Quantifying Non-Recurring Delay on New York City's Arterial Highways

Project C-01-29

FINAL REPORT

Prepared for:
New York State Department of Transportation
Region 11

Prepared by:
Rensselaer Polytechnic Institute
Polytechnic Institute of NYU
Rutgers University
City University of New York, City College

December 2008

DISCLAIMER STATEMENT

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle	5. Report Date
Quantifying Non-Recurring Delay on New York City's Arterial Highways	December 2008
	6. Performing Organization Code
7. Author(s) George List (RPI and later, NCSU), John Falcocchio (Polytechnic Institute of NYU), Kaan Ozbay (Rutgers) and Kyriakos Mouskos (CUNY – City College)	8. Performing Organization Report No. RF 55657-1215
9. Performing Organization Name and Address Region 2 - University Transportation Research Center The City College of New York New York, NY 10031	10. Work Unit No. 11. Contract or Grant No.
12. Sponsoring Agency Name and Address New York State Department of Transportation, 50 Wolf Road Albany, NY Federal Highway Administration U.S. Department of Transportation	13. Type of Report and Period Covered 14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This research project was undertaken for NYSDOT to provide a better understanding of the impacts of traffic incidents/accidents on traffic delays on New York City's Arterial Highways, and to better quantify and predict non-recurring traffic delay for the city's arterial highways.

The project had two basic goals: (1) the development of New York City input data for the New York City's application of the New York State DOT's delay prediction model "CNAM" (Congestion Needs Analysis Model), and (2) investigation of the published literature to identify models / methods that could improve the CNAM approach for estimating non-recurring delay.

The research activities consisted of six basic tasks as follows.

Task 1 developed the goals and objectives for the research project, and identified the performance measures to be used in the collection and analysis of traffic incident data for New York City. Task 2 contains a review of the models that have been developed for predicting non-recurring delay (NRD). Task 3 was started by searching for potential data sources that could used to identify nonrecurring incident characteristics and performance metrics. The task focused on agencies that are involved in highway incident management/monitoring as well as those that collect roadway attributes data, such as physical roadway and traffic flow characteristics. Task 4 was on Look-up Tables. In this task, the "look up tables" for CNAM's application in New York City were updated, consistent with data availability. Task 5 described how the new look up tables will change the structure of CNAM and alter its predictions of non-recurring delay. A new set of look up tables are recommended.

17. Key Words	18. Distribution Statement
Traffic delay, incidents, traffic model, incident management	

19. Security Classif (of this report)	20. Security Classif. (of this page)	21. No of Pages	22. Price
Unclassified	Unclassified	319	

Form DOT F 1700.7 (8-69)

Executive Summary

This research project was undertaken for NYSDOT to provide a better understanding of the impacts of traffic incidents/accidents on traffic delays on New York City's Arterial Highways, and to better quantify and predict non-recurring traffic delay for the city's arterial highways.

The project had two basic goals: (1) the development of New York City input data (look up tables) for the New York City's application of the New York State DOT's delay prediction model "CNAM" (Congestion Needs Analysis Model), and (2) investigation of the published literature to identify models / methods that could improve the CNAM approach for estimating non-recurring delay.

The research team comprised of Professors George List (RPI and later, NCSU), John Falcocchio (Polytechnic Institute of NYU), Kaan Ozbay (Rutgers) and Kyriakos Mouskos (CUNY – City College).

The research activities consisted of six basic tasks and products as described below.

Task 1 – Establishing Goals and Objectives

This task developed the goals and objectives for the research project, and identified the performance measures to be used in the collection and analysis of traffic incident data for New York City, and in the review of the published literature to assess the suitability of alternative models (to the CNAM) for estimating non-recurring delay.

Task 2 – Review of Non-Recurring Delay Models

This task contains a review of the models that have been developed for predicting non-recurring delay (NRD). The literature review did not uncover reliable NRD models in use in practice. The CNAM was found to be as good as or better than those that are in use today.

Task 3 – Data Collection and Analysis

The task was started by searching for potential data sources that could used to identify non-recurring incident characteristics and performance metrics. The task focused on agencies that are involved in highway incident management/monitoring as well as those that collect roadway attributes data, such as physical roadway and traffic flow characteristics.

The research team initially expected using detailed incident data from the Integrated Incident Management System (IIMS) Demonstration Project, sponsored by the USDOT. This source, however, could not make the necessary information available and a change in venue was made by deciding on the use of TRANSCOM as the primary source of data from incidents recorded for the I-278 Corridor in New York City, from February 1, 2004, to March 31, 2005.

The TRANSCOM incident data were complemented with data (volume, roadway geometry, etc.) from other sources to form the basis for incident analysis and CNAM model enhancement.

Incident data from TRANSCOM were summarized and cross tabulated for a number of variables, including (1) roadway facility type, (2) types of incident, (2) incident frequency, (3) number of lanes blocked by the incident, (4) incident duration, (5) time of day, (6) day of week, and (7) weather and pavement conditions. These data were used in the development of a model for predicting incident frequency and their duration as well as traffic delay. Because the incident frequency data set in the TRANSCOM database was found to be incomplete, the models' results must be interpreted with this constraint in mind.

Task 4 – Look-up Tables

In this task, the "look up tables" for CNAM's application in New York City were updated, consistent with data availability.

Task 5 – Strategy Assessment (Validation of Methodologies)

This task describes how the new look up tables will change the structure of CNAM and alter its predictions of non-recurring delay. A new set of look up tables are recommended. The change consists in changing the classification of incidents from what CNAM currently uses (i.e., number of lanes blocked) to one that defines incidents focusing on categories such as “property damage”, disabled vehicle”, personal injury”, etc. This new definition has the added advantage of using CNAM as a planning tool for reducing incident frequency and severity. This is because the delay information provided by the revised model can be directly related to the type of incident.

TABLE OF CONTENTS

TASK 2: Review of Non-Recurring Delay Models 2-1

TASK 3

Subtask 3.1: Data Assembly and Analysis 3.1-1

TASK 3

TASK 3

Subtask 3.2: Data Analysis

TASK 3

TASK 5: Strategy Assessment 5-1

Task 1: Goals and Objectives

Prepared by:

Hyunwoo Cho
George List
Department of Civil and Environmental Engineering
Rensselaer Polytechnic Institute
Troy, NY 12180

Tel: 518-276-6362
Fax: 518-276-4833
Email: listg@rpi.edu

March 14, 2004

1.0 Introduction

This task report presents the goals and objectives that will guide the development of performance measures for non-recurring delay (NRD) in the context of project C-01-29. Non-recurring delay amounts to 40-50% of all delay. It is caused by accidents, incidents, other unexpected events, special events (e.g., parades. sporting events), and construction.

The purpose of the project is to develop “look-up tables” that can predict non-recurring delay (NRD) for New York City’s freeway network. The “look-up tables” have to be capable of working in conjunction with CNAM, a network congestion analysis program.

In the sense that the goals and objectives of the project lead to a need for metrics for those goals and objectives, this technical memorandum is about performance measures as well.

A search has been conducted for reports and technical papers that present NRD goals, objectives and performance metrics. The papers of greatest interest have been those that present a philosophical discussion about what these goals and objectives should be, how they should be chosen, and what metrics should be employed. Of lesser interest are papers that present the results of an NRD analysis, in which, at least implicitly, a set of goals and objectives have been pursued by computing specific performance metrics.

Section 2.0 presents the findings from the literature review. Section 3.0 enumerates the goals and objectives suggested for developing the “look up tables”. Section 4.0 addresses next steps. The CD accompanying this report contains the papers that could be obtained in machine readable format.

2.0 Literature Review

Precedent is an important source of insight in establishing goals, objectives, and evaluation metrics for non-recurring delay. Although the topic is fairly new, many state and local agencies have wrestled with the problem of establishing ways to determine what actions should be taken (setting goals and objectives) and measuring the effectiveness of those actions (evaluation metrics). This section summarizes the findings from a review of those initiatives.

2.1 New York State DOT

NYSDOT (*31*) has established a number of goals for the transportation systems in the state, one of which relates to mobility:

> “GOAL: TO MOVE PEOPLE AND GOODS CONVENIENTLY, RELIABLY, SAFELY, AT A REASONABLE COST, AND IN AN ACCEPTABLE TRAVEL TIME ON THE STATE TRANSPORTATION SYSTEM BY IMPLEMENTING MOBILITY PROJECTS THAT ARE COST EFFECTIVE, ACCOMMODATE THE VARIOUS INTER-DEPENDENT MODES, AND ARE COMPATIBLE WITH AND ENHANCE ECONOMIC DEVELOPMENT, THE COMMUNITY, AND THE ENVIRONMENT.”

The second among several objectives related to this goal relates to non-recurring delay:

> "Reduce the growth of daily non-recurring person hours of delay (PHD) by 10 percent by the end of the first five years of the program period and by additional reductions within 20 years as the projects are fully implemented. Measure: Person-Hours of Delay and Person-Hours of Delay per centerline mile on the CMS network."

The highlighting has been added to emphasize the focus on NRD. Notice that the objective is to reduce *person hours of delay* (the evaluation metric) by 10% across five years. Notice also that the second evaluation metric being monitored is normalized based on centerline miles in the CMS (Congestion Monitoring System) network.

The narrative that follows the goal statement indicates:

> "The revised Goal Objectives build off the excellent early effort in the development of the Congestion Management System (CMS). Its modeling tools are available to estimate recurring and non-recurring delay (in terms of both VHD and PHD), ton hours of delay, and user costs for the entire State highway system (with the exception of the Thruway).
>
> While the previous Mobility Goal focused on vehicle hours of delay as its core measure, the revised Goal uses person hours of delay (PHD), which better addresses automobiles, high occupancy vehicles, transit, bicycles/pedestrians, and intermodal connections. It also uses ton hours of delay to address goods movements."

Clearly, NYSDOT determined that VHD was failing to reflect the true entities being affect by the NRD, namely the people and the goods being transported.

Explanation for the normalized metric is offered as follows:

> "There are separate Objectives for recurring delay and incident (or non-recurring) delay because each has separate and distinct characteristics. … The evaluation measure is person hours of delay (PHD). Analysis data for this Objective can be obtained from either CMS modeling tools or project-specific sources. While PHD is a useful statistic for measuring congestion, it is difficult for the public to relate to. Therefore, the rate, PHD per centerline mile on the CMS network has been included to make delay more understandable."

So, in sum, NYSDOT is focused on effectiveness / productivity of the transport system, as part of a mobility-related goal and desires to minimize PHD and PHD/center-line mile. This seems like a very reasonable perspective.

2.2 CalTrans

CalTrans (*5*) has a Freeway Management Program for which goals are presented related to freeway operations. Those goals are shown below in Table 2.1: improve safety, improve efficiency, and
Improve reliability.

Table 2.1 Freeway Operations Goals and Objectives

Goals	Objectives
Improve safety	• Reduce accident rates and severity • Improve incident management • Improve assistance to stranded motorists
Improve efficiency	• Minimize recurring traffic congestion • Reduce vehicle delay due to construction and event traffic • Improve incident management
Improve reliability	• Improve consistency of travel speeds and travel times • Improve incident management • Increase monitoring of freeway conditions • Provide information to motorists

Source: (*5*)

Also listed in the table are objectives for each of the goals and improving incident management appears for all three. The implication is that incidents, which are the root cause of NRD, are deemed to be something which should be managed with great care to minimize the negative impacts on the transportation system.

So this document shows that delays (with an unspecified metric) are again an important measure of the effects of NRD, that NRD is tied to an efficiency goal for the system and that there are two related goals, both of which are affected by incidents, which are the underlying cause for NRD. Those are improved safety and improved reliability. In the first case, the implicit metrics are accident rates (by severity type) and the quality of assistance provided to stranded motorists, and in the second, consistency in travel speeds and travel times, and the quality of information provided to travelers about current system conditions.

2.3 Other States

Arizona (*2*) has a document that describes its freeway management system, and included in that document are two objectives:

- Reduce the impacts of recurring congestion
- Minimize the duration and effects of non-recurring congestion

The implication is that NRD is a recognized source of delay and that one metric employed in monitoring and evaluation is the duration of NRD events.

Wisconsin (*46*) emphasizes safety, more than efficiency or reliability, with the following four objectives related to incident management:

- "Reduce the number of motor vehicle collisions and associated injuries and fatalities due to incidents and secondary effects.
- Ensure safety of responding personnel with improved incident site management.
- Improve and enhance the management of incidents involving hazardous materials.

- Improve the response time of emergency medical services for incidents involving injury."

New Jersey has a mindset similar to both New York and California, with an added dimension related to goods movement, as in a study conducted by NJIT (*30*):

> System Efficiency: System efficiency would consider the ability to move freight more quickly. Thus, travel delay as a function of miles traveled, or Ratio of Truck Delay per Mile Traveled would be measured. The delay component of this measure would be split into two forms: recurring delay and non-recurring delay. Recurring delay (delay due to congestion) is detrimental to the movement of freight on the highway. But when anticipated, recurring delay can be accounted for within trucking companies' cost calculations and scheduling. Non-recurring delay (delay caused by incidents and accidents) cannot be anticipated due to the unpredictable nature of its occurrence. As such, trucking companies cannot adjust cost calculations and schedules, giving less reliability in the delivery time of goods (especially in congested urban areas during peak hours).
>
> Additional Ton-miles Traveled under an analysis of freight movement without the facility in question. A small increase of ton-miles would indicate a high level of redundancy (goods easily rerouted) and a large increase in the number of ton-miles traveled would indicate a system critical facility that would be difficult to operate without."

2.4 Local Area Studies

The MPO in Washington DC (*29*) emphasizes the number of accidents and vehicle breakdowns that occur on roadway segments and transit routes. The Island of Oahu, in Hawaii, (*32*) has developed an ITS-based freeway management system that strives to have crashes and other non-recurring incidents identified and cleared away quickly. As with CalTrans, this reflects an emphasis on efficiency (specifically, quick detection of the occurrence of incidents, and their expeditious mitigation). Washoe County, in Nevada, (*35*) emphasizes clearance time. One of its documents focuses on removing disabled vehicles from traffic lanes and level of service (LOS). To this end, the document asks whether freeway service patrols might be useful in providing towing services to disabled vehicles. The claim is that these strategies will assist in preventing congestion caused by individual incidents.

2.5 FHWA Guidance

The Freeway Management and Operations Handbook (*16*) has a chapter focused on incidents and incident management. The Handbook's comment about incident management programs is that:

> "Reduced incident duration has proven the greatest contributor to the benefits of an incident management program. Reductions in duration can be achieved by:
>
> - Reduced detection time
> - Timely and appropriate response
> - Rapid clearance

Benefits of an effective program are both quantitative and qualitative. The quantitative benefits include:

- Increased survival rate of crash victims
- Reduced delay
- Improved response time
- Improved air quality
- Reduced occurrence of secondary incidents
- Improved safety of responders, crash victims and other motorists"

Two things to notice from this excerpt are a) the emphasis on reducing the *duration* of the incident, which implies an efficiency goal, and b) the metric by which performance will be assessed (incident duration). Also relevant are the actions that FHWA suggests can have an impact: prompt detection, astute response, and rapid clearance.

However, an additional observation is important. The excerpt also speaks to safety issues, such as the survival rate of crash victims, a reduced occurrence of secondary incidents, and improved safety for responders, crash victims, and other motorists. So, as with the CalTrans discussion in Section 2.3, there is a connection between concern over NRD and goals and objectives related to safety.

In another section of the document, reduced driver frustration is identified as a benefit. It is an interesting observation, because it is the only place in all of the articles where such an idea is identified. Implicitly, it ties back to the reliability goal described in the CalTrans document. It seems important to the research team, because the highway system is providing a service, and if the quality of that service is highly variable, the "customers" may be dissatisfied with performance.

The Handbook also suggests goals such as: reduce secondary incidents, increase safety for responders, increase and improve use of alternate routes, reduce liability for responding agencies. Supporting objectives include: decrease detection times, improve response times, increase motorist information, improve clearance procedures, decrease number of lanes closed, and decrease road and lane closure times.

The Handbook continues by identifying actions that are part of the incident management process. These include detection, verification, response, site management, traffic management, and clearance. For each, there is an objective which helps elicit FHWA's perspective on how incident management should be focused, including evaluation metrics that relate directly to NRD:

"*Detection* is determining that an incident has occurred. Rapid detection is necessary to minimize the period of time during which roadway capacity is reduced.

"*Verification* is determining the precise location and nature of an incident, as well as the display, recording, and communication of this information to the appropriate agencies. Proper verification is required to reduce the time required to deploy an appropriate response to the scene of an incident.

> “*Response* is the activation, coordination, and management of the appropriate personnel, equipment, and communication links and motorist information media as soon as it is reasonably certain that an incident has occurred. Timely and effective response reduces the incident's duration, and therefore, the time of roadway operation at reduced capacity.
>
> “Effective *site management* increases safety for crash victims, motorists and responders; coordinates responder activities; and decreases the impacts of incidents on the transportation system.
>
> ‘*Traffic management* is the application of traffic control measures at the incident site and on facilities affected by the incident. Effective traffic management minimizes traffic disruption while maintaining a safe workplace for responders.
>
> ‘*Clearance* is the removal of vehicles, wreckage, debris, spilled material and other items from the roadway and the immediate area to restore roadway capacity. Improving incident clearance procedures can:
>
> - Restore the roadway to its pre-incident capacity quickly and safely
> - Minimize motorist delay
> - Make effective use of all resources
> - Enhance the safety of responders and travelers
> - Protect the roadway and private property from unnecessary damage during the removal process

In another section of the document, focused on documentation and evaluation, “commonly used statistics” are identified that relate to the “cost” of responding to incidents:

- Number of service patrol assists
- Average elapsed time from incident occurrence to detection
- Average IRT response time
- Average elapsed time to restoration of capacity

These are all useful thoughts in the context of establishing evaluation metrics for NRD.

In a separate FHWA document focused on the impacts of adverse weather (*17*), the following comments are made:

> “Adverse weather is the second largest cause of non-recurrent congestion. Snow, ice and fog alone cause 15 percent of non-recurring delay. Likewise, a light rain can increase travel time delay by 12 to 20 percent. This translates into financial impacts. In metropolitan areas truckers lose about $3.4 billion (about 32 million hours) stuck in weather-related traffic delays. A one-day highway shutdown can cost a metropolitan area up to $76 million in lost time, wages, and productivity.”

This is very interesting because it speaks to an impact not identified anywhere else: the economic impacts to carriers and the economy that they support.

2.6 Other Studies and Papers

A few other studies have been conducted focused on NRD and its goals, objectives and performance metrics. The ideas presented seem consistent with the agency documents discussed in the previous sections

2.6.1 F-SHRP Study

A recent study focused on FSHRP (*43*) emphasizes an objective of providing highway users with more reliable travel times by reducing the impact of non-recurring incidents. It claims that if overall delay in urban areas were reduced by 5%, this would lead to savings of $3.5 billion in user costs. A 5% reduction in delay from non-recurring incidents alone would amount to savings of $2 billion a year. So the implication here is clear. A metric that should be considered is the monetary value of the time that would be saved through reduced NRD delay.

2.6.2 Integrated Public Safety and Highway Operations

In a study by Mitretek (*14*), the benefits to improved incident management were evaluated. Table 2.2 summarizes the findings. Metrics included lives saved and delays eliminated. It is interesting that the report does not mention injuries prevented.

Table 2.2 Benefits from Incident management

Operations Area	Potential Benefit	Benefit Value	Implementation Cost	Net	Ratio
Crash Prevention	5,000 lives	$20 billion/yr (1)	~$4 billion/yr (2)	$16B	1.0
Crash Mitigation	500 lives	$2 billion/yr	<$100 million/yr	$2B	0.1
Congestion Pricing	10% reduction in recurring congestion	$5 billion/yr (3)	~0 (4)	$5B	0.3
Traffic Incident Management	20% reduction in non-recurring congestion	$10 billion/yr (5)	~$1 billion (6)	$9B	0.6

Assumptions:
(1) $4 million per life saved
(2) 10,000 more highway patrol vehicles amortized at $20K/year; 4 officers/vehicle for 24/7 coverage at $100K/officer/year
(3) Based on $50 billion/year cost for recurring congestion
(4) Presumes low net public cost (revenue > expenses)
(5) Based on $50 billion/year cost for non-recurring congestion
(6) $10 million/year per metro area for 100 locations for staff, coordination, training, and information technology

2.6.3 Judycki and Robinson

Judycki and Robinson (*20*), in a review of NRD-related issues, make an interesting observation:

> "Delay due to incidents, however, cannot be planned for, and thus may have a greater impact on the movement of people and goods than does recurring congestion. As traffic volumes continue to increase, so do the frequency and impacts of incidents. More and more motorists are judging the integrity of our highway systems by how incidents are managed, or not managed, as the case may be."

They also make the observation that the impacts of incidents is not limited to traffic congestion:

> "Secondary accidents— those that occur at the end of a queue as high-speed traffic approaches an unexpected stopped or slow-moving backup—are a real and significant

direct result. Certainly more vehicle breakdowns and minor accidents occur in congested stop-and-go traffic, increasing the number of incidents and slowing the return of normal conditions after the initial incident is cleared. The danger to disabled motorists, police officers, and other responders is great, and the number of accidents involving vehicles abandoned or stopped on shoulders is also increasing rapidly, particularly in and near urban areas."

They continue by indicating that the time to clear an incident is critical:

> "An incident that blocks one lane of three on a freeway reduces capacity in that direction of travel by 50 percent, and even has a substantial impact on the opposing direction of travel because of rubbernecking. If traffic flow approaching the incident is high (near capacity), the resulting back-up can grow at a rate of about 8.5 miles per hour—that is, after one hour, the back-up will be 8.5 miles long.' Traffic also backs up on ramps and adjacent surface streets, affecting traffic that does not even intend to use the freeway. Observations in Los Angeles indicate that, in off-peak travel periods, each minute of incident duration results in 4 to 5 minutes of additional delay In peak periods, the ratio is much greater."

The implication is that using the time duration of an incident is an important performance metric in the context of NRD.

2.6.4 Lindley (1989)

Lindley (*25*) provides an assessment of the impacts of non-recurring incidents, and implicitly provides a sense of evaluation metrics that are important. And the percentage of all urban freeway delay caused by nonrecurring incidents, which was estimated at 61 percent in 1984, was estimated at 64 percent using 1987 data and is projected to increase to 72 percent in 2005. Table 2.3 shows the evaluation metrics considered in the study: delay in vehicle-hours, wasted fuel, and user costs.

Table 2.3 NRD Related Performance Measures

	1984	1987	2005 (using 1984 data)	2005 (using 1987 data)
Freeway Miles	15,335	16,097	15,335	16,097
Vehicle-Miles of Travel (billions)	276.6	337.4	411.0	492.5
Recurring Delay (million vehicle-hours)	485	728	2,049	3,030
Delay Due to Incidents (million vehicle-hours)	767	1,287	4,858	7,978
Total Delay (million vehicle-hours)	1,252	2,015	6,907	11,008
Total Wasted Fuel (million gallons)	1,378	2,206	7,317	11,638
Total User Costs (billion dollars)	9.2	15.9	50.5	88.2

2.6.5 Lomax, et al. (2001)

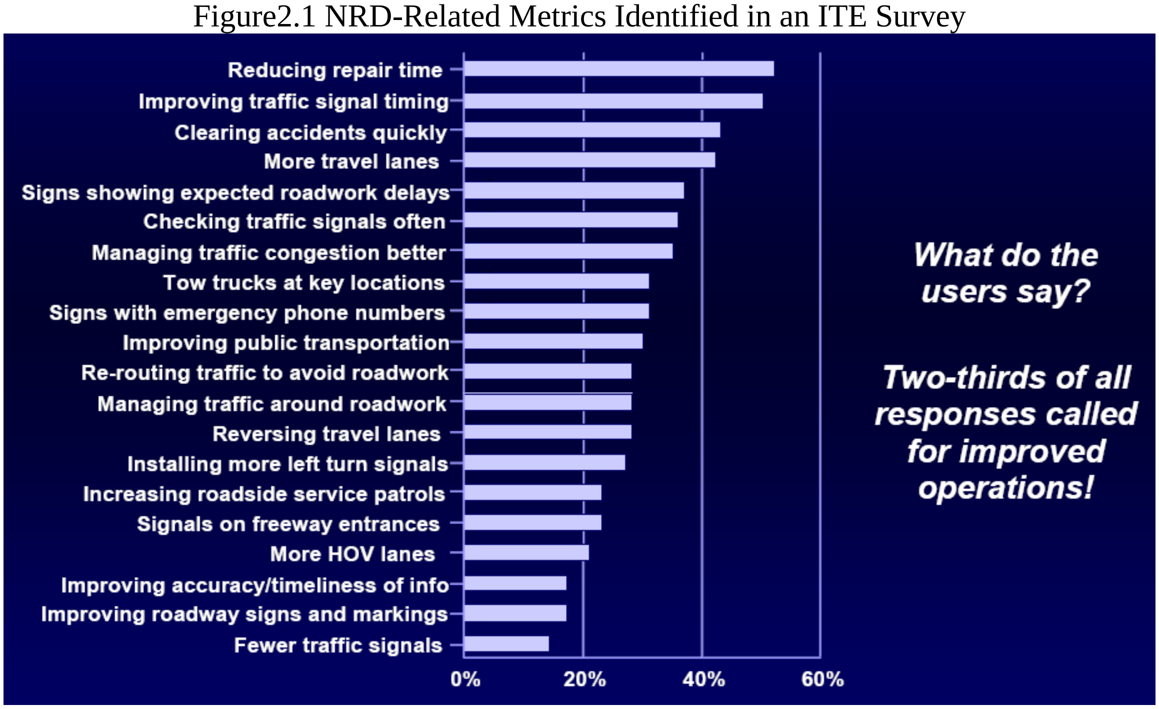
Lomax *et al.* (*27*) reinforce Lindley's ideas of important metrics.

> "In general, where roadways are already operating at capacity, even minor incidents can dramatically increase congestion. At these locations and times, an effective incident management system can significantly reduce congestion. Clearance of a single blocking car from a three-lane (one direction) freeway increases capacity by 50 percent. However, most motorists do not perceive these improvements because routine congestion (caused by geographical bottlenecks or simply too many vehicles) still exists. The biggest effect of incident management is in significantly reducing congestion and delay caused by major accidents. Although motorists still experience delays caused by these accidents, an effective incident management system can reduce what might be a 4-hour long, 5-mile traffic jam into a queue of less than 1 mile that lasts less than an hour."

2.6.6 Institute of Transportation Engineers

The ITE website (*19*) presents the findings from a survey that asked users what they thought would be the benefits from improved incident management. Figure 2.1 presents the findings. As can be seen, the most prevalent response was reducing the repair time for the incident and the third was clearing accidents quickly. Several other responses are similar, as in managing traffic congestion better, having tow trucks at key locations and re-routing traffic to avoid locations where roadwork is underway.

Figure2.1 NRD-Related Metrics Identified in an ITE Survey



2.6.7 Taylor, Saif, Sisiopiku (1999)

Taylor, Saif, Sisiopiku (*42*) tested the effectiveness of the ITS deployment in reducing off-peak incident-based congestion. Several evaluation metrics were employed. One measures incident-based congestion at a specific detector station. Another one measures the total delay over a freeway section. This metric is further categorized by incident frequency, queue length and duration. A third measures system-wide incident-related congestion. A fourth, called an incident-based congestion index, measures the amount of time that a station, segment, or the network is congested due to incidents. The index is used to estimate the average daily minutes of incident-based congestion by multiplying the index value times the number of off-peak daily minutes.

3.0 Project-Centered Thoughts

This section relates the findings from Section 2.0 to the current project. From the literature, it seems that *efficiency* is the main goal under which NRD-related concerns arise. Non-recurring delay is viewed as an impediment to system efficiency; a “cost” which should be minimized through good operational practice and careful geometric design.

3.1 Insights from the Literature

Three NRD-related objectives are mentioned most often: minimizing NRD-related vehicle-hours of delay, minimizing person-hours of delay, and minimizing ton-hours of delay. Sometimes, other objectives are also mentioned, like minimizing the cost of person-hour delay and minimizing the cost of goods movement delays. Occasionally, there is discussion about delays to other vehicle streams, like the one moving in the opposite direction that is delayed by rubber-necking.

NRD-related objectives also seem to be part of broader initiatives related to incident management. The *efficiency* concerns related to NRD are tied to *safety* and *reliability* concerns. Safety relates to the welfare of the people and goods involved in the incident, as well as those involved in secondary incidents, and emergency response personnel. Reliability relates to the fact that incidents cause unpredictable delays and system disruptions.

NRD is seen as a “cost” that can be mitigated by well-planned incident management. NRD delay can be reduced by clearing incidents quickly, by having well-thought-out maintenance of traffic plans for construction areas, and pro-active plans for responding to weather conditions.

3.2 New Thoughts

There do seem to be some new thoughts that can be added to the NRD discussion. One is a consideration of the cost of responding to incidents. It is not that such costs should be a major focal point, nor should they be minimized, but rather they should somehow be accounted for in developing NRD minimization strategies. There needs to be recognition of the fact that there is a tradeoff between minimizing NRD and expending resources to make that happen.

Another consideration is *security*. In addition to *efficiency*, *safety*, and *efficiency*, especially in the wake of 9-11, *security* should be a focus of incident management programs. This is not only the security of the transportation network given that an incident has occurred but also the security of the people and goods involved in the incident, especially if it was caused by a terrorist event. There is also the security of the environment around the incident if, as with a truck explosion, normal security is disrupted.

A third consideration is stakeholder perspectives. Although a set of generic goals and objectives may be sufficient, it is nice to find that those goals and objectives are consistent with the interests of all the stakeholders involved. This one deserves attention here, in the narrow sense of non-recurring delays.

At a minimum, there are stakeholders from six different groups: the travelers, the freight movers, the emergency response personnel, the adjacent land occupants, the transportation agencies, and society. The perspective of each should be considered to see what goals and objectives emerge as well as evaluation metrics.

The travelers are clearly interested in the quality of service provided. Unexpected delays are undesired. To maximize their daily productivity, and minimize the loss of leisure time, they would want to see non-recurring delays minimized. Person-hours of NRD per day per person is a plausible evaluation metric. They are also interested in consistency in the service provided. While one very long incident may have the same person-hours of NRD as a number of much shorter ones, probably, the larger number of shorter ones is preferred. So minimizing the duration of incidents is another defensible concern.

The freight movers are like the travelers. Unexpected delays are undesired. Maximizing profit means and minimizing lost time; making deliveries on time; maximizing the number of shipments handled per day. The freight movers would want to see non-recurring delays minimized. Dollar-hours of NRD per day per shipment is a plausible evaluation metric. Shipment-hours and ton-hours are other options. It depends upon the emphasis desired. Dollar-hours captures the economic value of the delays incurred. Shipment-hours and ton-hours are probably easier to estimate.

The emergency response personnel are probably interested in minimizing the length of incidents. That affects resource requirements, staffing needs, and costs. Shorter clearing times consume less resource-hours, and free up people, material, etc. sooner to respond to other incidents. A reasonable objective is minimizing the duration of incidents.

The adjacent land occupants are probably most interested in how long the incident lasts and how much interference the incident produces in terms of activities disrupted. A major truck accident that produces an evacuation is an excellent example. Minimizing the duration of incidents is probably a good objective. Another is minimizing the “size” of the incident, measured in a variety of ways, such as lane-miles, the number of vehicles delayed, and the number of lane-mile-hours.

The transportation agency probably has several goals and objectives. One is *efficiency*. It relates to the quality of the service provided. Delays are unwanted, whether they are NRD related or otherwise. Any kind of delay degrades the quality of service. So minimizing non-recurring delay would be an objective; so is minimizing person-hours of delay, or minimizing the shipment-hours of delay. These thoughts align with the interests of the travelers and the freight carriers.

The transportation agency is also interested in *safety*, *reliability*, and *security*. These aren't considered here in a significant way, but they are important. They are part of a multi-objective space in which tradeoffs must be made, collinearities exist, etc. For example, minimizing the average duration of an incident is also likely to enhance reliability and improve safety and security (less exposure time).

The transportation agency is also interested in controlling the cost of responding to the incidents that cause the non-recurring delays. So, as was said before, from the agency's perspective, there is a "tradeoff" between minimizing the non-recurring delays and minimizing the cost of responding to the incidents.

Society at large is probably most concerned with the social cost of the non-recurring delays. This includes the non-monetary metrics like person-hours of delay and $$-hours of delay for the goods shipments as well as the "costs" of secondary accidents and incidents, the costs of injured emergency response personnel (as well as those injured in the incidents themselves), the costs to the adjacent land users, the costs to the shippers and receivers whose goods are tied up or lost in the incident, and the economic losses in general. This is too broad a set of concerns for this specific project. It is going to be necessary to assume that these impacts can be estimated based on things like the duration of the incident and the number of people and shipments involved.

3.3 Summary

Thus, from the project team's perspective, the most relevant thoughts about NRD can be summarized as follows:

- The principal system goal that relates to NRD is *efficiency*
  - This includes how efficiently the system operates, as perceived by the users (both people and goods)
  - And the efficiency and effectiveness with which incidents are cleared by the system owner/operator.
- Thus, the NRD-related objectives seem to be
  - Minimize NRD-related delay
  - Minimize the duration of NRD-related incidents.
- Evaluation (performance) metrics which measure the extent to which NRD-related objectives are achieved include:
  - Incident duration
  - Vehicle-hours of delay
  - Person-hours of delay
  - Ton-hours of delay (for goods)
  - The cost of passenger delays (given a value for the cost of time)

    - The cost of goods-related delays (This includes the primary impact of an increase in inventory carrying cost, from the additional time that goods spend in inventory, and the indirect impacts of lost economic productivity due to the fact that goods are delayed in reaching the manufacturing plant, the point of potential sale, etc.)
- There is very little, if any, focus on *efficiency* as it relates to the cost of minimizing the NRD delay. This seems like a serious shortcoming in the analyses conducted to date. Obviously, it is implicit to "everyone" involved, that these costs exist, and they should be minimized, within the context of achieving NRD objectives, but the absence of attention to what the costs are and how they should be estimated, is unfortunate.
- NRD-related issues are part of a much broader focus on incident management
- The goals that relate to incident management are *efficiency*, *safety*, and *reliability*.
- The study team argues that *security* should be added to this list.
- The implication is that, especially from a "cost" perspective, focusing just on the costs of unexpected delays to people and goods is insufficient to characterize the impacts of the non-recurring delays. There are potentially other objectives, like minimizing the number of fatalities, that are of greater significance than minimizing non-recurring delay.
- Any plan for measuring NRD needs to take these issues into account in articulating a set of goals, objectives, and performance metrics to be employed in assessing NRD impacts.

4.0 Conclusions, Recommendations, and Next Steps

The project team's conclusions and recommendations are the following:

- Treat *efficiency* as the main system goal under which non-recurring delay is addressed.
- Also recognize that *reliability* is a system goal that is adversely affected by NRD.
- Treat the NRD-related objectives as being
    - Minimize the duration of NRD incidents
    - Minimize the person-hours of NRD-delay and the goods-hours of NRD-delay. (The latter could be measured in either ton-hours or $$-hours, with the latter metric reflecting the value of the commodities affected.)
- Focus on incident duration as being the principle evaluation metric estimated by the look-up tables.
    - It is measured in the field.
    - It is recognized as being an important objective by many studies.
- Other performance measures should be predicted analytically based on incident duration.
    - The field data will not support development of a validated, calibrated procedure for estimating such measures directly
    - These impacts increase proportional to the duration of the incident. Such impacts do not decrease in extent as the duration of the incident increases.
- Performance measures that ought to be derived from the incident duration include:
    - Person-hours of delay
    - Goods-hours of delay (either ton-hours or $$-hours, with the latter being preferable)
    - Cost of the person-hours of delay
    - Cost of the goods-hours of delay

- It must be recognized that these delays and costs can be attributable to the primary, causal incident as well as secondary incidents and "sympathetic" events, like rubbernecking. To the extent possible, these secondary impacts should also be estimated.
- NRD-related delays should always be thought-of in the much broader context of incident management. That is, non-recurring delays, while significant, may not be the over-riding concern in incident management.
- The other goals that seem to be important in incident management are *safety* and *reliability*.
    - Safety relates to the welfare of the affected parties in the primary incident, the welfare of additional affected parties from secondary incidents, and the welfare of emergency response personnel.
    - Reliability relates to the fact that incidents disrupt the systems operation and produce failures to deliver the service expected. People are late arriving at their intended destinations. Goods are delayed in transit.
- To these goals, the project team thinks *security* should be added.
    - To reflect the importance of the transportation network in terms of national security.
    - To recognize the fact that significant breaches in security can and do happen when incidents occur.
- In future efforts, it would be helpful to quantify the "costs" associated with these other incident-management related concerns, such as the cost of less-than-optimum safety, the cost of unreliable performance, and the cost of security failures.
- This project should provide estimating functions for non-recurring delay that would be compatible with such broader measures of incident impacts.

The next steps are to:

- Find models that have been developed to estimate the duration of NRD-related incidents.
- Identify ways to tie those models to the CNAM model.
- Find additional models that can estimate the other NRD-related performance measures based on incident duration and other factors (e.g., traffic volumes, auto occupancy, truck percentages, value of goods transported)
- Begin calibrating these models for NYC conditions.

REFERENCES

[1] Abbasi, M. F., V. P. Sisiopiku, and W. C. Taylor, "Development and Testing of a Freeway Recurring Congestion Measure", Proceedings of the 78th Annual Meeting of the Transportation Research Board, Washington D.C., Jan 1999.

[2] Arizona Department of Transportation, Freeway Management System: Infrastructure Design Guidelines, Sep 2002.

[3] Banks, J. H., "Performance Measurement for Traffic Management Systems", Transportation Research Record 1634, pp 39-45, 1998.

[4] Boarnet, M. G., E. J. Kim, and E. Parkany, "Measuring Traffic Congestion", Transportation Research Record 1634, 1998.

[5] Caltrans, Goods Movement Action Evaluation Methodology, State Goods Movement Strategy working draft issue paper #8, Jan 1998.

[6] Caltrans, California Highway Patrol, MTC, Bay Area Freeway Concept of Operations: Vision, Goals, and Objectives, Apr 2001.

[7] Central Arkansas Transit Authority, Arkansas State Highway and Transportation Department, and US DOT, Intelligent Transportation Systems Plan, Jun 2002

[8] Chu, L., H. X. Liu, and W. Recker, "Evaluation of Potential ITS Strategies under Non-recurrent Congestion Using Microscopic Simulation", Proceedings of the 83rd Annual Meeting of the Transportation Research Board, Washington D.C., Jan 2004.

[9] Chu, L., H. X. Liu, W. Recker, and S. Hague, Evaluation of Potential ITS Strategies Under Non-recurrent Congestion Using Microscopic Simulation, PATH Working Paper UCB-ITS-PWP-2003-2, Jan 2003.

[10] Cottrell, W. D., "An Empirical Freeway Queuing Duration Model", Proceedings of the 78th Annual Meeting of the Transportation Research Board, Washington D.C., Jan 1999.

[11] Cottrell, W. (Visalia, CA), "Estimating the Probability of Freeway Congestion Recurrence", Proceedings of the 77th Annual Meeting of the Transportation Research Board, Washington D.C., Jan 1998.

[12] Derr, B. R., "Delay Analysis for Freeway Corridor Surveillance, Communication, and Control Systems", Transportation Research Record 1132, pp 77-81, 1987.

[13] Dey, S. S., "'Hours of Congestion' as a Transportation Measure of Effectiveness Under Capacity Constrained Conditions", ITE Journal, Dec 1998.

[14] Dopart, K., Integrated Public Safety and Highway Operations: A Policy Framework and Analysis, Mitretek Systems, Inc., Aug 2001.

[15] Dowling, R., A. Skabardonis, M. Carroll, and Z. Wang, “Methodology for Measuring Recurrent and Non-Recurrent Traffic Congestion”, Proceedings of the 83rd Annual Meeting of the Transportation Research Board, Washington D.C., Jan 2004.

[16] Federal Highway Administration, Freeway Management and Operations Handbook: Chapter 10 Traffic Incident Management, Jan 2003

[17] Federal Highway Administration, Road Weather Management , found at http://ops.fhwa.dot.gov/Weather/index.asp, October 9, 2003.

[18] Forbes, G. J., “Identifying Incident Congestion”, ITE Journal, pp 17-22, Jun 1992.

[19] Institute of Transportation Engineers. “Getting at Congestion”, http://www.ite.org/management/ppoint/1getting.ppt, 2004.

[20] Judycki, D. C., and J. R. Robinson, “Managing Traffic during Nonrecurring Congestion”, ITE Journal, pp 21-26, Mar 1992.

[21] Kramer, R. F., “The A to **Z** Techniques to Reduce Congestion and Increase Capacity of Streets and Intersections”, ITE Journal, pp 49- 51, Mar 1988.

[22] Levinson, D., D. Gillen, and E. Chang, “The Value of Advanced Traveler Information Systems for Route Choice”, Draft, Nov 1999.

[24] Lindley, J. A., “A Methodology for Quantifying Urban Freeway Congestion”, Transportation Research Record 1132, 1987.

[25] Lindley, J. A., “Urban Freeway Congestion Problems and Solutions: An Update”, ITE Journal, pp 21-23, Dec 1989.

[26] Lindley, J. A., “Urban Freeway Congestion: Quantification of the Problem and Effectiveness of Potential Solutions”, ITE Journal, pp. 27-32, Jan 1987.

[27] Lomax , T., S. Turner, M. Hallenbeck, and R. Margiotta, “Traffic Congestion and Travel Reliability: How Bad is the Situation and What is Being Done About It”, found at http://www.ite.org/nationalsummit/vision/congestion.doc, Sep 2001.

[28] Louisiana Department of Transportation and Development, Louisiana Statewide Transportation Plan Update, Jul 2002.

[29] Metroplan (A Council of local Governments), ITS Plan: Central Arkansas Regional Transportation Study, Jun 2002.

[30] New Jersey Institute of Technology, Freight Planning Support System, found at http://www.njtpa.org/planning/freight/documents_freight/FPSSarticle.html, March 2004.

[31] New York State DOT, 21st Century Goals, Performance Measures and Objectives, found at http://assetmanagement.transportation.org/tam/aashto.nsf/0/e3e4ffae381d683585256d0f00450698?OpenDocument, April 2003.

[32] Oahu Metropolitan Planning Organization (Honolulu, Hawaii), OAHU Regional ITS Plan: Intelligent Transportation Systems Architecture and Integration Strategy, Apr 2003.

[33] Pal, R., S. P. Latoski, and K. C. Sinha, "Investigation of Freeway Incident Characteristics and Their Influence in Planning an Incident Management Program", Transportation Research Record 1634, pp 46-55, 1998.

[34] Parham, A. H., K. Fitzpatrick, M. D. Wooldridge, and D. Fenno, "Strategies to Facilitate Incident Management on New or Reconstructed Freeways", Proceedings of the 80th Annual Meeting of the Transportation Research Board, Washington D.C., Jan 2001.

[35] Regional Transportation Commission of Washoe County (Nevada), Chapter 10: Congestion Management, found at http://www.rtcwashoe.com/planning/downloads/pdfs/Chapter10.pdf, March 2004.

[36] Ritchie, S. G., and R. L. Cheu, Neural Network Models for Automated Detection of Non-Recurring Congestion, PATH Research Report UCB-ITS-PRR-93-5, Jun 1993.

[37] Schrank, D., and T. Lomax, The 2003 Annual Urban Mobility Report, Texas Transportation Institute, The Texas A&M University System, http://mobility.tamu.edu, Sep 2003.

[38] Shaw, T., and D. Jackson, "Reliability Performance Measures for Highway Systems and Segments", Proceedings of the 82nd Annual Meeting of Transportation Research Board, Washington D.C., Jan 2003.

[39] Skabardonis, A., P. P. Varaiya, and K. F. Petty, "Measuring Recurrent and Non-Recurrent Traffic Congestion", Proceedings of the 82nd Annual Meeting of Transportation Research Board, Washington D.C., Jan 2003.

[40] Smith, B. L., L. Qin, and R. Venkatanarayana, "Characterization of Freeway Capacity Reduction Resulting from Traffic Accidents", Journal of Transportation Engineering, pp 362-368, Jul/Aug 2003.

[41] Stamatiadis, C., N. H. Gartner, J. Winn, and R. Bond, "Evaluation of the Massachusetts Motorist Assistance Program: Assessment of Congestion and Air Quality Impacts", Transportation Research Record 1757, pp 1-9, 1998.

[42] Taylor, W. C., M. A. Saif, V. P. Sisiopiku, "Development of Freeway Incident-Based Congestion Measures", Proceedings of the 78th Annual Meeting of the Transportation Research Board, Washington D.C., Jan 1999.

[43] Transportation Research Board, Study for a Future Strategic Highway Research Program (F-SHRP): Draft Staff-Authored Overview, found at http://www4.nationalacademies.org/trb/newshrp.nsf/web/draft_staff-authored_overview?OpenDocument, Mar 2001.

[44] Transportation Research Board, Management of Freeway Emergencies and Special Events, Transportation Research Circular Number 326, Dec 1987

[45] Treasury Board Secretariat of Canada (TBS), IM/IT Investment Evaluation Guide, found at http://www.cio-dpi.gc.ca/cioscripts/cu-cn/contact_e.asp?who=/emf-cag/, 2000

[46] Wisconsin Department of Transportation, Traffic incident management enhancement, found at http://www.dot.wisconsin.gov/travel/smartways/time.htm, January 2004.

[47] Yang, H., K. K. Lo, and W. H. Tang, "Travel Time Versus Capacity Reliability of A Road Network", Proceedings of the 79th Annual Meeting of the Transportation Research Board, Washington D.C., Jan 2000.

Appendix A
More Thoughts on Evaluation Metrics

This appendix presents additional thoughts from the Freeway Operations Management Handbook about metrics that can be used to assess the impacts from NRD.

The delay rate (in minutes per mile) is the rate of time loss for vehicles operating in congested conditions on a roadway segment or during a trip (Equation 6). It is calculated as the difference between the actual travel rate and the acceptable travel rate. The delay rate can also be calculated as the difference (in minutes) between the actual travel time and the acceptable travel time divided by the segment length (in miles). The quantity can be used to estimate the difference between system performance and the expectations for those system elements, which is the key to prioritizing alternative improvements.

Mobility Aspect—Level, Location
Geographic Area—Section, Corridor

$$\begin{array}{c}\text{Delay Rate}\\\text{(minutes per mile)}\end{array} = \begin{array}{c}\text{Actual}\\\text{Travel Rate}\\\text{(minutes per mile)}\end{array} - \begin{array}{c}\text{Acceptable}\\\text{Travel Rate}\\\text{(minutes per mile)}\end{array}$$

$$= \frac{\text{Actual Travel Time} - \text{Acceptable Travel Time}}{\text{Trip or Segment Length (miles)}} \qquad (6)$$

Congested roadway describes the extent of congestion (in lane-miles) on the roadway or transit system. It is calculated by summing the congested segment lengths (Equation 12). The Urban Mobility Study statistics uses this information to illustrate the decline in roadway system operations.

Mobility Aspect—Location
Geographic Area—Section, Corridor, Subarea, Region

$$\begin{array}{c}\text{Congested}\\\text{Roadway}\\\text{(lane miles)}\end{array} = \sum \left[\begin{array}{c}\text{Congested Segment}\\\text{Length (lane miles)}\end{array}\right] \qquad (12)$$

The roadway congestion index (RCI) is calculated with the daily VMT per lane-mile of freeways and principal arterial streets. It is an empirically derived formula that uses HPMS data to quantify the relative congestion levels in urban areas. The RCI equation (Equation 13) weights the daily VMT per lane-mile values for the two functional classes by its respective amount of daily VMT. The functional class conditions are normalized by daily traffic volume per lane values that indicate the public perception of the beginning of undesirable congestion levels (values of 14,000 for freeways and 5,500 for principal arterial streets). RCI values greater than 1.0 represent an undesirable level of roadway congestion within an area.

$$\text{RCI} = \frac{\left[\begin{array}{c}\text{Freeway DVMT / Ln.– Mi. x}\\\text{Fwy. DVMT}\end{array}\right] + \left[\begin{array}{c}\text{Prin. Art. DVMT / Ln.– Mi. x}\\\text{Prin. Art. DVMT}\end{array}\right]}{\left[14{,}000 \text{ x Fwy. DVMT}\right] + \left[5{,}500 \text{ x Prin. Art. DVMT}\right]} \qquad (13)$$

The Lane-Mile Duration Index (LMDI) is a measure of the extent and duration of roadway congestion. The LMDI was developed by Cottrell in a study of congestion in 35 urban areas (*8*). The LMDI value for each urban area is the sum of the product of congested lane-miles and congestion duration (hours) for individual roadway segments (Equation 14), and is calculated using the indicator of average annual daily traffic volume per hourly capacity (AADT/C). The formula can be used for several roadway classes, but was used in a freeway analysis by assuming a v/c ratio greater than 1.0 (LOS F), or AADT/C ratio greater than 9.0 represented congested travel conditions. The study found this threshold of congestion to be consistent with the public's tolerance.

Mobility Aspect—Time, Location
Geographic Area—Subarea, Region

$$\text{LMDI}_F = \sum_{i=1}^{m} \left[\begin{matrix} \text{Congested} \\ \text{Lane - Miles}_i \end{matrix} \; x \; \begin{matrix} \text{Congestion} \\ \text{Duration}_i \text{ (hours)} \end{matrix} \right] \qquad (14)$$

where i equals an individual freeway segment, and m equals the total number of freeway segments in an urban area.

The Lane-Mile Duration Index (LMDI) is a measure of the extent and duration of roadway congestion. The LMDI was developed by Cottrell in a study of congestion in 35 urban areas (*8*). The LMDI value for each urban area is the sum of the product of congested lane-miles and congestion duration (hours) for individual roadway segments (Equation 14), and is calculated using the indicator of average annual daily traffic volume per hourly capacity (AADT/C). The formula can be used for several roadway classes, but was used in a freeway analysis by assuming a v/c ratio greater than 1.0 (LOS F), or AADT/C ratio greater than 9.0 represented congested travel conditions. The study found this threshold of congestion to be consistent with the public's tolerance.

Mobility Aspect—Time, Location
Geographic Area—Subarea, Region

$$\text{LMDI}_F = \sum_{i=1}^{m} \left[\begin{matrix} \text{Congested} \\ \text{Lane - Miles}_i \end{matrix} \; x \; \begin{matrix} \text{Congestion} \\ \text{Duration}_i \text{ (hours)} \end{matrix} \right] \qquad (14)$$

where i equals an individual freeway segment, and m equals the total number of freeway segments in an urban area.

The Congestion Severity Index (CSI) is used by the Federal Highway Administration in reporting the results of system analyses using Highway Performance Monitoring System data. It is included as a measure of progress toward the Mobility Goal in the Performance Plan (http://www.fhwa.dot.gov/aap/fhwapprt.htm) (9) and as one measure for economic growth and trade factors in the Department of Transportation's Strategic Plan (http://www.dot.gov/hot/dotplan.html) (10). The congestion severity index (CSI) has units of roadway delay per million vehicle-miles of travel (VMT) (Equation 15). The CSI methodology can use several different computer models or estimating procedures to estimate recurring delay. Nonrecurring delay is either estimated using an assumed distribution of incidents (based on VMT) and an incident model or with incident ratios. The user can specify the roadway classes included and the congestion threshold, but it should be noted in the analysis.

$$CSI = \left(\frac{\text{Total Delay (veh.-hrs.)}}{\text{VMT (1000)}} \right) \qquad (15)$$

TEA-21 permits use of local parameters for measuring facility performance and defining acceptable performance levels. Local performance measures for non-recurring congestion caused by incidents like construction, accidents and weather have not been identified in this CMS.

Task 2: Review of Non-Recurring Delay Models

Prepared by:

Kyriacos C. Mouskos, PhD, Research Professor
Camille Kamga, Ph.D. Candidate
Institute for Transportation Systems
University Transportation Research Center for Region II
Convent Ave, Bldg Y-220
NY, NY 10031
Tel: 212-650-8047
Fax: 212-650-8374
Email: mouskos@utrc2.org
ckamga@utrc2.org

And

Alice Huajing Shi, PhD
Weidlinger Associates, Inc.
375 Hudson Street
New York, NY 10014

And

George List, PhD, PE
Department of Civil and Environmental Engineering
Rensselaer Polytechnic Institute
Troy, NY 12180
Tel: 518-276-6362
Fax: 518-276-4833
Email: listg@rpi.edu

August 1, 2005

1.0 Introduction

This report presents a review of the models that have been developed for predicting non-recurring delay (NRD) or incident-related delay as it is also sometimes called. Reduction of NRD is one of the main objectives of transportation officials in every state. In many states the estimation of NRD for freeway systems is an integral part of their traffic management program. The report reviews the models currently in use, and those that have been developed in the past, and discusses their inputs, outputs, prediction methodologies, strengths, weaknesses, and underlying assumptions. The report includes matrices that show how the predictive models differ in the way they predict the NRD-related metrics identified in the Task 1 report.

Models for predicting incident duration and delay have been under development for nearly 30 years. While a few of these models have been brought into practice, most have only been hypothesized and tested in research projects. Even though this report includes descriptions of models that are both in use and the products of research, the ultimate goal, in Tasks 4 and 5 is to select a model for practical use. Consequently, the findings and conclusions need to be pragmatic and practical.

The ultimate recommendations have to be focused on implementation and take account of the fact that the Congested Network Analysis Model (CNAM) is presently the analytical tool being used by NYSDOT to assess non-recurring delay levels in the urban areas of the State and the Best Practice Model (BPM) is presently the one being used by the New York Metropolitan Planning Council (NYMTC) to do its planning assessments of capital investments.

With these objectives in mind, Section 2 reviews basic concepts about NRD, Section 3 describes NRD models that are in use in practice, Section 4 discusses other models that have been generated in research projects, and Section 5 presents recommendations for what models to move forward to Tasks 4 and 5. Appendix A provides greater detail on selected models.

2.0 Basic Concepts

Non-recurring delay arises because unexpected events occur that reduce a facility's capacity. If the demand flow exceeds the reduced capacity, a queue results and delays accrue to the detained vehicles. The causing event is often an accident or an incident, but it can also be construction work or inclement weather.

In a basic sense, NRD models can be classified based on five characteristics;

- *Link or network*: does the model focus on a single link or a network of links and nodes.
- *Deterministic or stochastic*: is the model deterministic, which means that all the uncertainty is left out, or is it stochastic
- *Static or Dynamic*: does the model attempt to predict NRD-related metrics on the basis of a static representation of the traffic flow phenomena or does it use a dynamic representation that allows for variations in the traffic and the traffic flow conditions
- *Modeling paradigm*: the basic technique on which the model is based like queuing analysis (e.g., Morales, 1986), shockwave analysis (e.g., Messer *et al*., 1973; Wirasinghe,

1978; Chow, 1974), and traffic simulation (e.g., Wickes and Lieberman, 1980; Stamatiades *et al*., 1998; Cambridge Systematics and Cohen (1998))

- *Single or multiple vehicles*: does the model focus on estimating delay for all vehicles (e.g., Morales, 1986; Messer et al., 1973; Wirasinghe, 1978; Chow, 1974; Wickes and Lieberman, 1980) or just for one vehicle (e.g., (Fu *et al*., 1997).

The notion of whether a model is focused on a link or a network is fundamental to its classification. Link-based models can address issues of NRD for isolated incidents that have no upstream or downstream impacts. Network-based models may be simpler in structure, but they can capture the upstream and downstream effects, as well as route diversion.

Static or dynamic is critically important because a static model has to impute what the NRD might be for a given incident given the dynamics that do take place based on a static representation of the incident and the system. Dynamic models are typically more complex, they involve more inputs, but they attempt to represent the changes in demand, queue length, capacity, etc, that occur in conjunction with the incident so that both the spatial and temporal aspects of the incident response can be seen.

Figure 1 presents a “simple” diagram that is often presented to describe the general ideas involved in NRD, the notion of reduced capacity, the duration of the incident and the resulting delays (see, for example, Morales, 1986):

- *Incident detection time* is the interval from the occurrence to the detection of the incident.
- *Incident response time* is the time between detection to the time the first response unit arrives.
- *Clearance time* is the time it takes for the incident to be removed from the road, and
- *Incident duration* which is the sum of these three above times, D
- *Recovery time* or residual delay is the time for the queue to dissipate and the demand flow rate to be restored after the incident has been cleared from the road, D_r
- *Incoming demand* is the upstream arrival rate, q_{a1}
- *Roadway capacity* is the maximum processing rate in the absence of the incident, q_c
- *Incident departure rate* is the maximum discharge rate while the incident is present, q_d
- *Incident departure rate after recovery* is the maximum discharge rate that is possible after the incident has been cleared, q_r
- *Incident departure rate after recovery* is the maximum discharge rate after the incident has been cleared, q_r
- *Incident delay* is the area inside the difference between the cumulative arrival and cumulative departure curves

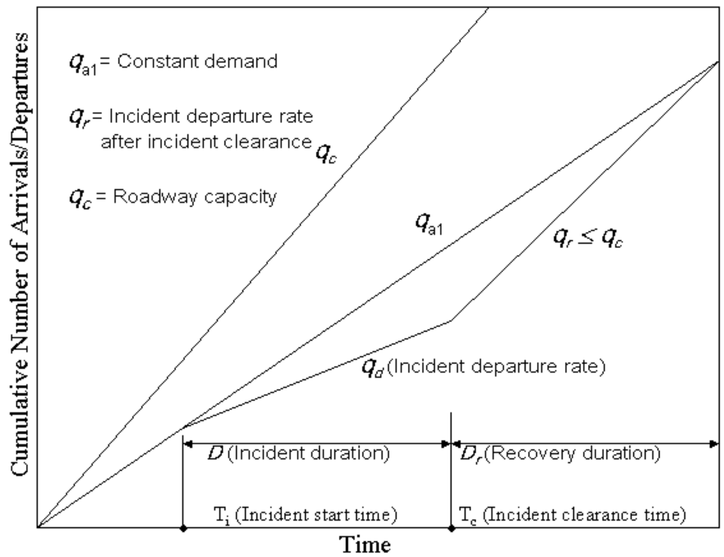


Figure 1- Incident Delay - Deterministic Model (one arrival rate)

Nominally, the incident duration starts when the incident happens, but that time is hard to document. Consequently, the start time that is recorded in most instances is the time when the incident is first detected / reported. Hence, the incident duration reported in the literature is a "modified incident duration", i.e., the incident duration based on the detection time.

Incident duration is the sum of incident detection time, incident response time, and incident clearance time. Most incident duration models are concerned with these three components.

The last component, residual delay or recovery time assesses the efficiency of the traffic control strategies used to recover from the event, such as traffic diversion and early traveler information systems. Not every incident involves all four components.

The delay equation that expresses total incident delay based on the simple model shown in Figure 1 is as follows:

$$Delay = \frac{D^2(q_r - q_d)(q_{a1} - q_d)}{2(q_r - q_{a1})} \qquad (1)$$

where D is the incident duration, q_{a1} is the rate of traffic flow just before the incident occurs, q_c is the saturation flow rate (prevailing roadway capacity) of the road segment where the incident occurs, q_d is the departure flow rate while the incident is present, and q_r is the departure flow rate (also called getaway rate) once the incident has been cleared..

A slightly more complex model takes into consideration the potential reduction in demand that might occur due to the incident (see Derr, 1987). As Figure 2 illustrates, the initial demand is the same as q_{a1}, the expected demand for the point in time when the incident occurs. That demand lasts for a certain time period D_1. The subsequent demand q_{a2} is smaller in magnitude and lasts for a time period ($D - D_1$). The reduced demand is presumed to occur due to information on the existence and nature of the incident and the availability of other alternative routes.

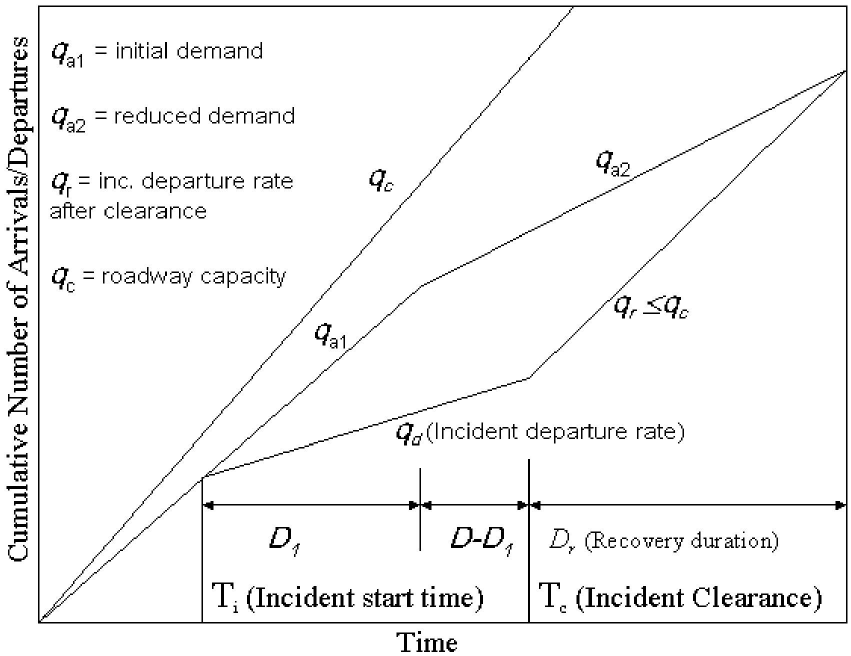


Figure 2- Incident Delay – Deterministic Model (Two-demand regime) Derr (1987)

3.0 Models in Practice

Based on a comprehensive review of the existing published literature, the project team has found that there are very few NRD models in use in practice. That is, if CNAM is taken as a benchmark against which to seek evidence that other states or other metropolitan planning agencies (MPOs) have developed similar models, the conclusion seems to be that if they exist, they have not been described in the published literature. In fact, CNAM has not been reported in the published literature either.

It seems reasonable to assume that every state and MPO has a method for estimating NRD. The work being done by XXXX at TTI makes it clear that this is likely to be the case. But it also appears, based on that work that in most instances, NRD is assumed to be estimable based on a multiplier applied to the recurring delay estimate.

Clearly, for specific projects, especially highway reconstruction efforts, estimates of NRD have been developed. Undoubtedly, too, a variety of methods have been used to compute the NRD ranging from simple multipliers to microscopic simulation models. But these efforts are not really relevant to the focus of this task, or this research project. The aim is to produce an enhanced version of CNAM, or a “new” model that can estimate NRD for an entire urban network on a segment-by-segment basis.

The three models that have been identified through the state-of-the-practice and state-of-the-art investigation are xXXX (TTI), IDAS, and CNAM. Each of these is described and evaluated next.

3.1 TTI Model - Shrank and Lomax (2003)

Shrank and Lomax (2003) present the results of their assessment of recurring and non-recurring delay for 2003. The publication has become a quasi-yearly document. Included with the 2003 report is an appendix that describes the methodology by which the delay estimates are developed.

Estimation of the recurring delay is quite detailed. It involves equations, nomographs, charts and tables. The procedure is applied to a large number of metropolitan areas. Location specific parameters are developed, defended, and employed.

Non-recurring delay is developed by applying a multiplier to the recurring delay. As the report states: "Another type of delay encountered by travelers is incident delay. This is the delay that results from an accident or disabled vehicle. Incident delay is related to the frequency of crashes or vehicle breakdowns and how easily those incidents are removed from the traffic lanes and shoulders. The basic procedure used to estimate incident delay in this study is to multiply the recurring delay by a ratio (Equation B-1)."

Equation B-1 is:

$$\begin{array}{ccccc} \text{Daily Freeway} & & \text{Daily Freeway} & & \text{Freeway} \\ \text{Incident} & = & \text{Recurring} & \times & \text{Recurring} \\ \text{Vehicle - Hours} & & \text{Vehicle - Hours} & & \text{to Incident} \\ \text{of Delay} & & \text{of Delay} & & \text{Delay Ratio} \\ \text{(Ex.) } 148{,}500 & = & 165{,}000 & \times & 0.9 \end{array} \qquad \text{(Eq. B-1)}$$

where the numerical values are those that pertain to a specific metropolitan area.

3.2 CNAM - NYSDOT (2003)

NYSDOT has developed the Congestion Needs Assessment Model (CNAM). It is a link-based incident delay model. It utilizes a modified queue diagram, shown in Figure 3, to estimate delays for each type of incident for each hour of the day. The CNAM model calculates incident delay per roadway segment independently. The roadway segments are defined as in the NYSDOT's Highway Sufficiency File or the Local Inventory File.

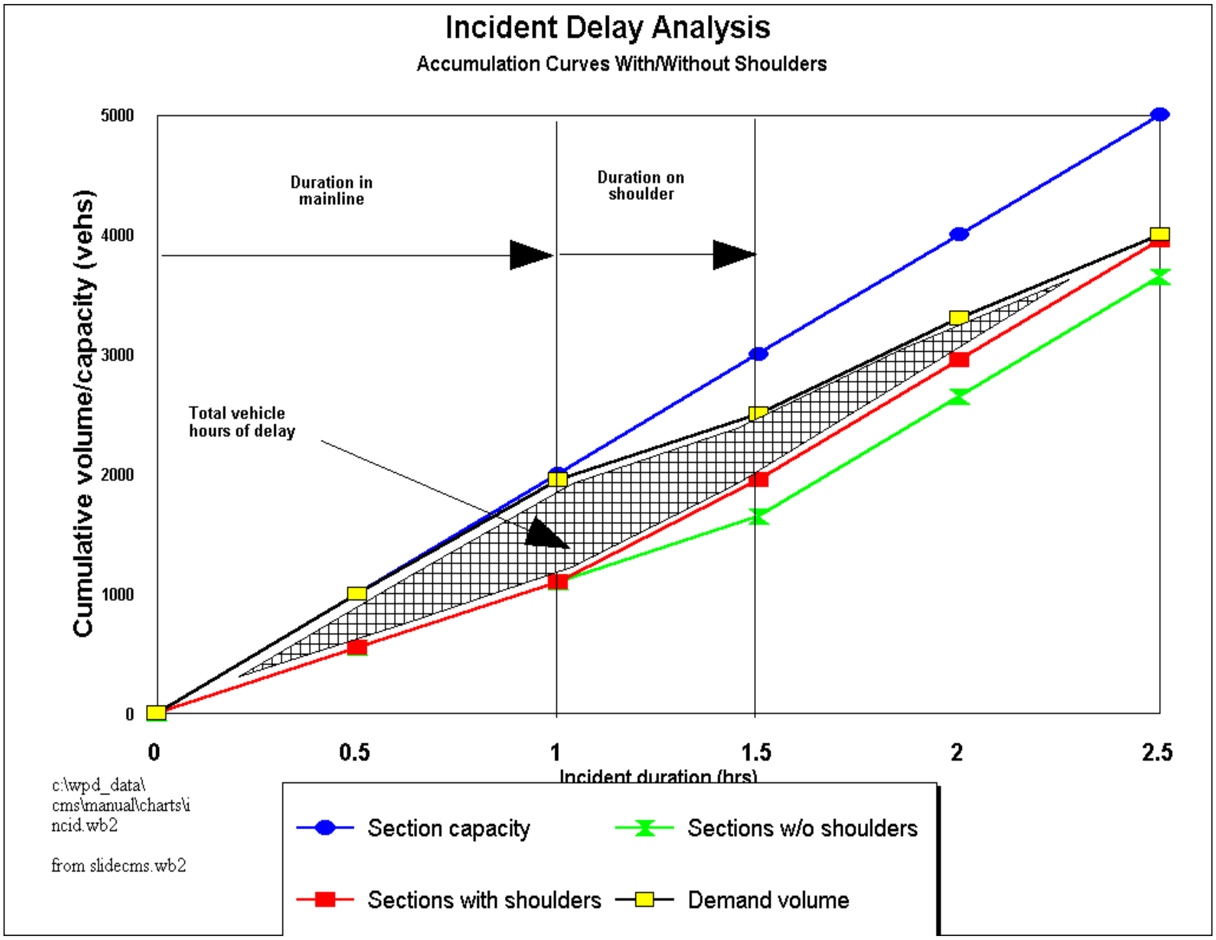


Figure 3- CNAM Incident Delay Analysis Diagram (source: CNAM Manual-Incident Analysis Module)

The incident delay model is based on a queuing model that utilizes the following parameters:

- *Incident occurrence time*: An incident is predicted to occur at a specific hour of the day. The mid-point of the predicted hour is also the estimated time of occurrence of the incident.
- *Demand* (*arrival rate*): For each incident type a demand (volume) accumulation function is developed based on the demand during and after the incident. CNAM produces estimates of hourly volumes for each hour of the day. These estimates are based on either traffic counts (if they exist) or on hourly estimates from factors based on the Annual Average Daily Traffic (AADT) - the AADT is recorded in the Highway Sufficiency File. A five-hour demand profile is generated for each incident type and for each hour of the day, which is based on the incident occurrence type. Each of the estimated hourly traffic volumes is used as the demand arriving at the incident location within the segment. Then the model estimates when the queue due to the incident dissipates. If the queue dissipates within four hours, then the incident demand profile is developed for the corresponding time (e.g. from the incident occurrence hour to the time that the queue dissipated) otherwise the incident demand accumulation function has duration of 5 hours total.

- *Getaway rate*: Given the incident type for the specific roadway segment, the capacity available subsequent to the incident is identified through the use of a look-up table, as shown in Table 1. It is useful to note that 1) the number of lanes is for both directions as defined in the sufficiency file of NYSDOT and 2) the factors are for accidents only.

Table 1- Incident Available Capacity: Default Percent (AVALCAP.DBF)

# Lanes	Attribute	Lanes Blocked			
		On Shoulder	1 Lane	2 Lanes	3 Lanes
4-5 Lanes	Field name	CAP_04	CAP_14	CAP_24	CAP_34
	Shoulder	0.0 (N/A)	42.5	0	0.0 (N/A)
	No Shoulder	0.0 (N/A)	42.5	0	0.0 (N/A)
6-7 Lanes	Field name	CAP_06	CAP_16	CAP_26	CAP_36
	Shoulder	87.5	56.7	23.3	0
	No Shoulder	0.0 (N/A)	56.7	23.3	0
8+ Lanes	Field name	CAP_08	CAP_18	CAP_28	CAP_38
	Shoulder	87.5	63.8	42.5	21.3
	No Shoulder	0.0 (N/A)	63.8	42.5	21.3

- *Type of incident*. Incidents are classified based on the number of lanes blocked: (a) Accidents on shoulder, (b) Accidents blocking one lane, (c) Accidents blocking two lanes, and (d) Accidents blocking more than two lanes.
- *Incident factor*: Defined as the percentage of the total incidents on the road segment for each incident type. Two types of incident factors are used: either the shoulder can be used during the incident (more than 6 feet) or it cannot (less than 6 feet). The incident factor is classified based on the number of lanes, incident type and location of the incident. Table 2 shows the default values of incident factors that are based on the above characteristics for each roadway segment. It is useful to note that N/A means not applicable and INC* is an abbreviation for INCFTR.

Table 2- Incident Factors: % of Accident Rate

# of Lanes	Attribute	On			
		Shoulder	1 Lane	2 Lanes	3 Lanes
4 - 5 Lanes	Field name	INC*_A04	INC*_A14	INC*_A24	INC*_A34
	Shoulder	0.48	0.44	0.08	0.00 (N/A)
	No Shoulder	0.00 (N/A)	0.85	0.15	0.00 (N/A)
6 - 7 Lanes	Field name	INC*_A06	INC*_A16	INC*_A26	INC*_A36
	Shoulder	0.47	0.43	0.08	0.02
	No Shoulder	0.00 (N/A)	0.81	0.15	0.04
8 or more Lanes	Field name	INC*_A08	INC*_A18	INC*_A28	INC*_A38
	Shoulder	0.47	0.43	0.08	0.02
	No Shoulder	0.00 (N/A)	0.81	0.15	0.04

Table 3 shows the estimates of incident duration used by CNAM for each area type. It should be noted that the number of lanes is for both directions as used in the sufficiency file of NYSDOT and where no shoulder exists, the duration of the incident on the shoulder is set to zero.

Table 3- Incident Duration: Default Minutes (INCIDUR.DBF)

Location	Attribute	Lanes Blocked			
		On Shoulder	1 Lane	2 Lanes	3 Lanes
Rural (Area Type 1)	Field name	LANE_01	LANE_11	LANE_21	LANE_31
	Lane clearance time	61	62	63	67
	Field name	SHLD_01	SHLD_11	SHLD_21	SHLD_31
	Shoulder clearance time	0	15	15	20
Pop < 5000 (Area Type 2or 3)	Field name	LANE_03	LANE_13	LANE_23	LANE_33
	Lane clearance time	37	38	39	43
	Field name	SHLD_03	SHLD_13	SHLD_23	SHLD_33
	Shoulder clearance time	0	15	15	20
Suburban (Area Type 4)	Field name	LANE_04	LANE_14	LANE_24	LANE_34
	Lane clearance time	51	52	53	59
	Field name	SHLD_04	SHLD_14	SHLD_24	SHLD_34
	Shoulder clearance time	0	15	15	20
City / Large Village (Area Type 5 or 6)	Field name	LANE_05	LANE_15	LANE_25	LANE_35
	Lane clearance time	43	44	45	51
	Field name	SHLD_05	SHLD_15	SHLD_25	SHLD_35
	Shoulder Clearance time	0	15	15	20

Several other observations are useful:

- *Incident Rate*: This is the number of incidents per million vehicle miles traveled (VMT) for roads by access control, area-type, number of lanes, and number of roadways.
- *Traffic flow growth factors*: The CNAM model utilizes traffic flow growth factors in estimating the projected volumes into the future for all roadway segments.
- *Queue Dissipation Time*: The queue dissipation time is the time required for the queue developed during the incident (from the time of occurrence up to the time that the incident is cleared) to dissipate. CNAM utilizes the following procedure to produce the queue dissipation time:
  - The time period is divided into 15-minute (default) time intervals - the model allows the user to select a smaller time interval (e.g. 5, 10 minutes) if desired.
  - At each time interval, the model compares the total volume serviced with the cumulative demand.
  - If the total volume serviced is larger than or equal to the cumulative demand then the specific 15-minute time interval is considered to be the queue dissipation time.
- *Incident Delay*: The incident delay is estimated based on the cumulative arrivals, the service volume during the incident, the service volume after the incident is cleared (also known as getaway rate), the incident duration and the queue dissipation time.
- *Annual Incident Delay*: The total incident delay for each roadway segment is calculated based on the accident rates and the incident delay calculation per incident type and hour of the day. Total annual incident delay for each specific hour per roadway segment is the total incident delay (sum of the four incident types) times 260 (number of days).

3.3 IDAS - Cambridge Systematics (1998)

One of the most widely used incident analysis models in practice for incident delay that is the Incident Delay Analysis System (IDAS). The IDAS model is also a single link based model that produces estimates of incident delay based on a single arrival rate for the time period of analysis. The model has received mixed reviews and is not always perceived as being a valuable tool, but it is reviewed here because it does address NRD in a comprehensive fashion. IDAS is primarily based on Figure 1 presented earlier.

Incident delays are analyzed for two, three, and four lane freeways as a function of the volume-to-capacity ratio. The model produces estimates of incident delay based on the above parameters and the basic parameters involved in the deterministic incident delay model presented with Figure 1: arrival volume (q_{a1}), roadway capacity (q_c), incident departure rate (q_d), the average incident departure rate during recovery (q_r), and the incident duration (D).

In addition, the following parameters are defined:

- r = capacity reduction factor due to the incident.
- $q_d = r\ q_c$, the incident departure rate. If r = 0, the freeway is completely blocked by the incident (all lanes closed).
- g = average "getaway" volume from the queue after the incident is cleared, expressed as a fraction of q_c, where $g = q_r/\ q_c$.
- Q = maximum queue length (in vehicles)
- d_i = delay incurred during the incident (in vehicle hours)
- d_r = delay incurred during the "getaway" period (in vehicle hours)
- d = total delay incurred as a result of the incident (in vehicle hours)

The queue growth rate q_{ad} during the incident (in vehicles per hour) is equal to the rate at which vehicles arrive at the end of the queue (q_{a1}) minus the rate at which they get past the incident (q_d):

$$q_{ad} = q_{a1} - q_d \qquad (3)$$

The maximum queue length, Q, occurs at that point in time when the incident is cleared:

$$Q = (q_{a1} - rq_c)D \qquad (4)$$

The average queue length (Q/2) is based on the deterministic incident delay. The queue grows from a length of zero (when the incident occurs) to a length of Q (when the incident is cleared). The corresponding total delay incurred by vehicles during the incident is calculated as follows:

$$d_i = (\tfrac{1}{2})QD = (\tfrac{1}{2})(q_{a1} - rq_c)D^2 \qquad (5)$$

After the incident is cleared, the queue gradually dissipates, at a rate depending on the getaway capacity and the volume:

$$T_r = \frac{Q}{(gq_c - q_{a1})} \qquad (6)$$

Hence, the delay incurred by vehicles is:

$$d_r = (\tfrac{1}{2})QT_r = (\tfrac{1}{2})Q^2/(gq_c - q_{a1}) = (\tfrac{1}{2})(q_{a1} - rq_c)^2 D^2/(gq_c - q_{a1}) \qquad (7)$$

The total delay due to the incident is:

$$d = d_i + d_r = (\tfrac{1}{2})q_c D^2 (q_{a1} / q_c - r)(g - r)/(g - q_{a1} / q_c) \qquad (8)$$

A spreadsheet applies these equations to estimate the mean and variance of delays due to incidents as a function of volume-to-capacity ratio (noted as v/c in the following equations). In the spreadsheet, incidents are classified by type (abandoned vehicle, accident, debris, mechanical/electrical, stalled vehicle, flat tire, and other) and severity (shoulder; one, two, three, or four lanes blocked). For each class of incident, the data from Sullivan et al. (1995) is used, supplemented by FRESIM runs, to estimate the following quantities:

- Frequency (number of incidents per million vehicle miles),
- *r* and *g*, as defined above,
- $\overline{D}$ and $Var[D]$
- $E[D^2] = Var[D] + \overline{D}^2$

The model produces estimates of the mean and variance of delays for v/c ratios ranging from 0.05 to 1.0 for freeways with two, three, and four lanes in each direction.
The following curves are fitted to the results:

- Freeways with Two Lanes in Each Direction:
    - $\overline{d} = 0.0154(v/c)^{18.7} + 0.00446(v/c)^{3.93}$
    - $\text{var}(d) = 0.00401(v/c)^{18.7} + 0.00163(v/c)^{3.93}$
- Freeways with Three Lanes in Each Direction:
    - $\overline{d} = 0.0127(v/c)^{22.3} + 0.00474(v/c)^{5.01}$
    - $\text{var}(d) = 0.00275(v/c)^{22.3} + 0.00148(v/c)^{5.01}$
- Freeways with Four or More Lanes in Each Direction:
    - $\overline{d} = 0.00715(v/c)^{32.16} + 0.00653(v/c)^{7.05}$
    - $\text{var}(d) = 0.00152(v/c)^{32.16} + 0.00178(v/c)^{7.05}$

where:

- $var(d)$ is the variance of delay due to incidents in hours^2 per vehicle mile,
- $\overline{d}$ is the average delay due to incidents in hours per vehicle mile,
- *v* is the arrival volume in vehicles per hour (q_{a1}),
- *c* is the roadway capacity in vehicles per hour (q_c).

The equations are not applicable when the demand-to-capacity ratio is less than 1.0 in the time period preceding the incident. That is to say that the methodology is applicable only to incidents that start during uncongested conditions. It is a combined model that is based both on the analytical results of the deterministic incident delay model and utilizes empirical data to determine the incident departure rate, incident duration, the incident severity, and the incident type as defined above.

The IDAS model was developed by the Cambridge Systematics for the Federal Highway Administration (FHWA). This is the main reason that it has been used widely. However, it suffers from the same deficiencies that single link delay models suffer from – it ignores the spatial characteristics of incident delay that produces statistically deficient results for incident delay.

3.4 Concluding Thoughts

It seems clear that the TTI model is not well suited to NYSDOT needs. It lacks the detailed representation of the congested urban network which is central to the objectives of the project. It is most likely quite appropriate for the purpose it presently serves, but adopting it as a replacement for CNAM seems ill-advised.

CNAM appears to have the general properties needed to produce estimates of NRD for congested urban networks. It needs look-up tables that are valid for the area under study, which is the purpose of this project, in the context of Region 11. It also would benefit from having an enhanced ability to capture the "system-level-impacts" that almost invariably arise in when accidents and incidents occur – the queuing delays that accrue on adjacent upstream, and sometimes, downstream facilities, due to detained traffic and the "secondary" or "external" delays that arise on nearby facilities due to diverted traffic. These aspects of CNAM can be enhanced in the future.

IDAS seems useful as a post-processor to be used in conjunction with a standard planning model, like the BPM, but as a stand-alone entity, it does not appear to offer significant advantages over using CNAM. It has about the same functionality, perhaps with greater "elegance", but not significant methodological features lacking in CNAM. Moreover, if it were to be adopted, an interface would have to be created between, most likely, the BPM so that the BPM could feed IDAS the traffic volumes, *v/c* ratios, etc., that IDAS needs to have to develop its estimates of NRD and then those estimates would have to be validated against field data, which is the same as what needs to be done for CNAM.

The one unanswered question is whether it might be possible to use the BPM directly to estimate NRD, but the sense of the project team, from discussions with NYMTC and NYSDOT is that the BPM may be too resource intensive to make it a convenient if not convincing basis for developing NRD estimates.

4.0 Other Models

Numerous other models have been developed, and to some degree tested, through research aimed at finding ways to estimate NRD. This section provides an overview of these models. Appendix A provides more details for those models where the project team felt it was useful to provide that additional information.

Given the institutional barriers associated with the reporting and archiving the incident data by the emergency teams as well as the transportation management teams, it is not surprising that it is a challenge to retrieve the relevant data for the development of incident duration estimation models. Several datasets have been used repeatedly to develop different models, by the same or different researchers.

The models that have been developed seem to fall into three categories. Either they predict incident duration, capacity reduction, or delay. Some models predict duration and delay, obviously using duration to help estimate delay. No model seems to address all three issues.

4.1 Incident Duration

Two approaches have been followed in creating duration prediction models: 1) analysis of historical incident databases and 2) surveys of experts. This section focuses primarily on models developed from historical data. However, there is evidence that experienced incident response teams or highway police officers can give credible estimates of the time needed to remove an incident and recover the road capacity after they determined what the incident conditions are and where it is located. They seem to be able to do this hypothetically and on site.

Generally, incident duration is affected by location, severity, and the response procedures of the local emergency services. Total duration seems to have a large variation. For example, Giuliano (1989) reports that the mean duration is about 37 min with a standard deviation of 30 min while Cohen and Nouveliere (1997) indicate that the mean duration is 26 min with a standard deviation of 23 min based on the respective incident databases they used.

An early study on incident duration was conducted as part of the John Lodge Freeway project. Duration data were collected from logs kept by CCTV observers between June 1962 and June 1963. A total of 927 lane-blocking incident durations were analyzed. It was found that accidents had an average duration (not incident duration) that vehicles remained in the travel lanes (not the same as incident duration) for 6.14 minutes. Vehicle disablements averaged 5.24 minutes (DeRose, 1964).

Goolsby (1971) collected incident duration data from police and reported an average duration time of 45 minutes for non-injury accidents and 18 minutes for vehicle stalls. In this case, duration is measured from the time of detection (by observers) to the time when the incident was cleared. Large standard deviations of 19 minutes for accidents, 15 minutes for stalls were observed in both cases, correspondingly. The authors cite weather conditions, incident severity, and police workload as contributing factors to the observed high variances.

Data on incident duration was collected by the California Department of Transportation (Juge, Kennedy, and Wang, 1974). Duration is measured from the time the incident is observed until it is cleared. A total of 196 incidents are recorded by time-lapse photography over a 17-month period in 1973-1974. The mean reported duration of all incidents is 42 minutes.

Sullivan (1997) generated an empirical model to estimate the number of freeway incidents and their associated delays. Sullivan computes incident durations as weighted averages of the clearing times reported in many study data sets. A fixed time is added in increments representing detection and response times determined from judgment based on data sets where total incident durations are documented. The incident duration distribution is determined by the incident type, existing incident management and the incident location. Moreover, using the expected mean and standard deviation of each incident's duration, Sullivan found it was possible to characterize the entire duration distribution. Sullivan's model uses the 20^{th}, 55^{th}, 80^{th}, and 95^{th} percentile duration corresponding to each accident situation to estimate the overall duration. The 20^{th} percentile values range from 4 to 20 minutes and the 95^{th} percentile values range from 60 to 130 minutes.

As research on incident duration has evolved, regression-based models (or truncated variants) and hazard-based models have emerged as the two primary methodological choices:

- Regression models offer the advantage that they are more easily understood and interpreted.
- Hazard-based models have the advantage that they allow the explicit study of duration effects (i.e. the relationship between how long an incident has lasted and the likelihood of it ending soon).

Regression models are commonly used because incident duration is a continuous variable. Hazard-based models are often selected because incidents can be assumed to have an increasing or decreasing hazard function (e.g., see Jovanis and Chang, 1989; Jones *et al.*, 1991).

4.1.1 Regression-Based Analyses

Using data from the California Highway Patrol dispatch logs, Giuliano (1989) developed a statistical model to estimate incident duration based on incident characteristics. The study stops short of developing a general model to predict incident duration because of data limitation, but two separate models are developed, one for all incidents and the other just for accidents. The study finds that the impact of incidents on freeway operation depends on many factors including their frequency, location, severity, time of day, and the level of usage of the facility. In these two models, qualitative categories are used representing incident types, lanes closed, the time of day, accident type, and truck involvement as independent variables. One of the most important findings is that there is a highly significant effect from truck involvement, which means that the presence of a truck in an accident increases the total duration. This in essence means that an incident duration model should consider trucks as an independent variable.

Golob *et al*. (1987) confirmed the findings of Giuliano through an analysis of freeway accidents that involved trucks. The incident durations are found to follow a lognormal distribution. Golob *et al*. theorized that the duration of the incident comprises sequential stages, each of which may be influenced by the preceding activity. They model the incident duration as a random variable within a traditional deterministic queuing model approach with a known probability density function. Kolmogorov-Smirnov tests of the truck data support an hypothesis that the lognormal distribution can be used to describe all incidents and each specific incident type. Using this model, they found it possible to estimate the probability of an accident in any group resulting in a duration greater than a fixed time - in effect this fixed time captures the incident response time plus some of the incident clearance time.

Golob *et al*. (1987) also noted that the immediate consequences of an accident differ according to the collision type. They measured consequences in terms of the numbers of injuries and fatalities, duration of the incident (the elapsed time from accident occurrence to the clearing of hazards and obstacles), and the number of lanes or ramps closed, if any.

In a 1991 unpublished paper, Khattak *et al.* (1991) studied incident clearance data for 121 incidents that occurred in the Chicago area. They found 9 statistically significant variables: severe injuries, the number of heavy vehicles involved, the presence of heavy loads, liquid or uncovered broken loadings in heavy vehicles, freeway facility damage, the presence of sand and salt operations, the use of a heavy wrecker, extreme weather conditions, and assistance from other response agencies. Two other variables seemed to have an effect but were not statistically significant: response time and incident reporting.

Khattak *et al*. (1995) published a study of incident data collected from 1989 and 1990 for the Chicago area. They found that a one-minute reduction in response time of the first rescue vehicle decreases the incident duration by slightly more than one-half minute. Considering that, in their study, the average response time is 7.5 minutes and the average incident duration is 71.6 minutes, so the potential for reducing the duration through a reduction in response time is limited. However, response time is critical nonetheless for injuries.

Khattak *et al*. (1995) also develop a conceptual structure for describing the relationship between incident duration and contributing factors. The most important variables in the incident duration prediction are incident characteristics and the consequent emergency response actions. The analysis showed that incident durations are longer when:

- the response times are higher,
- the incident information is not disseminated through the public media,
- there are severe injuries,
- trucks are involved in the incident,
- the trucks have heavy loads,
- State property is damaged, and
- the weather is bad.

More details are provided in Appendix A. These are not significant revelations, but they are affirmations of intuition.

Garib *et al*. (1997) developed a model based on six variables: the number of lanes affected, the presence of trucks (yes/no), the time of day, the police response time, and the weather conditions. The model was able to explain 81% of variation contained in the durations within the dataset. No other variables were found to be significant. The model does not include variables for property damage, driver age, injury accident, and number of injuries. The model indicates that: 1) the police response time is a highly significant factor in predicting the incident duration followed by 2) weather condition, 3) time of day, 4) truck involvement, and finally 5) the joint effect of the number of vehicles involved and the number of lanes affected. This model supports the premise that the incident duration follows the lognormal distribution.

Ozbay *et al*.(1997) were among the first researchers to recognize that the non-homogenous nature of the incident duration data interferes with the ability to use traditional linear regression for model estimation. Ozbay and Kachroo (1999) reported that the duration values did not follow either a lognormal or log-logistic distribution. This created a need for using non-parametric estimation techniques such as the classification tree approach that was employed in Ozbay and Kachroo (1997, 1999).

Hall (2000) focused on response and dispatch time and the contribution of these two times to congestion and delay. The total clearance time (T) is defined as the sum of the incident detection time (I), waiting time from incident detection until clearance vehicle is dispatched (W), the response time from dispatch until arrival (R), and the service time to clear the incident, subsequent to arrival of response vehicle (S). ($T = I + W + R + S$). Incident response time is modeled as a function of distance from the response vehicle to the incident, along with their relative direction of travel, the positioning of interchanges, and the presence of congestion that may slow incident response.

Smith *et al*. (2001) investigated various models for forecasting the clearance time of freeway accidents. Nonparametric regression was employed, along with Weibull and lognormal distributions and a classification tree. The dependent variable is the incident duration. The independent variables are divided into three categories:

- physical variables: accident time of day, the day of the week and the weather;
- vehicle variables: include number of vehicles, truck involvement, and passenger bus involvement; and
- response variables: binary variables of emergency agencies responding to the scene.

The classification tree is based on clearance time: short (1-15 min.), medium (16-30 min), and long (31 or greater – they are considered to be detrimental to traffic operations). These definitions are chosen based on practical experience. The independent variables are listed in Table 4 along with their weights.

The main MOE is the prediction accuracy. That is the percentage of test accidents in which the clearance time is predicted correctly. Other MOEs are:

- The percentage of long clearance times that are predicted correctly and
- The accuracy of the predicted durations for short clearance time accidents.

Table 4- Independent Variables used in the Classification Tree Model
(Source: Smith *et al*. 2001)

Variable	Name	Weight W(*x*)	Value
PEAK	A	3.43	1 = Peak (6-8am, 4-6pm)
			2 = Off-peak
WEEKDAY	B	3.9	1 = Weekday
			2 = Weekend
EMS	C	16.07	1 = Yes
			2 = No
FIRE	D	15.28	1 = Yes
			2 = No
HAZMAT	E	97.27	1 = Yes
			2 = No
POLICE	F	9.17	1 = Yes
			2 = No
VDOT	G	24.78	1 = Yes
			2 = No
TOW	H	20.83	1 = Yes
			2 = No
NUMVEH	I	WI12 = 7.44	1 = Single vehicle
		WI13 = 6.39	2 = Two vehicles
		WI23 = 13.83	3 = Three or more vehicles
TRUCK	J	16.1	1 = Yes
			2 = No
BUS	K	11.01	1 = Yes
			2 = No

4.1.2 Hazard Function-Based Analyses

Hazard-based models are mostly used in the biometrics and industrial engineering fields. They use conditional probabilities to find the likelihood that an incident will end in the next short time period given its continuing duration.

Jones *et al*. (1991) described multivariate statistical models for accident frequency and duration prediction. They used the hazard function as central to the accident duration statistical estimation method. For example, they indicate – based on their data set - that on average there is a 9-minute time lag between accident occurrence and trooper arrival at the accident scene.

Jones *et al*. (1991) also designed their data collection effort so as to study the factors that affect accident frequency (number of occurrences) and duration. Accident frequency data are derived from two sources: (1) Washington state accident records and (2) special events information. To account for the differing geometries, the routes are subdivided into six roughly homogeneous zones. The data were collected from all accidents that occurred in the six study zones from April 1987 to March 1989. In all, 5,637 accident reports are in the dataset. The primary source of accident duration data is State Patrol dispatch records. Since the analysis of accident duration requires information available from the accident reports as well as from the dispatch data, Jones *et al*. found it necessary to match the two data sets. Of the 5,637 accident reports, only 2,156 or roughly 50% could be matched with corresponding dispatch data - the accident duration data were drawn only from accidents that could be matched. This matched data is found to be biased toward severe accidents, because of more careful reporting. This is expected to have a major contribution to the amount of error in the models. One of the conclusions that could be made from this study is that the incident data sets have inherent data quality problems.

Jones *et al*. (1991) also assumed that the number of traffic accidents in a given day follows a Poisson distribution. The accident frequency model has four classes of independent factors: seasonal effects, weekly trends, special events, and environmental. Due to the susceptive reliability of the volume data obtained from magnetic loop detectors, the traffic volume is not directly used as an independent variable. Instead, most of the independent variables included in the model are proxies for traffic volume, and they are used to implicitly capture traffic variability. More details are in Appendix A.

Nam and Mannering (2000) used hazard-based incident duration models to statistically evaluate the incident duration (including detect/report, response, and clearance of incidents) using two-years of data from Washington State's incident response team program. In general, referred to as duration dependence, the probability of an incident ending is dependent on the length of time the incident has lasted. Similar to Jones *et al*. (1991), Nam and Mannering also use a hazard function to formulate the incident duration model in terms of the conditional probabilities of interest and provided insights into duration dependence.

Instead of using an accelerated lifetime model as Jones *et al*. (1991) did, Nam and Mannering (2000) used a proportional hazard model to include covariates that affect the incident duration time. The proportional hazard model operates on the assumption that covariates act multiplicatively on an underlying or baseline hazard. They apply the proportional hazard-based duration models to statistically evaluate the time it takes to detect, report, respond to, and clear the incident. The variable of interest in the duration analysis is the length of time that elapses from the beginning of an event until its end.

Nam and Mannering estimate the duration models using exponential, Weibull, lognormal, log-logistic, and Gompertz distributions. They select the model that provides the best fit as measured by the likelihood ratio statistic. The model estimation shows that a wide variety of factors significantly affect incident duration, and that different distributional assumptions for the hazard function are appropriate for the different incident times (i.e., detection/reporting, response, and clearance times). It is also found that the estimated coefficients are not stable between the two years of data used in model estimation. The stability of incident durations over time is assessed, by using likelihood-ratio tests.

Qi (2002) considered the application of hazard based models and conditional probabilities to an incident database obtained from NYSDOT. Qi (2002) considers the following variables to be included in the log-logistic incident duration model:

- Temporal characteristics (weekday, am peak, pm peak, night);
- Weather characteristics (snow, rain);
- Incident location (On-ramp, at the entrance of Off-ramps, at the exit of On-ramps); This information was not available within the incident database so this variable was not included in the model.
- Incident characteristics (incident type, number of vehicles involved, number of lanes blocked);

- Involved vehicle characteristics (types of vehicles involved, occupants or vehicle load); The occupants or vehicle load were not available so they were not considered in the development of the model.
- Incident clearance source (response time of first rescue vehicle, agency involved, tool used).

The incident type is also estimated using the corresponding probability that an incident *n* will be of type *i*. This is then used to calibrate the incident duration model based on the incident type. Under the incident type category, incidents that involved property damage, injury or fatality yield higher incident durations than do disabled vehicles. Another finding is that the incident duration is shorter when police is involved than when only NYSDOT or a tow truck is involved. A third finding is that rain does not have an impact on incident duration. As expected, snow shows a significant impact.

From a temporal standpoint, Qi finds that the AM and PM peak periods produce higher incident duration times than the off-peak periods. Interestingly, the night time-period produces the highest incident duration times.

Qi (2002) also developed an on-line incident duration prediction model using hazard-based duration regression models. Three separate incident duration models are developed. They pertain to: 1) incident report time to incident verification, 2) incident verification to incident clearance, and 3) incident clearance to the end of the incident (e.g., queue dissipation).

The incident duration prediction model estimates of the remaining incident clearance time given that the incident has lasted for some time period. The truncated median incident duration distribution is used as in the off-line model. The three models use the following characteristics to describe the incident: weather (snow, rain), time period (am peak, pm peak, night), incident type (property, injury/fatality, disabled vehicle), and incident clearance agency (police, NYCDOT, tow truck). Only some of these variables are used in each model. More details can be found in Appendix A.

4.2 Capacity Reduction

An important factor in incident duration is the capacity that pertains during and after the incident occurs. Goolsby (1971) studied capacity reductions caused by incidents using detailed logs coupled with video surveillance along the Gulf Freeway in Houston, Texas. By comparing the volumes under normal conditions to the volumes under incident conditions, Goolsby was able to determine the capacity reduction along the three-lane Gulf Freeway. He found that a 1-lane blockage by a minor accident or stall reduce flow by 50% even though the physical reduction is just 33%. An accident that blocks two lanes reduces the capacity by 79%. The presence of an accident on the freeway shoulder reduces the capacity by 33% of normal flow because of the effect of “apers-block” phenomenon.

Smith *et al*. (2003) analyze capacity reduction using a larger sample size. Their incident data is from the three-lane Hampton Roads region of Virginia. The flows during the incident conditions are compared to normal conditions in order to determine the reduction in capacity due to incidents. Smith et. al determine that incidents cause a greater reduction in capacity than what Goolsby and Smith establish.

McShane *et al*. (1990) presented an example to illustrate the effect of capacity reduction on the v/c ratio. They consider three different values for the v/c ratios and then they simulated the losses in capacity due to an incident by changing these three values by different percentages. They conclude that decreasing capacity by 10% or more may change freeway operation from a functional system to an oversaturated system.

Stamatiadis *et.al*. (1998) reported a 15% capacity reduction for shoulder minor incidents and a range between 39 to 69% for lane accidents. In contrast they reported that for incidents that were not served by the Massachusetts Motorist Assistance Program (MAP) the corresponding capacity reduction for shoulder minor incidents was 19%. The difference is attributed to the presence of a police vehicle with flashing lights for non-MAP assisted incidents versus the presence of an MAP van with flashing lights for MAP assisted incidents. However, there is no clear description of the methodology followed in determining the roadway capacity reduction.

Lindley (1987) prepared reduction-in-capacity guidelines / thumb rules that can be used to determine how much capacity remains depending on how many lanes are blocked. This table appears in the Traffic Control Systems Handbook and is reproduced as Table 5 below.

Table 5- Percentage of Freeway Capacity Available Under Incident Conditions
(Source: Traffic Control Systems Handbook, 1996, see also Lindley (1987))

Lanes per Direction	Shoulder Disablement	Shoulder Accident	Lanes Blocked		
			One	Two	Three
2	0.95	0.81	0.35	0	N/A
3	0.99	0.83	0.49	0.17	0
4	0.99	0.85	0.58	0.25	0.13
5	0.99	0.87	0.65	0.4	0.2
6	0.99	0.89	0.71	0.5	0.25
7	0.99	0.91	0.75	0.57	0.36
8	0.99	0.93	0.78	0.63	0.41

4.3 Incident Delay

Estimation of incident delay is in many ways the end objective of all NRD research. The models described here have the estimation of delay as the primary objective.

Morales (1986) used queuing analysis (a modified version of Figure 2, presented earlier) to calculate the average incident delay per vehicle. The model is static and focuses on a single link. In one of the case studies, the incident clearance time explicitly takes into consideration the potential that the roadway may be closed by the incident management team upon arrival at the scene. The model is developed for different types of incidents that have different arrival (demand) rates, capacity reductions, durations and discharge flow rates. The model assumes constant flows, which is an approximation to dynamic conditions. This kind of deterministic queuing model has been widely used throughout the US.

Stamadiadis *et al*. (1998) developed models for various routes in Massachusetts. Cohen (1999) developed a sketch model that produces estimates of incident delay and variance per vehicle class. Presley and Wyrosdick (1998) developed an incident delay model that incorporates the average incident duration and the incident severity (e.g. number of lanes closed).

Lindley (1987) studied recurrent and incident induced delay using traffic counts from 37 cities across the United States. Volume-to-capacity (v/c) ratios and travel demands are calculated for the cities. Highway Capacity Manual (1985) procedures are used to determine average speeds and travel times based on calculated v/c ratios. Incident delay is calculated using the deterministic queuing diagram presented in Figure 1. Incident delays are estimated using an expected number of incident types per facility to estimate the incident induced delay per facility.

Sullivan (1997) developed a methodology for determining incident-induced delay also using the queuing theory. The concept is akin to Morales (1986), but used in a way more similar to Lindley. Sullivan generates an empirical model to estimate the expected number of freeway incidents and their associated delays. Further, Sullivan computes incident durations as weighted averages of the clearing times. Fixed time increments are added to represent detection and response times determined from judgment based on data sets where total incident durations are documented. The model uses the 20^{th}, 55^{th}, 80^{th}, and 95^{th} percentile duration corresponding to each accident situation to estimate the overall expected delay for the entire duration distribution. Sullivan's delay model uses the percentage of incident type and the associated incident rate to determine the corresponding capacity reduction for each incident type. Each incident type is then matched to an incident duration to formulate weighted delay averages. The delay estimates at the studied facilities are then extrapolated to further produce delay estimates for a region.

Skabardonis *et al*. (1996) determined the amount of delay caused by incidents in order to evaluate the freeway service patrol. Loop detectors are used to determine the speed of vehicles through the segment and probe vehicles are used to detect the incidents. A formula is used to predict delay, which calculated delay as a function of traffic volume, time of congestion, length of impacted freeway segment, average incident travel speed, and normal travel speed. More details are in Appendix A.

The I-880 database consists of 276 hours of field data that are uniquely linked to provide a complete representation of the freeway operating conditions. A total of 1,616 incidents were observed during the field study. The methodology developed in this study to estimate incident delay is based on the travel time difference, but it uses data from loop detectors that are continually recorded at close spacings.

Garib *et al*. (1997) provided statistical models for estimating incident delay and a model for predicting incident duration, using the data collected from the study of the freeway service patrol (FSP) evaluation project (Skabardonis *et al*., 1995) conducted on Interstate 880 in Alameda County, Oakland, California. The section length was 7. 3 miles, the duration of the study was 3 months and the number of incidents recorded was 205 (accidents (38) and breakdowns (167)). The dependent variable in the model is the "cumulative incident delay" in vehicle-hours. The incident delay models show that up to 85% of the variation in incident delay can be explained by incident duration, number of lanes affected, number of vehicles involved, and traffic demand before the incident. Garib *et al*. develop two models to predict incident induced delay. The first model involves four variables: the number of lanes involved, the number of vehicles involved, the incident duration, and the traffic demand upstream of the incident. The second model uses only three variables because it ignores the traffic demand upstream of the incident. The models are facility specific. Calibration is recommended for their use at other facilities. More details are in Appendix A.

Pierce *et al*. (2005) used video detection to estimate incident delays by taking the difference between the actual travel times and the normal travel times. The normal travel times are estimated based on the conditions before or after the incident or from average conditions for the segment from non-incident days. Their process is time consuming and labor intensive because each individual vehicle must be tracked from upstream to downstream. Such studies are useful to evaluate models that produce estimates of incident delay and they should be performed periodically.

Fu and Rilett (1997) developed a dynamic and stochastic model that produces estimates of individual vehicle delay within a traditional deterministic queuing model approach. The model explicitly considers the incident duration to be a random variable. A mixed discrete and continuous vehicle-delay model is devised and the mean and variance are estimated for individual vehicle delay. Three delay models were crafted:

- *No Delay Regime* - In the no-delay regime the vehicle arrives after the incident has been cleared and the associated queue has been dissipated.
- *Fixed-Delay Regime* - In the fixed-delay regime the vehicle arrives at the incident location at time and joins the queue, where the incident queue has not dissipated, and will not be dissipated until after the vehicle traverses the link. In this situation the queue dissipates at a rate of and consequently the vehicle experiences the maximum delay: .
- *Variable-Delay Regime* - In the variable-delay regime the vehicle arrives at the link and either (a) the incident has been cleared but some portion of the queue remains or (b) the incident has not been cleared but will be cleared before the vehicle exits the link. In this situation, the dissipation rate of the standing queue is some combination of c and c*, and consequently the queue delay will lie between the previous two cases.

Ideally, anticipated and quantitative information such as time dependent delay caused by an incident should be estimated and provided to drivers. Procurement of such information is however not a trivial task because of the complex interactions among various factors such as incident location and severity, incident response capability, demand fluctuation and diverse driver responses to information. Moreover, most of these factors are subject to high uncertainty and information available to quantify them is often incomplete and subjective in nature. Consequently, provision of crisp values of expected delays to drivers would inevitably lower drivers' trust in the accuracy of the provided information because the actual delays they would experience will be either larger or smaller than what were suggested.

To provide a better delay estimation for each individual driver approaching the incident site, Fu and Rilett (1997) developed a dynamic and stochastic model within a traditional deterministic queuing model approach for predicting the delay that a vehicle would experience traveling through an incident location. The model explicitly considers the stochastic attributes of incident duration. They derived a mixed discrete and continuous vehicle-delay model and estimated the mean and variance of vehicle delay.

The use of a facility-specific incident duration distribution produces both the corresponding vehicle delay distribution as well as the individual vehicle delay based on the vehicle arrival time. The model's main deficiency is that it ignores the spatial characteristics of incident delay.

4.4 Simulation-Based Techniques

Simulation offers excellent mechanisms for representing, modeling, and predicting what the NRD might be for specific incidents, and through results integration, for incidents more generally. However, as of today, simulation is not an analysis "tool-of-choice". The project team perceives that this is largely due to three reasons. The first is that simulation is still a very resource-intensive mechanism for examining traffic flow phenomena, including incident response; the second is that simulation models are not yet able to deal with network-wide analyses, major corridors are the upper limit; and the third is that there are not integrated and practical ways to examine and tabulate the results of incident analyses so that comprehensive estimates of NRD can be obtained.

Traffic simulation is a promising tool that transportation agencies are beginning to use extensively to analyze incidents and other traffic phenomena. The models presently available include PARAMICS, VISSIM, and CORSIM. The fundamental relationships are based on traffic flow theory concepts such as car following, lane changing, and gap acceptance. They can be used to study both surface streets and freeways. PARAMICS and VISSIM are today overshadowing CORSIM due to better user interfaces and output graphics. In addition, both PARAMICS and VISSIM are path based while CORSIM is based on interchange/intersection turn percentages to direct traffic.

The microscopic traffic simulation model FRESIM (now embedded into CORSIM) was used for the development of the IDAS methodology. The macroscopic traffic simulator QSIM was used by Cambridge Systematics and Cohen (1998) to develop a comprehensive model to analyze incident impacts that includes a model that estimates incident delay for freeway and surface streets. The FREQ11 macroscopic traffic simulator was used by Stamatiades *et.al* (1998) to develop estimates of network incident delay.

Dynamic Traffic Assignment (DTA) models use a slightly more mesoscopic representation of the traffic flow phenomena so that much larger networks can be examined and so that the interactions between network control and vehicle path choice can be examined. The main DTA models used in the US are: DYNASMART, DYNAMIT, VISTA, INTEGRATION, TRANSIMS. Static and dynamic user equilibrium conditions can be examined. The main challenge in using a DTA model, besides developing the network and the control data, is estimation of the dynamic OD matrix. Currently there are very limited dynamic OD data as most OD surveys are rather aggregated to 15 minutes or more (usually one hour). As an interim measure, dynamic OD matrices are estimated using algorithms that try to match actual traffic counts. The resulting OD matrix may not be close to the actual one many OD matrices may fit the observed traffic counts.

The distinction between microscopic models and DTA models should be made. Microscopic models usually run either a static or dynamic traffic assignment to produce the vehicle paths for each OD pair. In contrast, DTA models identify these paths based on the User Equilibrium travel behavior rule for each OD pair – At equilibrium no traveler can switch to a new non-used path and improve his/her travel time for each assignment time interval of the day.

Advantages of traffic simulators:

- They can produce all necessary traffic flow characteristics (network-wide, OD-based, link-based) for the desired time interval;
- They can be used for planning, operational and traveler information studies;
- Once the first model is deployed it is rather easy to be kept updated;
- Its implementation will improve the skills of the transportation engineers and planners of the corresponding transportation agencies in transportation modeling. They will be able to conduct themselves alternative scenario analyses that can lead to better incident management procedures. Collaborative programs between all emergency agencies involved can be established to train their personnel based on simulated incidents and various alternative actions can be tried through the traffic simulator(s).
- The availability of such models will bring consistency among the studies conducted by all modelers that work on the same transportation network (consultants, researchers, DOT engineers and planners) since they will be using and continuously calibrating the same model. The data that they collect for a specific project will be used to produce a newer version of the traffic simulator.

The drawbacks are:

- High initial cost for the development of the first calibrated model;
- The current versions of these traffic simulators require rather high execution times that rise exponentially with the size of the network. Consequently, they cannot be used for real-time traffic forecasting;
- The personnel must be adequately trained to use them properly, as they may become a dangerous tool to inexperienced users, yielding erroneous results.

4.5 Concluding Remarks

To briefly summarize the models that are reviewed in this section:

- These additional models are interesting. They provide a sense of where the state of the practice might be going in the future.
- The fundamental examinations of incident duration provide a sense of how long incidents are and how much variation can be expected.
- Most of the incident duration models are complex and probably beyond the bounds of what NYSDOT can do and/or might want to do.

- The work by Qi (2002) looks like it should be explored to see if it represents a building block for enhancement of CNAM.
- The work by Kaan *et.al*. (1997, 1999) should also be explored as well as Skabardonis (1996).

Task 4 will examine these issues in greater detail.

5.0 Summary and Conclusions

This task report presents a review of models that have been developed to predict NRD. A basic summary of the findings are that: 1) there are not many NRD models in use in practice,(This statement needs additional support i.e. the survey to be done in task 3) 2) CNAM appears to be as good as or better than those that are in use, and 3) there are additional models that have been developed through research projects but they have not been brought forward into practice. Conclusions drawn from these findings are that building on CNAM as an analysis platform probably is the most sensible way in which to proceed; however, in addition to developing locally valid "look-up" tables it would be useful that give CNAM an ability to more credibly reflect the system-level impacts of incidents. This idea is slightly out-of-scope relative to the original project intent, but it seems like a useful additional capability that ought to be included. One unanswered question is whether it is useful to consider the use of the BPM for purposes of estimating NRD, but the project team believes that the use of the BPM would be too resource intensive to be of practical use on a day to day basis.

Summarizing the findings in a different way:

- The TTI model should not be considered, but the methodology it employs may be the state of the practice for most states and metropolitan planning organizations.
- IDAS does not look like a worthwhile option to consider, but it does appear to be the most comprehensive modeling tool that is in use in practice.
- CNAM has functional capabilities that are comparable enough to the state of the practice, if not the state-of-the-art, that NYSDOT should feel confident that its continued use is appropriate and prudent.
- A number of other models have been developed in research projects, but these models have not been incorporated into practice. These models tend to focus on estimating incident duration, reduced capacity, and incident delay.

Summaries of the various models that have been developed can be found in Tables 6, 7, and 8.
Related to moving forward in the project, specifically in Tasks 4 and 5, the sense of the project team is that:

- The hazard-based model using the log-logistic distribution proposed by Qi (2002) should be considered for implementation since it is found to produce a good fit for the NYSDOT facilities that are under investigation for the NRD study.
- The classification tree methodology (Kaan *et al*. 1997, 1999) should also be explored Its main advantage is that it classifies incidents into various groups thereby reducing the variance due to incidents that may exhibit higher or lower incident durations based on

their specific characteristics (e.g. trucks). Another advantage of this method is that it produces look-up tables that can be directly used as input values into CNAM.

- The project team needs to learn from the Task 3 effort what kinds of outputs (incident durations, speed reductions, delays, etc.) the models should be predicting for NYC conditions
- The project team needs to find a way to incorporate into the model some recognition for system-level effects. Ideally, that would be through a network model, at a minimum the segment-specific model parameters need to recognize that the system-level impacts occur
- There is also a need to factor in shifts in travel patterns that are produced in response to the incident. This might be done through a traffic assignment model
- The project team needs to use the TRANSCOM and IIMS data to gain a sense of how the reduced capacity is affected by the characteristics of the incident.
- The categorical breakdowns suggested by the project team need to be sufficiently rich that they allow accurate estimates of NRD.
- Skabardonis (1996) might be useful for incorporating system-level effects. Steps that might be involved in doing that are:
    - Step 1. Identify the boundaries of the roadway section that it is affected by the incident. Given an incident type estimate the propagation of the queue upstream of the incident location. Identify all detection areas that are affected by the incident at each 15-minute time period. Include a set of roadway detection areas downstream from the incident based on the observed traffic volumes during the incident.
    - Step 2. Develop an Origin-Destination (OD) estimation algorithm for the roadway section that is affected by the incident. This algorithm will be based only on the traffic counts observed at the on-ramps, off-ramps and main line. In the future a more comprehensive network wide OD estimation algorithm could be developed in collaboration with NYMTC as part of their future enhancement for the BPM model.
    - Step 3. Estimate the incident delay distribution (OD-ID) for each OD pair (within this freeway

Combinations of various types of variables including location, incident, and response specific variables should be considered in generating duration models based on the approaches proposed above. The evaluation of these models should be based on the newest incident data set of the last year, and preferably the last few months for the selected test corridors.

Table 6: Incident Duration Models										
	Sing/	Prac/	Stat/	Stoch/	Use/	Database			Test	
Model	Ntwk	Theo	Dyna	Deter	Not	Size	Tframe	Source	(Y/N)	Comment
Regression Based Models										
Goolsby (1971)	S	T	S	S	N	1217 *	17 Mo	CA	M	Early effort to develop a model *Not all data used
Juge, Kennedy, and Wang (1974)	S	T	S	S	N	196	17 Mo	CA	M	
	S	T	S	S	N					Sequential stages
Guliano (1989)	S	T	S	S	N	270	2 Yrs	CA	M	Models by category
Golob *et al*. (1987)	S	T	S	S	N	332	2 Yrs	CA	M	Sequential stages
Khattak *et al*. (1995)	S	T	S	S	N	109	2 Yrs	Chicago	M	Four time-sequential models
Cohen and Nouveliere (1997)										
Fu and Rilett (1997)	S	T	S	S	N	N/A	N/A	N/A	N	Assumed distribution
Sullivan (1997)	S	T	S	S	N	?	few mo.	National	M	Emphasis on Percentiles
Garib *et al*. (1997)	S	T	S	S	N	205	1 Yr	CA	Y	Simple, possibly useful, small sample
Ozbay and Kachroo (1997)	S	T	S	S	N	650	2 Yrs	VA	Y	Incorporates decision trees
Smith *et al*. (2000, 2001)	S	T	S	S	N	1707	3 Yrs	I-95	N	classification / regression trees
Hazard Based Models										
Jones *et al*. (1991)	S	T	S	S	N	5637	2 Yrs	WA	M	Linked to a delay model
Nam and Mannering (2000)	S	T	S	S	N	610	2 Yrs	WA	N	Stability difficulties
Qi (2002)	S	T	S	S	N	858	1 Yr	NY	M	Includes time of day
CNAM	S	P	S	D	U	N/A	N/A	NY	M	Average per region; Look-up table

Table 7: Capacity Reduction Models										
	Sing/	Prac/	Stat/	Stoch/	Use/	Database			Test	
Model	Ntwk	Theo	Dyna	Deter	Not	Size	Tframe	Source	(Y/N)	Comment
Goolsby and Smith (1971)	S	T	S	S	N	196		TX	M	Video data, Gulf Freeway
Urbanek and Rogers (1978)	S	T	S	S	N					
Mcshane *et al*. (1990)	S	T	S	S	N					Emphasis on v/c ratios
Smith *et al*. (2003)	S	T	S	S	N			VA	M	

Table 8: Incident Delay Models										
	Sing/	Prac/	Stat/	Stoch/	Use/	Database			Test	
Model	Ntwk	Theo	Dyna	Deter	Not	Size	Tframe	Source	(Y/N)	Comment
All Vehicles, Single Facility										
Messer *et al. (* 1973) ???										
Chow (1974) ????										
Wirasinghe (1978) ???										
Morales (1986)	S	T	S	Deter	N	N/A	N/A	N/A	M	Early effort to develop a model
Garib *et al*. (1987)	S	T	S	Deter	N	205	3 mo	CA?	M	Reghression models; Sequential stages
Lindley (1987)	S	T	S	Deter	N	N/A	1983-84	DC	M	Sequential stages; 1985 HCM based; incident rates used
Skabardonis *et al*. (1996)	Ntwk	T	S	Deter	N	1616	276 hrs	CA	M	Total delay based on multiple sequential detection sites
Sullivan (1997)	S	T	S	S	Use/	?	few mo.	CA	M	Based on the incident duration model
Stamadiadis *et.al.* (1998)	Ntwk	Prac/	S	deter	Use/	N/A	1995	MA	M	Use of simulation-FREQ11; types of incidents modelled
Pressley and Wyrosdick (1998) ???										
Cambridge Systematics and Cohen (1998)	Ntwk	Prac/	S	Deter	?			37 cities	M	QSIM based Recurring and NRD non-linear regression models; FHWA study
NYSDOT (2001) - CNAM	S	Prac/	S	Deter	Use/			NY	N	Sequential steps;
Pierce *et al*. (2005)	S	T	S	S	N				Y	Simple, possibly useful, small sample
One Vehicle										
Fu and Rilett (1997)	S	T	S	S	N	5637	2 Yrs	WA	M	Linked to an incident duration distribution; Individual veh. delay model
Nam and Mannering (2000)	S	T	S	S	N		2 Yrs	WA	N	Stability difficulties
Qi (2002)	S	T	S	Deter	N	N/A	N/A	NY	M	Includes time of day
Cambridge Systematics - IDAS	S	T	S	Deter	Y				M	IDAS; developed using FRESIM
Fu and Hellinga (2004)	S	T	S	S	N	N/A	N/A	N/A	M	Fuzzy logic model
All Vehicles, Network										
Wickes and Lieberman (1980)	S	T	S	Deter	N				M	FRESIM Microscopic Simulation
Mahmassani (?) - DYNASMART	S	T	Dyna	Deter	N				M	DTA; Pilot implementations; very promising in the future
May (?) - FREQ	S	T	S	S	N				M	Sequential stages

REFERENCES

Abbasi, M. F., V.P. Sisiopiku, and W.C. Taylor, "Development and Testing of a Freeway Recurring Congestion Measure", *Proceedings of the 78th Annual Meeting of the Transportation Research Board*, Washington D.C., Jan 1999.

Al-Deek, H.M., A.M. Garib, and A.E. Radwan, "Methods for Estimating Congestion Part-II: New Method," *ASCE Transportation Engineering Journal*, August 1994.

Al-Deek, H.M., A.M. Garib, and A.E. Radwan, "New Method for Estimating Freeway Incident Congestion," *Transportation Research Record 1474*, pp 30-39, 1995.

Banks, J.H., "Performance Measurement for Traffic Management Systems," *Transportation Research Record 1634*, pp 39-45, 1998.

Boarnet, M.G., E J. Kim, and E. Parkany, "Measuring Traffic Congestion," *Transportation Research Record 1634*, 1998.

Boyce, D., N. Rouphail, and A. Kirson, "Estimation and Measurement of Link Travel Times in the ADVANCE Project," *Proceedings of the IEEE-IEE Vehicle Naviation and Information Systems Conference*, Ottawa, Ontario, 1995.

Breiman, L., J. Friedman, and C. Stone, *Classification and Regression Trees*, Wadsworth & Brooks/Cole Advanced Books & Software, 1984.

Cambridge Systematics, H. Cohen, and Science Applications International Corp., *Sketch Methods for Estimating Incident-Related Impacts*, Federal Highway Administration, Washington, DC, 1998.

Chang, G. L. and J.Y. Point-Du-Jour, *Performance and Benefit Evaluation for CHART, Incident Management Program in 1997*, Final Report to Maryland State Highway Administration, September 2000.

Chang, G.L.,A. P. Petrov, N. Lin, N. Zou, and J.Y. Point-Du-Jour, *Performance Evaluation of CHART -The Real-Time Incident Management System-year 2000*, Final Report to Maryland State Highway Administration (MSHA), March 2002.

Chow, W.M. "Study of Traffic Performance Models Under an Incident Condition." *Highway Research Record 567*, TRB, National Research Council, Washington, D.C., pp. 31–36, 1976.

Chu, L., H. X. Liu, and W. Recker, "Evaluation of Potential ITS Strategies under Non-recurrent Congestion Using Microscopic Simulation," *Proceedings of the 83rd Annual Meeting of the Transportation Research Board*, Washington D.C., January 2004.

Chu, L., H.X. Liu, W. Recker, and S. Hague, *Evaluation of Potential ITS Strategies Under Non-recurrent Congestion Using Microscopic Simulation*, PATH Working Paper UCB-ITS-PWP-2003-2, January 2003.

Cohen S. and C. Nouveliere, “Modeling incident duration on an urban expressway,” M. Papageorgiu and A. Pouliezos (eds) *Proceeding of the 8th International Federation of Automatic Control (IFAC) Symposium on Transportation Systems*, 16-18 June, Chania, Greece, 1997.

Cottrell, W., “Estimating the Probability of Freeway Congestion Recurrence”, *Proceedings of the 77th Annual Meeting of the Transportation Research Board*, Washington D.C., Jan 1998.

Cottrell, W.D., “An Empirical Freeway Queuing Duration Model,” *Proceedings of the 78th Annual Meeting of the Transportation Research Board*, Washington D.C., Jan 1999.

DeRose Jr., F. “An analysis of random freeway traffic accidents and vehicle disabilities,” *Highway Research Record 59*, 53-65, 1964.

Derr, B.R., “Delay Analysis for Freeway Corridor Surveillance, Communication, and Control Systems,” *Transportation Research Record 1132*, pp 77-81, 1987.

Dey, S.S., “‘Hours of Congestion as a Transportation Measure of Effectiveness Under Capacity Constrained Conditions”, *ITE Journal*, Dec 1998.

Dowling, R., A. Skabardonis, M. Carroll, and Z. Wang, “Methodology for Measuring Recurrent and Non-Recurrent Traffic Congestion”, *Proceedings of the 83rd Annual Meeting of the Transportation Research Board*, Washington D.C., Jan 2004.

Federal Highway Administration, *FRESIM User Guide*, Version 4.5, Office of Research and Development," U.S. Dept. of Transportation, Washington, DC, 1984.

Forbes, G.J., “Identifying Incident Congestion,” *ITE Journal*, pp 17-22, Jun 1992.

Fu, L. “A Fuzzy Queuing Model for Real-Time, Adaptive Prediction of Incident Delay for ATMS/ATIS,” *Transportation Planning and Technology*, Vol. 27, No. 1, pp. 1-23, 2004.

Fu, L. and L.R. Rilett, “Real-Time Estimation of Incident Delay in Dynamic and Stochastic Networks,” *Transportation Research Record 1603*, pp. 99-105, 1997.

Fu, L.and B. Hellinga “A Fuzzy Queuing Model for Real-Time, Adaptive Prediction of Incident Delay for ATMS/ATIS,” *Transportation Planning and Technology*, Vol. 27, No. 1, pp. 1-23, 2004.

Garib, A., A.E. Radwan, and H. Al-Deek, “Estimating Magnitude and Duration of Incident Delays,” *Journal of Transportation Engineering*, v 139 n 6 November/December pp 459-465, 1997.

Giuliano, G., “Incident Characteristics, Frequency, and Duration on a High Volume Urban Freeway,” *Transportation Research A*, 23A, No. 5, 1989.

Golob, T.F., W.Recker, and J.D. Leonard, “An Analysis of the Severity and Incident Duration of Truck-Involved Freeway Accidents,” *Accident Analysis & Prevention*, Vol. 19, 1987.

Goolsby M.E., “Influence of incidents on freeway quality of service,” *Highway Research Record 349*, TRB, Washington D.C., pp. 41-46, 1971.

Greene, W.H., *Econometric Analysis*. Prentice Hall, 2000.

Hall, R., *Incident Dispatching, Clearance and Delay*, California PATH Working Paper UCB-ITS-PWP-2000-14, California Path program, ITS, University of California, Berkeley, CA, 2000.

Jones, B., L. Janssen and F. Mannering (1991), “Analysis of the Frequency and Duration of Freeway Accidents in Seattle,” *Accident Analysis and Prevention*. Vol. 23, No. 4, pp. 239-255, 1991.

Jovanis, P., and H.L. Chang, “Disaggregate model of highway accident occurrence using survival theory,” *Accident Analysis and Prevention*, Vol. 21, pp.445-458, 1989.

Judycki, D.C., and J.R. Robinson, “Managing Traffic during Nonrecurring Congestion,” *ITE Journal*, pp 21-26, Mar 1992.

Juge J., K. Kennedy, and T. Wang, *Early Detection and Road Removal of Disabled Vehicles and Other Hazards from the Freeway*, Final Report, California Department of Transportation and Department of California Highway Patrol, 1974.

Khattak, A., J. Schofer, J., and M-H Wang, “A Simple time sequential procedure for predicting freeway incident duration,” *IVHS Journal*, Vol. 2, No.2, pp. 113-138, 1995.

Kramer, R.F., “The A to Z Techniques to Reduce Congestion and Increase Capacity of Streets and Intersections,” *ITE Journal*, pp 49- 51, Mar 1988.

Lindley, J.A., “Urban Freeway Congestion: Quantification of the Problem and Effectiveness of Potential Solutions,” *ITE Journal*, v 57 n 1 pp 27-32, Jan 1987.

Lindley, J.A., “A Methodology for Quantifying Urban Freeway Congestion”, Transportation Research Record 1132, 1987.

Lindley, J.A., “Urban Freeway Congestion Problems and Solutions: An Update,” *ITE Journal*, pp 21-23, Dec 1989.

Lomax , T., S. Turner, M. Hallenbeck, and R. Margiotta, *Traffic Congestion and Travel Reliability: How Bad is the Situation and What is Being Done About It*, found at http://www.ite.org/nationalsummit/vision/congestion.doc, Sep 2001.

Masinick, P., and H. Teng. *An Analysis on the Impact of Rubbernecking on Urban Freeway Traffic*. Research Report No. UVACTS-15-0-62, 2004.

McShane, W.R., and R.P. Roess, *Traffic Engineering*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1990.

Messer, C.J., C.L. Dudek, and J.D. Friebele, "Method for Predicting Travel Time and Other Operational Measures in Real-Time During Freeway Incident Conditions," *Highway Research Record 461*, TRB, Washington, D.C., pp. 1–11, 1973.

Morales, M.J., "Analytical Procedures for Estimating Freeway Traffic Congestion," *ITE Journal*, Vol. 57, No. 1 pp 45-49, Jan 1987.

Morales, M.J., "Analytical Procedures for Estimating Freeway Traffic Congestion," *Public Roads*, Vol. 50, No. 2, 1986.

Mouskos, K.C., E. Niver, L.J. Pignataro, S. Lee, L. Papadopoulos, and N. Antoniou, *Evaluation of the TRANSMIT System*, Final Report to TRANSCOM, Institute for Transportation, NJIT, 1998.

Nam, D. and F. Mannering, "An Exploratory Hazard-Based Analysis of Highway Incident Duration," *Transportation Research A*, Vol. 34A, No. 2, pp 85-102, 2000.

Ozbay K., A.G. Hobeika, and Y. Zhang, "Estimation of Duration of Incidents in Northern Virginia" Proceedings of the 77th Annual Meeting of the Transportation Research Board, Washington D.C., Preprint# 971293, January 1997.

Ozbay, K. and Pushkin K., *Incident Management in Intelligent Transportation Systems*, Boston, MA, Artech House. 1999.

Pal, R., S.P. Latoski, and K.C. Sinha, "Investigation of Freeway Incident Characteristics and Their Influence in Planning an Incident Management Program," *Transportation Research Record 1634*, pp 46-55, 1998.

Parham, A.H., K. Fitzpatrick, M.D. Wooldridge, and D. Fenno, "Strategies to Facilitate Incident Management on New or Reconstructed Freeways," *Proceedings of the 80th Annual Meeting of the Transportation Research Board*, CD-ROM, Washington D.C., Jan 2001.

Pierce, R.A., C.C. Sun, and J.D. Lemp, "Incident Delay Modeling with the Use of Video Reidentification," *Proceedings of the 84th Annual Meeting of the Transportation Research Board, CD-ROM*, Washington, DC, Jan 2005.

Qi, Y., *Incident Delay Analysis and Estimation, PhD Dissertation*, Polytechnic University, Brooklyn, NY, 2002.

Ritchie, S.G., and R.L. Cheu, *Neural Network Models for Automated Detection of Non-Recurring Congestion*, PATH Research Report UCB-ITS-PRR-93-5, Jun 1993.

Schrank, D., and T. Lomax, *The 2003 Annual Urban Mobility Report*, Texas Transportation Institute," The Texas A&M University, http://mobility.tamu.edu, 2003.

Shaw, T., and D. Jackson, "Reliability Performance Measures for Highway Systems and Segments," *Proceedings of the 82nd Annual Meeting of Transportation Research Board*, Washington D.C., Jan 2003.

Skabardonis, A., P. Varaiya, and K.F. Petty, "Measuring Recurrent and Non-Recurrent Traffic Congestion," *Proceedings of the 82nd Annual Meeting of Transportation Research Board*, Washington D.C., Jan 2003.

Skabardonis, A., K. Petty, H. Noeimi, D. Rydzewski, and P. Varaiya, "I-880 Field Experiment: Data-Base Development and Incident Delay Estimation Procedures," *Transportation Research Record 1554*, pp 204-212, 1996.

Smith, B.L. and Smith K.W., *An investigation into incident duration forecasting for Fleetforward*, Final report of the ITS center project: Fleetforward traffic forecasting, UVA Center for Transportation Studies, UVA-CE-ITS_01-5, 2000.

Smith, B.L., L. Qin, and R. Venkatanarayana, "Characterization of Freeway Capacity Reduction Resulting from Traffic Accidents," *Journal of Transportation Engineering*, pp 362-368, Jul/Aug 2003.

Smith, B.L., L. Qin, and R. Venkatanarayana, "Characterization of Freeway Capacity Reduction Resulting from Traffic Accidents." *Journal of Transportation Engineering*, Vol. 129, No 4, July/August pp. 362-368, 2003.

Smith, K., and B. Smith, *Incident Duration Forecasting*, Final Report, ITS Implementation Center, George Mason University, Virginia, http://www.gmupolicy.net/its/incidentduration/

Smith, K.W., and B.L. Smith, *Forecasting the Clearance Time of Freeway Accidents*, Smart Travel Lab Report No. STL-2001-01, Center for Transportation Studies, University of Virginia, 2001.

Stamatiadis, C., N.H. Gartner, J. Winn, and R. Bond, "Evaluation of the Massachusetts Motorist Assistance Program: Assessment of Congestion and Air Quality Impacts," *Transportation Research Record 1757*, pp 1-9, 1998.

Sullivan, E., "New Model for Predicting Freeway Incidents and Incident Delays," *Journal of Transportation Engineering*, Vol. 123, No. 4, pp 267-275, July/Aug 1997.

Sullivan, E., Taff, and Daly, *A Methodology for Measurement and Reporting of Incidents and Prediction of Incident Impacts on Freeways*, prepared for FHWA by Ball Systems Engineering Division, 1995.

Taylor, W.C., M.A. Sarif, and V.P. Sisiopiku, "Development of Freeway Incident-Based Congestion Measures," *Proceedings of the 78th Annual Meeting of the Transportation Research Board*, Washington D.C., Jan 1999.

Thill, J.F., G.L. Rogova and J. Yan, "Delay Modeling for Corridor-Level Benefits Evaluation of VMS Deployments," *Proceedings of the 80th Annual Meeting of the Transportation Research Board*, CD-ROM, Jan 2001.

Urbanek, G.L. and R.W. Rogers, *Alternative surveillance concepts and methods for freeway incident management, Executive Summary*, FHWA, Report # FHWA-RD-77-58, Vol. 1, Washington, DC, 1978.

Wickes, D., and E. Lieberman, *Development and testing of INTRAS, a microscopic freeway simulation model, Vol. 1, Program Design, Parameter Calibration and Freeway Dynamic Component Development*, Report No. FHWA/RD-80/106, FHWA.,Washington DC, 1980.

Wirasinghe, C.S., "Determination of Traffic Delays from Shock Wave Analysis," *Transportation Research*, Vol. 12, 1978.

Yang, H., K.K. Lo, and W.H. Tang, "Travel Time Versus Capacity Reliability of A Road Network," *Proceedings of the 79th Annual Meeting of the Transportation Research Board*, Washington D.C., Jan 2000.

APPENDIX A
Model Descriptions

This appendix contains more technical detail about several of the models reviewed in the main body of the report. Models are included if they are particularly relevant and the presentation of the detail is helpful in understanding nuances.

INCIDENT DURATION MODELS

Khattak et al. (1991)
The regression model developed by Khattak *et al*. (1991) is:

Clearance Time = 14.03 + 35.57(HEAVY) + 16.47(WX) + 18.84(SAND)
– 2.31(HAR) + 0.69(RESP) + 27.97(OTHER) + 35.81(RDSIDE) + 18.44(NTRUCK)
+ 32.76(NONCON) + 22.90(SEVINJ) + 8.34(WRECKER) (A-1)

Where:

- HEAVY: presence of heavy loadings
- WX: existence of extreme weather conditions
- SAND: sand/salt pavement operations
- HAR: was incident notification promulgated (highway advisory radio)
- RESP: incident response time
- OTHER: assistance from other response agencies
- RDSIDE: freeway facility damage caused by incident
- NTRUCK: number of trucks involved
- NONCON: liquid or uncovered broken loadings in heavy vehicles
- SEVINJ: occurrence of severe injuries
- WRECKER: the use of a wrecker for incident clearance

Khattak et al. (1995)
Khattak et al. (1995) developed a truncated regression model for estimating / predicting incident duration, because small values of incident duration are not observed. The truncation point is arbitrarily chosen to be 10 minutes based on authors' examination of the data. The model is as follows:

$$Y = X\beta + \varepsilon \qquad \text{(A-2)}$$

where:

- Y = Vector of n dependent variable observations on incident duration,
- X = Matrix of k independent variables and n observations,
- β = Vector of k parameters,
- ε = The error term with expected value zero and variance σ^2.

Therefore Y is distributed normally with mean $X\beta$ and variance σ^2, i.e. $Y \sim N(X\beta, \sigma^2)$.

Instead of using a unique model, a four time-sequential prediction models is used to support earlier but probably less-accurate duration predictions with fewer variables. The model is updated in a series of steps as new information arrived. During subsequent stages, more information is acquired, and consequently better predictions become possible. Also, certain previously acquired information becomes irrelevant. A 5-minute time-interval is selected between each two sequential models. The minimum acceptable level of variable accumulated probability is 70% for each model. As time progresses the successive estimations of these models improves according to the values of standard error and R^2 for the four models.

The models show that incident durations are longer when:

- the response times are higher,
- the incident information is not disseminated through the public media,
- there are severe injuries,
- trucks are involved in the incident,
- the trucks have heavy loads,
- State property is damaged, and
- the weather is bad.

The most important variables in the incident duration prediction are incident characteristics and the consequent emergency response actions. This sequential approach offers a framework for building other, more locally relevant models that reflect the increasing availability of detailed incident information as the events proceed.

Sullivan (1997)
Sullivan (1997) generates an empirical model to estimate the number of freeway incidents and their associated delays. The incident durations are found to closely follow the lognormal distribution. Sullivan computes incident durations as weighted averages of the clearing times reported in many study data sets. A fixed time is added in increments representing detection and response times determined from judgment based on data sets where total incident durations are documented. The incident duration distribution is determined by the incident type, existing incident management and the incident location. The standard deviations are determined by observing the field data and data from previously published sources – they are based on the observation that there is a strong tendency for the standard deviations of incident duration distributions to be proportional to their means. That is, using the expected mean and standard deviation of each incident's duration, it becomes possible to characterize the entire duration distribution. Sullivan's model uses the 20^{th}, 55^{th}, 80^{th}, and 95^{th} percentile duration corresponding to each accident situation to estimate the overall duration. The 20th percentile values range from 4 to 20 minutes and the 95^{th} percentile values range from 60 to 130 minutes.

Garib et al. (1997)
Garib *et al*. (1997) develop a lognormal-based linear regression model based on six variables:

$$Log(IncDuration) = 0.87 + 0.027x_1x_2 + 0.2x_5 - 0.17x_6 + 0.68x_7 - 0.24x_8 \quad \text{(A-3)}$$

where:

- x_1 is the lanes affected,
- x_2 is the number of vehicles involved,
- x_5 is truck involvement – binary,
- x_6 is time of day – binary,
- x_7 is police response time (natural log of this value), and
- x_8 is weather conditions – binary.

The adjusted R^2 value of this regression model is 0.81. The incident duration prediction model showed that 81% of variation in incident duration could be predicted by number of lanes affected, number of vehicles involved, truck involvement, time of day, police response time, and weather condition. No other variables tested either individually or jointly were found to be significant. The model does not include variables for property damage, driver age, injury accident, and number of injuries. The model indicates that: 1) the police response time is a highly significant factor in predicting the incident duration followed by 2) weather condition, 3) time of day, 4) truck involvement, and finally 5) the joint effect of the number of vehicles involved and the number of lanes affected. This model supports the premise that the incident duration follows the lognormal distribution.

Smith et al. (2000)

Smith et al. (2000) develop a nonparametric regression model. Data from SmartRoute's SmarTraveler system and the I-95 Corridor Coalition's Information Exchange Network (IEN) are employed. Nonparametric regression is a forecasting technique that requires no strict assumptions regarding a functional relationship between dependent and independent variables.

Unlike traditional regression models that define a relationship for all ranges of dependent variables, nonparametric regression focuses on a specific area, or neighborhood, of past system states that are similar to the current system state. The past instances in this neighborhood are then combined (usually a weighted average) to predict the dependent variable value. This method relies heavily on having a wide range of quality data to make predictions. The key to an effective nonparametric model is effectively defining a neighborhood of past instances.

Smith et al. (2000) find that the predicted incident duration differs from the actual incident duration by an average of over 100%. They claim that the modeling approach itself is not likely to be the cause, but rather the independent variables that are employed. They comment that from a statistical standpoint, the independent variables may not be significant. The independent variables used are:

- type of highway incident (accident, construction, debris, etc.),
- time of day,
- day of week,
- incident duration,
- location of the incident (state), and
- number of lanes closed during the incident.

The location variable is supposed to capture the incident management capability of the local authorities. A more representative variable might be the specific assistance given to the incident. It is possible that the states along the I-95 Corridor under the study have similar response plans to highway incidents and that the same procedures are used.

The lane closure variable is supposed to capture the effects of incident severity. Preferable variables might be the number of vehicles involved, the presence of personal injuries, the presence of trucks, and damage to roadway.

Smith et al. (2001)
Smith et al. (2001) also investigate various models for forecasting the clearance time of freeway accidents. Nonparametric regression is employed, along with Weibull and lognormal distributions and a classification tree. The dependent variable is the incident duration. The independent variables are divided into three categories:

- *physical variables*: accident time of day, the day of the week and the weather;
- *vehicle variables*: include number of vehicles, truck involvement, and passenger bus involvement; and
- *response variables*: binary variables of emergency agencies responding to the scene.

The classification tree is based on clearance time: short (1-15 min.), medium (16-30 min), and long (31 or greater – they are considered to be detrimental to traffic operations). These definitions are chosen based on practical experience.

The independent variables are listed in Table A-1 and they are the same as the ones used for the non-parametric regression model.

Table A-1. Independent Variables used in the Classification Tree Model
(Source: Smith *et al*. 2001)

Variable	Name	Weight W(*x*)	Value
PEAK	A	3.43	1 = Peak (6-8am, 4-6pm)
			2 = Off-peak
WEEKDAY	B	3.9	1 = Weekday
			2 = Weekend
EMS	C	16.07	1 = Yes
			2 = No
FIRE	D	15.28	1 = Yes
			2 = No
HAZMAT	E	97.27	1 = Yes
			2 = No
POLICE	F	9.17	1 = Yes
			2 = No
VDOT	G	24.78	1 = Yes
			2 = No
TOW	H	20.83	1 = Yes
			2 = No
NUMVEH	I	WI12 = 7.44	1 = Single vehicle
		WI13 = 6.39	2 = Two vehicles
		WI23 = 13.83	3 = Three or more vehicles
TRUCK	J	16.1	1 = Yes
			2 = No
BUS	K	11.01	1 = Yes
			2 = No

The main MOE is the prediction accuracy. That is the percentage of test accidents in which the clearance time is predicted correctly. Other MOEs are:

- The percentage of long clearance times that are predicted correctly and
- The accuracy of the predicted durations for short clearance time accidents.

The methodology is implemented using the software Classification and Regression Tree (CART) (see Breiman *et al*., 1984). For each level of the tree where a decision node is present, CART considers each independent variable as the splitting criteria. The one split that provides the best results is selected for that decision node. This process continues until the largest possible tree has been created. Each new split creates a new classification tree that is a candidate for the optimal tree. Next, CART uses a pruning technique to determine the optimal tree. Starting from the largest tree, the testing sample is run through the classification tree to find the prediction accuracy. This continues for each smaller tree until the one with the best prediction accuracy is identified. This tree growing and pruning technique assure that the best possible splits and sizes are found.

The classification tree for incident duration uses the variables listed in Table A-1, sequentially. At each tree level it finds which variables should be included in the model or not (based on the binary values of the fourth column). For example if a tow truck is involved in the incident clearance, that variable is assigned a value of 1.

In the end, a path is identified from the top of the tree to the bottom. The tree structure does not follow the chronological steps of the actual incident clearance but it rather starts from the parameters that are expected to incur most of the delay, based on past experience.

The path followed determines where the incident falls in one the three categories (short, medium, long). The classification tree predicted correctly 76.73%. 19.14% and 64.48% for the short, medium and long clearance times, respectively. This result indicates that the model performs well for short-duration events, poorly for the medium-duration events, and OK for the long-duration events. The fact that none of the models perform particularly well raises the question as to whether the data were of good quality – the authors report that for a rather serious accident where police, EMS, fire department personnel were involved in a three-car accident, the reported incident clearance time was only 9 minutes.

The study tries to evaluate a stochastic model using probability density functions to describe clearance time. The Weibull and lognormal distributions are rejected based on the available clearance time data. None of the forecasting models produces results that are accurate enough to warrant implementation in an operational incident management system. The poor results are attributed to the choice of forecasting models and/or the quality of the accident data.

Hazard Based Incident Duration Models

Hazard-based models are mostly used in the biometrics and industrial engineering fields. They use conditional probabilities to find the likelihood that an incident will end in the next short time period given its continuing duration.

Jones et al. (1991)

Jones *et al*. (1991) assume that the number of traffic accidents in a given day follows a Poisson distribution. The accident frequency model has four classes of independent factors: seasonal effects, weekly trends, special events, and environmental. Due to the susceptive reliability of the volume data obtained from magnetic loop detectors, the traffic volume is not directly used as an independent variable. Instead, most of the independent variables included in the model are proxies for traffic volume, and they are used to implicitly capture traffic variability.

The incident duration distribution $F(t) = \Pr(T < t)$ specifies the probability that the duration will be less than t, where T is the incident duration. Correspondingly, a survivor function $S(t) = \Pr(T \geq t) = 1 - F(t)$ specifies the probability that the accident will have a duration greater than or equal to t.

Jones *et al*. (1991) analyze the accident duration using a hazard function, $h(t) = f(t)/S(t)$, where $f(t) = dF(t)/dt$. This function represents the ratio of the rate at which $F(t)$ is changing relative to the value of $S(t)$. In a simple case, $F(t)$ will be small initially and grow while $S(t)$ will be large initially and decrease. Since $h(t)$ can be positive, negative, or zero, if $h(t)$ is positive, the incident duration is likely to be at or near the current value of t, if it is negative, the duration is likely to be considerably longer, and if it is zero, the duration is independent of the present time t being considered.

The duration model is based on the accelerated lifetime model $h(t,\beta,X) = h_0[ty(\beta,X)]y(\beta,X)$, where X is a vector of explanatory variables, β is a vector of the parameters to be estimated, $h_0[t]$ is the baseline hazard, and $y(\beta,X)$ is a scaling factor defined as $y(\beta,X) = \exp(\beta X)$. The accelerated lifetime model uses explanatory variables to rescale time directly. Although, the accident duration distribution is found to be approximately normal, Jones *et al*. indicate that their duration data do not perfectly fit either the normal or the lognormal distributions, but rather the log-logistic distribution.

The classes of variables used by Jones *et al*. in their incident duration model are different from those used for their frequency model. The new classes include driver characteristics and accident severity measures. Different duration models are estimated for different geographic zones, since the independent variables are observed to have different levels of significance in different zones. The hazard function of every estimated duration model decreases with increasing t, so Jones *et al.* suggest that the longer an accident lasts, the less likely it is that it will be cleared soon.

Qi (2002)
A likelihood is found that an incident will end in a particular number of minutes given its current duration. A Kolmogorov-Smirnov test demonstrates that the log-logistics distribution can be used to describe the duration. In addition Qi tests the Weibull and lognormal distributions, however they fail the Kolmogorov-Smirnov test for the distribution of generalized residuals.

The hazard function of the log-logistic distribution h(t) is: $h(t) = \lambda p(\lambda t)^{p-1} / [1 + (\lambda t)^p]$ and the corresponding survival function $S(t)$ is: $1/[1 + (\lambda t)^p]$. Qi uses an accelerated lifetime model where $\lambda = \exp(-\beta X)$. The parameter λ is the mean of incident duration. The corresponding incident duration is estimated using the median as portrayed by Greene (2000):

$$\hat{t} = Median(t) = \exp(\beta X) \tag{A-4}$$

"The median is used instead of the mean since the log-logistic incident duration distribution is skewed and has a long tail," Qi (2002). Qi (2002) considers the following variables to be included in the log-logistic incident duration model:

- Temporal characteristics (weekday, am peak, pm peak, night);
- Weather characteristics (snow, rain);
- Incident location (On-ramp, at the entrance of Off-ramps, at the exit of On-ramps); This information was not available within the incident database so this variable was not included in the model.
- Incident characteristics (incident type, number of vehicles involved, number of lanes blocked);
- Involved vehicle characteristics (types of vehicles involved, occupants or vehicle load); The occupants or vehicle load were not available so they were not considered in the development of the model.
- Incident clearance source (response time of first rescue vehicle, agency involved, tool used).

The incident type is also estimated using the corresponding probability $P_n(i)$ that an incident n will be of type i. This is then used to calibrate the incident duration model based on the incident type.

Qi (2002) also develops an on-line incident duration prediction model using hazard-based duration regression models. As in the off-line incident duration model, the log-logistic distribution is used. Three separate incident duration models are developed. They pertain to: 1) incident report time to incident verification, 2) incident verification to incident clearance, and 3) incident clearance to the end of the incident (e.g., queue dissipation).

The incident duration prediction model estimates of the remaining incident clearance time given that the incident has lasted for some time period. The truncated median incident duration distribution is used as in the off-line model. The three models use the following characteristics to describe the incident: weather (snow, rain), time period (am peak, pm peak, night), incident type (property, injury/fatality, disabled vehicle), and incident clearance agency (police, NYCDOT, tow truck). Only some of these variables are used in each model.

DELAY MODELS

Skabardonis et al. (1996)

Skabardonis et al. (1996) determine the amount of delay caused by incidents in order to evaluate the freeway service patrol. A formula is used to predict delay, which calculated delay as a function of traffic volume, time of congestion, length of impacted freeway segment, average incident travel speed, and normal travel speed.

The I-880 database consists of 276 hours of field data that are uniquely linked to provide a complete representation of the freeway operating conditions. A total of 1,616 incidents were observed during the field study. The methodology developed in this study to estimate incident delay is based on the travel time difference, but it uses data from loop detectors that are continually recorded at close spacings. The freeway section upstream of the incident is divided into k segments of approximately equal length (L_k). The speeds and volumes on each segment are assumed to be constant and equal to the values provided by the loops within the segment. The average incident-free speed is based on the loop detector data throughout the study period. The delay is then calculated for each time slice i and each segment k upstream of the incident as follows:

$$D_{ki} = L_k\,(t/60)\,Q_{ki}\,(1/V_{ki} - 1/V_{kif}) \qquad \text{for } 0 < V_{ki} < V_{kif} \qquad \text{(A-5)}$$

Where:

- t is the length of the time slice, typically 1 to 5 min
- V_f = average travel speed under prevailing incident-free conditions (Km/hr)
- Q = traffic volume (vph)
- D = Incident delay (vehicle-hours)

The total incident delay is then:

$$D = \sum_{k=1}^{n} \sum_{i=1}^{m} D_{ki} \qquad \text{(A-6)}$$

Where n is the number of the freeway segments upstream affected by the incident (i.e., the end of the queue because of the incident) and m is the number of congested time slices (i.e., the incident duration plus the time it takes for the queue to clear).

This methodology tries to capture the impact of an incident on an expanded roadway section upstream of the incident location. The total delay estimate, D, is more comprehensive that the traditional deterministic queuing model. A similar approach could be employed in CNAM since it does not require additional data. What is missing, is the corresponding average vehicle speed downstream of the incident location, since it is expected to be different (potentially higher but under non-congested conditions) than the V_f . By considering a more expanded roadway section that includes the incident affected section upstream and downstream of the incident location a more comprehensive approach for incident delay could be developed. The analysis could further consider the establishment of Origin Destination (OD) pair estimates of incident delay for the expanded roadway section, based on the traffic volumes recorded at the on-ramps, off-ramps and mainline.

Garib et al. (1997)
Garib et al. (1997) provided statistical models for estimating incident delay and a model for predicting incident duration, using the data collected from the study of the freeway service patrol (FSP) evaluation project (Skabardonis et al., 1995) conducted on Interstate 880 in Alameda County, Oakland, California. The first model involves four variables: the number of lanes involved, the number of vehicles involved, the incident duration, and the traffic demand upstream of the incident. The second model uses only three variables because it ignores the traffic demand upstream of the incident. The models are facility specific.

Model 1 depicts incident delay as a function of incident duration, traffic demand, and capacity reduction represented by number of lanes affected and number of vehicles involved.

Model 1: $$Delay = -4.26 + 9.71 x_1 x_2 + 0.5 x_1 x_3 + 0.003 x_2 x_4 + 0.0006\,(x_2)^3 \qquad \text{(A-7)}$$

Model 2 predicts the cumulative incident delay as a function of incident duration, number of lanes affected, and number of vehicles involved.

Model 2: $$Delay = -0.288 + 3.8 x_1 x_2 + 0.51 x_1 x_3 + 0.06 x_3 + 0.356\,(x_2)^3 \qquad \text{(A-8)}$$

For both models:

- *Delay* = cumulative incident delay (vehicle-hours)
- x_1 = number of lanes affected by the incident
- x_2 = number of vehicles involved in the incident

- x_3 = incident duration (the difference between the corrected incident start time and its corrected end time in minutes
- x_4 = traffic demand upstream of the incident in the last 15 minutes before the incident starting time (vehicle per hour per lane).

Fu and Rilett (1997)
Fu and Rilett (1997) assume that the incident duration is a random variable with known distribution and all other parameters are assumed to be deterministic. The incident occurs at time T^* and lasts for D^*. The incident duration D^* is a random variable with a known probability density function $f_{D^*}(x)$. The traffic arrival rate q (q_{a1}), the normal capacity, c (q_c) and the incident capacity c^* (q_d) are known and constant. One problem with the model is that if more than one parameter is a random variable, its solution becomes mathematically intractable. And since, in the real world, q, c, and c^* are not deterministic, the model has limited value.

The probability distribution of the incident delay of a given vehicle a, (d_a) depends on the probability distribution pattern of the incident duration and the time the vehicle arrives at the link (T_a), where T_a is assumed to be fixed. Three delay models are provided:

- *No Delay Regime* - In the no-delay regime the vehicle arrives after the incident has been cleared and the associated queue has been dissipated.
- *Fixed-Delay Regime* - In the fixed-delay regime the vehicle arrives at the incident location at time and joins the queue, where the incident queue has not dissipated, and will not be dissipated until after the vehicle traverses the link. In this situation the queue dissipates at a rate of and consequently the vehicle experiences the maximum delay:

$$d_m = \frac{q - c^*}{c^*}(T_a - T^*) \qquad \text{(A-9)}$$

- *Variable-Delay Regime* - In the variable-delay regime the vehicle arrives at the link and either (a) the incident has been cleared but some portion of the queue remains or (b) the incident has not been cleared but will be cleared before the vehicle exits the link. In this situation, the dissipation rate of the standing queue is some combination of c and c^*, and consequently the queue delay will lie between the previous two cases.

Ideally, anticipated and quantitative information such as time dependent delay caused by an incident should be estimated and provided to drivers. Procurement of such information is however not a trivial task because of the complex interactions among various factors such as incident location and severity, incident response capability, demand fluctuation and diverse driver responses to information. Moreover, most of these factors are subject to high uncertainty and information available to quantify them is often incomplete and subjective in nature. Consequently, provision of crisp values of expected delays to drivers would inevitably lower drivers' trust in the accuracy of the provided information because the actual delays they would experience will be either larger or smaller than what were suggested.

To provide a better delay estimation for each individual driver approaching the incident site, Fu and Rilett (1997) developed a dynamic and stochastic model within a traditional deterministic queuing model approach for predicting the delay that a vehicle would experience traveling through an incident location. The model explicitly considers the stochastic attributes of incident duration. They derived a mixed discrete and continuous vehicle-delay model and estimated the mean and variance of vehicle delay.

The use of a facility-specific incident duration distribution produces both the corresponding vehicle delay distribution as well as the individual vehicle delay based on the vehicle arrival time. The model's main deficiency is that it ignores the spatial characteristics of incident delay.

Task 3: Data Collection and Analysis
Subtask 3.1 (Data Assembly and Analysis)

Prepare by:

Ali Afshar, Research Associate
John C. Falcocchio, Ph.D. PE
Urban ITS Center
Polytechnic University
Brooklyn, NY 11201

Tel: 718.260.3612
Fax: 718.260.3433
Email: jfalcocc@poly.edu

January 31st, 2006

1.0 Background

1.1 Purpose

Task 3 of scope of the work is divided into two subtasks: *Subtask 3.1* deals with data collection and summary of available data for model development; while *Subtask 3.2* covers data analysis/model development. This TM reports on the results of *Subtask 3.1*.

1.2 Subtask 3.1

Data assembly was the main objective of *Subtask 3.1* which was performed by Polytechnic University. The task started by searching for potential data sources that could be used to identify non-recurring incident characteristics and associated attributes. The task focused on agencies that are involved in highway incident management/monitoring as well as those that collect roadway attributes data, such as physical characteristics, and traffic flow.

2.0 Incident Data Sources

The scope of work identified the Integrated Incident Management System (IIMS) database as the primary source of data acquisition. But during the initial steps IIMS data were not readily available, and other potential data sources such as TRANSCOM database were identified and considered.

Content investigation of TRANSCOM databases convinced team members that the TRANSCOM dataset is the most comprehensive dataset available for the region although it does not include all desired data variables. In order to complement the TRANSCOM database, other related data sources such as traffic volumes and roadway characteristics were obtained from other resources as described below.

The following section provides information about the type of operation and data gathering procedures for various types of incident.

2.1 Directly from the Field

IIMS is a real time incident management system that enhances the communication of incident data among incident managers at operations centers and incident response personnel at the incident scene deployed by NYSDOT. IIMS is a multi-agency project managed and sponsored by the New York State Department of Transportation (NYSDOT), in partnership with New York City Department of Transportation (NYCDOT), the New York City Police Department (NYPD) and the New York City Office of Emergency Management (NYCOEM).

After an incident is detected, it is common to get verification from the field personnel to distinguish real incident from false alarms, and dispatch appropriate response. The trained field personnel will be able to provide accurate incident status and scene information for better management of the incident.

IIMS is unique in its ability to transmit incident scene data in real time to incident operations centers and mobile units. These data include the information required to effectively select the appropriate responders and equipment for clearance. Using IIMS, incident responders with mobile computers collect incident information and transmit it to inter-connected agencies. This information includes quick data reports, accurate location based on GPS, and digital pictures taken at the incident scene. During a major incident, IIMS can be used to coordinate a multi-agency response at and near the incident scene. Field responders can report infrastructure damage, environmental hazards, and incident severity. IIMS can also be used to support incident command by identifying staging areas, as well as emergency response and evacuation routes.

At present time, mobile equipment is installed in 19 NYPD, 3 NYCDOT, 3 NYCOEM and 3 MTA Police vehicles. NYPD vehicles are operated by highway patrol officers and NYCDOT vehicles are operated by emergency response supervisors. Multiple centers are equipped and many are operated 24 hours/day, 7 days/week by incident managers.

The IIMS data dictionary contains valuable data elements that could be used as input for incident duration and delay model estimation. However, the IIMS database project is in the developing stage and its outputs were not ready for use at this time.

2.2 From Remote Centers

A. TRANSCOM

The Transportation Operations Coordination Committee (TRANSCOM) is a coalition of 16 transportation and public safety agencies in the New York – New Jersey – Connecticut metropolitan region. It was created to provide a cooperative, coordinated approach to regional transportation management. The member agencies are:

New York State Department of Transportation,
New York City Department of Transportation,
New York State Police,
New York City Police Department,
New York State Thruway Authority,
MTA New York City Transit,
New York State Bridge Authority,
MTA Bridge and Tunnels,
Metropolitan Transportation Authority,
Port Authority of New York and New Jersey,
New Jersey Department of Transportation,
New Jersey State Police,
New Jersey Turnpike,
New Jersey Transit Corporation,
Port Authority Trans-Hudson Corporation,
Connecticut Department of Transportation,

One of the TRANSCOM's responsibilities is to collect and disseminate real-time incident and construction information, 24-hours-a-day, to over 100 member agencies and affiliates through the Operations Information Center (OIC). TRANSCOM maintain a database of reported incidents for the coverage area. Each member agency has a workstation networked with the TRANSCOM server with an interface to add incident data into the system. The TRANSCOM database comprise of a variety of incident events on major roadways in the NYSDOT Region 11.

B. JTOC

The Joint Traffic Operation Center (JTOC) monitors traffic and special events from CCTV on Interstate Highways in the New York City as well as incoming calls from highway maintenance crew and yard personnel. The JTOC staff enters information into the TRANSCOM workstation. Meanwhile they report and share the information with NYPD TMC, NYCDOT TMC and other agencies as needs arise.

C. NYPD

The New York City Police Department (NYPD) has the authority and control over the traffic operation on transportation network. The highway patrols monitor traffic flow as well as other irregularity on the roadways and report that to the police dispatch. The communication between the highway patrol and dispatch center is verbal through radio frequencies. Operator at the dispatch center (NYPD TMC) input the verbal information into a database for record keeping and follow ups. In addition, NYPD TMC receives information from 9-1-1 dispatcher as well as video feed from JTOC and NYCDOT TMC. When they receive information about the highway incident/events they pass it to JTOC for incorporation into the workstation.

D. TMC

The New York City Department of Transportation Traffic Management Center (TMC) has CCTV coverage on the New York City owned transportation network. The staff operators identify traffic flow irregularity through video monitors and communicate with appropriate agencies to mitigate the problem by implementing proper strategies. The NYCDOT TMC by adjusting appropriate signal timing plan and providing traveler information through variable message signs help to control the traffic congestion. The staff at this TMC passes the information to JTOC as discussed above.

E. Long Island City TMC

The three traffic management centers in the Long Island City (TMC, NYPD TMC, JTOC) used to input the incident data in the TRANSCOM work station separately. Since February 2004, JTOC became the point of contact for all these TMCs to receive incident information and feed into the workstation. With this recent policy, the quality of input data relating to incident events has improved and missing data input has reduced.

F. TRANSMIT

The TRANSCOM's System for Managing Incidents and Traffic (TRANSMIT). The TRANSMIT system utilizes Electronic Toll and Traffic Management (ETTM) equipment, which is compatible with the E-ZPass system, for its traffic surveillance and incident detection purposes. The E-ZPass system is an electronic toll collection system, currently in operation the

New York metropolitan area as well as neighboring states. In the project study area tag reader antennas specific to the TRANSMIT system are installed at intervals of 0.4 to 2.2 miles.

Each time an E-ZPass tag equipped vehicle enters the capture zone of a Roadside Terminal (RST) location, its tag identification number (tag ID) and the detection time are recorded. Data containing tag ID, detection time, location and lane position is forwarded to the Operations Information Center (OIC) at Jersey City, NJ. The tag's ID is then encoded at the OIC into a random number to ensure the anonymity of the motorist. Such surveillance data is acquired at all locations continuously on a 24-hour basis and archived by 15 minute intervals. The vehicle travel times between successive readers are then determined from the stored data at the OIC.

3.0 Study Area and Time Period

The study team defined the study area for which boundary and extent of data coverage are described below.

The Interstate 278 from Goethals Bridge in the Staten Island to the Bruckner Expressway Interchange in the Bronx with 33.62 centerline miles was selected as the study corridor for the project. Further investigation along the corridor indicated that other highway branches which feed into the corridor should also be included in the review. For this reason the West Shore Expressway (R440) in the Staten Island, Prospect Expressway (R27) in Brooklyn, and Sheridan

Figure 1 – Study Area

Expressway (I-895) in the Bronx was also included in the study area to have a better understanding of the whole interactions among the traffic flow and its physical environment. The new corridor expansion added additional 13.33 miles to the study area and the total corridor length became 46.95 miles of arterial highways in the New York City.

The time frame for the data collection and evaluation was set from February, 2004 and beyond, for a period of at least one year. Starting in February 2004 the JTOC staffs were assigned the responsibility to accept calls from neighboring agencies to enter incident data into TRANSCOM workstation. The new policy has contributed to greater consistency and completeness of the incident event data.

4.0 Data Assembly

The first batch of TRANSCOM database covering all records in the NYSDOT Region 11 from February 1^{st} to September 31^{st} were received on October 25^{th}. The eight months of data were reviewed and initial tabulations were performed. The second set of incident data for October 1^{st} 2004 through March 31^{st} 2005 were received on April 19^{th} 2005.

The first step was to filter records in the region 11 datasets to include only roadways in the study area. Reviewing the data dictionary revealed that the event types are defined in 120 different terms, some of which do not represent non-recurring incidents.

As the next step, all incident records were reviewed and stratified into two major areas. The first group falls into non-recurring incidents and the second group represented recurring incidents such as rush hour traffic, special event, planned activities, road construction and repairs. The non-recurring related incident then consolidated into eight categories. These data were grouped according to the eight categories of non-recurring incident types as specified in the Incident Duration Estimation Trees of the scope of work.

This TM will cover non-recurring incidents reported by TRANSCOM for the period of fourteen months (424 days) along the I-278 corridor as defined before.

4.1 TRANSCOM Incident Data

After reviewing incident records for missing values and excluding recurring incidents, the non-recurring incidents were identified in eight categories. The incident dataset provides information on the following data fields:

Incident ID;
Reporting Agency;
Incident Type;
Incident Location (facility name, roadway direction, nearest crossing street or exit/entrance ramp)*;*
Create Time (first time reported by any reporting agency)*;*
Close Time (last time reported that the incident is cleared)*;*
Lane Blockage (affected lanes and partial information on lane blockage)*;*
Pavement Conditions (less than half of the cases have this description)*;*
Weather Conditions (less than half of the cases have this description)*;*

In addition to the above fields there is a field labeled “description” which provides descriptive text for the event. This field is parsed in order to make the data available for potential analysis use. Data extracted from this field were manipulated for consistency, which will provide additional insight about the events such as the extent of the affected lanes location and their status.

4.2 Highway Sufficiency Ratings File (R11-SUFF)

The GIS version of the R11-SUFF data became available to Polytechnic on March 10th 2005. The R11-SUFF is an annual survey of the surface condition and physical characteristics for each section of New York State Touring Routes and Thruway Systems. The R11-SUFF provides a description of the physical characteristics of the roadway. The following presents a list of variables that are relevant to the project study:

Route Number;
County Name;
Starting Milepoint;
End Milepoint;
Section Length;
Number of Roadways;
Number of Lanes;
Pavement Width;
Shoulder Width;
Shoulder Type;
Median Width;
Median Type;
Percent of Trucks;
Annual Average Daily Traffic Volumes (AADT) – Two Way;
Functional Classification;
Adjusted Rated Capacity – Calc program – One Way;

The segmentation of the roadway in the R11-SUFF represents the status of the pavement conditions. This segmentation may vary each year based on the annual roadway pavement rating survey.

4.3 New York City Street Segment GIS File (NYCSSG)

The New York City Street Segment GIS files which became available in March 10th 2005, covers all roadways in the city including State highways. Each roadway is defined as a link connecting nodes and presented as separate records. This file contains street attribute field names and description for the street address data used in the Street Layer file and Street Alternate Name File. A review of this file indicates that *Speed* is a useful variable which can be used to expand the incident database. This speed information is the actual posted Speed Limit for the roadway.

4.4 New York State DOT Traffic Volume Counts (NYSDOT TVC)

The NYSDOT provided traffic volumes for five locations along the study corridor (three locations along Staten Island Expressway, two locations on Brooklyn Queens Expressway (BQE) one in Brooklyn and one in Queens). These traffic counts obtained from Automatic Traffic Recorder (ATR) during two consecutive weeks in May and another two weeks in June of 2004. Traffic volumes provided hourly traffic counts for both directions. There is also, classification count average weekday for each location in the month of May 2004.

4.5 TRANSMIT Database

The roadside/overhead antennas (E-ZPass Tag Readers/TRANSMIT readers) along the study corridor have the capability to record and calculate travel time between successive readers. The main server, which receives the signals from the reader antennas, computes the average link speed for the associated links and store data for each direction of traffic movement. Historical *Link Average Speed* and *Link Average Travel Time* data from TRANSMIT database could be retrieved for each link by direction, and each day of the week for 24 hours. The link speed is being continuously recorded by overhead antennas and conveyed to the computer server to record real time observation. These data are being aggregated and archived in 15 minutes intervals for later use and review.

The TRANSMIT database provides information on:

Link ID;
Date and Time;
Day Type;
Volume (Number of detected Vehicles);
*Speed (*Link Average Speed), and
Average Travel Time;

The TRANSMIT link is defined as the distance between two consecutive antenna. This distance varies considerably in the coverage area. The length variation is related to factors such as existence of right of way, existence of overhead structure, availability of the conduit for communication, etc. The Figure 2 shows the location of the TRANSMIT antennas along the study corridor.

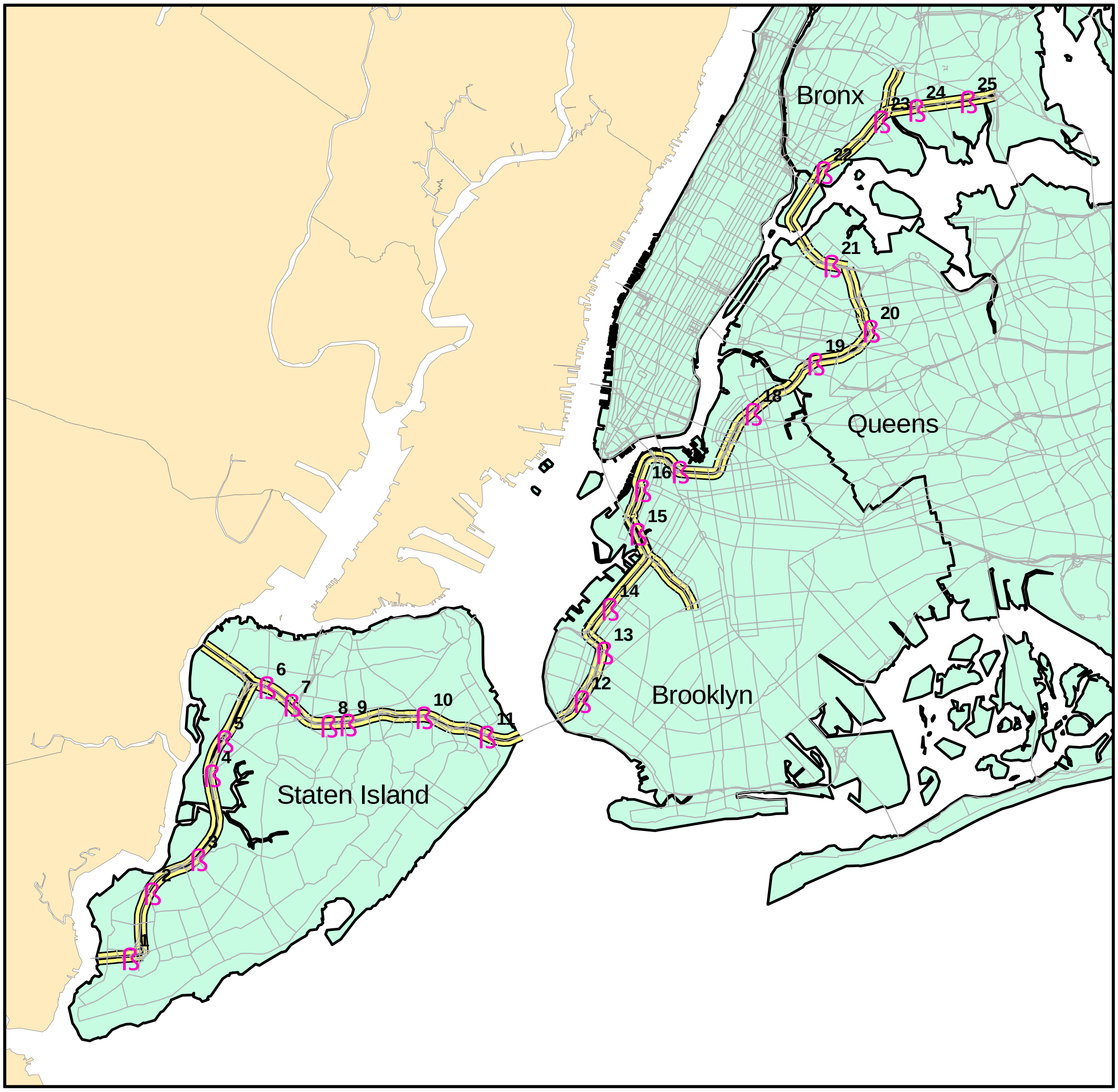


Figure 2 – Location of TRANSMIT Antennas

5.0 Data Fields

As mentioned before, data from other sources were utilized to expand the incident attributes for the analysis stage. The *Table 5.1* define the incident attributes as well its description and sources.

Table 5.1 – Incident attributes and Sources

	Data Fields		Data Source				
	Attributes	**Description**	**TRANSCOM**	**TRANSMIT**	**R11-SUFF**	**NYCSSG**	**NYSDOT**
1	**Incident ID**		X				
2	**Reporting Source**		X				
3	**Incident Time**	*Date and Time*	X				
4		*Day*		X			
5	**Incident Duration**	*Create Time*	X				
6		*Closed Time*	X				
7	**Event Types**		X				
8	**Incident Location**	*Facility Name*	X	X			
9		*County*	X				
10		*Direction*	X	X			
11		*Cross Street/Ramp*	X				
12		*Affected Lane/Shoulder*	X				
13		*Milepoint*			X		
14		*LinkID*		X			
15	**Incident Severity**	*Involved Lane(s)*	X				
16		*Blocked Lane(s)*	X				
17	**Roadway Characteristics**	*Number of Roadways*			X		
18		*Number of Lanes*			X		
19		*Pavement Width*			X		
20		*Shoulder Width*			X		
21		*Shoulder Type*			X		
22		*Median Width*			X		
23		*Median Type*			X		
24		*Percent of Trucks*			X		
25		*AADT Volume*			X		
26		*Adjusted Rated Capacity*			X		
27		*Volume (some locations)*					X
28		*Speed Limit (posted)*				X	
29		*Speed (calculated)*		X			
30		*Average Travel Time (calculated)*		X			
31	**Environment**	*Pavement Conditions*	X				
32		*Weather Conditions*	X				

The following section defines each incident attribute in detail.

Incident ID
The system generates an *ID* for each incident being input into the TRANSCOM workstation server.

Reporting Source
The first agency that provides information about the incident is being considered as the *Reporting Source*. If more than one responding unit provides support and assistance during the incident management process, still the first responding agency is the *Reporting Source* and none of the other responding units being acknowledged in the TRANSCOM data base.

Incident Time *Date and Time*
Day
Each incident in the TRANSCOM database has a time stamp which defines the *Date* and *Time of Day*.

The TRANSMIT database records information on *Date, Time* in 15 minute intervals and *Day* of the week.

Incident Duration *Create Time*
Closed Time
The time of incident occurrence is usually unknown, but the verification time by an authority is a reasonable value as starting time for incident duration calculation. It is a common practice that the first responding agency observing an incident through CCTV or arriving at an incident scene, to report to the higher authority responsible for incident management and dispatch. That "time" is being recorder in the TRANSCOM database as *Create Time*. Many other responding units may be required to be called for help and be involved in the incident management process. The TRANSCOM database is not recording the arrival or departure time of other responding units. The other time unit being recorded is the *Closed Time* which represents the closure of the incident and departure of the responding unit(s). The difference between these two times is a reasonable presentation of the incident duration. The incident occurrence time is neglected and roadway traffic recovery time is not included either, since the first one is hard to pinpoint and the latter is hard to measure.

Event Type	*1*	*Traffic Incidents*	*Disabled*	
	2		*Accident*	*Property Damage*
	3			*w. Personal Injury/Fatality*
	4	*Road Hazard*		
	5	*HAZMAT (Fuel/Cargo Spill)*		
	6	*Vehicle Fire*		
	7	*Weather Related*		
	8	*Other (Police/Fire/EMS/Tow Truck Activity)*		

In the TRANSCOM database *Event Type* are defined in 120 different terms (see Appendix 1) in order to provide a list from pull down menu for the user to select and identify an event type type. This list is long enough to cover a broad range of possible incidents; however, incidents are unique in nature and hard to stratify them in a representative manner. For the purpose of this project, the *Event Type* are broke down into recurring and non-recurring incidents. Then the non-recurring incidents categorized into eight tiers (see Appendix 2) for analysis purposes.

Incident Location *Facility Name*
County
Direction
Cross Street / Ramp
Affected Lane or Shoulder
Milepoint
Link ID

The TRANSCOM database provides information on incident location. However, the information provided in the database gives an approximation of the incident location not an exact point on the transportation network. The incident location is being recorded by the Facility and Direction, it also gives the nearest crossing road / ramp before the incident (From Loc) and after incident (To Loc) and in order to emphasize the proximity to each of these crossing roads/ramps ads a field called *Article* with 32 different terms (see Appendix 3) to give a better understanding of the location. Because of complexity of the roadway in the study area, pinpointing the incident location along the study corridor can not be exact but will be the best estimate based on the given description.

An assumption has been made for the above mentioned "Article" for geocoding the incident locations. The following *Table 2* depicts a recommendation made for this assumption collectively with Polytechnic and State DOT staff on February 22nd 2004 at NYMTC. Since transportation links are presented as vector files defining roadway as a link connecting nodes, the following policies has been made during geocoding stage:

Table 5.2 – Policy Assumptions for Location Description

Location Description			**Assumption for Geocoding**
Area of	At	From	Distributed around the node of that particular crossing road/ramp
Ramp from	Ramp to	to	
by	through	under	
over	into	out of	
entering at	exiting at	into	
out of	in		
East of	North of	South of	Distributed within 10% of the link length from the corresponding node
West of	Just past	Bypassing	
Approaching	Near	Before	Distributed within 20% of the link length from the corresponding node
In the vicinity of			
Between			Distributed in the middle of the corresponding link

The *Milepoint* provided in the R11-SUFF define the starting and ending points of the surface rating and that usually varies by year. The segment *Milepoint* associated with incident location does not provide enough information to accurately pinpoint the incident location.

Each incident falls within a link from TRANSMIT *Link ID*. The link in the TRANSMIT system varies from 0.4 to 2.2 miles long covering segments of the highway with more than one on/off ramp. The relationship between incident location and TRANSMIT is useful for vehicular speed and travel time derivation but not a guide for indication of incident location.

Incident Severity *Involved Lane(s)*
Blocked Lane(s)

The TRANCOM server interface provides various fields for roadway lanes description: “Number of Lanes”, “Lane Status”, “Total Lanes”, “Lanes”, and “Lane Details” for the incident period.

The combination of the “Number of Lanes” and “Lanes” fields provide a picture of the lane blockage during the incident period. However, the lane blockage is not for the whole duration of the incident, and there is no other means from the database to identify the lane blockage duration. It can be used to identify the impacted lanes by the incident.

“Lane Status” defines affected lanes as the result of the incident; this field is mostly incomplete.

“Total Lanes” provides information on the number of lanes which is a needed field for incident impact analysis. However, this is not fully populated by the operator, which can not be utilized in this manner.

For the locations that have more than on link connecting two nodes, the “Lane Details” provides information about the roadway location as: “Both Level”, “HOV Lane”, “Inner Roadway”, “Local Lanes”, “Lower Level”, “Outer Roadway”, “Ramp”, “Service Road”, and “Upper Level”

The information about the number of vehicles involved in the incident as well as number of responding units are not captured by this database. The information about injuries, fatalities and pedestrian involvement are also not available through this dataset. The recent data could be obtained from police accident reports which were not readily available to our team at present time.

Roadway Characteristics *Number of Roadways*
Number of Lanes
Pavement Width

Shoulder Width
Shoulder Type
Median Width
Median Type
Percent of Trucks
AADT Volume
Adjusted Rated capacity
Volume (Hourly Intervals)
Speed Limit (posted)
Speed (calculated)
Average Travel Time (calculated)

As mentioned before, R11-SUFF provides information on the roadway segments based on the pavement rating and sections does change by pavement condition at each year review. The information provided are valuable in the sense that roadway characteristics such as number of lanes, lane width, and roadway clearance does not usually change.

The “Percent of Trucks” is an average value over an average day and provide a general picture for the roadway flow. It does not provide accurate value to be used for roadway segment operational performance measure for a particular incident.

“Adjusted Rated Capacity” is a rough estimate for Level of Service E with no incident. LOS E represents operating conditions at or near the capacity level. This value will be used as best estimate of the roadway capacity under normal conditions and impact of incidents on capacity reduction will be measured against that.

The “Volume” in TRANSMIT database represent the number of vehicles equipped with the E-ZPass tag traversed the link. It just shows the portion of the traffic stream with the device which detected at the two consecutive roadside antenna and does not even include vehicle with tag who exit or entered in between.

The hourly “Volume” from NYSDOT is the most relevant traffic volume for the operational performance measures. These volumes are collected over four weeks and in the month of May and June 2004. The traffic volumes for this period can be expanded for the rest of the study period in combination with AADT.

The roadway “Speed” is the posted speed limit and provided in the New York City Street Segment GIS files, which is static and fixed.

The TRANSMIT system calculate the vehicular “Speed” within the two consecutive readers from equipped vehicles. This value is the average link speed and is the best estimate of the flow speed.

The “Average Travel Time” is also calculated in the TRANSMIT system and shows the link average travel time for every 15 minutes.

Environment	*Pavement Conditions*
Weather Conditions

The above values are given in the database and could be easily incorporated in the data analysis stage.

6.0 Incident Data Summary Analysis

The non-recurring incidents extracted from the TRANSCOM database for a fourteen months period from February 1st 2004 to the March 31st 2005 populated with other pertinent attribute information from other resources is the final dataset utilized for data analysis in this project. This includes 1907 number of non-recurring incident records. The actual roadway directions vary in the study area and for the purpose of clarity the following rule being applied to revised roadway direction in the database for consistency:

I-278 including Staten Island Expressway, Gowanus Expressway, Brooklyn Queens Expressway, and Bruckner Expressway is considered as east-west direction;

Sheridan Expressway I-895 and West Shore Expressway R-440 are considered as north-south direction;

Prospect Expressway is considered as east-west direction;

It should be noted here that, at present time, NYSDOT is utilizing Congestion Needs Assessment Model (CNAM) to estimate roadway delays. Any input for the CNAM model should be in a format compatible with the existing CNAM structure. For example, in the scope of the work incidents are grouped into eight categories based on the proposed incident tree. However, CNAM recognize incident types by their lane blockage in four categories (incident blocked one lane, incident blocked two lanes, incident blocked three or more lanes, and incident occurred on the shoulder).

For the potential utilization of the study findings in the CNAM model, we created additional summary tables for the CNAM model and indicated that they are for "CNAM".

6.1 Roadway Facilities and Their Ownership

The *Interstate 278* has different local names along its length in the New York City. It is called *Staten Island Expressway* in Richmond County, it turns to *Gowanus Expressway* in part of Brooklyn and then change to *Brooklyn Queens Expressway* and continues into Queens and after Triborough Bridge in the Bronx its name become *Bruckner Expressway* until it meets I-95 at Bruckner interchange. It has a three moving lanes on each direction and traverses in an east-west direction. This Expressway is highly utilized by vehicles and it carries total of 202,000 vehicles per day between Staten Island and Brooklyn (both directions combined). The heavy vehicles volumes vary from 4.5% to 25.5% on an average weekday and it reaches up to 43% at some segments. The lanes are 12 feet wide except in the Queens, which is 10 feet for the most of its length. The study corridor is lacking continuous shoulder except in the Staten Island. This corridor is mostly elevated in Brooklyn and there are not enough refuge areas along the roadway for disabled vehicles and emergency stops.

The *Route 440* in the Staten Island from Outerbridge Crossing is known as *Richmond Parkway*. Its name turns to *West Shore Expressway* until it merges with I-278. Again after diverging from I-278 it becomes *Martin Luther King Expressway* and continues north up to Bayonne Bridge. This section of R 440 is not included in the study corridor. The R 440 has two moving lane with right shoulder along its length. The AADT is 93,000 vehicles per day (both directions combined) and average heavy vehicle is 7%.

The highways in the study corridor are under NYSDOT ownership and NYCDOT is responsible for the operation and maintenance. The MTA Bridges and Tunnels owns the Verrazano Narrows Bridge, and PA NYNJ owns the Goethals Bridge and Outerbridge Crossing. *Table 6.1* shows the agencies responsible for incident reporting related to non-recurring incident in the study area. The NYSDOT with 96.3% of reported incidents is the main source of the incident input to the TRANSCOM database followed by NYPD, TRANSCOM, NYCDOT and MTA & PA NYNJ with 1.9%, 1.2%, 0.3% and 0.3% of all incidents during the study period respectively.

6.2 Incident Types and Roadway Facilities

The distribution of incidents along the corridor is presented in the *Table 6.2*. It shows that 49.3% of the incidents occurs along the Brooklyn Queens Expressway (BQE), it represents only 27.8% of the corridor length. This translates into 2.22 incidents per day and 1.71 incidents per 1,000,000 VMT.

Another 33.4% of incidents occurred along Gowanus Expressway with 15.2% of the corridor length. This translates into 1.50 incidents per day and 1.64 incidents per 1,000,000 VMT.

But the Staten Island Expressway with 19.4% of the corridor's length contains only 5.8% of all non-recurring incidents in the corridor, or 0.26 incidents per day and 0.22 incidents per 1,000,000 VMT. The Bruckner Expressway, which is mostly elevated along its length, experiences 0.25 incidents per day or 0.51 incidents per 1,000,000 VMT, while it accounts for 10.9% of the corridor mileage.

The Prospect Expressway experiences 73 non-recurring incidents or an average of 0.17 incidents per day (it had 1.18 incidents per 1,000,000 VMT). The West Shore Expressway/R440 with 37 recorded non-recurring incidents and Sheridan Expressway with only 2 incidents have the lowest incident rates in the study corridor.

Table 6.2 shows an interesting trend of incident rate along the corridor: Although I-278 covers 43% of the study corridor in Brooklyn and Queens it experiences almost 83% of all incidents; while the other 57% of the corridor length experiences only 17% of the incidents. There is a striking difference in incident rates between sections of the I-278 corridor. The VMT-based incident rate for the Gowanus and BQE section of the corridor is approximately 13 times higher than West Shore Expressway, and 8 times higher that of the Staten Island Expressway. These differences are partly attributed to the deficiencies in design standards prevailing at the time of their construction.

The *Table 6.3* depicts breakdown of incident types for each facility. The “property damage” accidents are accounted for 42.9% of all non-recurring incidents followed by “disabled vehicle”, “disabled truck”, “road hazard”, accident with “personal injury”, “Hazmat”, “vehicle fire” and “weather related” incidents with 36.4%, 9.4%, 4.2%, 3.5%, 1.5%, 1.5% and 0.5% of all incidents respectively. This table provides the insight about the characteristics of the non-recurring incidents in the study corridor. It shows that incidents with property damage and personal injury account for 45.4% of total incidents and disabled vehicles & trucks account for another 45.8%. The road hazard which includes pothole repairs, police activity, water-main break, missing manhole, sewer gate and downed pole account for 4.2%.

Table 6.4 shows that distribution of the incidents along the study corridor is not random. Based on the location information provided in the database, the table is organized according to the proximity of the incident location to the nearest crossing road or ramp. This table shows that there are more incidents at some locations than at other locations. The Brooklyn Queens Expressway/I-278 has the highest number of recorded non-recurring incidents in the vicinity of the Atlantic Avenue, Hamilton Avenue, and Kosciusko Bridge with 15.1%, 10.3% and 8.2% respectively followed by Gowanus Expressway/I-278 in the vicinity of the 39th Street, Prospect Expressway, and Gowanus Canal with 8.1%, 6.4% and 6.0% of incidents respectively. These differences in incident could be due to other contributing factors such as roadway vertical and horizontal alignments, sight distance, weaving sections, merge and diverge area, and etc. The existing database does not provide information on these variables and further investigation is not feasible with the database limitations.

6.3 Incident Types and Lane Blockage by Facility

Table 6.5 depicts that 90% of the non-recurring incidents in the corridor occurred along the sections with no shoulder. *Table 6.6* shows that 77.8% of the incidents blocked one travel lane while only 11.5% blocked two travel lanes and 2.6% blocked three or more travel lanes. And only 5.5% of all incidents occurred outside of travel lanes (shoulder area). *Table 6.7* identifies the facility of incidents blocking lanes, with and without shoulders. This table shows the number of incidents for each facility on sections with and without shoulders.

The accuracy of TRANSCOM database in regards to the lateral position of the incident location which this latter table was developed is questionable. As it can be seen in the table, 92 incidents locations were reported as “out of travel lane” (No Travel Lane) on section that there is no shoulder lane. Strange enough, 71 of them occurred along the Gowanus and BQE, which the highway is elevated and there is no shoulder lane according to the Highway Sufficiency Rating File.

This is a concern about the accuracy of the lateral location of the incidents in the TRANSCOM database as well as the shoulder lane information in the Highway Sufficiency Rating File. To clarify this issue the aerial photos for the study area and stated location of incident were investigated. The observation revealed that almost half of these incidents occurred on sections of the I-278 in Brooklyn and Queens with partial shoulder, which possibly used as refuge area for the involved vehicles and incident lateral location, were reported as on the shoulder. This issue may suggest further investigation and revision of the roadway segmentation in the Highway Sufficiency Rating file as well as more training for the reporting agencies by the State DOT.

The *Table 6.8* provides information on the relationship among incident type and lane blockages. It shows that almost 78% of incidents blocked only one lane. The range of one lane blockage varies for different incident types: 90.1% of disabled vehicles, 84.4% of disabled trucks, 79.3% of HAZMAT, 72.5% of property damages, 63.8% of road hazard, and 44.8% of personal injuries. However, for vehicle fire the possibility of blocking two lanes is higher than other incident types.

6.4 Incident Duration by Incident Type

The following section describes the finding related to the distribution of incidents duration. In order to have a clear picture of the distribution, the incident duration were depicted in 15 minutes intervals and two tables are developed. The *Table 6.9* shows the duration intervals by the incident type and the *Table 6.10* provide duration distribution by the lane blockage. The average duration of an incident varies from 27.8 minutes for disabled vehicle to 288.9 minutes for weather related incidents.

It is important to mention that the CNAM model divides the duration into three general categories: incidents with up to 15 minutes duration are considered as “Short”, and up to 30 minutes as “Medium.” Beyond 30 minutes the incident is considered of long duration. This general interval is also incorporated in these two tables. These tables show that 27.5% of all non-recurring incidents are considered as “Short”, 23.0% as “Medium” and 49.5% as long duration.

Table 6.10 shows incident duration by the lane blockage. The average duration of an incident varies from 38.5 minutes for three or more lanes blocked to 57.6 minutes for no travel lane blocked. The average duration by the lane blockage shows that incident with one lane blockage had an average of 40.7 minutes, with two lane blockage an average of 54 minutes, and 38.5 minutes for total closure (three or more lanes). It shows that 31% of incidents blocking one lane were cleared within 15 minutes, 24% up to 30 minutes and the other 45% took more than 30 minutes. At the same time, 19% of incidents blocking two lanes had short duration, 17% medium duration, and 64% long duration. The situation is some how different for the three or more lane blockage category, which actually means the roadway closure. It shows that 20% of this type was cleared within 15 minutes, 34% took up to 30 minutes and 46% are considered as incidents with long duration.

As mentioned earlier, each incident type is unique, which makes their duration range considerably wide. *Table 6.11* is created to depict some of these variations. This table relates the lane blockage with the incident type and provides duration statistics for each category. For example, the average incident duration for “disabled vehicle” is 27.8 minutes this average duration is then divided by lane blockage. The average duration for one lane blockage is 25.8 minutes, and for two lane blockage is 49.5 minutes (almost twice). The average duration for "disabled truck" is 51.6 minutes, which is 42.9 minutes for one lane blockage, and 130.1 minutes (more than three times) for two lane blockages.

The trend in incident duration from one lane to two lanes is not followed for incidents which blocked three or more lanes. It implies that the transportation users and authorities can not afford complete roadway blockage, and in the case of incident with roadway closure, all attempts are being made by responding agencies to open travel lanes and clear the incident scene. That is one of the attributing factors for lower duration for the incidents with three lanes blockage conditions. For example, the average duration for vehicle fire blocking three lanes is 32.8 minutes while it is 56.4 minutes for one lane blockage and 73.9 minutes for two lane blockage. This attribute will help the research team to have a better understanding of the factors affecting incident duration for each incident type during model development.

6.5 Incident Type by Day of Week and Time of Day

The study period covers fourteen months continuously, which will show a general picture of the distribution of incidents for each day of week. *Table 6.12* is generated to depict the incident trend. It shows that Saturdays have the lowest number of incidents and Tuesdays & Wednesdays have the highest number of recorded incidents. It also shows that 25% of all incidents occurred on weekends and 75% on weekdays. The two weekend days are 28.6% of the week but they experience 25% of the incidents. The lower number of incidents could be related with the lower traffic volumes and as well as lower truck traffic.

To complete the picture of the time distribution of incidents, *Table 6.13* was generated to depict the distribution of incidents type by the month of the year. The distribution of incident type and total incidents for each month is considerably consistent except for February 2004 and January 2005. As shown in the table, the number of property damage, disabled vehicle and weather related incidents are the highest during the study time period for January 2005 and the number of disabled vehicle and personal injuries incidents are the lowest for the February 2004. There might be other influencing factor for these two cases which require further investigation. Beyond that, this table does not shows significant changes in the number of incident based on the type and month.

6.6 Incident Frequency by Time of Day

The first table generated here is the distribution of the incidents during the course of the day. As it can be seen in *Table 6.14*, it fluctuates in accordance with traffic volume during peak and off peak periods. This table shows that on average, 4:00 to 5:00 PM has the largest number of reported incidents followed by 7:00 to 8:00 AM. The lowest number of reported incidents is after midnight from 1:00 to 3:00 AM.

The CNAM model differentiates only peak and off peak period. The peak period consists of 6:00 to 10:00 AM and 3:00 to 7:00 PM, and the other times are considered as off peak. The distribution of incidents according to the CNAM model is presented in *Table 6.15*. The interesting trend in this table is that 44.9% of all incidents are reported during one third of the time and 55.1% are during the two third of the time which is considered as off peak.

The last column of the table shows the general trend of reported incidents during the study time period. It shows that during the night time 12:00 to 6:00 AM the study area experienced the lowest number of incidents per hour. The highest number of incidents per hour is experienced during afternoon rush hour 3:00 to 7: PM followed by morning rush hour 6:00 to 10:00 AM.

6.7 Incident Frequency by Type of Incident and Time of Day

Table 6.16 is arranged to show the frequency distribution of incident type by time of day. Under each incident type there are three columns. The first column shows the hourly incident type; the second one shows the percentage out of the total incidents and the third one shows the percentage of the incident type for that time period.

As seen before, this table shows the number of incidents fluctuates according to the traffic volume during the course of a day but it does not show significant relationship between incident type and time of day. In other word, incident type is a random phenomena and it can occurs anytime and is hard to predict.

6.8 Weather and Pavement Conditions

The following two sections provide more detail information about the weather and pavement conditions during the time of incidents and incident type. The database base does not provide information in order to correlate these conditions as the contributing factors to the incident itself except those incidents were labeled as weather related incident in the incident type category.

6.8.1 Incident Type and Weather Conditions

The TRANSCOM database identifies weather conditions in twelve different terms and CNAM model utilize only two terms. The terms in the TRANSCOM database are matched with the terms in CNAM and the for the purpose of clarity, *Table 6.17* is created to show these terms side by side. It shows that out of 1907 incidents there are only 1251 records with weather related input. Out of these incidents, 85.1% occurred during normal/(clear/cloudy) conditions and only 14.9% during adverse conditions, of which the rain fall is accounted for 9.4%.

6.8.2 Incident Type and Pavement Conditions

The pavement conditions also have eight different terms in TRANSCOM database, while CNAM uses only normal and adverse to identify pavement conditions. *Table 6.18* shows the finding of 1246 incident with pavement condition information in the database; the rest of the TRANSCOM records did not provide pavement information for the incidents. The term Dry from TRANSCOM is considered by Poly as Normal pavement condition for CNAM and the other terms are considered as Adverse pavement condition for CNAM. As the table shows, 81.6% of incidents occurred during normal pavement condition and only 18.4% during the adverse pavement conditions.

A review of the above tables indicates that in the study area there is not significant correlation between incident type and the weather/pavement conditions.

TABLE 6.1- Source of TRANSCOM Non-Recurring Incident Data, by Facility

Facility Name	NYSDOT- R11 Field Crew	NYSDOT - JTOC	NYC PD - LIC TMC	NYC PD - HWY Staten Island	NYCDOT - OER	TRANSCOM	MTA Bridges & Tunnels	PA NYNJ - Staten Island Bridges	Total for Facility	%
	Incident Data Sources									
	NYSDOT		NYPD		NYCDOT		Bridges & Tunnels			
Brooklyn Queens Expressway/I-278	895	7	17		5	17			941	49%
Gowanus Expressway/I-278	623	4	3			4	3		637	33%
Staten Island Expressway/I-278	97	1	8	2		1		1	110	6%
Bruckner Expressway/I-278	102	1	4						107	6%
Prospect Expressway	72					1			73	4%
West Shore Expressway/R440	32		1	2	1			1	37	2%
Sheridan Expressway/I-895	2								2	0%
Total	1823	13	33	4	6	23	3	2	1907	100%
%	95.6%	0.7%	1.7%	0.2%	0.3%	1.2%	0.2%	0.1%	100%	
Total for Agency	1836		37		6	23	5		1907	
%	96.3%		1.9%		0.3%	1.2%	0.3%		100.0%	

Source: TRANSCOM Non-Recurring Incident Data from February 1st, 2004 through March 31st, 2005

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.2 - Incident Distribution by Facility (Incident Rate)

Facility Name	Recorded Incidents** #	Recorded Incidents** %	Facility's Length Mile	Facility's Length %	Incident Rate per Day**	Incident Rate per Mile per Day	Incident Rate per 1,000,000 VMT*
Brooklyn Queens Expressway/I-278	941	49.3%	12.76	27.8%	2.22	0.17	1.71
Gowanus Expressway/I-278	637	33.4%	6.96	15.2%	1.50	0.22	1.64
Staten Island Expressway/I-278	110	5.8%	8.88	19.4%	0.26	0.03	0.22
Bruckner Expressway/I-278	107	5.6%	5.02	10.9%	0.25	0.05	0.51
Prospect Expressway	73	3.8%	1.78	3.9%	0.17	0.10	1.18
West Shore Expressway/R440	37	1.9%	9.33	20.3%	0.09	0.01	0.13
Sheridan Expressway/I-895	2	0.1%	1.12	2.4%	0.00	0.00	0.11
Total	**1907**	**100%**	**45.85**	100.0%	**4.50**	**0.10**	**0.94**

Source:TRANSCOM Non-Recurring Incident Data from February 1st, 2004 through March 31st, 2005

* Weekday AADT and section length from Highway Sufficiency File used for VMT calculation

** 424 days of incident data

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.3 - Incident Types by Facility

Facility Name	Property Damage	Disabled Vehicle	Disabled Truck	Road HAZARD	Personal Injuries	HAZMAT	Vehicle Fire	Weather Related	Total	%
	Incident Types									
Brooklyn Queens Expressway/I-278	398	357	87	33	34	12	14	6	941	49.3%
Gowanus Expressway/I-278	250	264	70	23	15	9	4	2	637	33.4%
Staten Island Expressway/I-278	63	16	9	13	2	3	3	1	110	5.8%
Bruckner Expressway/I-278	64	23	6	8	5	1	0	0	107	5.6%
Prospect Expressway	26	29	5	0	10	0	3	0	73	3.8%
West Shore Expressway/R440	17	4	2	3	1	4	5	1	37	1.9%
Sheridan Expressway/I-895	1	1	0	0	0	0	0	0	2	0.1%
Total	819	694	179	80	67	29	29	10	1907	100%
%	42.9%	36.4%	9.4%	4.2%	3.5%	1.5%	1.5%	0.5%	100%	

Source:TRANSCOM Non-Recurring Incident Data from February 1st, 2004 through March 31st, 2005

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.4 - Incident Longitudinal Distribution by Facility

Facility Name	Location (Crossing Road/Ramp)	# of Recorded Incidents	%	%
Brooklyn Queens Expressway/I-278	Exit 27 - Atlantic Avenue	288	30.6%	15.1%
	Hamilton Avenue	196	20.8%	10.3%
	Kosciuszko Bridge	156	16.6%	8.2%
	Exit 28 - Brooklyn Bridge	41	4.4%	2.1%
	Exit 35 - I-495/Long Island Expressway	39	4.1%	2.0%
	Meeker-Morgan Avenues	34	3.6%	1.8%
	Exit 32 - Metropolitan Avenue	24	2.6%	1.3%
	Exit 28A - Cadman Plaza	22	2.3%	1.2%
	Exit 36B - NY 25/Queens Boulevard	20	2.1%	1.0%
	Exit 33 - McGuiness Boulevard/Humboldt Street	19	2.0%	1.0%
	Exit 30 - Flushing Avenue	18	1.9%	0.9%
	Williamsburg Bridge	14	1.5%	0.7%
	Exit 31 - Wythe/Kent Avenues	11	1.2%	0.6%
	Exit 29 - Manhattan Bridge	9	1.0%	0.5%
	Exit 37 - Broadway/Roosevelt Avenue	8	0.9%	0.4%
	Gowanus Expressway	8	0.9%	0.4%
	Exit 39E - Grand Central Parkway	6	0.6%	0.3%
	Exit 38 - NY 25A/Northern Boulevard	5	0.5%	0.3%
	Sackett Street	5	0.5%	0.3%
	Exit 40 - 30th Avenue/Grand Central Parkway split/m	3	0.3%	0.2%
	Kane Street	2	0.2%	0.1%
	Tillary Street	2	0.2%	0.1%
	27th Street	1	0.1%	0.1%
	58th street	1	0.1%	0.1%
	Bedford Avenue	1	0.1%	0.1%
	Congress Street	1	0.1%	0.1%
	Exit 41W - Grand Central Parkway/Triboro Bridge	1	0.1%	0.1%
	Flushing Avenue	1	0.1%	0.1%
	Vanderbilt Avenue	1	0.1%	0.1%
	Varick Avenue	1	0.1%	0.1%
	Water Street	1	0.1%	0.1%
	Union Street	1	0.1%	0.1%
	36th Street	1	0.1%	0.1%
	Total	**941**	**100%**	**49.3%**
	39th Street	155	24.3%	8.1%
	Exit 23 - NY 27/Prospect Expressway	122	19.2%	6.4%
	Gowanus Canal	115	18.1%	6.0%
	65th Street	78	12.2%	4.1%
	Hamilton Avenue	58	9.1%	3.0%
	Exit 25/26 - Hamilton Avenue	44	6.9%	2.3%
	Brooklyn Queens Expressway	20	3.1%	1.0%
	86th Street	7	1.1%	0.4%
	Belt Parkway	7	1.1%	0.4%
	20th Street	4	0.6%	0.2%
	Verrazano-Narrows Bridge	4	0.6%	0.2%
	92nd Street	3	0.5%	0.2%
	Fort Hamilton Parkway	3	0.5%	0.2%
	25th Street	2	0.3%	0.1%
	50th Street	2	0.3%	0.1%
	Mill Street	2	0.3%	0.1%
	17th Street	1	0.2%	0.1%
	23rd Street	1	0.2%	0.1%
	29th Street	1	0.2%	0.1%
	33th Street	1	0.2%	0.1%
	35th Street	1	0.2%	0.1%
	38th Street	1	0.2%	0.1%
	3rd Avenue	1	0.2%	0.1%
	40th Street	1	0.2%	0.1%
	45th Street	1	0.2%	0.1%
	46th Street	1	0.2%	0.1%
	56th Street	1	0.2%	0.1%
	Total	**637**	**100%**	**33.4%**

Source: TRANSCOM Incident Data from February 1st, 2004 through March 31st, 2005

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.5 - Incident Types Distribution and Availability of Shoulder Lane

	Shoulder Exists*					
Incident Type	**No**		**Yes****		**Total**	
Property Damage	708	86.4%	111	13.6%	819	42.9%
Disabled Vehicle	655	94.4%	39	5.6%	694	36.4%
Disabled Truck	166	92.7%	13	7.3%	179	9.4%
Road HAZARD	72	90.0%	8	10.0%	80	4.2%
Personal Injuries	62	92.5%	5	7.5%	67	3.5%
HAZMAT	22	75.9%	7	24.1%	29	1.5%
Vehicle Fire	21	72.4%	8	27.6%	29	1.5%
Weather Related	8	80.0%	2	20.0%	10	0.5%
Total	**1714**	**90%**	**193**	**10%**	**1907**	**100%**

Source: TRANSCOM Incident Data from February 1st, 2004 through March 31st, 2005

* NYSDOT Highway Sufficiency Rating file

** Includes Shoulder width of 8 feet or more

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.6 - Lane Blockage Distribution

Incident Type by Lane Blockage	Lane Blockage Description	Total		%	%
1 Lane	1 lane blocked	1461	1484	76.6%	**77.8%**
	HOV lane blocked	10		0.5%	
	1 lane and exit ramp blocked	5		0.3%	
	1 lane and right shoulder blocked	4		0.2%	
	1 lane and entrance ramp blocked	4		0.2%	
2 Lanes	2 Lanes blocked	220	220	11.5%	**11.5%**
3 or More Lanes	3 or more lanes blocked	49	50	2.6%	**2.6%**
	3 lanes and entrance ramp blocked	1		0.1%	
No Travel Lane	Right shoulder blocked	53	107	2.8%	**5.6%**
	Exit ramp blocked	22		1.2%	
	Entrance ramp blocked	16		0.8%	
	Left shoulder blocked	13		0.7%	
	Acceleration lane blocked	2		0.1%	
	Deceleration lane blocked	1		0.1%	
N/A*	Not Available	46	46	2.4%	**2.4%**
	Total	**1907**	1907	100%	**100%**

Source: TRANSCOM Incident Data from February 1st, 2004 through March 31st, 2005

* Refers to missing inputs in the TRANSCOM database

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.7 - Lane Blockage Distribution by Facility and Shoulder Existence

Shoulder Exists	Lane Blockage Classification (Incident Type)	Facility Name: Brooklyn Queens Expressway/I-278	Gowanus Expressway/I-278	Staten Island Expressway/I-278	Bruckner Expressway/I-278	Prospect Expressway	West Shore Expressway/R440	Sheridan Expressway/I-895	Total	%
No	**1 Lane**	750	517	6	49	32	5	0	1359	71.3%
	2 Lanes	102	66	0	9	2	1	0	180	9.4%
	3 or More Lanes	23	15	0	4	3	0	0	45	2.4%
	No Travel Lane*	40	31	1	13	7	0	0	92	4.8%
	N/A**	26	8	1	3	0	0	0	38	2.0%
	Subtotal	**941**	**637**	**8**	**78**	**44**	**6**	**0**	**1714**	89.9%
Yes	**1 Lane**	0	0	67	17	19	21	1	125	6.6%
	2 Lanes	0	0	25	10	3	2	0	40	2.1%
	3 or More Lanes	0	0	1	1	0	2	1	5	0.3%
	No Travel Lane*	0	0	4	1	7	3	0	15	0.8%
	N/A**	0	0	5	0	0	3	0	8	0.4%
	Subtotal	**0**	**0**	**102**	**29**	**29**	**31**	**2**	**193**	10.1%
Total		941	637	110	107	73	37	2	1907	100%
%		49.3%	33.4%	5.8%	5.6%	3.8%	1.9%	0.1%	100%	

Source: TRANSCOM Incident Data from February 1st, 2004 through March 31st, 2005

* Include incidents occurred on the ramp and off the travel lanes

* Refers to missing inputs in the TRANSCOM database

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.8 - Incident Lane Blockage by Incident Type

	Lane Blockage Type											
	1 Lane		2 Lanes		3 or More Lanes		No Travel Lane		N/A		Total	
Incident Types	#	%	#	%	#	%	#	%	#	%	#	%
Property Damage	594	72.5%	139	17.0%	29	3.5%	42	5.1%	15	1.8%	**819**	**42.9%**
Disabled Vehicle	625	90.1%	24	3.5%	7	1.0%	33	4.8%	5	0.7%	**694**	**36.4%**
Disabled Truck	151	84.4%	8	4.5%	7	3.9%	12	6.7%	1	0.6%	**179**	**9.4%**
Road HAZARD	51	63.8%	11	13.8%	1	1.3%	3	3.8%	14	17.5%	**80**	**4.2%**
Personal Injuries	30	44.8%	23	34.3%	1	1.5%	11	16.4%	2	3.0%	**67**	**3.5%**
Vehicle Fire	9	31.0%	11	37.9%	5	17.2%	2	6.9%	2	6.9%	**29**	**1.5%**
HAZMAT	23	79.3%	4	13.8%	0	0.0%	2	6.9%	0	0.0%	**29**	**1.5%**
Weather Related	1	10.0%	0	0.0%	0	0.0%	2	20.0%	7	70.0%	**10**	**0.5%**
Total	**1484**	**77.8%**	**220**	**11.5%**	**50**	**2.6%**	**107**	**5.6%**	**46**	**2.4%**	**1907**	**100%**

Source: TRANSCOM Incident Data from February 1st, 2004 through March 31st, 2005

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.9 - Incident Duration by Incident Type

Duration Intervals			Incident Type																Total	% for Duration Intervals		
Hourly	**CNAM***	**15-Min**	**Property Damage**		**Disabled Vehicle**		**Disabled Truck**		**Road HAZARD**		**Personal Injuries**		**Vehicle Fire**		**HAZMAT**		**Weather Related**		**Total**	**15-Min**	**CNAM**	**Hourly**
1 Hour	**Short**	00-14	196	24%	272	39%	38	21%	2	3%	7	10%	4	14%	5	17%	1	10%	525	27.5%	27.5%	78%
	Medium	15-29	182	22%	190	27%	46	26%	6	8%	9	13%	2	7%	3	10%			438	23.0%	23.0%	
	Long	30-44	149	18%	95	14%	29	16%	6	8%	17	25%	7	24%	3	10%			306	16.0%	49.5%	
		45-59	108	13%	63	9%	23	13%	7	9%	11	16%	4	14%	1	3%			217	11.4%		
2 Hours		60-74	71	9%	34	5%	15	8%	4	5%	11	16%	3	10%	3	10%	1	10%	142	7.4%		16%
		75-89	40	5%	23	3%	12	7%	4	5%	6	9%	3	10%	5	17%	1	10%	94	4.9%		
		90-104	21	3%	8	1%	3	2%	3	4%	4	6%	2	7%	3	10%			44	2.3%		
		105-119	11	1%	3	0%	5	3%	2	3%	1	1%			1	3%			23	1.2%		
3 Hours		120-134	7	1%	2	0%			3	4%					1	3%			13	0.7%		2%
		135-149	6	1%			1	1%	3	4%	1	1%	1	3%					12	0.6%		
		150-164	9	1%	1	0%			1	1%					1	3%	1	10%	13	0.7%		
		165-179			1	0%	2	1%	1	1%					1	3%			5	0.3%		
4 Hours		180-194	2	0%					3	4%			1	3%					6	0.3%		1%
		195-209	2	0%					4	5%					2	7%	1	10%	9	0.5%		
		210-224	2	0%					1	1%									3	0.2%		
		225-239	4	0%			1	1%	1	1%									6	0.3%		
5 Hours		240-254	1	0%			1	1%	4	5%			1	3%					7	0.4%		1%
		255-269	2	0%					4	5%									6	0.3%		
		270-284							5	6%									5	0.3%		
		285-299							2	3%									2	0.1%		
6 Hours +		300+	6	1%	2	0%	3	2%	14	18%	0	0%	1	3%	0	0%	5	50%	31	1.6%		2%
Total			819	100%	694	100%	179	100%	80	100%	67	100%	29	100%	29	100%	10	100%	1907	100%	100%	100%
Average Duration (min)			**44.1**		**27.8**		**51.6**		**174.7**		**48.8**		**71.7**		**73.2**		**288.9**		**46.6**			
25 Percentile			15.0		9.0		16.5		53.5		30.0		34.0		26.0		97.3					
50 Percentile			32.0		20.0		33.0		151.5		46.0		56.0		67.0		272.5					
75 Percentile			56.0		38.0		59.0		274.5		65.5		76.0		99.0		350.8					

* CNAM definition of intervals

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.10 - Incident Duration Distribution by Lane Blockage

Duration Intervals			Lane Blockage											% for Duration Intervals		
Hourly	CNAM*	15-Min	1 Lane		2 Lanes		3 or More Lanes		No Travel Lane		N/A		Total	15-Min	CNAM	Hourly
1 Hour	Short	00-14	454	31%	41	19%	10	20%	16	15%	4	8%	525	27.5%	27.5%	77.9%
	Medium	15-29	357	24%	38	17%	17	34%	17	16%	9	18%	438	23.0%	23.0%	
	Long	30-44	234	16%	43	20%	9	18%	18	17%	2	4%	306	16.0%	49.5%	
		45-59	160	11%	31	14%	6	12%	18	17%	2	4%	217	11.4%		
2 Hours		60-74	97	7%	25	11%	4	8%	10	10%	6	12%	142	7.4%		15.9%
		75-89	69	5%	11	5%	1	2%	8	8%	5	10%	94	4.9%		
		90-104	27	2%	10	5%			6	6%	1	2%	44	2.3%		
		105-119	12	1%	6	3%	1	2%	3	3%	1	2%	23	1.2%		
3 Hours		120-134	10	1%	1	0%	1	2%	1	1%	1	2%	14	0.7%		2.3%
		135-149	7	0%	4	2%							11	0.6%		
		150-164	9	1%	1	0%			3	3%			13	0.7%		
		165-179	4	0%	1	0%							5	0.3%		
4 Hours		180-194	3	0%					2	2%	1	2%	6	0.3%		1.3%
		195-209	5	0%	2	1%					2	4%	9	0.5%		
		210-224	2	0%	1	0%							3	0.2%		
		225-239	3	0%			1	2%			2	4%	6	0.3%		
5 Hours		240-254	4	0%					1	1%	2	4%	7	0.4%		1.0%
		255-269	4	0%	2	1%							6	0.3%		
		270-284	5	0%									5	0.3%		
		285-299	2	0%									2	0.1%		
6 Hours +		300+	14	1%	3	1%	0	0%	1	1%	13	25%	31	1.6%		1.6%
Total			1482	100%	220	100%	50	100%	104	100%	51	100%	1907	100%	100%	100%
Average Duration (min)			40.7		54.0		38.5		57.6		185.1					
25 Percentile			12.0		21.8		18.0		23.0		32.3					
50 Percentile			25.0		40.0		27.5		46.0		92.5					
75 Percentile			51.0		67.0		50.0		71.5		312.3					

Source: TRANSCOM Incident Data from February 1st, 2004 through March 31st, 2005

* CNAM definition of intervals

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.11 - Incident Duration by Lane Blockage for each Incident Type

Incident Type	Lane Blockage	Sample Size	Duration Statistics: 25 Percentile	50 Percentile	75 Percentile	Average	Standard Deviation
Property Damage	1 lane	594	14	31	55	**42.2**	47.5
	2 lanes	139	15	33	56	**43.3**	43.1
	3 or more lanes	29	18	27	50	**40.3**	45.7
	no travel lane	42	28	45	71	**55.4**	43.5
	Not available	15	22	61	92	**100.7**	117.7
Disabled Vehicle	1 lane	625	9	19	35	**25.8**	25.4
	2 lanes	24	22	35	73	**49.4**	4.3
	3 or more lanes	7	17	33	50	**43.4**	43.4
	no travel lane	33	17	36	56	**40.2**	26.5
	Not available	5	12	16	24	**76.0**	140.0
Disabled Truck	1 lane	151	16	29	53	**42.9**	72.4
	2 lanes	8	40	64	100	**130.1**	190.0
	3 or more lanes	7	25	34	45	**35.3**	23.1
	no travel lane	12	25	62	99	**103.5**	131.7
	Not available	1	-	-	-	**231.0**	
Road Hazard	1 lane	51	51	194	275	**172.9**	112.7
	2 lanes	11	64	95	144	**121.8**	86.9
	3 or more lanes	1	-	-	-	**15.0**	
	no travel lane	3	59	72	126	**99.3**	71.1
	Not available	14	94	221	323	**250.5**	233.0
Personal Injury	1 lane	30	20	39	60	**43.5**	30.4
	2 lanes	23	35	48	73	**53.7**	23.0
	3 or more lanes	1	-	-	-	**26.0**	
	no travel lane	11	28	53	79	**55.4**	33.2
	Not available	2	40	49	57	**48.5**	23.3
HAZMAT	1 lane	23	38	80	106	**80.1**	60.1
	2 lanes	4	49	64	75	**60.0**	34.7
	3 or more lanes	-	-	-	-		
	no travel lane	2	18	21	23	**20.5**	7.8
	Not available	-	-	-	-		
Vehicle Fire	1 lane	9	38	56	76	**56.4**	28.1
	2 lanes	11	35	42	69	**73.9**	88.2
	3 or more lanes	5	19	22	56	**32.8**	27.2
	no travel lane	2	69	80	92	**80.0**	32.5
	Not available	2	206	218	229	**217.5**	33.2
Weather Related	1 lane	1	-	-	-	**375.0**	
	2 lanes	-	-	-	-		
	3 or more lanes	-	-	-	-		
	no travel lane	2	40	80	119	**79.5**	111.0
	Not available	7	141	340	348	**336.4**	19.9
	All Incidents	**1907**	**13**	**29**	**55**	**46.6**	**67.4**

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.12 - Incident Distribution by Day of Week

			Incidents		Total	
Day Type	**Day Type %**	**Day**	**#**	**%**	**#**	**%**
Weekday	71.4%	Monday	282	14.8%	1431	75.0%
		Tuesday	299	15.7%		
		Wednesday	303	15.9%		
		Thursday	270	14.2%		
		Friday	277	14.5%		
Weekend	28.6%	Saturday	231	12.1%	476	25.0%
		Sunday	245	12.8%		
Total			1907	100%	1907	100%

Source: TRANSCOM Incident Data from February 1st, 2004 through March 31st, 2005

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.13 - Incident Type Distribution by Month

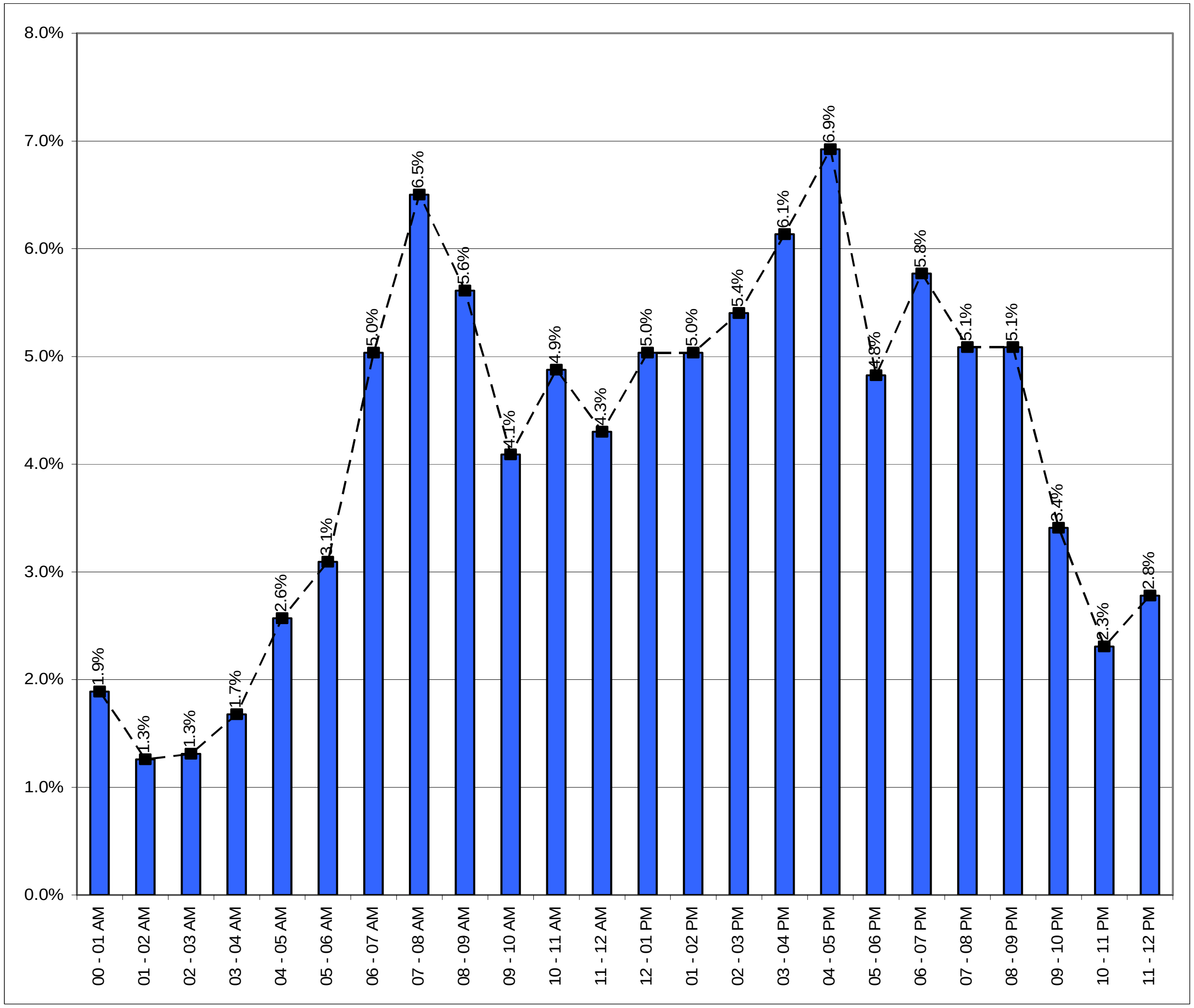
Year	Month	Incident Type: Property Damage	Disabled Vehicle	Disabled Truck	Road HAZARD	Personal Injuries	HAZMAT	Vehicle Fire	Weather Related	Total	%
2004	**February**	45	23	7	2	2	-	1	-	80	4.2%
	March	43	37	22	2	7	-	-	-	111	5.8%
	April	65	33	17	3	3	1	1	-	123	6.4%
	May	57	48	14	3	4	1	2	-	129	6.8%
	June	68	47	9	1	7	2	2	-	136	7.1%
	July	55	58	14	4	7	2	3	1	144	7.6%
	August	48	60	5	-	4	6	2	2	127	6.7%
	September	57	62	8	5	8	2	3	1	146	7.7%
	October	65	42	19	4	5	1	1	1	138	7.2%
	November	67	54	14	12	3	3	1	-	154	8.1%
	December	64	48	16	2	5	3	5	-	143	7.5%
2005	**January**	84	69	18	13	4	-	2	4	194	10.2%
	February	46	55	6	10	3	2	2	-	124	6.5%
	March	55	58	10	19	5	6	4	1	158	8.3%
	Total	819	694	179	80	67	29	29	10	**1907**	100%
									Monthly Average	**136.2**	7.1%

Source: TRANSCOM Incident Data from February 1st, 2004 through March 31st, 2005

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.14 - Incident Distribution by Time of Day

Time of Day	Total	%
00 - 01 AM	36	1.9%
01 - 02 AM	24	1.3%
02 - 03 AM	25	1.3%
03 - 04 AM	32	1.7%
04 - 05 AM	49	2.6%
05 - 06 AM	59	3.1%
06 - 07 AM	96	5.0%
07 - 08 AM	124	6.5%
08 - 09 AM	107	5.6%
09 - 10 AM	78	4.1%
10 - 11 AM	93	4.9%
11 - 12 AM	82	4.3%
12 - 01 PM	96	5.0%
01 - 02 PM	96	5.0%
02 - 03 PM	103	5.4%
03 - 04 PM	117	6.1%
04 - 05 PM	132	6.9%
05 - 06 PM	92	4.8%
06 - 07 PM	110	5.8%
07 - 08 PM	97	5.1%
08 - 09 PM	97	5.1%
09 - 10 PM	65	3.4%
10 - 11 PM	44	2.3%
11 - 12 PM	53	2.8%
Total	1907	100.0%



Source: TRANSCOM Incident Data from February 1st, 2004 through March 31st, 2005

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.15 - Incident Distribution by Time Period

Period	Time Period*	Duration (hours)	Total	% per Period	% per hour
Peak	06AM-10AM	4	405	21.2%	5.3%
	03PM-07PM	4	451	23.6%	5.9%
	Subtotal	8	856	44.9%	5.6%
Off Peak	12AM-06AM	6	225	11.8%	2.0%
	10AM-03PM	5	470	24.6%	4.9%
	07PM-12AM	5	356	18.7%	3.7%
	Subtotal	16	1051	55.1%	3.4%
Total		24	1907	100%	4.2%

Source: TRANSCOM Incident Data from February 1st, 2004 through March 31st, 2005

* Time periods defined in CNAM

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.16 - Incident Frequency by Type of Incident and Time of Day

Time of Day	Personal Property			Disabled Vehicle			Disabled Truck			Road HAZARD			Personal Injuries			HAZMAT			Vehicle Fire			Weather Related			Total
00 - 01 AM	21	1.1%	2.6%	5	0.3%	0.7%	1	0.1%	0.6%	2	0.1%	2.5%	6	0.3%	9.0%			0.0%	1	0.1%	3.4%				36
01 - 02 AM	15	0.8%	1.8%	7	0.4%	1.0%	1	0.1%	0.6%				1	0.1%	1.5%										24
02 - 03 AM	7	0.4%	0.9%	9	0.5%	1.3%	2	0.1%	1.1%				4	0.2%	6.0%				1	0.1%	3.4%	2	0.1%	20.0%	25
03 - 04 AM	16	0.8%	2.0%	7	0.4%	1.0%	5	0.3%	2.8%	2	0.1%	2.5%	1	0.1%	1.5%							1	0.1%	10.0%	32
04 - 05 AM	28	1.5%	3.4%	12	0.6%	1.7%	5	0.3%	2.8%				2	0.1%	3.0%	1	0.1%	3.4%				1	0.1%	10.0%	49
05 - 06 AM	30	1.6%	3.7%	18	0.9%	2.6%	4	0.2%	2.2%	1	0.1%	1.3%	2	0.1%	3.0%	2	0.1%	6.9%	2	0.1%	6.9%				59
06 - 07 AM	45	2.4%	5.5%	34	1.8%	4.9%	9	0.5%	5.0%	1	0.1%	1.3%	3	0.2%	4.5%	1	0.1%	3.4%	3	0.2%	10.3%				96
07 - 08 AM	54	2.8%	6.6%	46	2.4%	6.6%	12	0.6%	6.7%	2	0.1%	2.5%	4	0.2%	6.0%	6	0.3%	20.7%							124
08 - 09 AM	47	2.5%	5.7%	27	1.4%	3.9%	20	1.0%	11.2%	7	0.4%	8.8%	2	0.1%	3.0%	2	0.1%	6.9%	2	0.1%	6.9%				107
09 - 10 AM	33	1.7%	4.0%	24	1.3%	3.5%	9	0.5%	5.0%	6	0.3%	7.5%	1	0.1%	1.5%	2	0.1%	6.9%	3	0.2%	10.3%				78
10 - 11 AM	27	1.4%	3.3%	31	1.6%	4.5%	8	0.4%	4.5%	22	1.2%	27.5%	2	0.1%	3.0%	3	0.2%	10.3%							93
11 - 12 AM	34	1.8%	4.2%	29	1.5%	4.2%	9	0.5%	5.0%	7	0.4%	8.8%	2	0.1%	3.0%				1	0.1%	3.4%				82
12 - 01 PM	37	1.9%	4.5%	31	1.6%	4.5%	10	0.5%	5.6%	6	0.3%	7.5%	6	0.3%	9.0%	3	0.2%	10.3%	3	0.2%	10.3%				96
01 - 02 PM	43	2.3%	5.3%	34	1.8%	4.9%	8	0.4%	4.5%	6	0.3%	7.5%	3	0.2%	4.5%	1	0.1%	3.4%	1	0.1%	3.4%				96
02 - 03 PM	52	2.7%	6.3%	30	1.6%	4.3%	9	0.5%	5.0%	3	0.2%	3.8%	5	0.3%	7.5%	2	0.1%	6.9%				2	0.1%	20.0%	103
03 - 04 PM	54	2.8%	6.6%	49	2.6%	7.1%	5	0.3%	2.8%	4	0.2%	5.0%	1	0.1%	1.5%	1	0.1%	3.4%	2	0.1%	6.9%	1	0.1%	10.0%	117
04 - 05 PM	61	3.2%	7.4%	47	2.5%	6.8%	14	0.7%	7.8%	1	0.1%	1.3%	6	0.3%	9.0%				3	0.2%	10.3%				132
05 - 06 PM	33	1.7%	4.0%	41	2.1%	5.9%	11	0.6%	6.1%	3	0.2%	3.8%	1	0.1%	1.5%				2	0.1%	6.9%	1	0.1%	10.0%	92
06 - 07 PM	39	2.0%	4.8%	51	2.7%	7.3%	13	0.7%	7.3%	1	0.1%	1.3%	3	0.2%	4.5%	1	0.1%	3.4%	1	0.1%	3.4%	1	0.1%	10.0%	110
07 - 08 PM	38	2.0%	4.6%	44	2.3%	6.3%	9	0.5%	5.0%	1	0.1%	1.3%	2	0.1%	3.0%	1	0.1%	3.4%	1	0.1%	3.4%	1	0.1%	10.0%	97
08 - 09 PM	31	1.6%	3.8%	53	2.8%	7.6%	7	0.4%	3.9%	3	0.2%	3.8%	2	0.1%	3.0%	1	0.1%	3.4%							97
09 - 10 PM	30	1.6%	3.7%	25	1.3%	3.6%	2	0.1%	1.1%	1	0.1%	1.3%	5	0.3%	7.5%	1	0.1%	3.4%	1	0.1%	3.4%				65
10 - 11 PM	18	0.9%	2.2%	22	1.2%	3.2%	2	0.1%	1.1%	1	0.1%	1.3%	1	0.1%	1.5%										44
11 - 12 PM	26	1.4%	3.2%	18	0.9%	2.6%	4	0.2%	2.2%				2	0.1%	3.0%	1	0.1%	3.4%	2	0.1%	6.9%				53
Total	819	42.9%	100%	694	36.4%	100%	179	9.4%	100%	80	4.2%	100%	67	3.5%	100%	29	1.5%	100%	29	1.5%	100%	10	0.5%	100%	1907

Source: TRANSCOM Incident Data from February 1st, 2004 through March 31st, 2005

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.17 - Incident Type Distribution by Weather Conditions

	Weather Conditions*																
	Normal			Adverse													
Incident Type	**Clear**	**Cloudy**	Total (Normal)	**Rain**	**Light Rain**	**Heavy Rain**	**Snow**	**Foggy**	**Flurries**	**Sleet**	**Blizzard**	**Overcast**	**High Wind**	Total (Adverse)	**Subtotal**	**N/A****	**Total**
Property Damage	397	51	448	24	19	5	6	10	7	2	4	2	2	81	529	290	819
Disabled Vehicle	351	48	399	24	17	5	7	1	2	4	4	3	1	68	467	227	694
Disabled Truck	92	21	113	1	8		2	2	1			2		16	129	50	179
Road HAZARD	32	7	39					1	1	2				4	43	37	80
Personal Injuries	28	5	33	3	2			1						6	39	28	67
HAZMAT	15	2	17			1								1	18	11	29
Vehicle Fire	12	3	15	1	2					1	1			5	20	9	29
Weather Related					2	4								6	6	4	10
Total	927	137	1064	53	50	15	15	15	11	9	9	7	3	187	1251	656	1907
%	74.1%	11.0%	85.1%	4.2%	4.0%	1.2%	1.2%	1.2%	0.9%	0.7%	0.7%	0.6%	0.2%	14.9%	100%		

Source: TRANSCOM Incident Data from February 1st, 2004 through March 31st, 2005

* Weather conditions definition in CNAM

** Weather information were missing in the database

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Table 6.18 - Incident Type Distribution by Pavement Conditions

	Pavement Conditions*											
	Normal	Adverse										
Incident Type	**Dry**	**Wet**	**Snow covered**	**Flooding**	**Icy**	**Slick**	**Slushy**	**Black Ice**	**Total (Adverse)**	**Subtotal**	**N/A****	**Total**
Property Damage	431	83	5	1	4	1		1	95	526	293	819
Disabled Vehicle	381	75	4	4	1		1		85	466	228	694
Disabled Truck	106	21		1					22	128	51	179
Road HAZARD	36	6				1			7	43	37	80
Personal Injuries	32	7							7	39	28	67
HAZMAT	16	2							2	18	11	29
Vehicle Fire	15	5							5	20	9	29
Weather Related		5		1					6	6	4	10
Total	1017	204	9	7	5	2	1	1	229	1246	661	1907
%	81.6%	16.4%	0.7%	0.6%	0.4%	0.2%	0.1%	0.1%	18.4%	100%		

Source: TRANSCOM Incident Data from February 1st, 2004 through March 31st, 2005

* Pavement conditions definition in CNAM

** Pavement information were missing in the database

Note: All frequency results are considerably smaller than the real frequencies mainly due to the fact that TRANSCOM data contained only 8% of all accident in the study area.

Appendix 1, Different terms for event description in TRANSCOM database

Element_Value	Element_Text
0	Unknown Other
1	Known Other
2	Stalled vehicle
3	Vehicle fire/explosion
4	Roadway non-hazmat spill
5	Hazmat spill
6	Transit related incident
7	Overweight vehicle
8	Earthquake
9	Landslide
10	Flooding
11	Tornado
12	Hurricane
13	Unplanned demonstration
14	Rollover/overturn/jackknife
15	Non-vehicle fire/explosion
16	Pothole repairs
17	Pedestrian accident
18	Plowing and salting
19	Road sweeping
20	Late running construction
21	Gridlock alert day
22	Thru street program
23	Rubbernecking delays
24	Sinkhole
25	Security Check Point
26	Security related
27	Earlier Incident
28	Blackout
29	Malfunctioning traffic light
30	VIP visit
31	Parking related
32	Sewer collapse
33	Steam leak
34	Icicle removal
35	No water available
36	No food available
37	No diesel available
38	No fuel available
39	Motorcycle rally
40	Sewer main break
41	Single occupancy vehicle rules
42	Falling ice
43	Truck restrictions
44	Funeral procession
45	Accident
46	Overturned truck
47	Overturned tractor trailer
48	Overturned vehicle
49	Accident investigation
50	Jack-knifed tractor trailer
51	Split tractor trailer
52	Separated tractor trailer
53	Misplaced tractor trailer
54	Overheight tractor trailer
55	Truck fire
56	Bus fire
57	Tractor trailer fire
58	Vehicle fire
59	Brush fire

Element_Value	Element_Text
60	Building fire
61	Fire department activity
62	Police department activity
63	EMS activity
64	Disabled truck
65	Disabled tractor trailer
66	Disabled vehicle
67	Disabled bus
68	Emergency construction
69	Roadwork
70	Roving repairs
71	Operational Activity
72	Icing
73	Fog
74	Snow removal
75	Wet pavement
76	High winds
77	Weather related
78	Downed tree
79	Downed wires
80	Downed pole
81	Nearby building collapse
82	Debris spill
83	Cargo spill
84	Fuel spill
85	Watermain break
86	Capacity related
87	Heavy traffic
88	Delays
89	Power problems
90	Signal problem
91	Amber alert
92	Stuck gates
93	Drawbridge open
94	Transformer fire
95	Gas main break
96	Job action
97	Demonstration
98	New traffic Pattern
99	Sun glare
100	Special event
101	Construction
102	Accident with Injuries
103	Accident with Property Damage Only
104	Accident Road Closed
105	Spinout
106	Shifted Plates
107	Road Collapse
108	Speed restriction
109	Rough road
110	Pockets of Volume
111	Missing Manhole
112	HOV rules
113	Collapsed Manhole
114	Collapsed Sewer Gate
115	Missing Sewer Grate
116	Ozone Alert
117	Test Message
118	Collapsed Scaffolding
119	Falling Debris

Appendix 2, Non-recurring incidents definitions extracted from the TRANSCOM database

Incident Type	Event description in the database selcted as non-recuring Incident
Property Damage	Accident
	Accident & Accident Investigation
	Accident & Debris Spill
	Accident & Delays
	Accident & Delays
	Accident & Disabled Truck
	Accident & Disabled Vehicle
	Accident & Fire Department Activity & Delays
	Accident & Fuel Spill
	Accident & Heavy Traffic
	Accident & Jack-Knifed Tractor Trailer
	Accident & Overturned Vehicle
	Accident & Police Department Activity
	Accident & Pothole Repairs
	Accident & Vehcile Fire & Delays
	Accident & Vehicle fire
	Accident Investigation
	Accident Road Closed
	Accident with Property Damage Only
	Accident with Property Damage Only & Delays
	Accident with Property Damage Only & Fire Department Activity
Road HAZARD	Downed Pole
	Icicle Removal
	Missing Manhole
	Missing Sewer Grate
	Police Department Activity
	Police Department Activity & Delays
	Pothole Repairs
	Pothole Repairs & Delays
	Pothole Repairs & Roving Repairs
	Road Sweeping
	Roadwork
	Roadwork & Delays
	Roving Repairs
	Sinkhole & Emergency Construction
	Watermain Break
Weather Related	Flooding
	Icing

Incident Type	Event description in the database selcted as non-recuring Incident
Disabled Vehicle	Disabled Vehicle
	Disabled Vehicle & Delays
	Disabled Vehicle & Heavy Traffic
	Disabled Vehicle & Other
	Disabled Vehicle & Road Collapse
	Disabled Vehicle & Roadwork
	Overturned Vehcile
	Overturned Vehcile & Delays
	Overturned Vehicle
	Overturned Vehicle & Accident with Injuries &Heavy Traffic
	Overturned Vehicle & Delays
Disabled Truck	Disabled Bus
	Disabled Tractor Trailer
	Disabled Truck
	Disabled Truck & Delays
	Jack-Knifed Tractor Trailer
	Jack-Knifed Tractor Trailer & Fuel Spill
	Misplaced Tractor Trailer
	Overturned Tractor Trailer
	Overturned Tractor Trailer & Fuel Spill
	Overturned Truck
	Overturned Truck & Delays
Vehicle Fire	Brush Fire
	Building Fire
	Fire Department Activity
	Fire Department Activity & Police Department Activity
	Tractor Trailer Fire
	Transformer Fire
	Truck Fire
	Vehicle Fire
	Vehicle Fire & Delays
Personal Injuries	Accident with Injuries
	Accident with Injuries & Accident
	Accident with Injuries & Debris Spill
	Accident with Injuries & Delays
	Accident with Injuries & Heavy Traffic
	Accident with Injuries & Overturned Vehicle
HAZMAT	Debris Spill
	Fuel Spill

Appendix 3, Terms being used in TRANSCOM database to proximate location

Article_Display	Article
approaching	1
area of	2
at	3
between	4
east of	5
from	6
near	7
north of	8
ramp from	9
ramp to	10
south of	11
to	12
west of	13
before	14
by	16
just past	17
through	18
under	19
over	20
into/out of	21
ramps from	22
ramps to	23
entering at	24
exiting at	25
into	26
out of	27
bypassing	28
in	29
in the vicinity of	30
ramp to/from	31
ramps to/from	32

Task 3: Data Collection and Analysis
SUBTASK 3.1 (Additional Scope)
Measuring the Delay Impacts of Traffic Incidents from TRANSMIT Data

Prepared by:

Ali Afshar, Research Associate
John C. Falcocchio, Ph.D., P.E.
Urban ITS Center
Polytechnic University
Brooklyn, NY 11201

Tel: 718.260.3612
Fax: 718.260.3433
Email: jfalcocc@poly.edu

June 20th, 2007

1.0 Background

The purpose of this memorandum is to quantify the impact of a traffic incident on delay using TRANSCOM and TRANSMIT data. This is accomplished by combining the TRANSCOM incident dataset with the corresponding periods in the TRANSMIT database, thereby enabling the measurement of delay experienced by an average vehicle due to an incident of a given duration. The delay is determined for traffic directly impacted by the incident, and for traffic going in the opposite direction.

2.0 Data Sources

In this section the TRANSCOM and TRANSMIT data sources and data field utilized are identified and described. Linkage between data sources and limitation of the existing dataset are presented and the sample selection method is described.

2.1 Incident Data

The incident database was obtained from TRANSCOM for the study period (February 1st 2004 to March 31st 2005), as reported in subtask 3.1. These data provided relevant information about an incident such as location, incident type, created time, closed time, and lane blockages.

2.1.1 Incident Data Fields

- Incident Location: provides facility name, direction of travel, and proximity of the incident to the nearest ramp or crossing roadway or land mark.

- Incident Type: incidents in the database were recorded as events and included recurring and non-recurring incidents. Non-recurring incidents were consolidated into eight tiers as "incident type" which includes: property damage, personal injuries, disabled vehicle, disable truck, vehicle fire, road hazard, hazmat, and weather related. These eight categories were used in subtask 3.1. For the purpose of the impact analysis, the sampled incidents were aggregated into three simpler tiers as: accident, disabled vehicle, and non-vehicle related incident.

- Created Time: represents the time of incident detection and/or notification (by any source).

- Closed Time: represents the time when the incident has been cleared. This means that all responding personnel and equipments have been removed from the scene.

- Lane Blockage: refers to the lane(s) impacted by the incident. The lane blockage data, however, does not specify the length of time the lane(s) were blocked.

2.1.2 Incident Duration

Incident duration is obtained by subtracting “closed time” (the time when the incident is cleared by all vehicles and personnel) from “create time” (the time when the incident was first reported). The actual time of the incident, however, is typically earlier than that reported.

2.2 Link Travel Time and Speed Data

The TRANSMIT database consists of archived link travel time and speed for the various sections of the corridor. These data, collected over a period from February 1st, 2004 to March 31st, 2005, are summarized for each 15-minutes period from the E-ZPass readers located at the beginning and end of each link.

The TRANSMIT system receives signals from the readers and stores data for each link. Link information is being continuously collected by overhead antennas and conveyed to the computer server to record real time observation. The “Validation Travel Time Study of the TRANSMIT System” (2006) for New York State DOT, concluded that the system performs well for the TRANSMIT links where travel time data are available and captured.

Overhead antennas (transponder readers/TRANSMIT readers/E-ZPass antenna) have the capability to read transponder signals for successive readers and compute travel time between the two points of references. The roadside antennas capture dynamic vehicular travel attributes by detecting vehicles with E-ZPass transponders. The system with predefined algorithm calculates the travel time and speed between the reference points. The numbers of detected vehicles are labeled as volume in the database but are not the actual link traffic volume.

2.2.1 Data Coverage and Constraints

The TRANSMIT system is designed to monitor the average travel time and speed. These data are archived in 15 minutes intervals for every link in the coverage area. After extracting the TRANSMIT data for the study corridor into Excel, gaps and missing data were identified. The missing data records were as short as one interval (15 minutes) and as long as three weeks and in some cases for the duration of the study period.

2.3 Combining Incident (TRANSCOM) and Speed Data (TRANSMIT)

Because there is no common reference in the TRANSCON and TRANSMIT datasets, a common field had to be created to identify incident data common to both datasets.

The “Link ID” in TRANSMIT dataset identifies a section of the highway between two consecutive antennas. This was the only identifier that could be associated with the location of the incident in the TRANSCOM dataset. The “Link ID” for each incident was manually entered in the TRANSCOM database according to roadway direction as well as the nearest ramps or crossing roadway.

2.4 Sample Selection Criteria

All TRANSCOM incidents were plotted on GIS as point feature and superimposed on the roadway network. This activity helped in visualizing the incident distribution along the study corridor. The following criteria were used in selecting the study sample: (1) the highway section should have at least 30 incidents; (2) the incident duration should be greater than 30 minutes; (3) the highway section should have working TRANSMIT antennas; (4) links on both directions should be represented in the sample.

Table 1 shows the sample selection summary and Table 2 identifies the links and the 430 incidents in the study sample.

Figure 1 show the links listed in Table 2. It depicts the consecutive links with their link ID and direction, the location of the antennas, and the distribution of the incident in the corridor. This figure highlights the concentration of the incidents at merge or diverge areas close to on/off ramps. Depicted incident location in Figure 1 shows the concentration of incidents in the vicinity of 9th Street (Gowanus Expressway & BQE merge and diverge), Atlantic Avenue on/off ramps, BQE at Brooklyn and Manhattan bridges, BQE and Exit 33 (Leonard St and Mc Guinness Blvd), and on Kosciuszko Bridge.

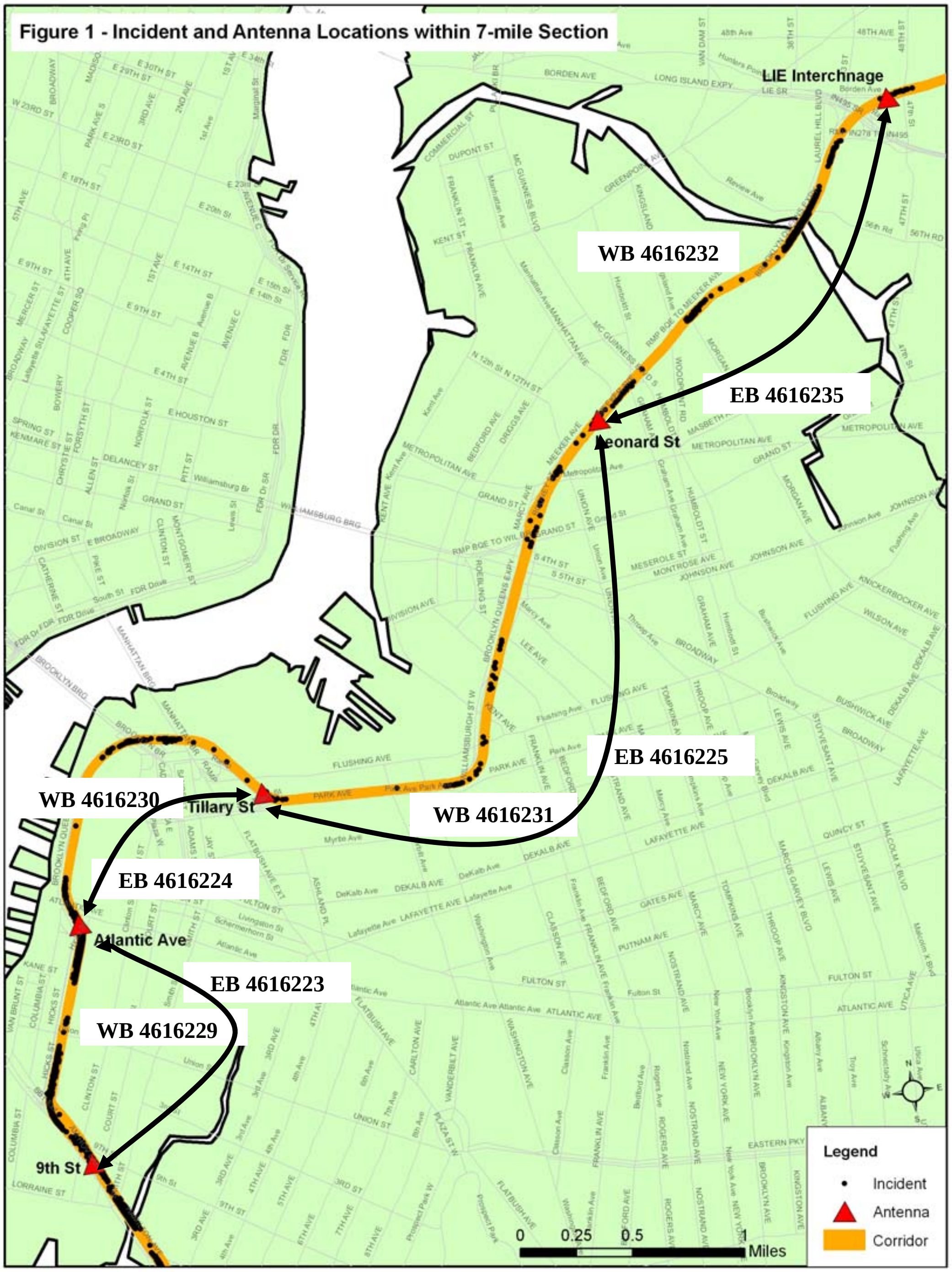
Table 1 - Sample Selection Summary

Description	# of Link	# of Incident
All Incidents	49 + *	1907
Highway Links with TRANSMIT Coverage	49	1810
Links with working TRANSMIT and Recorded Incidents	40	1317
Links with 30 or more Incidents	8	1056
Incidents with Duration of 30 min or More	**8**	**430**

* Indicates highway sections without TRANSMIT Link coverage (Prospect and Sheridan Expressways)

Table 2 - Selected TRANSMIT Links for Analysis

Link ID	Dir	Link Name/Coverage	Length (mile)	Incidents with 30 min Duration or Longer
4616229	WB	Atlantic Ave to 9th St	1.4	86
4616230		Tillary St to Atlantic Ave	1.6	69
4616231		Leonard St to Tillary St	1.9	39
4616232		46th St to Leonard St	2.1	58
		Subtotal	***7.0***	***252***
4616223	EB	9th St to Atlantic Ave	1.4	73
4616224		Atlantic Ave to Tillary St	1.6	23
4616225		Tillary St to Leonard St	1.9	19
4616235		Leonard St to 46th St	2.1	63
		Subtotal	***7.0***	***178***
		Total	**14.0**	**430**



As it is shown in Table 2 and Figure 1 links have different length and using TRANSMIT link travel time without any conversion would be misleading. In order to simplify the process and bring all travel time into a common denominator, all TRANSMIT link travel times are divided by its length to have travel time rate (minute/mile) and apply that value for analysis purposes which is easy to compare and simple to convert into speed (mph).

2.5 Link Speed Patterns

Figure 2 and 3 depict the westbound and eastbound average link speed (mph) for each link under normal conditions (without incident).

The link travel speed calculated by TRANSMIT is the average speed of a 15 minute sample of vehicle link speeds. For an incident of a known type the resultant link average speed may vary with the location of the incident on the link. For example, if the incident occurs at the downstream end of the link, the impact of the incident on the link speed will be highest, but if it happens at the other end of the link, the speed reduction will be smaller. To address this issue, an average link speed was calculated from the sampled observations.

Figure 2 depicts links on the westbound direction. The top graph on the figure describes link 4616229, from Atlantic Avenue to 9th Street. The on/off ramps are located at the both ends of the link and there is no interruption in roadway characteristics along this link. For the off peak period (8:00 PM to 6:00AM) the average link speed is close to the posted 50 mph speed limits. However, the average speed on the link is worst from 8:00 to 11:30 AM which reaches to 15-20 mph and from 3:30 to 7:30 PM at about 20-25 mph for all weekdays. On the weekends speed fluctuate smoothly around the speed limits.

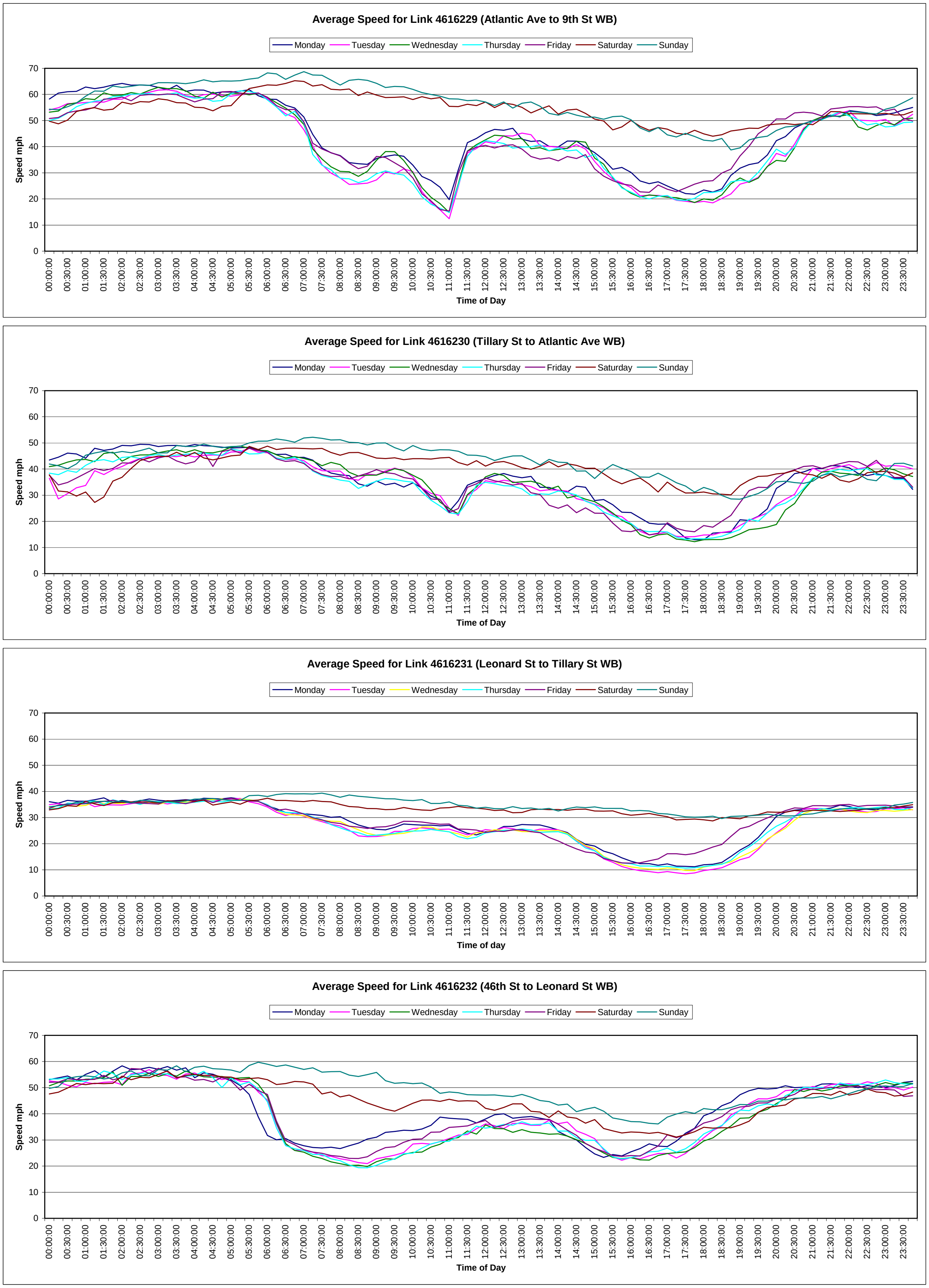
The second graph depicts link 4616230, from Tillary Street to Atlantic Avenue. The on/off ramps are located at the both ends of the link as well as on/off ramps for Brooklyn and Manhattan bridges between its two ends. The general speed profile is lower than the posted speed limits and lowest speed is experienced between 2:00 to 8:30 PM at 15-20 mph and 10:30 to 11:30 AM within 25-30 mph.

The third graph depicts link 4616231, from Leonard Street to Tillary Street with a major junction with Williamsburg Bridge as well as two on/off ramps connections. The posted speed limits is 50 mph however the average travel speed is lower under 40 mph and it reaches to 10 mph from 4:00 to 7:00 PM and reaches to 25 mph from 8:00 AM to 2:00 PM.

The bottom graph represents link 4616232, from 46th Street in Queens (Long Island Expressway/I-495 interchange with BQE) to Leonard Street in Brooklyn. It covers Kosciuszko Bridge crossing over Newtown Creek and two other ramp junctions. The average speed fluctuates in the course of the day and reaches to 20-30 mph from 6:30 to 11:00 AM and 2:30 to 6:00 PM.

Saturday and Sunday traffic speeds are higher than weekday traffic speeds from 5:00 AM to 8:30 PM for the following links: Atlantic Avenue to 9th Street; Tillary Street to Atlantic Avenue; and Leonard Street to Tillary Street; and from 5:30 AM to 6:00 PM for 46th Street in Queens to Leonard Street.

Figure 2 - Average Link Speed on Westbound Links



Source: TRANSMIT database from February 1st, 2004 to March 31th, 2005

Figure 3 depicts average link speed for the eastbound direction. While all on/off ramps and bridge junctions are the same as in the westbound direction, as described in Figure 2, on the eastbound direction average speed decreases as early as 6:30 AM and begins to increase to earlier values at 8:00 PM. Speed fluctuations in the eastbound direction in the course of the day is less pronounced compared to the westbound direction for all four links.

The top graph on this figure depicts link 4616223 from 9th Street to Atlantic Avenue. Link speed reaches 25-35 mph as early as 6:30 AM to 6:30 PM and general speed profile is much lower than the 50 mph posted speed limit.

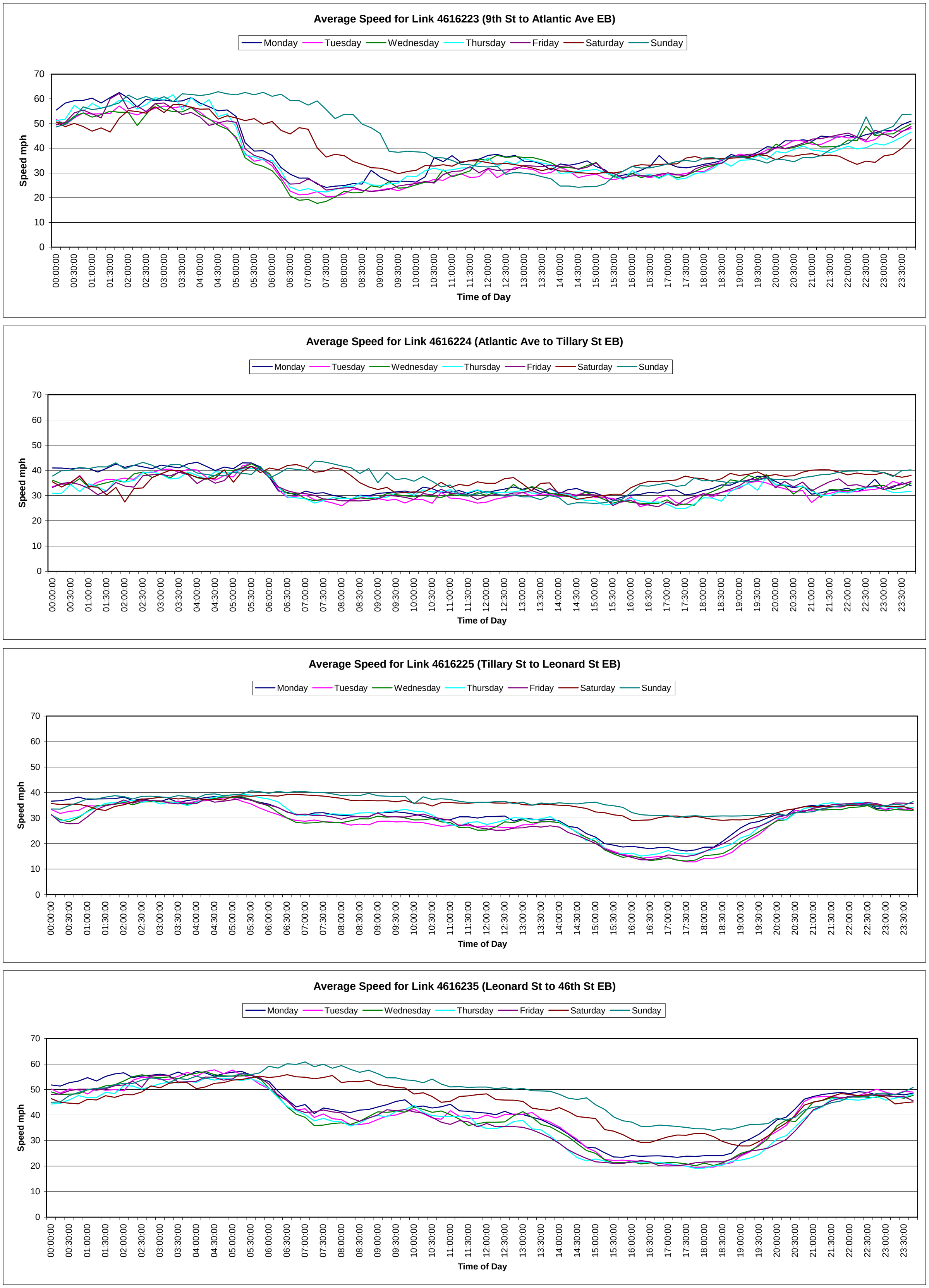
The second graph depicts link 4616224 from Atlantic Avenue to Tillary Street. The speed profile smoothly fluctuates between 30-40 mph for the course of the day.

The third graph shows link 4616225 from Tillary Street to Leonard Street. The average link speed is within 30-40 mph and it falls below 20 mph from 3:00 to 7:00 PM.

The bottom graph depicts link 4616235 from Leonard Street in Brooklyn to 46th Street in Queens. The free flow speed is around 50 mph and it falls to 40 mph from 6:30 AM to 1:00 PM and then drops further to 20 mph from 3:00 to 7:00 PM

Saturday and Sunday traffic speeds are somewhat higher than weekday traffic speed from 3:30 AM to 10:30 AM at the 9th Street to Atlantic Avenue; from 6:00 AM to 1:00 PM at the Atlantic Avenue to Tillary Street; and from 5:30 AM to 8:00 PM between Tillary Street and 46th Street in Queens.

Figure 3 - Average Link Speed on Eastbound Links



Source: TRANSMIT database from February 1st, 2004 to March 31th, 2005

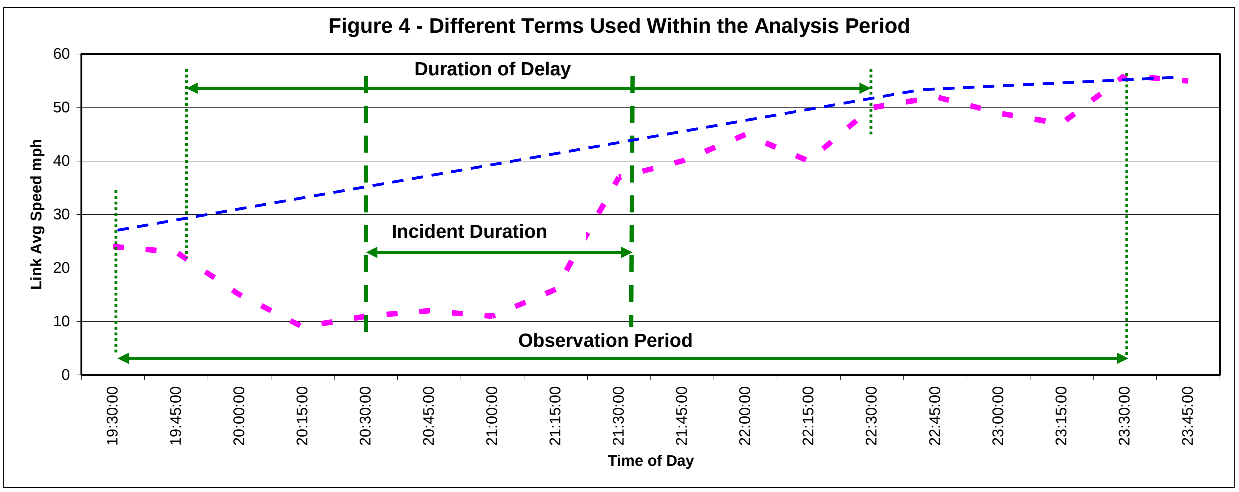
3.0 Method of Analysis

This section describes the method for estimating the delay impacts for individual incidents.

3.1 Terminology

Incident Duration, the difference between closed time and crated time based on TRANSCOM database,
Observation Period, from one hour before “created time” up to two hours after “closed time”,
Duration of Delay, the difference between graph resuming normal level after the closed time and curve drops from prevailing condition before incident created time
Delay Magnitude, is the difference between the incident curve and the without incident curve for the duration of delay,

Figure 4 illustrates how these terms are used in analysis of delay.



TRANSCOM data were combined in with the TRANSMIT database to measure the following performance criteria:

- Vehicle Delay Rate by Incident Type
- Vehicle Delay Rate by Lane Closures Characteristics
- Vehicle Delay Rate by Duration of Delay
- Vehicle Delay Rate by Pavement Conditions
- Vehicle Delay Rate by Weather Conditions

To account for the effect of traffic volume, each of these measures was summarized for an average 24 hour day, and for the average weekday PM peak period (3:00 to 7:00 PM).

Graphs similar to Figure 4 were plotted all incidents to identify the speed pattern with and without the incident for the reported incident duration, plus up to one hour before the incident was reported, and up to two hours after the incident was reported to have been cleared. The magnitude of delay created by an incident experienced by an average vehicle for the duration of delay (see Figure 4) is the difference between the average vehicle travel time rate (minute per mile) impacted by the incident, and the estimated average vehicle travel time rate without incident.

3.2 Illustrative Example

We have selected one incident to explain the assumptions made and the steps involved in the calculation of the various delay measures resulting from incidents occurring in the study sections.

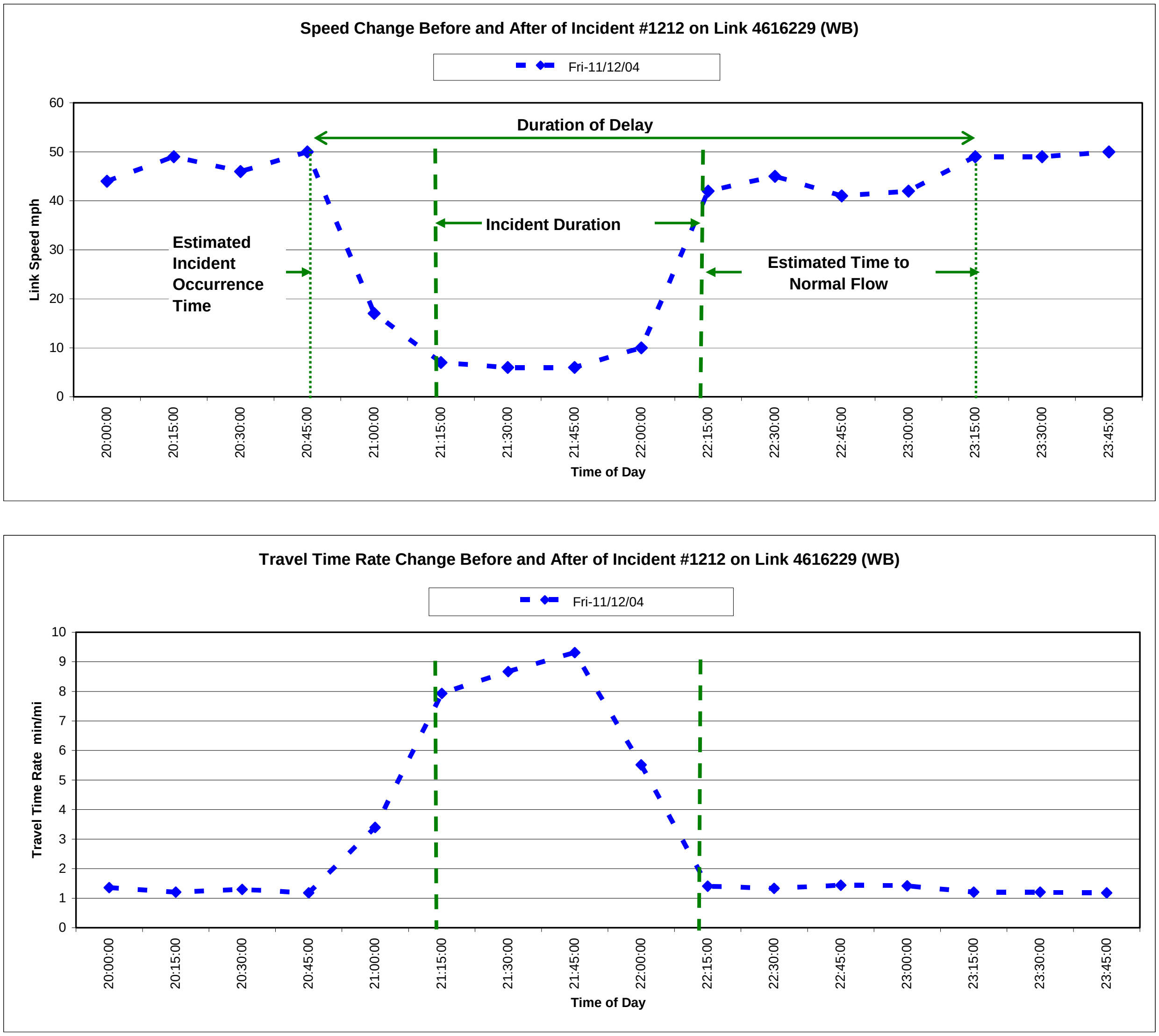
This example entails a property damage incident that was reported on Friday November 12, at 9:11 PM and cleared at 10:11 PM, on the westbound link (4616229) between Atlantic Avenue and 9th Street. This incident blocked one travel lane for an unknown length of time.

Figure 5 presents the travel speed (mph) and the travel time rate (min/mi) for traffic directly impacted by the incident. It should be noted that the prevailing traffic conditions before and after the incident are similar to those shown in Figure 2 for this link 4616229.

The incident created a drop in speed from approximately 45 mph to 6.5 mph (or an increase in the travel time rate from 1.3 min/mi, to 9.2 min/mi). While the incident was reported at 9:11, the speed profile indicates that this incident occurred at about 8:45 PM – approximately 30 minutes earlier. And while the reported duration of the incident was one hour [(10:11) – (9:15)], the time period of traffic delay was much longer (from 8:45 PM to 11:15 PM). After the incident was cleared, it took almost one hour for traffic to return to normal.

The average travel time rate over the entire period of traffic delay was calculated by taking the average of the delay rates reported in Figure 5, over the ten 15-minute periods (from 8:45 to 11:15 PM). Thus the average travel time rate during this three-hour period of congestion was 3.18 vehicle minutes per mile. And when compared with the travel rate before the incident of 1.24 vehicle minutes per mile, the vehicle delay due to the incident is calculated as 1.94 vehicle miles (3.18 – 1.24). Or the incident increases the average travel time rate by 156%.

Figure 5 - Speed and Travel Time Rate



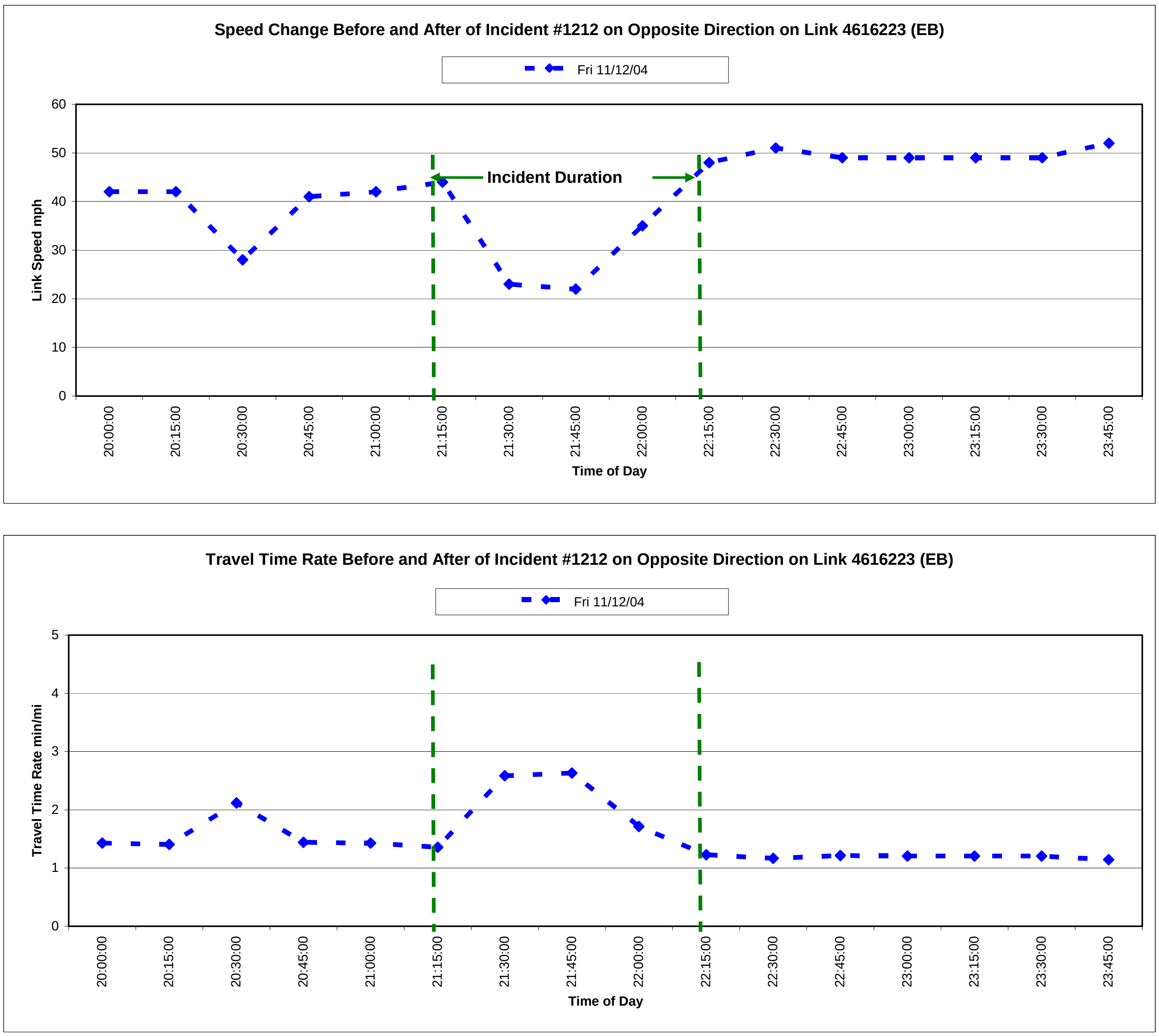
Source: TRANSCOM and TRANSMIT database

3.3 Impacts of Incident on the Speed of Opposing Traffic

Traffic speed in the opposite direction was also extracted from the TRANSMIT database. Figure 6 shows the traffic speed pattern in the opposite direction (Link 4616223).

Usually motorist out of curiosity reduce their speed and look at the incident scene. Figure 6 depicts prevailing conditions before and after the incident and for the duration of the incident (duration of delay). It shows a decrease in speed from 43 to 22 mph from the time the incident was created (9:11 PM) to the time when it was cleared (10:11 PM). It is to be noted that for this incident, the reduction in speed in the opposite direction ended at the 10:15 PM close to the time the incident is cleared. This confirms that the presence of responding agencies and flashing lights of the emergency vehicles at the incident scene contributes to traffic delays on the opposite direction. Table 3 shows the resulting delay from this graph observation for both direction of traffic. This incident increased the average west bound travel time rate from 1.22 minutes per mile, to 4.49 minutes per mile. And the rubbernecking delay (in the eastbound direction) changed the average travel time rate from 1.05 minutes per mile, to 1.90 minutes per mile.

Figure 6 - Impact of Incident on Speed and Travel Time Rate on the Opposite Direction (Rubbernecking Effect)



Source: TRANSCOM and TRANSMIT database

Table 3 - Summary of Average Incident Delay	
Incident Attributes (TRANSCOM data)	
ID	1212
Type	Property Damage
Day	Friday
Created Time	11/12/04 9:11 PM
Closed Time	11/12/04 10:11 PM
# of Lane(s)	1 lane blocked
Travel Time Rates for Same Direction (WB)	
1.22	Avg Travel Time rate without incident (min/mi)
4.49	Avg Travel Time Rate with incident (min/mi)
3.27	*Change in TT Rate (min/mi) caused by incident*
Travel Time Rates Opposite Direction (EB)	
1.05	Avg Travel Time rate (min/mi)
1.90	Avg Travel Time rate (min/mi) with rubbernecking effect
0.84	*Change in TT Rate (min/mi) because of rubbernecking*

4.0 Findings

The following section describes the delay impacts for various categories of incidents and selected characteristics. These findings are tabulated for traffic delays on both sides of the roadway and are reported in two parts: for an average day and for the weekday PM peak period (3:00 to 7:00 PM).

Tables 4 to 13 provide the travel time rates without and with an incident for traffic on the same side of the road as the incident, and for traffic in the opposite direction.

4.1 Average Day

4.1.1 Delay Rate by Incident Type

Table 4 summarizes the travel time rates before and after an incident for both directions of traffic flow. Incidents area categorized as accidents, disabled vehicles, and non-vehicle incidents.

An accident, such as property damage with/without personal injury, will increase the travel time rate of motorists directly impacted by the incident, from 1.70 to 3.36 min/mi (or a reduction in speed from 35 mph to 18 mph) - an increase of 97% over the travel time rate without the incident.

Disabled vehicles and non-vehicle incidents are also reported to significantly impact delay rates (an increase of 66% from disabled vehicles, and 61% from non-vehicle incidents). These results clearly show that any incident - especially an accident – imposes a larger increase in travel time.

Rubbernecking effect of the incident is also significant as it is shown that accidents, disabled vehicles and non vehicle related incidents cause 65%, 61% and 56% additional travel time rate over the prevailing conditions without the incident.

Table 4 - Vehicle Delay by Incident Type

						Opposite Direction			
Incident Type	Sample Size	Avg Travel Time Rate without Incident* (min/mi)	Travel Time Rate with Incident* (min/mi)		% Increase on Avg Travel Time Rate Because of Incident (min/mi)	Avg Travel Time Rate without Incident* (min/mi)	Avg Travel Time Rate with Rubbernecking Effect* (min/mi)		% Increase on Avg Travel Time Rate Because of Rubbernecking Effect (min/mi)
Accident (property damage, personal injury)	215	1.70	Avg	3.36	**97%**	1.44	Avg	2.37	**65%**
			Max	8.53			Max	8.22	
Disabled (disabled vehicle and truck, vehicle fire)	183	1.71	Avg	2.84	**66%**	1.44	Avg	2.31	**61%**
			Max	6.29			Max	7.79	
Non-Vehicle (road hazard, hazmat, weather related)	32	1.84	Avg	2.96	**61%**	1.39	Avg	2.18	**56%**
			Max	7.69			Max	4.00	
Total	430	1.71	Avg	3.11	**81%**	1.43	Avg	2.33	**63%**
			Max	8.53			Max	8.22	

* For the Durartion of Delay (From One Hour before to Two Hours after Incident Duration)

4.1.2 Delay Rate by Lane Blockage

Table 5 summarizes the travel time rates impacted by incidents as a function of the number of lanes blocked. It should be noted that I-278 in the sample area is a three lane highway on each direction with no shoulder. However, in some sections there is sufficient roadside space/width to hold vehicles involved in an incident temporarily in order to clear the traveling lanes. In the data reduction task, all incidents reported as “shoulder” or “on/off ramp” blockages were coded as “None Lane Blockage”.

Incidents resulting in none lane blockage cause an increase in travel time rate from 1.55 to 2.46 min/mi, a 59% additional delay. For one-lane blockage, the increase is from 1.74 to 3.09 min/mi equal to 77% additional delay; and for two-lane blockage the increase is equal to 112% additional delay. It is shown that the three-lane blockage causes 101% additional delay over normal traffic conditions. However this value may be not reliable due the small sample size.

The delay impact in the opposite direction follows the same pattern but it is less severe: 41%, 59%, 94% and 89% additional travel time rate for none, one, two, and three or more lane blockages respectively. This indicates that the lane blockage in one side of the road has a large rubbernecking effect on the other side of the road in our study area with the existing conditions and roadway characteristic.

Table 5 - Delay by Lane Blockage

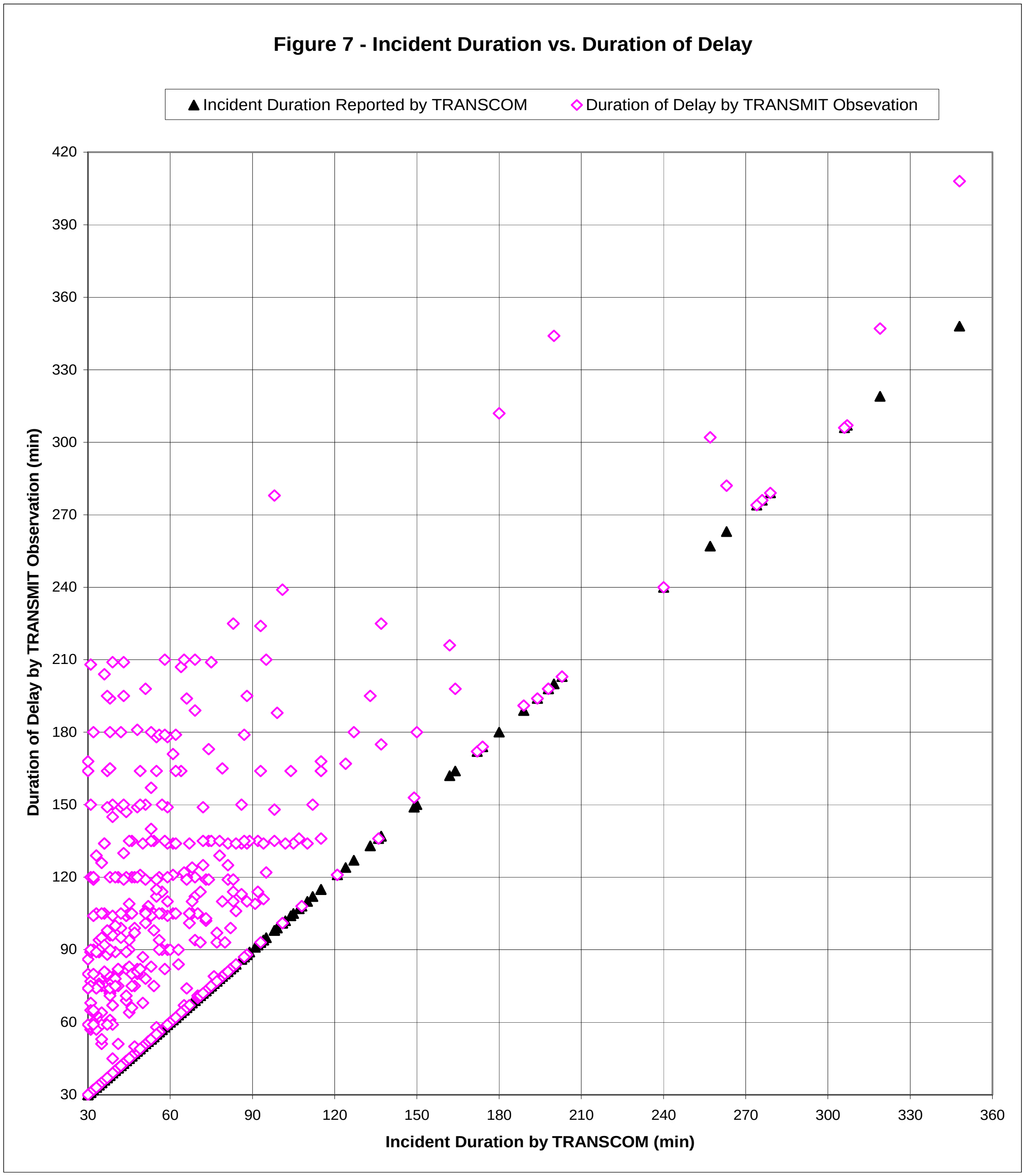
						Opposite Direction			
Lane Blockage	Sample Size	Avg Travel Time Rate without Incident* (min/mi)	Travel Time Rate with Incident* (min/mi)		% Increase on Avg Travel Time Rate Because of Incident (min/mi)	Avg Travel Time Rate without Incident* (min/mi)	Avg Travel Time Rate with Rubbernecking Effect* (min/mi)		% Increase on Avg Travel Time Rate Because of Rubbernecking Effect (min/mi)
None**	28	1.55	Avg	2.46	59%	1.63	Avg	2.28	41%
			Max	6.40			Max	4.24	
One	330	1.74	Avg	3.09	77%	1.44	Avg	2.30	59%
			Max	8.53			Max	8.22	
Two	62	1.62	Avg	3.44	112%	1.28	Avg	2.49	94%
			Max	7.93			Max	7.79	
Three or more	10	1.82	Avg	3.66	101%	1.29	Avg	2.45	89%
			Max	5.98			Max	3.90	

* For the Durartion of Delay (From One Hour before to Two Hours after Incident Duration)

** Incidcates Shoulder or Off ramp Blockage

4.1.3 Delay Rate by Duration of Delay

The delay rate was measured for three levels of incident duration: 30 to <60 minutes; 60 to < 90 minutes; and greater than 90 minutes. The number of observations within each interval were extracted from the TRANSMIT database. Records from TRANSCOM show that they do not correspond to TRANSMIT measurements (see Figure 7). In fact, the duration of delay reported by TRANSCOM is consistently equal to or lower than the duration of delay measured from TRANSMIT data.



Incident Duration by TRANSCOM = Closed Time minus (-) Created Time

Duration of Delay by TRANSMIT Observation = Time after "Closed Time" When Traffic Flow Return to Normal Operation minus (-) Time the Traffic Flow Reduced due to Incident before "Created Time"

Table 6 shows that an incident of 30 to 60 minutes in duration causes 43% additional delay; while incidents with a delay duration between 60 and 90 minutes, cause 67% additional delay, and incidents lasting longer than 90 minutes, cause an average 93% additional delay. The delay rate impact of an incident tends to increase as the duration increases.

The rubbernecking effect in the delay rate for the opposite traffic direction varies from 46% to 58% indicating the presence of an incident has a significant impact. This impact does not appear to be related to the duration of an incident.

Table 6 - Delay by Duration of Delay

Duration of Delay by TRANSMIT Observation	Sample Size	Avg Travel Time Rate without Incident* (min/mi)	Travel Time Rate with Incident* (min/mi)		% Increase on Avg Travel Time Rate Because of Incident (min/mi)	Opposite Direction: Avg Travel Time Rate without Incident* (min/mi)	Opposite Direction: Avg Travel Time Rate with Rubbernecking Effect* (min/mi)		Opposite Direction: % Increase on Avg Travel Time Rate Because of Rubbernecking Effect (min/mi)
30 to < 60 min	77	1.73	Avg	2.47	43%	1.56	Avg	2.28	46%
			Max	7.49			Max	5.87	
60 to < 90 min	106	1.69	Avg	2.82	67%	1.40	Avg	2.21	58%
			Max	7.93			Max	8.22	
90 min or more	247	1.78	Avg	3.43	93%	1.54	Avg	2.37	54%
			Max	8.53			Max	7.79	

* For the Durartion of Delay (From One Hour before to Two Hours after Incident Duration)

4.1.4 Delay Rate by Pavement Conditions

Table 7 shows that only 265 incidents out of 430 samples had pavement condition (wet or dry) information. These data show that although some differences in delay impacts could be attributable to pavement conditions, these differences are not materially different. In fact the reported differences may also be attributable to differences in volume, driver's behavior changes under different pavement conditions, night and daytime observations, etc. The travel time rate for the opposite direction follows a similar pattern.

It is noted that the percentage increase on the travel time rate for incidents with no pavement information is in the same value range than that for which there is no information about pavement conditions. This implies that, from the TRANSMIT data, if the pavement is dry or wet, is not a significant factor in explaining a change in the delay rate caused by an incident.

Table 7 - Delay by Pavement Conditions

Pavement Conditions	Sample Size	Avg Travel Time Rate without Incident* (min/mi)	Travel Time Rate with Incident* (min/mi)		% Increase on Avg Travel Time Rate Because of Incident (min/mi)	Opposite Direction: Avg Travel Time Rate without Incident* (min/mi)	Opposite Direction: Avg Travel Time Rate with Rubbernecking Effect* (min/mi)		Opposite Direction: % Increase on Avg Travel Time Rate Because of Rubbernecking Effect (min/mi)
Dry	230	1.80	Avg	3.26	81%	1.47	Avg	2.44	66%
			Max	8.53			Max	6.80	
Wet	35	1.83	Avg	3.05	67%	1.56	Avg	2.40	54%
			Max	6.96			Max	5.87	
Not Avaiable	165	1.57	Avg	2.91	85%	1.36	Avg	2.17	60%
			Max	7.93			Max	8.22	

* For the Durartion of Delay (From One Hour before to Two Hours after Incident Duration)

4.1.5 Delay Rate by Weather Conditions

Table 8 shows that only 266 incident records with weather condition data were available. Similar to the pavement condition findings, this table shows that weather conditions have little impact on the changes in travel time rates. Adverse weather conditions added 75% to the delay, while the impact of incidents occurring during dry weather, was to add 81% to the delay rate.

The difference for adverse and clear weather conditions for the opposite direction is smaller - at 63% to 64% range. These findings imply that for the present dataset, weather condition is not an influencing factor in the delay rate changes caused by an incident.

Table 8 - Delay by Weather Conditions

						Opposite Direction			
Weather Conditions	Sample Size	Avg Travel Time Rate without Incident* (min/mi)	Travel Time Rate with Incident* (min/mi)		% Increase on Avg Travel Time Rate Because of Incident (min/mi)	Avg Travel Time Rate without Incident* (min/mi)	Avg Travel Time Rate with Rubbernecking Effect* (min/mi)		% Increase on Avg Travel Time Rate Because of Rubbernecking Effect (min/mi)
Clear	204	1.79	Avg	3.23	81%	1.46	Avg	2.39	64%
			Max	8.53			Max	6.80	
Adverse	62	1.87	Avg	3.27	75%	1.59	Avg	2.59	63%
			Max	7.49			Max	5.99	
Not Avaiable	164	1.57	Avg	2.90	85%	1.35	Avg	2.16	60%
			Max	7.93			Max	8.22	

* For the Durartion of Delay (From One Hour before to Two Hours after Incident Duration)

4.2 Weekday PM Peak Period (3:00 PM – 7:00 PM)

The following sections summarize the tabulated values for the 75 incidents occurring during an average weekday PM peak period (3:00 PM to 7:00 PM). This time frame coincides with the peak traffic volumes in the study corridor as depicted in Figure 2 to 3. Because of the higher traffic demand, the magnitude of delay for nearly all categories of incidents is larger compared to the 24 hour period analysis. This shows that the impact of an incident on the delay rate is highly correlated to the volume using the highway at the time of the incident.

4.2.1 Delay Rate by Incident Type for Weekday PM Peak

Table 9 shows that an average accident increases the travel time rate from 1.89 to 3.64 minutes/mile and a disabled vehicle causes an increase from 2.08 to 3.40 minutes/mile. When compared to delays experienced on an average weekday (Table 4), an incident in the peak period has grater delay impacts. Accidents tend to create more delay than disabled vehicles.

It is interesting to note that in the opposite direction the rubbernecking effect portrays no sensitivity to incident type as experienced in an average day and confirms that any has a significant impact on rubbernecking effect.

Table 9 - Delay by Incident Types for Weekday PM Peak

Incident Type	Sample Size	Avg Travel Time Rate without Incident* (min/mi)	Travel Time Rate with Incident* (min/mi)		% Increase on Avg Travel Time Rate Because of Incident (min/mi)	Opposite Direction: Avg Travel Time Rate without Incident* (min/mi)	Opposite Direction: Avg Travel Time Rate with Rubbernecking Effect* (min/mi)		Opposite Direction: % Increase on Avg Travel Time Rate Because of Rubbernecking Effect (min/mi)
Accident (property damage, personal injury)	30	1.89	Avg	3.64	92%	1.81	Avg	2.75	52%
			Max	7.49			Max	4.55	
Disabled (disabled vehicle and truck, vehicle fire)	42	2.08	Avg	3.40	64%	1.59	Avg	2.45	54%
			Max	6.29			Max	4.75	
Non-Vehicle** (road hazard, hazmat, weather related)	3	N/A	Avg	N/A	N/A	N/A	Avg	N/A	N/A
			Max	N/A			Max	N/A	
Total	75	2.03	Avg	3.55	75%	1.64	Avg	2.55	55%
			Max	7.69			Max	4.75	

* For the Durartion of Delay (From One Hour before to Two Hours after Incident Duration)

** Sample Size is too Small

4.2.2 Delay Rate by Lane Blockage for Weekday PM Peak

Table 10 shows that one-lane blockage adds 72% and two-lane blockages add 101% to the prevailing travel time rate. There were no data showing the impact of incidents that did not block the travel lanes. The delay impacts incidents occurring in the PM peak period are significantly greater than those for an average day (Table5).

Table 10 - Delay by Lane Blockage for Weekday PM Peak

Lane Blockage	Sample Size	Avg Travel Time Rate without Incident* (min/mi)	Travel Time Rate with Incident* (min/mi)		% Increase on Avg Travel Time Rate Because of Incident (min/mi)	Opposite Direction: Avg Travel Time Rate without Incident* (min/mi)	Opposite Direction: Avg Travel Time Rate with Rubbernecking Effect* (min/mi)		Opposite Direction: % Increase on Avg Travel Time Rate Because of Rubbernecking Effect (min/mi)
One	66	2.03	Avg	3.49	72%	1.68	Avg	2.57	53%
			Max	7.69			Max	4.75	
Two	7	1.96	Avg	3.94	101%	1.21	Avg	2.50	106%
			Max	5.17			Max	4.55	
Three or More**	2	N/A	Avg	N/A	N/A	N/A	Avg	N/A	N/A
			Max	N/A			Max	N/A	

* For the Durartion of Delay (From One Hour before to Two Hours after Incident Duration)

** Sample Size is too Small

The rubbernecking delay impacts in the opposite direction are also significant, and greater than those occurring on an average day.

4.2.3 Delay Rate by Incident Duration for Weekday PM Peak

Table 11 shows the delay rates three incident duration intervals. It can be seen that as the duration of delay increases the delay impact of an incident increases. And this effect is more pronounced for the peak period than for an average day (Table 6).

The rubbernecking effect of an incident in the peak period is also very significant and tends to be higher than that for an average day.

Table 11 - Delay by Duration of Delay for Weekday PM Peak

						Opposite Direction			
Duration of Delay by TRANSMIT Observation	Sample Size	Avg Travel Time Rate without Incident* (min/mi)	Travel Time Rate with Incident* (min/mi)		% Increase on Avg Travel Time Rate Because of Incident (min/mi)	Avg Travel Time Rate without Incident* (min/mi)	Avg Travel Time Rate with Rubbernecking Effect* (min/mi)		% Increase on Avg Travel Time Rate Because of Rubbernecking Effect (min/mi)
30 to < 60 min	16	2.21	Avg	3.22	45%	1.75	Avg	2.64	51%
			Max	7.49			Max	4.02	
60 to < 90 min	23	1.75	Avg	2.97	69%	1.26	Avg	2.30	82%
			Max	5.21			Max	4.75	
90 min or more	36	2.15	Avg	4.07	89%	1.82	Avg	2.67	46%
			Max	7.69			Max	4.55	

* For the Durartion of Delay (From One Hour before to Two Hours after Incident Duration)

4.2.4 Delay by Pavement Conditions for Weekday PM Peak

Table 12 shows that there are not enough sample observations to compare the effect of pavement conditions on the delay rate impacts of an incident occurring in a pavement that is dry vs. one that is wet.

Table 12 - Delay by Pavement Conditions for Weekday PM Peak

						Opposite Direction				
Pavement Conditions	Sample Size	Avg Travel Time Rate without Incident* (min/mi)	Travel Time Rate with Incident* (min/mi)		% Increase on Avg Travel Time Rate Because of Incident (min/mi)	Avg Travel Time Rate without Incident* (min/mi)	Avg Travel Time Rate with Rubbernecking Effect* (min/mi)		% Increase on Avg Travel Time Rate Because of Rubbernecking Effect (min/mi)
Dry	53	2.03	Avg	3.60	77%	1.64	Avg	2.45	49%
			Max	7.69			Max	4.42	
Adverse**	5	N/A	Avg	N/A	N/A	N/A	Avg	N/A	N/A
			Max	N/A			Max	N/A	
Not Avaiable	17	2.02	Avg	3.56	77%	1.70	Avg	2.91	72%
			Max	5.17			Max	4.75	

* For the Durartion of Delay (From One Hour before to Two Hours after Incident Duration)

** Sample Size is too Small

4.2.5 Delay Rate by Weather Conditions for Weekday PM Peak

Table 13 shows that weather conditions have a likely significant impact on the delay rate of traffic impacted by an incident in the PM peak period. For the ten incidents occurring during adverse weather conditions, an incident increased the delay rate by 81%, compared with a 73% increase during clear weather condition. The results for the opposite direction also show a pattern similar to that of the average weekday.

Table 13 - Delay by Weather Conditions for Weekday PM Peak

						Opposite Direction			
Weather Conditions	Sample Size	Avg Travel Time Rate without Incident* (min/mi)	Travel Time Rate with Incident* (min/mi)		% Increase on Avg Travel Time Rate Because of Incident (min/mi)	Avg Travel Time Rate without Incident* (min/mi)	Avg Travel Time Rate with Rubbernecking Effect* (min/mi)		% Increase on Avg Travel Time Rate Because of Rubbernecking Effect (min/mi)
Clear	48	1.99	Avg	3.44	73%	1.66	Avg	2.48	49%
			Max	7.69			Max	4.42	
Adverse**	10	2.25	Avg	4.07	81%	1.49	Avg	2.27	53%
			Max	7.49			Max	4.55	
Not Avaiable	17	2.02	Avg	3.56	77%	1.70	Avg	2.91	72%
			Max	5.17			Max	4.75	

* For the Durartion of Delay (From One Hour before to Two Hours after Incident Duration)

** Sample Size is too Small

Task 3.2: DATA ANALYSIS

Prepared by:

Kaan Ozbay, Ph.D.
Associate Professor
Rutgers University, Civil and Environmental Engineering
623 Bowser Road, Piscataway, NJ 08854-8014
Tel: 732.445.2792
Fax: 732.445.0577
Email: kaan@rci.rutgers.edu

M. Anil Yazici
Graduate Assistant
Rutgers University, Civil and Environmental Engineering
623 Bowser Road, Piscataway, NJ 08854-8014
Tel: 732.445.5496
Fax: 732.445.0577
Email: yazici@eden.rutgers.edu

June 20, 2007

1.0 Introduction

In order to best predict delays due to non-recurrent accidents, a two-part model is needed. The first part is the estimation of the number (frequency) and duration of different types of incidents over a given time period. In other words, this is the estimation of the probability that incidents of a given type and duration will occur at a given location. The second part of the model is the delay estimation based on the previous frequency and duration estimates. In this section, statistical analysis of the incident data described in Task 3.1 is performed to determine the descriptive statistics and probability distributions of the incident frequencies and durations. For this purpose an extensive statistical analysis is performed for different parameters of incident duration records.

Firstly, incident durations are analyzed and distribution fits for different facilities and parameters are performed. Secondly, a statistical analysis is performed for incident frequencies, however the results of this analysis are not valid because the available TRANSCOM data covers less than 10% of the total accidents in the study area. The detailed explanations about the reliability of TRANSCOM data are mentioned in the following section. Hence, the results of incident rate analysis are given in APPENDIX as a guidance for model development and do not represent the real dynamics of the incident frequencies.

2.0 Some Important Remarks about the Available TRANSCOM Dataset

TRANSCOM incident data studied in Task 3 covers a variety of incident types that consists of 1907 records for the study area for 14 months. However, a recently available dataset obtained from a source titled *"NYSDOT Safety Information Management System, Accident Verbal Description Report For I-287 Intersection & Non-Intersection Accidents"* contains 16489 records for the same study area for a time period of 24 months. Major qualitative difference between the two datasets is that the new NYSDOT data only have collision records whereas the TRANSCOM data has also the disablement accidents. Table 1 shows NYSDOT data fields from which the accident type is extracted for comparing with TRANSCOM data.

Table 1- NYSDOT Data Fields According to Accident Types (Source: *"NYSDOT Safety Information Management System, Accident Verbal Description Report For I-287 Intersection & Non-Intersection Accidents)*

Accident Class	Type of Accident
• Property Damage (3022 records /16849) • Injury (5331 records /16849) • Fatal (25 records /16849) • Property Damage & Injury (625 records /16849) • Non-reportable = Property Damage < $1,000 (7846 records /16849)	• Coll.W/Earth Ele./Rock Cut/Ditch • Coll.W/Light Support/Utility Pole • Collision with Animal • Collision with Bicyclist • Collision with Bridge Structure • Collision with Building/Wall • Collision with Crash Cushion • Collision with Culvert/Headwall • Collision with Curbing • Collision with Fence • Collision with Fire Hydrant • Collision with Guide Rail • Collision with Guiderail - End • Collision with Median/Barrier • Collision with Median/Barrier - End • Collision with Motor Vehicle • Collision with Other • Collision with Other Barrier • Collision with Other Fixed Object • Collision with Pedestrian • Collision with Sign Post • Collision with Tree • Fire/Explosion • Not Entered • Other Non-Collision • Overturned • Unknown

As seen in Table 1, in NYSDOT dataset:

- Majority of "Accident Class" fields are either "Property Damage" or "Injury" type.
- Some records have 2 identifiers: "Property Damage and Injury".
- "Type of Accident" field has more variety, however all being of collision type.
- No disablement information can be gathered from this dataset. "Disabled Vehicle" and "Disabled Truck" incidents in TRANSCOM constitute about 47% of overall data, thus it can be said that disablement incidents are as important as collisions, assuming that the incident reports are unbiased and properly kept.
- There is no information regarding HAZMAT operations.
- Also, the cause of the incident is not specifically mentioned as "road hazard", or "weather related".

It should also be noted that TRANSCOM data covers a period of 14 months between February 2004 and March 2005, and NYSDOT dataset covers the 2-year period between January 2000 and December 2002. On the quantitative side, for the 14 month period, only 886 of 1907 records in TRANSCOM dataset are accidents, whereas NYSDOT data has a 14 month average of 11493 accidents. In Table 2 summary of two datasets in terms of incident types along with corresponding percentages is presented. Please note that "Fatal" accidents in NYSDOT are categorized as "Injury" and "Property Damage and Injury" accidents are categorized as "Property Damage" following the order of reporting, so that a comparison can be made with TRANSCOM.

Table 2- Summary of TRANSCOM and NYSDOT Databases

Category	Incident Types	# of Incidents		%		Total # of Incidents	
		TRANSCOM (02.04 – 04.05)	**NYSDOT (01.00-12.02)**	**TRANSCOM (02.04 – 04.05)**	**NYSDOT (01.00-12.02)**	**TRANSCOM (02.04 – 04.05)**	**NYSDOT (01.00-12.02)**
Accident	**Property Damage**	819	11493	42.9%	68.2%	886	16849
	Personal Injuries	67	5356	3.5%	31.8%		
Disabled	**Disabled Vehicle**	694	**N/A**	36.4%	0%	902	N/A
	Disabled Truck	179	**N/A**	9.4%			
	Vehicle Fire	29	**N/A**	1.5%			
Non-Vehicle Related	**Road HAZARD**	80	**N/A**	4.2%	0%	119	N/A
	HAZMAT	29	**N/A**	1.5%			
	Weather Related	10	**N/A**	0.5%			
TOTAL		1907	16849	100%	100%	1907	16849

Table 2 shows that:

- NYSDOT accident database cannot be simply compared with whole TRANSCOM database, since TRANSCOM has both accidents and disablements whereas NYSDOT has accidents only.
- New York State Department of Motor Vehicles "The Vehicle and Traffic Law" was amended in April, 1997 to provide statutory authority to police officer to report accidents mentions that reporting criteria for property damage relies on police officer judgment, if the accident appears to meet the criteria of damage in excess of $1,000 to the property of any one individual. Thus, "Non-Reportable" records are in fact property damage accidents, with a damage of less than $1,000, and they are treated as "Property Damage" accordingly.
- Table 3 shows the statistics for TRANSCOM and NYSDOT datasets, with the corresponding percentage of each accident class.
- Overall those assumptions are made for accident types to be able to compare two datasets which is shown in Table 3:
  - Fatal → Injury
  - Non-Reportable → Property Damage
  - Property Damage and Injury → Property Damage

Table 3- Summary of Dataset Records

	Property Damage	Personal Injury	Total
TRANSCOM	818 (92.4%)	67 (7.6%)	885
NYSDOT	11493 (68.2%)	5356 (31.8%)	16849

Major conclusions from the comparison of two datasets are as follows:

- Data dorm the two different datasets are collected during different time periods limiting unbiased comparison of the two datasets. .
  - TRANSCOM covers 14 months of data between February 2004 and March 2005, and NYSDOT dataset covers the 2-year period between January 2000 and December 2002.

- Analysis in percentages is more relevant for comparison of the two datasets since length of their collection periods is not the same.
- As seen from Table 2, the percentage of “Property Damage” in TRANSCOM dataset is significantly low than the NYSDOT data.
- It should be remembered that accidents having “Property Damage and Injury” identifiers were also accepted as “Property Damage”, and if the opposite is done, the difference will be even larger.
- Thus, it can be said that either there are reporting criteria differences between the two datasets or one of the datasets is biased in terms of the number of reported accidents.
- Since there are no disablement records in NYSDOT data, no similar inference can be made.

Overall, as the conclusions of the dataset comparison reveal, these two datasets cannot be merged to obtain a more complete dataset. The discrepancy in the number of records implies that TRANSCOM dataset underestimates the number of accidents by almost 93%, assuming that NYSDOT data has the complete number of accident records. In this respect, the results of this study are mainly limited by the incompleteness of the available TRANSCOM dataset.

Nevertheless, interpretations of the duration analysis can still be valuable mainly due to the fact that a sample of the total accident data is processed and the results of this analysis can be extended to represent the overall reality for that small sample size. However, the frequency analysis cannot be used to draw final rates that can be used by the State since, unlike the duration analysis, complete accident record set is needed for a reliable incident frequency analysis that can be used to determine accidents rates for different types of accidents. For this reason, as mentioned above, the frequency analysis section is moved to APPENDIX.

3.0 Duration Analysis

First, all incident data are analyzed without using any incident specific information. Data statistics for different expressways and incident types are also discussed seperately, to obtain facility (Brooklyn Queens Expressway, Gowanus Expressway etc.) and incident type specific information. The estimation of probability distributions is performed for different incident types, namely “Property Damage”, “Disabled Truck”, “Disabled Vehicle” because they are the most frequently observed incident types. The analysis of “Road Hazard”,”Personal Injury”, “Vehicle Fire”, “HAZMAT” and “Weather Related” incidents as individual categories cannot be performed is because of the small number of data points corresponding to each category. Total number of data points at all facilities for “Road Hazard”and “Personal Injury” are 80 and 67, “Vehicle Fire” and “HAZMAT” have 29 records and “Weather Related” has only 20 entries over 1907 total records. These numbers are not sufficient for general analysis and get smaller in facility spesific analysis. However effects of these relatively infrequent accident types are not negligible because, their durations can sometimes be very long (up to 16 hours). Some incidents which are reported to be road hazard are determined to be pothole repairs or road maintenance, thus the use of these “non-recurring” have to be further evaluated. However, the long incident durations contribute to overall delay and are important components of the analysis. Thus, when appropriate, they are used as part of the the overall data set used to estimate probability distributions for the complete data. In addition to very long durations, incident duration data also include very low values, even zero durations. Unlike the long durations, short durations are not expected to cause major delay and can be neglected to avoid unreasonable estimates. In the analysis, instead of ignoring the short duration incidents, only zero durations (which have nine entries in complete dataset) are eliminated and the rest are kept in the data set regardless of how short these durations are to ensure reasonable sample size.

Various probability distribution types for the complete and sub-categories of the incident dataset are estimated. MATLAB , is used for curve fitting and hypothesis testing. The following distributions are considered primarily for curve fitting:

- Beta
- Chi-Squared
- Exponential
- Extreme Value Type A (Gumbel)
- Extreme Value Type B
- Gamma
- Laplace (Double Exponential)
- Lognormal
- Normal
- Rayleigh
- Student's t
- Uniform
- Weibull
- Phase Bi-Exponential
- Phase Bi-Weibull
- Wakeby

Apart from the last three distributions, all other distributions are commonly used in statistical analysis. However, weibull, gamma and lognormal distributions are of special interest since past work in incident data analysis shows that those distributions are found to better represent probabilistic distribution of incident ferquencies and durations. These 3 distributions are also found to fit better NYC data, compared to the others. For each data set and probability distribution, relevant mean, variance and goodness-of-fit (GOF) measures are calculated for each incident type.

Below are some quick facts[4] about the distributions that are used in this study.

Weibull Distribution[4]

The formula for the probability density function of the general Weibull distribution is:

$$f(x) = \frac{\gamma}{\alpha}\left(\frac{x-\mu}{\alpha}\right)^{(\gamma-1)} \exp\left(-\left(\frac{x-\mu}{\alpha}\right)^{\gamma}\right) \qquad x \geq \mu; \gamma, \alpha \succ 0$$

where γ is the shape parameter, μ is the location parameter and α is the scale parameter. The case where μ = 0 and α = 1 is called the standard Weibull distribution. The case where μ = 0 is called the 2-parameter Weibull distribution.

The formula for the cumulative distribution function of the Weibull distribution is:

$$F(x) = 1 - e^{-\left(x^{\gamma}\right)} \qquad x \geq 0; \gamma \succ 0$$

Gamma Distribution[4]

The general formula for the probability density function of the gamma distribution is:

$$f(x) = \frac{\left(\frac{x-\mu}{\beta}\right)^{\gamma-1} \exp\left(-\frac{x-\mu}{\beta}\right)}{\beta\,\Gamma(\gamma)} \qquad x \geq \mu; \gamma, \beta \succ 0$$

where γ is the shape parameter, μ s the location parameter, β is the scale parameter, and Γ is the gamma function which has the formula:

$$\Gamma(a) = \int_0^{\infty} t^{a-1} e^{-t} dt$$

The case where μ = 0 and β = 1 is called the standard gamma distribution.

Lognormal Distribution[4]

A variable X is lognormally distributed if Y = LN(X) is normally distributed with "LN" denoting the natural logarithm. The general formula for the probability density function of the lognormal distribution is

$$f(x) = \frac{e^{-\left(\frac{\left(\ln\left(\left(x-\theta\right)/m\right)\right)^2}{\left(2\sigma^2\right)}\right)}}{(x-\theta)\sigma\sqrt{2\pi}}, \quad x \geq \theta; m, \sigma \succ 0$$

where σ is the shape parameter, θ is the location parameter and *m* is the scale parameter. The case where θ = 0 and m = 1 is called the standard lognormal distribution. The case where θ equals zero is called the 2-parameter lognormal distribution.

The formula for the cumulative distribution function of the lognormal distribution is:

$$F(x) = \Phi\left(\frac{\ln(x)}{\sigma}\right) \quad x \geq 0; \sigma \succ 0$$

where Φ is the cumulative distribution function of the normal distribution.

Overall, among those three distributions the Lognormal and Weibull distributions are probably the most commonly used distributions in reliability applications aimed at modeling failure times. The Weibull distribution has a relatively simple distributional form. The shape parameter allows the Weibull to assume a wide variety of shapes. Thus, it is no surprise that those distributions are also employed for incident durations which is analogous to component failures. If we consider traffic accidents as some kind of failure, it is thus not surprising that these distributions are found to perform well when used to model the duration of traffic accidents

For probability distribution fitting, two goodness of fit tests, namely Kolmogorov-Smirnov and Anderson-Darling tests, are used in this section. Below, some brief information about those tests are given for the interested reader.

Kolmogorov-Smirnov Test (K-S)[4]

The Kolmogorov-Smirnov test is used to decide if a sample comes from a population with a specific distribution. The Kolmogorov-Smirnov (K-S) test is based on the empirical distribution function (ECDF). Given *N ordered* data points $\boldsymbol{Y_1, Y_2, ..., Y_N}$, the ECDF is defined as:

$$E_N = {n(i)}/{N}$$

where *n(i)* is the number of points less than Y_i and the Y_i are ordered from smallest to largest value. This is a step function that increases by 1/*N* at the value of each ordered data point.
The Kolmogorov-Smirnov test statistic is defined as:

$$D = \max_{1 \leq i \leq N}\left(F(Y_i) - \frac{i-1}{N}, \frac{i}{N} - F(Y_i)\right)$$

where *F* is the theoretical cumulative distribution of the distribution being tested which must be a continuous distribution

An attractive feature of this test is that the distribution of the K-S test statistic itself does not depend on the underlying cumulative distribution function being tested. Another advantage is that it is an exact test (the chi-square goodness-of-fit test depends on an adequate sample size for the approximations to be valid). Despite these advantages, the K-S test has several important limitations:

1. It only applies to continuous distributions.

2. It tends to be more sensitive near the center of the distribution than at the tails.
3. Perhaps the most serious limitation is that the distribution must be fully specified. That is, if location, scale, and shape parameters are estimated from the data, the critical region of the K-S test is no longer valid. It typically must be determined by simulation.

Anderson-Darling Test (A-D)[4]

The Anderson-Darling test is used to test if a sample of data came from a population with a specific distribution. It is a modification of the Kolmogorov-Smirnov (K-S) test and gives more weight to the tails than does the K-S test. The K-S test is distribution free in the sense that the critical values do not depend on the specific distribution being tested. The Anderson-Darling test makes use of the specific distribution in calculating critical values. This has the advantage of allowing a more sensitive test and the disadvantage that critical values must be calculated for each distribution[4]. The Anderson-Darling test is an alternative to the chi-square and Kolmogorov-Smirnov goodness-of-fit tests.

The Anderson-Darling test statistic is defined as

$$A^2 = -N - S$$

where

$$S = \sum_{i=1}^{N} \frac{(2i-1)}{N} \left[\ln F(Y_i) + \ln\left(1 - F(Y_{N+1-i})\right) \right]$$

F is the cumulative distribution function of the specified distribution. Note that the Y_i are the *ordered* data.

The Anderson-Darling test is used to test if a sample of data came from a population with a specific distribution. It is a modification of the Kolmogorov-Smirnov test and gives more weight to the tails than does the K-S test. The K-S test is distribution free in the sense that the critical values do not depend on the specific distribution being tested. The Anderson-Darling test makes use of the specific distribution in calculating critical values. This has the advantage of allowing a more sensitive test and the disadvantage that critical values must be calculated for each distribution[4].

For all the goodness-of-fit tests, the following hypothesis stated below is used:

H_0: The data follow the specified distribution.

H_a: The data do not follow the specified distribution.

The hypothesis regarding the distributional form is rejected at the chosen significance level (alpha) if the test statistic, D, is greater than the critical value.

4.0 Overall Data Statistic

As mentioned before, the available data has 1907 records covering February 2004 through March 2005 period. An extensive descriptive analysis of the data is given in Task 3.1. Briefly, majority of the incidents (77.9 %) have less than 1-hour duration, 50.5% of the overall being less than half an hour. Overall, percentages show a decreasing trend as the duration increases, however for the incident types other than “Property Damage” and “Disabled Vehicle”, this trend does not hold in a consistent manner. Nevertheless, these two types, which constitute almost 80% of the total incident record, thus will affect the final lookup tables as much. Analysis also shows that majority of the incidents (77.7%) blocks one lane having duration less than an hour with a percentage of 50.5%. In general, it can be concluded that incidents in this database are mostly “Property Damage” (79.3%) or “Disabled Vehicle” incidents, blocking one lane (77.7%) with durations less than one hour (77.9%) for the complete dataset.

Distribution Analysis of Incident Durations

Below, the outcome of the data analysis is shown for three major incident categories, namely “Property Damage”, “Disabled Vehicle” and “Disabled Truck”. The order of the types is arranged according their frequency of occurrences. Other incident types, namely “Road Hazard”, “Weather Related”, “HAZMAT”, “Personal Injury”, “Vehicle Fire” are only included in the overall analysis since the number of data points for these types is large enough to draw statistically reliable conclusions. The analyses are presented firstly the complete dataset to understand the general picture. Then, more focused, facility specific distributions are presented.

Analysis of System-Wide Incident Durations

The histogram plot of incident durations can be seen in Figure 1. Descriptive incident duration statistics for all types of incidents is shown in detail in Table 4. Table 5 shows the incident durations according to the facility. It can be said that the incident duration is affected by both the incident type and the facility. “Disabled Vehicle” type of incident has the smallest average and standard deviation value, whereas Disabled Truck has a higher duration and deviance. Among facilities, GE has the smallest average value and SIE has the largest average. BE shows the largest deviance (about 30 mins more than SIE) although its average is 22 min less than SIE. Briefly, every facility exhibits different duration characteristics. Note that zero durations were eliminated from the data set and thus the total number of data points is less than 1907, as previously mentioned.

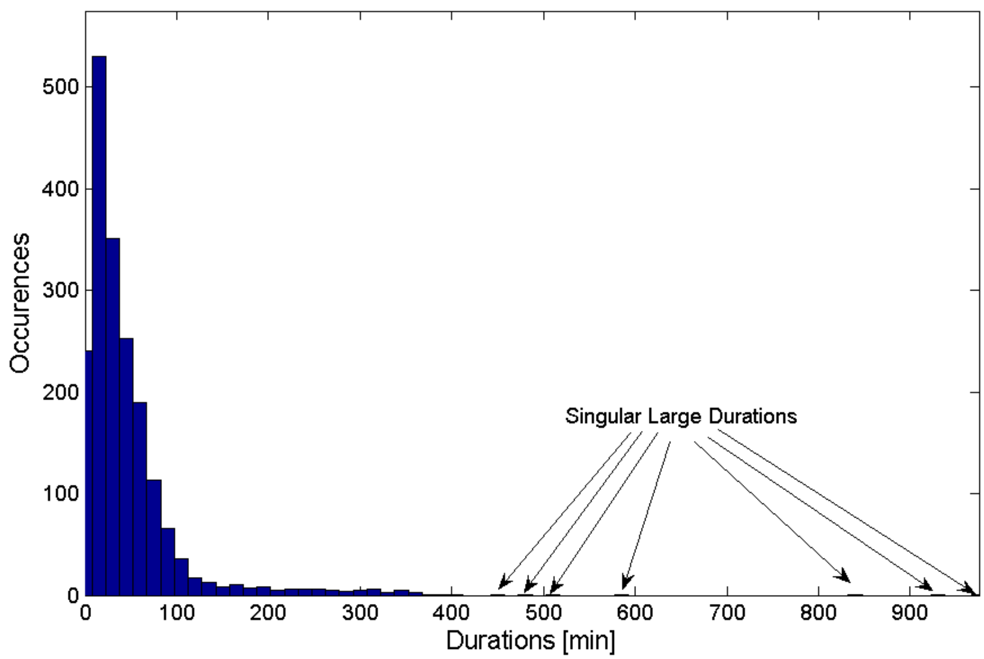


Figure 1- Histogram of Incident Durations for All Facitilies -All incident Types

Table 4- Descriptive Statistics of Incident Durations According to Incident Type

Incident Type	# of Data Points	Mean [min]	Std Deviation [min]	95% Duration	Max Duration

				[min]	[min]
Property Damage	768	43	46	119	515
Disabled Vehicle	654	28	29	77	327
Disabled Truck	170	52	88	117	839
Total*	1898	47	68	141	969

*Road Hazard, Weather Related, HAZMAT, Personal Injury and Vehicle Fire incident types are included

Table 5- Descriptive Statistics of Incident Durations According to Facility

Incident Type	# of Data Points	Mean [min]	Std Deviation [min]	95% Duration [min]	Max Duration [min]
BQE	936	47	68	127	969
GE	634	37	47	98	356
SIE	110	88	84	240	363
BE	107	66	111	100	933

The summary of the KS and AD tests for Weibull, Gamma and Lognormal distributions with relevant GOF statistics are presented in Table 6,

Table 7, and Table 8 respectively. All the tests are performed for 95% confidence level.

Table 6- Weibull Fit and GOF Test Statistics at 95% Confidence Level for Complete Incident Duration Data

Incident Type	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Good Fit	Test Statistics	Critical Value	Good Fit
Property Damage	44.499 1.0956	0.043305	0.048784	**Yes**	2.5412	0.757	No
Disabled Vehicle	29.504 1.1267	0.039405	0.052844	**Yes**	1.6239	0.757	No
Disabled Truck	50.347 0.93411	0.097258	0.10313	**Yes**	3.3839	0.757	No
All Incidents	45.89 0.94316	0.059606	0.031084	**No**	17.431	0.757	No

Table 7- Gamma Fit and GOF Test Statistics at 95% Confidence Level for Complete Incident Duration Data

Incident Type	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Fit	Test Statistics	Critical Value	Good Fit
Property Damage	1.0106 46.86	0.037555	0.048784	**Yes**	1.4809	0.844	No
Disabled Vehicle	1.037 50.515	0.037785	0.052844	**Yes**	1.0268	0.844	No
Disabled Truck	1.3165 21.381	0.096949	0.10313	**Yes**	2.8625	0.844	No
All Incidents	1.2493 34.315	0.064329	0.031084	No	17.046	0.844	No

Table 8- Lognormal Fit and GOF Test Statistics at 95% Confidence Level for Complete Incident Duration Data

Incident Type	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Good Fit	Test Statistics	Critical Value	Good Fit
Property Damage	3.2872 1.1115	0.07266	0.048784	No	1.8727	0.735	No
Disabled Vehicle	3.4043 1.0685	0.058991	0.052844	No	3.2206	0.735	No
Disabled Truck	2.9118 0.99221	0.086395	0.10313	**Yes**	7.6459	0.735	No
All Incidents	3.3073 1.0441	0.048313	0.031084	No	8.1295	0.735	No

First point to be noted is that the Lognormal distribution is not capable of capturing the duration distributions of these incidents at all. Second point is the relatively poor performance of all the distributions in case of the Anderson-Darling test. None of the distributions have succeeded to pass this test at 95% confidence level. This is mainly due to the fact that A-D test gives more weigh to tails, and since the tail values in incident duration data is widespread and scarce compared to main body, the test results are poor. If the test statistics for individual incident categories are investigated (Table 6 - Table 8), it can be observed that for incident types, which do not show infrequent long durations, test statistics are closer (or less) to the critical value. This inability to detect the long duration incidents can also be seen in Figure 2. In the logarithmic cumulative probability plot shown in Figure 2, if the data points (blue) fall on/near the red line, which represents the good prediction range, the assumption that the data comes from a Weibull distribution is assumed to be reasonable. As can be observed from Figure 2, the distribution follows the Weibull distribution almost perfectly for accident durations between 10 and 100 minutes, but fails at the tails (for incident durations>100 minutes). Weibull and Gamma distributions both exhibit very similar performance for the overall incident duration data. They can be assumed to represent the data fairly accurately, since the main body of the data lies in the "so-called" good-prediction range rather than the tails. However, the poor performance of tail prediction should be kept in mind for real-life implementation.

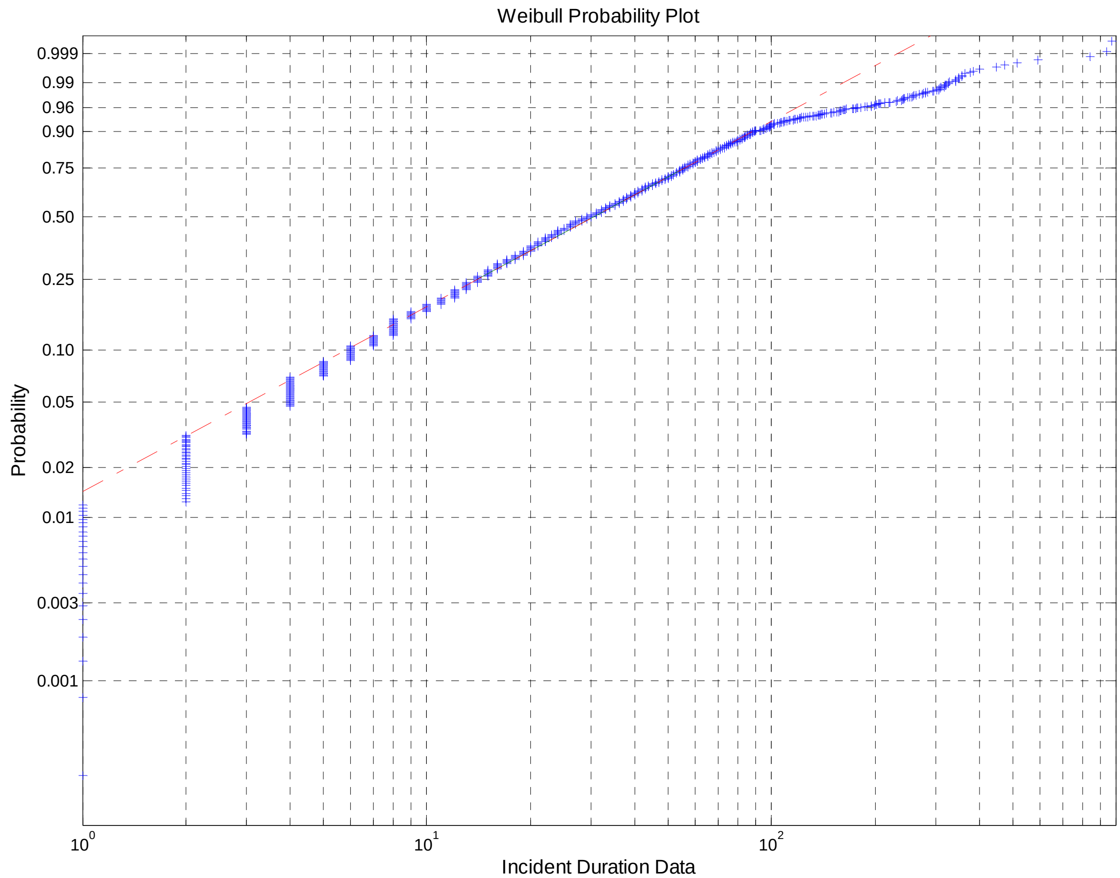


Figure 2- Weibull Probability Plot for Overall Incident Duration Data

Analysis of Facility Based Incident Durations

For location specific results, each facility is analyzed separately; both distributional and descriptive statistics are gathered for each facility in the dataset.

<u>Analysis of Brooklyn-Queens Expressway (BQE) Incident Durations</u>

The incident duration distributions for this facility can be seen in Figure 3 and descriptive incident duration statistics for all types of incidents can be found in detail in Table 9.

Table 9- Descriptive Statistics of Incident Durations for BQE

Incident Type	**# of Data Points**	**Mean [min]**	**Std Deviation [min]**	**95% Duration [min]**	**Max Duration [min]**
Property Damage	398	43	48	108	515
Disabled Vehicle	353	28	28	77	326
Disabled Truck	87	60	104	115	839
Total	936	47	68	127	969

*Road Hazard, Weather Related, HAZMAT, Personal Injury and Vehicle Fire incident types are included

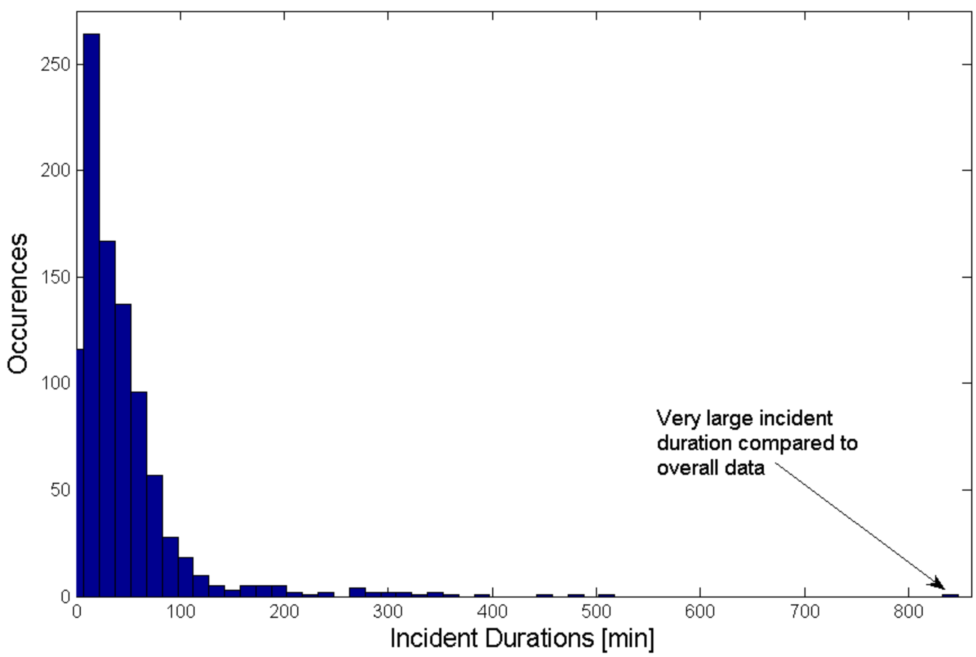


Figure 3- Histogram of Incident Durations for BQE-All incident Types

The summary of the K-S and AD tests for Weibull, Gamma and Lognormal distributions with relevant GOF statistics are presented in Table 10,

Table 11, and Table 12 respectively. All the tests are performed within 95% confidence level.

Table 10- Weibull Fit and GOF Test Statistics at 95% Confidence Level for BQE Incident - Duration Data

Incident Type	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Good Fit	Test Statistics	Critical Value	Good Fit
Property Damage	44.4580 1.0843	0.0541	0.0676	**Yes**	1.6452	0.757	No
Disabled Vehicle	29.9038 1.1464	0.0415	0.0718	**Yes**	0.803	0.757	No
Disabled Truck	57.0278 0.9124	0.1155	0.1436	**Yes**	2.2462	0.757	No
All Incidents	45.4809 0.9547	0.0651	0.0442	No	9.0804	0.757	No

Table 11- Gamma Fit and GOF Test Statistics at 95% Confidence Level for BQE Incident Duration Data

Incident Type	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Fit	Test Statistics	Critical Value	Good Fit
Property Damage	1.2396 34.6712	0.0481	0.0676	**Yes**	1.0612	0.844	No
Disabled Vehicle	1.3487 21.0408	0.0482	0.0718	**Yes**	0.50046	0.844	**Yes**
Disabled Truck	0.9919 60.7703	0.1310	0.1436	**Yes**	2.2491	0.844	No
All Incidents	1.0416 44.7604	0.0688	0.0442	No	8.1126	0.844	No

Table 12- Lognormal Fit and GOF Test Statistics at 95% Confidence Level for BQE Incident Duration Data

Incident Type	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Good Fit	Test Statistics	Critical Value	Good Fit
Property Damage	3.3060 1.0444	0.0865	0.0676	No	4.5844	0.735	No
Disabled Vehicle	2.9311 0.9825	0.0553	0.0718	**Yes**	1.7912	0.735	No
Disabled Truck	3.5165 1.1267	0.0948	0.1436	**Yes**	1.3964	0.735	No
All Incidents	3.2905 1.0869	0.0579	0.0442	No	4.6134	0.735	No

Like the overall case, none of the distributions succeed to represent all incident data distributions for BQE. This is not surprising since BQE constitutes the main portion of the total dataset and their characteristics are quite similar to the complete dataset. Descriptive statistics for all incident and BQE incident durations are shown in Table 4 and Table 9. However, unlike the overall case, for BQE, Lognormal distribution is found to capture the incident distributions, with the exception of “Property Damage” incident type. Gamma fit for “Disabled Vehicle” incident type succeeds to pass both tests since the BQE duration data do not contain very large values, which results in a smooth tail (see Figure 4). However, in overall, the Anderson-Darling test performances are poor, again because of the tails. This can be observed in Figure 5. BQE overall incident data can be modeled by all three distributions, with the error of badly predicting the tails. Among all, Gamma distribution is slightly better than the other two since it passed A-D test for one incident category.

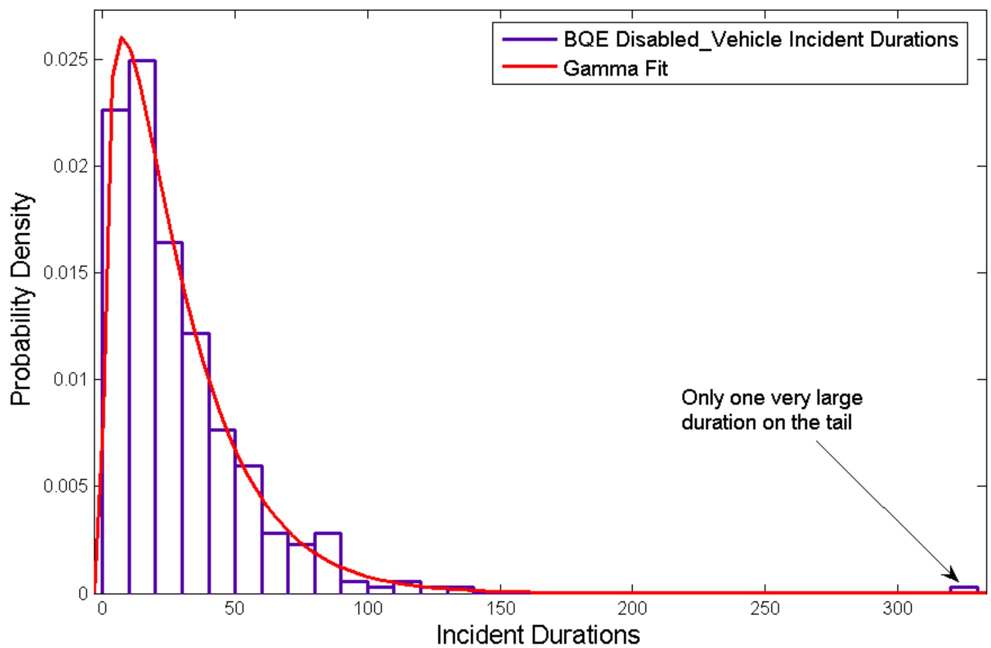


Figure 4- Gamma Distribution Fit for BQE Disabled Vehicle Incident Durations

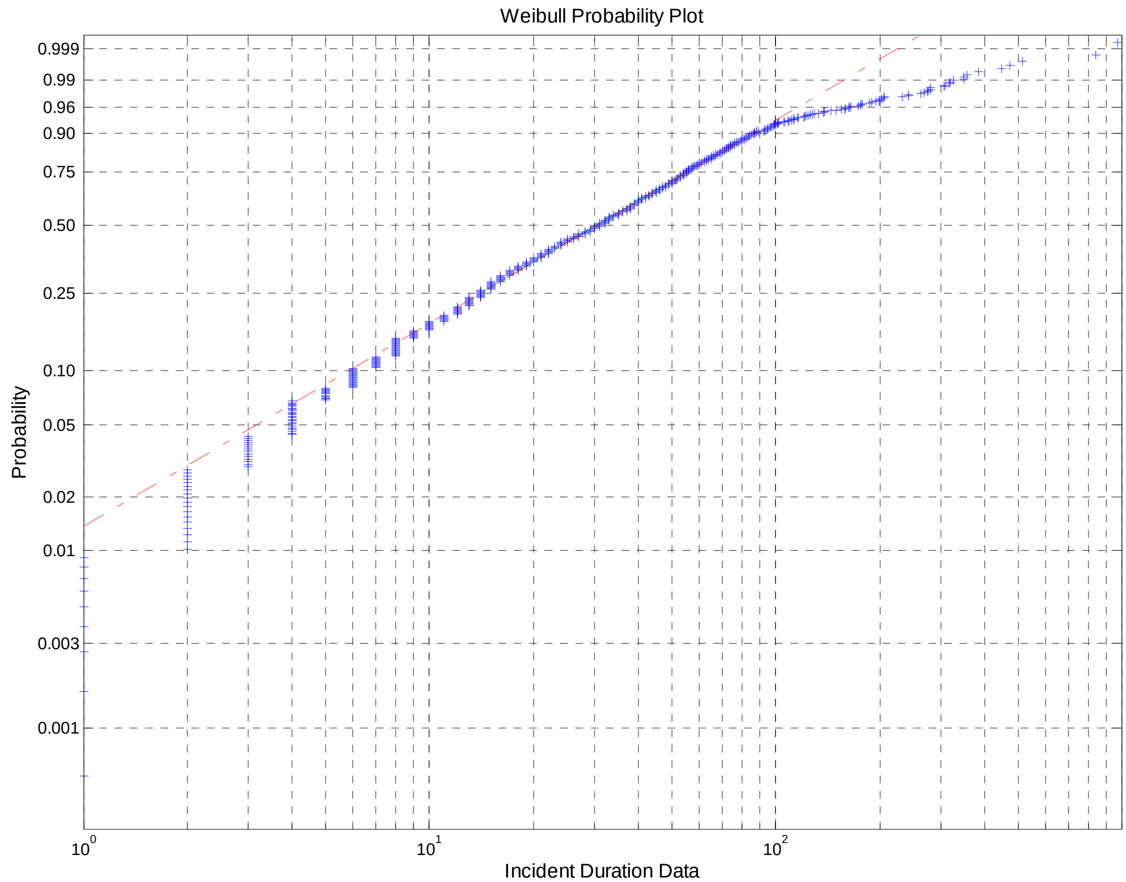


Figure 5- Weibull Probability Plot for BQE Incident Durations (legend, x and y coordinates)

Analysis of Gowanus Expressway (GE) Incident Durations

The descriptive incident duration statistics for all types of incidents can be found in detail in Table 13. Histogram of durations can be seen in Figure 6. Compared to both overall and BQE statistics (Table 4 and Table 9), GE mean duration for every incident type is lower, with also a lower standard deviation. This reduced variability can be detected in A-D test results below. With decreasing tail size, for some incident types, tested distributions pass A-D tests at 95% confidence level.

Table 13- Descriptive Statistics of Incident Durations for GE

Incident Type	# of Data Points	Mean [min]	Std Deviation [min]	95% Duration [min]	Max Duration [min]
Property Damage	243	37	38	100	257
Disabled Vehicle	262	24	21	64	158
Disabled Truck	70	35	27	84	119
Total*	634	37	47	98	356

*Road Hazard, Weather Related, HAZMAT, Personal Injury and Vehicle Fire incident types are included

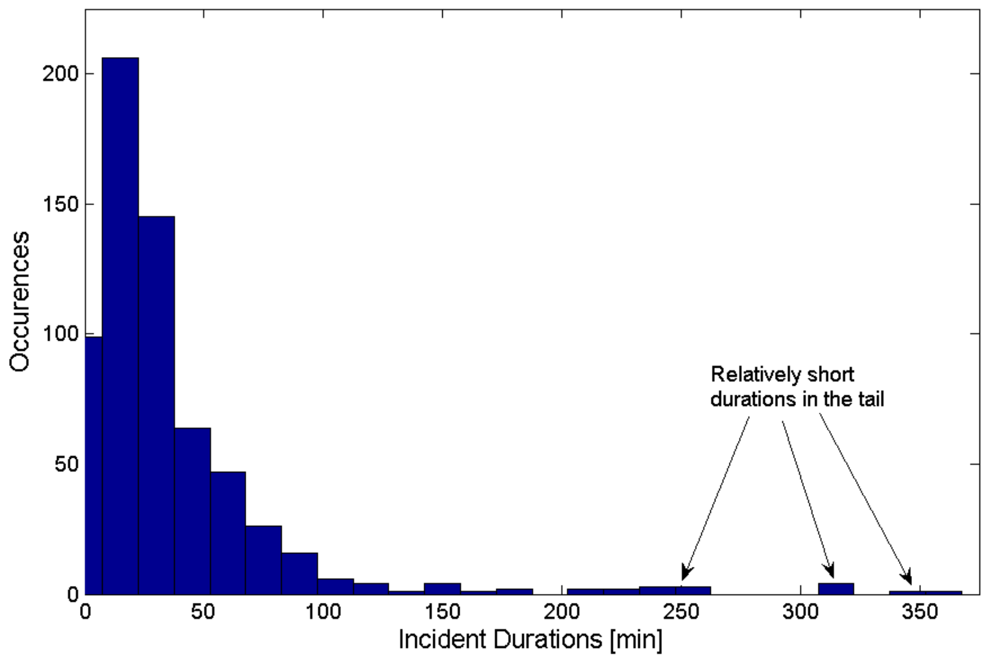


Figure 6- Histogram of Incident Durations for GE-All incident Types

The summary of the KS and AD tests for Weibull, Gamma and Lognormal distributions with relevant GOF statistics are shown in Table 14, Table 15, and Table 16, respectively. All tests are performed at 95% confidence level.

Table 14- Weibull Fit and GOF Test Statistics at 95% Confidence Level for Overall GE Incident Duration Data

Incident Type	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Good Fit	Test Statistics	Critical Value	Good Fit
Property Damage	37.8880 1.0940	0.0661	0.0864	**Yes**	1.4465	0.757	No
Disabled Vehicle	25.7526 1.2226	0.0631	0.0832	**Yes**	0.9605	0.757	No
Disabled Truck	37.1238 1.2610	0.0535	0.1598	**Yes**	0.2369	0.757	**Yes**
All Incidents	36.3614 0.9842	0.0702	0.0537	No	6.5059	0.757	No

Table 15- Gamma Fit and GOF Test Statistics at 95% Confidence Level for GE Incident Duration Data

Incident Type	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Fit	Test Statistics	Critical Value	Good Fit
Property Damage	1.2539 29.1181	0.0625	0.0864	**Yes**	1.0089	0.844	No
Disabled Vehicle	1.4751 16.2803	0.0536	0.0832	**Yes**	0.6840	0.844	**Yes**
Disabled Truck	1.4347 24.0562	0.0554	0.1598	**Yes**	0.2645	0.844	**Yes**
All Incidents	1.0861 33.7469	0.0787	0.0537	No	5.9178	0.844	No

Table 16- Lognormal Fit and GOF Test Statistics at 95% Confidence Level for GE Incident Duration Data

Incident Type	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Good Fit	Test Statistics	Critical Value	Good Fit
Property Damage	3.1485 1.0330	0.0845	0.0864	**Yes**	2.1342	0.735	No
Disabled Vehicle	2.8029 0.9353	0.0754	0.0832	**Yes**	1.4221	0.735	No
Disabled Truck	3.1540 1.0157	0.1095	0.1598	**Yes**	1.1905	0.735	No
All Incidents	3.0750 1.0662	0.0653	0.0537	No	2.7286	0.735	No

As mentioned before, decreasing standard deviation makes it possible for the proposed distributions to pass A-D test at 95% confidence level. In this respect, since all the distributions are same in terms of K-S test performance, Gamma distribution can be mentioned as being a better choice superior for modeling the overall incident duration distributions for GE.

<u>Analysis of Staten Island Expressway (SIE) Incident Durations</u>

The descriptive incident duration statistics for all types of incidents can be seen in detail in Table 17 and can be visually inspected in Figure 7. The mean incident durations are higher compared to other facilities discussed above, and standard deviations are also larger. However, the GOF statistics show that for the first time, there is a good fit for overall facility incidents, whereas in previous facilities no good fit could be achieved for overall incident durations. The reason for this can be the maximum durations of each incident type being similar and not introducing a major change when the other incident types (Road Hazard, Weather Related, HAZMAT, Personal Injury and Vehicle Fire) in the overall data set. Also note that the small number of data points makes it hard for a reliable conclusion, especially while studying individual incident types. For example, in “Disabled Truck” category, there are only 8 data points, making it unnecessary to test.

Table 17- Descriptive Statistics of Incident Durations for SIE

Incident Type	**# of Data Points**	**Mean [min]**	**Std Deviation [min]**	**95% Duration [min]**	**Max Duration [min]**
Property Damage	63	63	57	141	342
Disabled Vehicle	16	75	77	88	327
Disabled Truck	8	66	39	87	139
Total*	110	88	84	240	363

*Road Hazard, Weather Related, HAZMAT, Personal Injury and Vehicle Fire incident types are included

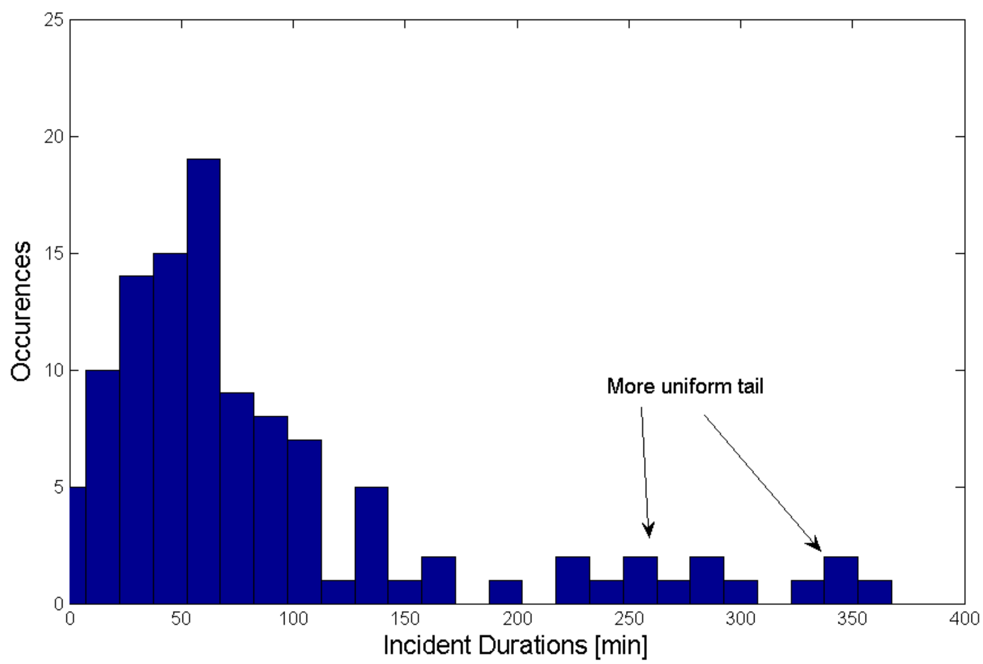


Figure 7- Histogram of Incident Durations for SIE-All incident Types

The summary of the K-S and A-D tests for Weibull, Gamma and Lognormal distributions with relevant GOF statistics are presented in Table 18, Table 19, and Table 20, respectively. All tests are performed at 95% confidence level.

Table 18- Weibull Fit and GOF Test Statistics at 95% Confidence Level for SIE Incident Duration Data

Incident Type	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Good Fit	Test Statistics	Critical Value	Good Fit
Property Damage	66.5749 1.1864	0.0786	0.1682	**Yes**	0.4842	0.757	**Yes**
Disabled Vehicle	76.9955 1.0577	0.2673	0.3273	**Yes**	1.2653	0.757	No
Disabled Truck	73.5784 1.8148	0.2591	0.4543	**Yes**	0.4108	0.757	**Yes**
All Incidents	90.7912 1.0972	0.1032	0.1279	**Yes**	1.4336	0.757	No

Table 19- Gamma Fit and GOF Test Statistics at 95% Confidence Level for SIE Incident Duration Data

Incident Type	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Fit	Test Statistics	Critical Value	Good Fit
Property Damage	1.3670 45.8643	0.0715	0.1682	**Yes**	0.4266	0.844	**Yes**
Disabled Vehicle	1.0672 70.6291	0.2736	0.3273	**Yes**	1.2787	0.844	No
Disabled Truck	2.4078 27.2550	0.2921	0.4543	**Yes**	0.4720	0.844	**Yes**
All Incidents	1.1965 73.1852	0.0995	0.1279	**Yes**	1.3161	0.844	No

Table 20- Lognormal Fit and GOF Test Statistics at 95% Confidence Level for SIE Incident Duration Data

Incident Type	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Good Fit	Test Statistics	Critical Value	Good Fit
Property Damage	3.7300 1.0574	0.1194	0.1682	**Yes**	1.4515	0.735	No
Disabled Vehicle	3.7856 1.3873	0.3371	0.3273	No	1.8811	0.735	No
Disabled Truck	3.9622 0.8070	0.3231	0.4543	**Yes**	0.5928	0.735	**Yes**
All Incidents	3.9995 1.1330	0.1203	0.1279	**Yes**	2.2689	0.735	No

Both Gamma and Weibull distributions can be selected because of their good performance with respect to K-S, and fairly good performance with respect to A-D test. However, as mentioned before, due to scarcity of data points is not possible to make a decisive conclusion.

<u>Analysis of Bruckner Expressway (BE) Incident Durations</u>

The descriptive incident duration statistics for all types of incidents can be found in detail in Table 21. The histogram of incident duration distribution can be seen in Figure 8. BE poses the hardest challenge in terms of distribution choice since GOF measures changes considerably for each distribution. However, just like SIE, number of data points is not sufficient to draw reliable conclusions.

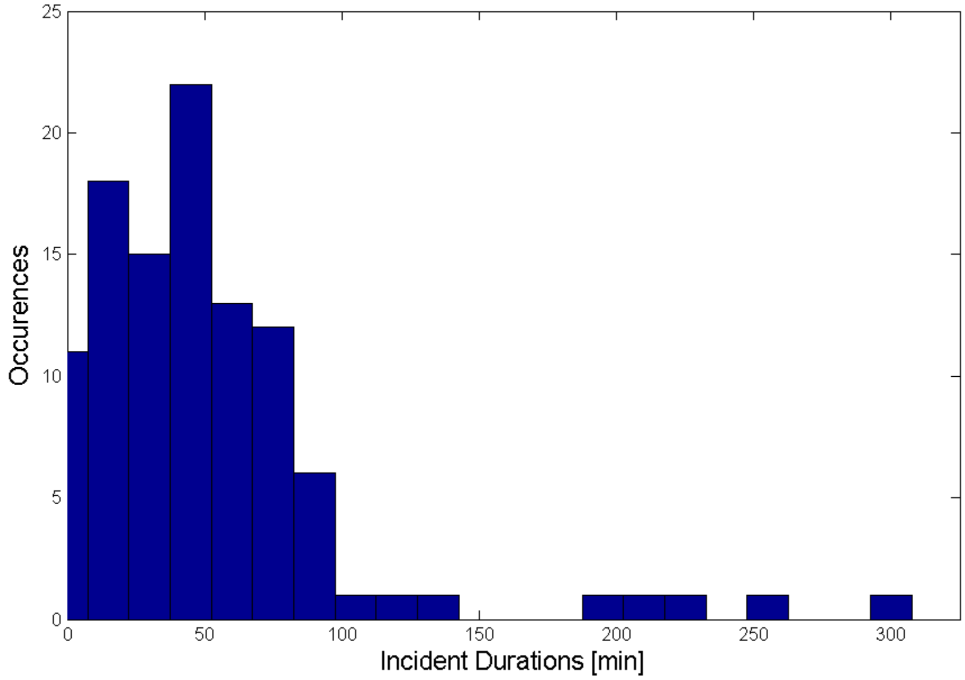


Figure 8- Histogram of Incident Durations for BE-All incident Types

Table 21- Descriptive Statistics of Incident Durations for BE

Incident Type	# of Data Points	Mean [min]	Std Deviation [min]	95% Duration [min]	Max Duration [min]
Property Damage	64	47	38	94	230
Disabled Vehicle	23	39	27	81	85
Disabled Truck	6	120	232	52	591
Total*	107	66	111	100	933

*Road Hazard, Weather Related, HAZMAT, Personal Injury and Vehicle Fire incident types are included

The summary of the K-S and A-D tests for Weibull, Gamma and Lognormal distributions with relevant GOF statistics are presented in Table 22, Table 23, and Table 24 respectively. All the tests are performed within 95% confidence level.

Table 22- Weibull Fit and GOF Test Statistics at 95% Confidence Level for BE Incident Duration Data

Incident Type	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Good Fit	Test Statistics	Critical Value	Good Fit
Property Damage	50.8756 1.3105	0.0577	0.1669	**Yes**	0.2601	0.757	**Yes**
Disabled Vehicle	41.7836 1.3065	0.1262	0.2749	**Yes**	0.4570	0.757	**Yes**
Disabled Truck	60.7477 0.5186	0.2308	0.5193	**Yes**	0.2862	0.757	**Yes**
All Incidents	61.1024 0.8914	0.1349	0.1296	No	2.4397	0.757	No

Table 23- Gamma Fit and GOF Test Statistics at 95% Confidence Level for BE Incident Duration Data

Incident Type	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Fit	Test Statistics	Critical Value	Good Fit
Property Damage	1.6005 29.2584	0.0684	0.1669	**Yes**	0.2811	0.844	**Yes**
Disabled Vehicle	1.3598 28.5536	0.1459	0.2749	**Yes**	0.4899	0.844	**Yes**
Disabled Truck	0.3862 310.3014	0.2936	0.5193	**Yes**	0.4037	0.844	**Yes**
All Incidents	0.9302 70.4277	0.1446	0.1296	No	2.3955	0.844	No

Table 24- Lognormal Fit and GOF Test Statistics at 95% Confidence Level for BE Incident Duration Data

Incident Type	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Good Fit	Test Statistics	Critical Value	Good Fit
Property Damage	3.5026 0.9407	0.1195	0.1669	**Yes**	1.1506	0.735	No
Disabled Vehicle	3.2483 1.1247	0.1824	0.2749	**Yes**	0.9590	0.735	No
Disabled Truck	3.0725 2.2700	0.1827	0.5193	**Yes**	0.2272	0.735	**Yes**
All Incidents	3.5564 1.1712	0.1167	0.1296	**Yes**	2.1914	0.735	No

5.0 Further Discussion about the Duration Distribution

In this section, two more aspects of the estimation of probability distributions will be discussed. One of them is the analysis of incident durations according to number of lanes at the incident location, and the other is the determination of the duration distribution according to the number of lanes closed as a consequence of an incident.

Incident Duration Distribution According To the Number Of Lanes In The Incident Segment

Investigation of incident durations for specific road segments is not always feasible since detailed section by section data may not be available for every segment, and it makes both the analysis procedure and any policy treatment of the system more cumbersome. Thus general road characteristics, such as number of lanes on a roadway, should be investigated for their affect on incident duration. For this purpose, the incident data is sorted according to the number of lanes of the segment over which the incident had occurred. Table 21 shows the corresponding GOF results for incident duration distribution according to the number of lanes at the incident location.

Please note that in the current dataset there are officially no segments with 4 lanes in case of the facilities of interest. However, with the help of aerial photos, some incident segments are found to have 4 lanes. These are either merging or diverging sections. Although these subjective additions are prone to errors, they enrich the existing dataset. It should also be remembered that the existing data is also collected and recorded by human operators that are already prone to commit perception errors.

Table 25- Descriptive Statistics of Incident Durations According to Number of Lanes at the Incident Location

# of Lanes	# of Data Points	Mean [min]	Std Deviation [min]	95% Duration [min]	Max Duration [min]
2-Lane Section	104	37	35	86	181
3-Lane Section	1740	48	69	150	969
4-Lane Section	54	42	48	88	376

The summary of the K-S and A-D tests for Weibull, Gamma and Lognormal distributions with relevant GOF statistics are presented in Table 26, Table 27, and Table 28, respectively. All the tests are performed at 95% confidence level.

Table 26- Weibull Fit and GOF Test Statistics at 95% Confidence Level for Complete Incident Duration Data

# of Lanes	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Good Fit	Test Statistics	Critical Value	Good Fit
2-Lane Section	42.9488 1.0693	0.0662	0.1315	**Yes**	0.8397	0.757	No
3-Lane Section	46.2751 0.9343	0.0626	0.0325	No	16.1867	0.757	No
4-Lane Section	39.0497 1.1770	0.1148	0.1814	**Yes**	0.7139	0.757	**Yes**

Table 27- Gamma Fit and GOF Test Statistics at 95% Confidence Level for Complete Incident Duration Data

# of Lanes	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Fit	Test Statistics	Critical Value	Good Fit
2-Lane Section	1.2107 34.4756	0.0604	0.1315	**Yes**	0.6141	0.844	**Yes**
3-Lane Section	0.9945 48.2869	0.0683	0.0325	No	16.4314	0.844	No
4-Lane Section	1.4253 25.7781	0.1129	0.1814	**Yes**	0.5947	0.844	**Yes**

Table 28- Lognormal Fit and GOF Test Statistics at 95% Confidence Level for Complete Incident Duration Data

# of Lanes	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Good Fit	Test Statistics	Critical Value	Good Fit
2-Lane Section	3.2647 1.0832	1.6994	0.0956	**Yes**	1.6994	0.735	No
3-Lane Section	3.2908 1.1181	6.9959	0.0453	No	6.9959	0.735	No
4-Lane Section	3.2137 0.9548	0.4570	0.1112	**Yes**	0.4570	0.735	**Yes**

There are 104, 1740 and 54 data points for 2, 3 and 4 lane road segments, respectively. Naturally, 3-lane segment distribution governs the overall behavior, as it constitutes the main body of the dataset. However, none of the distributions succeed to pass 95% confidence level tests for 3 lane sections. For 2 and 4 lane sections, gamma distribution fits the distribution best, passing both tests, where Weibull and Lognormal distributions fail the A-D test. Nevertheless, this fit represents only 8% of the overall data. In conclusion, number of lanes of the lane segments cannot be used as a part of the location specific analysis.

Incident Duration Distribution According To Number Of Lanes Blocked

In the literature, the number of blocked lanes during an incident is mentioned to be one of the components of non-recurrent delays. For this analysis, the incident data is sorted according to the number of lanes that are closed, and relevant durations resulting in corresponding lane blockages are computed. Table 29 shows the descriptive statistics for the data, and GOF statistics are presented in Table 30, Table 31, and Table 32 for Weibull, Gamma and Lognormal distributions, respectively. As can be seen in Table 29, data is mainly governed by 1-lane blockage, having a very large maximum value compared to its mean and standard deviation.

Table 29- Descriptive Statistics of Complete Incident Duration Data According to Number of Lanes at the Incident Location

# of Lanes	# of Data Points	Mean [min]	Std Deviation [min]	95% Duration [min]	Max Duration [min]
1-Lane Blocked	1440	39	50	121	839
2-Lane Blocked	223	54	61	139	591
3-Lane Blocked	50	45	57	108	326

Table 30- Weibull Fit and GOF Statistics at 95% Confidence Level for Complete Incident Duration Data

# of Lanes	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Good Fit	Test Statistics	Critical Value	Good Fit
1-Lane Blocked	38.3755 0.9918	0.0432	0.0357	No	7.6165	0.757	No
2-Lane Blocked	56.0561 1.0805	0.0733	0.0902	**Yes**	1.8786	0.757	No
3-Lane Blocked	46.1162 1.0320	0.1108	0.1884	**Yes**	1.0951	0.757	No

Table 31- Gamma Fit and GOF Statistics at 95% Confidence Level for Complete Incident Duration Data

# of Lanes	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Fit	Test Statistics	Critical Value	Good Fit
1-Lane Blocked	1.0843 35.5360	0.0419	0.0357	No	5.7901	0.844	No
2-Lane Blocked	1.2469 43.5024	0.0677	0.0902	**Yes**	1.3336	0.844	No
3-Lane Blocked	1.2197 37.2383	0.1238	0.1884	**Yes**	0.9749	0.844	No

Table 32- Lognormal Fit and GOF Statistics at 95% Confidence Level for Complete Incident Duration Data

# of Lanes	Distribution Parameters	Kolmogorov-Smirnov Test			Anderson-Darling Test		
		Test Statistics	Critical Value	Good Fit	Test Statistics	Critical Value	Good Fit
1-Lane Blocked	3.1240 1.1065	0.0507	0.0357	No	8.3161	0.735	No
2-Lane Blocked	3.5416 1.0365	0.1010	0.0902	No	3.0259	0.735	No
3-Lane Blocked	3.3530 0.9696	0.0766	0.1884	**Yes**	0.3215	0.735	**Yes**

The results show that only the Lognormal distribution passes both tests for 3 lane segments, and for the rest of data there is no good fit. 1-lane blockage, which needs special attention, cannot be fitted using any of the distributions.

Analysis of Incident Durations According to Physical Road Characteristics and Weather/Pavement Conditions

In the complete data set, there are several physical features of the road section:

- Lane width
- Shoulder existence
- Pavement conditions
- Weather conditions

Among those 4 factors above, lane width and shoulder existence exist in the original data set, where as longitudinal and lateral characteristics are added to the data. Lane width is almost uniform for all road sections in the study network, regardless of the facility. There are only 16 entries for 11 feet and 34 entries for 10 feet over total of 1907 entries. Rest is 12 feet. Thus no analysis was performed in terms of lane width. Shoulder width, on the other hand, can be analyzed for its effect on duration. However analysis cannot be performed for each facility because some facilities either have shoulder or not. There are very few exceptions in the facility-based data having both with-shoulder section incidents and without-shoulder section incidents. There are a total number of only 193 incident records with no shoulder over 1907 data points. Please note that the shoulder existence information is based on the facility that the incident had occurred. Although some partial shoulders exist in some portions of the facilities, the existence of shoulder is based on "Highway Sufficiency Rating File" and those partial shoulder/refuges are neglected. In the current analysis, the official records given in TRANSCOM dataset are used and no further personal judgments about the shoulder existence were employed. Pavement and weather condition data do not exist for all records. There are 661 missing records for pavement conditions and 656 missing records for weather conditions. Table 29 shows the descriptive statistics for incidents with shoulder information and Table 34 shows duration data analyzed according to pavement and weather conditions.

Table 33- Descriptive Statistics of Complete Incident Duration Data According to Existence of Shoulder

Shoulder	# of Data Points	Mean [min]	Std Deviation [min]	95% Duration [min]	Max Duration [min]
Shoulder Exists	193	44	78	231	402
No Shoulder	1714	75	66	121	968

Table 34- Descriptive Statistics of Complete Incident Duration Data According to Weather and Pavement Conditions

		# of Data Points	Mean [min]	Std Deviation [min]	Max Duration [min]
Pavement Conditions	Dry	1018	44	60	932
	Wet	205	36	41	276
	Other*	25	60**	75**	968
Weather Conditions	Clear	898	42	51	448
	Cloudy	35	36	45	240
	Rain	54	25	21	72
	Light Rain	49	41	39	158
	Heavy Rain	15	59	53	205
	Other***	32	35	31	136

* Slick, Snow Covered, Slushy, Flooding, Icy

** Mean and standard deviation is found by excluding the max value only.

*** Foggy, Flurries, Overcast, Sleet, Snow, Heavy Wind

Incidents at sections with shoulder show a smaller mean value for duration compared to no-shoulder sections. This can be due to the fact that sometimes the incident can be moved to the shoulder and assumed to be cleared. Also, sections with shoulders allow police and emergency cars to reach the area more quickly. Sections with no shoulder show less variability in duration. This can be explained with the fact that longer durations are more tolerable for places with shoulders since the traffic can flow with a relatively small disturbance if the incident is moved to shoulders. However, if there is no shoulder the incident should be completely cleared as soon as possible to prevent major delay due to lane closure. Overall, incidents at sections with shoulder are accessed easily and their official clearance time can be anywhere between time the incident is carried to shoulder and completely removed from the incident scene. On the other hand, no shoulder sections make it hard to access the incident site and clearance times are more exact in the sense that when an incident is cleared, no sign of it stays in the area. Nevertheless, the resulting delay for a short duration incident in a no-shoulder section can still be larger than a long incident occurred compared with a section with shoulder. Thus, shoulder information needs more attention for the delay analysis.

Regarding the pavement and weather conditions, the results do not exhibit significant and consistent differences. The mean duration for heavy rain is larger than the light rain. However, light rain shows a higher mean duration compared to "rain". On the other hand, average duration is less for rainy weather compared to clear weather. Since pavement conditions are related to weather conditions, similar interpretations can be made for pavement conditions. Overall, pavement and weather conditions do not change the durations significantly.

Analysis of Incident Durations According to Time of Day and Weekday/Weekend

Table 35 shows the descriptive statistics of incident durations at different times of week/weekend days. One thing to be noted is, although number of records is all the entries from the data set, zero-duration incidents are eliminated from the data while calculating mean and standard deviation.

Table 35- Descriptive Statistics of Complete Incident Duration Data According to Time of Day and Weekday/Weekend

		# of Records	Mean [min]	Std Deviation [min]	Max Duration [min]
Day	**Weekday**	1431	47	70	932
	Weekend	476	47	60	276
Time of Day	**Morning Peak (6 AM - 10 AM)**	405	53	83	326
	Evening Peak (3PM - 7PM)	451	36	40	839
	Off-Peak (10AM - 3PM)	470	54	63	402
	Evening (7PM - 12AM)	356	37	61	968
	Night (12AM - 6AM)	225	58	90	932

One quick fact that can be extracted from Table 35 is that week and weekend days do not affect the duration. Both categories exhibit the same mean value with very close standard deviations. Likewise, the incident occurrence rates are also close for weekend and weekdays. So it can be concluded that weekday and weekend variation does not contribute to both duration statistics. On the contrary, durations differ slightly for time of day. However, there is no clear pattern for different TOD periods. Morning and evening peaks behave completely differently. Morning peaks have 56 minutes of average duration whereas evening peak average is 36 minutes. Afternoon off-peak period average is only a minute more than morning peak, but with a smaller variance. Incidents between 12AM – 6AM show the largest average and standard deviation. This is reasonable since there are probably less officers working during these times. The difference between morning and evening peak needs more detailed information. Nevertheless, distributions were fitted for each TOD period and the good fits for each period are shown in Table 36. Please note that no distribution was able to pass Anderson-Darling test (which means poor representation for the tail) and good fit conclusions are only based on Kolmogorov-Smirnov test.

Table 36- Proposed Distributions for TOD Periods

	Time of Day					
	Morning Peak (6 AM - 10 AM)	**Evening Peak (3PM – 7PM)**	**Off-Peak (10AM - 3PM)**	**Evening (7PM - 12AM)**	**Night (12AM - 6AM)**	**Morning Peak (6 AM - 10 AM)**
Good Fits	Log norm$^{(3.34, 1.15)}$	Weibull$^{(36.69, 1.01)}$ Gamma$^{(1.07, 33.94)}$	Lognorm$^{(3.45, 1.08)}$	Weibull$^{(36.08, 0.94)}$ Gamma$^{(1.01, 36.72)}$ Lognorm$^{(3.05, 1.14)}$	Lognorm$^{(3.50, 1.07)}$	Lognorm$^{(3.34, 1.15)}$

Unlike facility specific analysis, Lognormal distribution, overall, performs better than Weibull and Gamma distributions. The distribution parameters are mostly distinct. The reason is that as can be seen in Table 35, either average duration or the standard deviation is similar for two or more periods, but never both. Consequently, fitted curves also differ.

6.0 Conclusions and Discussion

The selected probability functions with estimated distribution parameters for each facility and incident type are shown in Table 37. The distributions that passed more tests were recommended as the distibutions to be used and if there is a tie in terms of number of tests passed, both distributions are included in the final list. For some data sets, none of the candidate distributions could pass the K-S and A-D tests, e.g. BQE-All Incidents, or number of data points were not sufficient to draw a statistically reliable conclusion, e.g. SIE-Disabled Truck.

Table 37- Proposed Distributions for All Facilities and Incident Types

	BQE	GE	SIE	BE	All Facilities
Property Damage	Weibull$^{(44.46, 1.08)}$ Gamma$^{(1.24, 34.67)}$	Weibull$^{(37.89, 1.09)}$ Gamma$^{(1.25, 29.12)}$ Lognorm$^{(3.15, 1.03)}$	Weibull$^{(66.57, 1.19)}$ Gamma$^{(1.37, 45.86)}$	Weibull$^{(50.88, 1.31)}$ Gamma$^{(1.60, 29.26)}$	Weibull$^{(44.50, 1.10)}$ Gamma$^{(1.01, 46.86)}$
Disabled Vehicle	Gamma$^{(1.35, 21.04)}$	Gamma$^{(1.48, 16.28)}$	*Insufficient number of data*	*Insufficient number of data*	Weibull$^{(29.50, 1.13)}$ Gamma$^{(1.04, 50.52)}$
Disabled Truck	Weibull$^{(57.03, 0.91)}$ Gamma$^{(0.99, 60.77)}$ Lognorm$^{(3.52, 1.13)}$	Weibull$^{(37.12, 1.26)}$ Gamma$^{(3.15, 1.02)}$	*Insufficient number of data*	*Insufficient number of data*	Weibull$^{(50.35, 0.93)}$ Gamma$^{(1.32, 21.38)}$ Lognorm$^{(2.91, 0.99)}$
All Incidents	N/A	N/A	Weibull$^{(90.79, 1.10)}$ Gamma$^{(1.20, 73.19)}$ Lognorm$^{(4.00, 1.13)}$	Lognorm$^{(3.56, 1.17)}$	N/A

Although some common distributions can be assumed for an incident type or facility, the distribution parameters are not consistent for all facility and incident types. Thus, one important conclusion of the analysis can be that duration distributions should be defined based on the facility (or location) and incident type. Unfortunately a single probability distribution that models all cases reliably could not be identified.

One way to unify the distributions for facility or incident types can be the use of other statistical tests to identify - if there is any- the existence outliers in the data. This may help in the analysis by giving better distribution fits, as well as allowing more uniform results in terms of duration distributions. Basically, a kind of filtering based on statistical reasoning can be performed on data for improved results.

Apart from the facility and incident type specific duration analysis, no statistically significant distribution that represents the distribution of incident duration according to resulting lane blockage or incident lane location could be identified.

Regarding the factors affecting the incident duration, facility and incident type can be said to be 2 significant factors. Among the road physical characteristics, having a shoulder or not is also found to affect the incident duration. Incidents occurred at sections with shoulder do not last as long as no-shoulder sections. Durations for different pavement and weather conditions do not exhibit much variation, thus claimed not to have significant effect on duration. Lane width was found to be the same for almost all sections and excluded from the analysis.

Weekday/weekend affect and Time of Day (TOD) was also analyzed and it can be said that weekday and weekend does not affect the incident durations. However incident durations show distinct values for different TOD periods. No general pattern was found between peak/off-peak and long/short incident durations.

OVERALL CONCLUSIONS AND DISCUSSION FOR INCIDENT DURATION ANALYSIS OF THE TRANSCOM DATA

The major results of duration analyses performed in Task 3.2 can be summarized as follows:

- TRANSCOM dataset contains only 8% of all the accidents in the region which makes it a clearly incomplete and inadequate dataset for frequency analysis and duration analysis to a lesser degree. Even for duration analysis, it is important to understand that the sample provided by TRANSCOM can be highly biased in terms of certain accident types since the research team did not have any input in creating this sample. For example, TRANSCOM data might contain certain types of accidents that are not relevant to their operations and this might bias the results of the data analysis. Thus, all the conclusions listed below have to be interpreted in the light of this major shortcoming related to the quality of the accident data used in this study.
- A single distribution is not found to represent all the incident durations.
    - Weibull, lognormal and gamma distributions yield good fit for different cases.
    - Lognormal distribution is found to be less representative of the data compared to the results of other studies in the literature
- Some factors affecting incident duration was found by descriptive analysis
    - Incident Type: Different types of incidents yield different distributional properties, generally consistent with the findings in the literature.
    - Shoulder Existence: Sections with shoulder are found to have less average durations, but larger standard deviations
    - Time of Day (TOD): Although no significant relationship between durations and TOD could be determined, TOD can be mentioned to affect the durations slightly.

REFERENCES

1. Nam, Doohee and Fred Mannering. "An Exploratory Hazard-Based Analysis of Highway Incident Duration." *Transportation Research – A*. Vol. 34A, No. 2, 2000. pp 85-102.
2. Wang W., Chen H., and Bell M.C., “Vehicle Breakdown Duration Modeling”, Journal of Transportation and Statistics, Vol.8, No.1. Can be accessed from: http://www.bts.gov/publications/journal_of_transportation_and_statistics/volume_08_number_01/html/paper_06/
3. Kevin Smith and Brian L. Smith, “Forecasting the Clearance Time of Freeway Accidents”, Final report of ITS Center project: Incident duration forecasting A Research Project Report, 2001.
4. *NIST/SEMATECH e-Handbook of Statistical Methods*, http://www.itl.nist.gov/div898/handbook/*, Accessed at 12/26/2005*
5. Agostino Tarsitano, “Fitting Wakeby model using maximum likelihood”, *Working Paper*, Dipartimento di Economia e Statistica - Università degli Studi della Calabria
6. Kent E. Cordes and Rolliri H. Hotchkiss, “Design Discharge of Culverts”, *NDOR* 1 *Research Project Number RES-1 (0099)*, University of Nebraska, December 1993.
7. Kevin Smith and Brian L. Smith, “Forecasting the Clearance Times of Freeway Accidents”, *Final report of ITS Center project: Incident duration forecasting*, Center for Transportation Studies, University of Virginia, 2001.
8. Golob, Thomas F., Wilfred W. Recker and John D. Leonard. “An Analysis of the Severity and Incident Duration of Truck-Involved Freeway Accidents.” *Accident Analysis & Prevention*. Vol. 19, No. 4, August 1987. pp 375-395.
9. Garib, A., A. E. Radwan and H. Al-Deek. “Estimating Magnitude and Duration of Incident Delays.” *Journal of Transportation Engineering*. Vol. 123, No. 6, November/December 1997. pp. 459-466.
10. Giuliano, Genevieve. “Incident Characteristics, Frequency, and Duration on a High Volume Urban Freeway.” *Transportation Research – A*. Vol. 23A, No. 5, 1989. pp. 387-396.
11. Nam, Doohee and Fred Mannering. “An Exploratory Hazard-Based Analysis of Highway Incident Duration.” *Transportation Research – A*. Vol. 34A, No. 2, 2000. pp 85-102.

12. Kaan Ozbay and Pushkin Kachroo. *Incident Management in Intelligent Transportation Systems*. Boston, MA: Artech House, 1999.

13. Lasse Fridstrom, Jan Ifver, Siv Ingebrigtsen, Risto Kulmala, and Lars Krogsgard Thomsen, "Measuring The Contribution Of Randomness, Exposure, Weather, And Daylight To The Variation In Road Accident Counts", *Accident Analysis and Prevention*, Vol.27 (1), (1995), pp.1-20.

14. Venkataraman Shankar, Fred Mannering, and Woodrow Barfield, "Effect Of Roadway Geometrics And Environmental Factors On Rural Freeway Accident Frequencies", *Accident Analysis and Prevention*, Vol.27 (3), (1995), pp.371-389.

15. K.W. Ogden, "The Effects of Paved Shoulders on Accident on Rural Highways", *Accident Analysis and Prevention*, Vol.29 (3), (1997), pp.353-362.

16. Charlesworth, K.D., Survey of Shoulder Usage. *Australian Road Research Board Internal Report 404-4*, p. 43. ARRB, Melbourne, 1985.

17. Jean-Louis Martin, "Relationship between crash rate and hourly traffic flow on interurban motorways", *Accident Analysis and Prevention*, Vol.34, (2002), pp.619-629.

18. Thomas F. Golob, Wilfred W. Recker, "A method for relating type of crash to traffic flow characteristics on urban freeways", *Transportation Research Part A*, Vol.38, (2004), pp.53–80.

19. Jovanis P & Chang H, "Modeling the relationship of accidents to miles traveled", *Transportation Research Record* 1068, (1986), pp. 42–51.

20. John Milton And Fred Mannering, "The Relationship Among Highway Geometrics, Traffic-Related Elements And Motor-Vehicle Accident Frequencies", *Transportation*, Vol.25, (1998), pp. 395–413.

APPENDIX

INCIDENT FREQUENCY ANALYSIS

In this part, first, general data statistics are given, partly referring to Task 3.1. Those statistics are presented according to facility, incident type, road section and weather characteristics, just like the duration analysis. Please note that, as mentioned under "Some Important Remarks about the Available TRANSCOM Dataset" section of this report, the frequency analysis presented below is an analysis based on a small sample (about 8% of the total accidents) of the overall accidents in the study area. Hence, accident rates or factors affecting the accident rates cannot be used to make final inferences about the overall accident statistics for the study area. The results of this study are strictly restricted to the available TRANSCOM records and should not be used as final recommendations of the study.

Descriptive Analysis Of Incident Frequencies

Some general descriptive statistics for incident frequencies are presented in Task 3.1. Number of occurences for different cases are identified, and further, some obvious incident spots are also determined. In this section further descriptive statistics will be presented and further interpretation of the data pertinent to incident frequencies will be given.

Analysis of Incident Frequencies According to the Facility

Table 6.2 of Task 3.1 gives the incident occurrence statistics for the facilities of interest. More specific information about the black spots is also mentioned. Analysis shows that there are more incidents at some locations compared with other locations. The Brooklyn Queens Expressway/I-278 has the highest number of recorded non-recurring incidents in the vicinity of the Atlantic Avenue, Hamilton Avenue, and Kosciusko Bridge with 15.1%, 10.3% and 8.2%, respectively followed by Gowanus Expressway/I-278 in the vicinity of the 39th Street, Prospect Expressway, and Gowanus Canal with 8.1%, 6.4% and 6.0% of incidents, respectively. In the database, incident frequencies for 7 different facilities are recorded, as shown in Table A-1 that is a slightly modified version of Table 6.2 with speed limits and number of exits of each facility.

Table A-1 Facility Specific Incident Frequency Statistics

(Based on Table 6.2 of Task 3.1)

Facility Name	# of Recorded Incidents	%	Daily VMT	Total VMT	Incident Rate per 1,000,000 VMT*	Facility's Length	Incident Rate per Day**	Incident Rate per Mile per Day	Speed Limit (mph)	# of Exits
Brooklyn Queens Expressway /I-278	941	49.3%	1,300,057	338,014,820	2.78	12.76	2.22	0.17	45	16
Gowanus Expressway /I-278	637	33.4%	914,686	237,818,360	2.68	6.96	1.50	0.22	50	10
Staten Island Expressway /I-278	110	5.8%	1,176,976	306,013,760	0.36	8.88	0.26	0.03	50	13
Bruckner Expressway /I-278	107	5.6%	491,681	127,837,060	0.84	5.02	0.25	0.05	50	6
Prospect Expressway	73	3.8%	145,650	37,869,000	1.93	1.78	0.17	0.10	35	6
West Shore Expressway /R440	37	1.9%	696,279	181,032,540	0.20	9.33	0.09	0.01	50	9
Sheridan Expressway /I-895	2	0.1%	41,863	10,884,380	0.18	1.12	0.00	0.00	50	0
Total	**1907**	**100%**	**4,767,192**	**1,239,469,920**	**1.54**	**45.85**	**4.50**	**0.10**		

Based on the results shown in Table A-1 obtained using the data set, conclusive recommendations in terms of modeling incident frequencies cannot be made. Facilities with comparable features sometimes display inconsistent incident statistics. For example BQE and SIE have comparable daily VMT and length, with almost the same speed limit and same number of exits along the facility, however incident rates and incident per day statistics differ in magnitude. On the other hand, SIE and BE shows the same incident per day rates, but have completely different facility features. The difference between BQE and SIE can be explained due to the existence of shoulders. BQE does not have shoulders and have a very high rate of incidents compared with SIE, which has shoulder lanes. This reasoning is also valid between SIE and BE, since BE also does not have shoulders. Thus existence of shoulder can be stated as a major factor affecting the incident frequency where its affect on duration was also stated at duration analysis section. Effect of shoulder existence is further studied with examples from literature in “Analysis of Incident Frequencies According to Geometric Characteristics of Roadways” section in detail**.** Another important point is that no incident records exist for some roadway portions (see Figure

A-1). These portions correspond to 2 bridges in the study area; namely Verrazano and Triborough Bridges.

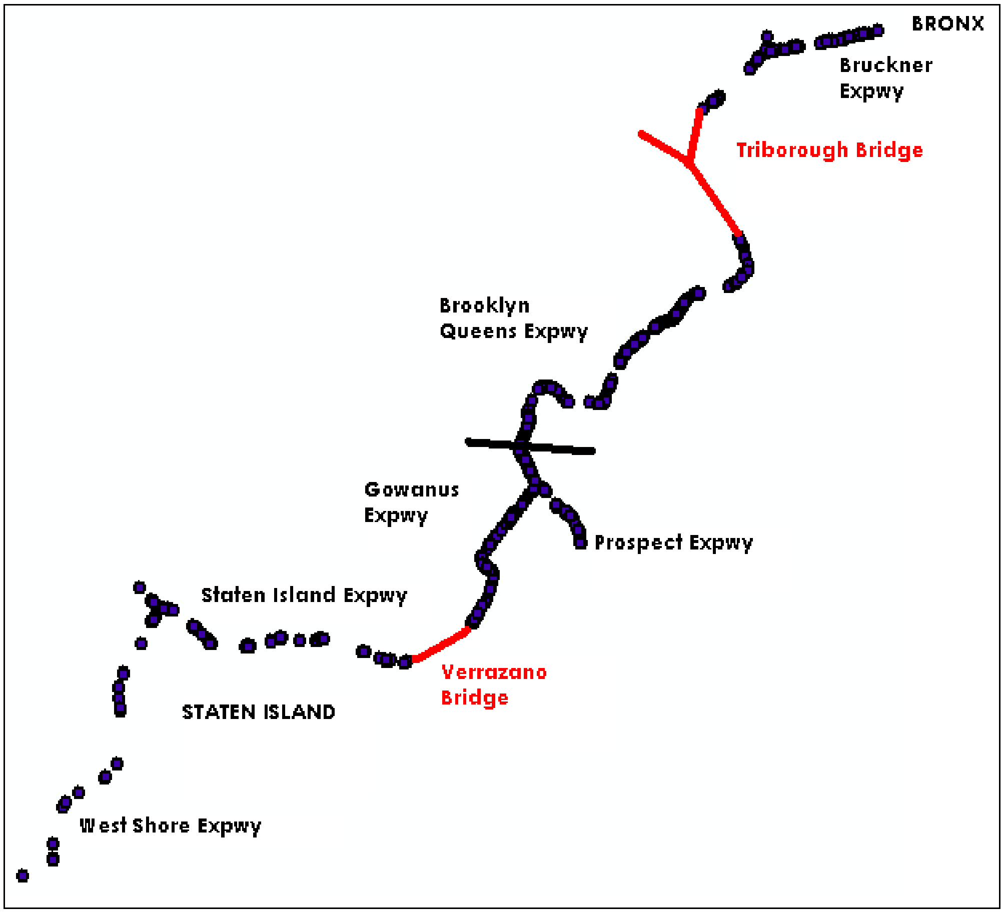


Figure A-1 Distribution of Incidents Along the Study Area

Table A-2 Incident Rates According to Incident Type and Facility

(Based on Table 6.3 of Task 3.1)

	Incident Types									
Facility Name	**Property Damage**	**Disabled Vehicle**	**Disabled Truck**	**Road HAZARD**	**Personal Injuries**	**HAZMAT**	**Vehicle Fire**	**Weather Related**	**Total**	**%**
BQE	398 (42%)	357 (38%)	87 (9%)	33 (4%)	34 (4%)	12 (1%)	14 (1%)	6 (1%)	941	49 %
GE	250 (39%)	264 (41%)	70 (11%)	23 (4%)	15 (2%)	9 (1%)	4 (1%)	2 (0%)	637	33 %
SIE	63 (57%)	16 (15%)	9 (8%)	13 (12%)	2 (2%)	3 (3%)	3 (3%)	1 (1%)	110	6 %
BE	64 (60%)	23 (21%)	6 (6%)	8 (7%)	5 (5%)	1 (1%)	0 (0%)	0 (0%)	107	6%
PE	26 (36%)	29 (40%)	5 (7%)	0 (0%)	10 (14%)	0 (0%)	3 (4%)	0 (0%)	73	4 %
WSE	17 (46%)	4 (11%)	2 (5%)	3 (8%)	1 (3%)	4 (11%)	5 (14%)	1 (3%)	37	2 %
SE	1 (50%)	1 (50%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2	0 %
Total	819	694	179	80	67	29	29	10	1907	100%
%	43 %	36 %	9 %	4 %	4 %	2 %	2 %	1 %	100%	

Table A-2, which is a slightly modified version of Table 6.3 of Task 3.1, shows the number of incidents at each facility according to the incident type. Please note that the percentages in paranthesis show the percentage of that specific incident type at that specific facility. As the total number of incidents decrease, the percentages get less significant, and even may become misleading since the percentages will change dramatically with small increases. However all percentages were included in this table to ensure consistency. Hence, the order of overall occurrence for incident types can be said to be valid for all individual facilities, e.g. Property damage, Disabled Vehicle, Disabled Truck, Road Hazard, Personal Injury and almost equally weighted HAZMAT, Vehicle Fire and Weather Related.

Analysis of Incident Frequencies According to Geometric Characteristics of Roadways

Besides the existing records, with the help of aerial photos, longitudinal (Curved/Straight) and lateral characteristics (Diverging/Merging/Weaving/Basic) were also extracted. As discussed before, this helps in addressing human perception errors (which should be not more than the human error in the original data records), Although these physical features gain even more importance for incident frequency analysis their effect on duration will also be briefly discussed.

Table A-3 shows the statistics for all facilities regarding the physical road attributes. Although some conclusions can be drawn from the table, these conclusions will be misleading since there is no reliable geometric information to normalize the number of any section over the whole roadway. For example, the numbers can be misleading unless the curved portion of the road segment is considered. Naturally, curved road sections constitute a lesser portion of the road compared to straight sections, thus the intuition that a curved section will be more dangerous is still valid, assuming all other factors are the same. Statistics are given in Table A-3 to illustrate the information that was extracted since the geometric characteristics are addressed as an important variable in the literature.

Shankar *et. al.* [14] study the effect of roadway geometrics and environmental factors by employing negative binomial regression. Geometric characteristics include number of horizontal curves, number of horizontal curves under designed (those curves with design speeds less than 112.6 kph, less than 96.5 kph, and less than 80.45 kph), maximum and minimum horizontal radii, number of vertical curves, and maximum and minimum grades. They find that in order to reduce accident likelihoods in areas that frequently experience adverse weather, the basis of establishing design criteria should be expanded beyond wet pavements and should be to avoid steep grades and horizontal curves with low design speeds in areas with adverse weather. They go further than finding this intuitive information and quantify the impacts such as reduction of monthly accident frequency by % 47.3 by eliminating all horizontal curves with a design speed less than 96.5 kph on a roadway section that experiences at least 5.1 cm of snowfall one or more days. In addition to modeling overall accident frequency on highway sections, separate regressions of specific accident types are performed. Unfortunately, the available dataset does not allow this kind on inferences since road geometrics are not available and weather data, as will be discussed separately, is not fully available.

Although there seems to be more accidents at diverging sections, a statistically valid conclusion cannot be drawn. Firstly, these cross-sectional features do not exist in the dataset and they are gathered from aerial photos with personal perception, so there might definitely be errors and biases. Secondly, there is no data about the overall roadway characteristics such as the percentage of segments for each kind of section (e.g. weaving, basic, diverging). That is the main reason for being able to normalize the number of incidents over all sections. This may generate misleading results. Overall, data reinforces some of intuitive assumptions, however lack of additional data makes it impossible to have a statistically reliable conclusion.

Table A-3 Road Section Specific Analysis of Incident Frequencies for the TRANSCOM data set

Factors		# of Incident Occurrences							
		BQE	**GE**	**SIE**	**BE**	**PE**	**WSE**	**SE**	**Total**
Longitudinal	**Curved Section**	354	205	57	26	9	16	1	668
	Straight Section	587	432	53	81	64	21	1	1239
Lateral*	**Merge**	57	18	18	3	6	2	0	104
	Diverge	503	375	31	54	38	26	1	1028
	Weaving	159	205	20	32	13	2	0	431
	Basic	192	42	31	45	16	6	1	333
Shoulder	**Yes**	0	0	102	29	29	31	2	193
	No	941	637	8	78	44	6	0	1714

*Data has some missing records thus numbers may not add up to number of total incidents

Having no shoulder is also a factor that is expected to increase incident rates and this effect is also verified by the data. Ogden[15] focuses on the effect of paved shoulders on accidents at rural highways in Victoria, Australia. Ogden finds that paved shoulders can decrease the accidents by 41%, which equivalent to 0.071 accidents per million vehicle kilometers. It is also found that accident types are also affected, being rear-end accidents having the largest variation followed by overtaking, out of control and off-carriages to both sides. One interesting information is given based on a survey of shoulder usage [16] that less than a quarter of vehicles that stopped on shoulders do this for emergency. Thus it is suggested that instead of full-width pavements, narrow shoulders can be build in a cost-efficient way, also decreasing the incident rates. This information is worth noting since the area of our NRD study, both accident rates (up to 8 times increase) and duration (59% increase) are affected by shoulder existence. This discussion can be extended further with the structural benefit issues of paved shoulders since the longest duration incidents are the road maintenance type of incidents for the study area.

Analysis of TRANSCOM Incident Frequencies According to Time of Day and Day of the week (Weekday/Weekend)

As mentioned in Task 3.1, there is no significant difference between week and weekend days. However, incident rates change for different times of the same day. An extensive representation of hourly incident rates can be found in Table 6.14 of Task 3.1. However, if the data is aggregated as morning-peak, evening-peak, off-peak, evening and night as shown in Table A-4, the incident rates become more comparable. Aggregation naturally makes the analysis less precise. Thus, a directional statistics was performed for different times of day to determine if there is any pattern on directionality, especially during peak time. Table A-4 shows the directional statistics.

Table A-4 Descriptive Statistics of Complete Incident Duration Data According to Time of Day and Facility with Direction Info (EB: East Bound, WB: West Bound)

		BQE		GE		SIE		BE		Total
		EB	WB	EB	WB	EB	WB	EB	WB	
Time of Day	Morning Peak (6 AM - 10 AM)	66	90	79	37	21	6	8	11	318
	Evening Peak (3PM - 7PM)	65	91	49	77	7	5	9	5	308
	Off-Peak (10AM - 3PM)	63	125	60	59	16	13	0	14	350
	Evening (7PM - 12AM)	51	81	36	57	5	3	7	6	246
	Night (12AM - 6AM)	50	38	24	18	4	8	4	8	154
	All Day	295	425	248	248	53	35	28	44	

Table A-4 summarizes the time of day variation of incident rates including the facility and direction information. Only 4 facilities with the highest number of incidents were analyzed. Note that some incidents like pothole repairs, road maintenance (as well as some accidents) are recorded as bi-directional and double counted for both directions. Also, the incident numbers are taken from weekdays since the effect of peak hours is for weekdays. Large number of incidents are expected for morning and evening peaks at opposite directions. This is observed for GE, however, BQE data which constitute the main body of the dataset do not meet this expectation. All these facilities lie on I-278, however there are obvious changes in directional patterns from facility to facility. From definitions shown in Table A-4 peak period lasts 4 hours while off-peak period is 5 hours. For all facilities, off-peak incident rates are almost the same as peak times. However, the evening and night rates are lower than peak/off-peak times, which is intuitive. Peak & off-peak times adds up to 13 hours and, evening & night constitutes almost the same amount of time; 11 hours. On the other hand, peak/off-peak have 2.5 times more incidents. This shows that time of day does not have an affect between peak and off-peak hours but makes a difference for evening and night. Nevertheless it should affect the delay at all times, since flow patterns are closely related to TOD. Overall, TOD can be said to have an effect on incident rate. It can be said that direction also is not a significant contributor to frequency within the whole dataset since there is no general pattern, but can affect the incident rate from a spatially detailed perspective.

Analysis of Incident Frequencies According to AADT

Dataset include AADT values for each incident. Figure A-2 shows the distribution of the number of incident on aggregate AADT values.

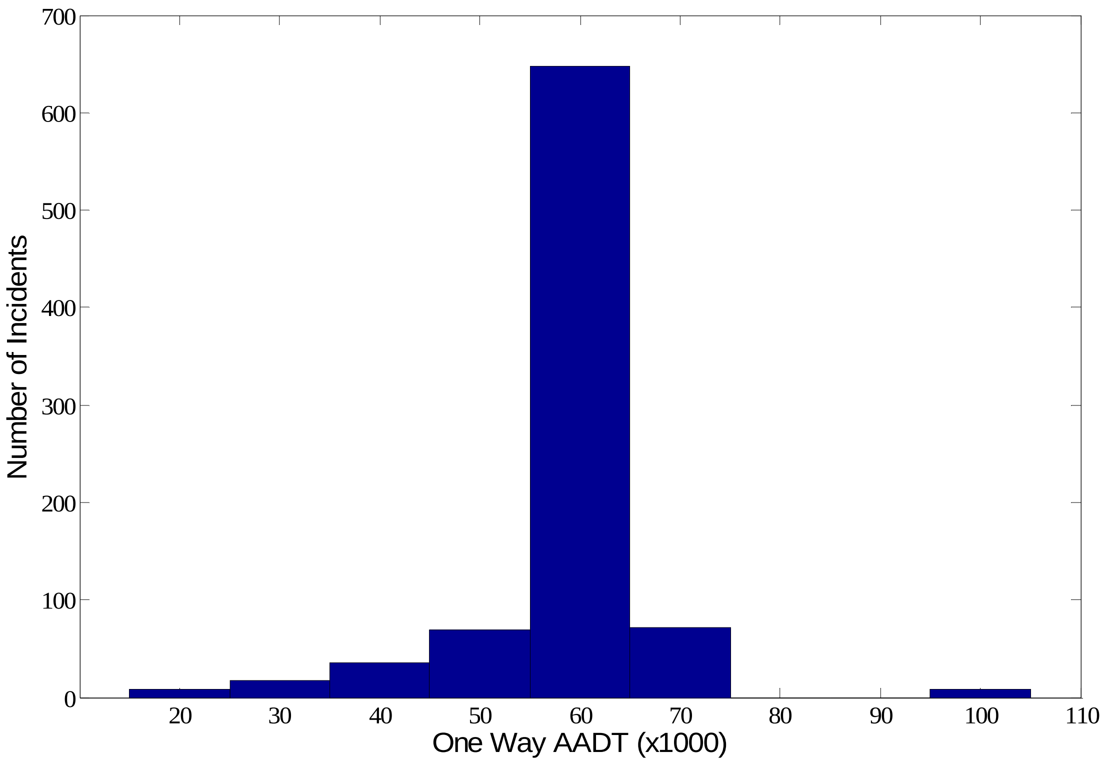


Figure A-2 Distribution of Number of Incidents with respect to One Way AADT

As shown in Figure A-2, the number of incidents are related to AADT. There are significantly more incidents for AADT's around 60000 vehicles per day. This is most probably due to the study area charactertics since the facilities used in this study have close AADT values, thus the overall distribution is narrow. However the change in AADT from 50000 to 60000 vehicles per day is considerable. A causal relationship between AADT and incident rates can be estbalished and thus it can be said that the incident rate is affected by the AADT. This is a reasonable result and agrees with the previous findings in the literature literature(7,11,12).

Analysis of Incident Frequencies According to Pavement and Weather Conditions

Pavement and weather conditions are also inter-related. As mentioned in Task 3.1 there are 1251 records out of 1907 including the necessary weather information. 85.1% of the incidents occurred during clear and cloudy weather conditions, and 14.9% occurred during adverse weather conditions.

Incidents during adverse weather conditions were also investigated as a separate category. For this analysis all the weather conditions except the clear weather, including blizzard, cloudy, flurries, foggy, heavy rain, high wind, light wind, light rain, overcast, rain, sleet and snow, were considered. The distributions of incident types on those days are given in Table A-5 with the values in parenthesis being the overall percentages. The order of the incident types were taken same as the system-wide occurrence order, and it can be seen that only HAZMAT incidents do not follow this order. Since the numbers become very small for this type, HAZMAT accidents can be neglected. Looking at the overall and adverse weather percentages of each incident type in Table A-5, it can be said that adverse weather conditions do not affect the type of the incident to occur.

Table A-5 Number of Incidents According to Incident Type During Adverse Weather Conditions

	Incident Types							
	Property Damage	**Disabled Vehicle**	**Disabled Truck**	**Road HAZARD**	**Personal Injuries**	**HAZMAT**	**Vehicle Fire**	**Weather Related**
Number of Incidents	132	116	37	11	11	3	8	6
Percentage	41% (43%)	36% (36%)	5.5% (9%)	3% (4%)	3% (4%)	1% (2%)	2% (2%)	2% (1%)

Likewise, the incident statistics for each facility during adverse weather conditions are also investigated to see if any facility is more vulnerable to weather conditions. The results are given in Table A-6. Based on the overall percentages shown in parenthesis in Table A-5 and the percentages under adverse weather conditions, it can be concluded that there is no significant pattern change in terms of incident rate for facilities.

Table A-6 Number of Incidents According to Facility Type During Adverse Weather Conditions

	Facility Types							
	BQE	**GE**	**SIE**	**BE**	**PE**	**WSE**	**SE**	**Total**
Number of Incidents	157	117	13	16	15	7	0	325
Percentage	48%	36%	4%	5%	5%	2%	0%	100%

Weather conditions are also mentioned to affect the incident frequency, especially in the presence of steep grades and sharp curves. As mentioned in the analysis of road geometry section, Shankar *et.al.*[14] find that in order to reduce accident likelihoods in areas that frequently experience adverse weather, the basis of establishing design criteria should be expanded beyond wet pavements and should be to avoid steep grades and horizontal curves with low design speeds in areas with adverse weather. However, using the existing data, no such inference can be drawn in this study.

Discussion on Descriptive Incident Frequency Analysis of the TRANSCOM data

Among the analyzed factors that are available in the dataset, the strongest factor affecting the incident frequency is determined to be the existence of the shoulder. Between two very similar facilities, BQE and SIE, the major difference that can be extracted from the dataset is the existence of the shoulder and incident rates. BQE which has no shoulders, has almost 6 times the incident rate (Table A-1) of SIE, which has shoulders. Physical characteristics of the road section can be mentioned to affect the frequency. Briefly, curved and merging sections are found to have higher incident rates in literature. This was also shown in the relevant section of descriptive analysis (Table A-3) although the results are not very reliable since there is no data to normalize the weight of curved or merging sections for the whole study area. In terms of the time of day, the data do not exhibit a significant variance between peak & off-peak, but for evening and night periods, incident rates decrease considerably. There are some subtle points about the impact of direction. BQE has more incidents on westbound regardless of the time of day. Although both facilities are on the same highway (I-278), GE shows the opposite statistics compared to BQE and has more incidents in the eastbound direction during morning peak and almost the same number incidents during evening peak (Table A-4). The other two facilities (SIE and BE) have more incidents in eastbound direction in total (like GE and opposite to BQE), however since the number of records is not very high, it is hard to reach a statistically significant conclusion.

OVERALL CONCLUSIONS AND DISCUSSION FOR FREQUENCY ANALYSIS OF THE TRANSCOM DATA

The major results of duration analyses performed in Task 3.2 can be summarized as follows:

- TRANSCOM dataset contains only 8% of all the accidents in the region which makes it a clearly incomplete and inadequate dataset for frequency analysis and duration analysis to a lesser degree. Even for duration analysis, it is important to understand that the sample provided by TRANSCOM can be highly biased in terms of certain accident types since the research team did not have any input in creating this sample. For example, TRANSCOM data might contain certain types of accidents that are not relevant to their operations and this might bias the results of the data analysis. Thus, all the conclusions listed below have to be interpreted in the light of this major shortcoming related to the quality of the accident data used in this study.
- Some factors affecting incident frequency are found by using descriptive analysis
    - AADT: Incident frequency is higher around AADT~50-60000 vehicles, and less elsewhere. Thus, it can be said that the relationship between AADT and incident frequency is not always linear.
    - Shoulder Existence: No shoulder regions have higher incident rates compared to sections with shoulder.
    - Direction and Time-of-Day were found to affect incident rates for some specific regions and some time periods.
- Dataset lacks very important data categories needed to model incident durations and frequency accurately.
    - Dataset includes no information about emergency response timings
    - Information about the severity of the accidents is not available.
    - Incident locations are not exact.
    - Physical characteristics (i.e. grade, curvature) do not exist
    - Some unreasonably high and very low durations exist in the data, producing poor statistics for analysis.

Task 3.3: MODEL DEVELOPMENT

Prepared by:

Kaan Ozbay, Ph.D.
Associate Professor
Rutgers University, Civil and Environmental Engineering
623 Bowser Road, Piscataway, NJ 08854-8014
Tel: 732.445.2792
Fax: 732.445.0577
Email: kaan@rci.rutgers.edu

M. Anil Yazici
Graduate Assistant
Rutgers University, Civil and Environmental Engineering
623 Bowser Road, Piscataway, NJ 08854-8014
Tel: 732.445.0579 (x119)
Fax: 732.445.0577
Email: yazici@eden.rutgers.edu

June 20, 2007

1.0 Introduction

Task 3.2 presented a detailed descriptive analysis based on the available dataset. In Task 3.3, information obtained in Task 3.2 is used for modeling the incident durations and frequencies. It is however important to re-emphasize the problem with frequency dataset. Recently, research team discovered that the dataset that has been provided to the team did not contain all the incidents in the study area. Detailed explanation of this problem with the incident data set is given in Task 3.2. Hence, only the existing duration models in the literature are investigated then the most promising models are estimated using the data. Due to the inadequate number of records in TRANSCOM dataset, the incident frequency analysis is moved to APPENDIX for modeling guidance purposes. As mentioned in the revised statement of work, the aims of ask 3.3 are:

- Listing the sources for the models considered,
- Describing the model structures selected as well as those considered,
- Performing an analysis of the factors that influence incident delay,
- Proposing a set of equations for estimating incident frequency, duration

Overall, Task 3.3 forms the basis for Task 4.1, where the chosen models will be used for duration estimations to be used in the lookup tables.

However, before studying these models in depth, a sectional analysis is performed to locate each incident along the study section. Only location information that is available the facility name. More specific geo-location information is not explicitly given in the dataset. However to see the distributional patterns of duration and incident number within the facilities along the study section and to be able to provide spatial information needed by the models, the “vicinity” records that are available in the dataset were used to approximately geo-code incident locations. Then sectional characteristics of individual sections were used to perform a clustering analysis where the study area was divided into High-Medium-Low incident frequency and duration sections. The estimation of selected models also benefited from this clustering analysis as shown in the section of this report where we discuss estimation results.

2.0 Sectional Analysis

The dataset provided for this study contains incident location records with street or intersection names at each facility. No detailed geographic location information, such as milepost, is available in the dataset. It was also shown in Task 3.2 that facility information does not provide insights to incident dynamics. To gather more specific information, the “verbal” vicinity information provided in the dataset is used. However, the available vicinity information in the data set include some comments that are not always very clear such as “near”, “east of” etc. Based on a subjective interpretation of these records, incidents were placed along the individual sections of facilities in the study section. These individual sections were categorized as “at” or “between” the exits along the roadway. It should be noted that these sections are not homogenous in terms of physical attributes and not evenly placed, however this approach still provides valuable input, especially for the existing dataset, which is incomplete in terms of detailed location information. The error introduced is also bounded since an incident can be assigned to 3 consecutive sections (before, at of after an exit). Thus, an erroneous assignment will still give a rough idea about the spatial distribution of the incidents and their durations.

Below, the figures show the duration and incident frequencies for each facility. Sheridan Expressway was not included since there are only 2 records for that facility. At the end overall study area is represented in the same way. The exit numbers shown in facility specific charts (Figure 3 - Figure 14) are the real exit numbers on that facility. For the overall duration charts (Figure 15 and Figure 16), the exit numbers were arranged as shown below to be able to represent all durations in a single chart. The study area can be represented as a single line since all the facilities form an almost continuous corridor.

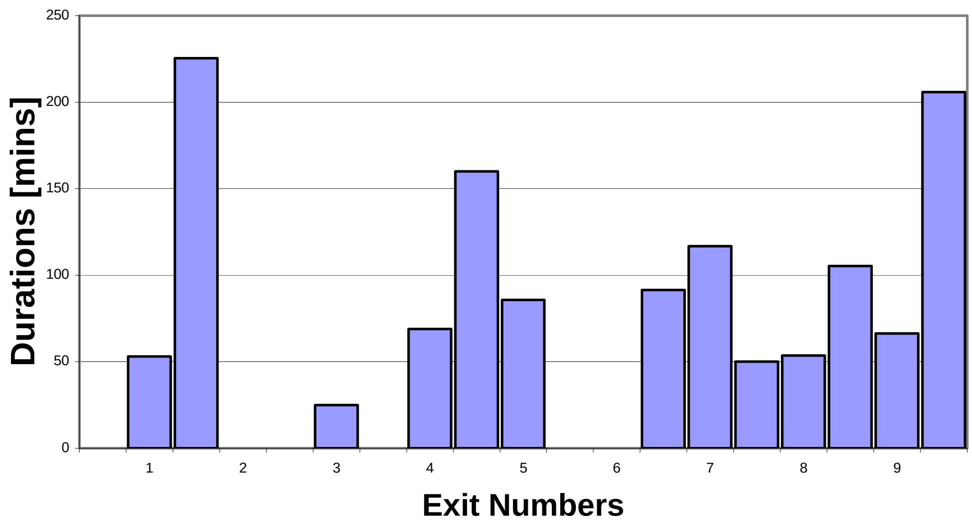


Figure 3- Incident Durations West Shore Expressway (WSE)

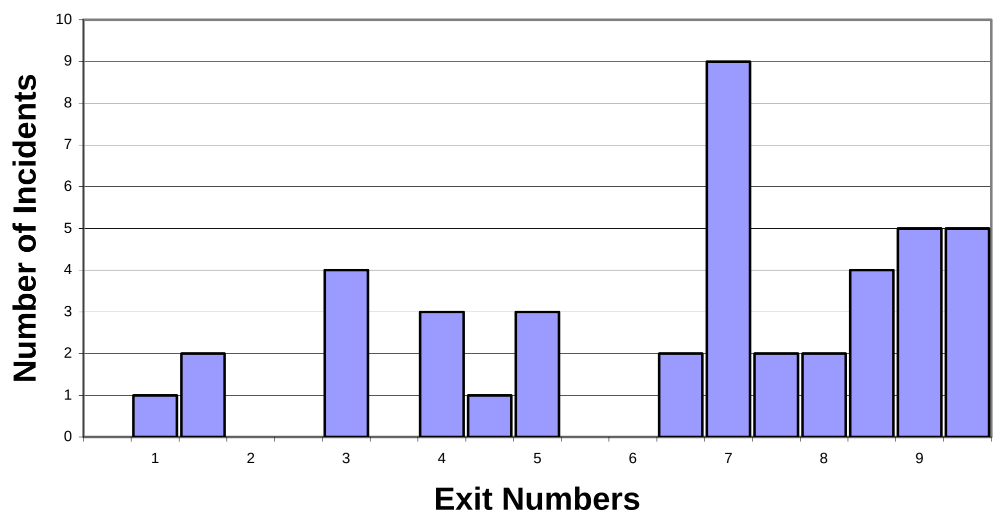


Figure 4- Number of Incidents Along West Shore Expressway (WSE)

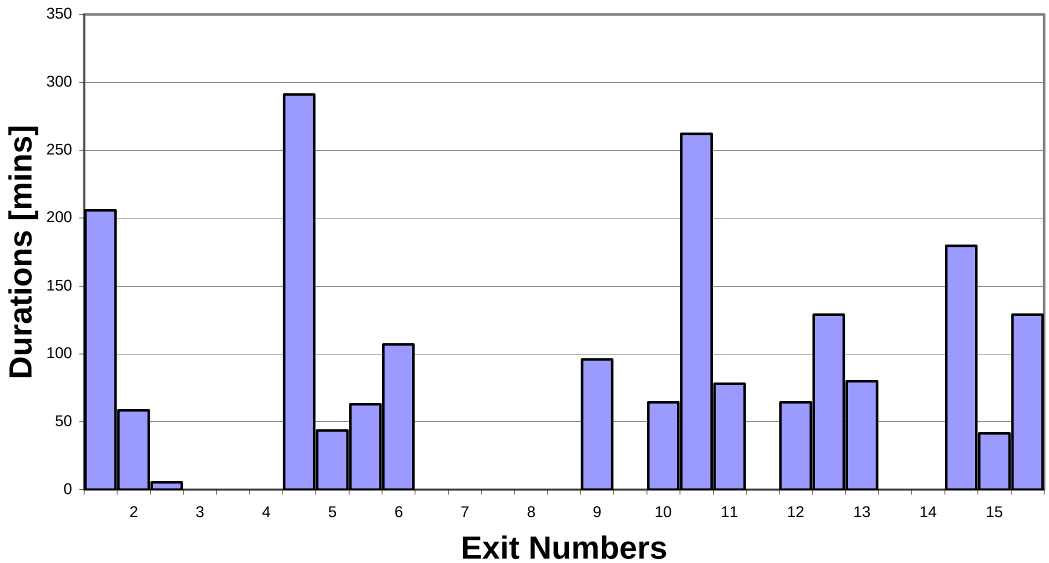


Figure 5- Incident Durations Along Staten Island Expressway (SIE)

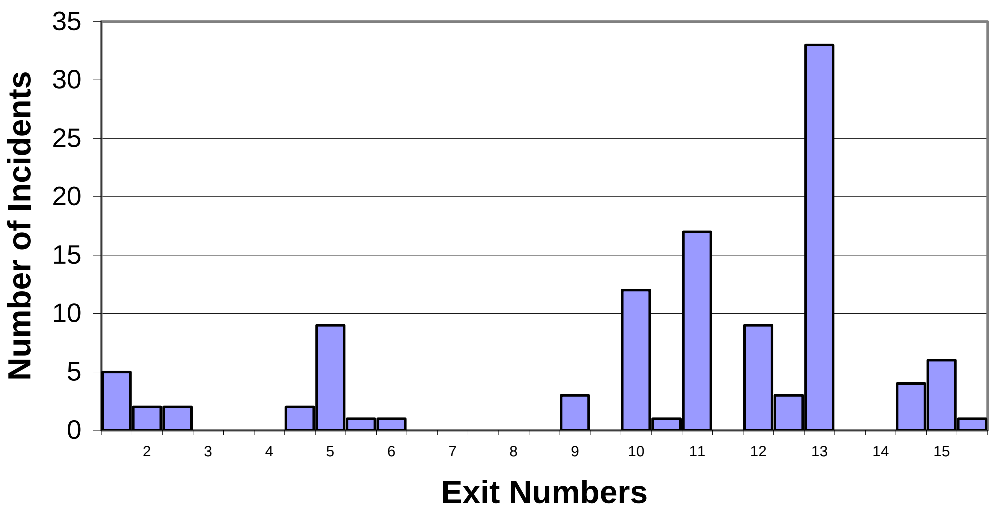


Figure 6- Number of Incidents Along Staten Island Expressway (SIE)

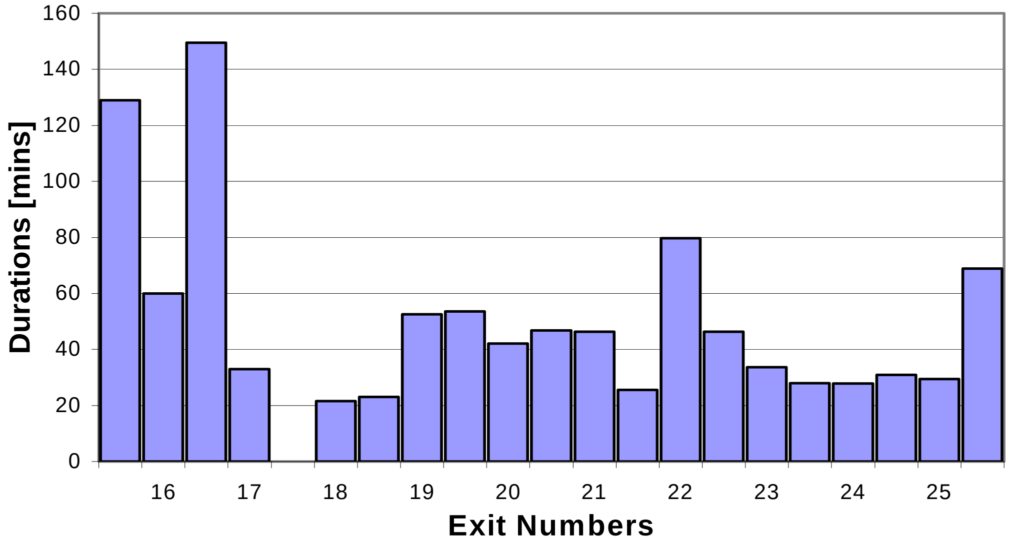


Figure 7- Incident Durations Along Gowanus Expressway (GE)

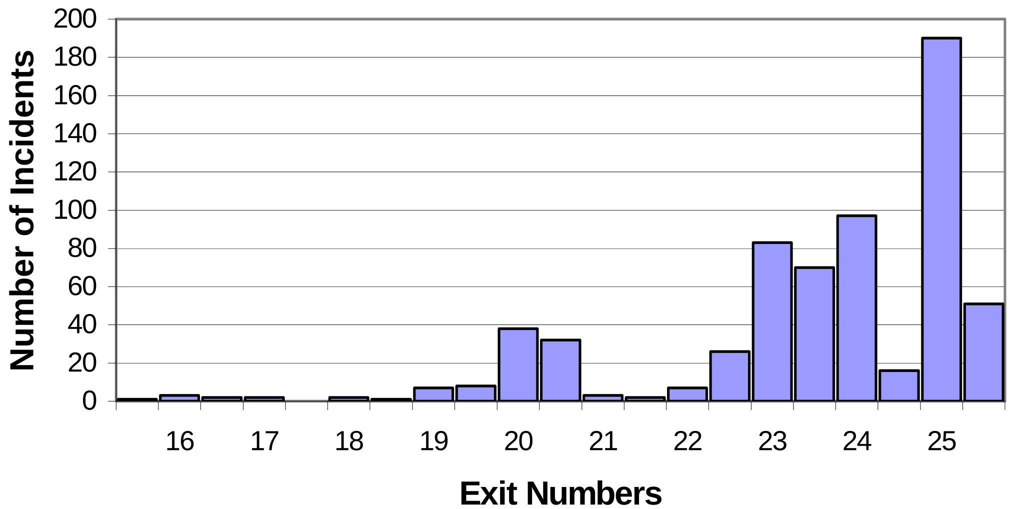


Figure 8- Number of Incidents Along the Gowanus Expressway (GE)

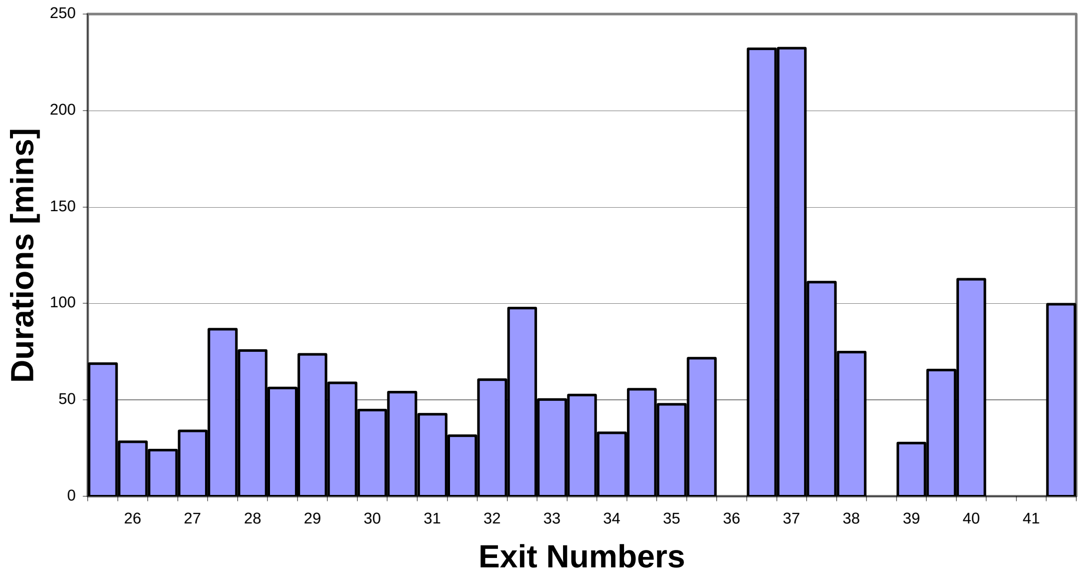


Figure 9- Incidents Durations Along Brooklyn Queens Expressway (BQE)

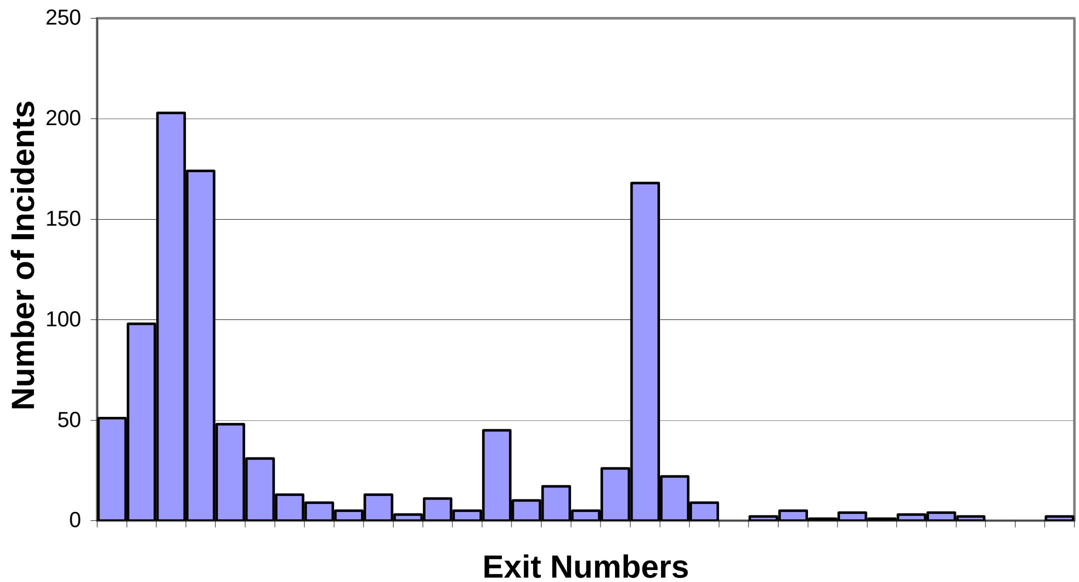


Figure 10- Number of Incidents Along Brooklyn Queens Expressway (BQE)

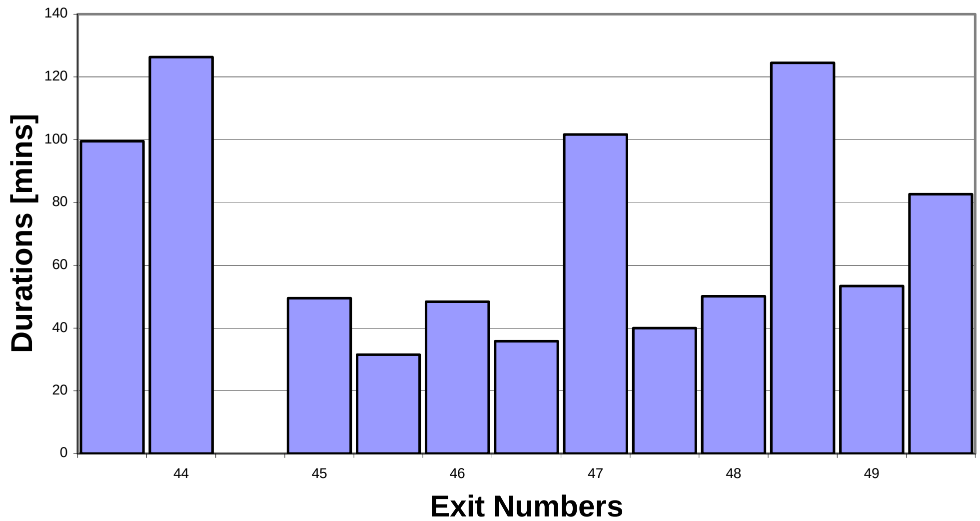


Figure 11- Incidents Durations Along Bruckner Expressway (BE)

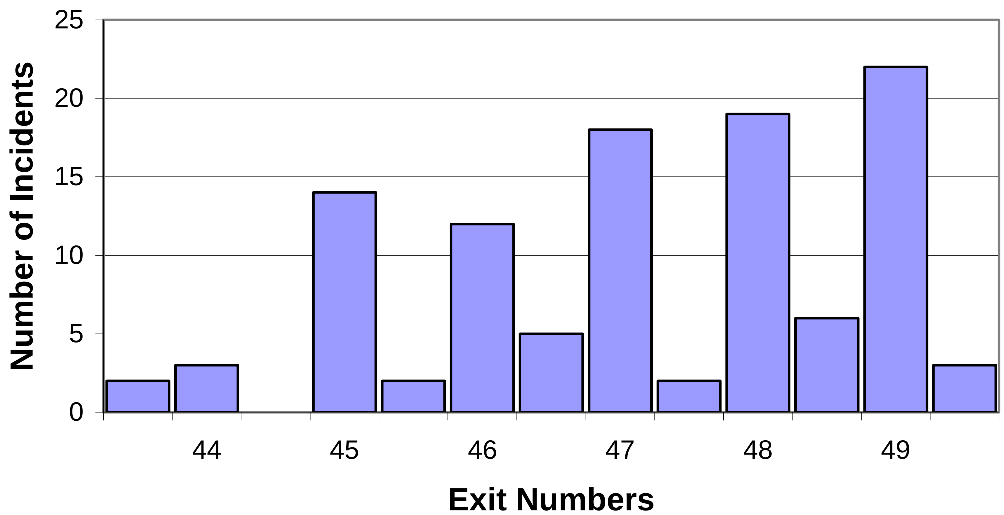


Figure 12- Number of Incidents Along Bruckner Expressway (BE)

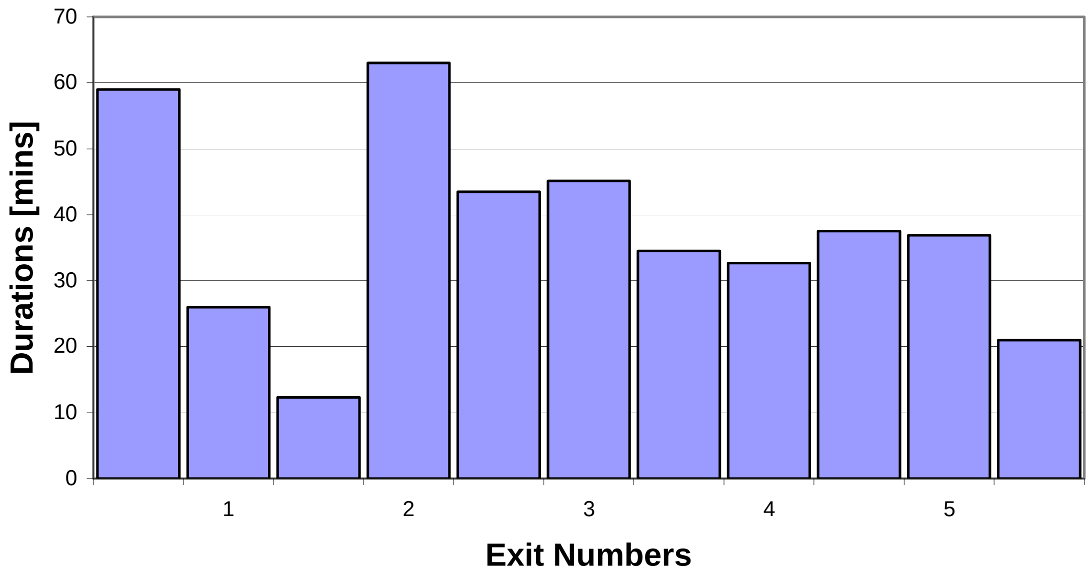


Figure 13- Incident Durations Along Prospect Expressway (PE)

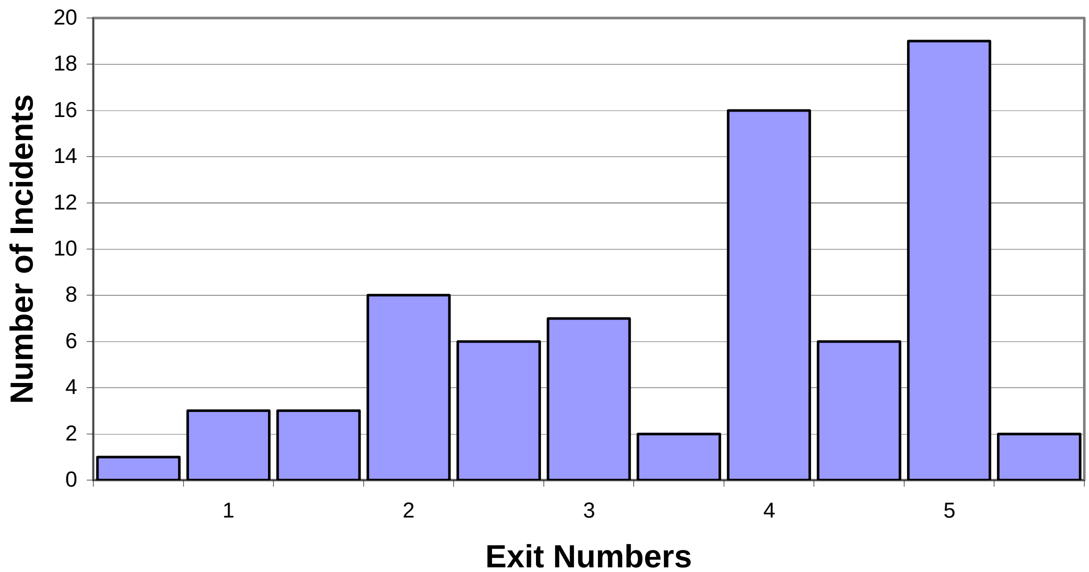


Figure 14- Number of Incidents Along Prospect Expressway (PE)

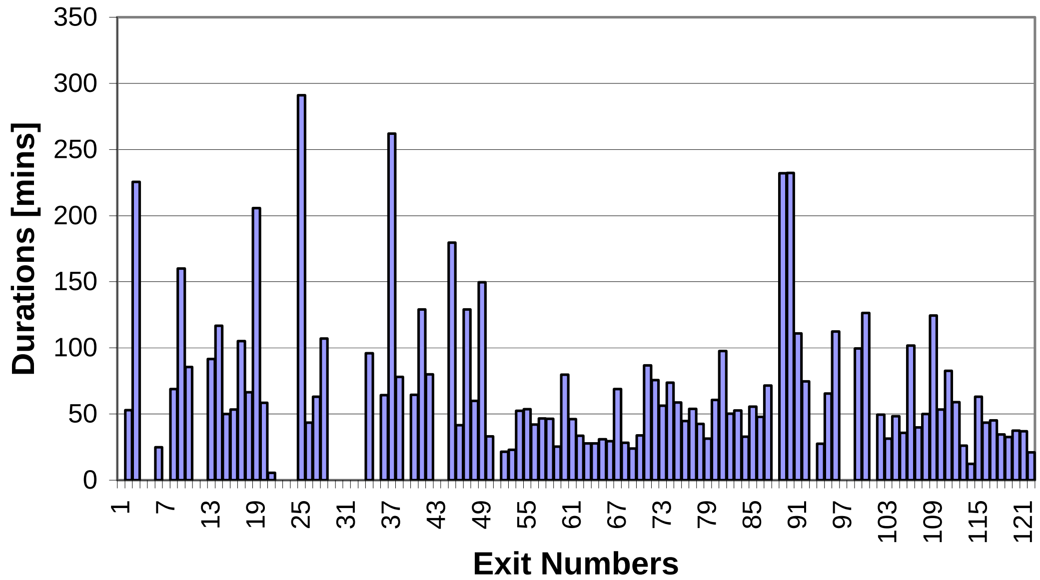


Figure 15- Incident Durations for the Whole Study Area

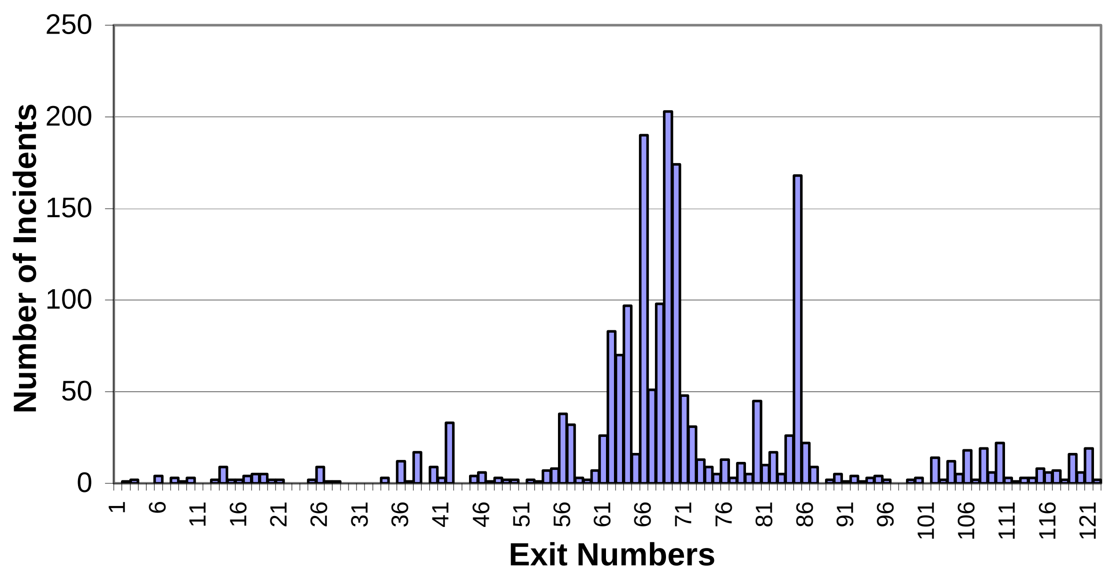


Figure 16- Number of Incidents for the Whole Study Area

Table 7- The Facilities and Corresponding Section Numbers

Facility	Exit Numbers	Section Numbers
WSE	Exit 1 – Exit 9	1 – 19
SIE	Exit 2 – Exit 15	19 – 47
GE	Exit 16 – Exit 25	47 – 67
BQE	Exit 25 – Exit 41	67 – 99
BE	Exit 41 – Exit 49	99 – 111
PE	Exit 1 – Exit 5	112-122

Based on the facility specific incident frequency figures that are obtained from the original dataset which is a sub-set of all the incidents on these sections, it can be said that the incident rate distribution is not uniform between and within the facilities[1]. If the general picture of the study area is remembered where each facility is a continuation of the previous one, it can be said that there is a concentration of incidents between Exits 23 and 27, being partly on Gowanus and Brooklyn Queens Expressways. The section between Exits 34 and 35 in BQE also shows high frequency of incidents (155 incidents) compared to the neighboring exits. This section corresponds to Kosciuszko Bridge and the narrowness of the bridge can be the reason of the relatively higher incident rates. Similar higher rates are also observed also along WSE (Exit 7 with 9 incidents) and SIE (Exit 13 with 33 incidents), however their significance is relatively less since those facilities do not exhibit incident rates as high as GE and BQE.

The inhomogeneous distribution of durations and incident numbers lead the research team to define new spatial parameters instead of using just the facility name. Then, a clustering analysis was performed to divide the study area into relatively more homogenous sections for which more accurate models can be estimated.

Clustering Analysis for Incident Rates and Durations

To define a new spatial variable that can identify "similar" sections over the study area, spatial the K-means clustering algorithm is used. The algorithm determines consecutive clusters, which show the minimum variation of duration or incident numbers within each cluster, and show distinct values compared to neighboring clusters. This algorithm basically assigns temporary borders to each cluster and adjusts those borders so that the variation among each cluster is minimized. For the sake of simplicity, 3 clusters (High-Medium-Low) that are extracted using this algorithm are shown in Table 8 and
Table 9 with the corresponding cluster boundaries, for duration and frequency values respectively. The clusters are ordered according to South-to- North (or West-to-East) direction.

[1] The discussions of the discrepancy between TRANSCOM and NYSDOT datasets given in Task 3.2 show that TRANSCOM dataset does not have the same percentages for types of accidents as in NYSDOT dataset. Hence, it can be argued that there is a bias in TRANSCOM dataset regarding the accident types that were recorded, which consequently introduces bias on the accident durations. There is no information about disablements in NYSDOT dataset and no comparison can be made with TRANSCOM, however it is possible that same kind of bias is also valid for disablement durations.

Table 8- Incident Duration Clusters for the Whole Study Area

Duration Cluster	Cluster Boundaries
High	WSE Exit 1 – SIE Exit 11
Low	SIE Exit 11 – BQE Exit 36
Medium	BQE Exit 36 – BE Exit 49 + PE Exits1-5

Table 9- Incident Rate Clusters for the Whole Study Area

Frequency Clusters	Cluster Borders
Low Incident Rate	WSE Exit 1 – GE Exit 22
High Incident Rate	GE Exit 22 – BQE Exit 28
Medium Incident Rate	BQE Exit 28 –BE Exit 49 + PE Exits1-5

The duration and frequency pattern changes shown in Figure 15 and Figure 16 are also captured by the clustering approach. It is important to note that the clusters shown in
Table 9 have different boundaries compared to the ones shown in Table 8. This is an expected result since duration and frequency are two distinct variables and clustering performed using either one of them is expected to be distinct. High incident frequency area mostly falls into low incident duration clusters, which can be interpreted as the better level of incident response at those areas. It can be because of the preparedness and awareness for high rate of incidents at that area, or other concerns, such as close proximity of response team etc. Another possible reason can be the possible incident record bias in the data as well. This is a subject of further investigation, but for time being this is skipped due to lack of information.

INCIDENT DURATION ESTIMATION FOR THE TRANSCOM DATASET

In the literature, several methods are employed for the incident duration prediction. These methods can be summarized as (7):

- Probability Distributions
- Linear Regression
- Weibull Regression
- Conditional Probabilities
- Time Sequential Models
- Decision Trees

The probability distributions for various assumptions were presented previously in Task 3.2. The probability of an incident lasting less (or more) than a specific time can be extracted from these estimated distributions. In the literature lognormal (8,9,10) and Weibull (11) distributions are generally employed for incident durations. Our dataset show that both Weibull and Gamma distributions fit the duration data for different cases/facilities. Lognormal distribution is also found to represent the data at an acceptable level of significance. However, the fitted distributions exhibit variations for different incident types and facilities. In this current task, data was filtered and distributions were re-fitted to the data to see if there is any improvement due to filtering. Besides it is worthwhile to investigate other methods listed above. Among these, alternative methods conditional probability is used for finding the accident characteristics that have an impact on the incident duration instead of direct forecasting of durations (7). Time sequential model is proposed in the literature but has not been applied on real life cases to show its effectiveness for duration prediction (7). Thus, in this current study linear regression, weibull regression, and decision trees will be further investigated for the task of duration forecasting.

Data Filtering

In Task 3.2, all analyses were performed with the raw data since the objective of the task was to analyze the raw data made available to the research team. It was shown using both histograms and distribution fits that duration data have outliers that are far larger than the overall duration values. These outliers sometimes caused "poor" distribution fits. Considering that almost 80% of the incidents are below 1 hour, it is appropriate to filter the long durations from the data for better estimations. Thus, the durations were truncated using an upper bound for all types independently. The upper bound is set to be two standard deviations away from the mean value. For the overall data set, the number of filtered points is 74, which is not a large number given that there are 1907 in total. 37 of the filtered incidents are found to be "Road Hazard" incidents. This corresponds to 50% of all the filtered accidents whereas the overall share of "Road Hazard" is only 4%. Second biggest share, 26%, among the filtered records is "Property Damage" incidents, which constitute 43% of all the accidents in the dataset. Moreover, if the incident description field is investigated for "Road Hazard" incidents, it can be seen that most of them are road maintenance, pothole repair, etc., which are scheduled events rather than probabilistic events. Those facts reinforce both the need to have a separate analysis for each incident type as well as the need for data filtering for more homogeneous samples. In other words, deleting those records do not decrease the quality of the data. On the contrary, it eliminates the inclusion of non-relevant records in the analysis. Overall reduction in the number of data points is negligible if the improvement on the prediction power is considered.

Probability Distributions

After the filtering scheme is employed, the distribution fits show significant improvements in terms of test statistics and goodness of fit results. Table 10, Table 11 and Table 12 show the comparison between the fits for overall data from Task 3.2 and new fits after filtering.

Table 10- GOF Measures Weibull Regression Before and After Data Filtering

Incident Type	Kolmogorov-Smirnov Test						Anderson-Darling Test				
	Old Test Statistics	Old Critical Values	New Test Statistics	New Critical Values	Good Fit? Old	Good Fit? New	Old Test Statistics	New Test Statistics	Critical Values	Good Fit? Old	Good Fit? New
Property Damage	0.043305	0.048194	0.032133	0.048784	Yes	Yes	2.5412	0.80669	0.757	No	**No**
Disabled Vehicle	0.039405	0.052251	0.040854	0.052844	Yes	Yes	1.6239	1.4091	0.757	No	No
Disabled Truck	0.097258	0.10224	0.050489	0.10313	Yes	Yes	3.3839	0.24176	0.757	**No**	**Yes**
All Incidents	0.059606	0.031707	0.025365	0.031084	**No**	**Yes**	17.431	1.0078	0.757	No	**No**

Table 11- GOF Measures Gamma Regression Before and After Data Filtering

Incident Type	Kolmogorov-Smirnov Test						Anderson-Darling Test				
	Old Test Statistics	Old Critical Values	New Test Statistics	New Critical Values	Good Fit?		Old Test Statistics	New Test Statistics	Critical Values	Good Fit?	
					Old	New				Old	New
Property Damage	0.037555	0.048794	0.046764	0.048194	Yes	Yes	2.5412	1.609	0.844	No	No
Disabled Vehicle	0.037785	0.052844	0.035425	0.052251	Yes	Yes	1.6239	1.1153	0.844	No	No
Disabled Truck	0.096949	0.10313	0.041744	0.10224	Yes	Yes	3.3839	0.28958	0.844	**No**	**Yes**
All Incidents	0.64329	0.031084	0.024077	0.031707	**No**	**Yes**	17.431	1.0057	0.844	No	**No**

Table 12- GOF Measures Lognormal Regression Before and After Data Filtering

Incident Type	Kolmogorov-Smirnov Test						Anderson-Darling Test				
	Old Test Statistics	Old Critical Values	New Test Statistics	New Critical Values	Good Fit?		Old Test Statistics	New Test Statistics	Critical Values	Good Fit?	
					Old	New				Old	New
Property Damage	0.07266	0.048784	0.085388	0.048194	No	No	2.5412	11.794	0.735	No	No
Disabled Vehicle	0.058991	0.052844	0.06456	0.052251	No	No	1.6239	4.636	0.735	No	No
Disabled Truck	0.086395	0.10313	0.093433	0.10224	Yes	Yes	3.3839	2.2567	0.735	No	No
All Incidents	0.048313	0.031084	0.063991	0.031707	No	No	17.431	16.205	0.735	No	No

As seen in Table 10, Table 11 and Table 12, the test statistics improve for weibull and gamma distributions. Fit statistics for the Lognormal distribution get worse with filtering nevertheless the performance in terms of passing the tests remains the same. On the other hand, for gamma and weibull distributions, some incident types that previously failed to pass tests gives better test statistics and pass the tests. For the first three major incident types and overall data, distribution fits pass Kolmogorov-Smirnov test. Anderson-Darling test, as mentioned in Task 3.2, is harder to pass since it gives weight to tails and our data shows uneven portioning for the tails. However, improvements for the Anderson-Darling test is observed obtained after performing the filter. These changes are marked in bold. Some improvements cannot be seen by only looking at pass/fail results of the tests. For instance, Anderson-Darling test statistic for property damage at 95% significance level do not drop under the critical value in the case of weibull distribution, however it gets very close to critical value which means that it will be accepted for a less strict confidence interval that is less than 95%. For ensuring reliability of distribution fits, confidence level was always kept at 95% but this kind of improvements were also marked in bold. Overall, almost all the test statistics, except lognormal, exhibited some level of improvement after data filtering.

Linear Regression

For the linear regression, powerful statistics software "R" [35] was used. The independent variable set was selected to be as large as possible, while taking into account the results of the previous descriptive and statistical analysis. Linear regression analysis was performed for each incident type separately. First, system wide regression was done using all the parameters that were found to affect the duration in Task 3.2. These are:

- Number of Lanes
- Number of Closed Lanes
- Shoulder existence (No shoulder=0, Shoulder exists=1)
- Time of day (Off-peak=0, Peak=1)
- Weekday/weekend (Weekend=0, Weekday=1)

Naturally, incident duration was chosen to be the independent variable. Both raw and log of durations were used in the regression. Using log of the duration implicitly assumes that the duration is distributed according to a lognormal probability density function. This assumption was employed in (9) and found to have a high predictive performance. Unfortunately, very poor R-square values were obtained in that system-wide regression. Lastly, the cluster information was used instead of facility information. Nonetheless, the regression still produced poor results. Our analysis also shows that using the log of the duration data does not make much difference both in terms of statistically significant variables and R-squared values. Using the clustered sections that were found in sectional analysis caused a small improvement in estimation compared to using facility type alone. Overall, the linear regression yields very poor regression results consistent with the findings of Ozbay *et. al.*[12] and Smith *et. al.*[7]. Hence, no further analysis was performed using linear regression.

Weibull Regression

The Weibull distribution is widely used to study lifetime data in reliability engineering. It can also exhibit the characteristics of other types of distributions because of its parametric flexibility. Weibull distribution can also be parameterized to represent the survival distribution. Survival analysis aims to find explanations to the issues like the fraction of a population which will survive past a certain time. The probability of an incident to be cleared increases as time passes, just like the increasing probability of a mechanical failure or failure of a device when time between the last and expected failure increases. In that sense, the analogy between incident duration and a life time of an event can explain incident duration being Weibull distributed.

The formula for the probability density function of the general Weibull distribution is:

$$f(x) = \frac{\gamma}{\alpha}\left(\frac{x-\mu}{\alpha}\right)^{(\gamma-1)} \exp\left(-\left(\frac{x-\mu}{\alpha}\right)^{\gamma}\right) \qquad x \geq \mu; \gamma, \alpha > 0$$

where γ is the shape parameter, μ is the location parameter and α is the scale parameter. The case where μ = 0 and α = 1 is called the standard Weibull distribution. The case where μ = 0 is called the 2-parameter Weibull distribution.

To be able to address the poor data quality, Bayesian approach is also considered for Weibull regression besides classical Weibull regression. Bayesian analysis briefly handles the problem using distributional properties of the data rather than individual data points, and makes re-sampling for the predictions. This introduces flexibility in the case of poor data quality as well as being able to deal with missing data entries. To test the usability of Bayesian approach, weibull regression is carried out with both classical and Bayesian approaches.

For the classical weibull regression built in function of statistical software "R" (35) is used. For the Bayesian regression, WINBUGS software (36) is employed. For both regressions, the chosen set of parameters is set to be the independent variable, where duration is chosen as the dependent variable. Some of these independent variables are already found to affect the duration as a result of the descriptive analysis. For example shoulder existence is found to have a major impact and TOD having a minor effect on the durations. The others are chosen among the variables that have enough diversity in the data and can be considered as factors that are expected to have an impact on duration. The lack of important fields, such as emergency response information, and lack of diversity in the data, such as weather information, limit the analysis to a relatively small number of variables. The variables are exploited as much as possible during regression analysis are:

- Number of Lanes
- Number of Closed Lanes
- Shoulder existence (No shoulder=0, Shoulder exists=1)
- Time of day (Off-peak=0, Peak=1)
- Weekday/weekend (Weekend=0, Weekday=1)

Classical Weibull Regression

The weibull regression is performed with both filtered and unfiltered data, using "Zelig" package developed for the R software. Just like the distribution fits, the filtered data gave better regression statistics. Below, the regression output of the "R" software for the filtered "property damage" incidents is given in Table 13.

Table 13- R Output for Property Damage Weibull Regression

Variables	**Value Std.**	**Error**	**z**	**p**
(Intercept)	3.76098	0.2758	13.638	2.37e-42
TOD	0.03120	0.0545	0.573	5.67e-01
# of Closed lanes	-0.00232	0.0553	-0.042	9.67e-01
Weekday/Weekend	0.00695	0.0622	0.112	9.11e-01
# of lanes	-0.04421	0.0905	-0.488	6.25e-01
Shoulder Existence	0.28827	0.0828	3.482	4.97e-04
Log(scale)	-0.28885	0.0286	-10.088	6.22e-24
Scale= 0.749 Weibull distribution Loglik(model)= -3521 Loglik(intercept only)= -3529 Chisq= 15.32 on 5 degrees of freedom, p= 0.0091 Number of Newton-Raphson Iterations: 7 n= 775`				

From *p* values, it can be inferred that shoulder existence is the only significant variable in addition to the intercept, and the other variables are insignificant. Same significant variables are identified for "disabled vehicle" and "disabled truck" incident types. This picture is also reinforced by the preliminary analysis, since none of these variables found to affect the duration considerably. Based on regression results it can be said, that the dataset do not include the statistically significant parameters that can be used to determine the duration (this is in accordance with the descriptive findings of Task 3.2).

Bayesian Weibull Regression

WINBUGS software is used for Bayesian weibull regression with the same independent variables used in classical regression. The non-informative priors, having normal distributions with zero mean and precision 0.001 are assigned for covariate coefficients used in shape parameter. Scale parameter is assigned gamma distribution with a mean of 1 and precision of 0.01. Markov Chain Monte Carlo samples for each chain are produced within WINBUGS and then parameters are estimated. The WINBUGS doodle module used for regression can be seen in Figure 17. To assure convergence, 3 separate chains each having 10000 updates are run and the results show that convergence is achieved for all variables.

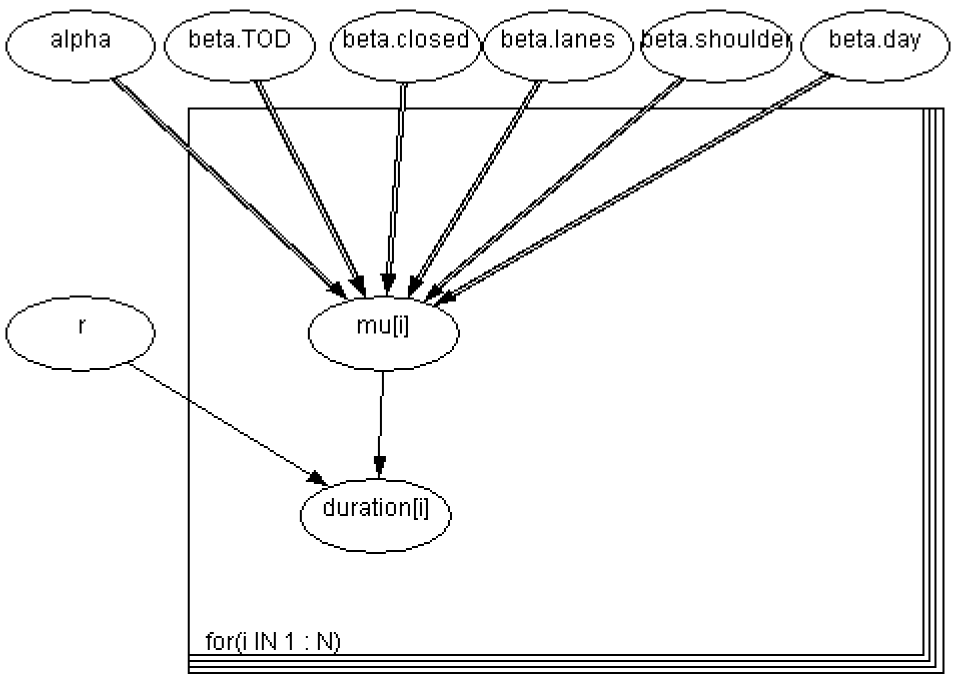


Figure 17- WINBUGS Model Used for Bayesian Weibull Estimation

In Table 14, WINBUGS output for property damage incident is given:

Table 14- WINBUGS Output for Property Damage Weibull Regression

node	mean	sd	MC error	2.5%	median	97.5%
Intercept	-5.034	0.3877	0.01925	-5.762	-5.05	-4.22
TOD	-0.0431	0.07281	7.208E-4	-0.1866	-0.04234	0.09749
# of Closed lanes	0.001388	0.07423	0.001281	-0.1463	0.001731	0.1444
Weekend/Weekday	-0.008491	0.08358	0.001171	-0.1721	-0.008735	0.1565
# of lanes	0.06242	0.1161	0.005682	-0.1788	0.06617	0.2821
Shoulder existence	-0.3899	0.1108	0.001358	-0.6092	-0.3886	-0.175
Scale Parameter	1.335	0.03761	0.001208	1.261	1.335	1.408

The convergence of coefficients is checked by tracing the MCMC updates. The MC error, which is basically an estimate of the difference between the mean of the sampled values are desired to be less than 5% of the sample standard deviation (sd), and this is satisfied by our regression.

Comparison of Classical and Bayesian Approaches

First point to be mentioned before comparison is that the default parameterization of weibull distribution employed in R and WINBUGS show slight differences. For instance they parameterize the scale and shape parameters differently. Hence, coefficients' signs for independent variables determined by using WINBUGS should be reversed and inverse of the shape parameter should be calculated to compare two software outputs. If done so, it can be seen that the order of coefficients are the same and the coefficient signs are the same for both analyses. The scale parameters are almost equal for both cases. Nevertheless other coefficients have relative differences. These facts do not reveal any important information about the goodness of the regression since the comparison of two approaches must be based on their predictive performances. It should be noted that for classical regression, the insignificant parameters are not used for the regression. For WINBUGS results, a random test sample that is 25% of the total sample is formed. Then a sensitivity analysis was performed by excluding variables and re-calculating the root mean square error (RMSE). The significance of the variables is decided based on the change in RMSE. The results show that, like in classical regression, shoulder is a significant factor. Besides, TOD is also found to affect the predictions, hence was kept in the WINBUGS analysis. Nevertheless, it can be said that the regression is mainly governed by the scale parameter which is not affected by the explanatory variables. Thus classical and Bayesian approaches are observed to perform in a similar manner as will be discussed in the following section.

Cross validation is employed to compare the model predictions. Random sub-sample, which is set to have about 20% of the overall data, is assigned as the test sample. The other half of the sample is set as the training data and prediction accuracy is measured using root mean square error (RMSE). The RMSE formula is given below:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=0}^{N}\left[\hat{F}(z_i) - F(z_i)\right]^2} \qquad (1)$$

where N is the test sample size, F and $\hat{F}$ are the real and the predicted duration values.

The cross validation performances of classical and Bayesian weibull regression for three incident types are shown in Table 15.

Table 15- RMSE Values of Classical and Bayesian Weibull Regression

Incident Type	Root Mean Square Error	
	R- Classical	WINBUGS- Bayesian
Property Damage	44.977	42.087
Disabled Vehicle	28.685	28.989
Disabled Truck	44.403	37.577

Table 15 shows that except the disabled vehicle incidents, Bayesian regression always performs better than classical regression for both disabled vehicle and truck. Number of significant variables are found to be only one (shoulder existence) for classical regression and two (shoulder existence and TOD) for the Bayesian analysis. The incident types are given in Table 15 according to their number of data points, property damage having the highest number of data points and disabled truck having the lowest. Property damage and disabled vehicle test samples have 387 and 334 records respectively, whereas disabled truck incident test sample has only 81 records. Hence, it can be said that the Bayesian estimation makes improvement in prediction when the sample size decreases. Nevertheless, the overall results show that the prediction error is not within an acceptable range

The same analysis shown above is also performed by incorporating the spatial information. Although spatial variables are found as significant, no major improvement is achieved.

<u>Decision Trees</u>

For decision trees, MATLAB statistics toolbox was used. The duration was predicted using the same independent variables used for the regression analysis. As done in regression analysis, each incident type was studied separately. If a CART tree has too many branches then there is a possibility that that it fits the data set well but predictions of the expected values will not be as good as the fit. Lower branches might be strongly affected by outliers and other properties of the current data set. A simpler tree that avoids these over-fitting related problems would be preferred. “The best tree size can be estimated by cross validation. First, a re-substitution estimate of the error variance for this tree and a sequence of simpler trees are computed, and plotted as the lower (blue) line shown in Figure 18. Then a cross-validation estimate of the same quantity is computed and plotted as the upper (red) line shown in Figure 18. The cross-validation procedure also provides an estimate of the “best” pruning level needed to achieve the optimal tree size” (Source: MATLAB Help). This level is found by choosing the minimum cost tree. Firstly, a general decision tree, incorporating incident type as the first node is formed with the same set of explanatory variables that are previously used:

- Number of Lanes
- Number of Closed Lanes
- Shoulder existence (No shoulder=0, Shoulder exists=1)
- Time of day (Off-peak=0, Peak=1)
- Weekday/weekend (Weekend=0, Weekday=1)

The resulting tree can be seen in Figure 19. Then, to be consistent and to be able to make comparisons with the previous regression analysis, each incident type is analyzed separately using the decision trees. The resulting trees for each incident type is given in Figure 20-Figure 22. It should be noted that cross validation is performed with random samples. Thus the size of the optimal tree can change according to random test and learning samples. For the overall best tree size, cross validation is performed numerous times and the tree size that occurs most frequently is chosen as the best tree size.

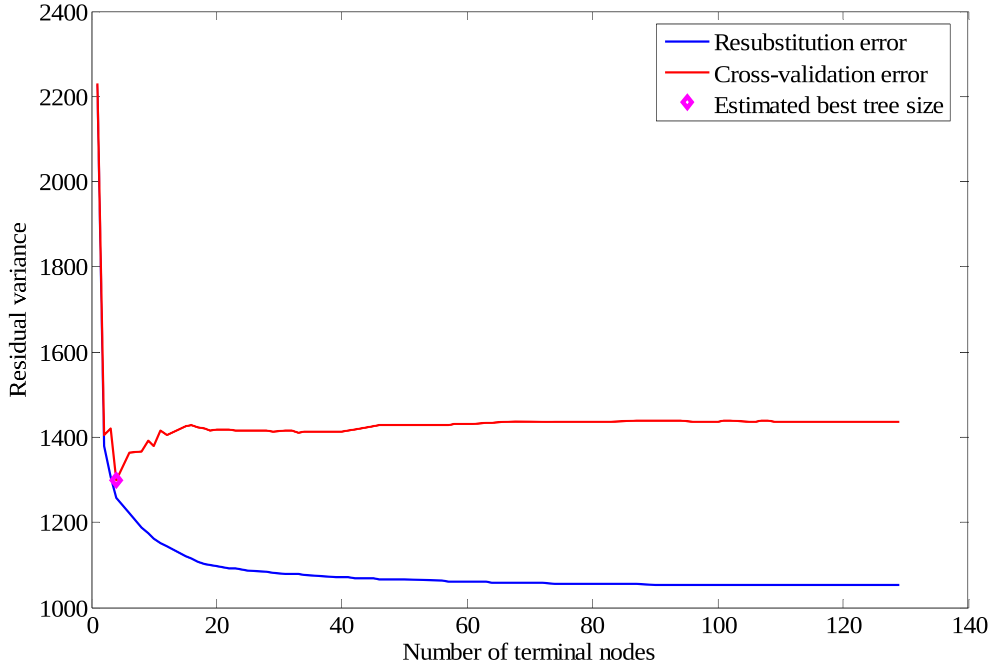


Figure 18- Best Tree Size Estimation for Overall Duration Data

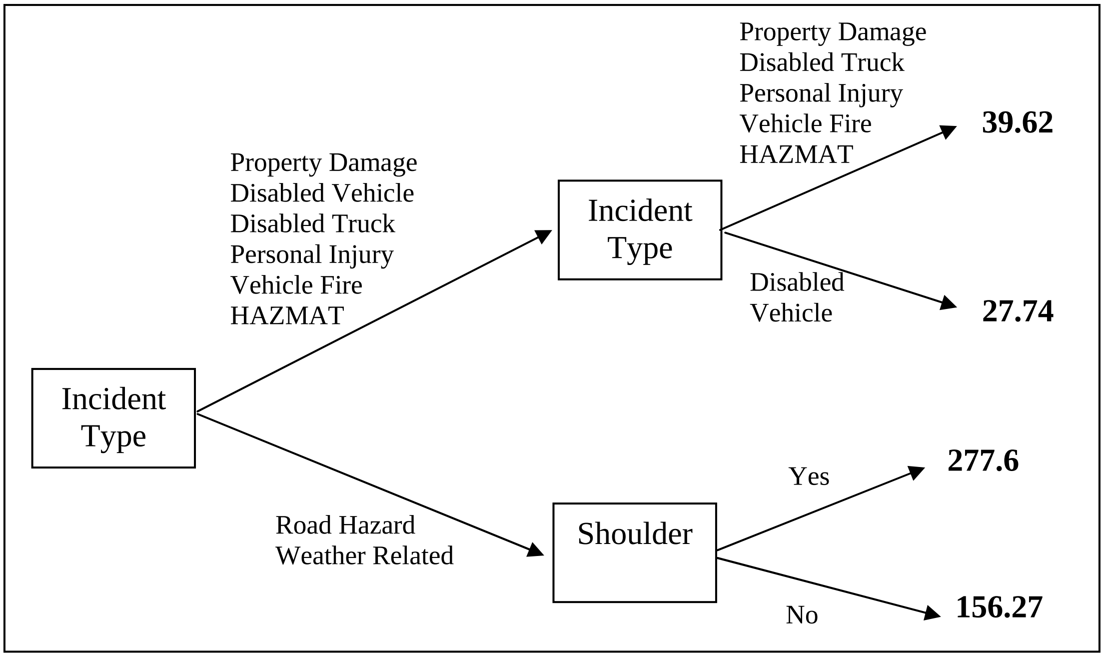


Figure 19- "Best" Tree for Overall TRANSCOM Duration Data

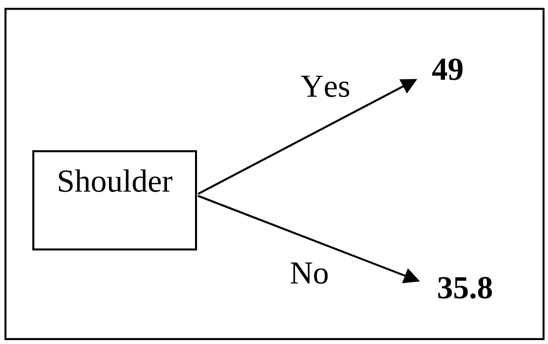


Figure 20- The Optimal Tree for TRANSCOM Property Damage Incidents

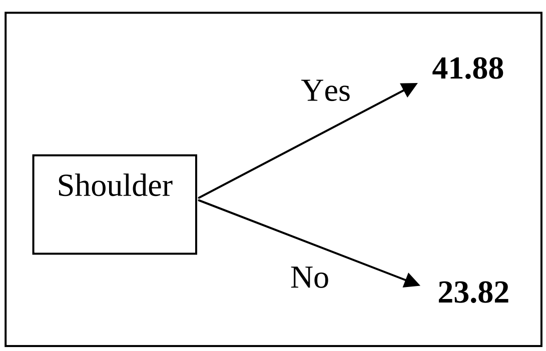


Figure 21- The Optimal Tree for TRANSCOM Disabled Vehicle Incidents

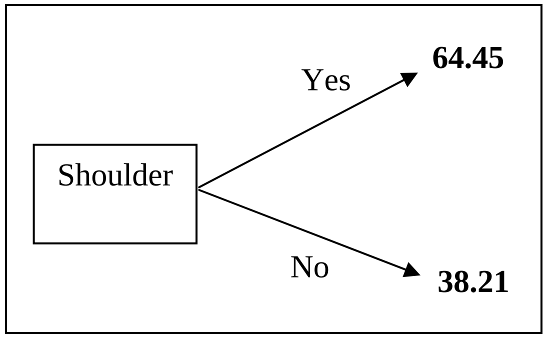


Figure 22- The Optimal Tree for Disabled Truck Incidents

It can be seen that the tree analysis also identifies incident type and shoulder existence as the most important parameters. The first node in the overall tree analysis (Figure 19) is the incident type, which also justifies the approach of treating each incident type separately for the regression analysis. When tree analysis is performed for each incident type, the shoulder existence is found to be the only parameter affecting the duration. This result is consistent with the previous regression results where shoulder existence is found to be the only significant variable. However, please note that the shoulder existence information is based on the facility that the incident had occurred. Although some partial shoulders exist in some portions of the facilities, the existence of shoulder is based on "Highway Sufficiency Rating File" and those partial shoulder/refuges are neglected. In the current analysis, the official records given in TRANSCOM dataset are used and no further personal judgments about the shoulder existence were employed.

To assess the impact of spatial information, new trees are formed based on the clusters obtained from the clustering analysis. The resulting trees are shown in Figure 23-Figure 25. Please note that for the "Disabled Truck" incidents, the optimal tree produces only node, thus only one value (~38.9 minutes). Another fact is that the tree fitted using the overall TRANSCOM data shows many branches however the trees found for individual incident types do not have many branches. Cluster information becomes the most and the only influential parameter if the incident cases are treated separately. This is not a surprising result since as it was mentioned in Task 3.2, the shoulder existence is highly related to the location. Some facilities, which equally means that some clusters, have shoulders all along but some does not. Thus, cluster parameter takes the place of shoulder existence when both parameters are present.

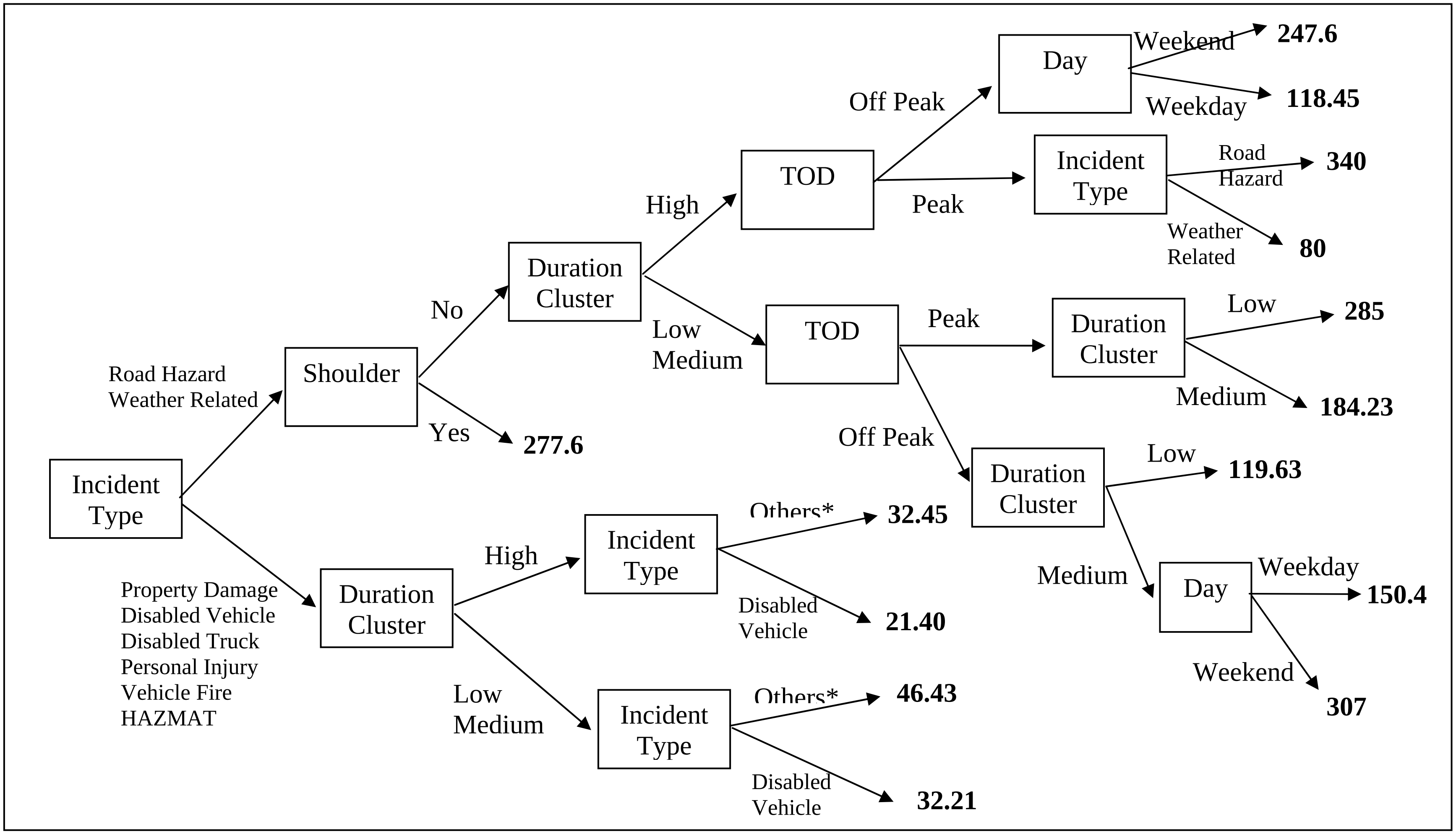


Figure 23- The Best Tree for the Overall TRANSCOM Duration Data Incorporating the Spatial Variation

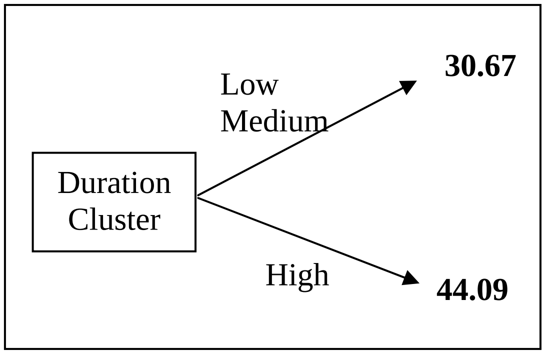


Figure 24- The Optimal Tree for TRANSCOM Property Damage Incidents Incorporating Spatial Information

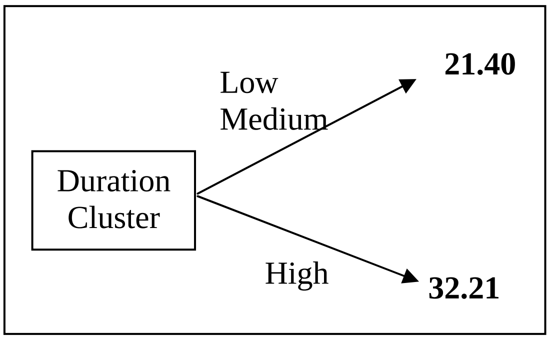


Figure 25- The Optimal Tree for TRANSCOM Disabled Vehicle Incidents Incorporating Spatial Information

Lastly, the same learning and test datasets as used before were used for cross validation and the resulting errors for both kinds of trees (spatial and non-spatial) are given in Table 16.

Table 16- Cross Validation Results of the Tree Analysis for the TRANSCOM Incident Duration Data

	Non-Spatial (secs)	Spatial (secs)
Overall Tree	35.3601	31.8473
Property Damage	27.9940	27.5796
Disabled Vehicle	19.4903	19.2076
Disabled Truck	30.8095	31.5608

If compared with the previous linear and weibull regression analysis, it can be seen that CART method gives considerably improved prediction results. However, addition of spatial information does not improve the prediction performance similar to the previous cross validation results of linear and weibull regressions. It can be also said that incident type specific analysis yields better predictions compared to overall prediction results. Overall decision trees give the best estimation results for the TRANSCOM duration data among tall of the models tested in this report.

3.0 Conclusions & Discussion for Incident Duration Estimation

As discussed in detail in Task 3.2 in detail, TRANSCOM dataset, which is the basis of the results presented below is found to cover only a small portion (around 8%) of the overall accident records for the study area. This conclusion is drawn after a comparison of NYSDOT accident dataset that cover the same corridor with the available TRANSCOM dataset. Nevertheless, interpretations of the duration analysis can still be valuable due to the fact that a relatively small sample of the total accident data was available to the research team the results of this analysis can be extended to represent the overall reality with sample size limitations. The results of this study are thus strictly restricted to available TRANSCOM records and should not be used as final recommendations for the study area.

The results of the incident duration estimation can be summarized as follows:

- None of the estimation models presented above succeeded to predict the incident durations within a good confidence level or with a desired detail.
  - Linear regression was found to perform poorly.
  - Weibull regression did not give promising prediction results both for classical and Bayesian approaches.
  - CART can be mentioned to produce results consistent with the findings in Task 3.2 and yielded the best cross validation results compared to linear and weibull regressions.
- The reason of poor performance of the prediction is more related to the dataset, than the used method. Unlike incident rate, which can be more related to physical factors, incident duration is also dependent on the emergency response rate, which the current dataset has no information about. Thus, the lack of more detailed information leads the study into using any reasonable parameter that is readily available but that might not be necessarily a major factor in terms of durations. For instance, below are two different studies using regression and decision trees for predicting the incident durations with the corresponding parameters that they employed.

Table 17 A Sample of Important Paramaters Found In the Literature to Affect The Incident Duration

Study	Parameters
Ozbay and Kachroo, (12) Tree Based Approach	▪ Heavy wrecker usage ▪ Assistance from response agencies ▪ Heavy vehicle involvement ▪ Severe injuries ▪ Extreme weather ▪ Caused freeway damage
Garip et.al. (9) Linear Regression	▪ Number of lanes affected ▪ Truck involvement ▪ Time of day ▪ Police response time ▪ Weather conditions

As shown in Table 17, both studies include an emergency agency response parameter. Regarding this current study, the factors that are used resembles to the ones adopted in Garib et.al.(9)'s, with the exception of police response factor. However our linear regression gives R-square value less than 0.1, where Garib et.al. (9) reported an R^2 equal to 0.81 for their dataset. On the other hand, there are very few common parameters with Ozbay and Kachroo's study. It can be concluded that the predictions that can be made with any model using this dataset will have problems mainly due to the limitations of the dataset. Finally, this dataset is determined to be a sub-set of a much larger accident dataset and it is important to emphasize that the failure to estimate statistically significant and reliable duration prediction models can also be due to the problems with the sample that contains TRANSCOM data.

Overall, for duration analysis, dataset is found to be incomplete for this study and does not allow for robust and reliable estimation of both parametric and non-parametric models. Including spatial information as an explanatory variable does not make big difference both for duration and frequency estimations. The CART model outperforms the other models in duration analysis.

RECOMMENDATIONS ABOUT COLLECTION OF ADDITIONAL TRANSCOM INCIDENT RECORD DETAILS

Although the TRANSCOM database provides valuable incident information, some additional records can yield significantly better explanatory power in terms of incident duration and frequency modeling and prediction. Thus, the research team suggests the following additional information to be collected to improve the TRANSCOM database[2]:

Milepost: As discussed in sectional analysis, the TRANSCOM database places the incident location by using relatively vague terms such as "near", "after" etc. with an intersection, or an exit name. Instead, milepost information can be recorded and the place of the incident can be determined more accurately and this prevents personal judgment errors. Then the incidents can be analyzed spatially to determine incident hotspots more accurately.

Detailed Information to Determine Accident Severity: Detailed accident related information such as property damage/injury/fatality and if injury/fatality, number of injuries and fatalities must be recorded.

Number and Types of Response Vehicles: The types and number of response and are frequently cited in the literature to affect the incident durations. Thus, this information has to be recorded. Arrival and departure times of response vehicles are also important data that will improve the existing data set.

Involvement of Heavy Vehicles: Heavy vehicle (e.g. heavy trucks) involvement is also cited in the literature to affect the incident duration, thus it is beneficial to know if an incident involves heavy vehicle (s).

[2] Although, in theory Item 2 to 7 are included in the IIMS database, they are listed here to emphasize the need for this data and to clearly indicate that in its current form the two databases are not well synchronized to enable the research team to get needed data points.

Incident Lane(s): The number of lanes affected is already recorded in TRANSCOM. However, the exact lane that the incident occurred, e.g. far left/right lane may change the response type and duration considerably. For instance, an incident in the right-most lane can be cleared easily in case there is shoulder, however an incident in the inner or far-left lane may need more time and effort.

Opening Sequence of Incident Lanes: When more than 1 lane is closed, then it would be very useful to know how long each lane stayed closed. This can be achieved if the incident management personnel who would respond the specific incident record the time each lane remained closed. This does not have to be done for all the incidents and can be done for a sample of incidents to avoid additional data collection burden by the response teams. A sample incident data collection form (Figure 26) shows what additional information can be useful in terms of the type of analysis done in this project.

Please note that the factors listed will provide several immediate details with respect to the TRANSCOM data mainly for better analysis of incident durations and frequencies. Additional driver related details such as age or gender can be used to model influence of driving habits or familiarity with the area on incident frequency. A reference template can be NJDOT crash records where detailed crash records are kept. However, while in NJDOT crash records duration information is missing and limits its usage for incident management studies. On the other hand, the TRANSCOM database has this valuable duration information. Hence, introducing some minor details to the existing TRANSCOM database will make it very useful in terms of incident management.

Additional Useful Information

Roadway characteristics (grade, curvature, merging/diverging section, number of lanes, shoulder existence etc.) are important features that are needed to model incident frequencies as well as durations. Thus, these fields must exist in any complete incident database. To collect this data, either a comprehensive sectional legend of the roadway should be provided to the incident management teams or police so that the person in charge can use the legends to enter the details. Or, a better solution which is less dependent on human perception and which will also reduce the reporting time, an offline database that can be created by using the milepost information and the roadway characteristics. The person in charge can be responsible for collecting only the milepost information and the corresponding roadway characteristics can be gathered automatically from this database based on the milepost. The procedure can be computerized easily. This requires a detailed sectional analysis of the roadway facilities. The majority of the information can be gathered from road construction design projects. However, use of such data collection technique needs precise milepost information, e.g. not using milepost intervals but the exact milepost or x-y coordinates that can be obtained by GPS. Of course most of these data already exist in GIS and is a matter of processing it for better use by incident management personnel.

Briefly, the needed fields are:

a. Number and width of lanes
b. Shoulder existence
c. Longitudinal features (Grade & curvature)
d. Lateral features (Merge/Diverge/Weaving/Basic sections)

Item a, and b are available through the Highway Sufficiency File. HSF is developed to record pavement conditions, and each link is created based on the pavement condition levels. If the definitions of the link are revised in order to match with physical characteristics of the roadway (item c, and d) that would better serve the purpose.

There is also a need for a clear guideline for recording incident type. At this time, more than 100 different inputs are given for incident types. Reporting authorities and TRANSCOM personnel should get training on defining incident types and minimize the number of incident type. The incident type description can still be recorded as text the incident descriptive field part of the TRANSCOM database.

FREEWAY INCIDENT CLEARANCE SURVEY

Please see on the back side of this page for Instructions

Please fill in the Following about the Incident Response

Date: ________ Your Last Name: ________________

Location: ______________________________

Agency Affiliated to: ______________________

Event *or* Case # ________ (OPTIONAL) Occurrence time of the Incident: ________

What is the Incident Type? Select One. Choose the type based on the Major Causative.

Disabled Vehicle *Property Damage* *Fatal Incident*
Road Hazard *Personal Injury* *Others*

Is an involved Vehicle on Fire ? *Yes* *No*

Is there a HAZMAT involved? *Yes* *No*

If *Yes*, select the type & Nature?

Hazmat State	Hazmat Nature	Hazmat Type
Solid	*Spilled Fuel*	*Poisonous*
Liquid	*Spilled cargo*	*Radioactive*
Gas	*Engine Fluid Spill*	*Inflammable*
	Cargo on Fire	*Others*
	Others	

Please fill in the Following about the Incident.

\# of Vehicles Involved → ☐
\# of Cars Involved → ☐
\# of Tractor Trailers → ☐
\# of Personal Injuries → ☐
\# of Fatalities Involved → ☐

Resources Used for Clearance of Incident

\# Sweepers → ☐ # Front End Loaders → ☐
\# Spreaders → ☐ # of Fire Engines → ☐
\# Cones → ☐ # Ambulances → ☐
\# Wreckers → ☐ # Arrow Boards → ☐
\# Sign Boards → ☐ Others (explain on back) → ☐
# of PERSONNEL USED → ☐

HAZMAT Crew. Select One. *Private* *Public* *None Used*

Fixed Post Traffic Control *Yes* *No*

Alternate Route Established *Yes* *No*

What is your Opinion on the Resource Availability for Clearing this Incident?

More than Adequate *Adequate*
Less than Adequate *Absolutely Inadequate*

What was the Time of Occurrence of the Incident? Please Select one.

6:00 AM - 10:00 AM *4:00 PM - 7:00 PM*
10:00 AM - 12:00 NOON *7:00 PM - 1:00 AM*
12:00 NOON - 2:00 PM *1:00 AM - 6:00 AM*
2:00 PM - 4:00 PM

Which one of the following Choices best Describes the Prevailing Weather?

Clear *Misty* *Sleet/Ice*
Cloudy *Rain* *Smoke/Dust*
Foggy *Snow* *Other*

Light & Temperature Select one each.

Bright *< 45 deg*
Satisfactory *45< t < 85*
Dark *> 85 deg*

Please Select the appropriate choice for the following aspects of the Location

Land Use Type

Open Land *Urban/Bldgs.* *Bridge/Tunnel*

Location Geometry *(Select ALL appropriate Choices)*

Straight *Up Hill* *Elevated Rwy.*
Curve *Down Hill* *Ramps*
Level

Right Shoulder *Left Shoulder*
Present *Absent* *Present* *Absent*

Lane Closure Information.
Select the Lanes Blocked by the Incident.
Edit Lane Geometry presented to suit the Location.

Left Shoulder	Lane 6	Lane 5	Lane 4	Lane 3	Lane 2	Lane 1	Right Shoulder

What was the CLEARANCE TIME for the Incident? Select the best Alternative.

Below 15 Mins.	Between 15 and 30 Mins.	Between 30 and 45 Mins.	Between 45 and 60 Mins.	Between 60 and 75 Mins.

If above 75 Mins., Indicate the exact time below.

☐ *Hours* and, ☐ *Mins.*

Please Return Completed forms to your Supervisor. Thank You for Your Cooperation.

Figure 26 Sample Incident Survey Form

REFERENCES

1. Nam, Doohee and Fred Mannering. "An Exploratory Hazard-Based Analysis of Highway Incident Duration." *Transportation Research – A*. Vol. 34A, No. 2, 2000. pp 85-102.
2. Kevin Smith and Brian L. Smith, "Forecasting the Clearance Times of Freeway Accidents", *Final report of ITS Center project: Incident duration forecasting*, Center for Transportation Studies, University of Virginia, 2001.
3. Golob, Thomas F., Wilfred W. Recker and John D. Leonard. "An Analysis of the Severity and Incident Duration of Truck-Involved Freeway Accidents." *Accident Analysis & Prevention*. Vol. 19, No. 4, August 1987. pp 375-395.
4. Garib, A., A. E. Radwan and H. Al-Deek. "Estimating Magnitude and Duration of Incident Delays." *Journal of Transportation Engineering*. Vol. 123, No. 6, November/December 1997. pp. 459-466.
5. Giuliano, Genevieve. "Incident Characteristics, Frequency, and Duration on a High Volume Urban Freeway." *Transportation Research – A*. Vol. 23A, No. 5, 1989. pp. 387-396.
6. Nam, Doohee and Fred Mannering. "An Exploratory Hazard-Based Analysis of Highway Incident Duration." *Transportation Research – A*. Vol. 34A, No. 2, 2000. pp 85-102.
7. Kaan Ozbay and Pushkin Kachroo. *Incident Management in Intelligent Transportation Systems*. Boston, MA: Artech House, 1999.
8. Lasse Fridstrom, Jan Ifver, Siv Ingebrigtsen, Risto Kulmala, and Lars Krogsgard Thomsen, "Measuring The Contribution Of Randomness, Exposure, Weather, And Daylight To The Variation In Road Accident Counts", *Accident Analysis and Prevention*, Vol.27 (1), (1995), pp.1-20.
9. Venkataraman Shankar, Fred Mannering, and Woodrow Barfield, "Effect Of Roadway Geometrics And Environmental Factors On Rural Freeway Accident Frequencies", *Accident Analysis and Prevention*, Vol.27 (3), (1995), pp.371-389.
10. K.W. Ogden, "The Effects of Paved Shoulders on Accident on Rural Highways", *Accident Analysis and Prevention*, Vol.29 (3), (1997), pp.353-362.
11. Charlesworth, K.D., Survey of Shoulder Usage. *Australian Road Research Board Internal Report 404-4*, p. 43. ARRB, Melbourne, 1985.
12. Jean-Louis Martin, "Relationship between crash rate and hourly traffic flow on interurban motorways", *Accident Analysis and Prevention*, Vol.34, (2002), pp.619-629.
13. Thomas F. Golob, Wilfred W. Recker, "A method for relating type of crash to traffic flow characteristics on urban freeways", *Transportation Research Part A*, Vol.38, (2004), pp.53–80.
14. Jovanis P & Chang H, "Modeling the relationship of accidents to miles traveled", *Transportation Research Record* 1068, (1986), pp. 42–51.
15. John Milton And Fred Mannering, "The Relationship Among Highway Geometrics, Traffic-Related Elements And Motor-Vehicle Accident Frequencies", *Transportation*, Vol.25, (1998), pp. 395–413.
16. Joshua S & Garber N, "Estimating truck accident rate and involvement using linear and Poisson regression models", *Transportation Planning and Technology*, Vol. 15, (1990), pp.41–58.
17. Miaou SP, Hu P, Wright T, Rathi A, & Davis S, *Relationships Between Truck Accidents and Highway Geometric Design: A Poisson Regression Approach.* Presented at the 71st Annual Transportation Research Board Meeting, Washington DC, 1992.

18. Miaou SP & Lum H, “Modeling vehicle, accidents and highway geometric design relationships”, *Accident Analysis and Prevention,* Vol.25(6), (1993), pp.689–709.
19. Miaou, S.P., “The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions”, *Accident Analysis and Prevention*, Vol.26 (4), (1994), 471–482.
20. Poch, M., Mannering, F., “Negative binomial analysis of intersection—accident frequencies”, *Journal of Transportation Engineering*, Vol.122 (2), (1996), pp.105–113.
21. Hadi, M., Aruldhas, J., Chow, L.-F., Wattleworth, J., “Estimating safety effects of cross-section design for various highway types using negative binomial regression”, *Transportation Research Record,* Vol.1500, (1995), pp.169–177.
22. McCarthy, P.S., “Public policy and highway safety: a city-wide perspective” *Regional Science and Urban Economics*, Vol.29 (3), (1999), pp.231–244.
23. Carson, J., Mannering, F., “The effect of ice warning signs on ice-accident frequency and severity”, *Accident Analysis and Prevention*, Vol.33 (1), (2001), pp.99–109.
24. Shankar, V.N., Milton, J., Mannering, F., “Modeling accident frequencies as zero-altered probability processes: an empirical inquiry”, *Accident Analysis and Prevention*, Vol.29 (6), (1997), pp.829–837.
25. Lee, J., Mannering, F., “Impact of roadside features on the frequency and severity of run-off-roadway accidents: an empirical analysis”, *Accident Analysis and Prevention,* Vol.34 (2), (2002), pp.149–161.
26. Lee, A.H., Stevenson, M.R., Wang, K., Yau, K., “Modeling young driver motor vehicle crashes: data with extra zero”, *Accident Analysis and Prevention,* Vol.34 (4), (2002), pp.515–521.
27. Jin, X., Cheu, R.L., Dipti, S., “Development and adaptation of constructive probabilistic neural network in freeway incident detection”, *Transportation Research Part C*, Vol.10 (2), (2002), pp.121–147.
28. Yuan, F., Cheu, R.L., “Incident detection using support vector machines”, *Transportation Research Part C* , Vol11 (3-4), (2003), pp.309–328.
29. Li-Yen Chang, “Analysis of freeway accident frequencies: Negative binomial regression versus artificial neural network”, Safety Science, Vol.43, (2005), pp.541–557.
30. Li-Yen Chang and Wen-Chieh Chen, “Data mining of tree-based models to analyze freeway accident frequency”, Journal of Safety Research, Vol.36, (2005), pp.365 – 375.
31. Karolien Geurts A, Isabelle Thomasb, Geert Wets A, “Understanding Spatial Concentrations Of Road Accidents Using Frequent Item Sets”, Accident Analysis And Prevention, Vol.37, (2005), pp.787–799.
32. Wen Cheng and Simon P. Washington, “Experimental Evaluation Of Hotspot Identification Methods”, Accident Analysis And Prevention, Vol.37, (2005), pp.870–881.
33. Geurts, K ; Wets, G ; Brijs, T ; Vanhoof, K, “Identification And Ranking Of Black Spots: Sensitivity Analysis”, Transportation Research Record, No. 1897, (2004), pp.34-42.
34. Dominique Lorda, Simon P. Washington and John N. Ivan, “Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory”, *Accident Analysis and Prevention*, Vol. 37, (2005), pp.35–46.
35. R Statistical Package. Available at: http://www.r-project.org/
36. BUGS Project; WINBUGS. Available at: http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml

APPENDIX

INCIDENT RATE ESTIMATION

We will first give a brief review of the most relevant studies to motivate the modeling approach adopted in our study[3]. In the literature, different models are proposed for incident frequency modeling. Early studies use linear regression. However, as pointed out by Jovanis and Chang (19) the increase in the variance of the accident frequency generally violates the homoscedasticity requirement of the linear regression. Linear regression is also not restrained from predicting negative accident frequency, which becomes an important point for low accident rate regions (20). As a result of these findings, follow-up studies focused on Poisson regression (17) and results show that the Poisson approach is superior to linear regression. One important point to note is that Miaou *et.al.*(0) find AADT, horizontal curvature and vertical grade to be significantly correlated with truck accident rate in North Carolina. These variables are also widely used in other accident frequency studies that are not limited to truck accidents. However, the dataset used for NRD study lacks curvature and vertical grade data and only have AADT values.

Recent practice for accident frequency analysis mainly relies on Poisson and Poisson-gamma models (negative binomial regression). Earlier studies used Poisson regression. However, the equal mean and variance requirement of Poisson regression prevented its further usage without important statistical problems. Thus, Poisson-gamma regression, which is a modified version of Poisson regression and also called negative binomial is used in most of the recent studies (0,0,21,22,23). However, negative binomial model also loses its power due to over dispersion caused by increasing number of zero accident regions. Thus zero-inflated (zero-altered) models are employed to overcome this problem (24,25,26). Both Poisson and negative binomial regressions are performed under zero-inflated conditions (namely ZIP for zero-inflated Poisson and ZINB for zero-inflated negative binomial). There are recent attempts to use Artificial Neural Networks (ANN) in incident prediction (27,28,29). Chang (29), compares ANN approach to negative binomial regression. Although it is stated that although ANN model is more time consuming and harder to build, it has additional advantages over negative binomial regression, such as having no distributional assumptions underlying the independent variables. Moreover ANN's ability to handle correlations between independent variables is another advantage over other approaches. ANN is also claimed to have better prediction results in one or more accidents, but performs slightly worse in zero incident regions. The underlying assumptions of Poisson, negative binomial, ZIP and ZINB models also addressed in Chang and Chen [30] suggests the use of Classification and Regression Tree (CART), which is a nonparametric tool that does not require any a priori distribution or variable selection. In the same study, the CART and negative binomial regression models are found to provide similar results in terms of incident frequency prediction performance on the training data and test data, which demonstrates that CART analysis, is an appropriate methodology for analyzing the frequency of traffic accidents. The overall model prediction accuracy of CART model, is found to be 58.2%, 52.6% for the training and the testing data, respectively. These measures are found to be 52.9% and 52.3% for the negative binomial model. Thus the CART model performs slightly better than the negative binomial regression model in predicting the training data. Also stated in the same study, in case of incident frequency prediction, the CART model performs better than the negative binomial

[3] A detailed literature review is conducted in Task 2 of this study.

model for the highway sections with one or more accidents, but the accuracy is relatively low. The negative binomial model performs slightly better than the CART model for the sections with zero accidents. One important feature of CART model mentioned in the study is that the CART model relies more on traffic and environmental variables than geometric and location variables (which does not exist in out dataset) to classify accident frequencies on the freeway sections.

Another approach for incident rate prediction is to pinpoint the black spots, which have the highest rate of incidents (31,32,33). However their usage is more towards dealing with certain sections rather than exploring the overall roadway incident statistics.

Summing up the discussion above, the following are the regression type models that are most frequently used in the accident frequency studies:

1. Linear regression
2. Poisson regression
3. Negative binomial regression
4. Zero-inflated Poisson regression (ZIP)
5. Zero-inflated negative binomial regression (ZINB)
6. Artificial Neural Networks (ANN)
7. Classification and Regression Trees (CART)

Linear regression is not a good candidate because previous studies clearly show linear regression's poor performance compared to Poisson and negative binomial due to the reasons mentioned above. The studies using zero-inflated regression models mostly deal with long, rural highway sections with lower AADT values compared to NYC roads studied in this report (Table A-1). In these cases, it is likely to have road sections with zero incidents. For our study, the region of interest is about 45 miles long and has a one-way AADT around 55000 vehicles. Figure 16 depicts the fact that there are not many roadway sections in our study area without incident points, except the unreported bridge incidents. Thus, our study area does not warrant the use of zero-inflated models.

Table A-1 Summary of Studies Using Zero-inflated Models

(Source: Lorda et.al (34))

Study	Location	Crashes per year				Exposure				Observations
		Min	Max	Mean	Std	Min	Max	Mean	Std	7610/653
Miaou	RPA	0	8	0.2	–	0.0008 [a]	5.03 [a]	0.25 [a]	–	N/A
Shankar et.al.	RPA	0	84	0.29	1.09	251	26415	4534	4255	N/A
	RMA	0	7	0.09	0.35	187	11016	1691	1537	N/A
	RCA	0	6	0.61	0.28	146	10438	982	931	N/A
Lee & Mannering	RPA	0 [b]	9[b]	0.11 [b]	0.37 [b]	988	65222	2194	998	120 [c]
Qin et.al.	RSV	0	61	0.68	–	240	40000	992	–	19480/10320
	RSD	0	23	0.15	–	240	40000	992	–	26640/3160
	ROD	0	7	0.04	–	240	40000	992	–	28609/1191
	RID	0	23	0.08	–	240	40000	992	–	28068/1132
Shankar et.al.	Suburban	0	4	0.16	0.476	10350 [d]	98875 [d]	25425 [d]	20475 [d]	440 [c]
Kumara & Chin	Urban Intersection	0	6	0.29	0.70	1093 [e]	66378 [e]	12992 [e]	9479 [e]	2238/542

* RPA:Rural Principal Arterial, RMA: Rural Minor Arterial, RCA:Rural Collector Arterial, RSV: Rural Single Vehicle,
RSD: Rural Same Direction, RSD: Rural Opposite Direction, RSD: Rural Intersection

a: Number of trucks in veh x miles x 10^6

b: Crashes per month

c: Total number of observations

d: AADT for vehicles

e: Total entering flow

Regarding the Poisson and negative binomial regression, major problem comes out to be the lack of detail in terms of the physical roadway characteristics such as grade, curvature etc. Original data set available to the research team does not contain any of these geometric characteristics, but similar to the approach adopted in the descriptive analysis section, subjective (observation based) information about the curvatures can be used. However, this information does not provide a reliable numeric value that can be used in the regression. Table A-2 shows some common parameters used in the regression studies found in the literature (Our study dataset has the italicized parameters only). In the literature, larger number of variables is employed for the analysis of incidents at intersections but they are not mentioned in this section because they are irrelevant to our study. As clearly seen in Table A-2, our study dataset lacks some important data. Some of the missing data like environmental conditions can be gathered by available weather reports, but the incident location specific data are almost impossible to gather accurately. Thus, a more general approach, where the generic and relatively less accurate section properties determined with the help of aerial photos can be employed. Non-parametric classification and regression tree analysis seem more appropriate for this type of analysis.

It should be noted that incident rate should have easily understandable units for better interpretation. Most common unit is the number of incidents per unit distance per unit time. This basic unit can be easily converted to incidents per day, or incident per a million vehicles traveled. Nevertheless, a section with a known length must be given to apply those converted units. Since our dataset does not include any milepost records, the lengths of sections are gathered from secondary sources. The NYSDOT charts, which were also used to extract the AADT values along the study area, include length of sections between exits. This data is used together with "From Location" field in the dataset. The length associated with the incident section having the exit name at “From Location” field is assumed to be the section length. In other words, the roadway is divided into smaller segments between each exit and incident counts are assigned to these segments accordingly. This is similar to what was done for cluster analysis, however for that analysis the sections were set to be “at” and “between” the exits and length data to be used with this kind of sectioning does not exist. On the other hand, the new sectioning is problematic since NYSDOT charts do not necessarily include lengths between the exits. For example, some fields cover the length between several exits at a time. In these cases, the length between two consecutive exits is assumed to be same. The rates at each section are calculated using these extracted distances.

Table A-2 Parameters Used in the Literature for Poisson, NB, ZIP, ZINB Regressions

Traffic Flow Factors	Road Characteristics	Environmental Conditions
Speed *AADT* *Frequency* Truck Volume	*Length of section* Horizontal curve radius Degree of horizontal curvature Total length between horizontal curves Sharp curve *Number of lanes* Flat section *Straight section* Narrow center right section # of horizontal curves for design speeds Number of horizontal curves in section Max/Min horizontal curve radius Number of vertical curves in section Max/ Min grade in section	Average monthly rainfall Maximum daily rainfall/month Number of rainy days/month Average monthly snowfall Maximum daily snowfall/month Number of snowy days/month

Linear Regression

Although the poor performance of the linear regression is frequently shown in the literature, it was still used in this study for completeness purposes. The statistics software R was used, to estimate the parameters of the following independent variables:

- Shoulder Existence (No shoulder=0, Shoulder exists=1)
- Weekend/Weekday (Weekend=0, Weekday=1)
- Time of Day (offpeak=0, peak=1)
- Number of Lanes
- AADT

The frequency clusters found in the clustering analysis were used as location parameters. Regression was performed for each individual incident type. Below, the “R” output for the most frequent incident type (property damage) in the highest incident rate region is presented. As it can be seen from the results, R-square value is very poor for this incident type and section, which constitutes a major portion of the overall records.

```
Residuals:
      Min        1Q    Median        3Q       Max
-0.018351 -0.010307 -0.006257  0.001729  0.177555

Coefficients:
                Estimate   Std. Error   t value  Pr(>|t|)
(Intercept)    9.814e-03   8.883e-03    1.105    0.26989
Shoulder      -7.240e-03   2.844e-03   -2.546    0.01126 *
NumberofLanes -9.084e-04   2.777e-03   -0.327    0.74375
TOD           -1.571e-03   2.082e-03   -0.755    0.45093
Day            6.814e-03   2.114e-03    3.224    0.00137 **
AADT           8.708e-08   7.901e-08    1.102    0.27105
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.02093 on 412 degrees of freedom
Multiple R-Squared: 0.03933,    Adjusted R-squared: 0.02767
F-statistic: 3.373 on 5 and 412 DF,  p-value: 0.005349
```

Shoulder existence and weekday/weekend are identified as significant variables within 1% and 0.1% confidence levels. However, similar to the duration study, linear regression did not perform well (Extremely poor R-Squared values). Linear regression is also employed with the adjusted dataset that includes the spatial information obtained from the clustering analysis. Nonetheless, regression still yields poor R-squared values. Overall, both with and without spatial information, the linear regression generates very poor R-Squared values and was dropped as a possible modeling methodology for Task 4.1.

Negative Binomial Regression

Following the practice in the literature, negative binomial regression was estimated using the same independent variables used in linear regression. Below is the *R* output for the negative binomial regression. The statistics for individual incident type (property damage) is given below.

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.0431  -0.9499  -0.4790   0.4981   2.0852

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)    2.006e+00  6.151e-01   3.261  0.00111 **
Shoulder      -7.540e-01  1.783e-01  -4.230 2.34e-05 ***
NumberofLanes -3.406e-01  1.852e-01  -1.839  0.06589 .
TOD           -6.826e-02  1.395e-01  -0.489  0.62472
Day            7.530e-01  1.437e-01   5.242 1.59e-07 ***
AADT           7.813e-06  4.891e-06   1.598  0.11015
Length        -2.544e-02  1.887e-01  -0.135  0.89275
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(2.0371) family taken to be
1)

    Null deviance: 187.16  on 144  degrees of freedom
Residual deviance: 141.45  on 138  degrees of freedom
AIC: 784.74

Number of Fisher Scoring iterations: 1

Correlation of Coefficients:
              (Intercept) Shoulder NumberofLanes TOD   Day   AADT
```

```
Shoulder      -0.29
NumberofLanes -0.84        0.23
TOD           -0.15        0.03     0.04
Day           -0.07       -0.13    -0.03          -0.05
AADT          -0.25        0.00    -0.20           0.02 -0.04
Length        -0.30        0.14     0.08           0.02 -0.05 -0.10

              Theta:  2.037
          Std. Err.:  0.319

 2 x log-likelihood:  -768.736
```

Negative binomial analysis for "property damage" incidents identifies "shoulder existence", "weekend/weekday" and "number of lanes" variables as significant. These results are not consistent with the literature where AADT is mentioned to be the most parameter for incident frequency estimation (7,12,0). However, same analyses for other incident types identify different sets of variables as significant. For instance, for "disabled vehicle" AADT is found to be significant whereas "number of lanes" is found insignificant. For "disabled truck" only shoulder existence and day are found to be significant. Overall, NB regression results do not yield same set of significant variables for each incident type.

Secondly, as done in linear regression, the NB regression is performed with using cluster information regarding low, medium and high incident rate regions. The R software results are given below:

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.3222  -0.8981  -0.3990   0.2969   2.3183

Coefficients:
                Estimate Std. Error z value Pr(>|z|)
(Intercept)   -2.601e-01  7.143e-01  -0.364  0.71576
Cluster        5.679e-01  1.301e-01   4.364 1.28e-05 ***
Shoulder      -2.385e-01  2.314e-01  -1.031  0.30268
NumberofLanes -9.143e-02  2.103e-01  -0.435  0.66379
TOD           -1.497e-02  1.552e-01  -0.096  0.92319
Day            6.750e-01  1.612e-01   4.188 2.81e-05 ***
AADT           1.427e-05  5.140e-06   2.776  0.00551 **
Length        -1.561e-01  2.115e-01  -0.738  0.46043
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(1.6284) family taken to be
1)

    Null deviance: 202.60  on 141  degrees of freedom
Residual deviance: 136.17  on 134  degrees of freedom
AIC: 767.08

Number of Fisher Scoring iterations: 1

Correlation of Coefficients:
              (Intercept)  Cluster  Shoulder  NumberofLanes  TOD      Day
AADT
Cluster       -0.26
Shoulder      -0.38        0.55
NumberofLanes -0.76       -0.17    0.07
TOD           -0.13        0.05    0.05     0.01
Day           -0.12        0.04   -0.08    -0.02         -0.05
```

```
AADT           -0.18       -0.01    0.05    -0.23          0.05  0.03
Length          -0.34       -0.04    0.09     0.17           -0.05 -0.08 -
0.13

              Theta:  1.628
          Std. Err.:  0.239

 2 x log-likelihood:  -749.075
```

Including spatial information in the dataset changes the regression considerably for "property damage" incidents. Log-likelihood which can be thought as the goodness of fit gets better. The number of variables identified as significant also changes. AADT is identified as significant which agrees with the literature. When cluster information is included, the significance of "shoulder existence" diminishes. This is mainly due to the fact that shoulder existence is related to location and shoulder existence parameter remains constant along consecutive sections. Nevertheless, just like in non-spatial analysis, the set of parameters that are found to be significant changes for different incident types. For instance, significant parameters for "disabled vehicle" agree with "property damage" incidents, however for "disabled truck" incidents, AADT is not found to be a significant parameter. However, since "property damage" and "disabled vehicle" incidents constitute a major portion of the overall TRANSCOM data, it can be said that analysis results are not consistent among incident types however consistent for the majority of the records.

Overall, NB regression is considered to be further analysis since it has identified the same as in Task 3.2 to have an effect on the frequency rates. Below are the cross validation results for both spatial and non-spatial NB regression analysis.

Table A-3 Cross Validation Results for NB Regression

	Non-Spatial	Spatial
Property Damage	0.026137	0.031147
Disabled Vehicle	0.034185	0.04489
Disabled Truck	0.012111	0.015707

It can be seen that including cluster information to the regression does not improve the cross-validation RMSEs. Thus, although a better log-likelihood result is obtained after including cluster information in the regression, the prediction performance for incident rates gets worse.

CART Analysis

The same kind of analysis performed for durations is also done for incident frequencies. Two analyses, one having no spatial information and one having cluster information, are performed. The parameters are chosen based on the descriptive analysis and the ones which are found to affect the frequencies the most (Shoulder existence, AADT) are used. Other factors are added to check the possible coupled effects. The main reason for adding these parameters is to maximize the use of the available data. Below are the parameters that are used in the estimation:

- Existence of Shoulder: As shown in descriptive analysis, found to affect the frequency considerably
- Number of lanes: Intuitively expected to affect the incident frequency.
- AADT: Identified in almost all the literature as a variable that affects the accident rates
- Time of day (TOD)
- Weekday/weekend: Found to have a small impact on incident frequencies, however still added to the analysis dataset.

It is worth noting that dataset contains another important information cited in the literature namely, lane width. However, all the facilities in the dataset have the same lane width and there are negligible number of records (10, 11 and 12 feet for 35, 16 and 1858 records respectively) reported to be different, thus lane width is not included in the analysis.

A MATLAB based toolbox was used to build the CART tree structure. With the same reasoning stated for the duration tree analysis, Figure A-1 is drawn with cross-validation (upper red curve) and re-substitution (lower blue curve). The resulting best tree is shown in Figure A-2. The optimal trees for "property damage" and "disabled vehicle" incident types are also given in Figure A-3 and Figure A-4 respectively. Please note that optimal tree for "disabled truck" incidents has only one node, which means that all disabled truck incidents are assigned the same rate; numerically 0.0087836 incidents per mile per day.

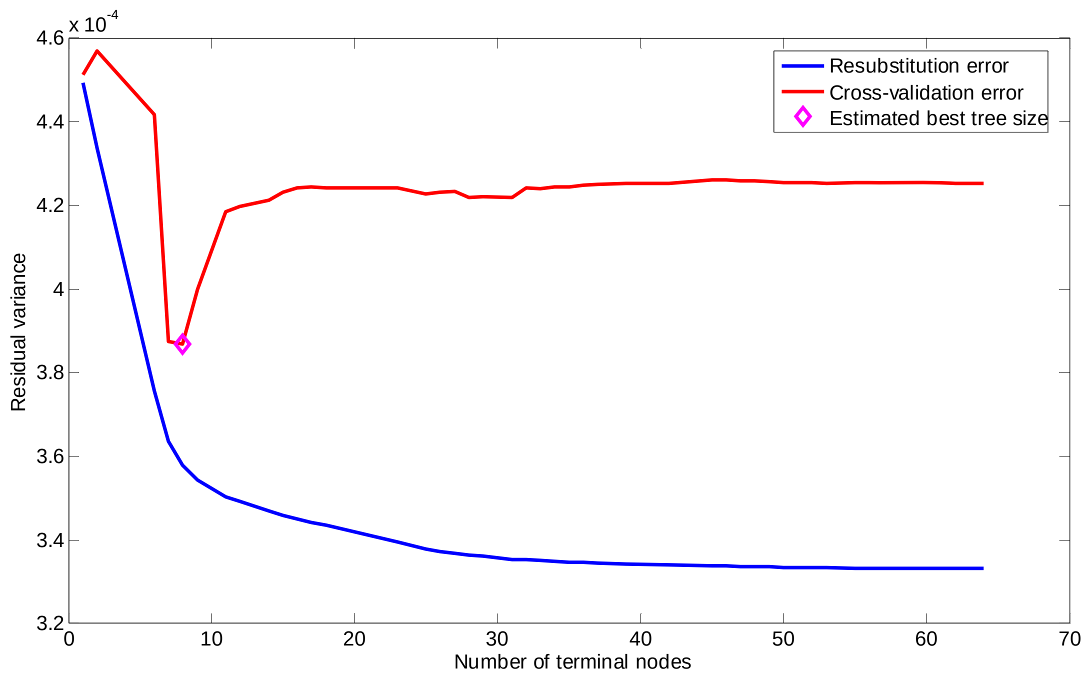


Figure A-1 Estimation of Optimal Tree Size for Overall TRANSCOM Incident Frequency

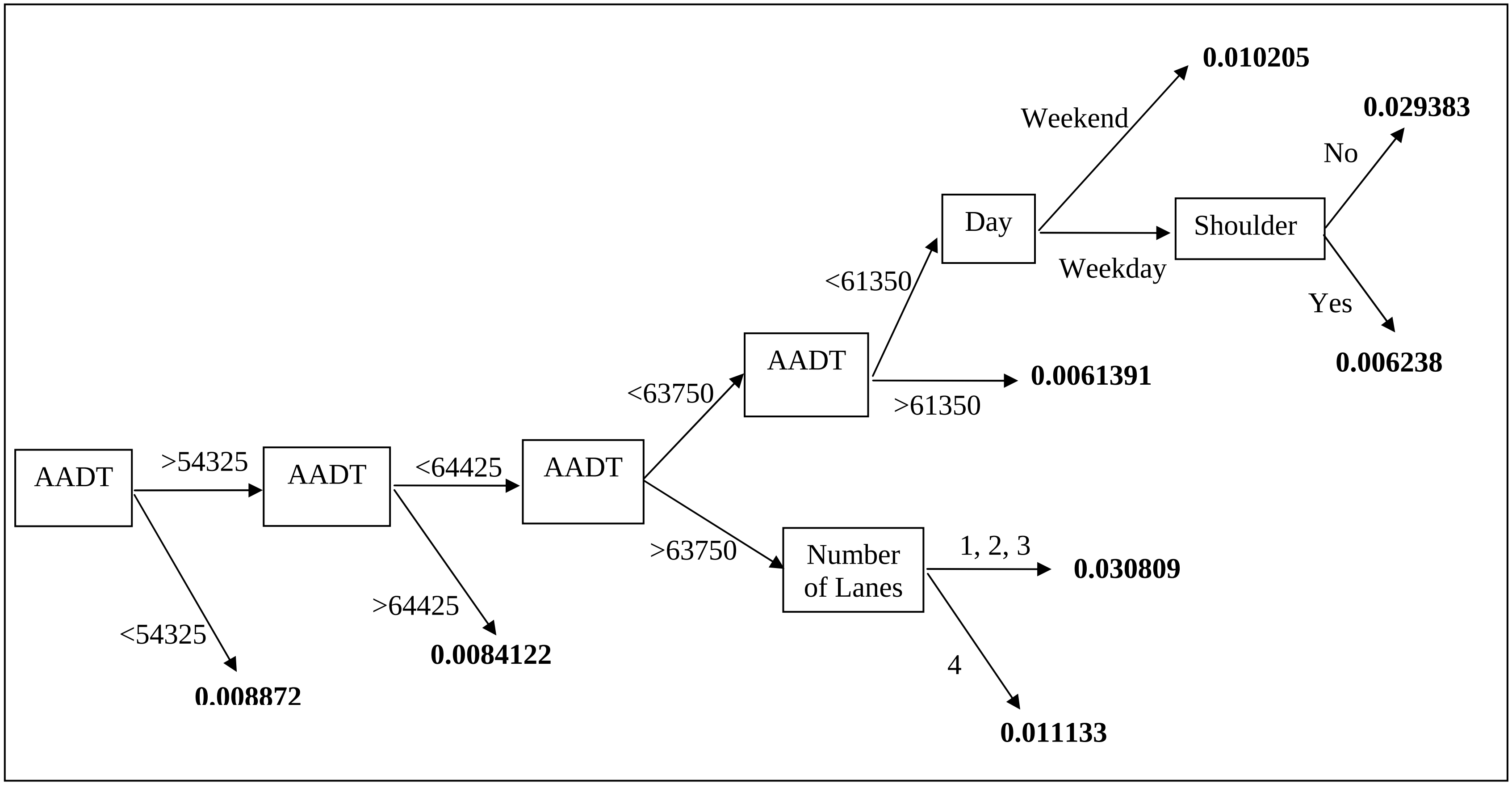


Figure A-2 Optimal Tree for Overall TRANSCOM Incident Frequency [Incidents/Mile/Day]

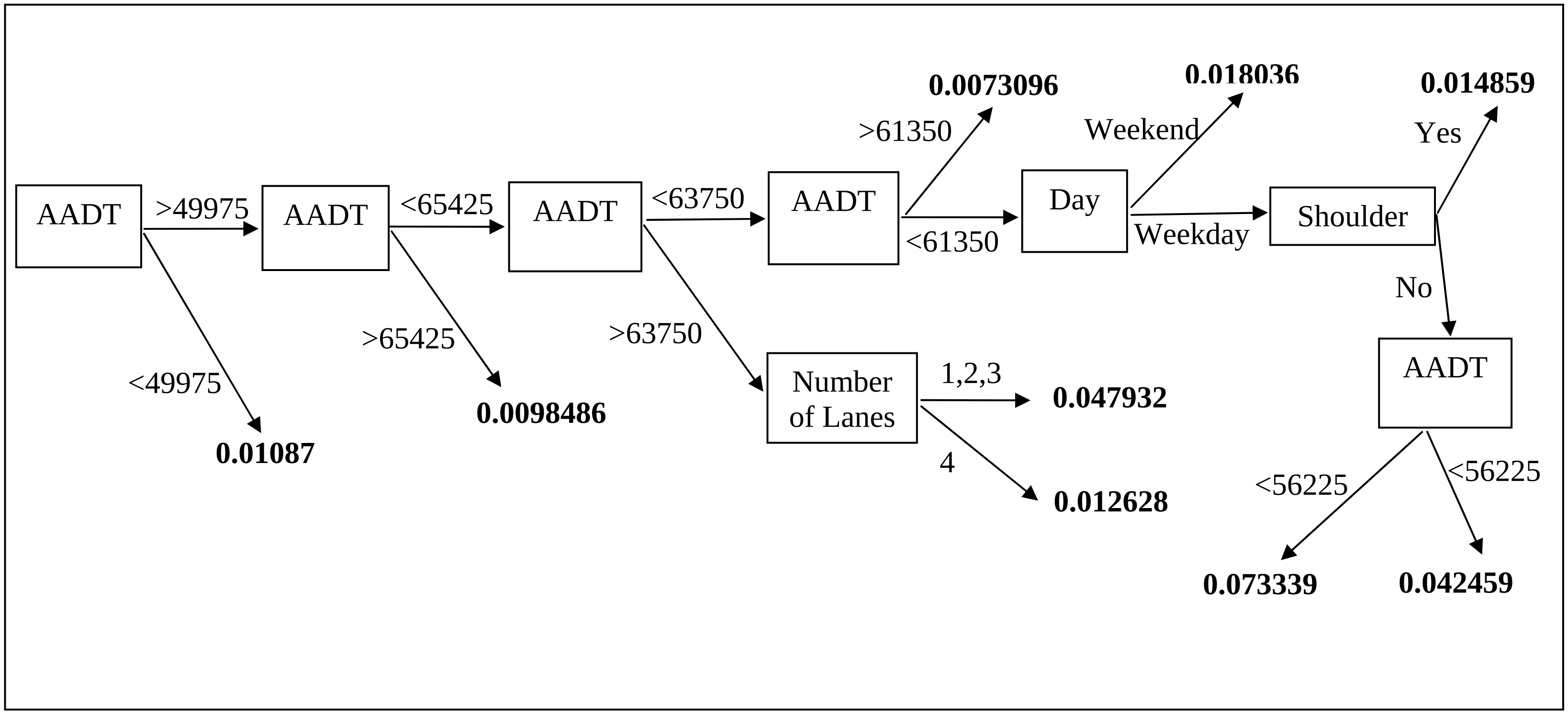


Figure A-3 Optimal Tree for TRANSCOM Property Damage Incident Frequency [Incidents/Mile/Day]

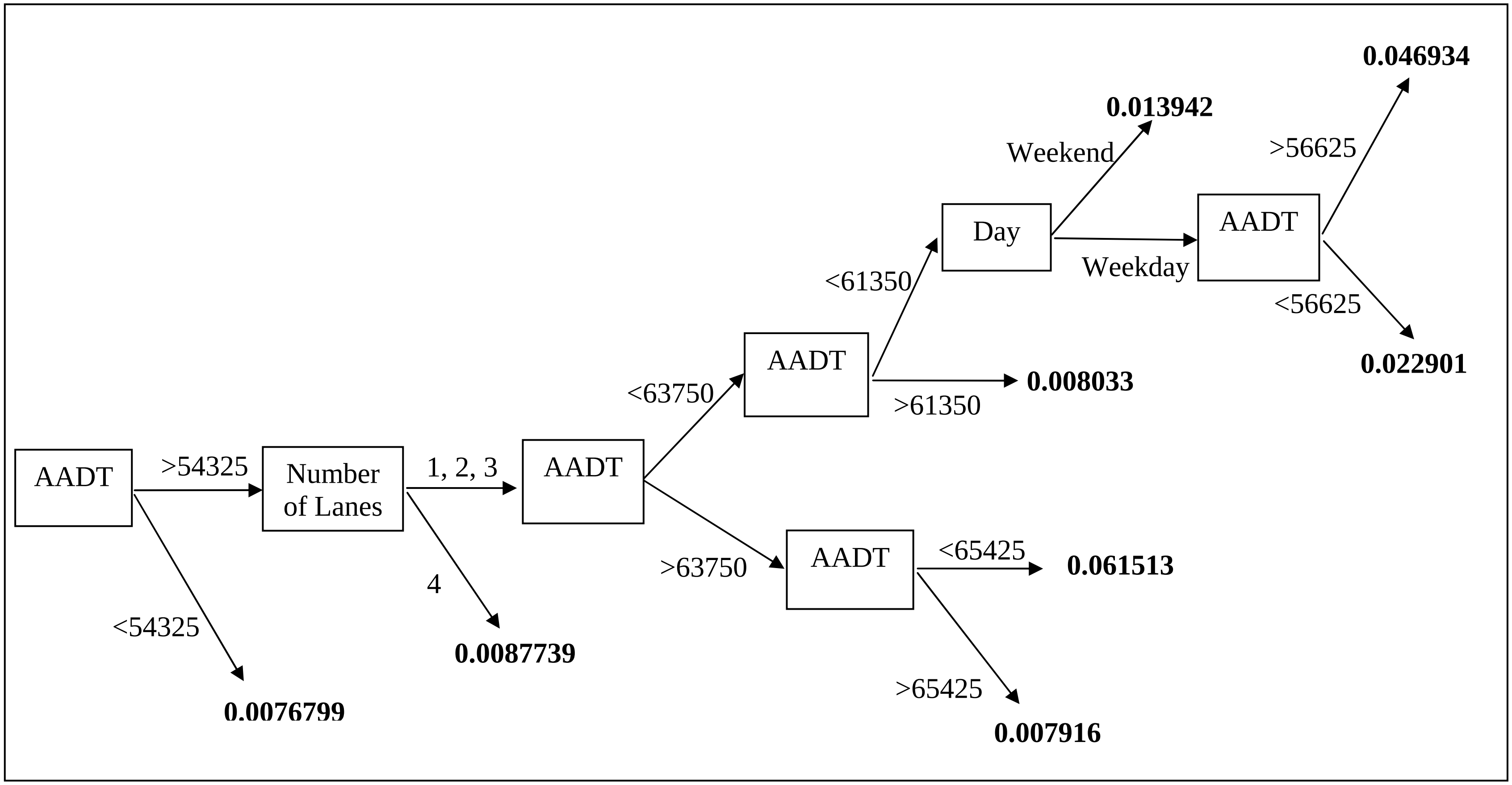


Figure A-4 Optimal Tree for TRANSCOM Disabled Vehicle Incident Frequency [Incidents/Mile/Day]

For overall data, as well as incident specific analysis, there are 4 factors in the optimal tree that affect the incident rates, namely, AADT, Weekday/Weekend, shoulder existence and Number of Lanes. Among all, AADT is found to be the most significant parameter for all incident types. First node, which is the most influential node is always AADT in the trees. It should be noted again that "number of lanes" have entries 4 lanes whereas officially there are at most 3 lane sections. 4-lane section records are extracted by visual inspection of aerial photos to enhance the location information and these sections are merging or diverging areas along the roadway. The trees show that this kind of data extraction can add to the explanatory power of the overall dataset since "number of lanes" are one of few factors left in the final optimum decision tree.
As for the duration analysis, cluster information is added to the incident data. The optimal trees are found with the same method where the tree with the minimum cost is chosen as the best tree. However, no significant improvement could be achieved in terms of predictive performance. Table A-4 shows the cross validation results of the tree analysis with and without spatial information to compare the prediction performance with the other regression models.

Table A-4 Cross Validation Results of the Tree Analysis for TRANSCOM Incident Frequency

	Non-Spatial	Spatial
Overall Tree	0.018529	0.023355
Property Damage	0.013487	0.02565
Disabled Vehicle	0.02114	0.037431
Disabled Truck	0.008443	0.0083632

First, it can been from Table A-4 that non-spatial incident type specific analysis yields better prediction results for "property damage" and "disabled truck", however the prediction error increases for "disabled vehicle". Nevertheless, the increase in the RMSE with incident type specific analysis is compensated with the relative improvement of prediction performance, thus incident specific analysis proves to be a useful approach just like in the duration analysis. Overall, the non-spatial analysis gives the best prediction results compared to spatial tree analysis as well as the negative binomial analysis.

CONCLUSIONS & DISCUSSION FOR INCIDENT FREQUENCY ESTIMATION

Conclusions & Discussion of TRANSCOM Incident Duration Prediction

As discussed in detail in Task 3.2 in detail, TRANSCOM dataset, which is the basis of the results presented below is found to cover only a small portion (around 8%) of the overall accident records for the study area. This conclusion is drawn after a comparison of NYSDOT accident dataset that cover the same corridor with the available TRANSCOM dataset. Nevertheless, interpretations of the duration analysis can still be valuable due to the fact that a relatively small sample of the total accident data was available to the research team the results of this analysis can be extended to represent the overall reality with sample size limitations. The results of this study are thus strictly restricted to available TRANSCOM records and should not be used as final recommendations for the study area.

The results of the incident frequency estimation can be summarized as follows:

- Linear regression performs very poorly. This is an expected result based on the literature and it is also an intuitive result since the frequency data does not have the characteristics that can be captured by a linear regression. However it was still tested to show its performance when applied to the current data set.
- Negative binomial regression succeeds to capture some important parameters that are frequently addressed in the literature. However, lack of physical road characteristics data does not allow us to conduct a meaningful comparison between our results and the ones given in the past studies.
- CART, a non-parametric modeling approach, is more successful in determining the important parameters. Moreover, the best cross validation results are obtained when using CART analysis.

Disregarding the inadequate number of records in TRANSCOM, similar to duration analysis, dataset is found to be incomplete for this study and does not allow for robust and reliable estimation of both parametric and non-parametric models for incident frequency analysis. Including spatial information as an explanatory variable does not make big difference both for duration and frequency estimations. The CART model again outperforms the other models in frequency analysis.

Task 4: LOOK-UP TABLES

Prepared by:

Kaan Ozbay, Ph.D.
Associate Professor
Rutgers University, Civil and Environmental Engineering
632 Bowser Road, Piscataway, NJ 08854-8014
Tel: 732.445.2792
Fax: 732.445.0577
Email: kaan@rci.rutgers.edu

M. Anil Yazici
Graduate Assistant
Rutgers University, Civil and Environmental Engineering
632 Bowser Road, Piscataway, NJ 08854-8014
Tel: 732.445.0579 (x119)
Fax: 732.445.0577
Email: yazici@eden.rutgers.edu

February22, 2008

1.0 Introduction

As previously mentioned in Task 3.2 (under section *Some Important Remarks about the Available TRANSCOM Dataset*), after comparison with "NYSDOT Safety Information Management System, Accident Verbal Description Report For I-287 Intersection & Non-Intersection Accidents" records, it was found that TRANSCOM dataset does not cover all the accidents in the study area. This caused no problem for incident duration calculations since duration analysis can be performed over a sample, and the results can be applied over all accidents with certain confidence level imposed by the sample size. With the same logic, the accident frequency model in which the factors affecting incident frequency are determined, can still accepted as valid since the analysis is based on a fairly large sample size (1907 records). However, the accident rate values cannot be used because estimation of actual accident frequency requires having a complete set of accident records. This completeness is satisfied by NYSDOT dataset only for accidents in terms of number of accidents. NYSDOT database do not include some major frequency modeling parameters and more importantly, do not cover any disablements. Although TRANSCOM and NYSDOT records do not show a compatible pattern in terms of accident type content, some assumptions can still be employed to use the information from both datasets to produce reasonable accident rate values for practical use. Some important problems about the compatibility and assumptions to overcome these problems are listed below:

1. **Accident Rate Model Validity:** NYSDOT records made available to the research team do not have parameters necessary to build an accident prediction model similar to the one proposed in Task 3.3. NYSDOT records also lack disablement incidents, so any model derived from NYSDOT records cannot be applied for look-up tables requiring rates for disablements. Hence, the accident rate model based on TRANSCOM data is still assumed to be valid, since it was derived from a sample of accidents that satisfy above requirements.
2. **Accident/Non-accident (Disablement) Ratio:** From TRANSCOM incident data set covering a 14-month study period (February 2004 – March 2005), a ratio of accidents versus non-accidents for the I-278 corridor is calculated. It is determined that for every accident there are (902 + 119) / 886 = 1.15 non-accidents. Thus conversion factor for NYSDOT accidents data for the I-278 corridor into total corridor incidents is assumed to be 2.15. CNAM uses the following incidents-to-accidents ratios: 2.3 for peak hour traffic, and 1.7 for

off peak traffic. On the other hand, TRANSCOM accident/non-accident ratios are found to be 2.21 and 2.11 for peak and off-peak periods, respectively. Although there is a slight difference for off-peak and peak periods, is the results are fairly close, and overall the accident/non-accident ratios can be assumed to be comparable. This assumption allows us to expand accident-only NYSDOT records to represent the entire incident dataset.

3. **Total Number of Incidents:** The main assumption is based on the determination of a base year to extract the total number of accidents to inflate the TRANSCOM accident rate estimations to represent actual incident rates. A review of three years (2000-2002) of NYSDOT accident reports, indicated a decrease in the number of accidents from 2000 to 2002, but there is a sharp decrease in the total number of accidents from June to December of 2002. Since there is no explanation for the sharp reduction in the number of accidents shown for the second half of 2002, the team collectively decided to use 2001 data to calculate the expansion factor based on the conference call January 13th 2008.

Basically, the new incident frequency rates are based on NYSDOT records, instead of incomplete TRANSCOM dataset, where the incident rate model based on TRANSCOM dataset is kept as before.

Based on the assumptions summarized above, the frequency tables given in Task 4.1 can be updated following the procedure below:

1. Total number of accidents in TRANSCOM database is substituted with NYSDOT records of year 2001,chosen as the representative year among 2000-2002. The total number of non-accidents (or disablements) are found by multiplying the total accident count by non-accident/accident ratio
2. The individual counts of accident types are found based on NYSDOT base year percentages (71% property damage and 29% injury) and counts of non-accidents are found based on TRANSCOM overall disablement/accidents ratio. That is found to be (902 + 119) / 886 = 1.15, or in other words, 46% accidents and 54% disablements. The percentages of the disablement types among all disablements are calculated from TRANSCOM, which is ~68% disabled vehicle, ~18% disabled truck, ~8% road hazard, ~3% vehicle fire, ~3% HAZMAT and ~1% weather related.

3. The rates given in the Task 4 Look-Up Tables (see Draft Final Report dated January 2008) are updated based on the total counts calculated in step3, keeping the rate estimation model structure the same as in Task 3.3 , e.g. based on AADT, number of lanes, day, and shoulder existence.

Updated number of accidents and disablements are given in Table 5:

Table 5 Updated Incident Counts

	Accidents		Disablements					
	Property Damage	Injury	Disabled Vehicle	Disabled Truck	Road Hazard	Vehicle Fire	HAZMAT	Weather Related
	4258	1740	4786	1234	552	200	200	69
TOTAL	5998		7041					

Table 6 Adjusted Look-up Table for “Property Damage” Incident Type for TRANSCOM Dataset

AADT	**# of Lanes**	**Day**	**Shoulder**	**Duration [min]**	**Frequency [Inc/mile/day]**
AADT<49975	→	→	Yes	49	0.065648
			No	35.8	
49975<AADT<56225	→	Weekend	Yes	49	0.10895
			No	35.8	
		Weekday	Yes	49	0.089746
			No	35.8	0.44293
56225<AADT<61350	→	Weekend	Yes	49	0.10895
			No	35.8	
		Weekday	Yes	49	0.089746
			No	35.8	0.25643
61350<AADT<63750	→	→	Yes	49	0.044148
			No	35.8	
63750<AADT<65425	1,2,3	→	Yes	49	0.28947
			No	35.8	
	4	→	Yes	49	0.076278
			No	35.8	
65425 < AADT	→	→	Yes	49	0.059488
			No	35.8	

Table 7 Adjusted Look-up Table for "Disabled Vehicle" Incident Type for TRANSCOM Dataset

AADT	# of Lanes	Day	Shoulder	Duration [min]	Frequency [Inc/mile/day]
AADT<54325	→	→	Yes	41.88	0.061524
			No	23.82	
54325<AADT<56625	1,2,3	Weekend	Yes	41.88	0.11167
			No	23.82	
		Weekday	Yes	41.88	0.18345
			No	23.82	
	4	→	Yes	41.88	0.070256
			No	23.82	
56625<AADT<61350	1,2,3	Weekend	Yes	41.88	0.11167
			No	23.82	
		Weekday	Yes	41.88	0.37596
			No	23.82	
	4	→	Yes	41.88	0.070256
			No	23.82	
61350<AADT<63750	1,2,3	→	Yes	41.88	0.064328
			No	23.82	
	4	→	Yes	41.88	0.070256
			No	23.82	
63750<AADT<65425	1,2,3	→	Yes	41.88	0.49276
			No	23.82	
	4	→	Yes	41.88	0.070256
			No	23.82	
65425<AADT	1,2,3	→	Yes	41.88	0.063447
			No	23.82	
	4	→	Yes	41.88	0.070256
			No	23.82	

Table 8 Adjusted Look-up Table for "Disabled Truck" Incident Type for TRANSCOM Dataset

AADT	# of Lanes	Day	Shoulder	Duration [min]	Frequency [Inc/mile/day]
→	→	→	Yes	64.45	0.070312
			No	38.21	

Table9 Adjusted Look-up Table for "Personal Injury" Incident Type for TRANSCOM Dataset

AADT	# of Lanes	Day	Shoulder	Duration [min]	Frequency [Inc/mile/day]
→	→	Weekend	→	58.28	0.26699
		Weekday	→	43.02	

Table 10 Adjusted Look-up Table for "Road Hazard" Incident Type for TRANSCOM Dataset

AADT	# of Lanes	Day	Shoulder	Duration [min]	Frequency [Inc/mile/day]
→	→	→	Yes	257.50	0.054344
			No	154.70	

Table 11 Adjusted Look-up Table for "HAZMAT" Incident Type for TRANSCOM Dataset

AADT	# of Lanes	Day	Shoulder	Duration [min]	Frequency [Inc/mile/day]
→	→	→	→	63.78	0.011937

Table 12 Adjusted Look-up Table for "Vehicle Fire" Incident Type for TRANSCOM Dataset

AADT	# of Lanes	Day	Shoulder	Duration [min]	Frequency [Inc/mile/day]
→	→	→	→	56.30	0.011937

Table 13 Adjusted Look-up Table for "Weather Related" Incident Type for TRANSCOM Dataset

AADT	# of Lanes	Day	Shoulder	Duration [min]	Frequency [Inc/mile/day]
→	→	→	→	213.44	0.0040878

Table 14 A Simplified Adjusted Frequency Look-up Table for Overall Incidents for RANSCOM Dataset

AADT	**Frequency [Incidents / Mile / Day]**
AADT<49975	0.45551
49975<AADT<54325	1.2127
54325<AADT<56625	1.5139
56625<AADT<61350	1.4599
61350<AADT<63750	0.49967
63750<AADT<65425	1.3477
65425<AADT	0.51945

Task 5: Strategy Assessment (Validation of Methodologies)

Prepared by:

George List, Ph.D., P.E.
Professor and Head
NC State University
Civil, Construction and Environmental Engineering
Campus Box 7908, Raleigh, NC, 27695-7908
Tel: 919.515.7212
Fax: 919.515.7908
Email: gflist@ncsu.edu

October 15, 2008

1.0 Introduction

The Task 4 report presents a new set of freeway-based look-up tables for the Congested Network Analysis Model (CNAM). The intent is to make CNAM better for studying non-recurring delay in the New York City network. This report describes how the new tables will change the structure of CNAM and alter its predictions.

2.0 Overview of the Congested Network Analysis Model

(Abstracted from the CNAM Users Manual)

CNAM (Congestion Needs Analysis Model) is a computer model intended for use by NYSDOT, MPO and county officials in conjunction with CMS (Congestion Management System) or other congestion-related activities. CMS is the process of managing congestion mandated by 23 CFR 500. CNAM is specifically designed to estimate present and future congestion/mobility conditions. It estimates congestion as a result of implementing congestion-relief projects (strategies). Hence, it satisfies some of the steps of the CFR process.

CNAM focuses principally on estimating the amount of delay occurring in the state network. This includes vehicle-hours of delay as well as person-hours and ton-hours of delay. In December, 1989, NYSDOT adopted a Corporate Mobility Management Goal which identifies the objective of reducing the predicted number of vehicle hours of delay (VHD) at Level of Service *E* or *F*. (Travel at LOS *D* is acceptable.)

CNAM provides output for each hour of the day and for each highway link, including:

- Vehicle Hours of Delay (VHD), Passenger (PHD) , and Truck (THD)
- Costs attributed to VHD, PHD, THD
- Average modeled speed
- GIS-compatible files of all data.

CNAM determines whether travel at each hour is at Level of Service *E* or *F*, based on comparing actual volumes to the service volumes for LOS *E* and *F*. In the first step of calculating delay, for those segments that are at LOS *E* or *F*, travel time is equal to the segment length divided by the average operating speed. VHD is then calculated as the hourly volume multiplied by the difference between travel time at LOS *E* or *F* and travel time at LOS *D*. The formulas used are:

Travel Time = Section Length/Average Operating Speed
VHD/Hr. = Vol./Hr. x (*Travel Time @ LOS E* or *F - Travel Time @ LOS D*)
Daily VHD = Sum of Hourly Delays, at 260 working days/year.

CNAM is designed to act as a central storage location for all congestion-related data (including travel speeds, hourly volumes, average occupancy, accident rates, accident durations, etc). It employs 30 input tables that represent data on the state highway system. It uses both the State Touring Route System (STR) and the Local Highway Inventory (LHI) to develop its description of the state highway network.

CNAM performs its functions using several sub-models:

- *Recurring Model*: for both freeways and arterials. Uses the Bureau of Public Roads formula to calculate recurring delay based on volume. The formula is bypassed if the model is provided real travel speeds.
- *Incident Model*: (the focus of this task report) for freeways only. Uses a queue dissipation formula segment-by-segment to estimate non-recurring delay.
- *Strategy Model*: uses user-input estimates of decreases in traffic volume (or accident duration for incident strategies) to calculate delay savings of each proposed project.
- *Intersection Model*: is complete but not in use due to the large amount of input data needed (and not available).
- *Traffic Count Viewer*: allows users to view all traffic counts and travel speeds.

A separate Manual for the CNAM Incident Delay Model is one of three CNAM documents. It is included with this report as Attachment C and is summarized in the text that follows.

CNAM uses a deterministic queuing model to estimate total, network-level incident-related delay. One highway section at a time, the appropriate accident rate is converted into a number of accidents per hour (using the hourly counts). These are then distributed among accident types based on the likelihood that a given type of accident[4] will occur. Each type of accident has a specific duration[5] and diminished capacity[6]. From the midpoint of each hour, and moving forward, changing the traffic demand as time progresses, the cumulative diminished capacity (supply) is compared with the cumulative demand at 15-minute intervals, searching for a time when the queue returns to zero. Total vehicle delay is then the area between the cumulative demand and cumulative supply (capacity) curves.[7]

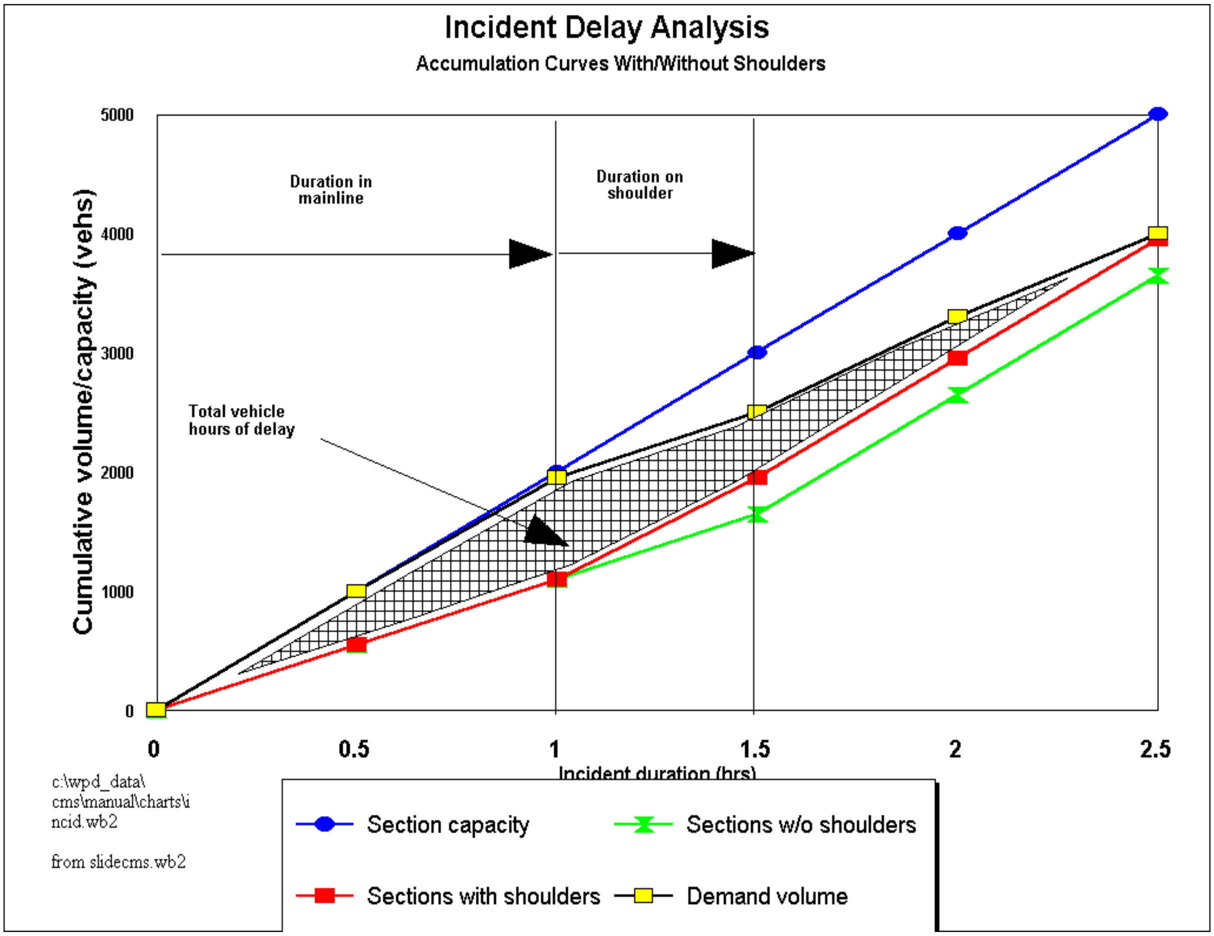
This idea is illustrated in Figure 1. The black line with yellow boxes is the demand. The blue line is the normal section capacity while the red and green lines are the capacity that remains after the incident starts at t = 0 for sections with a shoulder (red) and without (green)[8]. For the red case of a highway with a shoulder, the incident causes a decrease in capacity for one hour and after 2.5 hours, the queue is gone. For the green case where there is no shoulder, the diminished capacity lasts 1.5 hours and the queue is not yet gone at the end of 2.5 hours. Another ½ hour of queue clearance time is required.

---

[4] An accident type blocks X number of lanes, from 0 (i.e., the shoulder) to two-or-more.
[5] Duration is measured in minutes.
[6] This is defined as the percentage of the unimpeded capacity that remains.
[7] The annual incident delay for the hour is computed by multiplying the total incident delay by 260, the assumed number of normal use days in a year.
[8] A useable shoulder is deemed to be at least 6-feet wide.

Figure 1: Computation of Vehicle Hours of Delay for a Single Highway Segment

Source: CNAM Incident Delay Model User's Guide



3.0 Look-Up Table Changes – Purpose of the Project

The new look-up tables are significantly different from those presently used in CNAM. The new tables use a new definition of an incident as well as a new methodology to compute incident-related delays.

CNAM's Present Structure

CNAM's present structure involves an accident table, an incident table, and an incident duration table. The accident table shows the accidents per million vehicle miles of travel broken down by access control, divided or undivided highway[9], area type[10], and number of lanes, as shown in Table 1. To provide an illustration, for a 4-lane undivided highway with partial or full access control in area type 5 the accident rate is 2 accidents (of any type) per million vehicle miles of travel.

Table 1: Accidents per Million Vehicle Miles Traveled

# Of Lanes*	No Access Control Area Type 1	2	3	4	5	6	Partial or Full Access Control Area Type 1	2	3	4	5	6
	Undivided Roadway (median < 4-feet)											
1	3.62	3.62	3.62	3.62	3.62	3.62	3.62	3.62	3.62	3.62	3.62	3.62
2	2.73	4.78	4.78	4.73	3.07	4.78	1.18	1.57	1.57	1.39	2.00	1.57
3	2.93	6.20	6.20	4.87	4.87	6.20	1.18	1.57	1.57	1.39	2.00	1.57
4	3.56	6.58	6.58	6.57	5.48	6.58	1.18	1.57	1.57	1.39	2.00	1.57
5	3.56	6.58	6.58	6.57	6.57	6.58	1.24	1.27	1.27	1.39	2.00	1.27
6	5.57	5.57	5.57	5.31	5.31	5.57	1.01	1.87	1.87	1.15	1.59	1.87
7	3.33	5.57	5.57	5.31	5.31	5.57	1.01	1.87	1.87	1.15	1.59	1.87
8	5.57	5.57	5.57	5.31	5.31	5.57	1.01	1.87	1.87	1.15	1.59	1.87
9+	5.57	5.57	5.57	5.31	5.31	5.57	1.01	1.87	1.87	1.15	1.59	1.87
	Divided Roadway (median >= 4-feet)											
1	3.62	3.62	3.62	3.62	3.62	3.62	3.62	3.62	3.62	3.62	3.62	3.62
2	2.73	4.78	4.78	4.73	4.73	4.78	1.18	1.57	1.57	1.39	2.00	1.57
3	2.93	6.20	6.20	4.87	4.87	6.20	1.18	1.57	1.57	1.39	2.00	1.57
4	3.20	6.27	6.27	5.40	4.26	6.27	1.18	1.57	1.57	1.39	2.00	1.57
5	3.20	6.27	6.27	5.40	5.40	6.27	1.24	1.27	1.27	1.39	2.00	1.27
6	5.57	5.57	5.57	5.31	4.76	5.57	1.01	1.87	1.87	1.15	1.59	1.87
7	5.57	5.57	5.57	5.31	5.31	5.57	1.01	1.87	1.87	1.15	1.59	1.87
8	5.57	5.57	5.57	5.31	5.31	5.57	1.01	1.87	1.87	1.15	1.59	1.87
9+	5.57	5.57	5.57	5.31	5.31	5.57	1.01	1.87	1.87	1.15	1.59	1.87

* = Total of both directions

These *accident rates* are distributed among *incident types* based on the number of lanes blocked.[11] The probabilities for each type of incident are shown in Table 2. To illustrate the implications, on a 4-lane freeway that has shoulders, there is a 48% chance that the incident will only block the shoulder, a 44% chance that it will block one lane, and an 8% chance that it will block 2. If the highway has no shoulders, then there is an 85% chance that one lane will be blocked and a 15% chance that both lanes will be blocked.

[9] Highways with a 4-foot median or more are considered divided.
[10] The area types are 1 (rural), 2 or 3 (pop < 5000), 4 (suburban), 5 or 6 (city or large village)
[11] There does not appear to be a distinction between accident rates and incident rates.

Table 2: Probabilities of Incidents by Type *								
	Lanes Blocked / With Shoulder				Lanes Blocked / No Shoulder			
#lanes	0	1	2	>2	0	1	2	>2
4	48.0	44.0	8.0	-	-	85.0	15.0	-
6	47.0	43.0	8.0	2.0	-	81.0	15.0	4.0
8	47.0	43.0	8.0	2.0	-	81.0	15.0	4.0
* = Number of lanes blocked								

Each incident type has a specific diminished capacity as shown in Table 3. To illustrate, if an incident blocks one lane of a 6-lane freeway with shoulders, the diminished capacity is 56.7% of the original capacity.

Table 3: Percent of Capacity Remaining after an Incident								
	With Shoulder / Lanes Blocked*				Without Shoulder / Lanes Blocked*			
#lanes	0	1	2	>2	0	1	2	>2
	Default Values							
4	85.0	42.5	-	-	-	42.5	-	-
6	87.5	56.7	23.3	-	-	56.7	23.3	-
8	87.8	63.8	42.5	21.3	-	63.8	42.5	21.3
	User-Supplied							
4	83.2	42.5	-	-	-	42.5	-	-
6	87.5	56.7	23.3	-	-	56.7	23.3	-
8	87.8	63.8	42.5	21.3	-	63.8	42.5	21.3
* Incident type defined by the number of lanes blocked, 0, 1, 2, >2								

The incident also has a specific duration as shown in Table 4. For example, for an incident in area type 5 [12] where two lanes are blocked, the clearance time is 45 minutes plus 15 minutes to clear the shoulder. If the incident occurs on the shoulder, it takes 43 minutes to clear and, since, it was on the shoulder, there is no incremental shoulder clearance time. The decreased capacity pertains during the incident clearance time but not during the shoulder clearance time.

[12] Major urban areas are of type 5. This includes New York City.

Table 4: Times (minutes) for Shoulder and Accident Clearance								
# Lanes	Shoulder Clearance by Area Type*				Incident Clearance by Area Type**			
Blocked	1	3	4	5	1	3	4	5
0	0	0	0	0	61	37	51	43
1	15	15	15	15	62	38	52	44
2	15	15	15	15	63	39	53	45
>2	20	20	20	20	67	43	59	51

* detection, verification, response, and clearance times
** shoulder clearance time only

To summarize in a stepwise fashion:

- For each specific highway segment, one obtains
  - Area type
  - Number of lanes and shoulder (y/n)
  - Access Control
  - Divided or undivided
- This permits identification of an accident rate (Table 1)
- The number of lanes (4, 6, or 8) determines a suite of incident likelihoods (Table 2)
- The likelihoods and attendant capacity reductions (Table 3) and durations (Table 4), permit development of a composite supply (capacity) curve
- Total demand, from the mid-point of the hour, is compared with total composite supply, every 15 minutes, until the queue reaches 0.
- The total vehicle-hours of delay is the difference between total vehicle-hours for the demand curve and total vehicle-hours for the composite supply curve (Figure 1).

Proposed New Structure

The proposed new structure is very different. As described in the Task 4 report, it defines incidents more traditionally, focusing on categories like “property damage”, “disabled vehicle”, disabled truck”, etc. as shown in Table 5.[13]

Each type of incident has a frequency of occurrence, measured in incidents per mile per day (one way), which in six cases is a fixed value, but for “Property Damage” and “Disabled Vehicle”, is a function of the AADT.[14] To illustrate one of the fixed ones, for “Disabled Trucks”, Table 5 shows an incident rate of .07374 incidents per mile per day (one-way), which means for a facility like I-278, which is 33.6 miles long according to the CNAM sufficiency file, there should be 2.48 incidents per day each way, or 4.96 for both directions. Overall, since the average number of incidents per mile per day (each way) is 0.77913, this means I-278 should see 26.2 incidents each day per direction, or 52.4 incidents per day total.

[13] The Task 4 report contains slightly more detail than that shown here, as in accident rates for weekend days as well as weekdays. Weekday values are presented here.
[14] See Tables 6 and 7. The values shown in Table 5 are the averages for all facilities derived from an analysis based on NYSDOT accident and TRANSCOM accident and incident data.

Table 5: Incident Frequencies and Durations by Type*			
		Duration (min)	
Incident Type	Freq*	No Shoulder	Shoulder
Property Damage**	0.25443	35.80	49.00
Disabled Vehicle**	0.28598	23.82	41.88
Disabled Truck	0.07374	38.21	64.45
Personal Injury	0.10397	58.28	58.28
Road Hazard	0.03298	154.70	257.50
HAZMAT	0.01195	63.78	63.78
Vehicle Fire	0.01195	56.30	56.30
Weather Related	0.00412	213.44	213.44
Overall	0.77913	-	-
* = one-way incidents/mile/day			
** = The actual rates vary. See Tables 6 and 7.			

Compared with Table 1, which presents *accident rates* in accidents per million vehicle miles of travel (MVMT), if one assumes the rates in Table 5 are based on freeways that have AADTs in the range of 60,000 per direction (120,000 two-way), the overall incident rate of 0.77913 incidents/mile/day implies 12.98 *incidents* per MVMT. This compares to the range of 1.15 to 6.2 *accidents* per MVMT in Table 1.[15]

Each incident also has a specific duration that depends upon whether shoulders are present or not.[16] For example, a disabled vehicle incident lasts 41.88 minutes if a shoulder exists and 23.82 minutes if it does not. Incidents always take longer if shoulders are present.

Table 6: Incident Frequencies - Property Damage *		
AADT Range (one-way)	No Shoulder	Shoulder
AADT < 49,975	0.06565	0.06565
49,975 < AADT < 56,225	0.44293	0.08975
56,225 < AADT <61,350	0.25643	0.08975
61,350 < AADT < 63,750	0.04415	0.04415
63,750 < AADT < 65,425	0.07628	0.07628
65,425 < AADT	0.05949	0.05949
Simple Average	0.15749	0.07084
* = one-way incidents/mile/day		

The incident frequencies for property damage accidents depend on the AADT for the facility as shown in Table 6. To illustrate, for a freeway with a one-way AADT between 49,975 and 56,225, the incident rate is 0.44293 per mile per day if the freeway has no shoulders and 0.08975 if it does[17]. The simple averages of 0.15749 and 0.07084 are intended to be helpful, but they should not be over-interpreted. They simply give an average among the values shown, not a true weighted average

[15] This *incident* rate may be far too low. Some studies, discussed later, indicate that the *incident rate* should be four to five times the *accident rate*. There are many issues associated with the completeness of the TRANSCOM data. So, while these new results may be useful and informative, they are built on a database of questionable credibility.
[16] Interestingly, the incident durations are longer with shoulders than without.
[17] This is a huge difference. If the data are correct, NYSDOT should strive to add shoulders wherever possible.

that pertains based on distributions of AADT among highway segments. The value overall for property damage incidents is 0.25443 as shown in Table 5.

For disabled vehicle events, the incident rate depends on both the one-way AADT and the number of freeway lanes, as shown in Table 7. For the mid-range one-way AADT values there is a substantial difference between the incident rates for freeways with 1-3 lanes (in one direction) and those that have 4. Consistent with the comments about the property damage values, the simple averages are intended to be helpful, but they should not be over-interpreted. Those rates give an average among the values shown, not a true weighted average based on distributions of AADT among highway segments. That average was presented in Table 5.

Table 7: Incident Frequencies - Disabled Vehicles *		
AADT Range (one-way)	1-3 Lanes	4 Lanes
AADT < 54,325	0.06152	0.06152
54,325 < AADT < 56,625	0.18345	0.07026
56,625 < AADT <61,350	0.37596	0.07026
61,350 < AADT < 63,750	0.06433	0.07026
63,750 < AADT < 65,425	0.49276	0.07026
65,425 < AADT	0.06345	0.07026
Simple Average	0.20691	0.06880

* = one-way incidents/mile/day

As was the case with CNAM, there are probabilities that a number of lanes will be blocked, as shown in Table 8. “None” implies no lanes are blocked while “N/A” means the incident occurs without impact to the lanes or shoulder.

The final element of the new model relates to capacity reductions, for which the current CNAM values are employed since no information about this was available in the field data. This means that the values found in Table 3 are employed.

Table 8: Probabilities of Lane Blockage					
Incident Type	None	1 Lane	2 Lanes	3+	N/A
Property Damage	5.1%	72.5%	17.0%	3.5%	1.8%
Disabled Vehicle	4.8%	90.1%	3.5%	1.0%	0.7%
Disabled Truck	6.7%	84.4%	4.5%	3.9%	0.6%
Personal Injuries	16.4%	44.8%	34.3%	1.5%	3.0%
Road HAZARD	3.8%	63.8%	13.8%	1.3%	17.5%
HAZMAT	6.9%	79.3%	13.8%	0.0%	0.0%
Vehicle Fire	6.9%	31.0%	37.9%	17.2%	6.9%
Weather Related	20.0%	10.0%	0.0%	0.0%	70.0%

4.0 Calculation Comparisons

A true comparison of CNAM's predictions versus those of the new model would require changes to the CNAM program code; but that is beyond the scope of this project. Consequently, the study team used a workspace, developed in Excel and Excel/VBA, to mimic the current CNAM model and implement the new model.

The first calculation comparisons used worksheets with formulas to compare the delay predictions of CNAM with those of the new model for four highway segments in the CNAM Region 11 database. Table 9 presents the input data for the four segments.

Table 9: Segment Analyses				
Parameter	Highway Segment			
Road	I-278	I-278	I-278	I-278
Segment	9756	9782	9797	9790
#Lanes	4	6	6	8
Length (mi)	0.88	0.61	1.37	0.42
Capacity	3720	5100	5580	6990
AADT	76500	112300	93790	111300
Analysis Volume	1912	2807	2344	2782
Shoulder	Y	N	Y	N
Divided	N	Y	N	Y

Table 10: CNAM Analysis				
Parameter	Highway Segment			
Road	I-278	I-278	I-278	I-278
Segment	9756	9782	9797	9790
#Lanes	4	6	6	8
Length (mi)	0.88	0.61	1.37	0.42
Capacity	3720	5100	5580	6990
AADT	76500	112300	93790	111300
Volume	1912	2807	2344	2782
Shoulder	Y	N	Y	N
Divided	N	Y	N	Y
Accident Rate	2.00	1.59	1.59	1.59
0 Lanes Blocked				
Probability	48%	0%	47%	0%
Pct Capacity	85.0%	87.5	87.5%	87.8
Clearance Time	43	43	43	43
Done Time	0.0	0.0	0.0	0.0
Total Delay	0	0	0	0
Delay/Vehicle	0.0	0.0	0.0	0.0
1 Lanes Blocked				
Probability	44%	81%	43%	81%
Pct Capacity	42.5%	56.7%	56.7%	63.8%
Clearance Time	44	44	44	44
Done Time	52.1	0.0	0.0	0.0
Total Delay	6318	0	0	0
Delay/Vehicle	3.8	0.0	0.0	0.0
2 Lanes Blocked				
Probability	8%	15%	8%	15%
Pct Capacity	0.0%	23.3%	23.3%	42.5%
Clearance Time	45	45	45	45
Done Time	92.6	76.8	59.5	0.0
Total Delay	66386	46598	23297	0
Delay/Vehicle	22.5	13.0	10.0	0.0
>2 Lanes Blocked				
Probability	0%	4%	2%	4%
Pct Capacity	-	0%	0%	21.3%
Clearance Time	-	51	51	51
Done Time	-	113.4	87.9	66.7
Total Delay	-	135322	87608	36642
Delay/Vehicle	-	25.5	25.5	11.9
Prob Weighted Delay Before Incident Rate				
Total Delay	8091	12403	3616	1466
Delay/Vehicle	3.5	3.0	1.3	0.5
Prob Weighted Delay After Incident Rate				
OneWay Min/Hr	62.6	77.7	42.5	6.3
TwoWay Min/Hr	125.2	155.3	84.9	12.5

The analysis assumes that an incident occurs at $t = 0$ and lasts as long as each individual model suggests is appropriate. The demand is assumed to be 5% of the directional AADT, which in CNAM terms is the average 4-5 hour volume, compensating for demand fluctuations; it continues until the incident ends. Each incident scenario is analyzed, as defined by the two models, and then applicable probabilities and incident rates are used to determine how many vehicle-minutes of delay per hour could be expected for each segment.

CNAM Analysis

Table 10 presents the predictions that the study team obtained from the Excel/VBA representation of CNAM. In the case of Segment 9756, because it is a 4-lane, undivided facility with shoulders, the accident rate is 2.0 accidents per million-vehicle-miles. Then, taking into account the 2.3 multiplier that CNAM uses to convert the accident rate into an incident rate, the incident rate is 4.6 incidents per MVMT.

In CNAM there are four “incident types”: none, one, two, and more than two lanes blocked. Since this is a 4-lane facility with shoulders, only the first three pertain. For “no lanes blocked”, the reduced capacity is 3,162 vph (85% of normal); this is a value large enough, given the demand of 1912 vph (5% of the directional AADT of 38,250), to keep a queue from developing; so no incident-related delay occurs. For “one lane blocked”, the reduced capacity is 1581 vph, which is smaller than the demand, so a queue develops. The queue grows for 44 minutes and then begins to shrink. At 52.1 minutes, the queue is gone and delay accumulation stops. The total delay is 6,318 vehicle-minutes, which is 3.8 minutes per queued vehicle. Similarly, for “two lanes blocked”, the clearance time is 45 minutes; the queue disappears at 92.6 minutes; and 66,386 vehicle-minutes of delay accumulate; an average of 22. 5 minutes per queued vehicle. Combining the results for these latter two incident types given their likelihoods (44% and 8% respectively) yields a probability weighted delay of 8,091 vehicle-minutes. Then, taking the incident rate into account (4.6 incidents/MVM), and the fact that the segment is 0.88 miles long, the probability weighted delay is 62.6 minutes/hour one-way; or 125.2 minutes of delay/hr two-way.

To back up a bit, the total delay values for each incident type are computed as follows. As Figure 2 shows, delay accumulates because the demand arrival rate exceeds the reduced capacity, c_1, and stops accumulating when the renewed normal capacity c catches up with demand d. The time t_2 at which the queue returns to zero is given by:

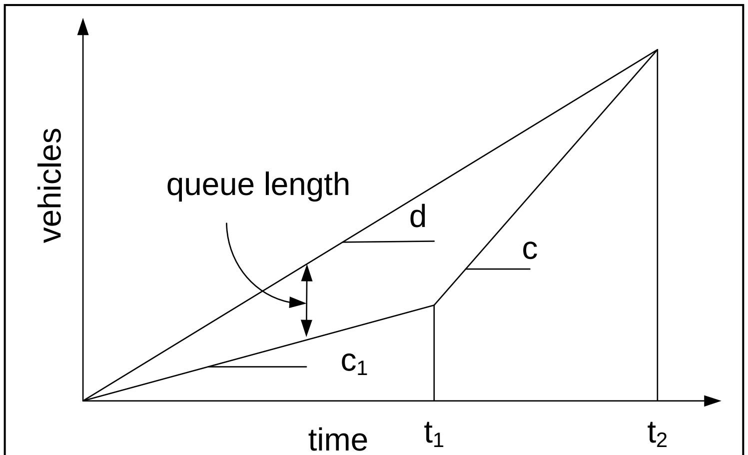


Figure 2: Delay Accumulation

EMBED Equation.DSMT4

which solves to:

EMBED Equation.DSMT4

Then, the total delay is given by:

EMBED Equation.DSMT4

The delay calculation can be simplified by taking the difference between the demand curve and the capacity curve. This results in the diagram shown in Figure 3. The queue length continues to increase until t_1 and then decreases.

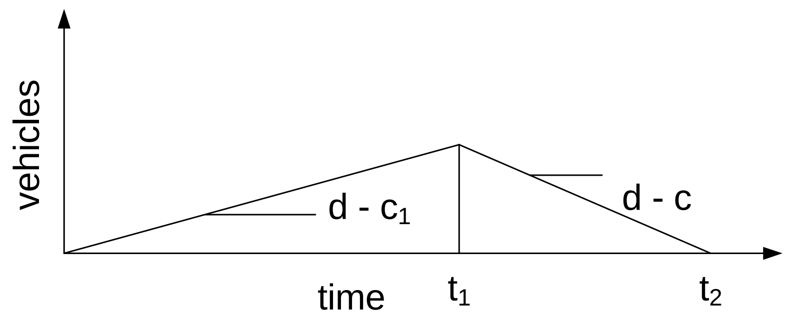


Figure 3: Delay Accumulation - 2

In this case, the delay is given by the sum of the two triangular areas on either side of t_1:

EMBED Equation.DSMT4

For segment 9756 in the incident condition where two lanes are blocked, d = 31.9 veh/min (1912 veh/hr), c_1= 0, c= 62 veh/min (3720 veh/hr), and t_1 = 45 min, which results in t_2 = 92.6 min, and D = 66,386 veh-min.

As can be seen from Table 9, the overall probability-weighted two-way vehicle minutes of delay per hour are 125.2, 155.3, 84.9, and 12.5 minutes/hour respectively for the four segments.

New Model Analysis

For the new model, the analysis procedure is very similar, but more calculations are involved. On the similar side, the input parameters are the same, as shown by Table 11. However, eight incident types are involved, starting with Property Damage and ending with Weather. And for each of these, there are lane blockage conditions to examine, from none to 3+.

Table 11: New Model Examples

Parameter	Highway Segment			
Road	I-278	I-278	I-278	I-278
Segment	9756	9782	9797	9790
#Lanes	4	6	6	8
Length (mi)	0.88	0.61	1.37	0.42
Capacity	3720	5100	5580	6990
AADT	76500	112300	93790	111300
Volume	1912	2807	2344	2782
Shoulder	Y	N	Y	N
Divided	N	Y	N	Y

Overall				
OneWay Min/Hr	996	282	471	13
TwoWay Min/Hr	1993	565	943	26
Incident Rate	0.3220	0.5277	0.5013	0.3158

Property Damage				
Incident Freq	0.06565	0.44293	0.06565	0.44293
Clearance Time	49.0	35.8	49.0	35.8
0 Lanes Blocked				
Probability	5.1%	5.1%	5.1%	5.1%
Pct Capacity	85.0%	87.5%	87.5%	87.8%
Done Time	0.0	0.0	0.0	0.0
Total Delay	0	0	0	0
Delay/Vehicle	0.0	0.0	0.0	0.0
1 Lanes Blocked				
Probability	72.5%	72.5%	72.5%	72.5%
Pct Capacity	42.5%	56.7%	56.7%	63.8%
Done Time	58.0	0.0	0.0	0.0
Total Delay	7835	0	0	0
Delay/Vehicle	4.2	0.0	0.0	0.0
2 Lanes Blocked				
Probability	17.0%	17.0%	17.0%	17.0%
Pct Capacity	0.0%	23.3%	23.3%	42.5%
Done Time	100.8	61.1	64.8	0.0
Total Delay	78712	29493	27623	0
Delay/Vehicle	24.5	10.3	10.9	0.0
3+ Lanes Blocked				
Probability	3.5%	3.5%	3.5%	3.5%
Pct Capacity	0.0%	0.0%	0.0%	21.3%
Done Time	100.8	79.6	84.5	46.8
Total Delay	78712	66680	80871	18055
Delay/Vehicle	24.5	17.9	24.5	8.3
Prob Weighted Delay After Incident Rate				
OneWay Min/Hr	63	99	34	6
TwoWay Min/Hr	126	199	68	12

Disabled Truck				
Incident Freq	0.07374	0.07374	0.07374	0.07374
Clearance Time	64.5	38.2	64.5	38.2
0 Lanes Blocked				
Probability	6.7%	6.7%	6.7%	6.7%
Pct Capacity	85.0%	87.5%	87.5%	87.8%
Done Time	0.0	0.0	0.0	0.0
Total Delay	0	0	0	0
Delay/Vehicle	0.0	0.0	0.0	0.0
1 Lanes Blocked				
Probability	84.4%	84.4%	84.4%	84.4%
Pct Capacity	42.5%	56.7%	56.7%	63.8%
Done Time	76.2	0.0	0.0	0.0
Total Delay	13555	0	0	0
Delay/Vehicle	5.6	0.0	0.0	0.0
2 Lanes Blocked				
Probability	4.5%	4.5%	4.5%	4.5%
Pct Capacity	0.0%	23.3%	23.3%	42.5%
Done Time	132.6	65.2	85.2	0.0
Total Delay	136175	33597	47789	0
Delay/Vehicle	32.2	11.0	14.4	0.0
3+ Lanes Blocked				
Probability	3.9%	3.9%	3.9%	3.9%
Pct Capacity	0.0%	0.0%	0.0%	21.3%
Done Time	132.6	85.0	111.1	50.0
Total Delay	136175	75959	139910	20568
Delay/Vehicle	32.2	19.1	32.2	8.9
Prob Weighted Delay After Incident Rate				
OneWay Min/Hr	74	10	38	1
TwoWay Min/Hr	148	20	77	2

Road Hazard				
Incident Freq	0.03298	0.03298	0.03298	0.03298
Clearance Time	257.5	154.7	257.5	154.7
0 Lanes Blocked				
Probability	3.8%	3.8%	3.8%	3.8%
Pct Capacity	85.0%	87.5%	87.5%	87.8%
Done Time	0.0	0.0	0.0	0.0
Total Delay	0	0	0	0
Delay/Vehicle	0.0	0.0	0.0	0.0
1 Lanes Blocked				
Probability	63.8%	63.8%	63.8%	63.8%
Pct Capacity	42.5%	56.7%	56.7%	63.8%
Done Time	304.6	0.0	0.0	0.0
Total Delay	216378	0	0	0
Delay/Vehicle	22.3	0.0	0.0	0.0
2 Lanes Blocked				
Probability	13.8%	13.8%	13.8%	13.8%
Pct Capacity	0.0%	23.3%	23.3%	42.5%
Done Time	529.8	263.9	340.6	0.0
Total Delay	2173730	550715	762845	0
Delay/Vehicle	128.8	44.6	57.3	0.0
3+ Lanes Blocked				
Probability	1.3%	1.3%	1.3%	1.3%
Pct Capacity	0.0%	0.0%	0.0%	21.3%
Done Time	529.8	344.1	444.0	202.2
Total Delay	2173730	1245111	2233349	337146
Delay/Vehicle	128.8	77.4	128.8	36.0
Prob Weighted Delay After Incident Rate				
OneWay Min/Hr	677	93	303	3
TwoWay Min/Hr	1353	185	607	6

Disabled Vehicle				
Incident Freq	0.06152	0.18345	0.06152	0.07026
Clearance Time	41.9	23.8	41.9	23.8
0 Lanes Blocked				
Probability	4.8%	4.8%	4.8%	4.8%
Pct Capacity	85.0%	87.5%	87.5%	87.8%
Done Time	0.0	0.0	0.0	0.0
Total Delay	0	0	0	0
Delay/Vehicle	0.0	0.0	0.0	0.0
1 Lanes Blocked				
Probability	90.1%	90.1%	90.1%	90.1%
Pct Capacity	42.5%	56.7%	56.7%	63.8%
Done Time	49.5	0.0	0.0	0.0
Total Delay	5724	0	0	0
Delay/Vehicle	3.6	0.0	0.0	0.0
2 Lanes Blocked				
Probability	3.5%	3.5%	3.5%	3.5%
Pct Capacity	0.0%	23.3%	23.3%	42.5%
Done Time	86.2	40.6	55.4	0.0
Total Delay	57500	13057	20179	0

Personal Injury				
Incident Freq	0.10397	0.10397	0.10397	0.10397
Clearance Time	43.0	43.0	43.0	43.0
0 Lanes Blocked				
Probability	16.4%	16.4%	16.4%	16.4%
Pct Capacity	85.0%	87.5%	87.5%	87.8%
Done Time	0.0	0.0	0.0	0.0
Total Delay	0	0	0	0
Delay/Vehicle	0.0	0.0	0.0	0.0
1 Lanes Blocked				
Probability	44.8%	44.8%	44.8%	44.8%
Pct Capacity	42.5%	56.7%	56.7%	63.8%
Done Time	50.9	0.0	0.0	0.0
Total Delay	6039	0	0	0
Delay/Vehicle	3.7	0.0	0.0	0.0
2 Lanes Blocked				
Probability	34.3%	34.3%	34.3%	34.3%
Pct Capacity	0.0%	23.3%	23.3%	42.5%
Done Time	88.5	73.4	56.9	0.0
Total Delay	60673	42588	21292	0

HAZMAT				
Incident Freq	0.01195	0.01195	0.01195	0.01195
Clearance Time	63.8	63.8	63.8	63.8
0 Lanes Blocked				
Probability	6.9%	6.9%	6.9%	6.9%
Pct Capacity	85.0%	87.5%	87.5%	87.8%
Done Time	0.0	0.0	0.0	0.0
Total Delay	0	0	0	0
Delay/Vehicle	0.0	0.0	0.0	0.0
1 Lanes Blocked				
Probability	79.3%	79.3%	79.3%	79.3%
Pct Capacity	42.5%	56.7%	56.7%	63.8%
Done Time	75.5	0.0	0.0	0.0
Total Delay	13275	0	0	0
Delay/Vehicle	5.5	0.0	0.0	0.0
2 Lanes Blocked				
Probability	13.8%	13.8%	13.8%	13.8%
Pct Capacity	0.0%	23.3%	23.3%	42.5%
Done Time	131.2	108.8	84.4	0.0
Total Delay	133358	93609	46801	0

Looking at the "Property Damage" incidents, one can see that the incident frequency is 0.06565 incidents per day per mile for Segment 9756. This is because Segment 9756 has shoulders and its one-way AADT is less than 49,975. The same logic pertains to Segment 9797. For Segment 9782, the incident frequency is 0.44293 because it does not have shoulders and its one-way AADT is between 49,975 and 56,225. The same pertains for Segment 9790. The clearance times are 49.0 minutes for the facilities with shoulders and 35.8 minutes for those without. Regardless of the number of lanes or the presence or absence of shoulders, the probabilities of lane blockages are 5.1% for none, 72.5% for 1, 17% for two, 3.5% for 3+ and 1.8% for not applicable. The percent capacities remaining, given the lane blockages, are the same as for CNAM. Once these values have been established, the calculations of the “done times”, total delays, and delays per vehicle are the same as for CNAM. Figures 2 and 3 and the formulas apply the same as they did before. The one-way delays are also computed in the same way: (the probability of none, 1, two, and 3+ lanes being blocked) times (the total delay or the delay/vehicle). The overall delays are then computed by summing delays for the eight incident types.

The total incident-weighted two-way vehicle minutes of delay/hour are 1993, 565, 943, and 26 vehicle hours/hr for the four segments respectively. These compare to the values of 125, 155, 85, and 12 minutes/hour obtained from CNAM. Not only are the values bigger, but the relative sizes are different. This is because the two models treat these situations differently. Since no empirical data are available to ascertain which model produces estimates closer to the truth, asking which model is “right” or “better” is impossible. It is only possible to observe that the estimates are different.

5.0 NETWORK-LEVEL COMPARISONS

Two network-level comparisons were conducted. The first was for the segments that lie along the corridor from which the incident data were obtained. The other is for all the freeway segments in the CNAM database for Region 11.

The assumptions used in conducting these analyses are the same as for the four example segments. The incident occurs at t = 0 and lasts as long as each model suggests is appropriate. The demand is 5% of the directional AADT, which in CNAM terms, is the average 4-5 hour volume, intended to compensate for demand fluctuations; it continues until the incident ends.

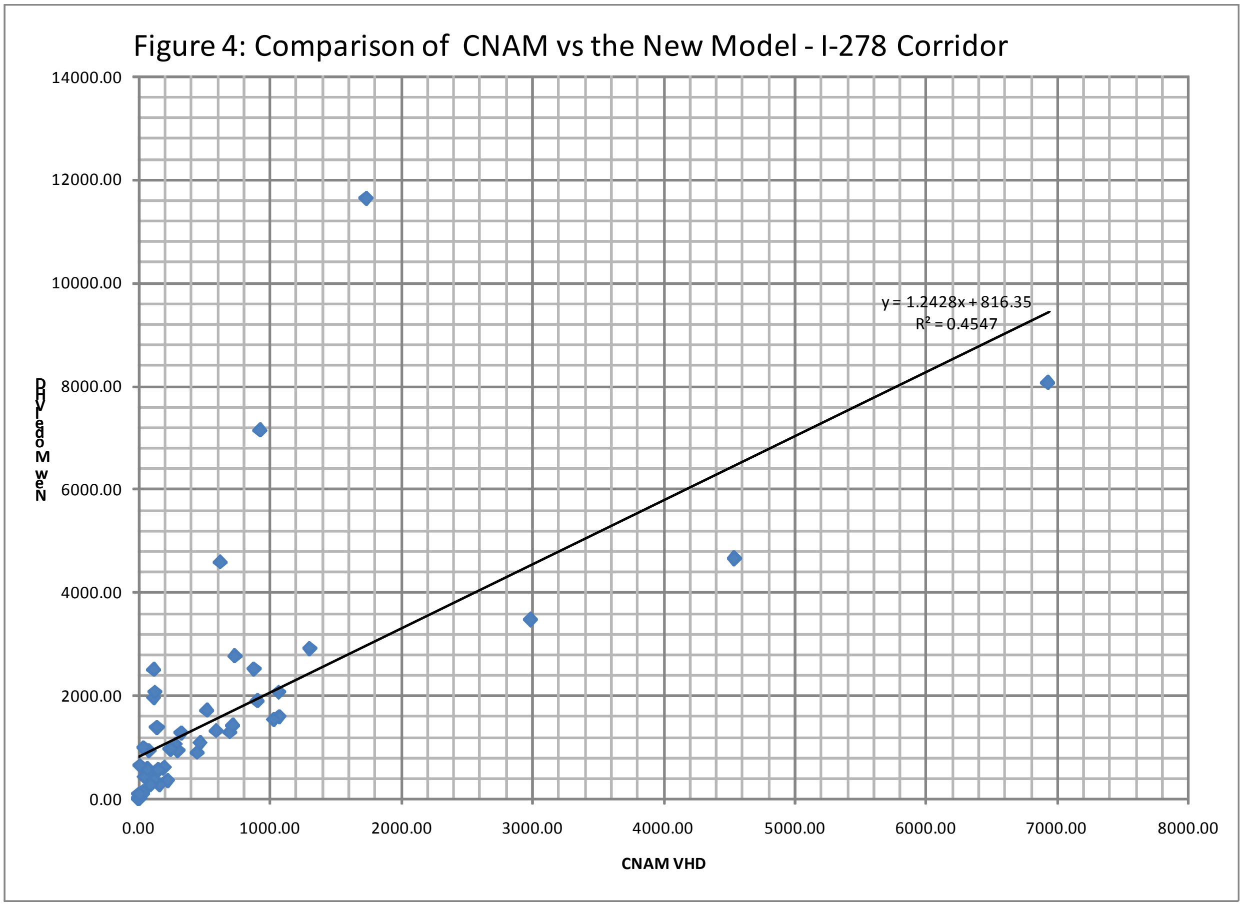
One difference is that, in these analyses, instead of using formulas embedded in the Excel worksheet, a VBA program is employed.[18]

I-278 Corridor

The first case study to which the workspace was applied focuses on the links in the Region 11 that lie along the corridor from which the incident data were obtained. This includes the Gowanus Expressway, the BQE, and the Staten Island Expressway. There are 54 segments in the CNAM database for these facilities, totaling 46.4 miles of freeway. For these segments, CNAM predicts a total of 31,868 vehicle-minutes of delay while the new model predicts 83,688.This is 2.63 times more delay.

A scatterplot of the delays for the individual segments is presented in Figure 4. The delay predicted by CNAM is plotted on the horizontal axis while the delay predicted by the new model is plotted on the vertical axis. The regression line shows that the new model estimates are typically 25% larger than the CNAM estimates, but the correlation is low, with an R^2 of 0.4547. There also appear to be "rays" of proportionality, with one set of segments, toward the bottom right, having about a 1:1 correspondence between the predictions and another set, further toward the top left, having a 6:1 correspondence. A detailed review of the delays for the individual segments does not show any obvious trends in the differences.

[18] Checks were performed to ensure that the worksheet formulas and the VBA code obtain the same results for the four segments analyzed earlier.



<u>All Region 11 Freeways</u>

The second case involves all the limited access links in the Region 11 CNAM dataset (Access = 2). These are the links to which the current CNAM incident delay procedure applies.

The segments are all freeways and they have an average AADT of 113,000[19]. Table 12 presents an overview of these links.

[19] This compares very closely to the 124,040 which the sufficiency file shows for I-278.

Table 12: Overview of the Segments in the Region 11 Sufficiency File

Functional Class	# Seg
11	128
12	126
14	13
Total	267

Route Type	# Seg
A	30
B	7
C	16
D	3
F	5
G	3
H	6
I	134
J	
L	11
M	20
P	4
V	
Other	28
Total	267

# of Roadways	# Seg
1	11
2	256
Total	267

# of Lanes	# Seg
2	1
3	2
4	44
5	6
6	185
7	
8	23
9	6
Total	267

Shoulder Type	# Seg
0	11
1	192
4	60
5	4
Total	267

Shoulder Width	# Seg
0	203
6	1
7	1
8	1
9	2
10	56
12	2
16	1
Total	267

AADT *	# Seg
0	0
1000	0
2000	3
3000	13
4000	4
5000	11
6000	12
7000	10
8000	12
9000	17
10000	19
11000	16
12000	18
13000	20
14000	22
15000	30
16000	21
17000	11
18000	8
19000	3
20000	8
21000	0
22000	0
23000	2
24000	3
25000	0
26000	1
27000	0
28000	0
29000	0
30000	3
Total	267

* Upper Bound of AADT Class
#Seg = Number of segments

To help with readability, here are some example readings of the data in Table 12:

- There are 267 records (specific highway segments) in the file.
- 128 segments are in FHWA *functional class* 11[20]

[20] Functional classes are Principal Arterial Interstate (11), Principal Arterial Expressway (12),

- 134 of the segments are interstates[21]
- 256 of the freeways are separated by a median at least 4-feet wide
- 185 of the segments pertain to 6-lane freeways
- 192 of the segments have a shoulder type 1[22]
- 203 of them have no shoulder
- The AADTs range from 10,000 to 300,000 (multiply by 10) with the most common AADT being between 140,000-150,000.

A picture of the distribution of the AADTs is also presented in Figure 5 below.

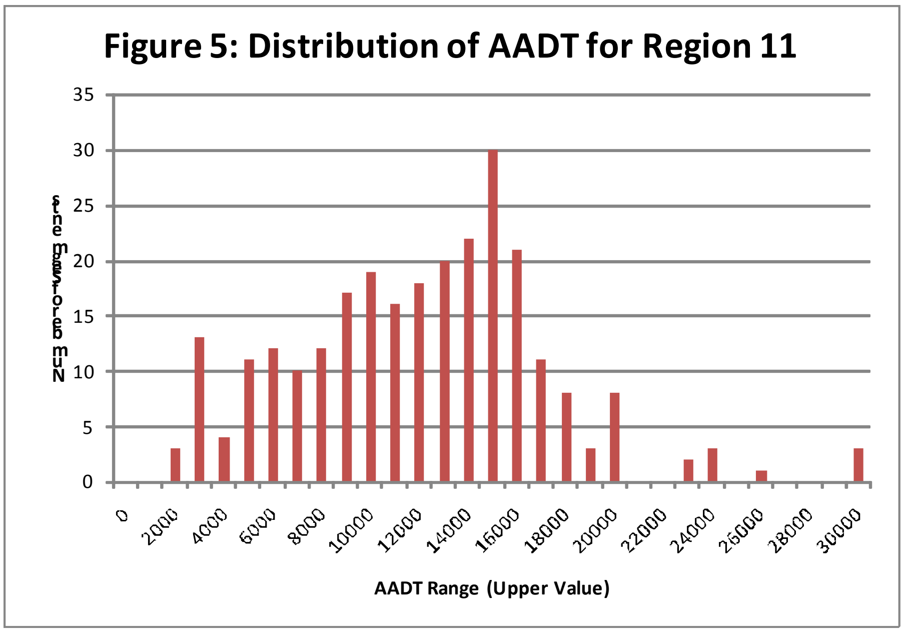


Figure 5: Distribution of AADT for Region 11

The VBA/Excel model uses the same hypothetical situation described earlier:

- The demand is 5% of the AADT.
- The incident occurs at the beginning of the analysis time
- The incident occurs in one of the two directions
- For the analyses based on CNAM look-up tables, an incident of each type is analyzed (number of lanes blocked) for each section and the results are weighted based on the probability that the incident would have occurred.
- For the analyses based on the new look-up tables, an incident of each type – a total of 8 - is analyzed for each section and the results are weighted based on the likelihood that the incident might occur.

Principal Arterial Other (14), Minor Arterial (16), Collector (17), and Local (19)

[21] The other route types are various suffixes, like 9A and 25A, or parkways, like 907F and 907M. The guide to the sufficiency file has more details.

[22] Shoulder types are: Curbed, mowing (0), Curbed, no mowing (1), Stabilized, mowing (4), and Stabilized, no mowing (5). There are others.

- The incident is deemed to end when the queue length returns to zero.
- The incident delay is doubled to reflect the sum of both directions

It is interesting to note that the analysis time can last as long as 8 hours[23].

One minor difference is that the incident is analyzed using deterministic simulation, on a minute-by-minute basis, instead of calculating the incident ending time and total delay.

Attachment D contains the VBA code. Most if not all of the detailed questions about how the analysis has been conducted can be answered by reviewing the code.

Accident and Incident Rates – An Issue

One issue in conducting the analysis relates to accident and incident rates. It is important to talk about the differences, the ratios between them, the default accident rates built into CNAM, the actual accident rate results obtained by the study team from NYSDOT data, and the incident rates derived from the TRANSCOM data.

Accidents are often defined as events involving fatalities, injuries, or significant property damage. Incidents comprise a broader category including vehicle breakdowns, weather events, HAZMAT events, etc. Table 13 compares accident rates and incident rates for several recent studies. Accidents seem to comprise 10-30% of all incidents, except in the case of the 40-year-old study by Goolsby *et al*. (1970) where the split was closer to even.

---

[23] Road hazard incidents, as shown in Table 5 can have durations up to 257.5 minutes

Table 13: Distributions of Incidents among Accidents, Disablements, and Other Events				
Study	Year & Place	Accidents	Disablements	Other
DeRose (1)	MI (06/1962-06/1963)	25%	75%	-
Goolsby et.al. (2)	Houston, TX (1968-1969)	49%	48%	3%
Guiliano (3)	Los Angeles, CA (9/83-6/84 & 9/84-6/85)	11%	80%	9%
Cambridge Systematics (4)	Chicago, Los Angeles, Fort Worth, Minneapolis, NY/NJ - 1990	10%	80%	10%
Skarbadonis et. al. (5)	San Francisco, CA -1990	5%	61%	34%
Garib et.al. (6)	Oakland, CA (02/1993-10/1993)	19%	81%	-
Sullivan (7)	Charlotte, Chicago, Houston, Los Angeles, Orlando, SanFran - 1993	33%	62%	5%
Ozbay & Kachroo (8)	VA (1994-95)	33%	60%	7%
Skarbadonis et. al. (9)	CA (Spring 1993)	10%	89%	1%
Chang G-L. and Rochon (10)	MD, DC (1996-2005)	37%	63%	-

The analysis of NYSDOT accident data and TRANSCOM incident data suggested that for the NYC area about 46% of the incidents are accidents.[24] This means that for the CNAM data presented in Table 1 and the NYSDOT data about to be presented, the incident rates are about twice the accident rates.

With hopes of developing detailed accident and incident rates for the NYC region, the study team obtained 14 months of accident and incident data for 45.85 miles of freeway facilities from TRANSCOM through the diligent efforts of NYSDOT. Table 14 summarizes these data. The dataset contains information about 1907 incidents, half of which are on the BQE, producing incident rates as high as 0.216 incidents per mile per day and 1.71 incidents per million VMT.

---

[24] See Ozbay, K. and M. A. Yazici, *Revised Task 4: Adjusted Accident Frequency Look-Up Tables*, February 22, 2008.

Table 14: Comparisons between the TRANSCOM and NYSDOT Data								
	# of Incidents		Incidents/Day***		Incidents/mile/day*****		Incidents/MVMT****	
Incident Type	TRANSCOM* (02.04 – 04.05)	NYSDOT** (01.00-12.02)	TRANSCOM* (02.04 – 04.05)	NYSDOT** (01.00-12.02)	TRANSCOM* (02.04 – 04.05)	NYSDOT** (01.00-12.02)	TRANSCOM* (02.04 – 04.05)	NYSDOT** (01.00-12.02)
Property Damage	819	11493	1.93	15.74	0.0421	0.3434	0.405	3.303
Personal Injury	67	5356	0.16	7.34	0.0034	0.1600	0.033	1.539
Disabled Vehicle	694	-	1.64	-	0.0357	-	0.343	-
Disabled Truck	179	-	0.42	-	0.0092	-	0.089	-
Vehicle Fire	29	-	0.07	-	0.0015	-	0.014	-
Road Hazard	80	-	0.19	-	0.0041	-	0.040	-
HAZMAT	29	-	0.07	-	0.0015	-	0.014	-
Weather	10	-	0.02	-	0.0005	-	0.005	-
TOTAL	1907	16849	4.50	23.08	0.0981	0.5034	0.943	4.842

Notes

* The TransCom data appear to be missing many accidents. The same may be true for incidents.
** The NYSDOT data contain no incidents. Based on other cities, the incident rate could be 4-5 times larger
*** 424 days for the TransCom data, 730 for the NYSDOT data
**** 4,767,192 Daily VMT for the area studied
***** 45.85 miles

Unfortunately, the study team had to conclude that many incidents, including accidents, were missing from this database.

Accident data from a separate source were then provided by NYSDOT which allowed the study team to generate both accident and incident rates (per day per mile) as shown in Table 15.[25]

Table 15: Comparisons between the TRANSCOM and the Hybrid NYSDOT/TRANSCOM Data								
	# of Incidents		Incidents/Day***		Incidents/mile/day*****		Incidents/MVMT****	
Incident Type	TRANSCOM*	Hybrid**	TRANSCOM*	Hybrid**	TRANSCOM*	Hybrid**	TRANSCOM*	Hybrid**
Property Damage	819	4258	1.93	11.67	0.0421	0.2544	0.405	2.447
Personal Injury	67	1740	0.16	4.77	0.0034	0.1040	0.033	1.000
Disabled Vehicle	694	4786	1.64	13.11	0.0357	0.2860	0.343	2.751
Disabled Truck	179	1234	0.42	3.38	0.0092	0.0737	0.089	0.709
Vehicle Fire	29	200	0.07	0.55	0.0015	0.0120	0.014	0.115
Road Hazard	80	552	0.19	1.51	0.0041	0.0330	0.040	0.317
HAZMAT	29	200	0.07	0.55	0.0015	0.0120	0.014	0.115
Weather	10	69	0.02	0.19	0.0005	0.0041	0.005	0.040
TOTAL	1907	13039	4.50	35.72	0.0981	0.7791	0.943	7.494

Notes

* These values are directly from the TRANSCOM data
** The property damage and personal injury values are from 2001 NYSDOT data. The remaining values are estimated using the following formula: (PD+PI)*1.174*(TC(i)/Sum(TC(i) for non-accidents), e.g., (4258+1740)*1.174*(694)/(694+179+29+80+29+10)
*** 424 days for the TransCom data, 365 for the Hybrid data
**** 4,767,192 Daily VMT for the area studied
***** 45.85 miles

One immediately sees two things in these tables. The first is that the incident rates in Table 15 are significantly higher than those in Table 14. For example, the "Hybrid" model value shown in

[25] See Ozbay, K. and M. A. Yazici, *Revised Task 4: Adjusted Accident Frequency Look-Up Tables*, February 22, 2008 for a more complete description of how the accident and incident rates were developed.

Table 15 (the value used by the new model) is 0.7791 incidents per day per mile while the corresponding value from the TRANSCOM model is 0.0981. (In terms of incidents per million VMT, the contrast is the same, 7.494 incidents per million VMT versus 0.943.) This means the "Hybrid" values are 7.94 times larger than the TRANSCOM values. This is no small difference.

The second observation is that the new incident rate of 7.494 incidents per million VMT is 1.6-2.0 times larger than the 3.6-4.6 values employed by CNAM[26]. So it is not reasonable to expect that the non-recurring delay estimates are likely to be similar for a given network.

Analysis Results

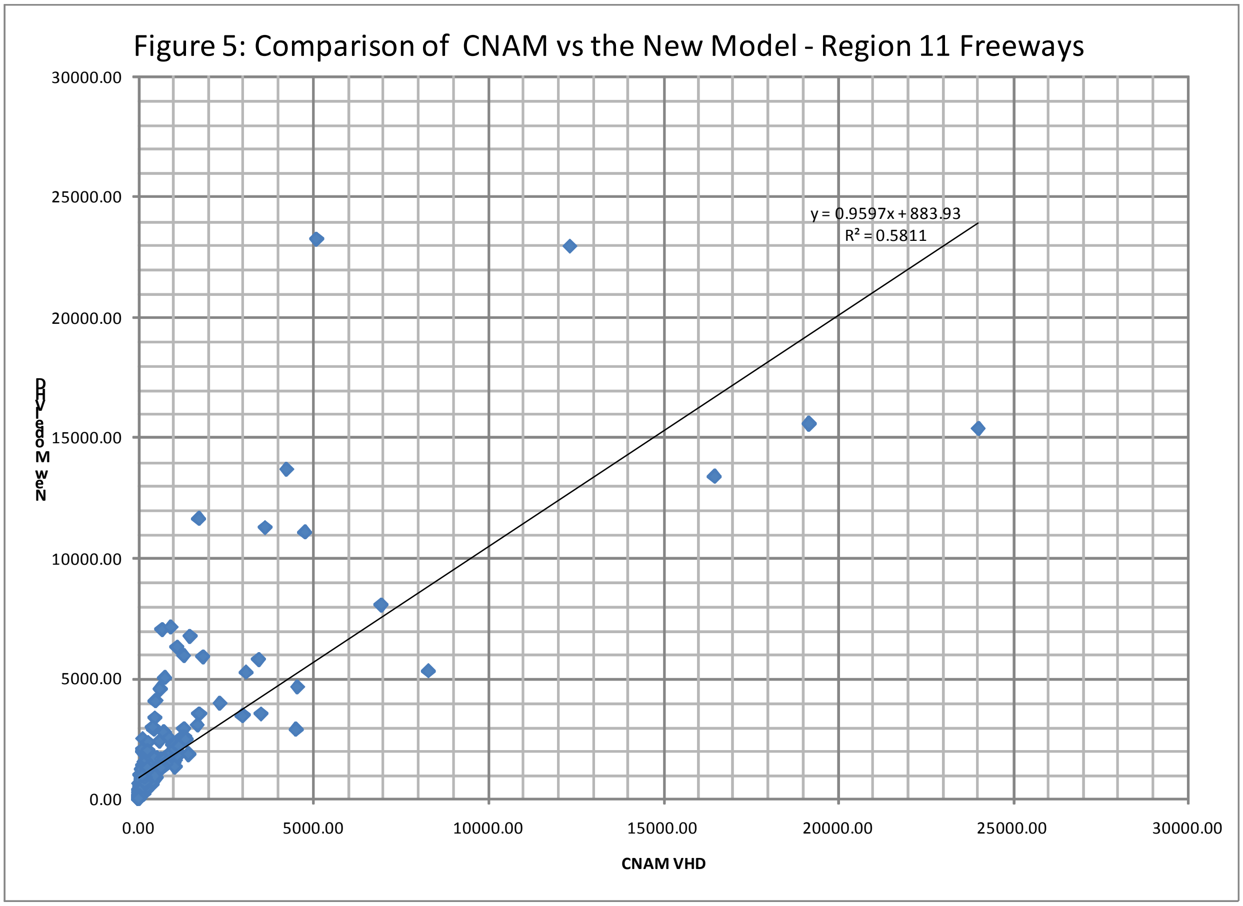
The analysis produces a prediction of 214,727 vehicle-minutes of delay based on the CNAM look-up tables versus 442,078 vehicle-minutes of delay based on the new look-up tables. The new values are about twice as large.

The differences stem from several sources. The most significant are:

- The average *incident rate* for the new model (0.424 incidents/day/mile) is twice that of the average *incident rate* for CNAM (0.211 incidents/day/mile). In a macro sense, this difference carries over into the delay estimates.
- The blockage times used in the new model are longer than those employed by CNAM (please compare Table 5 with Table 4). This means the queues produced are longer, the clearance times, longer, and the total delays, much larger.
- The probabilities of lane blockage used by the new model are higher than those employed by CNAM (please compare Table 8 with Table 2). This produces longer queues and more delay.

Figure 5 presents a scatterplot that compares the delay on each segment for the two models. CNAM's predictions are on the X-axis while the new model's predictions are on the Y-axis. The trend line suggests that the new model is predicting link-specific incident delays that, for the largest values, are about the same as CNAM in that the slope of the line is 0.9597.

[26] Based on 1.59 to 2.0 accidents per MVMT and a 2.3 multiplier for incidents compared to accidents.



6.0 REFLECTIONS

Three significant observations are prompted by this analysis. They all suggest a need for follow-on work.

The first relates to the delay estimates. Unfortunately, no direct observations of delay exist, so it is not possible to validate, verify, or calibrate the estimates of either the new model or CNAM. Only anecdotal comparisons are possible, along with common sense and experience. Hence, while it is possible to observe the significant difference between the delay estimates of the new model and that of CNAM, it is difficult to determine which one is right or better. The strategy used here has been to compute values of the average delay per vehicle and determine that the estimates provided by the new model are likely to be much better.

The second relates to the accident and incident rates. More specificity would be very helpful, for both models. In the case of CNAM, the "problem" is that only two accident rates are driving all of the estimates. They are: 2.0 accidents per million vehicle miles traveled for 2-5 lane undivided and divided highways in an area type 5 and 1.59 for highways of 6 lanes or more. The transformation to accidents/day/mile occurs by post-multiplying by the traffic volume; in this case 5% of the directional AADT. In the case of the new model, five of the seven values (all but property damage and disabled vehicle) are constant regardless of any attribute of the highway (e.g., number of lanes, shoulders, AADT, etc.), so each of those values pertain to every highway segment. The other two are sensitive to AADT, but in a fairly erratic pattern, as shown in Tables 6 and 7, with some incident rates that are 4-6 times the size of the values elsewhere. This means the incident rates have a largely bi-modal distribution, hovering around 0.170 incidents per day per mile most of the time and jumping to around 0.500 when one of the larger values in Tables 6 and 7 are selected because the AADT falls into one of the ranges with the larger incident frequency.

The last issue relates to the look-up tables in the current CNAM model. CNAM may be underestimating the duration of incidents. The TRANSCOM data suggest that the “with shoulder” durations should be 40% larger while the “without shoulder” durations should be 10% smaller. Someone should check to see if the current durations are defensible. The TRANSCOM data suggest otherwise.

7.0 Summary and Conclusions

This report describes how the new look-up tables presented in the Task 4 report would change the manner in which CNAM develops predictions of non-recurring delay and the results obtained.

The two sets of look-up tables are fundamentally different in that the ones presently in use in CNAM are based on incident types that relate to the number of lanes blocked while the new ones are based on categories like “property damage”, “disabled vehicle”, etc. The incident rates employed by CNAM are based on accident rates (accidents per million vehicle miles traveled) prepared by NYSDOT based on statewide crash data while the ones in the new look-up tables are based on incident rates in the TRANSCOM data for the I-278 corridor.[27] Tables 1-4 describe the look-up tables on which the present CNAM methodology depends. Tables 5-8 describe the new look-up tables.

The new model and look-up tables are a good idea for Region 11 to continue pursuing. The incident categories map more closely to traditional definitions of incidents. The delay information the model provides can be very helpful in identifying actions that Region 11 should

[27] 43 of the 267 segments in the controlled access portion of the sufficiency file relate to I-278.

take to reduce the frequency and magnitude of the incidents. The next step is to generate more comprehensive data that can be used to better calibrate and validate the model.[28]

The simple analysis of a hypothetical situation provides a contrast between the predictions provided by the current look-up tables and those that might be expected from the new ones. While the delay estimates are quite different, the trend in comparing the two suggests a similarity. Generally speaking, highway segments that have a large amount of delay based on the current look-up tables are also predicted to have a large amount of delay with the new look-up tables. This suggests that further development and refinement would be worthwhile.

The VBA code developed for the analysis of the hypothetical situation provides an illustration of how CNAM would have to be altered to accommodate the new look-up tables. The finding in that regard is that the change is straightforward and the network dataset already has the fields needed by the new look-up tables. Consequently, implementing the new look-up tables depends on a policy decision and an investment of programmer time. NYSDOT will at some future point in time benefit from evaluating the merits of that investment.

[28] This objective was supposed to be part of this project, but the quality of the available datasets would not support that intent.

ATTACHMENT A
EXCERPT FROM THE CNAM USER MANUAL
XIV. EDIT FREEWAY INCIDENT MODEL TABLES

For advanced users:

A separate manual (inmanual.doc) is available for discussing the incident model and its tables. If this chapter is insufficient to answer your questions please refer to the incident manual.

There are two types of changes that can be made to incident input tables: changing incident parameters and changing accident rates.

INCIDENT PARAMETER TABLES

You may choose to enter data into your own (highway-segment-specific) incident parameter tables for incident duration and capacity available. For these two parameters there are one set of tables for “Do nothing Runs” and one set for “Strategy Runs”. To change these tables choose:

1. Select 'Tools' / 'Edit Incident Parameter Tables' from the CNAM main menu.
2. Select either: >For Do Nothing Run= (Tables are: AVALCAP.DBF and NCIDUR.DBF.)

- OR -

2. 'For Strategy Runs' (Tables are HCAP_AVL.DBF and HINC_RES.DBF.)

Appendix Tables 1 and 2 define the fields in each file you can edit from this screen. Note that on the editing screen in the lower-left hand corner is the table names that are edited by that screen. Please document all changes to any table: reason for changing, new values, locational limits of data (roadway, begin mile and end mile.) If you change these tables and later need to return to the original, please call the CMS team for a new copy.

In the future there will be a screen to edit the other incident parameter table -- for incident factors. Fields that you can edit in dBASE on the incident factor tables are listed on Table 3.

ACCIDENT RATE TABLES

To add new (user-supplied) accident rates (to HACC_RTE) there are two basic methods. Both methods require that you to:

1. Select > Tools > Edit Accident Rate Table.
2. Select the highway segments that you have data for by selecting route, region, county, access control, and number of roadways (this is Screen #1.)

Note 1: These data items are how come regional traffic & safety groups have statistically-grouped data into valid accident rate groups.

3. Select a “regional accident rate table” as an input table.
4. Select either > Rates (accidents/million VMT) or > numbers (numbers of accidents/year for chosen location). See the appropriate subsection for further steps.

FOR RATES

5. The next screen shows the data found in the input table (e.g., DACC_RTE) for the selected parameters (screen #1.) You then can change the accident rate that is displayed (especially if you chose the default table (DACC_RTE as the input table.) Do this by typing over the existing one--keeping in mind that each record may have a different combination of parameters--but same route and county.) Alternatively, you can accept all the information on the screen to be written to your highway-specific accident rate file (HACC_RTE)--especially if you input a file that contains information specifically for the chose route or the data matches that for your chosen route.

Warning: There is no reason for copying the data from the DACC_RTE directly to HACC_RTE with no changes--this will cause the model to use the same data, but force it to do much more searching and slow the model greatly.

Warning: Be sure to check this file occasionally to ensure it has the data you want. If you don=t want the data to be used by the model: copy HACC_RTE to another name (to save data in case it is needed in the future). Zap the original file (removes all data from table). For further information on the accident rate tables see Table 4.

HINT: You must perform this operation for each desired parameter combination and each route and county that you have regionally-obtained data for.

FOR NUMBERS

CNAM allows you to enter the number of accidents per year at each location. The program then uses the AADT and the section length to convert this to the accident rate.

Note 2: AADTs displayed on screen are actually AADT/10 (which is how it is stored on the mainframe sufficiency file.)

5. The next screen shows the data found in the input table (e.g. DACC_RTE) for the selected parameters (screen #1). You then type in the number of accidents/year found on that highway segment.

After you click on the “OK to Update Accidents” box, the conversion to rates is performed and these rates are written to HACC_RTE for those road segments and other parameters selected in (screen #1).

HINT: You must perform this operation for each desired parameter combination and each route and county that you have regionally-obtained data for.

ATTACHMENT B

TABLE DESCRIPTIONS FROM THE CNAM USER MANUAL

TABLE 1: CAPACITY AVAILABLE FACTORS (AVALCAP.DBF & HCAP_AVL.DBF)

Percent of service capacity remaining for vehicle use during an incident--by road segment. Note the table is grouped by number of lanes for the highway or road.

RD_SEQ	Road Sequence
SHLDUSE	Presence or absence of shoulder {ie. Y or N}

4-lane freeway

CAP_04	Percent capacity remaining when incident blocks the shoulder (0).
CAP_14	Percent capacity remaining when incident blocks 1 lane.
CAP_24	Percent capacity remaining when incident blocks 2 lanes.
CAP_34	NOT POSSIBLE (enter 0.0)

6-lane freeway

CAP_06	Percent capacity remaining when incident blocks the shoulder (0).
CAP_16	Percent capacity remaining when incident blocks l lane.
CAP_26	Percent capacity remaining when incident blocks 2 lanes.
CAP_36	Percent capacity remaining when incident blocks > 2 lanes.

8-lane freeway

CAP_08	Percent capacity remaining when incident blocks the shoulder (0).
CAP_18	Percent capacity remaining when incident blocks 1 lane.
CAP_28	Percent capacity remaining when incident blocks 2 lanes.
CAP_38	Percent capacity remaining when incident blocks > 2 lanes.

Notes:

1. Number of lanes is for both directions.
2. To edit: **CMS menu > Tools > Edit Incident Parameter Tables**.

TABLE 2: INCIDENT DURATION (INCIDUR.DBF and HINC_RES.DBF)

This table records the minutes of duration for each "type" of accident (type = by number of lanes blocked and by the location's area-type). For fields starting with LANE the duration includes detection, verification, response, and clearance (D, V, R, L). For fields starting with SHLD the duration includes only shoulder clearance (S). This table is grouped by area-type.

Area Type 1

LANE_01	Duration (min) for incidents blocking the shoulder.
LANE_11	Duration (min) for incidents blocking 1 lane.
LANE_21	Duration (min) for incidents blocking 2 lanes.
LANE_31	Duration (min) for incidents blocking > 2 lanes.
SHLD_01	Duration (min) for incidents blocking the shoulder.
SHLD_11	Duration (min) for incidents blocking 1 lane.
SHLD_21	Duration (min) for incidents blocking 2 lanes.
SHLD_31	Duration (min) for incidents blocking > 2 lanes.

Area Type x (x = 3, 5)

LANE_0x	Duration (min) for incidents blocking the shoulder.
LANE_1x	Duration (min) for incidents blocking 1 lane.
LANE_2x	Duration (min) for incidents blocking 2 lanes.
LANE_3x	Duration (min) for incidents blocking > 2 lanes.
SHLD_0x	Duration (min) for incidents blocking the shoulder.
SHLD_1x	Duration (min) for incidents blocking 1 lane.
SHLD_2x	Duration (min) for incidents blocking 2 lanes.
SHLD_3x	Duration (min) for incidents blocking > 2 lanes.

TABLE 3: INCIDENT FACTORS
(INCIFAC.DBF & HINC_FRQ.DBF)

These incident factors are the percentages of the accident rate that occur for each lane of a road or highway. Note table is grouped by number of lanes on the highway.

SHLDUSE	Presence or absence of shoulder {ie. Y or N}
INCFTR_04	Factor for accidents blocking the shoulder on a 4-lane highway.
INCFTR_14	Factor for accidents blocking 1 lane on a 4-lane highway.
INCFTR_24	Factor for accidents blocking 2 lanes on a 4-lane highway.
INCFTR_34	NOT POSSIBLE (enter 0.0)
INCFTR_06	Factor for accidents blocking the shoulder on a 6-lane highway.
INCFTR_16	Factor for accidents blocking 1 lane on a 6-lane highway.
INCFTR_26	Factor for accidents blocking 2 lanes on a 6-lane highway.
INCFTR_36	Factor for accidents blocking 3 lanes on a 6-lane highway.
INCFTR_08	Factor for accidents blocking the shoulder on an 8-lane
INCFTR_18	Factor for accidents blocking 1 lane on an 8-lane highway.
INCFTR_28	Factor for accidents blocking 2 lanes on an 8-lane highway.
INCFTR_38	Factor for accidents blocking >2 lanes on an 8-lane highway.

Note: Number of lanes is for both directions.

ATTACHMENT C

CNAM's INCIDENT DELAY SUBMODEL MANUAL

INMANUAL.DOC (slightly edited)

4/20/04 VERSION

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Notes to Readers of this Manual

This is the third of three CNAM Model Manuals. The first manual *CNAM Manual*, which is nwmanual.doc, gives specific instructions on how the model operates and how users should operated it. The second is *Recurring Submodel Manual*. The third is the *CNAM's Incident Delay Submodel* This manual covers the theoretical underpinnings of the incident submodel, and Appendix I has user instructions. For user instructions on the overall model, manual "CNAM Manual" should be the main source of information.

Please note that the Powerpoint document (CN04R11.ppt) titled *CNAM Incident Delay Framework* is a short explanation of the incident portion of CNAM.

.

A. OVERVIEW

Introduction

This (CNAM Incident Delay) Manual will discuss the theory, structure, and use of the incident delay model. As with all CNAM components, it is hoped that users will add to the model's usefulness by submitting comments, suggestions, and/or programs. We will be happy to discuss and adopt these suggestions or programs into future versions of the model.

As with the entire CNAM model, no detailed step-by-step instructions are needed to run the model (simply click on applicable buttons), but these steps are included in Appendix I.

The incident delay model is one component of the Congestion Needs Analysis Model (CNAM.) Most CNAM components are written in dBASE 5.5 and accessed through the CNAM's user-friendly menu. The remaining CNAM components -- utilize Arc View to view the delay results in GIS. All CNAM components are discussed in the *CNAM Model Manual*.

Purpose of the Model

The purpose of the incident model is to estimate the magnitude and cost of incident delay on limited-access roads (i.e., freeways.) Another purpose is to be able to assess the effectiveness (on delay) of strategies implemented to reduce incidents. Incidents include accidents, disablements, or non-collisions (such as spills, maintenance, police activities.) Because of the lack of non-accident data, incident delay analysis is performed only for accidents. However, adjustment factors (2.0 for peak and 1.5 for off-peak periods) are then applied to account for disablements and non-collisions. These adjustment factors were established through a comparative analysis by the CMS Team of the share of incident delay attributable to disablements and non-collisions.

Use of Results

The importance of incident delay is many times underestimated, but the *ITE - A Toolbox for Alleviating Traffic Congestion* (1989) estimates that 60% of all freeway congestion is incident related (i.e., those non-recurring delays due to break-downs or accidents.) With recent emphasis on maximizing usage of existing highways, and in an era of decreased funds for highway investments, the need for a CMS to quantify benefits and costs of incident-related congestion is especially clear. This model quantifies the magnitude and identifies the locations of incident delay in terms of vehicle, person, and freight-ton hours of delay (VHD, PHD, FHD.) The model summarizes these performance measures by region, route, by hour of day, and other ways, and with the use of Arc View these results can be directly viewed in GIS.

The incident delay model is designed for use by the New York State Department of Transportation (NYSDOT), Metropolitan Planning Organizations (MPOs), and other operators of major transportation facilities. As for the NYSDOT, the incident delay model will be used to help prepare its capital program (the Five-Year Goal Oriented Program, or GOP) and MPOs will use this model to fully assess the impact of their proposed Transportation Improvement

Programs (TIPs.) It will be used by all transportation providers to predict and assess the effectiveness of projects that aim to reduce incidents.

Basic Theory

The basic theory of the model is based on a queue dissipation model that is nationally accepted. The accident rate of each highway section is the basic input and converted to accidents per hour (using hourly counts), which is then converted to type of accident (blocking X number of lanes.) Using an average duration to clear -- the dissipation model tries to clear the roadway based on the volume and the reduced capacity (due to the incident.) The model sums the delay during the accident.

Input Tables

The incident delay model is designed to work with Highway Sufficiency or Local Highway Inventory road sections. Data from these two files are used to create the CNAM input file HWY_SECT.DBF. This file is maintained by the Planning Bureau. The incident (and recurring) model calculates delay on each road section before moving to the next record.

The input data tables are of two classes: (a) user-supplied tables and (b) program-default tables. The user-supplied tables provide users the opportunity to use local or field-measured data where available. The program uses the program-default tables only when data on the road segment under study are not available in user-supplied table. Table 1 lists the main input tables used by the incident delay model. Some default tables were created from existing Department files (hwy_sect, cnt_stn, cstn_des, trf_cnts) others were produced by the CMS team based on NYSDOT or US DOT data.

Figure 1 shows how the Data Structure is organized; how the tables are related to each other in as a relational database.

Figure 1

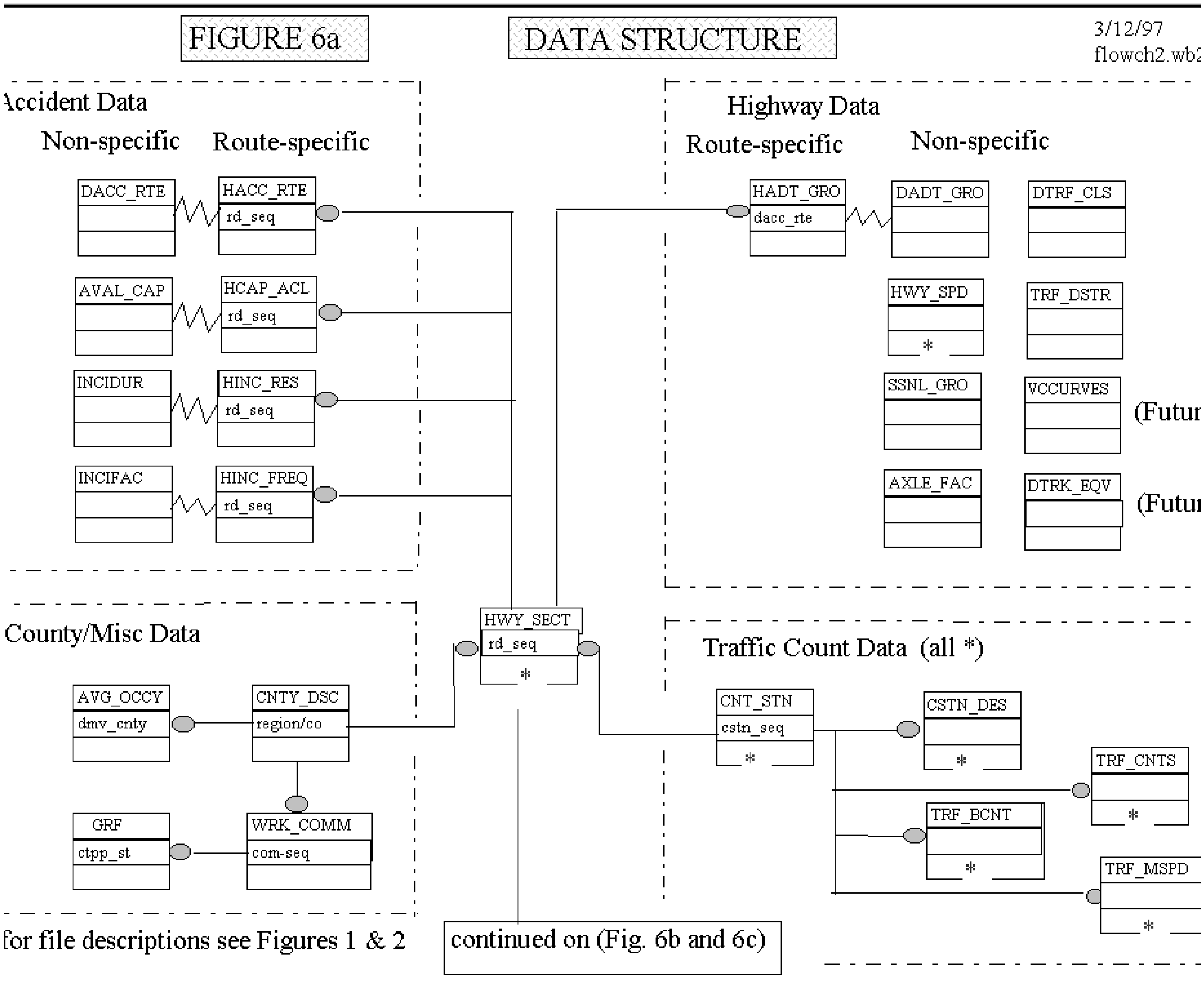


FIGURE 6a
DATA STRUCTURE
3/12/97
flowch2.wb2
Accident Data
Non-specific
Route-specific
DACC_RTE
HACC_RTE
rd_seq
AVAL_CAP
HCAP_ACL
rd_seq
INCIDUR
HINC_RES
rd_seq
INCIFAC
HINC_FREQ
rd_seq
Highway Data
Route-specific
Non-specific
HADT_GRO
dacc_rte
DADT_GRO
DTRF_CLS
HWY_SPD
*
TRF_DSTR
SSNL_GRO
VCCURVES
(Futur
AXLE_FAC
DTRK_EQV
(Futur
County/Misc Data
HWY_SECT
rd_seq
*
AVG_OCCY
dmv_cnty
CNTY_DSC
region/co
GRF
ctpp_st
WRK_COMM
com-seq
Traffic Count Data (all *)
CNT_STN
cstn_seq
*
CSTN_DES
*
TRF_CNTS
*
TRF_BCNT
*
TRF_MSPD
*
for file descriptions see Figures 1 & 2
continued on (Fig. 6b and 6c)

TABLE 1
TABLES: FOR RECURRING & INCIDENT MODEL

	Default Tables	Description	Proc # which uses file	User- Supplied Tables
Roadway/location input files (region-specific):				
1.	HWY_SECT.DBF	Section characteristics	5	No
2.	HWY_SPD.DBF	Speed limit data	6	No
3.	CNTY_DSC.DBF	County description data	14	No
4.	HSRV_FLO.DBF**	Highway service volume data	17	No
Traffic input files (region-specific) :				
5.	CNT_STN.DBF	Count station information	13	No
6.	CSTN_DES	Count station description	13	No
7.	TRF_CNTS.DBF	Data on traffic counts (axles)	15	No
Traffic input files:				
8.	TRF_DSTR.DBF	Hourly distribution of AADT	6	No
9.	DADT_GRO.DBF	Highway AADT growth rates	7	HADT_GRO.DBF
9b.	N/A	VMT-based growth rates		HADT_VMT.DBF
10.	AXLE_FAC.DBF	Axle-correction factors	11	No
11.	DTRF_CLS.DBF	Traffic classification data	12	No
12.	SSNL_GRO.DBF	Seasonal growth factors	16	No
13.	AVG_OCCY.DBF	Auto occupancy data	19	HAVG_OCC.DBF
Accident input files:				
14.	INCIFAC.DBF	Incident factors	8	HINC_FRQ.DBF
15.	INCIDUR.DBF	Incident duration intervals	9	HINC_RES.DBF
16.	DACC_RTE.DBF	Accident rates	10	HACC_RTE.DBF
17.	AVALCAP.DBF	Capacity available data	18	HCAP_AVL.DBF
Base output files:				
18.	HWY_CONG.DBF	Congestion characteristics	23	No
19.	HWY_DLA.DBF	Highway delay results	24	No
20.	RESULTS.DBF	Summary delay results	26	No
21.	RUN_MSGS	Messages from last run	numerous	No
Misc files				
21.	SCEN.DBF	Scenario data	2	No
22.	INDX_TBL.DBF	Index table	5	No

* For a description of all tables use following menu selections in CNAM: HELP/VIEW DEFINITIONS/TABLE NAME.

** Also used as an output file when STG or RTE run is run (creates servol and sequence numbers for these runs.)

Note 1: Most procedures are found in CMSPROC1.PRG, others are in FREEWAY or INCIDENT.PRG.

Input tables can be categorized into: roadway, traffic, and accident-related tables. For roadway data, the data is obtained from NYSDOT mainframe files (Highway Sufficiency File for State

Touring Routes and the Local Highway Inventory for statewide local roads, and the NYC LHI for local roads within NYC.) Mainframe files are obtained for the traffic counts on state routes, on National Highway routes and HPMS sections. Local traffic counts must be organized by the MPO=s starting with the establishment of permanent traffic count stations (last digit of station number = C, T, V or G is reserved for local counts within each county of the state.) Data from each of these mainframe files is then transferred to CNAM's dBASE tables. CNAM has corresponding user-input tables to hold data that Regions may have that concern certain sections of roadways.

Outputs

All outputs are identical to those for the recurring model--with the field delay type = "I" instead of "R". As a review of the two basic outputs, the following is offered, (a full description is found in the *CNAM Model Manual*.

After each analysis run the summary table is displayed (RESULTS.DBF.) The hourly data is stored in two files (HWY_CONG and HWY_DLA.)

Because of the relational database structure of the CNAM, one can not draw complete picture from either of these hourly tables; they must be linked together (related.) These tables can be linked by the CNAM Model by selecting the menu **>GIS** > **CREATE MASTER THEME.** The tables linked are:

HWY_SECT.DBF - parent,
HSRV_FLO.DBF - child,
HWY_CONG.DBF - grand-child, and
HWY_DLA.DBF - great-grand child.

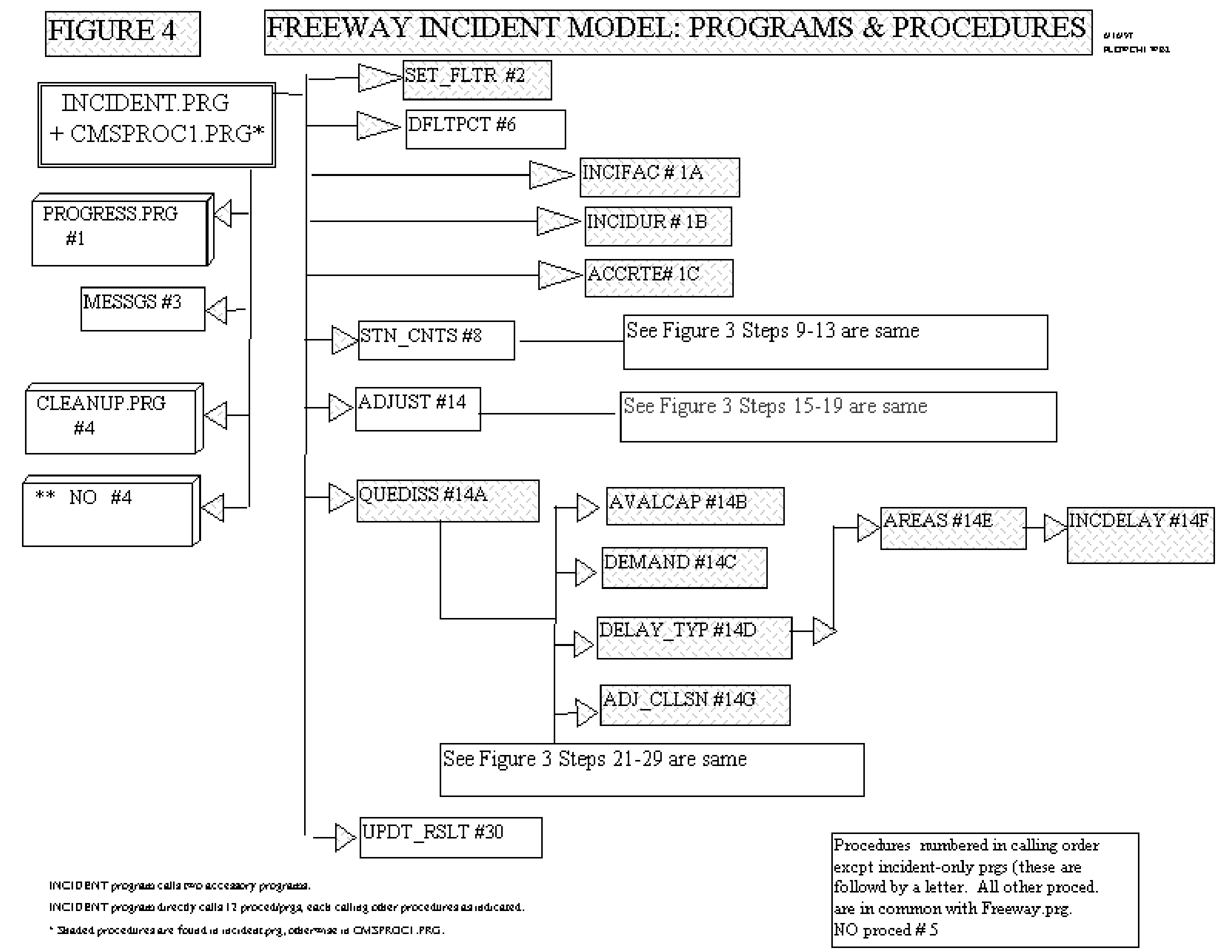
Note: Parent is the base file. The child file is linked to the parent file, and the grandchild is linked to the child file etc.

This creates a GIS compatible file, with one record for each type of delay, each performance measure, for each hour that has delay.

B. CALCULATION OF INCIDENT DELAY

Most of the calculations that are completed are in the QUEDISS.PRG (Program # 20.) Each step of the analysis is explained below. The four accident-related tables listed in this section are described in more detail Section C. The overall structure of the model (procedures called in CNAM) are shown in Figure 2.

Figure 2



Routes Segments

Each road segment is treated independently and in order. The road segments are the same as reported in the Highway Sufficiency File or the Local Highway Inventory. For those sections where no hourly volumes exist, a temporal (hourly) distribution of AADT is used to compute the hourly traffic volumes in each direction. The AADT used is the annually-estimated AADT found in the Highway Sufficiency File. If hourly counts are used, they are seasonally factored, axle corrected, and growth factored to the scenario year.

Demand/Volume Curves

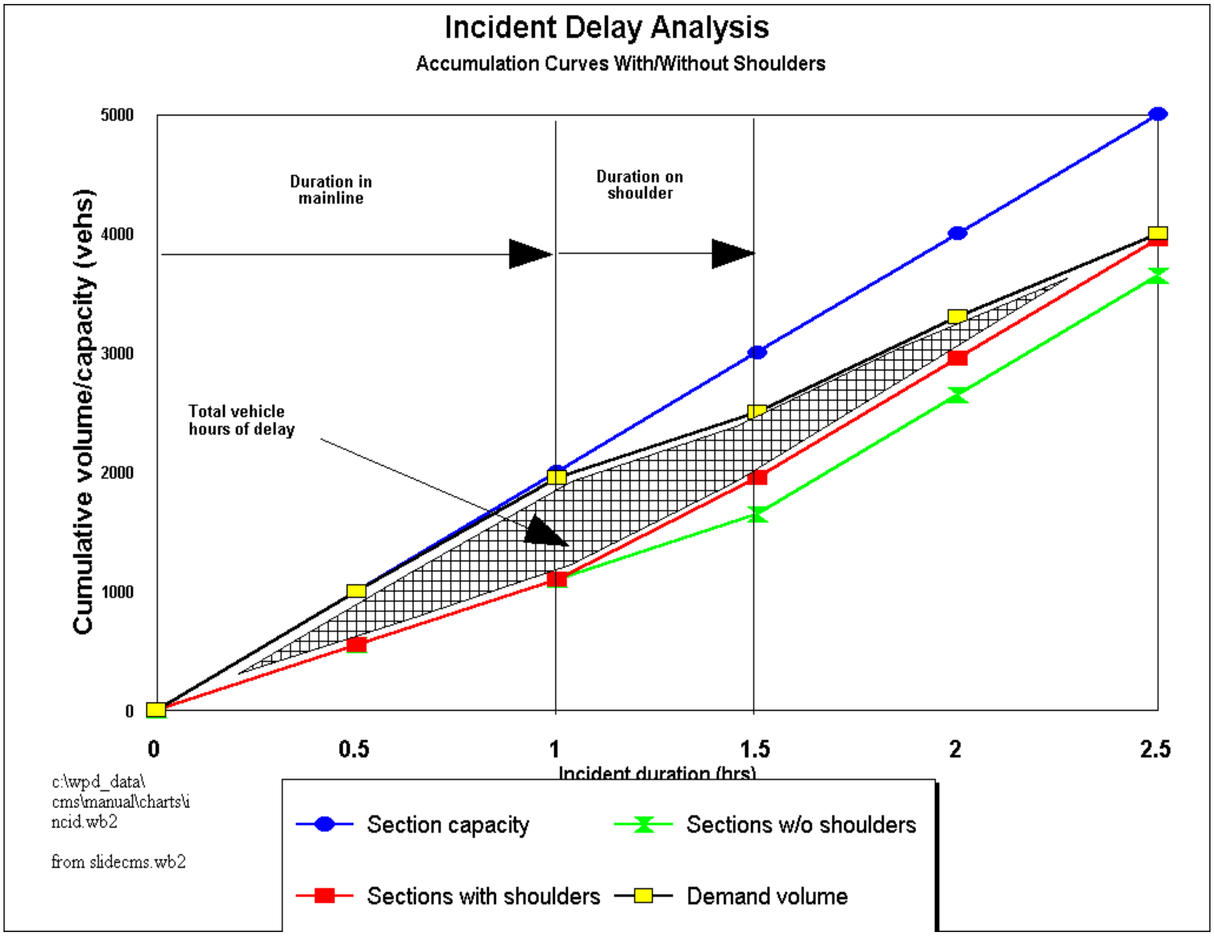
An incident is assumed to occur on average at the mid-point of the hour under consideration. The mid-point assumption is based on stochastic (probabilistic) process. A demand (volume) accumulation function is developed based on the demand during and after the incident. The demand accumulation is carried until the build-up queue dissipates or up to four hours after the incident, whichever is smaller. In developing the demand curves the model looks ahead 5 hours in time and uses the hourly demand which would be occurring during this period in computing the queue build-up and dissipation. In other words, when an incident occurs, the fluctuation in traffic is accounted by the model.

The number of accidents is based on the volume and the accident rate (DACC_RTE.) Also these accidents are classified as to the number of lanes blocked (incident type) and then using the lookup table (INCIFAC) the percent of accidents on that road are calculated for each incident type.

Service/Supply Curves

A service accumulation curve is developed to measure the cumulative volume of vehicles serviced during and after the incident (AVALCAP.) The service accumulation curve is based on the capacity available during the incident as well as capacity of the road section under prevailing traffic and environmental conditions. It should be recognized that service volumes are constrained by the section capacity. This is particularly important for future scenarios where future traffic demand is projected to exceed capacity of the road section. Figure 3 shows the volume and service accumulation functions.

FIGURE 3



Note that the red curve and green curve show the flow past an incident location. Concerning the red curve the first hour is the flow before the accident is cleared off the travel lanes. The remainder hours show the optimal getaway flow. The green curve has a ½ hour of additional lane clearance time. This is similar to the graph shown in 1. Gordon 1996.

Queue Dissipation

The time to restore the highway (segment) to normal flow is obtained by finding the queue dissipation time. The program computes the queue dissipation time by comparing the service volume to the demand volume at 15 minute intervals -- dependent on the incident duration (INCIDUR.) Theoretically, the queue ends when the total volume serviced exceeds cumulative demand. Decreasing the interactive time interval to say 5 minutes increases accuracy, however, it increases the program run time considerably. The 15 minutes intervals give optimal results without sacrificing accuracy. The queue dissipation times are estimated by the routine "*quediss*".

Delay Calculation

To avoid the complex mathematical equations involved a series of areas under the service and volume accumulation curves are used to determine when a queue dissipates and in which interval the queue dissipation occurs. This task is accomplished by the procedure "*areas*". Figure 3 shows a series of such areas under the volume (demand) curve used in the incident delay model.

The incident delay is computed by subtracting the area under the service volume accumulation curve (capacity-minutes) from the area under the demand accumulation curve (volume-minutes.) The total incident delay is computed by adding together the incident delay attributable to the four incident types. This task is performed by the routine "*incdelay*." The annual incident delay occurring within the hour is computed by multiplying the total incident delay by 260.

C. ACCIDENT-RELATED TABLES

When the incident program runs, a number of procedures (small programs within a program) are called to calculate the major parameters used by the incident model: Incident factor, accident duration, accident rate, and available capacity for each incident type. Before each of the tables is discussed, the term incident type needs to be defined.

Incident Type (not a table)

Is a classification of accidents that three of the four accident tables utilize (the accident rate table does not use this variable). The classification is based on the number of lanes which are blocked during an incident:

* accidents that block the shoulder,
* accidents blocking one travel lane,
* accidents blocking two travel lanes, and
* accidents blocking more than two travel lanes.

1. Incident Factor Table

The first procedure that the main incident program (incident.prg) calls is incifac, which looks up the percentage of the total accidents on the road that occur for each "incident type". These percentages depend on the roadway width (in number of lanes) and whether the road has a usable shoulder or not. A shoulder is considered useable if it is wider than six feet. See table C-1A for the field definitions of this table.

The program-default incident factors are stored in the file **INCIFAC.DBF** see Table C-1B. The user-supplied incident factor table is **HINC_FRQ.DBF** which has the same structure.

Concerning obtaining user input values, these data are not collected at site. Users are cautioned to carefully inspect and validate user-supplied data. Under certain conditions (especially high volume roads) invalid incident factors may lead to dramatic over- or under-estimation of incident delay.

The following are some useful guidelines to aid users in the supply of local incident factors:

1. The presence of shoulder increases the likelihood for drivers to pull onto shoulders especially for disablements.
2. The proportion of incidents blocking one lane for sections without shoulders should be greater than the proportion of incidents blocking one lane on sections with shoulders.
3. The proportion of incidents occurring on the shoulder should be greater than mainline incidents blocking one lane.
4. The proportion of incidents blocking one lane should be greater than the proportion of incidents blocking two lanes, and likewise, the proportion of

incidents blocking two lanes should be greater than the proportion of incidents blocking three or more lanes.

TABLE C-1A
INCIDENT FACTORS: FIELD DEFINITIONS
(INCIFAC.DBF & HINC_FRQ.DBF)

The following are these two table's fields defined:

SHLDUSE Presence or absence of shoulder {i.e., Y or N}

INCFTR_04 shoulder on a 4-lane hwy.
INCFTR_14 1 lane on a 4-lane hwy.
INCFTR_24 2 lanes on a 4-lane hwy.
INCFTR_34 **NOT POSSIBLE** (enter 0.0)

INCFTR_06 shoulder on a 6-lane hwy.
INCFTR_16 1 lane on a 6-lane hwy.
INCFTR_26 2 lanes on a 6-lane hwy.
INCFTR_36 3 lanes on a 6-lane hwy.

INCFTR_08 the shoulder on an 8-lane
INCFTR_18 1 lane on an 8-lane hwy.
INCFTR_28 2 lanes on an 8-lane hwy.
INCFTR_38 >2 lanes on an 8-lane hwy.

TABLE C-1B
INCIDENT FACTORS: DATA in % OF ACCIDENT RATE
(INCIFAC.DBF)

Section Characteristics		On Shoulder	Blocking 1 Lane	Blocking 2 Lanes	Blocking 3 Lanes
4 - 5 Lanes	Field name	**INC*_A04**	**INC*_A14**	**INC*_A24**	**INC*_A34**
	Shoulder	0.48	0.44	0.08	0.00 (N/A)
	No Shoulder	0.00 (N/A)	0.85	0.15	0.00 (N/A)
6 - 7 Lanes	Field name	**INC*_A06**	**INC*_A16**	**INC*_A26**	**INC*_A36**
	Shoulder	0.47	0.43	0.08	0.02
	No Shoulder	0.00 (N/A)	0.81	0.15	0.04
8 or more Lanes	Field name	**INC*_A08**	**INC*_A18**	**INC*_A28**	**INC*_A38**
	Shoulder	0.47	0.43	0.08	0.02
	No Shoulder	0.00 (N/A)	0.81	0.15	0.04

Note: 1. N/A = Not applicable
2. **INC*** was shortened from INCFTR (to save space.)
3. Edit this field in CNAM with the *BROWSE ANY TABLE* menu selection.

Source: see ref. # 11
To programmers: The code that combines lanes is found in incident.prg:
If num_lnes <= 5; sl = 4.

2. Incident Duration Table

The second procedure (incidur) looks up the incident duration (in minutes) for each incident type, for each *area-type* of the selected roadway section. Area-type is defined by the Highway Sufficiency Report and is obtained from Table C-2B.

Incident duration usually is composed of five intervals, namely, detection, verification, response, lane-clearance, and shoulder-clearance. The detection and verification intervals depends on the available technology in the region or route under study such as CCTV, loop detectors, HAR, probes, and patrols teams. The response interval on the other hand depends on the degree of accessibility to the location of the incident and the time of the day (peak or off-peak). The degree of accessibility is difficult to measure, but it depends on the spacing of exit and entry ramps, availability of useable shoulders, and frontage roads. The lane-clearance and shoulder-clearance intervals depend on severity of the incident. It should be recognized that the shoulder-clearance interval for those sections of highways with no usable shoulders are set to zero. Because incident duration in disaggregate form is not necessary for computing incident delay, processing disaggregate response times considerably increases program run-time.

TABLE C-2A
INCIDENT DURATION: FIELD DEFINITIONS
(INCIDUR.DBF and HINC_RES.DBF)

A CNAM screen is available to edit this table. LANE durations include detection, verification, response, and clearance. SHLD durations include only shoulder clearance (s). The incidur table is statewide default values; the other is user-input values. The name of the field reflects three characteristics: 1) whether in lane or on shoulder, 2) the incident type, 3) the area-type as shown:

Blockage	Fieldnames relating to lane blockage	Fieldnames relating to to shoulder blockage
	Area Type 1	
Shoulder	LANE_01	SHLD_01
1 Lane	LANE_11	SHLD_11
2 Lanes	LANE_21	SHLD_21
>2 Lanes	LANE_31	SHLD_31
	Area Type x (*x* = 2, 3, 4, 5, 6)	
Shoulder	LANE_0x	SHLD_0x
1 Lane	LANE_1x	SHLD_1x
2 Lanes	LANE_2x	SHLD_2x
>2 Lanes	LANE_3x	SHLD_3x

Note: Area-type is defined in Table C-3B, which also gives the incident durations for each classification.

TABLE C-2B
INCIDENT DURATION: DATA IN MIN (INCIDUR.DBF)

Incident Location		On Shoulder	Blocking 1 Lane	Blocking 2 Lanes	Blocking 3 Lanes
Area Type 1 (rural)	Field name	**LANE_01**	**LANE_11**	**LANE_21**	**LANE_31**
	Lane clearance time	61	62	63	67
	Field name	**SHLD_01**	**SHLD_11**	**SHLD_21**	**SHLD_31**
	Shoulder clearance time	0	15	15	20
Area Type 3 or 2 (<5000 pop)	Field name	**LANE_03**	**LANE_13**	**LANE_23**	**LANE_33**
	Lane clearance time	37	38	39	43
	Field name	**SHLD_03**	**SHLD_13**	**SHLD_23**	**SHLD_33**
	Shoulder clearance time	0	15	15	20
Area Type 4 (suburban)	Field name	**LANE_04**	**LANE_14**	**LANE_24**	**LANE_34**
	Lane clearance time	51	52	53	59
	Field name	**SHLD_04**	**SHLD_14**	**SHLD_24**	**SHLD_34**
	Shoulder clearance time	0	15	15	20
Area Type 5 or 6 (city/lg vill.)	Field name	**LANE_05**	**LANE_15**	**LANE_25**	**LANE_35**
	Lane clearance time	43	44	45	51
	Field name	**SHLD_05**	**SHLD_15**	**SHLD_25**	**SHLD_35**
	Shoulder Clearance time	0	15	15	20

Notes: 1. Number of lanes is for both direction as used in the sufficiency file of NYSDOT.
2. Where no shoulder exists, the duration of incident on shoulder is set to zero.

I believe this table could do with a good updating.

My quick search of the national research (see Ref #22), shows that Lindley 1986 states the he created his Table 3, and 4 (p. 13 and 14) based on three other studies (see ref # 22). These same tables were used in a consultant report for NYSDOT PIN 8729.30 “Freeway Delay Analysis Software Users Guide”, Nov 1992 by Howard Needles Tammen & Bergendoff.

Lindley’s tables seem much more logical than CNAM’s table as far as the differences between the duration of accidents blocking 1, 2, or 3 lanes. As this report finds there is a 10 minute per lane increase in duration as the number of lanes blocked increase by each lane. CNAM’s table shows no such large increase – especially going from 1 lane to 2 lanes. I have combined rows in the Lindley’s tables to correspond to CNAM’s organization of data as follows:

EVENT DURATIONS
(values in minutes)

Location	Shoulder Disablement	Shoulder Accident	Disablement with lanes blocked: One	Two	Accident with lanes blocked One	Two	Three
FREEWAYS WITH ADEQUATE SHOULDERS							
In lane	20	20	25	30	30	35	40
On shoulder*	10	20	15	15	20	25	30
Total	30	40	40	45	50	60	70
FREEWAYS WITH NO SHOULDERS							
In lane	-	-	10	25	30	40	50
On shoulder*	-	-	0	0	0	0	0
Total**	-	-	10	25	30	40	50

*after response
** corrected from original version

3. Accident Rate Table

DACC_RTE is the base table provided by the Main Office and its source is data received from Traffic Engineering and Safety Div (based on raw accident data from NYS Dept. of Motor Vehicles). This table has accident rates for roads by access control, area-type, number of lanes, and number of roadways. Since the Department records do not have accident rates for all possible permutations of these fields, the model defaults to a statewide average accident rate for the missing permutations. Also note that this data is on reportable accidents (accidents requiring an accident report be filled out). Some single vehicle and small damage accidents do not need to have an accident report be filed. Thus CNAM results are adjusted to account for all accidents and other incidents (see next Section).

If regions have better accident rates, either by the fields listed above or by route or by rd_seq, the region is encouraged to add to or modify the respective table. The HACC_RTE table is organized by road sequence number (rd_seq), if you have a route-specific accident rate then the model is set up to look up each rd_seq on the route and enter the data for each rd_seq number—when you select **region** and **route** on first screen. Note to change either of these tables select **CNAM menu >Tools > Edit Accident Rate Table.**

TABLE C-3
ACCIDENT RATE: FIELD DEFINITIONS
(DACC_RTE.DBF)

ACCS_CNTL: This field defines the extent of access available to get on the roadway facility:

>N= = No control of access
>P= = Partial control of access
>F= = Full control of access.

AREA_TYP : Area type defines the geographic and population size characteristics of the area around the highway section.

NUM_LANES : Number of lanes in both directions.

NUM_RDWYS : An undivided roadway is coded >1', and a divided highway is coded >2.' Divided highway has to have a median wider than four feet.

ACC_RATE : Number of accidents per million vehicle miles traveled (ACCID/MVMT.)

4. Available Capacity Table

Capacity available during an incident depends on the "incident type", width of the road (in number of lanes) and whether there is a usable shoulder available. In the presence of a shoulder, some capacity is regained when an incident is cleared onto the shoulder. On sections without shoulders, no capacity is regained until an incident is cleared from the highway.

The model calculates the remaining capacity (during an incident) of the selected roadway section. Available capacity is used to estimate when a queue dissipates and the highway returns to its normal flow and operating conditions. The default capacity available is expressed in percentages and is stored in **AVALCAP.DBF**. Users may enter local field data in the highway-segment-specific table **HCAP_AVL.DBF**. The field definitions are given in Table C-4A, and the data in Table C-4B. This data differs quite a bit from similar data found in the research (see ref # 44).

TABLE C-4A
AVAILABLE CAPACITY: FIELD DEFINITIONS
(AVALCAP.DBF & HCAP_AVL.DBF)

Percent of service capacity remaining for vehicle use during an incident--by road segment.

SHLDUSE Presence or absence of shoulder {i.e., Y or N}

Fieldnames relating to capacity on 4-lane hwy.	% capacity remaining on a:	Fieldnames relating to cap. on 6-lane hwy.
CAP_**04**	4-lane hwy. when incident blocks the shoulder (0)	CAP_**06**
CAP_**14**	4-lane hwy. when incident blocks 1 lane.	CAP_**16**
CAP_**24**	4-lane hwy. when incident blocks 2 lanes	CAP_**26**
CAP_**34**	**NOT POSSIBLE** (enter 0.0)	CAP_**36**

Fieldnames relating to capacity on 8-lane hwy.	
CAP_**08**	8-lane hwy. when incident blocks the shoulder (0.)
CAP_**18**	8-lane hwy. when incident blocks 1 lane.
CAP_**28**	8-lane hwy. when incident blocks 2 lanes.
CAP_**38**	8-lane hwy. when incident blocks > 2 lanes.

Notes:
1. The table is grouped by number of lanes for the highway or road.
2. Number of lanes is for both direction
3. To change this table use **CNAM menu >Tools > Edit Incident Parameter Tables.**

TABLE C-4B
AVAILABLE CAPACITY: DATA: PERCENT AVAILABLE (AVALCAP.DBF)

Incident Location		On Shoulder	Blocking 1 Lane	Blocking 2 Lanes	Blocking 3 Lanes
4-5 Lanes	Field name	CAP_04	CAP_14	CAP_24	CAP_34
	Shoulder	85.0	42.5	0.0	0.0 (N/A)
	No Shoulder	0.0 (N/A)	42.5	0.0	0.0 (N/A)
6-7 Lanes	Field name	CAP_06	CAP_16	CAP_26	CAP_36
	Shoulder	87.5	56.7	23.3	0.0
	No Shoulder	0.0 (N/A)	56.7	23.3	0.0
8 or more Lanes	Field name	CAP_08	CAP_18	CAP_28	CAP_38
	Shoulder	87.8	63.8	42.5	21.3
	No Shoulder	0.0 (N/A)	63.8	42.5	21.3

Notes:
1. Number of lanes is for both direction as used in the sufficiency file of NYSDOT
2. Factors are for accidents only
3. N/A = Not Applicable
4. Rubbernecking reduces capacity in dir opposite of the incident (model ignores this).

D. PROGRAM MODULES (PROCEDURES AND PROGRAMS) -- for programmers.

The incident delay model is composed of procedures and programs. Procedures act like programs, but reside within programs (called subroutines in other programming languages.) A flow chart showing the relationships between the programs and procedures is shown in Figure 2.

The following is a discussion of the incident-specific programs & procedures (for all other programs listed in Figure 2 please see the ACNAM Model Manual@.) Numbers in parenthesis match the numbering system used in Figure 2.

1. *INCIFAC* (1A) — Reads from **INCIFAC.DBF** the incident factors (or % of all accidents that are of each incident type (block shoulder, block one lane, block 2 lanes, block 3 or more lanes). This table has a value for each of the 4 incident types , 3 incid-specific-number-of-lanes, 2 types of lanes (ie. shoulder present, not present) which adds to 24 different values. Users can enter their own factors into **HINC_FRQ.DBF**.

2. *INCIDUR* (1B) — Reads the incident duration intervals (minutes of duration) for each incident type from **INCIDUR.DBF**. This table is based on 4 incident types, 4 incid-specific-area-types, 2 lane types (shoulder, no shoulder) which adds to 32 different values. These intervals are used in computing queue build-up and time to restore the segment to normal flow. Users can enter their own intervals into **HINC_RES.DBF**.

3. *ACCIRTE (1C)* — Reads from **ACCIRTE.DBF** the accident rate (number of accidents per million VMT) for each type of roadway. Table has a values based on: access control, area type, number of lanes, number of roadways. The accident rate is multiplied by the incident factor and also the volume of traffic to obtain the corresponding annual number of incidents which are accidents. Users can enter their own rates into **HACC_RTE.DBF**.

4. QUEDISS (14A) — Computes time to normal flow (i.e., when the road is restored to normal flow pattern.) This is done by comparing volume-minutes and capacity-minutes at 15-minutes intervals until queue dissipates. Procedure uses five hourly traffic volumes starting from the hour under study to calculate when the queue ends. These hourly traffic volumes, however, are constrained in the procedure to the maximum service volume.

5. *AVALCAP (14B)* — Reads % available capacity after an incident from AVALCAP.DBF. This table has a value for each of the 4 incident types (ie. shoulder, block 1 lane, etc.) , 3 incid-specific-number-of-lanes, 2 types of lanes (ie. shoulder present, not present) which adds to 24 different values. User can input their own available capacity table into **HCAP_AVL.DBF**.

6. *DEMAND (14C)* Looks at demand volumes for next five hours and calculates the duration for the four incident intervals.

7. *DELAY_TYP* (14D) Computes incident delay for each of the incident types.

8. *AREAS* (14E) Computes area under the volume (demand) and the capacity (supply) curves. It is the area sandwiched between these two curves which corresponds to total incident delay within the hour.

9. *INCDELAY (14F)* Computes incident delay attributable to each of the four incident types (accidents on shoulder, accidents blocking 1 lanes, accidents blocking 2 lanes, and accidents blocking 3 or more lanes.) All four incident types are summed together over a year to estimate annual incident delay occurring within the hour under study.

10. *ADJ_CLLSN (14G)* Accounts for non-reportable, non-collision, disablement incidents.

E. ADJUSTMENTS TO TOTAL INCIDENT DELAY

Because the accident rate used in CNAM reflects only reportable accidents by the police to the Department of Motor Vehicles, an adjustment is made to increase the delay due to the incidents reported is brought up to that which reflects total incidents (including breakdowns (tire changers, dead batteries, stalling, running out of fuel) and police activity, lost or tired motorist, stopping for

Cell calls, debris in the road, special events, storms and snowfall ????

The adjustment factor for the peak hours (7-9 am and 4-6 pm) is 2.3 for other times it is 1.70. Source of these values are not available at this time.

Code is found in incident.prg and is:

```
    CASE M->J = 8 .OR. M->J = 9 .OR. M->J = 17 .OR. M->J = 18
            DELAY[M->I,M->J] = DLAHRS * 2.30
   OTHERWISE
            DELAY[M->I,M->J] = DLAHRS * 1.70
```

F. STEPS TO RUN THE MODEL -- for new users

The incident delay model should be run on a 486 or Pentium. dBASE 5.5 is required and all tables and programs must be installed on the designated directories (e.g. C:\CNAM_MOD\1995DATA\R05STATE.) To run the incident delay model:

Step 1: Double-click the dBASE 5.5 for Windows icon (in the program manager) to call up dBASE 5.5.

Step 2: Navigate to the c:\cnam_mod directory.

Step 2: Click on the "program" icon to display all program files(.prgs.)

Step 3: Double-click 96MAINMN.prg which will display the CNAM screen.

Step 4: Click on ANALYSIS menu button to display pulldown menu of choices of available types of analyses: (a) Recurring Delay and (b) Incident Freeway Delay, and c) Recurring (project-level) Intersection Delay--this step requires detailed input tables--call CMS Team.)

Step 5: Click the Incident Freeway Delay choice to display the pulldown menu of choices: (a) Incident Delay for All Routes (b) Incident Delay for Specific Route.

Step 6: Click the AAll Routes@ choice to display the incident delay screen. This screen asks the user to input: (a) peak-hour factor, and (b) year to grow traffic to. These inputs are explained below.

Peak-Hour Factor
The peak-hour factor is chosen by using the "spinbuttons" (up and down arrows.) We recommend using 0.95 for all runs. The theory of peak-hour factor is found in the HCM, but it basically accounts for the peak 15-minutes period within the hour (because NYSDOT=s traffic counts are full hourly volumes and not broken to 15-minutes intervals.)

Year to Grow Traffic To (scenario year)
The scenario year is selected from a "list box" (see Figure 1.) The scenario year indicates which year the analysis is to be performed for, four scenario years are provided; (1) Base year (beginning of the capital program or GOP year), (2) Five years from the base year (end of the 5-year capital program or GOP interval), (3) Year 2015 (for air quality analysis) and (4) Year 2020 (25 years from the base year--for Long Range Plan purposes.) These years of analysis are stored in SCEN.DBF.

Each scenario must be run separately and analyzed independently. It must be pointed out that each of these four scenarios will run on only the base network (i.e., no improvement to the network.) An after treatment network can be outputted (see CNAM Model Manual).

REFERENCES

Goolsby, M. E., “Influence of Incidents on Freeway Quality of Service”, Highway Research Record No. 349, TRB, 1971.

Gordon, R.L., R.A. Reiss, H. Haenel, E.R. Case, R.L. French, A. Mohaddes, and R. Wolcott. Traffic Control Systems Handbook. Report FHWA-SA-95-032. Federal Highway Administration, U.S. Department of Transportation, 1996.

Lari, A. et. al., I-35W Incident Management and Impact of Incidents on Freeway Operations, Minnesota DOT, January 1982.

Lindley, Jeffrey A, “Quantification of Urban Freeway Congestion, Report: FHWA/RD-87/052, Oct 1986. This report

Owen J. R. and Urbanek, G. L., “Alternative Surveillance Concepts and Methods for Freeway Incident Management – Vol. 2: Planning and Tradeoff Analyses for Low Cost Alternatives, USDOT, Report No. GHWA-RD-77-59, March 1978.

Reiss, R.A., and W.M. Dunn, Jr. Freeway Incident Management Handbook. Report FHWA-SA-91-056. Federal Highway Administration, U.S. Department of Transportation, July 1991. Available online at http://www.library.unt.edu/gpo/ota/tech/traffic/ap153.html Also, it was updated in 97. In NYSDOT library, p. 4-17, Figure 4-6 Effect of an incident on motorist delay.

SG Associates, Inc. & Howard Needles Tammen & Bergendoff, “PIN 8729.30 Freeway Delay Analysis Software Users Guide”, Nov. 1992.

************** Numbered References **********************

Ref. # 11. (Incident Factor) These factors are the percentage of total accidents that occur for each “incident type”. The data found for “shoulder” with “6-7 lanes” and “8 or more lanes” are the same as found in the column labeled “Reportable Accidents” in Table 3, p 2 in Reiss 1991. Reiss cites the data source as Ontario Ministry of Transportation, Incident Management Project, January, 1988. The data is summarized from three months of data from Highway 401 in Toronto. The source for the remaining data in CNAM’s table is unknown.

The Reiss/Ontario table is also used in Gordon, R. L. et.al. (1996) Table 4-25 on p. 4-45. However since NYSDOT does not have the number of lanes by direction (only number of lanes in both directions), staff had to combine lanes into groupings roughly equivalent to Gordon’s table.

Regardless that numerous researchers base their publications on this data, because this table is based on only three months of data we believe that this table is subject to revision by the Regions/consultants.

Ref. # 22. (Accident Rates) Seth Asante, NYSDOT, unpublished. Developed this table based on 3 million accident reports from NYS Dept of Motor Vehicles. Using the latest available data in April 2004, Rodney Delisle is updating this table.

Ref. # 33. (Average Duration) Lindley 1986 p. 12, states " The average duration for incidents, ….. were estimated from several data sources… (Owen and Urbankek, 1978, Goolsby, 1971, Lari, A. 1982.

Ref. # 44. (Available Capacity) This percent of capacity is found in many places with various sources cited:

Found in Lindley, 1986 (Table 2, p 13). To create this table, Lindley started with data from Owen and Urbankek, 1978, and did additional work. Portions of his table that apply to CNAM are as follows:

Number of lanes			*** # of lanes blocked ***		
In each dir		on shoulder	1	2	3
2	Shoulder	.81	.35	0	0
3		.83	.49	.17	0
4		.85	.58	.25	.13

Found in HCM 2000, Exhibit 22-6, p 22-11, source cited is Reiss, 1991, Gordon, et. al. 1996.

Found in Gordon 1996, Table 4-9, p. 4-18, source cited is Reiss, 1991.

Found in SG Assoc, Inc, Howard Needles Tammen & Bergendoff consultants to NYSDOT, no source given.

ATTACHMENT D

Visual Basic Code used in the Hypothetical Analysis

```
Option Explicit

Private Type NetworkRecord
    Length As Single
    NRdwys As Integer
    NLanes As Integer
    Shldr As Integer
    AADT As Single
    Access As Integer
    Cap As Single
End Type
Dim Net() As NetworkRecord

Private Type CnamAccRecord
    NLanes As Integer
    NoCntrl(2) As Single
    FPCntrl(2) As Single
End Type
Dim CnamAcc() As CnamAccRecord

Private Type CnamIncidentRecord
    Lower As Integer
    Upper As Integer
    PrInc(2, 4) As Single
    PctCap(2, 4) As Single
    CTime(4) As Single
End Type
Dim CnamInc() As CnamIncidentRecord

Private Type NewIncRecord
    Freq As Single
    Dur(2) As Single
    ProbLB(5) As Single
End Type
Dim NewInc() As NewIncRecord

Private Type NewPctCapacityRecord
    Lower As Integer
    Upper As Integer
    PctCap(2, 4) As Single
End Type
Dim NewPctCap() As NewPctCapacityRecord

Private Type PropDamageRecord
    Lower As Single
    Upper As Single
    Freq(2) As Single
End Type
Dim PropDam() As PropDamageRecord
```

```
Private Type DisabledVehicleRecord
    Lower As Single
    Upper As Single
    Freq(2) As Single
End Type
Dim DisVeh() As DisabledVehicleRecord

Dim CnamInciToAcc As Single

Public Sub Main()

Call Initialize
Call CnamEvaluate
Call NewEvaluate
Call FinishUp

End Sub

Public Sub Initialize()

Dim i As Integer, j As Integer, k As Integer
Dim Test As Variant
Dim Xpand As Single

'The AADT values in the network dataset are in 10's, so are the capacities
'The AADT values are two-way
i = 1
Test = Worksheets("Network").Cells(i + 1, 1)
Do While (Test <> "")
  ReDim Preserve Net(i)
  Net(i).Length = Worksheets("Network").Cells(i + 1, 4)
  Net(i).NRdwys = Worksheets("Network").Cells(i + 1, 5)
  Net(i).NLanes = Worksheets("Network").Cells(i + 1, 6)
  Net(i).Shldr = Worksheets("Network").Cells(i + 1, 7)
  Net(i).Access = Worksheets("Network").Cells(i + 1, 12)
  Net(i).AADT = 10# * Worksheets("Network").Cells(i + 1, 9)
  Net(i).Cap = 10# * Worksheets("Network").Cells(i + 1, 13)
  i = i + 1
  Test = Worksheets("Network").Cells(i + 1, 1)
Loop

'CNAM Model Data
'Note: Accidents rates are per million vehicle miles for CNAM

'CNAM Accident Rates
ReDim CnamAcc(9)
For i = 1 To 9
  CnamAcc(i).NLanes = i
  CnamAcc(i).NoCntrl(1) = Worksheets("CNAM").Cells(i + 6, 3)
  CnamAcc(i).NoCntrl(2) = Worksheets("CNAM").Cells(i + 16, 3)
  CnamAcc(i).FPCntrl(1) = Worksheets("CNAM").Cells(i + 6, 4)
  CnamAcc(i).FPCntrl(2) = Worksheets("CNAM").Cells(i + 16, 4)
Next i

'CNAM Incident probabilities and lane blockage times
```

```
ReDim CnamInc(3)
For i = 1 To 3
  CnamInc(i).Lower = Worksheets("CNAM").Cells(i + 5, 5)
  CnamInc(i).Upper = Worksheets("CNAM").Cells(i + 5, 6)
  For j = 1 To 2
    For k = 1 To 4
      CnamInc(i).PrInc(j, k) = Worksheets("CNAM").Cells(i + 5, 6 + 4 * (j -
1) + k).Value / 100#
      CnamInc(i).PctCap(j, k) = Worksheets("CNAM").Cells(i + 13, 6 + 4 * (j -
1) + k).Value / 100#
      CnamInc(i).CTime(k) = Worksheets("CNAM").Cells(20, 6 + 4 * (j - 1) +
k).Value
    Next k
  Next j
Next i

'CNAM Incident to Accidents Ratio
CnamInciToAcc = Worksheets("CNAM").Range("InciToAcc").Value

'New Model

'Note: Incident rates are per day per mile for the new look-up tables

'The accident rate expansion factor is employed.
'It is read from the "New" tab.
'Cell (21,11) contains the factor based on distance-weighted average rates.
'Cell (22,11) contains the factor based on trend line analysis.
'Cell (23,11) contains the factor based on accidents alone.
'Cell (24,11) contains the factor currently being used.
Xpand = Worksheets("New").Cells(24, 11).Value

ReDim NewInc(8)
For i = 1 To 8
  NewInc(i).Freq = Worksheets("New").Cells(i + 4, 3).Value * Xpand
  NewInc(i).Dur(1) = Worksheets("New").Cells(i + 4, 5).Value
  NewInc(i).Dur(2) = Worksheets("New").Cells(i + 4, 4).Value
Next i

'The AADT values in Tables 6 and 7 are one-way
ReDim PropDam(6)
For i = 1 To 6
  PropDam(i).Lower = Worksheets("New").Cells(i + 3, 9).Value
  PropDam(i).Upper = Worksheets("New").Cells(i + 3, 10).Value
  PropDam(i).Freq(1) = Worksheets("New").Cells(i + 3, 12).Value * Xpand
  PropDam(i).Freq(2) = Worksheets("New").Cells(i + 3, 11).Value * Xpand
Next i

ReDim DisVeh(6)
For i = 1 To 6
  DisVeh(i).Lower = Worksheets("New").Cells(i + 12, 9).Value
  DisVeh(i).Upper = Worksheets("New").Cells(i + 12, 10).Value
  DisVeh(i).Freq(1) = Worksheets("New").Cells(i + 12, 11).Value * Xpand
  DisVeh(i).Freq(2) = Worksheets("New").Cells(i + 12, 12).Value * Xpand
Next i
```

```
'Percent Lane Blockage Data for the New Procedure
For i = 1 To 8
For j = 1 To 5
  NewInc(i).ProbLB(j) = Worksheets("New").Cells(i + 17, j + 2).Value
Next j
Next i

'Percentage Capacities for the New Procedure
ReDim NewPctCap(3)
For i = 1 To 3
  NewPctCap(i).Lower = Worksheets("New").Cells(i + 29, 1)
  NewPctCap(i).Upper = Worksheets("New").Cells(i + 29, 2)
  For j = 1 To 2
    For k = 1 To 4
      NewPctCap(i).PctCap(j, k) = Worksheets("New").Cells(i + 29, 3 + 4 * (j
- 1) + k).Value / 100#
    Next k
  Next j
Next i

End Sub

Sub CnamEvaluate()

'The subroutine simulates an incident in one direction.
'It then doubles that result to get the VHD for the link.

Dim i As Integer, j As Integer, k As Integer

Dim Queue As Single, CaseDlyMin As Single, Length As Single
Dim TotDlyMin As Single, CaseProb As Single, CaseDur As Single
Dim DirAADT As Single, NLanes As Integer, ShWidth As Integer, NRdwys As
Integer
Dim PctCap As Single, CTime As Integer, Dem As Single, Access As Integer
Dim CaseCap As Single, NL As Integer, ShType As Integer, Cap As Single
Dim AccRate As Single, IncRate As Single

Dim t As Integer

For i = 1 To UBound(Net)
  TotDlyMin = 0
  NLanes = Net(i).NLanes
  Length = Net(i).Length
  ShWidth = Net(i).Shldr
    'The CNAM manual says 6' is the minimum to be useful
    If (ShWidth >= 6) Then ShType = 1 Else ShType = 2
  Access = Net(i).Access
  NRdwys = Net(i).NRdwys
  DirAADT = Net(i).AADT / 2#

  'Values of capacity and demand for the study hour
  Cap = Net(i).Cap 'Some capacities are less than 5% AADT
  Dem = min(0.99 * Cap, 0.05 * DirAADT) '5% of Directional AADT, and less
than cap
```

```
  If (Access = 2) Then
    AccRate = CnamAcc(NLanes).FPCntrl(NRdwys)
  Else
    AccRate = CnamAcc(NLanes).NoCntrl(NRdwys)
  End If

  'Write out the basic CNAM accident rate
  Worksheets("Network").Cells(i + 1, 21).Value = AccRate

  'Accident rate per mile per day
  '((Accidents/MVMT)/(1e6))* DirAADT = Accidents/Mi/Day
  AccRate = (AccRate / 1000000#) * DirAADT

  'Write Out Incident and Accident Rates
  IncRate = CnamInciToAcc * AccRate
  Worksheets("Network").Cells(i + 1, 15).Value = IncRate
  Worksheets("Network").Cells(i + 1, 22).Value = AccRate

  'Convert the Accident Rate and the Incident Rate into values for the study
hour
  'Assume it is proportional to the traffic being studied (5% of DirAADT)
  IncRate = IncRate * 0.05
  AccRate = AccRate * 0.05

  For k = 1 To 3
    If ((NLanes >= CnamInc(k).Lower) And (NLanes <= CnamInc(k).Upper)) Then
      NL = k
      Exit For
    End If
  Next k

  'Convert demand and capacity into values per minute
  Cap = Cap / 60#
  Dem = Dem / 60#

  'Conduct the analysis
  For j = 1 To 4
    CaseProb = CnamInc(NL).PrInc(ShType, j)
    If (CaseProb > 0) Then
      CaseDlyMin = 0
      Queue = 0
      PctCap = CnamInc(NL).PctCap(ShType, j)
      CaseDur = CnamInc(NL).CTime(j)
      CaseCap = Cap * PctCap
      For t = 1 To 480
        If (t >= CaseDur) Then CaseCap = Cap
        CaseDlyMin = max(0, CaseDlyMin + 0.5 * Queue)
        Queue = max(0, Queue + Dem - CaseCap)
        CaseDlyMin = max(0, CaseDlyMin + 0.5 * Queue)
        If (Queue <= 0) Then Exit For
      Next t
      TotDlyMin = TotDlyMin + (CaseProb * CaseDlyMin) * Length * IncRate
    End If
  Next j
```

```
  'Double the total delay to reflect both directions
  Worksheets("Network").Cells(i + 1, 17).Value = TotDlyMin * 2

Next i

End Sub

Sub NewEvaluate()

Dim i As Integer, j As Integer, k As Integer, n As Integer

Dim ShType As Integer, NomCap As Single, DirAADT As Single
Dim NLanes As Integer, NL As Integer, Access As Integer
Dim InciRate As Single, ShWidth As Integer
Dim CaseRate As Single, TotDlyMin As Single
Dim PctCap As Single, CTime As Integer, Dem As Single
Dim Queue As Single, Length As Single, DlyMin As Single
Dim Cap As Single, ProbLB As Single, Dur As Single
Dim AccRate As Single, LaneFlag As Integer

Dim t As Integer

For i = 1 To UBound(Net)
  Length = Net(i).Length
  ShWidth = Net(i).Shldr
    If (ShWidth >= 6) Then ShType = 1 Else ShType = 2
  NLanes = Net(i).NLanes
    If (NLanes >= 8) Then LaneFlag = 2 Else LaneFlag = 1
  For k = 1 To 3
    If ((NLanes >= NewPctCap(k).Lower) And (NLanes <= NewPctCap(k).Upper))
Then
      NL = k
      Exit For
    End If
  Next k

  'AADT, Capacity, and Demand per Hour
  NomCap = Net(i).Cap
  DirAADT = Net(i).AADT / 2# 'Network AADTs are for both directions
  Dem = min(0.99 * NomCap, 0.05 * DirAADT) 'Less than the DIrAADT or 5% of
directional AADT

  'Covert capacity and demand to values per minute
  NomCap = NomCap / 60#
  Dem = Dem / 60#

  TotDlyMin = 0
  InciRate = 0
  AccRate = 0
  For j = 1 To 8
    If (j = 1) Then
      For k = 1 To 6
        If (DirAADT >= PropDam(k).Lower) And (DirAADT <= PropDam(k).Upper)
Then
          CaseRate = PropDam(k).Freq(ShType)
```

```
          Exit For
        End If
      Next k
    ElseIf (j = 2) Then
      For k = 1 To 6
        If (DirAADT >= DisVeh(k).Lower) And (DirAADT <= DisVeh(k).Upper) Then
          CaseRate = DisVeh(k).Freq(LaneFlag)
          Exit For
        End If
      Next k
    Else
      CaseRate = NewInc(j).Freq
    End If

    'Update the total incident rate
    InciRate = InciRate + CaseRate

    'Update the accident rate from the incident rate
    If ((j = 1) Or (j = 4)) Then
      AccRate = AccRate + CaseRate
    End If

    'Convert CaseRate to the hour under study
    CaseRate = CaseRate * 0.05 'Assume consistent wtih 5% traffic

    'SubCases for each # of lanes blocked
    For n = 1 To 4
      DlyMin = 0
      Queue = 0
      Dur = NewInc(j).Dur(ShType)
      ProbLB = NewInc(j).ProbLB(n)
      PctCap = NewPctCap(NL).PctCap(ShType, n)
      Cap = PctCap * NomCap
      For t = 1 To 480
        If (t >= Dur) Then Cap = NomCap
        DlyMin = max(0, DlyMin + 0.5 * Queue)
        Queue = max(0, Queue + Dem - Cap)
        DlyMin = max(0, DlyMin + 0.5 * Queue)
        If (Queue <= 0) Then Exit For
      Next t
      'Assume case rate for an hour is 5% based on demand
      TotDlyMin = TotDlyMin + (ProbLB * DlyMin) * Length * CaseRate
    Next n
  Next j

  'Write out the incident and accident rates
  Worksheets("Network").Cells(i + 1, 16).Value = InciRate
  Worksheets("Network").Cells(i + 1, 23).Value = AccRate

  'Multiply VHD times 2 to account for both directions
  Worksheets("Network").Cells(i + 1, 18).Value = TotDlyMin * 2#

Next i

End Sub
```

```
Sub FinishUp()

End Sub

Function min(X, y) As Variant

  If (X < y) Then min = X Else min = y

End Function

Function max(X, y) As Variant

  If (X > y) Then max = X Else max = y

End Function
```