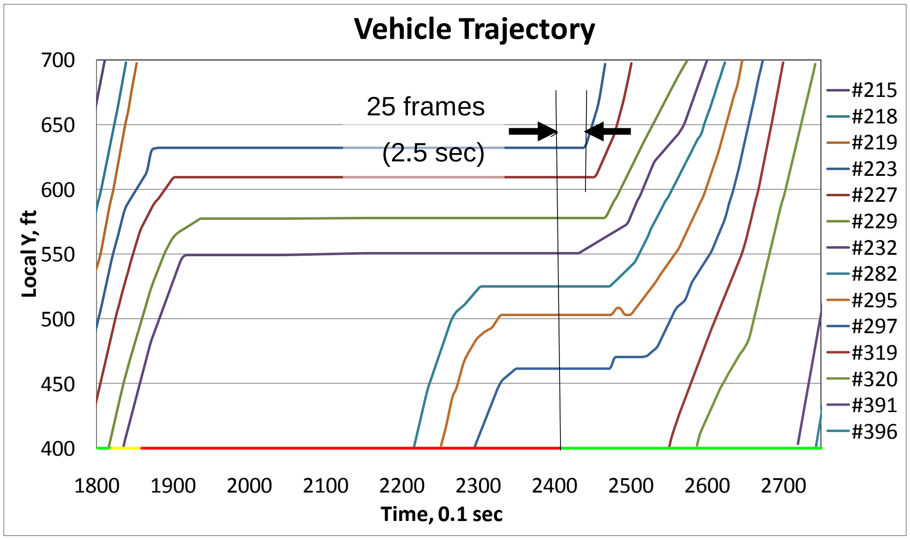



**Figure 3-5: Illustration of the start-up response.**

### 3.6.4 Start-Up Lost Time

According to Traffic Signal Timing Manual 2008, start-up lost time is the additional time, in seconds, consumed by the first few vehicles in a queue at a signalized intersection above and beyond the saturation headway, because of the need to react to the initiation of the green phase and to accelerate. Procedures for calculating start-up lost time from the NGSM dataset are as follows:

i. For the first vehicle in a queue, the headway was measured as the time difference between the start of green and the first vehicle's rear end crossing the stop bar.

ii. For the other vehicles in a queue, the headway was measured as the time difference between the preceding vehicle's rear end crossing the stop bar and the following vehicle's rear end crosses the stop bar.

iii. A vehicle's speed must be zero for a significant time when it was at the stop bar or near the stop bar. Besides, the video was also checked to make a cross check.

iv. Saturation headway is the average headways of the vehicles that are in queue starting from the 4th position of the queue to the last position of queue.

v. Start-up lost time was measured as the sum of excess amount of headways of the first three vehicles of the queue to the saturation headways.

The analysis of the start-up lost time in the northbound and southbound directions experienced by in-queue vehicles will be shown later in section 6.1.1.

### 3.6.5 Response Time

Response time, or reaction time, is the time between the initial movement of the leading vehicle and the initial movement of the following vehicle. In Figure 3-6: Illustration of the response timean illustration has been shown for response time. The leader vehicle 223 initially starts moving at frame ID 2435, whereas the immediate follower vehicle 227 starts moving at frame ID 2450. So, the response time for the immediate follower vehicle is ((2450-2435)*0.1=) 1.5 sec.

Again, the second vehicle of the queue starts moving at a frame ID of 2470. As a result, the response time for this vehicle is ((2470-2450)*0.1=) 2.0 sec.

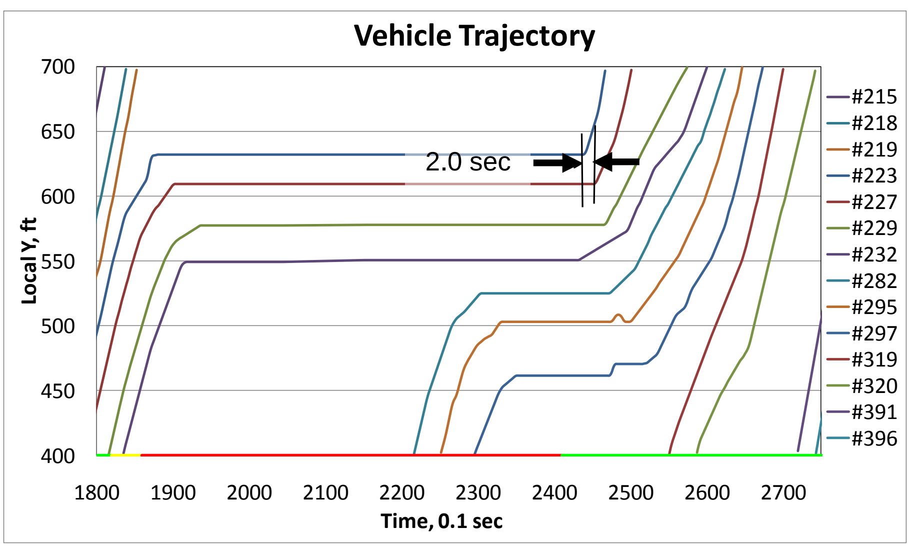



**Figure 3-6: Illustration of the response time.**

### 3.6.6 Follow-up Headway

Follow-up headway is the time between the moments when the leading and following vehicles discharge over the stop bar. When observed during queue discharge, follow-up headway is synonymous with saturation flow rate and is illustrated as shown in Figure 3-7. In this figure, vehicles 223 and 227 are crossing the stop bar at a follow-up headway of ((2480-2450)*0.1=)) 3.0 sec.

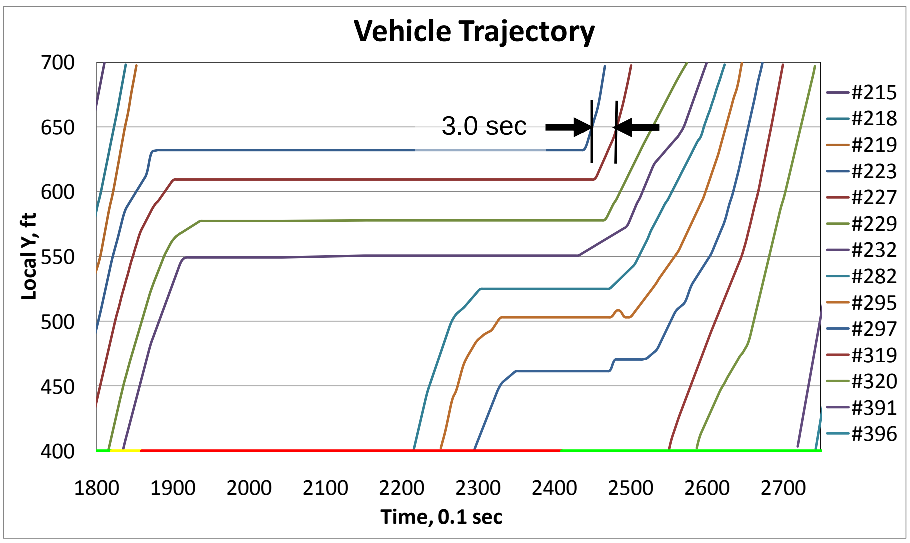



**Figure 3-7: Illustration of the follow-up headway.**

### 3.6.7 Velocity

The slope of the vehicle trajectory plot represents the velocity of the vehicle at that instant. An illustration of this slope is shown in Figure 3-8, where a line tangent to the vehicle 229 trajectory was drawn beginning at frame ID 1820.

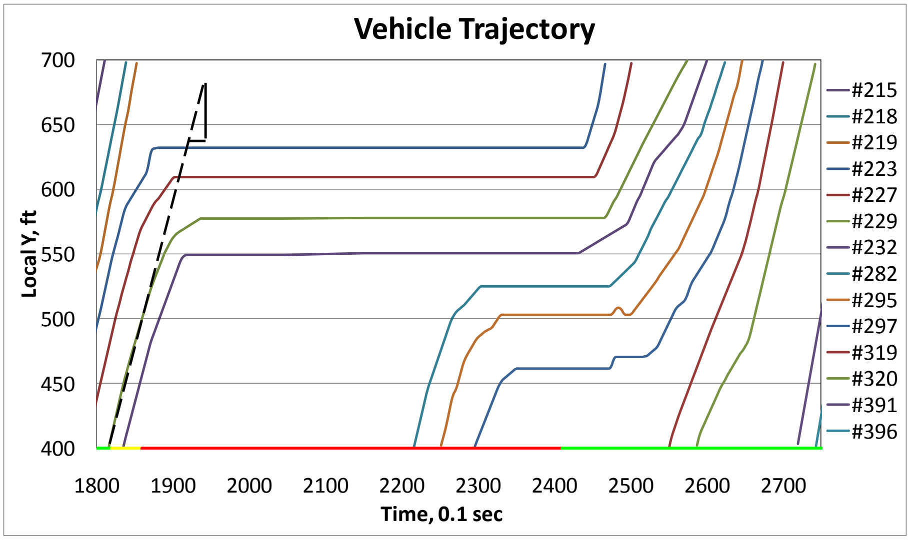



**Figure 3-8: Illustration of the velocity.**

### 3.6.8 Acceleration and Deceleration Rates

The second derivative of the slope of vehicle trajectory plot represents the acceleration/deceleration rate of the vehicle, or the slope of a speed curve, at a particular time. The illustration of acceleration is shown in Figure 3-9 for two vehicles, where vehicle 619 is the leading vehicle. The figure shows two data types, trajectory data and speed data, to more clearly identify acceleration periods. When the speeds change, some acceleration is occurring, and the acceleration at that time is the slope of the speed curve. At frame ID 4200, the vehicles discharge, accelerating in the midst of a discharging queue. In this case, it is interesting to note the correlation between the changes in speeds of the two vehicles and the step-like nature of the speed curve.

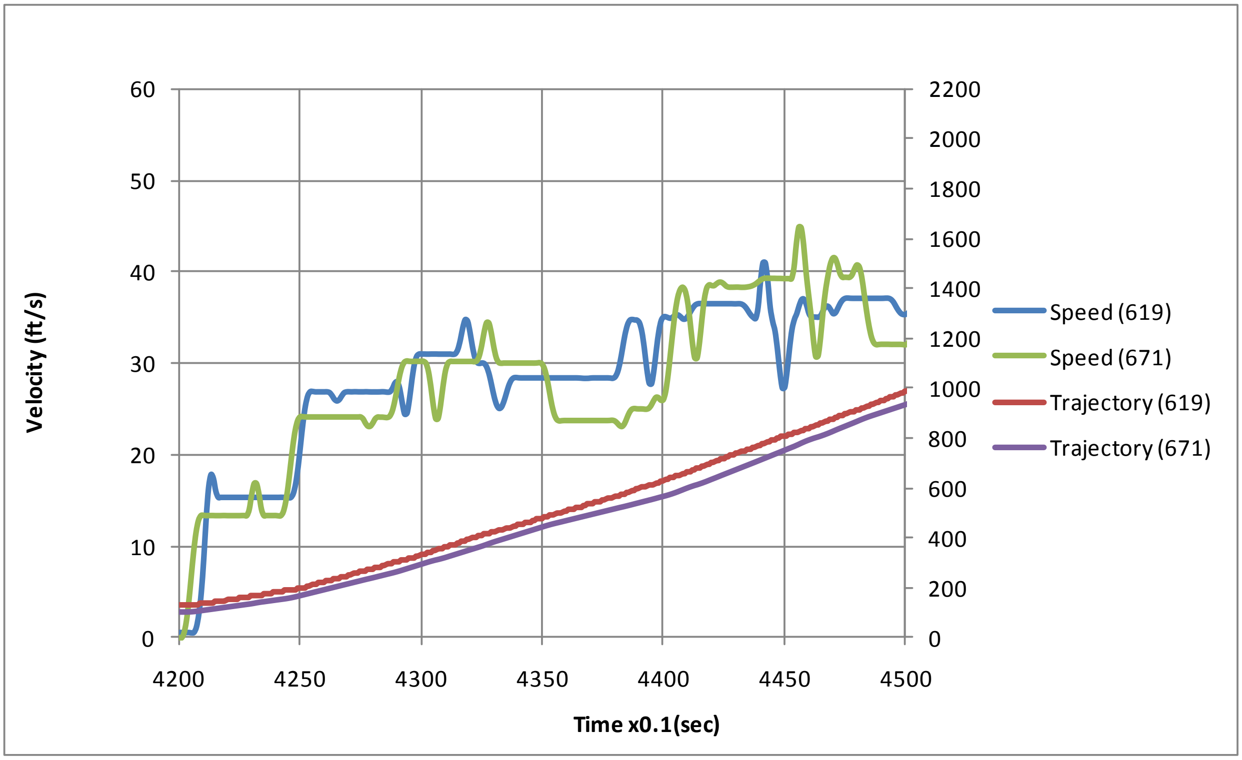



**Figure 3-9: Illustration of the acceleration.**

### 3.6.9 Summary of Car-Following Model Parameters Obtainable from the NGSIM Vehicle Trajectory Data

Table 3-4 lists parameters used in different car-following models, their definitions, and whether it is possible or not to obtain the parameter from the NGSIM vehicle trajectory dataset. Parameters for which Weidemann is cited are available in VISSIM and the remaining parameters are generally available in other traffic simulation packages.

**Table 3-4: Car Following Model's Parameters**

| Parameter | Description | From the Dataset |
|---|---|---|
| $x_{n-1}(t)$ or $x_{n-1}$ | **Position of the preceding vehicle at time t** | Yes |
| $x_n(t)$or $x_n$ | **Position of the following vehicle at time t; or the distance from the stop line when the vehicle reacts to the control** | Yes |
| $\Delta x$ | **Space headway, the distance d between two vehicles** | Yes |
| $\Delta v$ | **The relative speeds between the $n^{th}$ and n-1 vehicle, assessed at an earlier time $t-\tau$, where $\tau$ is the driver reaction time** | Yes |
| $v_n(t+\tau)$ or $v_n$ | **The speed for vehicle n during the time interval $(t+\tau)$; or speed for vehicle n.** | Yes |
| $v_n^{desired}$ | **Desired speed for vehicle n** | No |
| $v_n^a(t+\tau)$ | **The maximum speed a vehicle can accelerate to during one time step** | No |
| $v_n^b(t+\tau)$ | **The maximum safe speed for vehicle n with respect to the vehicle in front at time t** | No |
| $a_n$ | **The acceleration of vehicle n implemented at time t by a driver** | Yes |
| $a_n^+$ or $a_n^{max}$ | **Maximum acceleration rate for vehicle n, a function of vehicle type and current speed; or maximum desired acceleration for vehicle n** | No |
| $a_n^-$ | **Normal deceleration rate for vehicle n, a function of vehicle type and current speed** | Yes |
| $d_n^{max}$ or $\hat{d}_{n-1}$ | **Maximum desired deceleration for vehicle n; or estimation of maximum deceleration desired by vehicle n-1** | No |
| $L_{n-1}$ | **Length of the preceding vehicle** | Yes |
| $s_{n-1}$ | **The effective length of a vehicle, consists of the vehicles length and the user specified parameter min distance between vehicles** | Yes |
| T | **Perception reaction time** | Yes |
| $T_D$ | **The time gap desirable for driver to complete driving maneuver** | No |
| $T_r$ | **The risky time gap accepted by driver to complete driving maneuver** | No |
| $y_n$ | **The distance required for a vehicle with speed $v_n$ to decelerate to a stop by applying the normal deceleration rate $a_n^-$** | No |
| $y_{min}$ | **A lower bound of the normal stopping distance** | No |
| $b^*$ | **Assumed braking rate of the vehicle ahead** | No |
| $B$ | **Maximum braking rate** | No |
| $S$ | **Jam headway** | Yes |
| $h_{upper}$, $h_{lowerr}$ | **Thresholds to determine which regimes of driver behavior should be used** | No |
| $\alpha^\pm, \beta^\pm, \gamma^\pm$ | **Model parameters. The parameters $\alpha^+, \beta^+, \gamma^+$ are used if $v_n \leq v_{n-1}$ and $\alpha^-, \beta^-, \gamma^-$ if $v_n > v_{n-1}$** | No |
| $k_{PTN}$, $k_{PTP}$ and $f_x$ | **Model parameters** | No |
| $\epsilon$ | **Noise amplitude** | No |

| Parameter | Description | From the Dataset |
|---|---|---|
| *η* or *z* | **Random number in [0, 1]; or a value of range [0, 1] which is normal distributed around 0.5 with a standard deviation of 0.15 (VISSIM, 2005)** | No |
| | **Position of the queue** | Yes |
| *Ax* | **Average safety distance defines the average desired distance between stopped cars. It has no variation in Wiedemann 99, but a fixed variation of ± 1m in Wiedemann 74** | Yes |
| *bx_add* | **Additive part of desired safety distance (VISSIM, 2005)** | No |
| *bx_mult* | **Part of desired safety distance (VISSIM, 2005)** | No |
| CC1 | **Means the Headway time which is the time (in sec) that a driver wants to keep. The higher the value, the more cautious the driver is. (Wiedemann 99)** | No |
| CC2 | **'Following' variation restricts the longitudinal oscillation or how much more distance than the desired safety distance a driver allows before he intentionally moves closer to the car in front. If this value is set to e.g. 10m, the following process results in distances between *dx_safe* and *dx_safe* + 10m. The default value is 4.0m which results in a quite stable following process. (VISSIM, 2005)** | Yes |
| CC3 | ***Threshold for entering 'Following'* controls the start of the deceleration process, i.e. when a driver recognizes a preceding slower vehicle. In other words, it defines how many seconds before reaching the safety distance the driver starts to decelerate. (VISSIM, 2005)** | Yes |
| CC4 and CC5 | ***Following' thresholds* control the speed differences during the 'Following' state. Smaller values result in a more sensitive reaction of drivers to accelerations or decelerations of the preceding car, i.e. the vehicles are more tightly coupled. CC4 is used for negative and CC5 for positive speed differences. The default values result in a fairly tight restriction of the following process. (VISSIM, 2005)** | Yes |
| | **Position of the queue** | Yes |
| *Ax* | **Average standstill distance defines the average desired distance between stopped cars. It has no variation in Wiedemann 99, but a fixed variation of ± 1m in (VISSIM, 2005)** | Yes |
| *bx_add* | **Additive part of desired safety distance (VISSIM, 2005)** | No |
| *bx_mult* | **Part of desired safety distance (Wiedemann 74)** | No |
| CC6 | ***Speed dependency of oscillation;* Influence of distance on speed oscillation while in following process. If set to 0 the speed oscillation is independent of the distance to the preceding vehicle. Larger values lead to a greater speed oscillation with increasing distance. (VISSIM, 2005)** | No |
| CC7 | ***Oscillation acceleration;* Actual acceleration during the oscillation process. (VISSIM, 2005)** | No |
| CC8 | ***Standstill acceleration;* Desired acceleration when starting from standstill (limited by maximum acceleration defined within the acceleration curves) (VISSIM, 2005)** | Yes |
| CC9 | ***Acceleration at 80 km/h;* Desired acceleration at 80 km/h (limited by maximum acceleration defined within the acceleration curves). (VISSIM, 2005)** | Yes |

| Parameter | Description | From the Dataset |
|---|---|---|
| $\Delta x$ and $v_n(t)$ | **They are unit less quantities that can be obtained by dividing the headway and speed with a constant term. This constant term is allowed to vary from 6.5 to 15, which is basically the minimum distance between adjacent vehicles.** | Yes |
| $\alpha$ | **The sensitivity factor α is allowed to vary from 0.1 to 1** | No |
| $D_n$ | **An estimate of a constant distance measure varies from 1 to 10 m** | No |
| $T$ | **Time step (NETSIM, PRESIM and INTRAS, CORSIM)** | No |
| K | **The buffer space between the vehicles (CORSIM)** | No |

# 4 INTEGRATING DATASET COMPONENTS

Several database tables were created in Microsoft Access, and they can be connected to each other when querying the database. The query extracts the requested data from each of the tables using indices they hold in common. Specific criteria related to the query can be set by the user, based on any of the fields contained in the tables. It then produces the data needed for its user in a tabulated format based on the desired query criteria.

In order to retrieve information in an efficient way from trajectory data, the data were imported to a MS Access 2007 database from raw text files. The database, “NGSIM_ATL_DBMS.accdb”, can be downloaded from http://www.webs1.uidaho.edu/ngsim/. This database contains three tables of trajectories, signal timing, and stop bar locations. There are two trajectory tables, Trajectories1 and Trajectories2, for the time periods of 12:45 p.m. to 1:00 p.m. and 4:00 p.m. to 4:15 p.m., respectively. The primary key for both tables is a combination of the field’s vehicle ID and frame ID. No primary keys were used for the signal timings or stop bar locations tables. A brief description of the structures for these tables, signal timing and stop bar location, can be found in Table 4-1 and Table 4-2, respectively.

**Table 4-1: Data Dictionary of the Table Titled "Signal_Timing"**

| Column | Field Name | Data Type | Description |
|---|---|---|---|
| 1 | Sets | Number | The code for the datasets. 1 represents the 12:45 data and 2 is for the 4:00 data. |
| 2 | Intersection | Number | The number of connected intersections. It was coded from 1 to 5 without 4 since the 4th intersection was not a signalized intersection. |
| 3 | Direction | Number | The traffic direction the signal controlled. It was coded as: 1: EB; 2: NB; 3: WB; 4: SB. |
| 4 | Indication | Number | The indication of the connected signal. It was coded as: 1: Red; 2: Green; 3: Yellow; 4: Green for left; 5: Yellow for left; 6: Red for left. |
| 5 | Cycle | Number | The identification of cycle which began with yellow. 0 represents the incomplete cycle at the beginning. |
| 6 | Begin At | Number | The frame ID when the indication in column 4 began. |

**Table 4-2: Data Dictionary of the Table Titled "Stop_Bar"**

| Column | Field Name | Data Type | Description |
|---|---|---|---|
| 1 | Sets | Number | The code for the datasets. 1 represents the 12:45 data and 2 is for the 4:00 data. |
| 2 | Intersection | Number | The number of connected intersection. It was coded from 1 to 5 without 4 since the 4th intersection was not a signalized intersection. |
| 3 | Direction | Number | The traffic direction the signal controlled. It was coded as: 1: EB; 2: NB; 3: WB; 4: SB. |
| 4 | Lane | Number | The identification of the connected lane. It was coded as: 1: For the most inner, 2: For the outer lane, 11: For the left turn lane. |
| 5 | Global_X | Number | X coordinate of the center of the stop bar based on NAD83. |
| 6 | Global_Y | Number | Y coordinate of the center of the stop bar based on NAD83. |
| 7 | Local_x | Number | Lateral (X) coordinate of the center of the stop bar - perpendicular to the median of the Peachtree St. |
| 8 | Local_y | Number | Longitudinal (Y) coordinate of the center of the stop bar along the median of the Peachtree St. The start point is at the southern boundary of the study area. |

Since the trajectory tables were not normalized, advanced queries (i.e. coded by the SQL statements directly) are necessary for facilitating the data preparation. In an Access database, a query allows the user to present a question to the database by specifying specific criteria. Queries have two basic views: Design view and Datasheet view. In the Design view, the user specifies which tables or queries are going to be used, and the criteria that records has to meet in order to create a new query. In the Query Datasheet view, the user sees the records found to meet the given criteria.

SELECT queries in Access are the most useful type of queries to analyze data. A SELECT query can be used to select certain data from table(s) as well as the tables from other stored queries. It basically filters and sorts the data. In addition, SELECT queries can perform simple calculations such as counting, summing and averaging. An easy way to create a select query in the Access database is to use Simple Query Wizard. The Simple Query Wizard allows selecting fields from tables or other queries to include in another query. This Access application is useful for eliminating extra fields on the new query. Advanced queries have to be accomplished in Design view or SQL (Structured Query Language) view since the capability of the Simple Query Wizard is too limited for complex queries.

Traffic dynamics described in part by signal operations can be queried based on the integration of the signal and trajectory tables in Access. For instance, a SELECT query can select and display a vehicle's location relative to the stop bar at the moment the downstream signal turns yellow. Visual Basic programming skills would be helpful to process a query that must be applied repeatedly because the major tables are not normalized. For a complicated query, however, it is recommended to have a manual verification of the results to address measurement errors in the data. If a researcher is not familiar with Microsoft Access 2007, six step-by-step instructions are provided in Appendix A.

# 5 DATASET CORRECTIONS

To use the dataset in a correct format, some corrections were made in the current NGSIM Atlanta dataset. These are stop bar location, phase status time synchronization, unnumbered vehicles, vehicle location correction, and spurious stops. The original dataset had some discrepancies in the above mentioned areas. So, the team members found it necessary to correct the NGSIM Atlanta dataset. To do this, they had to match the video and the current dataset followed by some corrective measures. Each of the problems and the steps that were taken to resolve it are discussed in the following sub sections.

## 5.1 Dataset Continuity

A large initial investment was required to acquire the traffic data in the field and process it to a useful format. To preserve the initial integrity of this data, all changes to the data are made to a copy of the data.

## 5.2 Stop Bar Locations

In the original NGSIM Atlanta dataset, two sets of stop bar locations were provided for two different datasets. For both of the datasets, video data were collected for the same four intersections. The reason for the variation in the stop bar location for the two datasets may be a slight change in camera positions. As a result, for the two datasets, local Y for stop bar location was different. The original stop bar location for the two datasets can be seen in Table 5–1 and Table 5–2.

The stop bar locations were not sufficiently accurate to use as reference points for vehicle events. For example, for Atlanta dataset1, intersection1, and the northbound direction inside lane (A11N1), the stop bar location was provided at 112.7. From the trajectory dataset, it was seen that vehicle 120 was located (local Y) at 133.845 at frame ID 1532. According to the original stop bar location table, this means that the vehicle had already passed the stop bar. But from the video it was seen at frame ID 1532, that the vehicle had not crossed the stop bar yet (see Figure 5-1), indicating an incorrect stop bar location.


**Figure 5-1: Stop bar location correction from video.**

To reconcile the problem, 10 random vehicles were used for each of the intersections. For each of the 10 random vehicles, the following steps were taken to correct the stop bar location:

1. For each stop bar, a vehicle was chosen that crossed it.
2. The vehicle's (from step 1) trajectory data was accessed from the DBMS.
3. The vehicle (from step 1) was observed in the video and paused when the front of the vehicle arrived at the stop bar (see Figure 5-1).
4. At the paused time (from step 3), the video frame ID was noted.
5. The corresponding time frame (from step 4) was found in the vehicle's trajectory and the local Y associated with this frame ID was used as a sample stop bar location measurement.

Step 1 to step 5 were followed for each of the 10 randomly selected vehicles for each intersection of each dataset. After getting the location of 10 vehicles from the video, these 10 locations were averaged and used as the new updated stop bar location. The local Y variations among the 10 updated stop bar locations for Atlanta dataset1 were as follows:

- NB
  - Intersection 1: 10.85 (lane1) and 19.0 (lane2)
  - Intersection 2: 17.49 (lane1) and 11.26 (lane2)
  - Intersection 3: 12.35 (lane1) and 21.08 (lane2)
  - Intersection 5: 15.797 (lane1) and 17.59 (lane2)
- SB
  - Intersection 1: 5.73 (lane1) and 9.24 (lane2)
  - Intersection 2: 13.70 (lane1) and 8.99 (lane2)
  - Intersection 3: 14.37 (lane1) and 7.08 (lane2)

And for Atlanta dataset2, the variations of local Y were as follows:

- NB:
  - Intersection 1: 3.343 (lane1) and 6.13 (lane2)
  - Intersection 2: 3.17 (lane1) and 11.92 (lane2)
  - Intersection 3: 8.947 (lane1) and 4.59 (lane2)
  - Intersection 5: 46.62 (lane1) and 7.95 (lane2)
- SB:
  - Intersection 1: 7.34 (lane1) and 8.75 (lane2)

- Intersection 2: 4.07 (lane1) and 2.91 (lane2)
- Intersection 3: 3.43 (lane1) and 4.61 (lane2)

The corrected stop bar location and the original stop bar location for each of the intersections are shown in Table 5-1 and Table 5-2 for Atlanta dataset 1 and Atlanta dataset 2, respectively.

**Table 5-1: Stop Bar Correction for Atlanta Dataset 1**

| Stop Bar Location Provided (Local Y) | Corrected Stop Bar Location (Local Y) | Stop Bar Location Provided (Local Y) | Corrected Stop Bar Location (Local Y) |
|---|---|---|---|
| A11N1[1] | | A11S1 | |
| 123.5 | 112.7 | 234.4 | 225.615 |
| A11N2 | | A11S2 | |
| 125.88 | 112.46 | 234.34 | 226.28 |
| A12N1 | | A12S1 | |
| 630.5 | 623.088 | 760.1 | 764.65 |
| A12N2 | | A12S2 | |
| 630.3 | 624.09 | 760.1 | 764.65 |
| A13N1 | | A13S1 | |
| 1165.53 | 1158.10 | 1226.1 | 1237.4 |
| A13N2 | | A13S2 | |
| 1168.61 | 1160.31 | 1230.73 | 1237.89 |
| A15N1 | | | |
| 1955.45 | 1948.6 | | |
| A15N2 | | | |
| 1955.45 | 1945.68 | | |

[1]A11N1 represents Atlanta Dataset1, intersection1, Northbound(N),Inside Lane(1)
A12S2 represents Atlanta Dataset2, Intersection2, Southbound(S), Outside Lane( 2)

**Table 5-2: Stop Bar Correction for Atlanta Dataset 2**

| Stop Bar Location Provided (Local Y) | Corrected Stop Bar Location (Local Y) | Stop Bar Location Provided (Local Y) | Corrected Stop Bar Location (Local Y) |
|---|---|---|---|
| A21N1 | | A21S1 | |
| 140.81 | 135.5 | 251.7 | 236.55 |
| A21N2 | | A21S2 | |
| 143.17 | 138.99 | 251.64 | 236.29 |
| A22N1 | | A22S1 | |
| 647.84 | 647.75 | 774.4 | 779.25 |
| A22N2 | | A22S2 | |
| 647.61 | 646.04 | 777.4 | 779.39 |
| A23N1 | | A23S1 | |
| 1182.83 | 1183.62 | 1243.4 | 1244.93 |
| A23N2 | | A23S2 | |
| 1185.90 | 1185.70 | 1248.02 | 1248.77 |
| A25N1 | | | |
| 1972.74 | 1966.63 | | |
| A25N2 | | | |
| 1966.71 | 1966.90 | | |

## 5.3 Phase Status Time Synchronization

In the NGSIM dataset, for each intersection, two video sets are available. From one video, signal data can be acquired and from another video (aerial video), traffic data can be observed. These two videos were not synchronized for these datasets. For example, it was seen from the signal video at one time that a signal was red. Whereas from the aerial video for the same frame ID, it was seen that vehicles were moving which indicated that these two videos were not synchronized.

The beginning of green times needed to be adjusted for closer synchronization with the vehicle trajectory data. To do this, heavy vehicles, such as buses, seen in both videos, were selected (see Figure 5-2). Figure 5-2 shows the aerial video and the signal data video side-by-side, for a time when a bus (vehicle 323) is in both videos. When the vehicle was judged to be in the same location in both videos, the time frames from both videos were recorded and used to synchronize the two videos.


**Figure 5-2: Snapshot for phase time synchronization (Atlanta dataset2, intersection 5, EB LT lane).**

Project team members repeated this process over time and for each intersection to arrive at an average index between the two videos. Table 5-3 shows the time synchronization for the two intersections (intersection 3 and intersection 5) of Atlanta dataset2.

**Table 5-3: Phase Time Synchronization**

| Vehicle type | Vehicle ID | Movement | Direction | Signal Timing Video Frame ID | Aerial View Video Frame ID |
|---|---|---|---|---|---|
| Intersection 5 Dataset 2 | | | | | |
| City Bus | 323 | LT | EB | 1988 | 1989 |
| Intersection 3 Dataset 2 | | | | | |
| White Utility Truck | 1087 | TH | NB | 6629 | 6630 |

As shown in Figure 5-2, only heavy vehicles can be seen in the signal video, and there are very few heavy vehicles in the corridor. As a result, there were few data points for synchronization limiting the precision of this correction.

In the future, all cameras should transmit to a server so that all of the cameras will start timing and recording at the same time. On the other hand, if the synchronization is done manually, a stop watch can be used as the common clock between cameras by briefly showing its display in the video field of view.

## 5.4 Unnumbered Vehicles

Occasionally, the vehicles in the traffic video were not numbered. In the Atlanta dataset, some unnumbered vehicles were observed at the beginning and at the end of the data collection period. In Figure 5-3, it can be seen that there are two unnumbered vehicles in front of a numbered vehicle (vehicle 15). This situation is called "Preceding U". In Figure 5-4, it can be seen that there is one unnumbered vehicle following a numbered vehicle (vehicle 42). This situation is denoted as "Following U".

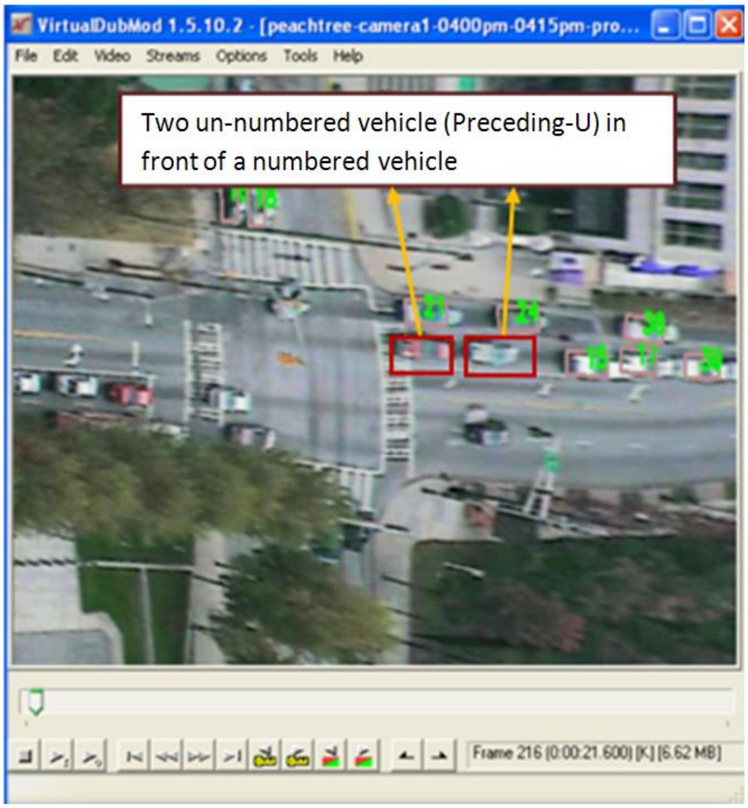



**Figure 5-3: Snapshot for unnumbered vehicle (Preceding U).**


**Figure 5-4: Snapshot for unnumbered vehicle (Following U).**

An unnumbered vehicle preceding or following a numbered vehicle causes serious database discrepancies when researching vehicle operations sensitive to neighboring vehicle positions and operations. To correct this information and communicate it more readily, two fields, i.e. "Preceding" and "Following", were added to the trajectory datasets to record the vehicle ID of unnumbered vehicles leading or following numbered vehicles.

The steps for processing unnumbered vehicles are listed as follows:

1. Watch the videos for the first 1000 frame IDs and the part after frame ID 9000 in order to find whether there were unnumbered vehicles leading (or following) a numbered vehicle. Unnumbered vehicles were not observed during other periods.
2. Given a vehicle from step 1, note the vehicle ID of the numbered vehicle and the frame ID.
3. Set the quantity (how many unnumbered vehicles) of unnumbered vehicles to the added column "Preceding U" for leading or "Following U" for following into the trajectory table for the frame ID from step 2.

Sample output are shown in Table 5-4, where the last two columns are the representation of how many unnumbered vehicles exist before or after one numbered vehicle. In the column "Preceding U", "1" represents that there is only one (1) unnumbered vehicle before vehicle 53.

**Table 5-4: Unnumbered Vehicle**

| Vehicle ID | Frame ID | Total Frames | Global Time | Local X | Local Y | Preceding U | Following U |
|---|---|---|---|---|---|---|---|
| 53 | 316 | 1248 | 1163050200 | -4.81 | 842.299 | 1 | |
| 53 | 317 | 1248 | 1163050300 | -4.886 | 839.894 | 1 | |
| 53 | 318 | 1248 | 1163050400 | -4.993 | 837.708 | 1 | |
| 53 | 319 | 1248 | 1163050500 | -5.063 | 835.794 | 1 | |
| 53 | 320 | 1248 | 1163050600 | -5.13 | 834.193 | 1 | |
| 53 | 321 | 1248 | 1163050700 | -5.159 | 832.865 | 1 | |
| 53 | 322 | 1248 | 1163050800 | -5.219 | 831.756 | 1 | |
| 53 | 323 | 1248 | 1163050900 | -5.282 | 830.836 | 1 | |
| 53 | 324 | 1248 | 1163051000 | -5.345 | 829.888 | 1 | |
| 53 | 325 | 1248 | 1163051100 | -5.408 | 828.914 | 1 | |
| 53 | 326 | 1248 | 1163051200 | -5.471 | 827.946 | 1 | |
| 53 | 327 | 1248 | 1163051300 | -5.534 | 826.977 | 1 | |
| 53 | 328 | 1248 | 1163051400 | -5.597 | 826.009 | 1 | |
| 53 | 329 | 1248 | 1163051500 | -5.66 | 825.04 | 1 | |
| 53 | 330 | 1248 | 1163051600 | -5.723 | 824.071 | 1 | |
| 53 | 331 | 1248 | 1163051700 | -5.787 | 823.103 | 1 | |
| 53 | 332 | 1248 | 1163051800 | -5.849 | 822.135 | 1 | |
| 53 | 333 | 1248 | 1163051900 | -5.912 | 821.166 | 1 | |
| 53 | 334 | 1248 | 1163052000 | -5.975 | 820.197 | 1 | |
| 53 | 335 | 1248 | 1163052100 | -6.038 | 819.226 | 1 | |
| 53 | 336 | 1248 | 1163052200 | -6.101 | 818.254 | 1 | |
| 53 | 337 | 1248 | 1163052300 | -6.165 | 817.29 | 1 | |
| 53 | 338 | 1248 | 1163052400 | -6.225 | 816.345 | 1 | |

In the 1st dataset (12:45 -13:00 hrs), the total records (total number of rows) in the DBMS are 391,646, and the total number of recorded vehicles is 1,115. Statistics for the unnumbered vehicles for the 1st dataset are listed as follows:

**Records:**

The number of records requiring unnumbered vehicle revisions are listed below and organized by whether the unnumbered vehicles are leading or following the numbered vehicle and by how many unnumbered vehicles there are:

- The records with at least one unnumbered leading vehicle or following: 4216
- The records with at least one unnumbered leading vehicle: 3664
- The records with at least one unnumbered vehicle following: 1432
- The records with at least one unnumbered vehicle leading and following: 880

**Recorded vehicles:**

Statistics relating the number of vehicles that were near unnumbered vehicles (preceding, following, or both) are listed below:

- The number of vehicles affected by at least one unnumbered leading or following vehicle: 14
- The number of vehicles affected by at least one unnumbered leading vehicle: 12
- The number of vehicles affected by at least one unnumbered leading vehicle: 4
- The number of vehicles affected by at least one unnumbered leading and following vehicle: 1

In the future, if high resolution traffic data are collected using video, care should be taken during data extraction. A method can be followed for numbering the vehicles so that no unnumbered vehicle exists:

- Count all the vehicles entering the system and currently in the system, and put numbered boxes on the top of each individual vehicle.

## 5.5 Vehicle location correction

At times, the database contained negative velocities reflecting instances where the database extraction procedure erroneously recorded a vehicle moving backwards. Figure 5-5 illustrates some examples indicated with a red circle. As illustrated, the changes in direction traveled are clearly incorrect for several reasons. First, it is unlikely that a driver would decide to make these maneuvers. Second, in all but one of these occurrences, it is physically impossible for a vehicle to decelerate and accelerate in this fashion. By checking the consistency of the fields of direction and local Y against the field of frame ID, negative velocity can be found in both Atlanta datasets. The logic to check whether there was a negative velocity is:

Let $y_i$ and $y_{i+1}$ represent local y at frame ID i and i+1.
If $y_i > y_{i+1}$ when the vehicle is northbound, then a negative speed occurred.
If $y_i < y_{i+1}$ when the vehicle is southbound, then a negative speed occurred.

For example, a negative velocity occurred when the local Y increased in the red circles in the following time-space diagram.

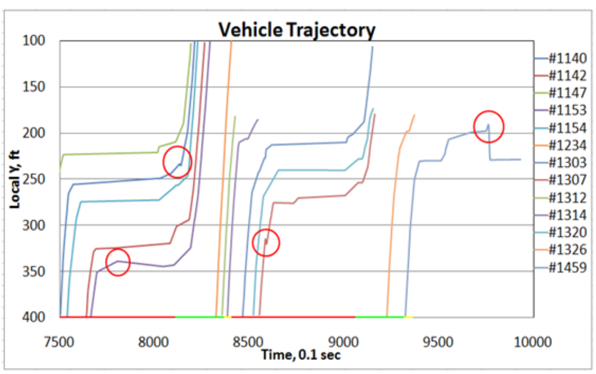



**Figure 5-5: Illustration for vehicle location before correction.**

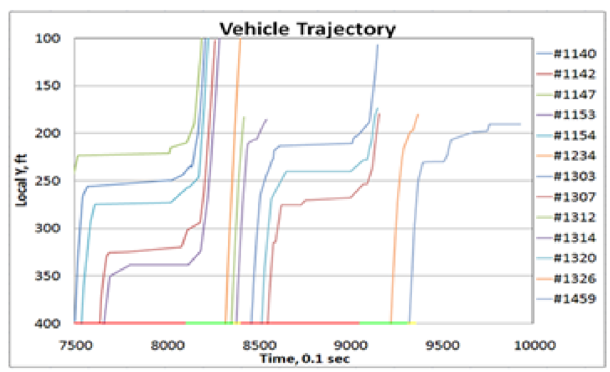



**Figure 5-6: Illustration of vehicle location after correction.**

Statistics for negative velocity occurrence are listed in Table 5-5, where one or more continuous records possessing a negative velocity is defined as an occurrence. The error percentages range from 2.48% to 7.01% for northbound and southbound traffic, respectively.

**Table 5-5: Statistics for Negative Velocity (Both Datasets)**

| | 1st Dataset | | 2nd Dataset | |
|---|---|---|---|---|
| | Northbound | Southbound | Northbound | Southbound |
| Error Counts | 7491 | 3703 | 14305 | 14438 |
| Total Records | 218823 | 149610 | 204003 | 253994 |
| Error Percentage | 3.42% | 2.48% | 7.01% | 5.68% |

By observing the videos, negative velocity may occur when the numbered box was pulled back suddenly to circumscribe the vehicle it represented. A box leading its corresponding vehicle resulted because the algorithm used to change the box location overestimated the vehicle's movement rate. Because negative velocity frequently occurred, manually correcting the error was not feasible. In order to keep the impact of modification on "real" trajectories as small as possible, the following logic was used to modify the data to eliminate negative velocity in the datasets:

- Let $y_i$ and $y_{i+1}$ represent local y at frame ID i and i+1.
- If $y_i > y_{i+1}$ when the vehicle is northbound, then $y_{i+1} = y_i$.
- If $y_i < y_{i+1}$ when the vehicle is southbound, then $y_{i+1} = y_i$.

After applying the aforementioned logic, it was found that over thirty thousand records were modified in the datasets; and the statistics of the modifications are listed in Table 5-6.

After modification for negative velocity, the trajectory diagram smoothens as shown in Figure 5-6. The modification percentages range from 17.01% to 32.95%. The results verified the argument that it is difficult to fix all negative velocity instances manually, because it occurred abundantly. In addition, no explanation for the error occurrence exists. The error percentages in Table 5-5 are much smaller than those shown in Table 5-6, because in many cases one occurrence required modifying more than one record. Figure 5-6: Illustration of vehicle location after correction shows the modified trajectories, where the modifications for negative velocity were made. Note that some negative velocity errors were not detected and still remain in the dataset, as illustrated by the persistence of the first error shown in Figure 5-5: Illustration for vehicle location before correction.

**Table 5-6: Statistics after Modification—Total Records Modified (Both Datasets)**

| | 1st Dataset | | 2nd Dataset | |
|---|---|---|---|---|
| | Northbound | Southbound | Northbound | Southbound |
| Modification Counts | 72110 | 34003 | 44469 | 43208 |
| Total Records | 218823 | 149610 | 204003 | 253994 |
| Modification Percentage | 32.95% | 22.73% | 21.8% | 17.01% |

A couple of snapshots are shown in Figure 5-7: Snapshots illustrating vehicle location correction to illustrate the problem. At frame ID 2881, the box is just above vehicle ID # 314. But the next frame ID # 2882 shows the box behind the vehicle, causing the negative speed. In the following frames, the box moves by itself towards the vehicle even though the vehicle has stopped. Finally, at frame ID 2902, the box fits exactly above the vehicle as it was in frame ID 2881.

Frame ID 2881 Frame ID 2882 Frame ID 2884

Frame ID 2885 Frame ID 2890 Frame ID 2902

**Figure 5-7: Snapshots illustrating vehicle location correction.**

The vehicle location corrections were needed to produce a cleaner dataset. The problem occurred due to miss allocating the box assigned to vehicle ID # 314. In this particular instance, the software would more easily and accurately place the box if vehicle 314 were not obscured by the trees. However, given that this error frequently occurs with unobscured views, the software algorithm for tracking vehicles with the box needs additional work.

## 5.6 Spurious stops

Instances were found where the video shows a vehicle stopping, but the dataset indicates that the vehicle did not stop. This error tended to occur while the vehicle is decelerating to a stop at the stop bar. No other evidence was found to gain more insight into this problem. To assess the prevalence of this problem, a random sample was taken of all of the vehicles tracked. Seventy vehicles were randomly selected and their data were compared to the video. Thirteen percent of the vehicles had discrepancies in their stop data, where the video shows they had stopped and the dataset indicated they had not. Speeds at which this mistake tended to occur were around 9 ft/s. Some snapshots are shown in Figure 5-8 to illustrate the problem. Table 5-7 shows that vehicle 11 did not stop from frame ID 499 to 529. But from Figure 5-8, it is clearly seen that the box was moving for each of these frame IDs, even though the vehicle was stopped.

**Table 5-7: Data for Vehicle ID #11**

| Vehicle ID | Frame ID | Local Y | Vehicle Velocity |
|---|---|---|---|
| 11 | 499 | 810.574 | 7.58 |
| 11 | 500 | 809.728 | 9.04 |
| 11 | 501 | 808.735 | 9.47 |
| 11 | 507 | 803.051 | 9.45 |
| 11 | 508 | 802.106 | 9.4 |
| 11 | 509 | 801.167 | 9.33 |
| 11 | 527 | 786.534 | 6.78 |
| 11 | 528 | 785.855 | 6.78 |
| 11 | 529 | 785.177 | 6.78 |


Frame ID 499 Frame ID 508 Frame ID 529

**Figure 5-8: Snapshots for the spurious stops.**

No steps were taken to correct the NGSIM dataset. Assuming the correct data for the trajectories of these vehicles lies in the video, there is no feasible method to extract the correct information other than reprocessing the video after calibrating parameters used to govern the vehicle tracking box's movements. However, to reduce the impacts of this dataset error, researchers should base their research on vehicle trajectories that were verified to be correct.

# 6 ANALYSIS

In this section, the beginning of green and end of green phenomena observed in the current NGSIM Atlanta dataset are illustrated and analyzed. Parameters mentioned in section 3.6 were targeted for analysis, but all were not available due to limitations in the traffic conditions present during the data collection periods. The dataset had the potential to describe or analyze the following parameters, which were mentioned in section 3.6 as examples of parameters potentially observed from vehicle trajectory data:

i. End-of-green-gain
ii. Safety clearance distance
iii. Start-up response time
iv. Start-up lost time
v. Response time
vi. Follow-up headway
vii. Velocity, acceleration, deceleration rates of vehicles

No end-of-green-gain data were analyzed, because the data were insufficient due to queues predominantly clearing before the green terminated. However, the probability of vehicles stopping or continuing at the onset of yellow was observed and included in this analysis.

## 6.1 Beginning of Green

The analysis of the beginning of green is an important aspect to study the driver behavior characteristics. In this section, the beginning of green is analyzed for the Atlanta dataset in three different fields such as: Saturation headway, start-up lost time, and start-up response time. The analyses and findings are reported in the following paragraphs:

### 6.1.1 Saturation Headway (Follow-up Headway) and Start-up Lost Time

Saturation headways and start-up lost time were measured in a fashion similar to that recommended in the HCM 2000. The procedure followed to assess the saturation headway and the start-up lost time was explained in section 3.6.4 and 3.6.6, respectively. One important variation from this is that the vehicle point of reference was the rear of the vehicle. This was necessary for two reasons. First, the first vehicle in queue frequently stopped on top of the stop bar, which results in negative recorded headway values for the corresponding vehicle. Second, no information was given to describe the location of any vehicle axles relative to the front or rear of the vehicles. Measurements were taken in through lanes only and represent ideal saturation headways, because measurements were isolated from heavy and turning vehicle effects. Illustrations of saturation headway and start-up lost time are shown in Figure 6-1 and Figure 6-2 for northbound and southbound direction.

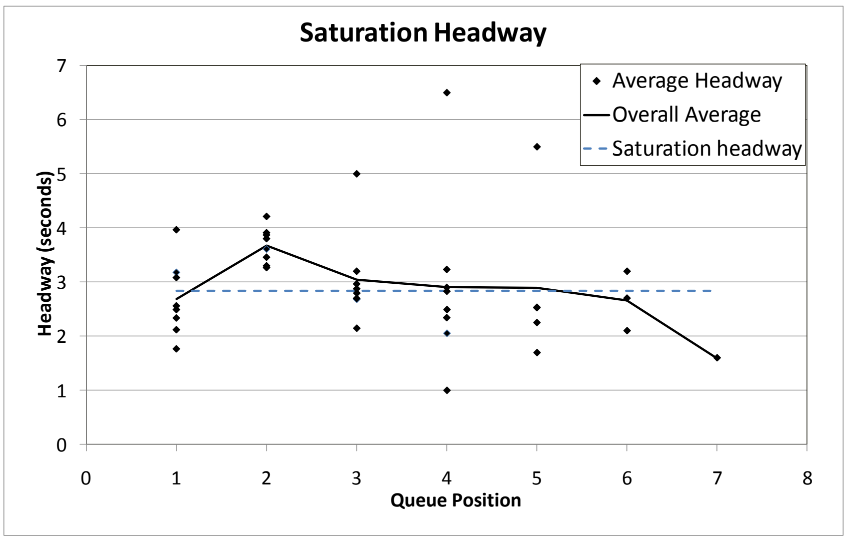



**Figure 6-1: Saturation headway and start-up lost time for northbound.**

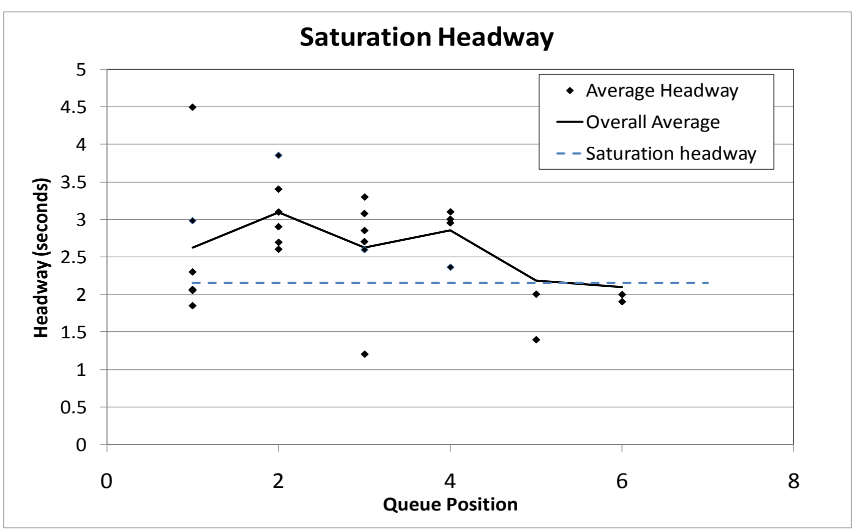



**Figure 6-2: Saturation headway and start-up lost time for southbound.**

As shown in Figure 6-1 and Figure 6-2, the saturation headway is reasonable with a value of 2.6 and 2.1 seconds for the northbound and southbound approaches, respectively. However, note that queue lengths were limited in this dataset as illustrated by the queue length frequency plotted in Figure 6-3. Figure 6-3 shows that only 8 headways were observed at queue position 5 and beyond. This is especially important for saturation headway, because it relies on average headways for these queue positions. The start-up lost time observed for the northbound and southbound directions were 0.6 and 0.7 seconds, respectively. These values are smaller than expected, which coincides with higher than normal saturation headways.

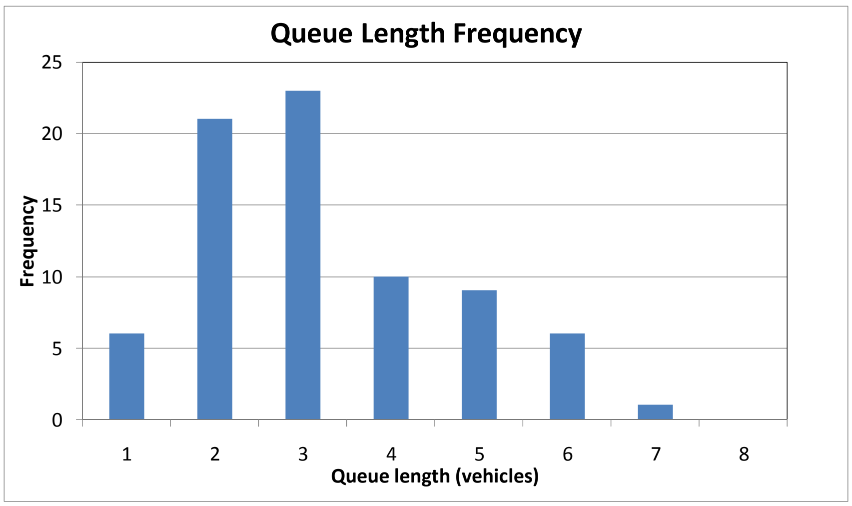



**Figure 6-3: Queue length frequency.**

### 6.1.2 Start-up Response Time

Response time is the time between the beginning of green and the first vehicle in the queue to start moving. It depends upon various factors such as: driver's age, driver's gender, driver's mentality, environmental conditions, etc. In this summary analysis, an investigation was performed to find the response time for vehicles in the NGSIM Atlanta dataset.

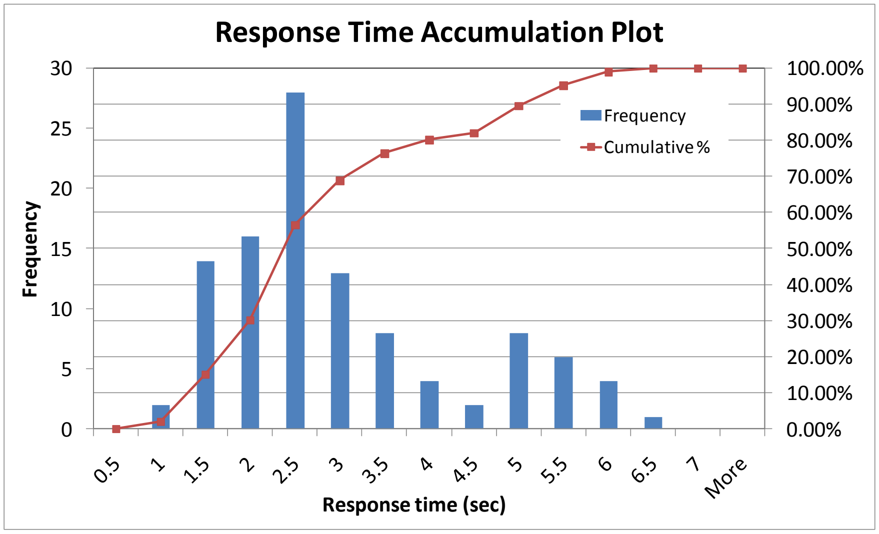



**Figure 6-4: Response time.**

At first, the lead vehicle of every queue was investigated to ensure they were in the front of the queue. To find out whether the vehicle is in front of the queue, an inventory was taken of vehicles in queue, describing the positions of every vehicle according to their vehicle ID and queue positions. Cross checks were made watching the videos.

A query was made in MS Access to find the green begin time for every queue for a particular intersection.

The response time was then calculated as the time difference between the beginning of green and the time the first vehicle in queue starts moving.

A cumulative frequency plot was drawn for all the response times of the vehicles for both directions, as shown in Figure 6-4. Here it clearly shows that the most frequent category of the vehicles response time is about 2.5 sec and 80% of the response times are equal to or less than 4.5 sec.

## 6.2 End of green

When a phase transition occurs from green to yellow, the drivers approaching the signalized intersection must make a decision to either stop or proceed through the intersection. The probability of stopping or going through is an important variable used to determine the dilemma zone or indecision zone. Assessing the dilemma zone is necessary in order to locate detectors and determine other signal control parameters.

At first, a query was made in the MS Access DBMS for the vehicles that were within a distance of 500 ft upstream of the stop bar for the onset of yellow. The query was built as a function of the signal timing, trajectory, and stop bar tables. The total number of vehicles found was 302. Figure 6-5 shows the percentage of stopping and non-stopping vehicles at the onset of yellow, where the percentage not stopping increases the closer the vehicle is to the intersection. From the figure, 70% are willing to stop at a distances equal to 300 ft upstream of the stop bar, while 30% of drivers will go through the intersection. At 200 ft, it is equally likely that drivers will stop or go through the intersection. These numbers seem valid, however closer inspection reveals that the rate at which the curves reach 100% or 0% is too low. The approach speed for this arterial is 35 mph. At this approach speed, the AASHTO stopping distance is 250 ft and the ITE recommended time for the change and clearance intervals is 4 seconds and 2 seconds, respectively. This means that a vehicle beyond 6 seconds from the intersection has ample time to stop, a distance of approximately 310 ft. However, Figure 6-5 shows that at a distance of 310 ft roughly 30% of the drivers do not stop, suggesting that the signal timing table frame IDs are too early relative to the trajectory table frame IDs. These results emphasize the importance of accurate time synchronization between datasets and that the NGSIM dataset will not yield a valid analysis for driver response to the onset of yellow.

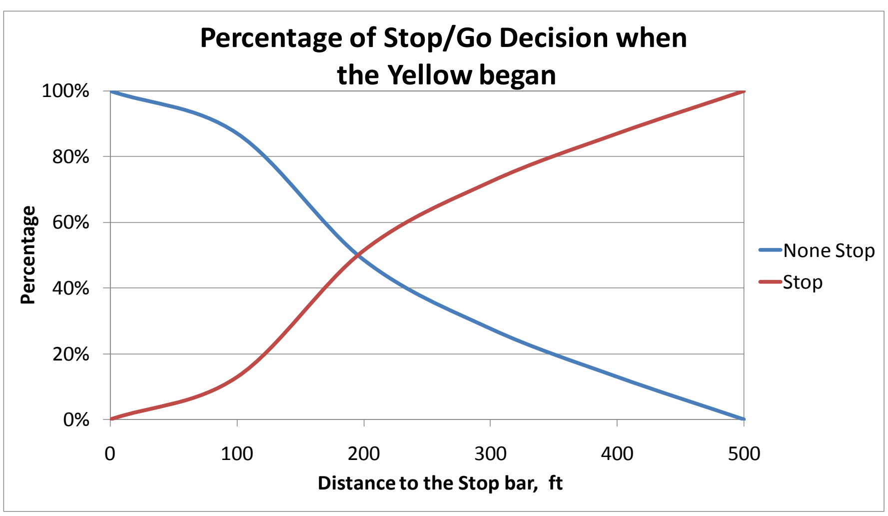



**Figure 6-5: Illustration of percentage accumulation of stop/go decision.**

## 6.3 Vehicle Velocity

To study the velocity profile of vehicles referenced to distance, a sample of 6 vehicles was selected for the same direction, lane, and section between two stop bars. Figure 6-6 shows the velocity profiles for these vehicles, five of which were stopped at the upstream intersection. While the vehicles accelerate, the dataset shows a velocity range of 0 to 20 ft/sec (0-13.6 mph). After the acceleration region, the flow velocity varies from 28 to 52 ft/sec (19.0-35.4 mph). Some vehicles stop at the downstream intersection, while others proceed through the intersection. Notice that the variation in speed for any given 100 ft increment can be remarkable, and this is likely due to vehicle-vehicle interactions.

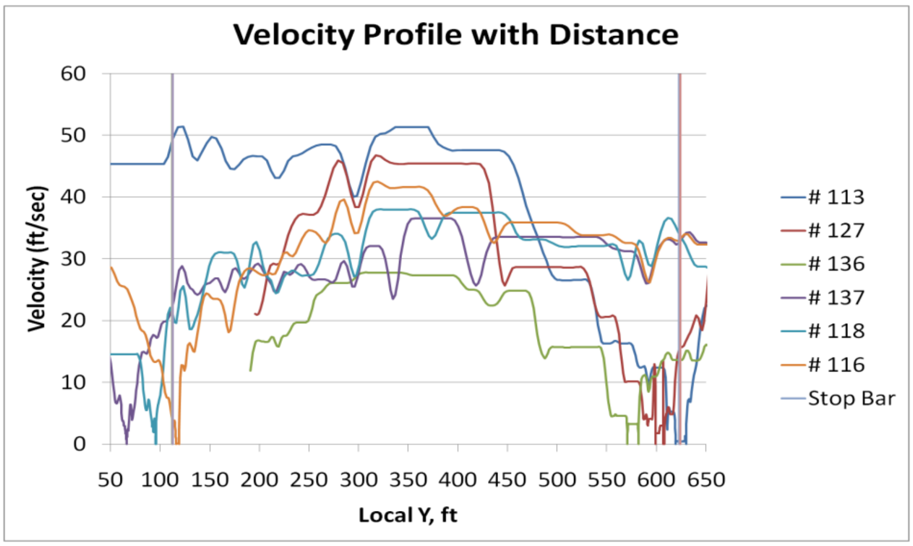



**Figure 6-6: Vehicle velocity profile with distance.**

## 6.4 Conclusion

This project found the data to describe vehicle operations in great detail, where vehicle trajectories are given in 0.1 second resolution. The database is accessible on-line and can be opened with a variety of database software. The project expanded the existing NGSIM dataset further by linking together other tables of data to directly connect vehicle data with signal control data. Some errors were found and corresponding corrections were made, resolving most of the errors. Situations of negative velocity were mostly resolved, with a few cases still remaining. Instances where vehicles are missing were also recorded in the database. Stop bar locations, phase status times, vehicle location, and spurious stops were all errors found in the database. When possible, steps were taken to correct these errors; however some of them remain in the database. Finally, researchers analyzed the dataset to determine basic stop bar and approach traffic behavior.

Aggregate behavior appeared very clean and reasonable; however this was not the case for individual driver data. It was found that the primary cause for this shortcoming is the accuracy of the phase status time stamps. Efforts to correct these time stamps did not remove errors and, as a result, driver behavior measurements, such as the response to the onset of yellow, were unrealistic. For each error found, researchers provide suggestions to control these errors in future data collection efforts. Fortunately, the data do seem to generally relate vehicle-vehicle interactions realistically. This is especially the case after filtering out spurious changes in velocity that tend to happen because of measurement errors documented in this report.

# 7 REFERENCES

VISSIM 4.0 User Manual," PTV Planung Transport Verkehr AG, 2005

# APPENDIX A

In order to introduce Access for beginners, six of the most helpful steps were chosen to get an essential understanding of the basic Access tools for manipulating.

## A-1. Composing a Basic Query

Task: To retrieve trajectories (in local Y) for the vehicle numbered 1182 in the 2nd dataset from frame ID 6750 to 7400.

Step 1. Open the Access database titled as NGSIM_ATL_DBMS.accdb.

Step 2. Click [Create] and then [Query Design] at the tool bar (top of the Access screen) (see A-1.1).

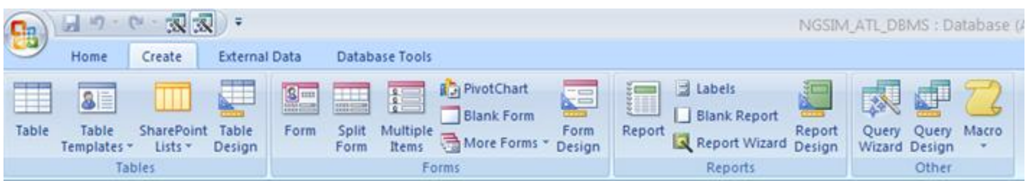



**Figure A-1.1: MS Access database.**

Step 3. Select [Trajectories2] under the tag [Tables] in the [Show Table] dialog, then click [Add] button (see A-1.2). Click [Close] button to close the dialog.

Step 4. Select [Vehicle ID] from the list of field in the 1st column, [Frame ID] from the list of field in the 2nd column, and then [Local Y] from the list of field in the 3rd column (see A-1.3).

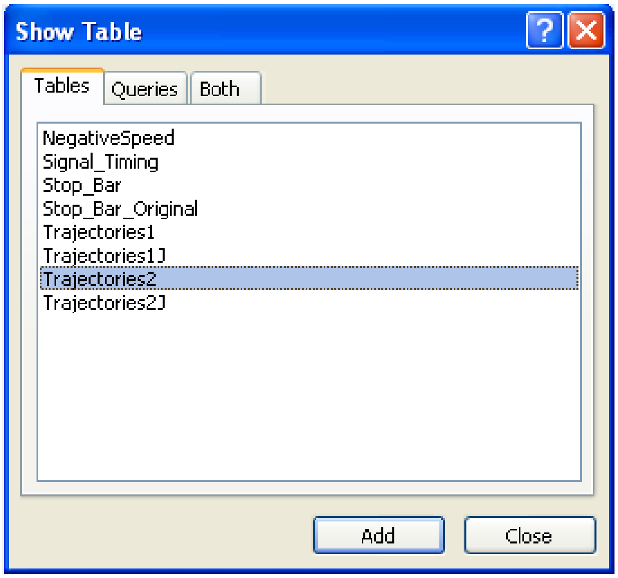



**Figure A-1.2: Trajectories under table.**

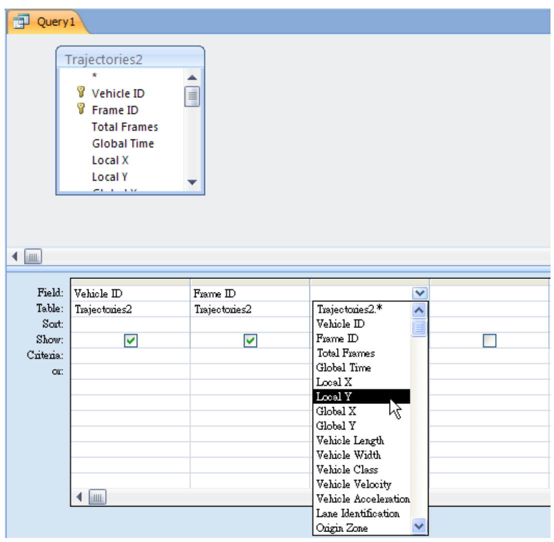



**Figure A-1.3: Query design view.**

Step 5. Type “1182” (A-1.4) into the row of criteria under the column vehicle ID and type “>=6750 and <=7400” into the row of criteria under the column frame ID.

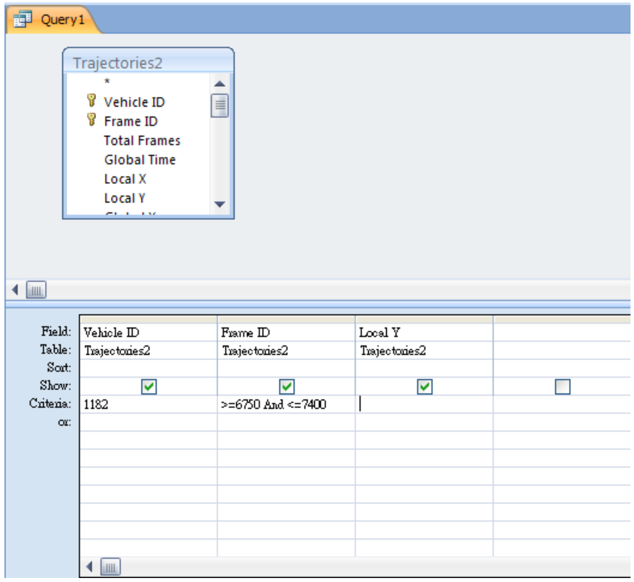
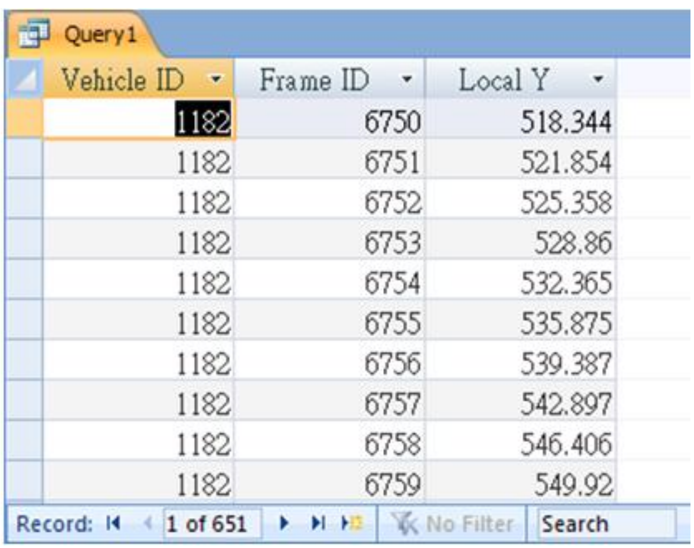



**Figure A-1.4: Query design view.**

Step 6. Click [Run] from the tool bar at the top of the Access screen.

Step 7. Get the result! (see A-1.5)



| Vehicle ID | Frame ID | Local Y |
|---|---|---|
| 1182 | 6750 | 518.344 |
| 1182 | 6751 | 521.854 |
| 1182 | 6752 | 525.358 |
| 1182 | 6753 | 528.86 |
| 1182 | 6754 | 532.365 |
| 1182 | 6755 | 535.875 |
| 1182 | 6756 | 539.387 |
| 1182 | 6757 | 542.897 |
| 1182 | 6758 | 546.406 |
| 1182 | 6759 | 549.92 |

Record: 1 of 651 No Filter Search

**Figure A-1.5: Results of the query.**

## A-2. Composing a More Advanced Query That Involved Two or More Tables

Task: to find the distance to the downstream stop bar for the vehicle numbered 1182 in the 2nd dataset from frame ID 6750 to 7400.

Step 1. After the Access database titled as NGSIM_ATL_DBMS.accdb is opened Click [Create] and then [Query Design] at the tool bar of the Access screen.

Step 2. Select [Trajectories2] under the tag [Tables] in the [Show Table] dialog, then click the [Add] button (see A-2.1). Select [Stop_Bar] under the tag [Tables] in the [Show Table] dialog, then click the [Add] button. Click the [Close] button to close the dialog.

Step 3. Click [Lane Identification] from the [Trajectories2] table, and then drag it to the [Lane] in the [Stop_Bar] table on the upper window of the query 1 (see A-1.2). Click [Direction] from the [Trajectories2] table, and then drag it to the [Direction] in the [Stop_Bar] table on the upper window of the query 1. The result should be similar as follows;

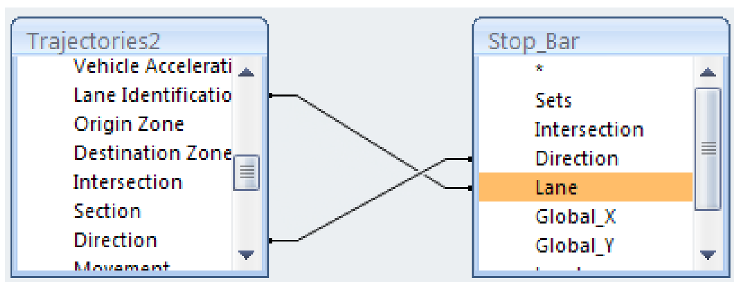



**Figure A-2.1: Joined tables, Trajectories2 and Stop Bar.**

Step 4. Select [Trajectories2.Vehicle ID] from the list of field in the 1$^{st}$ column, [Trajectories2.Frame ID] from the list of field in the 2$^{nd}$ column, [Stop_Bar.Intersection] from the list of field in the 3$^{rd}$ column [Stop_Bar.Direction] from the list of field in the 4$^{th}$ column, and [Stop_Bar.Lane] from the list of field in the 5$^{th}$ column (see A-2.2).

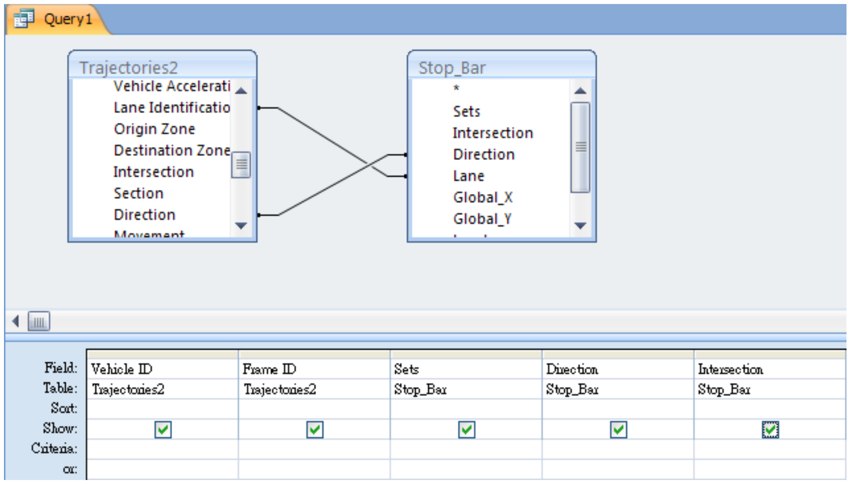



**Figure A-2.2: Query design view, window to join tables.**

Step 5. Uncheck the show box for column 3, 4, and 5. Type “1182” into the row of criteria under the column vehicle ID; type “>=6750 and <=7400” into the row of criteria under the column frame ID; type “2” into the row of criteria under the column sets; type “2” into the row of criteria under the column direction; type “2” into the row of criteria under the column intersection (see A-2.3).

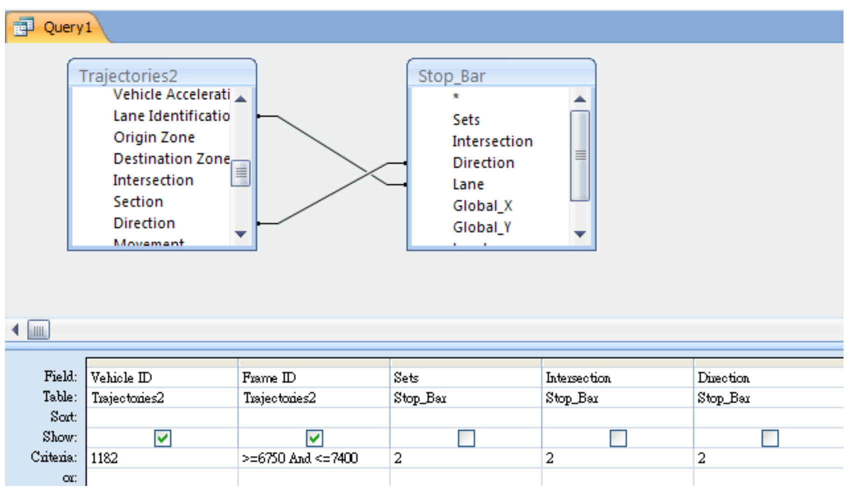



**Figure A-2.3: Query design view.**

Step 6. Type “Distance to Bar: [Stop_Bar]![Local_y]-[Trajectories2]![Local Y]” into the row of field at the 6th column, and then check the show box at this column (see A-2.4).

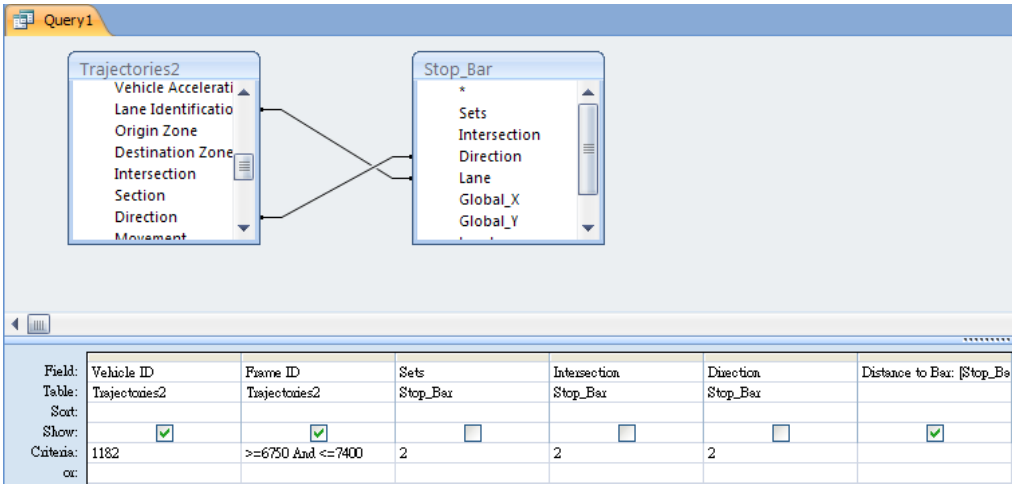
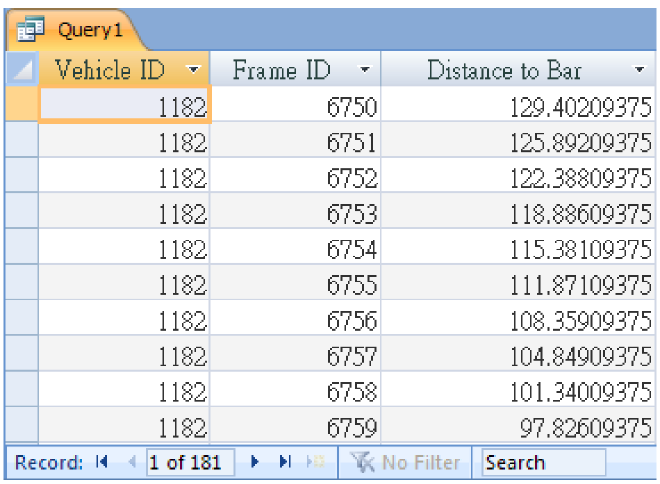



**Figure A-2.4: Query design view.**

Step 7. Click [Run] from the Access tool bar.

Step 8. Get the result! (see A-2.5)



| Vehicle ID | Frame ID | Distance to Bar |
|---|---|---|
| 1182 | 6750 | 129.40209375 |
| 1182 | 6751 | 125.89209375 |
| 1182 | 6752 | 122.38809375 |
| 1182 | 6753 | 118.88609375 |
| 1182 | 6754 | 115.38109375 |
| 1182 | 6755 | 111.87109375 |
| 1182 | 6756 | 108.35909375 |
| 1182 | 6757 | 104.84909375 |
| 1182 | 6758 | 101.34009375 |
| 1182 | 6759 | 97.82609375 |



**Figure A-2.5: Query results.**

## A-3. Composing a Join and an Explanation Why This Is Advantageous to Relying on Queries

A join in a relational database defines a certain relation between two tables. It provides an efficient way to access the wanted data. Such a relation also offers the flexibility of retrieving information based on different criteria. That is, we don't have to put everything into an extremely huge table, but just a portion of it is needed for the analyses.

Task: Check the joint properties in the previous example.

Step 1. After the Access database titled as NGSIM_ATL_DBMS.accdb is opened, click [Create] and then [Query Design] at the Access tool bar.

Step 2. Select [Trajectories2] under the tag [Tables] in the [Show Table] dialog, then click the [Add] button. Select [Stop_Bar] under the tag [Tables] in the [Show Table] dialog, then click the [Add] button. Click the [Close] button to close the dialog.

Step 3. Click [Lane Identification] from the [Trajectories2] table and then drag it to the [Lane] in the [Stop_Bar] table on the upper window of the query 1. Click [Direction] from the [Trajectories2] table and then drag it to the [Direction] in the [Stop_Bar] table on the upper window of the query 1. The result should be similar as follows (see A-3.1).

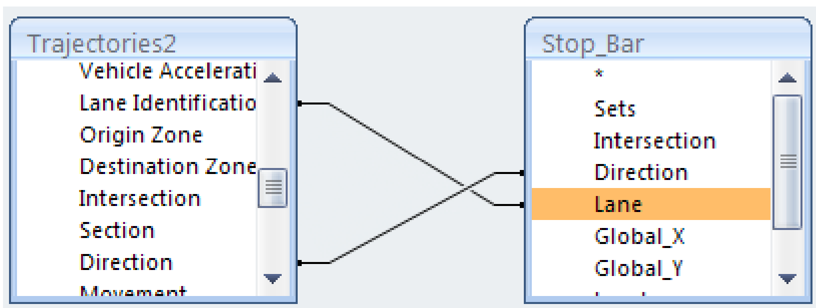



**Figure A-3.1: Joined tables trajectories and stop bar.**

Step 4. Right click on the join line between [Trajectories2].[Lane Identification] and [Stop_Bar].[Lane], and then a menu connected to this join will pop up (see A-3.2).

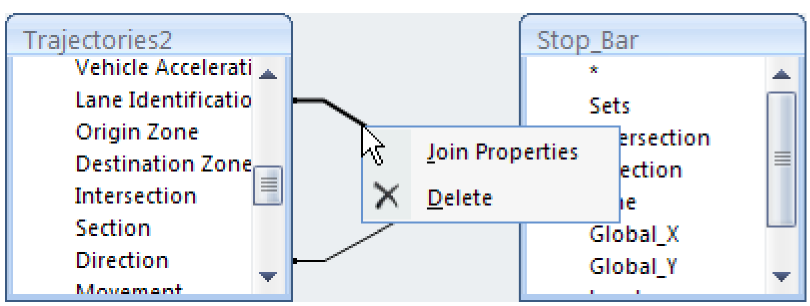



**Figure A-3.2: Properties of joined tables.**

Step 5. Select [Join Properties] from the popup menu in the previous step, and then modify (if it is necessary) the connected properties in the join properties window (see A-3.3).

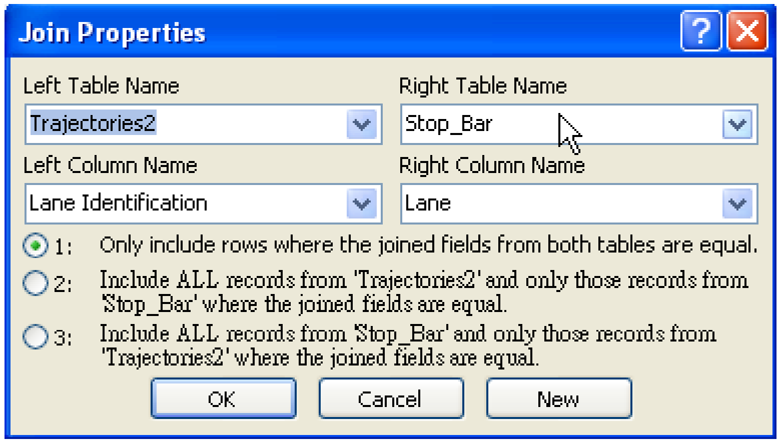



**Figure A-3.3: Join properties window.**

## A-4. Step-By-Step Instruction on Exporting a Table in an Excel File

The major concern of exporting results from Access to an Excel file is the number of records. It depends on the capability of Excel. Check the help in Excel to figure out how many rows can be processed in an Excel file. Since the same procedure could be used to export a table or a query to an Excel file and the lateral is more complicated, a query (instead of a table) would be used to instruct the export process as follows:

Task: Output the result from the 1st example task to an Excel file.

Step 1. Before the result of a query can be exported to an Excel file, the query has to be stored. Once the result is ready, click the [Excel] icon in the [Export] group under the [External Data] tag from the Access tools (see A-4.1).

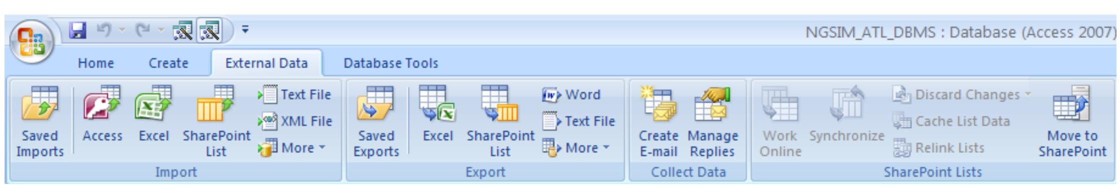



**Figure A-4.1: MS Access tool bar.**

Step 2. On the [Export – Excel Spreadsheet] dialog, specify the destination file name. Since the file format is already in Excel workbook, it is not necessary to change it. After specifying the destination file name, click the [OK] button to close this dialog (see A-4.2).

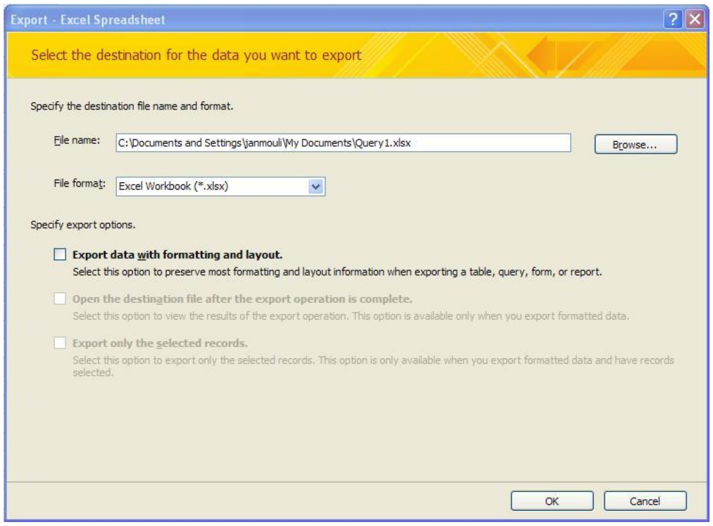



**Figure A-4.2: Export Excel spreadsheet's window.**

Step 3. If the result of query is stored, a confirm message would show in a dialog to ask whether you would like to save the export steps. However, if the result of query is not stored yet, an error message would pop up as follows: You have to click its [OK] button to close the dialog and re-do the steps here after the result is stored.

## A-5. Step-By-Step Instructions on Importing Additional Data into the Existing Tables

Task: Import additional trajectory data for the 1st dataset.

Step 1. Click [External Data] and then [Text File] after the Access database titled as NGSIM_ATL_DBMS.accdb is opened.

Step 2. Click the [Browse] button to select the data file you would like to import; check the radio button for [Append a copy of the records to the table]; and then select [Trajectories1] from the table list followed by the radio button you just checked. Click the [OK] button to close the dialog and begin to get the external data which is stored in the file titled as "additional.txt" (see A-5.1)

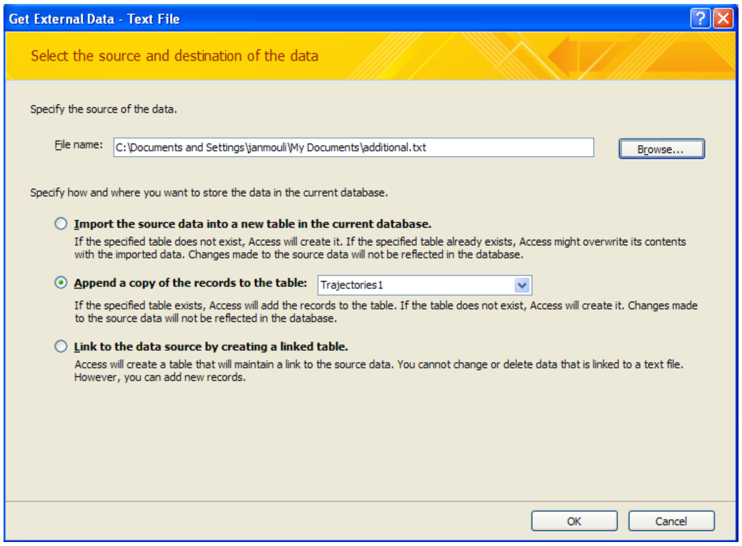



**Figure A-5.1: Get external data window.**

Step 3. Tell the Import Text Wizard about the format of the importing file. For example, how the fields are defined in the file, etc. “Delimited” is the format for the example file used here. After the selection, click the [Next] button for the successive step (see A-5.2).

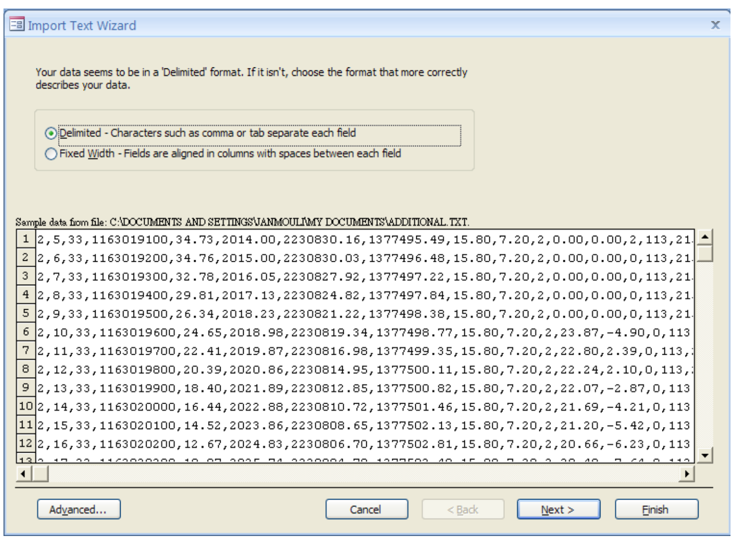



**Figure A-5.2: Import text window.**

Step 4. According to the format of the importing file, choose the delimiter that separates the fields. [Comma] is selected for the example file here. If there are field names in the first row of the importing file, check [First Row Contains Field Names]; otherwise, keep it unchecked. Click the [Next] button for the successive step if there is nothing to change for the importing file (see A-5.3).

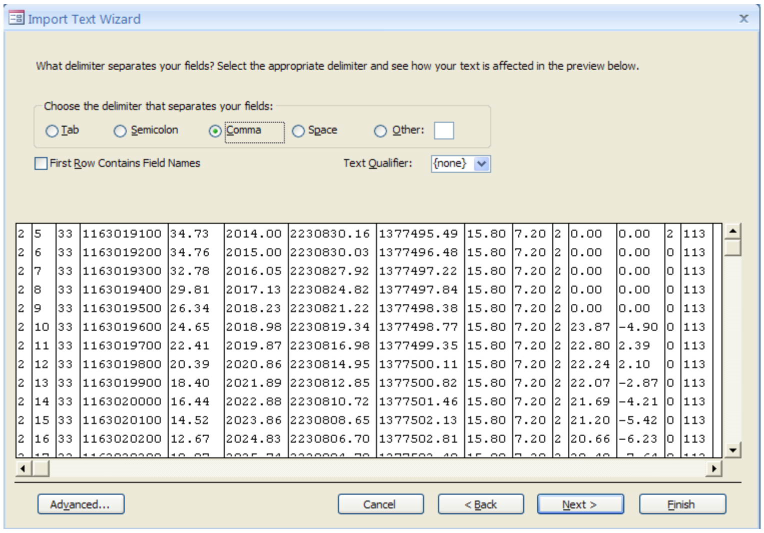



**Figure A-5.3: Import text window.**

Step 5. Click the [Finish] button to start the import if the table name in the “Import to Table:”(see A-5.4) is correct; Otherwise, revise the table name as it should be and then click the [Finish] button.

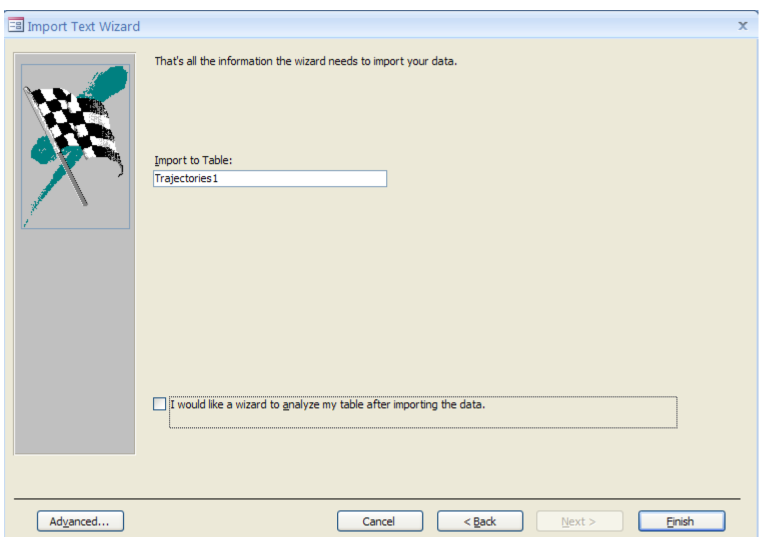



**Figure A-5.4: Import text window.**

Step 6. Click the [Close] button to close the last dialog of “Get External Data - Text File” (see A-5.5), and then all the data in the text file is imported and ready to be used.

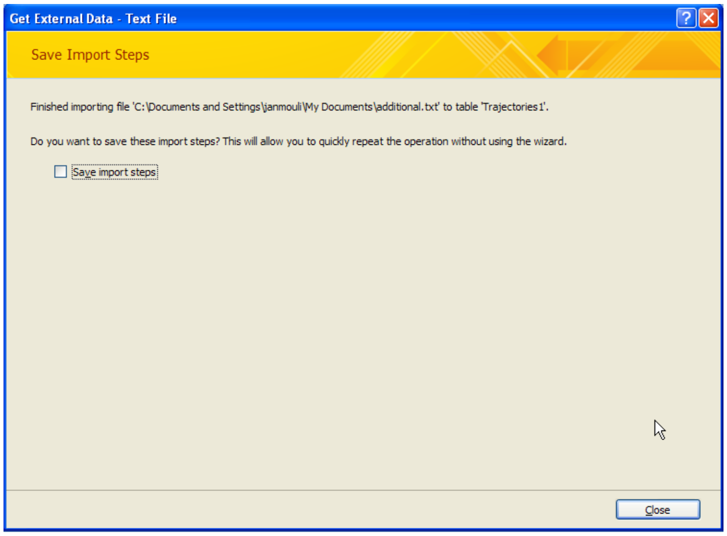



**Figure A-5.5: Get external data.**

## A-6. Importing Additional Data into a New Table

Task: Import the trajectory data for the 1st dataset, assuming it did not exist in the database yet.

Step 1. After the Access database titled as NGSIM_ATL_DBMS.accdb is opened, Click [External Data] and then [Text File] at the Access tool bar.

Step 2. Click the [Browse] button to select the data file you would like to import; check the radio button for "Import the source data into a new table in the current database" (see A-6.1); and then click the [OK] button to close the dialog and begin to import additional data into a new table.

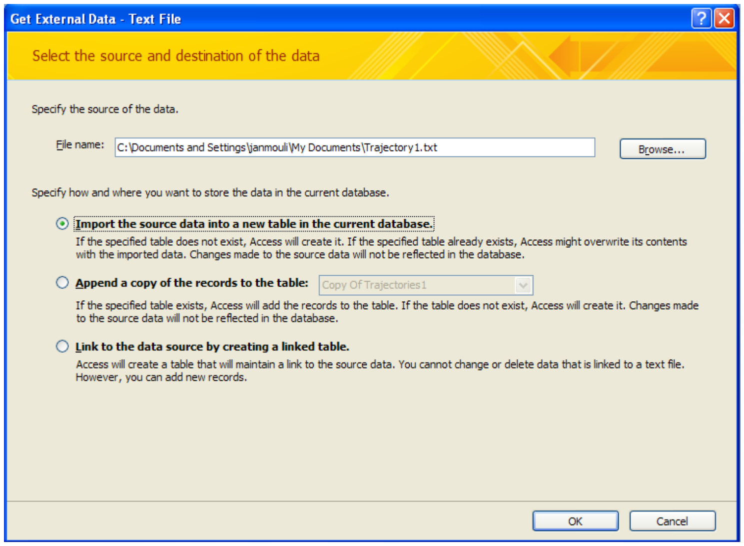



**Figure A-6.1: Get external data window.**

Step 3. Tell the Import Text Wizard about the format of the importing file. For example, how the fields are defined in the file, etc. "Delimited" is the format for the example file used here. After the selection, click the [Next] button for the successive step (see A-6.2).

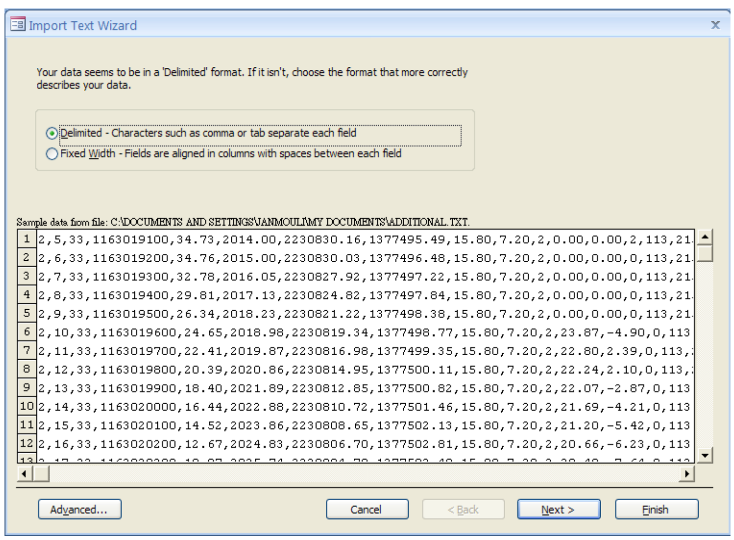



**Figure A-6.2: Import text window.**

Step 4. According to the format of the importing file, choose the delimiter that separates the fields. [Comma] is selected for the example file here. If there are field names in the first row of the importing file, check [First Row Contains Field Names]; Otherwise, keep it uncheck. Click the [Next] button for the successive step if there is nothing to change for the importing file (see A-6.3).

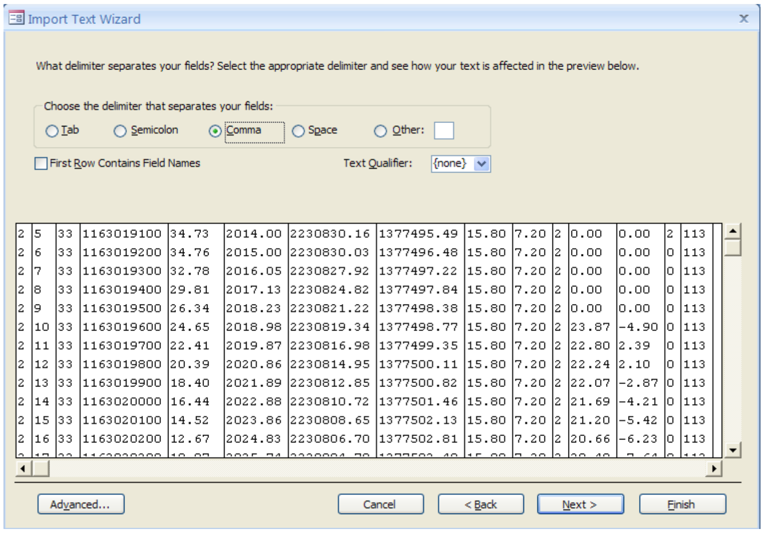



**Figure A-6.3: Import text window.**

Step 5. Specify information about each of the fields you are importing. Once the setting is done, click the [Next] button for the successive step (see A-6.4).

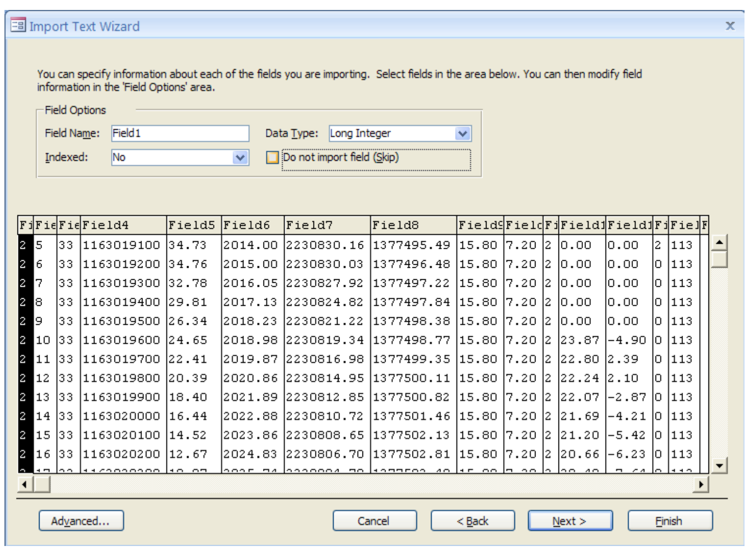



**Figure A-6.4: Import text window.**

Step 6. Set the primary key. Keep the option of “Let Access add primary key” as checked (it is a default setting) and then click the [Next] button for the final step (see A-6.5).

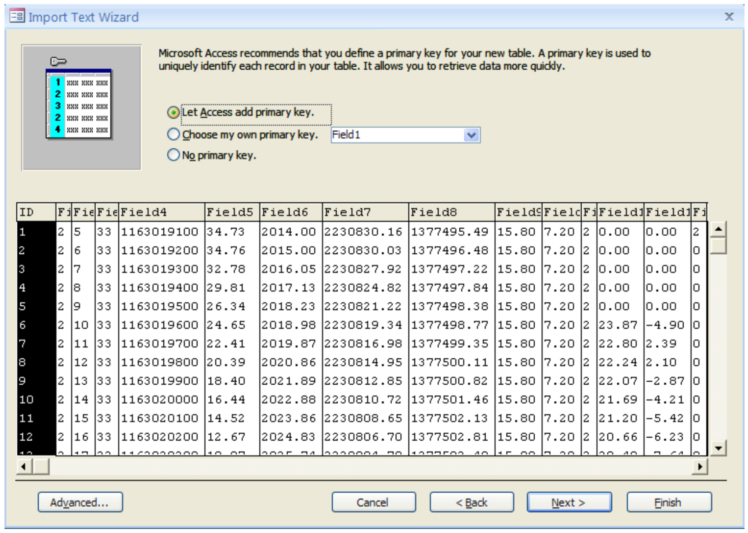



**Figure A-6.5: Import text window.**

Step 7. Click the [Finish] button to start the import if the table name in the "Import to Table:" is correct; otherwise, revise the table name as it should be and then click the [Finish] button (see A-6.6).

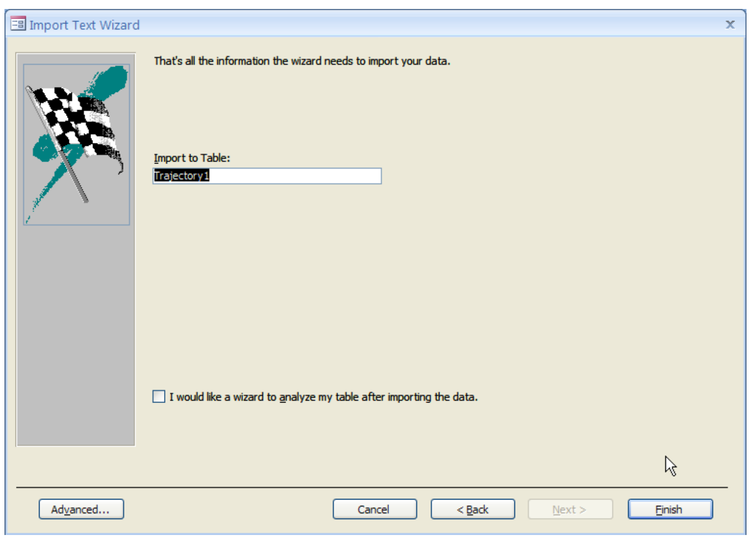



**Figure A-6.6: Import text window.**

Step 8. Click the [Close] button to close the dialog of "Get External Data - Text File", and then all the data in the text file is imported and ready to be used (see A-6.7).

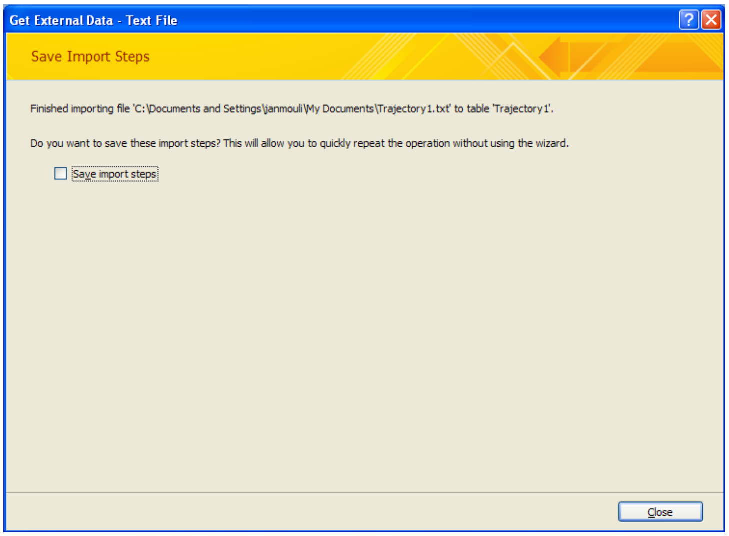



**Figure A-6.7: Get external data window.**